Geometría y Álgebra Lineal 1

Parte 2

Instituto de Matemática y Estadística “Prof. Ing. Rafael Laguardia”

Abril de 2005

Oficina de Publicaciones del Centro de Estudiantes de Ingeniería
Julio Herrera y Reissig 565 – Tel. 711 79 10 – Fax 711 06 21
Correo electrónico publicaciones@cei.fing.edu.uy

Puesta en página: Omar Gil (IMERL - Facultad de Ingeniería)
Colaboración técnica: Gabriel Santoro (DI.R.A.C. - Facultad de Ciencias)

Montevideo – Uruguay – 2005

Impresión y encuadernación: mastergraf srl
Gral. Pagola 1727 – Tel. 203 4760*
Correo electrónico mastergraf@netgate.com.uy

Depósito Legal 335.588 – Comisión del Papel
Edición amparada al Decreto 218/96

0.1. Introducción

Este texto pretende servir como referencia básica para el curso Geometría y Álgebra Lineal 1 (GAL1) que la Facultad de Ingeniería de la Universidad de la República ofrece a sus estudiantes en el primer semestre de sus carreras. Trabajaremos con la versión que tenés entre tus manos en la edición 2005 de ese curso.

En esta introducción comentaremos algunas de las razones por las que un curso como GAL1 aparece en el currículo de la Facultad de Ingeniería; daremos una descripción general, ligeramente técnica, de los contenidos del curso y el texto, y comentaremos algunos detalles acerca de la elaboración de este último.

0.1.1. ¿Por qué estudiar Geometría y Álgebra Lineal?

Nos gustaría compartir contigo algunas razones que justifican que un curso como GAL1 aparezca al principio de las carreras de Ingeniería.

Comenzaremos por mencionar la enorme utilidad y aplicabilidad que tiene el Álgebra Lineal, que provee la aritmética básica para tratar problemas en los que interviene más de una variable. Está en los fundamentos de la Matemática contempóranea, y en el corazón de muchas de sus aplicaciones a nuestra tecnología.

Algunas aplicaciones del Álgebra Lineal

Hagamos una lista no exhaustiva y un tanto abigarrada de problemas en cuya solución el Álgebra Lineal desempeña un papel relevante:

- la emplean los buscadores de Internet para ofrecer al usuario una lista de sitios ordenada de una forma útil;
- analizar cómo se distribuyen los esfuerzos en cualquier estructura sometida a cargas (la estructura reticulada de un puente, el fuselaje de un avión, etcétera);
- permite ordenar la información de muchos problemas combinatorios para tratarlos. Por ejemplo, con su ayuda se puede dar una demostración simple de un criterio para determinar si es posible casar a todas las mujeres de un pueblo de forma tal que cada una de ellas esté satisfecha con el marido que le tocó;
- proteger la integridad de nuestra información cuando la almacenamos en un disco compacto o la transmitimos por cualquier canal;

- desplegar un gráfico sobre la pantalla de una computadora;
- nos ayuda a describir cómo depende el costo de una serie de productos del costo de cada uno de los insumos necesarios para fabricarlos;
- permite describir los movimientos del brazo de un robot para controlarlo y ponerlo a realizar alguna tarea que nos interese;
- provee algunas de las herramientas necesarias para ajustar un conjunto de observaciones a los modelos teóricos disponibles;
- trabaja dentro de muchos algoritmos de compresión de la información que son imprescindibles para hacer posibles las comunicaciones modernas (por ejemplo, se despliega ante nuestros ojos cada vez que miramos un archivo `.jpg`);
- interviene en los procesos de optimización necesarios para una gestión eficiente;
- permite describir el esquema de conexiones en cualquier red de cualquier naturaleza;
- etcétera, etcétera, etcétera, requete-recontra-etcétera.

Discutiremos algunos de estos ejemplos a lo largo de este curso. Encontrarás otros más adelante en tu carrera, o a lo largo de tu vida profesional. En particular, hemos omitido en la lista aplicaciones relevantes cuya presentación requeriría explicaciones previas demasiado extensas para esta introducción[1]. Estamos seguros de que existen aplicaciones bonitas e interesantes que los autores de este texto ignoramos, pero que bien podrían formar parte de él. También que en el futuro aparecerán otras que hoy no existen. Es posible incluso que tu trabajo consista en desarrollarlas.

Por el lugar que ocupa en la ciencia contemporánea el Álgebra Lineal está entre los primeros cursos de las carreras científicas en cualquier lugar del mundo. En particular, en las carreras de Ingeniería. Te invitamos a hacer una recorrida por los sitios web de otras instituciones científicas y comprobarlo por tí mismo. Si el tema te interesa, también podés buscar información adicional sobre el Álgebra Lineal y sus aplicaciones en las bibliotecas y en Internet.

[1]Algunas de ellas están directamente relacionadas con la posibilidad de construir buenas aproximaciones lineales de fenómenos no lineales, un asunto que discutimos algunas páginas más adelante, en esta misma introducción.

Para cerrar estos comentarios citemos un párrafo tomado[2] de la introducción de [S]: *Estoy muy feliz porque la necesidad del Álgebra Lineal es ampliamente reconocida. Es* absolutamente *tan importante como el cálculo. No cedo nada al respecto, cuando miro cómo se usa actualmente la matemática. Muchas aplicaciones son hoy discretas más que continuas, digitales más que analógicas, linealizables más que erraticas y caóticas. Entonces los vectores y las matrices son el lenguaje a conocer.*

El papel de la Geometría

En el curso GAL1 también trataremos temas de Geometría. Quizás no parezca necesario detenerse a explicar por qué se justifica estudiar Geometría. Al fin y al cabo, muchos de los problemas que tiene que resolver un ingeniero están ubicados en el espacio, y la Geometría, rama antigua de la Matemática especializada en estos menesteres, lo modela adecuadamente. Sin embargo, hay otra razón, al menos tan valiosa como la que acabamos de dar, para estudiar Geometría: la Geometría va muchísimo más allá y es muchas más cosas que una rama de la Matemática que describe el espacio físico.

Algunas nociones geométricas básicas como distancia, ángulo, recta, plano, curva, etcétera, admiten generalizaciones a contextos muy variados, que probablemente no imaginás; con los que casi seguramente ni siquiera soñaron Pitágoras, Euclides, ni los antiguos egipcios preocupados por las crecidas del Nilo. Apuntamos uno de ellos: el tratamiento contempóraneo de la información (señales de radio y televisión, información digital) emplea elaborados conceptos geométricos.

La Matemática crea permanentemene nuevas geometrías, para nuevas situaciones.

Esperamos que estos comentarios sean útiles para todos los estudiantes, pero los hacemos teniendo especialmente presentes a los estudiantes de Informática, que suelen mostrar menos interés por la Geometría. Quizás la difusión de la creencia errónea de que la Geometría se refiere exclusivamente al mundo físico, que es distinto del mundo en el que a muchos informáticos les gusta vivir, explique este fenómeno. Enfaticemos entonces que Internet, las cadenas de bits y los códigos criptográficos, por poner algunos ejemplos, también tienen su Geometría. ¡¡¡Ni qué hablar de la computación gráfica!!!

Se trata, además, de una Geometría que se lleva muy bien con el Álgebra Lineal, y es su compañera inseparable en el tratamiento de muchos de los problemas que hemos incluido en nuestra lista de aplicaciones del Álgebra

[2]La traducción es de los autores.

Lineal.

0.1.2. ¿De qué se trata el curso Geometría y Álgebra Lineal?

En las próximas secciones daremos una breve descripción del curso GAL1, con un tono que es casi de divulgación. No esperamos que todo quede claro en una primera lectura, pero sí que el hacerla te estimule a aprender más. Todo lo que aquí se discute se desarrolla luego en el cuerpo principal del texto.

Álgebra Lineal

Intentemos una explicación breve. Escojamos, para empezar, entre varias opciones posibles, un punto de vista que creemos que puede ser interesante para un estudiante de Ingeniería: podemos decir que se trata de las matemáticas de los fenómenos en que los efectos son proporcionales a las causas. Las intensidades son proporcionales a las caídas de potenciales, los esfuerzos en una estructura son proporcionales a las cargas a las que se somete, lo que gastamos en bizcochos es proporcional a la cantidad que comemos. Pero cuando hay muchas causas y muchos efectos mezclados entonces las relaciones de dependencia se vuelven más complejas, y hay que ordenarlas para entenderlas.

Al hacerlo, emerge una estructura algebraica, una estructura absolutamente abstracta que en principio no hace referencia a ninguna aplicación particular, que ni siquiera las necesita para sostenerse y vivir, pero que permite tratar infinidad de ellas[3].

Esta estructura algebraica es la de los *espacios vectoriales de dimensión finita*. El Álgebra Lineal es la rama de la matemática que estudia, justamente, los espacios vectoriales de dimensión finita.

¿Qué significa esto?

La expresión "espacio vectorial" alude a un espacio formado por *vectores*. La palabra "espacio" hace referencia a un conjunto dotado de cierta estructura. De modo que estamos hablando de un conjunto de vectores, que tiene además la estructura de espacio vectorial que mencionamos antes.

¿Qué cosas son los vectores?

En algún sentido esta pregunta esta esencialmente desencaminada, fuera de lugar. Estamos tratando con una estructura algebraica **abstracta**. Para ella

[3]Arriesgamos aquí una opinión. Creemos que en esta frase se condensa buena parte de las dificultades que muchas personas tienen para el aprendizaje de las matemáticas: la estructura abstracta no significa nada, porque jamás han recorrido un camino que les diera contenido. Y terminan enfrentadas a una tarea que podría ser comparada con el intento de aprender un idioma sin entender el significado de ninguna de sus palabras.

no importa nada la naturaleza de sus elementos. Sólo **interesa cómo se relacionan entre sí**. Digamos entonces que los vectores pueden sumarse, y multiplicarse por números, y estas dos operaciones siguen una serie de reglas que son las que definen la estructura algebraica de un espacio vectorial.

Un vector es, pues, una noción abstracta. Llamaremos vector a cualquier elemento de un conjunto dotado de estructura de espacio vectorial, independientemente de cuál sea su naturaleza.

Dados dos vectores X e Y podremos sumarlos para producir un nuevo vector

$$X + Y,$$

o multiplicar uno de ellos, por ejemplo X, por un número λ para generar el vector

$$\lambda X.$$

Muchos objetos obedecen este esquema. Por ejemplo, los vectores geométricos y las funciones reales. En nuestro cursos trabajaremos mucho con listas de números

$$X = (x_1, x_2, \ldots, x_n), \tag{1}$$

donde x_1, x_2,..., x_n, son, por ejemplo, números racionales o reales.

Los vectores también pueden ser cosas como clases de equivalencia de curvas que pasan por el mismo punto de un objeto que es una generalización de una superficie. Ejemplo un tanto intimidante en una primera aproximación, pero que sirve para extender a las geometrías no euclidianas la noción de plano tangente a una superficie . Esta construcción es necesaria, por ejemplo, para formular la Teoría General de la Relatividad.

Los conjuntos formados por los vectores geométricos, las funciones reales y las listas de n números, también el espacio tangente a una superficie en uno cualquiera de sus puntos, son ejemplos de espacios vectoriales. Realizaciones concretas de la noción abstracta de espacio vectorial.

Álgebra Lineal y Geometría

A continuación explicaremos cuál es la geometría que vamos a usar más frecuentemente en este texto. Vale la pena subrayar que no es la única geometría posible.

Trabajaremos frecuentemente con vectores de la forma (1), listas de n números reales. Listas de dos o tres números representan, respectivamente, vectores geométricos en el plano (de dimensión 2) o el espacio (de dimensión 3), luego de que se escoge un sistema ortogonal de coordenadas. Considerar listas con más números no genera mayores dificultades, la geometría puede extenderse a espacios de dimensiones muy altas, y conservar allí todo su significado.

Por ejemplo, el procedimiento con el que las imágenes son codificadas en la forma de un archivo `.jpg` requiere una noción de perpendicularidad en un espacio de dimensión 64; el buscador Google ordena datos trabajando en un espacio que, en principio, tiene una dimensión igual a la cantidad de páginas web disponibles en Internet; para describir la vibración de una cuerda de guitarra analizando el peso de cada una de las frecuencias presentes en el sonido hace falta un espacio que contiene infinitas direcciones perpendiculares entre sí.

Todas estas extensiones de la geometría dependen de que las nociones de distancia y ángulo pueden introducirse a través de un procedimiento algebraico, el producto escalar, o producto interno. Lo que habilita a definirlas en cualquier espacio vectorial, independientemente de la naturaleza de sus vectores.

Dimensión finita

Resta explicar qué significa la expresión "dimensión finita". Quiere decir que cualquier vector del espacio queda perfectamente caracterizado por una cantidad finita de números. Por ejemplo, para las listas de números (1) hacen falta exactamente n. La cantidad de números que precisamos es lo que se llama la *dimensión* del espacio.

Hay otra imagen más geométrica para expresar el mismo concepto. Un espacio tiene dimensión n si es posible especificar n vectores

$$X_1, X_2, \ldots, X_n,$$

de modo que cualquier vector X en el espacio puede reconstruirse de manera única a partir de esos n usando las dos operaciones disponibles: multiplicando primero cada X_i por un número λ_i, $i = 1, 2, \ldots, n$, para producir nuevos n vectores

$$\lambda_1 X_1, \lambda_2 X_2, \ldots, \lambda_n X_n,$$

y luego sumando todo de modo que

$$X = \lambda_1 X_1 + \lambda_2 X_2 + \ldots + \lambda_n X_n. \tag{2}$$

En esta sencilla observación reside buena parte de la gracia de la teoría de los espacios vectoriales. Con sólo n vectores podemos reconstruir todos los vectores del espacio, que pueden ser infinitos, o más, muchos más, que n. La expresión de X en la forma (2) es lo que llamamos una *combinación lineal* de la familia formada por los vectores $X_1, X_2, \ldots, X_n$. Esta construcción es tan básica para la teoría, que podríamos resumir en una única frase la caracterización de los vectores como objetos que pueden sumarse y multiplicarse por escalares: **consideramos *vectores* a los objetos con los que podemos hacer combinaciones lineales**.

En los problemas prácticos la dimensión suele aparecer como la cantidad de variables independientes que podemos manipular o los parámetros disponibles para el diseño. También es una medida del "tamaño" del espacio vectorial.

0.1.3. Miradas cruzadas

De la geometría al álgebra, del álgebra a la geometría; de las aplicaciones a la teoría, de la teoría a las aplicaciones; de un problema práctico a su modelo matemático, del modelo a su teoría y al cálculo, de aquí a interpretar los resultados en el problema original.

En ese ir y venir, entre distintas maneras de aproximarse a un problema, entre distintas descripciones en distintos registros, está mucho de la capacidad de usar una teoría científica para modelar la realidad y transformarla. En este curso intentaremos ilustrar, aunque sea modestamente, estas posibilidades. Compartimos desde ya una de ellas con el lector.

Matrices

Buena parte de nuestro trabajo tendrá que ver con matrices. Una matriz es un arreglo de números en filas y columnas. Por ejemplo

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 3 & e & \pi & -1 \\ \sqrt{2} & -e^2 & 0 & -\sqrt{3} \end{pmatrix}$$

es una matriz de números reales, que tiene 3 filas y 4 columnas. A lo largo del texto trabajaremos con matrices con un número arbitrario de filas y de columnas, cuyas entradas podrán ser números de cualquier naturaleza.

Las matrices son un interesante ejemplo de como distintos puntos de vista pueden enriquecerse y complementarse. Por ejemplo, nuestra primera matriz podría provenir de algún problema práctico, y tener un significado propio en el contexto de ese problema; pero puede pensarse también como un conjunto de doce números; como una familia ordenada de tres filas de cuatro números; como cuatro columnas de tres números; también puede considerarse como una única cosa, que además es un elemento de una estructura algebraica; puede ser el objeto de trabajo de un algoritmo; o la descripción de una función de un espacio de dimensión 4 en uno de dimensión 3. Cada una de estas posibilidades tiene pleno sentido, dentro alguna estructura algebraica y/o geométrica.

Matrices y Geometría

El punto de vista geométrico es especialmente rico y fructífero. Tal como veremos, transformaciones geométricas como rotaciones, simetrías, proyecciones y dilataciones pueden describirse por medio de matrices. Esto es realmente notable. Pero más notable aún es que se puede dar la vuelta a esta observación. Si la geometría puede guardarse en unas matrices, las matrices y su acción pueden ser reinterpretadas en clave geométrica.

Cuando pensamos las cosas de esta manera ganamos un nuevo punto de vista para analizar distintas situaciones y problemas. Muchos problemas en que los datos son listas o matrices de números admitirán una representación geométrica. Ganaremos así una imagen clara, en la que nuestra intuición puede guiarnos, para entenderlos. Una cosa es pensar en términos de una larga lista de, digamos, 1.600 números reales. Y otra es pensar que, por ejemplo, esos números están ordenados en una matriz de 40 filas y 40 columnas, cuya acción puede ser comprendida como una proyección de un espacio de dimensión 40 sobre uno que tiene dimensión 3. Tal vez te sorprendas si te anunciamos que semejante proyección no se comporta muy distinto que la proyección usual sobre un plano del espacio. Y una vez comprendido este hecho la danza de los 1.600 números empieza a cobrar sentido. En consecuencia, puede describirse eficientemente con menos información. Por ejemplo, con 120 números. 40 por cada uno de los 3 vectores necesarios para recuperar todo un espacio de dimensión 3.

Este es el tipo de relaciones que esperamos ir iluminando y desentrañando a lo largo del texto.

0.1.4. La estructura conceptual del Álgebra Lineal

Una vez que hemos ilustrado brevemente algunos de los ejes conceptuales del curso GAL1 estamos prontos para corregir el exceso de los primeros párrafos de esta introducción, en el que se justifica el estudio del Álgebra Lineal por su gran cantidad de aplicaciones. Esta visión es esencialmente incompleta, insuficiente, y no hace justicia a la disciplina. ¡Ni siquiera a su gran versatilidad y aplicabilidad!

Él Álgebra Lineal constituye una teoría coherente, elegante, rica. Como dice el autor de [T] al analizar el papel que tiene en la actualidad el Álgebra Lineal: *¡El Álgebra Lineal es un modelo de lo que una teoría matemática* *debería ser!*

Está edificada alrededor de algunas pocas ideas centrales, simples y profundas. Y se puede estructurar a partir de un reducido conjunto de axiomas o postulados, a partir del cual la teoría se construye lógicamente siguiendo un desarrollo que bien puede exponerse en uno o dos semestres de cursos de Matemática al nivel de cursos finales de Bachillerato o iniciales en la Universidad. Ofrece pues un modelo para pensar con rigor, organizar información, intuir y luego demostrar, abstraer estructuras generales a partir de situaciones más o menos concretas, interpretar un resultado abstracto en un ejemplo, establecer conexiones entre distintas representaciones de un mismo concepto u objeto, etcétera. Tiene entonces un conjunto de cualidades, a las que se suma su capacidad para explicar o contribuir a hacer posible partes del mundo que nos rodea, que invitan a llevarla a las aulas.

Contra un utilitarismo exagerado y empobrecedor

Es importante que desarrolles, entre otras, las habilidades que enumerábamos sobre el fin del párrafo anterior. Es una cuestión que hace a tu formación intelectual, a tu bagaje cultural. Es quizás lo más importante que tenés para hacer en tu pasaje por los estudios de grado. Si alcanzás estos objetivos seguramente estarás capacitado para aprender luego lo que necesites saber para cualquier actividad futura.

Es una buena apuesta preocuparse más por entender, por desarrollar la capacidad de manejar conceptos con flexibilidad, que adquirir información detallada sobre una herramienta específica que permite resolver un buen conjunto de problemas. Digamos entonces que la Geometría y el Álgebra Lineal, en general la Matemática, son mucho más que una herramienta. Y tienen en el currículo un valor formativo[4] importante. Es cierto que pueden ser empleados con el punto de vista un tanto estrecho de considerarlos pura y exclusivamente como artificios útiles para resolver un problema, y aún así conseguir algunos resultados. Pero nuestra recomendación es que no te pierdas el resto de la historia.

Los comentarios anteriores son un buen pie para cambiar el punto de vista, mirar el conjunto de la sociedad, y expresar la convicción de que tampoco es una buena idea desarrollar la ciencia con un afán puramente utilitario en mente. Por ejemplo, es cierto que buena parte de la Matemática existente

[4]Sin embargo, arriesgamos aquí el punto de vista de que ninguna asignatura debería incurrirse en el currículo simplemente porque tenga un valor formativo. Si ésa es la única razón para incluir un curso entonces debería ser sustituido por otra actividad que tenga valor formativo y, al mismo tiempo, atienda otros objetivos.

se desarrolló alimentándose del interés por resolver problemas prácticos. Pero muchas otras partes no. Y es sorprendente ver como se repite una y otra vez la circunstancia de que ideas que aparecieron en algún campo del conocimiento son la clave para progresar luego en otros, a través de una conexión originalmente imprevisible. En particular, muchas soluciones prácticas que nuestra tecnología brinda están basadas en conocimientos que aparecieron buscando respuestas a preguntas que fueron formuladas por el afán de entender, no por el de solucionar ningún problema concreto.

Cerramos estos párrafos dedicados a prevenirte contra un excesivo utilitarismo, contra la exageración en el preguntar ¿para qué sirve esto? agregando algunas dimensiones más. Se puede acceder a muchas cuestiones interesantes con una perspectiva lúdica, explorando posibilidades por el placer de jugar con ellas. La referencia al placer no es casual, ni menor. Esperamos que recorrer este curso, y los que le seguirán, sean experiencias disfrutables, en las que vivas la alegría de descubrir y compartir ideas con tus docentes y compañeros. Es posible encontrar también un sentido estético en la Matemática. Para muchas personas esto es suficiente para dedicar sus vidas a ella.

En fin, hay muchas maneras de aproximarse a un curso de Matemática. Entre ellas, claro está, la de aceptar con resignación que forma parte del currículo e intentar sobrevivir a la experiencia. ¡Ojalá que encuentres alguna otra!

0.1.5. Linealizando lo no lineal

Hemos mencionado antes que, aunque relativamente simple, la estructura lineal del Álgebra Lineal (valga la redundancia) es suficientemente flexible como para adaptarse a muchas situaciones y tratar una gran variedad de aplicaciones.

Agreguemos otra razón para estudiar detenidamente los problemas lineales: cuando un problema no es lineal es habitual ¡transformarlo en uno lineal!

Seguramente conozcas ya un ejemplo destacado del uso de esta estrategia: el Cálculo Diferencial.

Cálculo Diferencial con una variable

Cuando se busca la derivada de una función en un punto se está buscando la mejor aproximación lineal a la función cerca de ese punto. Por ejemplo, consideremos el comportamiento de la función cuadrática x^2 cerca de $x = 1$, en que toma el valor 1. Si nos desplazamos desde 1 a $1 + h$, encontramos el valor

$$(1+h)^2 = 1 + 2h + h^2.$$

Para valores pequeños de h, es decir, cerca de 1, tenemos

$$(1+h)^2 \approx 1+2h,$$

y el incremento que sufre la función cuadrática es esencialmente igual a $2h$. Otra vez encontramos que los efectos (incrementos de la función) son proporcionales a las causas (incrementos en la variable independiente). La constante de proporcionalidad es 2, que es la derivada de x^2 en $x=1$. El Cálculo Diferencial de las funciones reales extiende estas consideraciones y las explota hasta generar una riquísima teoría, llena de hermosos resultados y aplicaciones de muy variada índole.

Cálculo Diferencial con varias variables

Ideas similares están por detrás del Cálculo Diferencial para funciones de muchas variables, y el Álgebra Lineal está presente. Ilustremos con un ejemplo. Supongamos que tenemos dos cantidades (f,g) que dependen de otras dos cantidades, dos variables, (x,y) siguiendo las reglas[5]

$$f(x,y) = 1+x^2+xy, \quad g(x,y) = (x+2y)^2.$$

Es decir, dados los valores de x e y todo lo que hay que hacer para conocer los valores de f y g es sustituir x e y es las fórmulas que acabamos de dar, y calcular. Por ejemplo, cuando x e y valen 1 y 2 respectivamente entonces f y g toman los valores 4 y 25.

¿Cuál es el incremento de f y g cuando x e y varían ligeramente desde $(1,2)$ a $(1+h,2+k)$? Hagamos los cálculos evaluando en estos nuevos valores.

$$\begin{aligned} f(1+h,2+k) &= 4+4h+k+h^2+hk, \\ g(1+h,2+k) &= 25+10h+20k+4hk+h^2+4k^2. \end{aligned}$$

Para valores pequeños de h y k los términos k^2, h^2 y kh son mucho más chicos que k y h, por lo que los incrementos Δf y Δg que sufren f y g respecto a los valores 4 y 25 pueden aproximarse por

$$\begin{aligned} \Delta f &= f(1+h,2+k) - 4 \approx 4h+k, \\ \Delta g &= g(1+h,2+k) - 25 \approx 10h+20k. \end{aligned}$$

[5]Decimos que f y g son funciones de las dos variables x e y. Hemos escogido expresiones polinómicas para que los cálculos fueran sencillos. Al igual que ocurre en el Cálculo Diferencial de una variable, estos cálculos pueden extenderse a una clase muchísimo más amplia de funciones.

Para justificar completamente estos cálculos hace falta incorporar algunas ideas geométricas a la discusión, e introducir una noción adecuada de límite para funciones de dos variables. No lo haremos aquí, de modo que invitamos al lector a aceptar momentáneamente nuestros argumentos informales.

Las conclusiones a las que hemos llegado pueden esquematizarse en la tabla

	h	k
Δf	4	1,
Δg	10	20,

en la que aparecen los coeficientes con los que los incrementos de h y k contribuyen a los incrementos Δf y Δg. La manera habitual de expresar estas relaciones es por medio de la matriz

$$\begin{pmatrix} 4 & 1 \\ 10 & 20 \end{pmatrix},$$

y una relación algebraica que devuelve el par de incrementos $(\Delta f, \Delta g)$ como el producto de la matriz por el vector (h, k) de incrementos en las variables. Se trata de una versión 2-dimensional de la fórmula que nos decía que el incremento de x^2 entre $x = 1$ y $x = 1 + h$ es esencialmente igual al producto

$$2 \times h$$

de la derivada en el punto 1 por el incremento de la variable independiente. La matriz desempeña el papel de la derivada en este cálculo.

0.1.6. Requisitos previos

Para poder avanzar en la lectura de este texto es necesario saber operar con los conjuntos numéricos $\mathbb{Q}$ y $\mathbb{R}$. Sólo hace falta manejar el álgebra en estos conjuntos: sumar, restar, multiplicar y dividir, que sigue las reglas comunes a cualquier cuerpo. En el apéndice A.2 aparece toda la información necesaria para este curso, y alguna más.

En algunas partes del texto realizaremos cálculos sobre el cuerpo $\mathbb{C}$ de los números complejos, y sobre el cuerpo $\mathbb{Z}_2$. No esperamos que estés familiarizado con estas estructuras. Encontrarás toda la información necesaria en el apéndice A.2.

Supondremos que conocés las propiedades más importantes del Cálculo Diferencial con funciones reales de una variable, y nos referiremos a ellas libremente. En algunos ejemplos utilizaremos el concepto de integral, que no supondremos conocido para tí, y para el que daremos alguna referencia a la amplia bibliografía existente sobre el tema.

Si tenés conocimientos previos de la Geometría Analítica del plano o del espacio te ayudarán a construir imágenes geométricas de los conceptos que desarrollaremos. Si no es así, encontrarás en el texto la información que necesites.

0.1.7. Sobre el uso de este material

Este texto fue concebido como referencia básica para el curso GAL1, y así lo usaremos en 2005. Está estructurado en cuatro capítulos: sistemas lineales, matrices, geometría y espacios vectoriales. Cada capítulo se abre con una introducción que describe el material que contiene. El texto completo se presentará en más de un tomo, presumiblemente dos, que irán apareciendo a lo largo del primer semestre de 2005. El volumen que estás leyendo es el primero de la serie.

Los capítulos están divididos en secciones. Cada sección está construida alrededor de una idea central y comienza con una breve introducción al material que allí se discute. Haciendo los énfasis adecuados, y seleccionando el material a presentar, lo esencial del contenido de cada sección debería poder cubrirse en el curso de una clase de teoría de hora y media de duración, y su correspondiente clase de práctico y tiempo de trabajo domiciliario.

La sección 2.5 dedicada a los determinantes es excepcional, porque es demasiado extensa, pero en su introducción se describen claramente cuáles son los puntos esenciales que esperamos que el lector asimile. Aparecen allí algunos resultados importantes para esa parte de la teoría cuya demostración sugerimos omitir en una primera lectura.

Cada sección termina con un punteo de las principales ideas y resultados que el lector debería asimilar. Te recomendamos no seguir avanzando antes de asegurarte de que estos objetivos han sido alcanzados.

Algunos ejemplos aparecerán una y otra vez a lo largo del texto. Como recomendación general te sugerimos trabajar ordenadamente y registrar tu trabajo, en particular los resultados de los ejercicios. Esta manera de trabajar seguramente te permitirá ahorrar esfuerzo, y lograr una mejor comprensión. Muchas ideas se van construyendo a lo largo de varios pasajes del texto, muchos problemas se tratan desde varias ángulos distintos y complementarios, y hay diversas referencias cruzadas. Algunos ejercicios contienen resultados preliminares, necesarios para la construcción de partes de la teoría. Los hay que se resuelven varias veces, con distintas técnicas. En algunos casos los resultados se reutilizan, y permiten resolver problemas nuevos prácticamente sin hacer nada.

La mayor parte de los ejercicios aparece en el cuerpo principal del texto. Para los ejercicios se usa una letra de un tamaño algo menor que la habitual. Los ejemplos finalizan dónde aparece el símbolo ♣, y las observaciones con ♠. El final de una demostración se indica así: □

En un apéndice hemos incluido algunos ejercicios algo más extensos y ambiciosos. Se trata de problemas que pueden dar lugar a un pequeño trabajo, preferentemente para realizar en grupo. En general combinan cuestiones de Geometría y/o Álgebra Lineal con ideas de alguna otra rama de la Matemática, para la resolución de un problema propio de otra disciplina. Esperamos que puedan servir a los docentes como guía para proponer actividades a sus estudiantes.

Hemos sido parcos en las citas a la bibliografía, pero cuando se da alguna referencia se hace con cierto detalle, invitando al lector a seguirla. En cualquier caso, digamos que hay una amplia literatura sobre los temas que aquí cubrimos, que incluye algunos textos excelentes. Te exhortamos con entusiasmo a consultar otras fuentes, descubrir otros enfoques y ampliar tus horizontes recorriendo críticamente las bibliotecas y la oferta de material disponible en Internet[6].

0.1.8. Sobre la preparación de este material

Este texto pretende ser un paso en la elaboración de un libro de Geometría y Álgebra Lineal, para el uso en los cursos de la Facultad de Ingeniería. En su redacción se tomó como base un conjunto preexistente de notas del curso [I], notas adicionales, repartidos de práctico, y enunciados de parciales y exámenes redactados por docentes del Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia"(IMERL).

La preparación del material en la forma que ahora tiene, se realizó por un equipo de docentes del IMERL integrado por Omar Gil, Nelson Möller, Juan Piccini y Cecilia Saint Martin, y que contó con el apoyo de las docentes Carolina Crisci, Marina Míguez, Ximena Otegui, Nancy Peré y Virginia Rodés, de la Unidad de Enseñanza de la Facultad de Ingeniería. Buena parte de este trabajo se hizo con financiación de un proyecto de la Comisión Sectorial de Enseñanza (CSE). El coordinador de este equipo fue Omar Gil. Otros colegas colaboraron con sus críticas y comentarios luego de leer versiones preliminares de algunas partes del texto y/o invirtiendo parte de su tiempo en intercambiar ideas con algunos de los miembros del equipo.

[6]Conviene redoblar el sentido crítico al recorrer internet, porque a un "click"del ratón se encuentrán allí tanto La Biblia como un calefón.

Es así que muchas personas participaron directa o indirectamente en la preparación de este material. Además de los ya nombrados, al menos[7] lo hicieron Laura Aspirot, Eduardo Canale, Marcelo Cerminara, Nelson Chaves, José Díaz, Pablo Fernández, Alfredo Jones, Juan José Manfredi, Roberto Markarian, Fernando Peláez, Alfredo Piria, Freddy Rabín, y Jana Rodríguez Hertz.

Al preparar este material fuimos haciendo diversas opciones que han dado al curso que aquí proponemos una fisonomía bastante diferente que el que se encuentra en las notas [I].

Respecto a [I] hay una diferencia en el tono general de la presentación: antes de introducir cualquier concepto hemos tratado de justificar ante el lector la necesidad de hacerlo. Por ejemplo, en vez de presentar las estructuras matemáticas confiando en que el desarrollo futuro de la teoría justificará para el lector su introducción, hemos intentado hacerlas aparecer como respuesta a problemas. Mostrando que cuanto más rica es la estructura que tengamos mejor podemos tratar e interpretar los datos de los problemas; las definiciones están casi siempre precedidas de alguna discusión que anticipa cómo deben formularse; algunas proposiciones y teoremas se demuestran primero, en el curso de una discusión, y luego los resultados se elaboran y ordenan en la forma de un enunciado preciso.

Este tipo de exposición puede ser más difícil de seguir que una más apegada a la estructura lógica del conocimiento. Pero creemos que es más adecuada para la comprensión de los conceptos.

Discutiremos en los próximos párrafos cuáles son las novedades que el lector encontrará en el contenido de los capítulos 1 y 2 del texto, que componen este volumen, lo que nos servirá además para brindar una breve descripción de ellos. Haremos lo propio con los capítulos siguientes, cuando los entreguemos al lector.

Capítulo 1

En el capítulo 1 hemos mantenido la propuesta de comenzar discutiendo los sistemas de ecuaciones lineales y el método de Gauss. Sin embargo, hemos incluido una sección preliminar en la que aparecen algunos problemas cuyo tratamiento conduce a construir un modelo matemático que implica la resolución de un sistema de ecuaciones lineales. Este tipo de información nos

[7] Es posible que falte el nombre de algún colega que hizo aportes para este trabajo. Desde ya pedimos disculpas por cualquier posible omisión, y pedimos a quien la detecte que nos comunique la información para corregirla.

pareció especialmente adecuada para un texto destinado a estudiantes de Ingeniería, a quienes presumimos especialmente interesados por las aplicaciones de la Matemática y su capacidad de construir buenos modelos de la naturaleza.

A partir de esta coincidencia inicial nuestra presentación diverge, ya que hemos decidido utilizar el tratamiento de los sistemas lineales para motivar la introducción de las nociones básicas del Álgebra Lineal. El espacio $\mathbb{K}^n$ de las n-uplas se introduce muy temprano, y luego se va avanzando en la construcción de la teoría hasta culminar con la noción de *rango*, que permite resumir en un único número mucha información acerca de un sistema lineal de ecuaciones dado. El vínculo entre estos conceptos y el cálculo con sistemas lineales es, todo el tiempo, el método de eliminación de Gauss. Más detalles de ese recorrido están esbozados en la introducción al capítulo 1, que complementa y extiende esta descripción. Nos ahorramos darlos aquí. Si todavía querés más información te decimos que el propio capítulo 1 contiene absolutamente todos los detalles acerca del capítulo 1.

Capítulo 2

Nuevamente coincidimos con las notas [I] en presentar las matrices en el capítulo 2. Pero también aquí hemos introducido una sección más o menos informal, con ejemplos en que las matrices se usan con diversos propósitos. Esperamos así ofrecerte una visión preliminar que acerque los temas a tratar a algunos de tus intereses personales. Vos podrás juzgar si hemos tenido éxito o no.

Discutimos luego el álgebra de matrices. En particular el producto de matrices. En nuestra presentación hemos escogido decir al lector qué se pretende con la definición del producto de matrices antes de formularla, en vez de dar la definición y anunciar que el desarrollo posterior de la teoría la justifica.

Hemos dado cierto peso a la descomposición LU de una matriz en dos factores, prácticamente equiparando la importancia de su tratamiento al cálculo de inversas. La razón es que las inversas de una matriz tienen gran interés teórico, pero ocupan un lugar menos destacado en el cálculo, una de las facetas más importantes del uso del Álgebra Lineal en la resolución de problemas tecnológicos. En cambio el cálculo con matrices, por ejemplo, para resolver numéricamente sistemas lineales, recurre permanentemente a la búsqueda de factorizaciones como la LU.

La descomposición LU es sólo un primer ejemplo en esta dirección. Se puede argumentar que los ejemplos más interesantes de factorizaciones son otros. Lo que es cierto. Pero requieren más estructura matemática, y su tratamiento debería hacerse más tarde. La descomposición LU es simple, y además depende del método de eliminación de Gauss, que ya será un viejo conocido a la hora de tratarla.

No hay un capítulo sobre determinantes en este texto, que se tratan en una única, extensa, sección. ¡Casi tan larga como el capítulo de determinantes en [I]! Hemos encontrado preferible enfatizar el cálculo matricial, de acuerdo a lo que ha sido la evolución histórica de la disciplina, y dejar los determinantes como una sección del capítulo de matrices. Encontrarás alguna información al respecto en las notas históricas.

Al igual que para el capítulo 1, si deseás más información sobre los contenidos del capítulo 2 podés consultar su introducción.

Una pequeña dosis de Matemática Discreta

Otra opción que hemos tomado es la de incluir ejemplos sobre cuerpos distintos del cuerpo $\mathbb{Q}$ de los números racionales y el cuerpo $\mathbb{R}$ de los números reales. La mayor parte del trabajo se hace sobre $\mathbb{R}$, pero también hay unos cuantos ejemplos sobre el cuerpo discreto $\mathbb{Z}_2$, que está formado por los dos números 0 y 1, con la regla de suma

$$1 + 1 = 0.$$

Una razón para trabajar de esta manera es que no hay que hacer casi ningún esfuerzo para incluir a $\mathbb{Z}_2$ y, muy brevemente, otros cuerpos $\mathbb{Z}_p$ en la discusión: la teoría depende de la estructura de cuerpo, no de cuál sea el cuerpo; la aritmética en $\mathbb{Z}_2$ es trivial y fácil de manejar.

Pero al tratar ejemplos en $\mathbb{Z}_2$ se abre una conexión muy natural y muy bonita hacia temas de matemática discreta que son interesantes en sí, y especialmente interesantes para los estudiantes de Computación[8] e Ingeniería Eléctrica, que tienen que trabajar mucho con *bits* y lógica: 0, 1; `verdadero`, `falso`. El cuerpo $\mathbb{Z}_2$ ofrece una estructura algebraica para manejarlos.

0.1.9. El texto, los estudiantes y nosotros

En la sección anterior hemos repasado parte del buen conjunto de opciones y decisiones que hay que tomar al preparar un curso, haciendo explícitas las más importantes y compartiéndolas con el lector. Creemos que esto ayudará a criticarlas y revisarlas. Para ratificarlas si son buenas, y para sustituir las malas por otras mejores. Esperamos que haya más de las primeras que de las segundas, pero por las dudas corresponde aclarar que el coordinador del grupo

[8] Aproximadamente la mitad de los estudiantes que ingresan cada año a esta Facultad se orientan hacia el área de Computación.

que preparó este material es el responsable por todos los desaciertos que pueda contener. También por las opiniones y recomendaciones que aparecen en esta introducción, por la mayoría de las erratas y errores, y por todos los chistes malos que esperan agazapados al lector para saltarle al cuello.

Más allá de cuál sea el resultado obtenido, este material fue hecho tratando de tener tan presentes a los estudiantes como nos fue posible. Preguntándonos muchas veces ¿cómo nos gustaría que nos contaran por primera vez estas ideas? Es un experimento tal vez condenado al fracaso, porque nuestra primera vez fue hace ya muchos años. Pero, como la intención de los autores es que este material siga madurando y mejorando, ahora que va a estar en manos de auténticos estudiantes, debutantes en la materia, aprovechamos la oportunidad para alentarte a formular todas las críticas y observaciones que creas conveniente hacernos llegar.

Este texto es para vos y tus compañeros, y sólo tiene sentido si les resulta de utilidad. Sea en esta edición, o en alguna venidera.

Cada ejemplar del libro está acompañado de una hoja de evaluación, que te invitamos a llenar y a devolver al IMERL una vez leído el material. Esperamos también que algunos de ustedes se tomen muy en serio la tarea de detectar y comunicar todo tipo de erratas, errores tipógraficos, etcétera. Se trata de errores pequeños, difíciles de detectar para los autores, que sólo pueden eliminarse con la ayuda de un corrector profesional, o de muchos lectores entusiastas. Para este trabajo hemos optado por esta última modalidad.

En el correr del primer semestre de 2005 colocaremos en la puerta del IMERL una urna para recibir las encuestas de evaluación. Cuando esto ocurra lo anunciaremos a través de la página web `http://imerl.fing.edu.uy/gal1`, a través de la que difundiremos información varia sobre el curso GAL1. Cualquier comunicación sobre este material puede entregarse en la Secretaría del IMERL, o enviarse por correo electrónico a `omargil@fing.edu.uy`. Todas serán calurosamente bienvenidas.

Las condiciones materiales en las que el curso se desarrollará no son buenas (seguramente ya lo hayas descubierto cuando leas estas líneas), por lo que depositamos cierta esperanza en que encontrarás en este texto la información que te perdiste cuando no oiste bien lo que el profesor decía porque estabas sentado muy lejos; o la respuesta a la pregunta que hubieras querido formular en un práctico.

Si no encuentran otra cosa, por lo menos que haya una palabra de aliento:

¡ánimo!

Los docentes podremos compartir contigo parte del trabajo, pero quizás te resulte escaso el acceso a nosotros. Tus compañeros estarán más cerca.

Apoyálos y buscá apoyo en ellos. Las ideas se entienden mejor cuando se comunican y se comparten con los demás, por lo que te recomendamos que trabajes en grupo, que discutas con otros tus resultados. Y se comprenden luego de trabajar con ellas, masticarlas, analizar sus consecuencias. Te recomendamos entonces que dediques el tiempo necesario para hacer los ejercicios, estudiar la teoría, consultar otras referencias y contrastar tus puntos de vista sobre lo que estás aprendiendo con otros estudiantes y los docentes[9].

Sólo nos queda decirte que hemos disfrutado mucho preparando este material. Esperamos de todo corazón que te resulte útil en caso de que dediques parte de tu tiempo a trabajar con él, y que te ayude en tu trabajo a lo largo del curso de Geometría y Álgebra Lineal 1.

Gracias por haber escogido acompañarnos en el viaje. La ciencia es una actividad muy hermosa, y confiamos en que el trayecto compensará con creces el esfuerzo que requiere. Éxito en el curso de Geometría y Álgebra Lineal 1, ¡y en todo lo que te propongas!

Montevideo, febrero de 2005

[9]También que no olvides el tiempo necesario para otras actividades. ¡Nos han comentado que hay otras cosas en la vida además de la Geometría y el Álgebra Lineal! Por ejemplo, el Cálculo Diferencial y l@s chic@s.

0.2. Índice de notaciones más corrientes

En esta sección presentamos al lector las notaciones que aparecen en el cuerpo del texto.

a_{ij}, b_{ij}, las entradas de matrices A y B respectivamente, o los coeficientes de un sistema lineal de ecuaciones;

$(a_{ij})_{j=1,\dots,n}^{i=1,\dots,m}$, matriz $m \times n$ con entradas a_{ij};

a_{ij}, b_{ij}, las entradas de matrices A y B respectivamente, o los coeficientes de un sistema lineal de ecuaciones;

$(a_{ij})_{j=1,\dots,n}^{i=1,\dots,m}$, matriz $m \times n$ con entradas a_{ij};

$a_1 a_2 \dots a_n$, el producto de los n números a_i;

$a_1 + a_2 + \dots + a_n$, la suma de los n números a_i;

A, en general designa matrices;

A_i, las columnas o las filas de una matriz A, según el contexto;

A_{ij}, la matriz adjunta del elemento a_{ij}, que se obtiene eliminando de A la fila i y la columna j;

$(A_1\ A_2\ \dots\ A_l)$, matriz con columnas o filas A_i;

A^{-1}, la matriz inversa de A;

A^t, la matriz traspuesta de A;

$A|B$, la matriz ampliada de un sistema lineal de ecuaciones;

b_j, los términos independientes de un sistema de ecuaciones lineales;

B, en general designa una matriz, frecuentemente es la columna con los términos independientes de un sistema de ecuaciones lineales;

$\mathbb{C}$, el cuerpo de los números complejos;

$\mathbb{C}^n$, es el espacio vectorial formado por las n-uplas de números complejos;

E, forma escalerizada de una matriz;

i, j, k, l, m, n, por lo general designan números naturales;

I, la matriz identidad. El tamaño de la matriz se desprende del contexto en el que el símbolo aparece;

$I_{n\times n}$, la matriz identidad de tamaño $n \times n$;

$\mathbb{K}$, un cuerpo cualquiera;

$\mathbb{K}^n$, el espacio vectorial formado por las n-uplas de números en un cuerpo $\mathbb{K}$;

$m \times n$, el tamaño de una matriz con m filas y n columnas;

$n \times n$, el tamaño de una matriz cuadrada, con n filas y columnas;

$\mathbb{N}$, el conjunto de los números naturales;

O, el vector nulo, o la matriz nula;

$\mathbb{Q}$, el cuerpo de los números racionales;

$\mathbb{Q}[\sqrt{2}]$, el cuerpo de los números reales de la forma $a+b\sqrt{2}$, con a y b racionales;

$\mathbb{R}$, el cuerpo de los números reales;

$\mathbb{R}^n$, el espacio vectorial formado por las n-uplas de números reales;

x, y, z, componentes de un vector de $\mathbb{R}^2$ o $\mathbb{R}^3$, o incógnitas de un sistemas con dos o tres incógnitas.

x_i, y_i, designan las entradas de un vector X o Y de algún $\mathbb{K}^n$, o las incógnitas de un sistema de ecuaciones lineales;

$(x_1, \dots, x_n)$, $(y_1, \dots, y_n)$, vectores de algún espacio $\mathbb{K}^n$;

X, Y, en general designan vectores de algún espacio $\mathbb{K}^n$;

$\mathbb{Z}$, el conjunto de los números enteros;

$\mathbb{Z}_2$, el cuerpo $\{0, 1\}$, con la suma y producto módulo 2. En esta estructura se satisface $1 + 1 = 0$;

$\mathbb{Z}_p$, el cuerpo $\{0, 1, \dots, p-1\}$, con las operaciones de suma y producto módulo el número natural p;

$\prod_{i=1}^n a_i$, el producto de los n números a_i que se obtienen cuando el índice i varía entre 1 y n.

$\sum_{i=1}^{n} a_i$, la suma de los n números a_i que se obtienen cuando el índice i varía entre 1 y n;

α, β, en general indican constantes escalares;

λ, una constante escalar. La mayor parte de las veces es un valor propio de una matriz.

λ_i, constantes escalares. Cuando aparecen indizadas suelen ser coeficientes en una combinación lineal.

0.3. Introducción a la segunda parte

Esta segunda parte contiene los capítulos de geometría y de espacios vectoriales del curso Geometría y Álgebra Lineal 1 (GAL1). En la redacción de estos capítulos hemos tenido en cuenta, además de los aportes de las personas mencionadas en la introducción a la primera parte, comentarios y sugerencias de Julio Borghi y Heber Enrich. Pero recordemos al lector que los desaciertos que contiene el texto son, de todos modos, responsabilidad del coordinador del equipo redactor.

0.3.1. Las novedades en este texto

Los lineamientos generales de este volumen son los que ya discutimos en la introducción general a la obra, pero nos interesa comentar aquí cuáles son las novedades que contienen los capítulos 3 y 4, dedicados al estudio de la geometría del espacio, y a la teoría abstracta de los espacios vectoriales, respectivamente.

Capítulo 3

Hemos optado por discutir la geometría de $\mathbb{R}^3$ directamente, sin intentar construir la geometría desde una presentación libre de coordenadas, basada, por ejemplo, en los postulados de la geometría euclidiana del espacio.

Este enfoque tiene importantes ventajas, que justifican que lo hayamos adoptado:

- es esencialmente autocontenido. Sólo se requiere cierta familiaridad con el sistema de los números reales para poder trabajar con la geometría de $\mathbb{R}^3$. Cierta preparación geométrica previa ayuda, pero no es indispensable;
- la complejidad técnica es muy baja, y las herramientas necesarias se reducen al cálculo con sistemas lineales y matrices;
- el contexto en el que se trabaja es el mismo que hemos usado al estudiar los sistemas lineales y matrices, y la geometría se vincula entonces naturalmente con los otros temas del curso. En particular, provee de una imagen muy útil para el tratamiento de los sistemas lineales;
- se generaliza con total naturalidad a cualquier espacio $\mathbb{R}^n$, y constituye una guía indispensable para calcular en problemas que involucran muchas variables;

- las propiedades métricas de $\mathbb{R}^3$ aparecen a través de la introducción del producto escalar, que induce una manera de medir distancias y ángulos. Este enfoque abre la posibilidad de introducir otras maneras de medir distancias y/o ángulos, que originen otras geometrías diferentes sobre el mismo conjunto $\mathbb{R}^3$. También induce a pensar con ideas geométricas a otros contextos, por ejemplo planos discretos construidos sobre la base de un cuerpo finito.

Esta última razón es muy importante, y verás aparecer frecuentemente en el futuro la geometría de la mano del cálculo.

Por supuesto, este enfoque tiene también algunas desventajas:

- requiere considerar un sistema de coordenadas, en principio ajeno a cualquier problema geométrico, para su interpretación;
- los elementos de $\mathbb{R}^3$ aparecen algunas veces como *puntos*, fijando una posición en el espacio; y otras veces como *vectores*, determinando una dirección;
- podría inducir a creer que cualquier terna de números debe ser siempre tratada como un vector.

Vale la pena hacer algunos comentarios al respecto.

Aunque no lo hagamos ahora, digamos que es posible construir un enfoque libre de coordenadas a partir de la geometría lineal de $\mathbb{R}^3$. De hecho, esto es lo que se hace en el capítulo 4: la teoría de los espacios de vectoriales es una teoría libre de coordenadas para los vectores. En este texto hemos preferido trabajar con problemas geométricos más cercanos a la intuición, antes de pasar a una teoría sin coordenadas, pero abstracta. Por otra parte, a la hora de calcular es casi siempre imprescindible escoger un sistema de coordenadas, y entonces puntos y vectores aparecerán representados como ternas de números, que requieren la misma interpretación geométrica que los elementos del espacio $\mathbb{R}^3$.

Para levantar definitivamente la confusión entre puntos y vectores, deberíamos introducir la noción de *espacio afín*. Otra vez hemos preferido posponer esta abstracción, y dejar que el contexto nos diga si tenemos que interpretar una terna de números como un punto o como un vector. La noción de espacio afín permite desarrollar toda la geometría, y liberarnos también de la representación en coordenadas. El lector interesado puede consultar el texto [H]. En la página 245 aparece la definición de espacio afín, en el capítulo 10 se introducen los espacios afines euclídeos, y se hace geometría en ese marco.

Tratar o no tratar una terna de números como un vector es una cuestión de escoger adecuadamente los modelos con los que se trabaja. En problemas de física, por ejemplo, no cualquier magnitud puede ser considerada un vector, y sólo merecen ese nombre las que se transforman frente a cambios de coordenadas siguiendo ciertas reglas bien determinadas. Pero tres números que no deben ser considerados un vector en el marco de una teoría física bien podrían ser tratados como vectores de un espacio vectorial en un problema de análisis de datos. Estos comentarios pueden parecer un poco confusos al lector, así que tratemos de aclarar la cuestión con una pregunta: ¿forman los números

$$1\ 0\ 1\ 1\ 0\ 1\ 1$$

un vector de $\mathbb{R}^7$? La respuesta es que el lector podrá considerarlos como un vector de $\mathbb{R}^7$ si ese modelo tiene algún sentido en algún contexto, pero también podría considerarlos como un vector de $\mathbb{Z}_2^7$. O tratarlos de otra manera, si resulta más conveniente.

Como resumen, digamos que el hecho de que dispongamos de una teoría en la que ternas de números tienen un sentido geométrico, no implica que esa teoría deba ser usada para interpretar cualquier cosa que se nos ponga por delante. Para decirlo de manera elocuente: a pesar de su evidente significado geométrico ¿pensarías en la terna 90-60-90 como en un vector de $\mathbb{R}^3$?

Los productos escalar y vectorial se introducen a partir de las componentes de los vectores de $\mathbb{R}^3$. La introducción del producto escalar está precedida de una discusión en la que analizamos cómo deben definirse las propiedades geométricas básicas para que teoremas clásicos como el teorema de Pitágoras, o el teorema del coseno, sean válidos en nuestra geometría. Para el producto vectorial recorremos el capítulo inverso: damos su definición por una fórmula basada en la representación de los planos, y recién después discutimos sus propiedades. Nos ha resultado difícil motivar la definición directamente a partir de as propiedades que debe tener este producto, y optamos por este camino. Puede ser un poco más difícil para el lector aceptar una definición cuyas propiedades quedarán claras recién más adelante, pero esta vía tampoco es ajena al desarrollo de la matemática: las definiciones aparecen en forma tentativa, se examinan sus consecuencias, se modifican, se cambian por otras mejores, y recién cuando una teoría está elaborada de adoptan definiciones definitivas. Reproducir este camino en el texto puede resultar confuso, y, por una vez, pedimos al lector que acepte una definición a cuenta de propiedades que demostraremos más adelante.

Hemos reducido mucho la presentación de ecuaciones de conos y cilindros, para enfatizar un aspecto importante de la geometría: curvas y superficies en el espacio son objetos que pueden caracterizarse por algunas ecuaciones, o que pueden recorrerse por medio de *parametrizaciones*. También hemos tratado de mostrar ejemplos y aplicaciones interesantes de las parametrizaciones más habituales de las esferas y los cilindros.

Además de analizar estas superficies tan importantes, pero tan particulares, también mostramos, a través de algunos ejemplos, que no es cierto que cualquier ecuación defina una superficie, ni que cualquier "parametrización" una curva en el plano o el espacio. No entramos en la discusión de los detalles, que son más bien propios de un curso de geometría diferencial, pero tratamos de dar al lector un panorama del tipo de problemas que se encuentran al elaborar una teoría general del curvas y superficies.

Capítulo 4

La estructura general de este capítulo es bastante cercana a la que tiene el tratamiento de los espacios vectoriales y las transformaciones lineales en las notas [I].

Sin embargo, el tratamiento de los espacios vectoriales se beneficia mucho de lo hecho en el capítulo 1 para los espacios $\mathbb{K}^n$. La teoría general tiene un desarrollo cercano al que dimos para el caso particular de los espacios $\mathbb{K}^n$, porque en el capítulo 1 tuvimos cuidado en presentar argumentos fácilmente generalizables. En particular, hemos puesto cuidado en mostrar que las técnicas de cálculo del capítulo 1 se adecuan perfectamente al tratamiento de problemas lineales en espacios vectoriales de dimensión finita cualesquiera, una vez que se representan los vectores del espacio por sus coordenadas respecto a una base.

En la parte de transformaciones lineales hemos introducido muy tempranamente la matriz asociada, y enfatizado la posibilidad de calcular por medio de esta matriz. La matriz asociada a una transformación lineal se hace aparecer directamente, utilizando la linealidad de la transformación y de la representación en coordenadas, y la importante caracterización del producto de una matriz por un vector como la combinación lineal de las columnas de la matriz con los coeficientes almacenados en el vector. Creemos que este enfoque facilita mucho la comprensión y el manejo de este conceptos.

Otra novedad es la inclusión de ejemplos, como problemas de interpolación y de ecuaciones diferenciales, que emplean la teoría en situaciones concretas, y el mayor énfasis en conectar los temas de este capítulo con los de capítulos anteriores. Por ejemplo, hemos dedicado más ejemplos y ejercicios a mostrar cómo la teoría de las transformaciones lineales permite tratar transformaciones geométricas como giros, simetrías y proyecciones.

Finalmente, hemos optado por terminar el curso con una sección dedicada al teorema de las dimensiones, del que ofrecemos al lector tres demostraciones diferentes. Nos pareció adecuado culminar la exposición con este resultado central, que admite además una interpretación muy clara en términos de los problemas de sistemas de ecuaciones lineales que iniciaron nuestra discusión.

0.3.2. Epílogo

Sólo nos queda recordar que este material no es un texto acabado, y que seguiremos trabajando sobre él para que estudiantes de próximas ediciones del curso cuenten con una referencia mejor para su trabajo. Esperamos entonces que participes de su elaboración, haciéndonos llegar tus comentarios y sugerencias. Reiteramos entonces la información que dábamos en la introducción a la primera parte de este texto, acerca de cómo ponerse en contacto con nosotros.

En el correr del primer semestre de 2005 colocaremos en la puerta del IMERL una urna para recibir las encuestas de evaluación. Cuando esto ocurra lo anunciaremos a través de la página web `http://imerl.fing.edu.uy/gal1`, a través de la que estamos difundiendo información varia sobre el curso GAL1. Cualquier comunicación sobre este material puede entregarse en la Secretaría del IMERL, o enviarse por correo electrónico a `omargil@fing.edu.uy`. Todas serán calurosamente bienvenidas.

Tal como decíamos en el comienzo de esta introducción, el Álgebra Lineal es una parte muy básica de la Matemática contemporánea, y seguramente la verás aparecer y reaparecer una y otra vez en tu carrera. Esperamos que este texto te resulte una referencia útil que te acompañe a lo largo de tu formación, y, por qué no, de tu vida profesional, y que puedas recurrir a él con confianza cuando necesites refrescar alguna de la información que contiene.

Montevideo, abril de 2005

0.4. Notaciones introducidas en la segunda parte

En esta sección presentamos al lector las notaciones que aparecen en el cuerpo del texto, y que no fueron introducidas en la lista de notaciones de la primera parte.

$\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}}$, sucesiones reales;

$\mathcal{A}, \mathcal{B}$, en general designan bases o generadores de espacios vectoriales, también representan familias de vectores cualesquiera;

$[\mathcal{A}]$, el espacio generado por $\mathcal{A}$;

$\mathcal{C}$, en general designa la base canónica de un espacio;

$\mathcal{C}(\mathbb{R})$, las funciones reales con variable real continuas;

$\mathcal{C}^1(\mathbb{R})$, las funciones reales con variable real que tienen derivada continua;

$\dim(\mathbb{V})$, la dimensión del espacio vectorial $\mathbb{V}$;

$d(X, Y)$, la distancia entre X e Y ;

e, versor, vector de longitud 1;

e_X, e_Y, versores en la dirección de X e Y respectivamente;

f, g, h, funciones cualesquiera;

$\mathcal{F}$, espacios de funciones;

$(\vec{i}, \vec{j}, \vec{k})$, la terna ortonormal directa con $\vec{i}$ en la dirección del eje Ox, $\vec{j}$ en la del eje Oy y $\vec{k}$ en la del eje Oz;

$\operatorname{im}(T)$, la imagen de una transformación lineal;

$\ker(T)$, el núcleo de una transformación lineal;

$M^{m\times n}(\mathbb{K})$, el espacio de matrices $m \times n$ con entradas en el cuerpo $\mathbb{K}$;

N, el vector normal a un plano;

O, el vector nulo de un espacio vectorial;

$O_{\mathbb{V}}$, el vector nulo del espacio vectorial $\mathbb{V}$;

P, Q, puntos del espacio $\mathbb{R}^3$ o $\mathbb{R}^2$;

$P_{\mathbb{S}}Y$, la proyeción ortogonal de Y sobre $\mathbb{S}$;

r, s, rectas en el espacio;

$\mathbb{R}[x]$, polinomios reales;

$\mathbb{R}_n[x]$, polinomios reales de grado menor o igual que n;

$\mathbb{S}, \mathbb{T}$, subespacios vectoriales;

$\mathbb{S} \cap \mathbb{T}$, la intersección de los subespacios $\mathbb{S}$ y $\mathbb{T}$;

$\mathbb{S} + \mathbb{T}$, la suma de los subespacios $\mathbb{S}$ y $\mathbb{T}$;

$\mathbb{S} \oplus \mathbb{T}$, la suma directa de los subespacios $\mathbb{S}$ y $\mathbb{T}$;

$T_{\mathbb{S}}$, restricción de T al subespacio $\mathbb{S}$;

$T : \mathbb{V} \to \mathbb{W}$, transformación lineal del espacio $\mathbb{V}$ al espacio $\mathbb{W}$;

Tv, $T(v)$, ambas notaciones tienen el mismo significado. La imagen del vector v por la transformación lineal T;

$(Tv)_{\mathcal{B}}$, coordenadas del transformado de v en la base $\mathcal{B}$;

${}_{\mathcal{B}}T_{\mathcal{A}}$, matriz asociada a T en las bases $\mathcal{A}$ y $\mathcal{B}$;

T^{-1}, la inversa de la transformación T;

$T \circ S$, la composición de las transformaciones lineales T y S;

u, v, w, los vectores genéricos de algún espacio vectorial;

U, V, W, X, Y vectores del espacio $\mathbb{R}^{\nvDash}$ o $\mathbb{R}^2$;

$U \wedge V$, el producto vectorial de U y V;

$v_{\mathcal{A}}, (v)_{\mathcal{A}}, \mathrm{coord}_{\mathcal{A}}(v)$, todas las notaciones tienen el mismo significado. Las coordenadas del vector v en la base $\mathcal{A}$;

$(v_1, \ldots, v_n)$, familia ordenada de n vectores de un espacio vectorial $\mathbb{V}$;

$\mathbb{V}, \mathbb{W}$, espacios vectoriales;

$(\mathbb{V}, \mathbb{K}, +, \cdot)$, espacio vectorial $\mathbb{V}$ sobre un cuerpo $\mathbb{K}$, con operaciones $+$ y $\cdot$ de suma, y producto por un escalar, respectivamente;

$|X|$, la longitud del vector X;

$X \cdot Y$, $\langle X, Y\rangle$, ambas notaciones tiene el mismo significado. El producto escalar de X e Y;

(X, Y, Z), el producto mixto de los vectores X, Y y Z;

λ, μ, parámetros en la descripción paramétrica de rectas y planos. También indican escalares cualesquiera.

λ_i, suelen ser coeficientes en una combinación lineal;

π, planos en el espacio;

Índice general

Capítulo 1

Sistemas de ecuaciones lineales

Gran parte de lo que estudiaremos en este texto se vincula de alguna manera con el estudio de *sistemas de ecuaciones lineales*, por lo que hemos dedicado este primer capítulo a su estudio. Los sistemas de ecuaciones lineales tienen en este texto un triple papel:

1. aparecen en primer lugar como *modelos* para distintos problemas, muchos de ellos propios de la Ingeniería. Invitamos al lector a recorrer los ejemplos que aparecen en la sección 1.1;

2. nos permitirán introducir las nociones fundamentales del *Álgebra Lineal* a partir del estudio de sus soluciones, y de los problemas de compatibilidad y determinación. Realizar ese recorrido es, en realidad, el corazón de este primer capítulo;

3. serán una herramienta a la que tendremos que recurrir frecuentemente a lo largo de este texto. A la hora de calcular, en gran variedad de dominios, se términa reduciendo los problemas a la resolución de sistemas de ecuaciones lineales. Los reencontraremos al tratar problemas de Geometría en el Capítulo 3, o problemas de Álgebra Lineal en los contextos más o menos abstractos que proponemos en el capítulo 4.

Pasamos ahora a describir brevemente el contenido de cada una de las secciones de este capítulo.

La primera de ellas, sección 1.1, está dedicada a mostrar al lector unos cuantos problemas que conducen al estudio de sistemas lineales de ecuaciones cuando se les formula matemáticamente.

En la sección 1.2 desarrollamos la técnica básica para la resolución de sistemas de ecuaciones lineales: el procedimiento de *escalerización* o de *eliminación* de Gauss. Se trata de un procedimiento sistemático, algorítmico, adecuado para el tratamiento de sistemas que tienen muchas variables y muchas ecuaciones. En esa misma sección introducimos la notación matricial para los sistemas, y la interpretación vectorial de los coeficientes e incógnitas del sistema. La introducción de matrices y vectores, que sumerge el estudio de los sistemas de ecuaciones lineales dentro de estructuras geométricas y algebraicas que serán de gran utilidad para analizarlos, será desarrollada a lo largo de todo el texto.

La sección 1.3 enfatiza el hecho de que un sistema de ecuaciones lineales puede interpretarse como una única ecuación vectorial en términos de las columnas de coeficientes que multiplican a cada una de las incógnitas del sistema. Este punto de vista nos ayudará a entender la *compatibilidad* o *incompatibilidad* de los sistemas de ecuaciones lineales, y nos llevará a introducir las importantes nociones de *combinación lineal* y de *subespacio generado* por las columnas de una matriz.

En la sección 1.4 avanzamos un poco más sobre la estructura algebraica asociada con los sistemas de ecuaciones lineales, y buscamos dar buenas descripciones de los subespacios generados por las columnas de la matriz de un sistema. Aparece así la noción de *generador*, a la que siguen en la sección 1.5 las de *independencia lineal* y *base*, y en la sección 1.6 la de *dimensión*.

Un sistema lineal compatible puede tener una única solución o muchas. Toda la teoría que se desarrolló en las secciones 1.4, 1.5 y 1.6 para tratar de comprender cuándo un sistema lineal es compatible o incompatible por medio del análisis del espacio de columnas de la matriz del sistema, se utiliza en la sección 1.7 con el objetivo de estudiar este problema de *determinación* o *indeterminación*. La respuesta está encerrada en otro subespacio asociado con la matriz del sistema: su *núcleo*.

La sección 1.8 es una especie de remate de toda la teoría que se expone en el capítulo. Se introduce la importante noción de *rango*, y se muestra que el comportamiento de un sistema de ecuaciones lineales está esencialmente determinado por el conocimiento de tres números: los dos primeros son el número de ecuaciones y el número de incógnitas, variables evidentes a primera vista; el tercero es el rango de la matriz del sistema.

Queremos advertir al lector que el desarrollo histórico del Álgebra Lineal no corresponde al recorrido que en este capítulo le proponemos. El tratamiento de los sistemas lineales condujo primero al concepto de *determinante*, y la estructura vectorial, aunque tiene gran importancia en nuestra época, es de aparición más tardía. Damos alguna información adicional al respecto en las notas históricas de las secciones dedicada a los determinantes y a la teoría abstracta de los espacios vectoriales. En cualquier caso, esperamos que el camino que hemos escogido para nuestra presentación resulte sugerente y atractivo.

1.1. Problemas que conducen a la consideración de sistemas lineales de ecuaciones

En esta sección presentamos algunas situaciones cuyo análisis nos lleva a formular y resolver sistemas de ecuaciones lineales. El espíritu es, básicamente, mostrar ejemplos en los que la modelización de un problema nos conduce a un sistema de ecuaciones lineales y a estudiar cómo son sus soluciones.

Por *modelización de un problema* entendemos la identificación de las principales variables en juego, y de las relaciones entre ellas. En este curso nos interesará especialmente el caso en que las variables pueden representarse por números de algún tipo, y sus relaciones puedan expresarse por medio de ecuaciones lineales que las vinculan.

No esperamos que el lector se familiarice con los detalles de la modelización, ya tendrá otros cursos en los que trabajar sobre esto. Tampoco esperamos que cada lector se interese en todas las aplicaciones que mostramos. Pero confiamos en que todos los lectores encontrarán que alguna de ellas es cercana a los problemas que más lo motivan y despiertan sus deseos de comprender más y mejor.

Entre nuestros ejemplos aparecerán problemas de ingeniería que discutiremos con cierto detalle, como

- la determinación del flujo de la circulación sobre una red vial;
- el estudio de circuitos eléctricos;
- el análisis de los esfuerzos a los que está sometida una estructura.

Esta breve lista puede extenderse para incluir además el cálculo de los insumos necesarios para la fabricación de bienes y/o el suministro de servicios (ver el ejercicio 2.1, en la página 161); la codificación y decodificación de la información para protegerla contra errores en su transmisión o almacenamiento (ver el ejercicio 1.28, en la página 60); el ordenamiento de una lista de datos según su importancia (ver, por ejemplo, el artículo de divulgación [M], o, el algo más risueño, [F]); el estudio de la dinámica de poblaciones (ver el ejercicio 1.10 en la página 21 de esta misma sección). Y podríamos continuar.

Ser exhaustivos sería imposible porque, seguramente dejaríamos fuera alguna aplicación relevante. También las aplicaciones que hoy no existen pero que aparecerán en el futuro. El lector seguramente ampliará la lista de ejemplos a lo largo de sus estudios y de su vida profesional.

No es el principal objetivo de esta sección la resolución de los sistemas lineales que aparecen, tarea que, de todos modos, dejamos planteada en algunos casos. Será estupendo si conseguimos resolverlos ya e interpretar los resultados en el contexto de nuestros modelos. Pero si algún sistema se resiste, o parece demasiado grande o complejo para abordarlo con las técnicas de que el lector dispone en este momento, puede ser una buena idea posponer el cálculo de las soluciones para después de la lectura de la sección 1.2.

1.1.1. Problemas de revistas de entretenimientos

El primer ejemplo es puramente recreativo: la edad de Pedro es el triple de la edad de Ana, pero dentro de 10 años su edad sólo será el doble de la de Ana. ¿Cuáles son las edades de Pedro y Ana?

Si a las edades de Ana y Pedro, expresadas en años, las llamamos respectivamente a y p, estos números deben satisfacer las dos relaciones

$$p = 3a, \quad p + 10 = 2(a + 10),$$

que son equivalentes al sistema de ecuaciones

$$\left\{ \ p - 3a = 0, \quad p - 2a = 10. \right. \tag{1.1}$$

Las variables a y p intervienen en forma lineal en estas ecuaciones, porque sólo aparecen multiplicadas por algunos coeficientes y luego sumadas (no hay expresiones como a^2 o $\log p$ en juego), por lo que decimos que (1.1) es un *sistema de ecuaciones lineales*. Este sistema tiene sólo dos incógnitas, y resulta de sencilla solución. Restando la segunda ecuación en (1.1) de la primera obtenemos

$$-a = -10,$$

que inmediatamente arroja el resultado

$$a = 10.$$

Sabemos entonces que Ana tiene 10 años, y concluimos rápidamente que Pedro tiene 30.

Ejercicio 1.1. Una madre es 21 años mayor que su hijo, y dentro de 6 años su edad será exactamente 5 veces del niño. ¿Dónde está el papá del pequeño? ♣

1.1.2. Circulación sobre una red vial

Veamos ahora un modelo para el estudio del tránsito. El mapa que aparece en la figura 1.1 muestra algunas de las calles del centro de Montevideo hacia el año 2050, con los nuevos nombres que tendrán por esas fechas. El sentido de

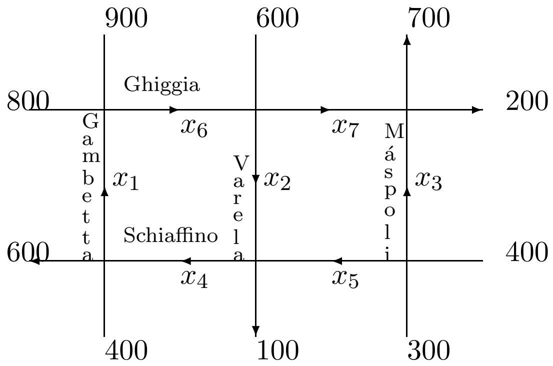


Figura 1.1: Mapa de parte del centro de Montevideo al 2050

circulación en cada una de ellas está indicado por las flechas. El mapa indica el flujo de tránsito que entra o sale a cada calle, en unidades de vehículos por hora.

Como el flujo de tránsito varía considerablemente durante el día, supondremos que los números mostrados representan el flujo de tránsito promedio a la hora de mayor circulación.

Algunas obras de reparación dificultarán la circulación en la calle Schiaffino entre Varela y Gambetta. ¿Es posible cortar completamente el tráfico allí y atender la demanda que plantea la circulación de vehículos en la hora pico? Si no es posible, ¿qué medida es conveniente adoptar para minimizar el tráfico por esa calle?

Intentemos comprender cómo es la circulación de vehículos en este trozo de la red vial. Para ello buscaremos calcular cómo puede distribuirse el tráfico en las calles de la ciudad que aparecen en el mapa, de modo de satisfacer las demandas de vehículos que entran y salen por cada calle. Introducimos variables x_i que representán la cantidad de vehículos por hora que circulan por cada una de las cuadras en la parte de la ciudad que estamos considerando, como se muestra en la tabla:

variable,	vehículos por hora que circulan por,
x_1,	Gambetta desde Schiaffino hacia Ghiggia,
x_2,	Varela desde Ghiggia hacia Schiaffino,
x_3,	Máspoli desde Schiaffino hacia Ghiggia,
x_4,	Schiaffino desde Varela hacia Gambetta,
x_5,	Schiaffino desde Máspoli hacia Varela,
x_6,	Ghiggia desde Gambetta hacia Varela,
x_7,	Ghiggia desde Varela hacia Máspoli.

En cada intersección el tráfico de entrada debe ser igual al de salida, no aparecen ni desaparecen misteriosamente autos en ninguna esquina, de modo que las circulaciones en cada cuadra deben satisfacer ecuaciones que reflejan esta propiedad. Por ejemplo, a la esquina de Schiaffino y Gambetta llegan cada hora x_4 vehículos por Schiaffino, 400 por Gambetta, y salen x_1 vehículos por Gambetta y 600 por Schiaffino. La ecuación que corrresponde a esta intersección es entonces

$$-x_1 + x_4 - 600 + 400 = 0,$$

que puede ser simplificada en

$$-x_1 + x_4 = 200.$$

Razonando en forma similar para todas las intersecciones concluimos que los valores de las siete circulaciones x_i deben satisfacer el sistema

$$\left\{\begin{aligned} -x_1 + x_4 &= 200,\\ x_2 - x_4 + x_5 &= 100,\\ x_3 + x_5 &= 700,\\ x_1 - x_6 &= 100,\\ x_2 - x_6 + x_7 &= 600,\\ x_3 + x_7 &= 900, \end{aligned}\right.$$

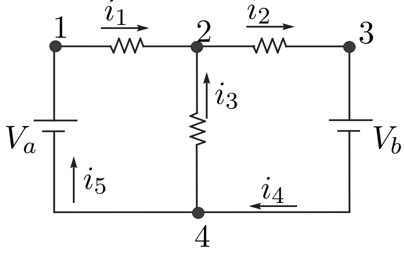
formado por las ecuaciones de las seis intersecciones que aparecen en el plano de la figura 1.1. Las soluciones de este sistema representan las posibles maneras de ordenar el flujo de vehículos en estas calles.

Si no estamos dispuestos a cambiar el sentido de circulación en ninguna calle entonces tendremos que considerar sólo las soluciones en que las siete variables x_i son mayores o iguales que cero.

Pospondremos el análisis de este ejemplo hasta tener algunas herramientas que nos permitan trabajar ordenadamente con las variables de un sistema de seis ecuaciones con siete incógnitas. Los cálculos se desarrollan en el ejemplo 1.2.27, página 50.

1.1.3. Voltajes y corrientes en un circuito

Vamos a estudiar cómo se distribuyen las corrientes y cuáles son los voltajes en el circuito de la figura. Hay dos tipos de magnitudes a determinar: las cuatro intensidades i_k, de las corrientes en los arcos,



y los cuatro voltajes v_k en los nodos. El
índice k toma valores entre 1 y 4.

¿Qué principios gobiernan los valores de estas variables? Debemos recurrir a otra disciplina, en este caso a la física, para hallar la respuesta. Esta viene en la forma de las leyes de Kirchoff y Ohm, que permiten construir un *modelo matemático* del circuito.

Ley de Kirchoff[1]. Es un principio de conservación, que asegura que no se genera ni se destruye carga en los nodos. Por lo tanto lo que entra en cada nodo debe ser igual a lo que sale. La suma de las intensidades en los arcos que confluyen en un nodo dado debe ser igual a cero. Entonces las cuatro ecuaciones que corresponden a los cuatro nodos del circuito son

$$\begin{aligned} i_5 - i_1 &= 0, \\ i_1 - i_2 + i_3 &= 0, \\ i_2 - i_4 &= 0, \\ -i_3 + i_4 - i_5 &= 0, \end{aligned} \tag{1.2}$$

que corresponden respectivamente a los nodos 1, 2, 3 y 4. En este cálculo debemos sumar las intensidades en los arcos dirigidos hacia el nodo, y restar la de los arcos que salen del nodo.

Ley de Ohm[2]. Las intensidades en los arcos no pueden ser cualesquiera y están determinadas por las diferencias de potencial entre los extremos de los arcos. La relación entre diferencia de potencial e intensidades debe cumplir la ley de Ohm. Esta expresa que la intensidad en el k-ésimo arco es igual a la caida de potencial entre sus extremos, dividida la resistencia R_k en el arco. Al tener en cuenta en qué nodo comienza y termina cada uno de los arcos en los que el circuito con el que estamos trabajando tiene una resistencia obtenemos

$$\begin{aligned} R_1 i_1 &= v_1 - v_2, \\ R_2 i_2 &= v_2 - v_3, \\ R_3 i_3 &= v_4 - v_2. \end{aligned} \tag{1.3}$$

Observación 1.1.1. La ley de Kirchoff considerada para los circuitos representa el mismo principio aplicado en el ejemplo del tránsito: **lo que entra en un nodo es igual a lo que sale**.

Se trata en realidad de un principio de conservación muy general, aplicable a redes de muy diversa naturaleza. Se satisface, por ejemplo, para los vehículos en una red vial y para la masa de fluido en una red de cañerías. Sin embargo, salvo en situaciones excepcionales en las que el esquema de conexiones de la

[1] Debe su nombre a Gustav Robert Kirchhoff, 1824-1887

[2] En honor a Georg Simon Ohm, 1789-1854

red es trivial, el sistema de ecuaciones que resulta de aplicar solamente estos principios de conservación resulta ser indeterminado.

Es decir, hay muchos flujos posibles sobre la red que satisfacen el principio de conservación. Por lo tanto no es posible determinar sólo a partir de la ley de Kirchoff de qué manera operará la red. En el caso de la red vial de la subsección 1.1.2 este hecho se pone en evidencia en el ejemplo 1.2.27. Ver también el ejercicio acerca de las ecuaciones 1.2

La consideración de los potenciales v resuelve este problema: al incorporarlos al modelo y relacionarlos con las intensidades por medio de la Ley de Ohm los potenciales permiten determinar completamente el flujo sobre la red y conocer la intensidad de corriente que circula por cada conexión. ♠

Las caídas de potencial entre los nodos 3 y 4, y entre 1 y 4 están determinadas por las dos fuentes:

$$\begin{aligned} V_a &= v_1 - v_4, \\ V_b &= v_4 - v_3. \end{aligned} \tag{1.4}$$

Las ecuaciones primera y tercera en (1.2) implican que $i_1 = i_5$ e $i_2 = i_4$. Combinando estas igualdad con (1.3) obtenemos

$$i_1 = i_5 = \frac{v_1 - v_2}{R_1}, \quad i_2 = i_4 = \frac{v_2 - v_3}{R_2}, \quad i_3 = \frac{v_4 - v_2}{R_3}.$$

Sustituimos estas expresiones de las intensidades en la segunda y cuarta ecuación de (1.2) y usamos (1.4) para obtener el siguiente sistema de ecuaciones sobre los voltajes:

$$\begin{aligned} \tfrac{1}{R_1}v_1 + \left(-\tfrac{1}{R_1} + \tfrac{1}{R_2} + \tfrac{1}{R_3}\right)v_2 - \tfrac{1}{R_2}v_3 - \tfrac{1}{R_3}v_4 &= 0 \\ \tfrac{1}{R_1}v_1 - \left(\tfrac{1}{R_1} + \tfrac{1}{R_2} + \tfrac{1}{R_3}\right)v_2 + \tfrac{1}{R_2}v_3 + \tfrac{1}{R_3}v_4 &= 0 \\ v_1 - v_4 &= V_a, \\ v_4 - v_3 &= V_b. \end{aligned} \tag{1.5}$$

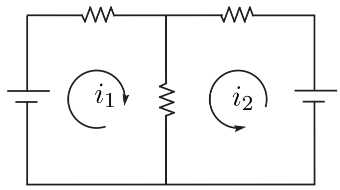
A partir de este sistema de cuatro ecuaciones con cuatro incógnitas intentaremos determinar los voltajes y las intensidades. Dejamos planteado el ejercicio al lector para un caso particular en el que hemos fijado los valores de las resistencias y de las diferencias de potencial introducidas por las fuentes.

Ejercicio 1.2. Tomando $V_a = V_b = 12V$, $R_1 = 1\Omega$, $R_2 = 2\Omega$ y $R_3 = 3\Omega$, determinar voltajes e intensidades en el circuito que estamos considerando. Interpretar los resultados.

Observación 1.1.2. En la construcción de un modelo matemático para un problema la elección adecuada de las variables ayuda a simplificar y sistematizar la deducción y escritura de las ecuaciones de los circuitos. Veremos un ejemplo de esto en la próxima sección, en que la introducción de las *corrientes de malla*

Más adelante, en el capítulo 2 destinado a las matrices, veremos como una buena manera de representar y ordenar las variables también es de gran ayuda. Encontraremos allí que la expresión de las ecuaciones en términos de corrientes de malla se reduce a determinar la *matriz de impedancias* del circuito (ver la página 173). Aprenderemos también que las leyes de Ohm y de Kirchoff admiten sencillas expresiones matriciales cuando un circuito se describe en términos de la *matriz de incidencia* que recoge la estructura de sus conexiones (ver el ejemplo 2.1.6, en la página 172). ♠

Corrientes de malla



Vamos a mostrar ahora una manera de sistematizar y simplificar la deducción de las ecuaciones de un circuito por medio de la introducción de otras variables, las llamadas *corrientes de malla*. Consideraremos dos corrientes i_1 e i_2 tal como se ilustra en la figura. Estas corrientes corresponden a las dos *mallas* formadas por las conexiones $(1, 2)$, $(2, 4)$ y $(4, 1)$ en el circuito, y por $(2, 3)$, $(3, 4)$ y $(4, 2)$. Sobre la conexión $(4, 2)$ entre los nodos 4 y 2, que es común a las dos mallas en el circuito, la corriente toma el valor $i_1 + i_2$, lo que asegura que se satisface la ley de Kirchoff de conservación de la corriente en los nodos.

La suma total de diferencias de potencial a lo largo de una malla debe ser nula. Es decir, se satisfacen las obvias igualdades

$$0 = (v_1 - v_2) + (v_2 - v_4) + (v_4 - v_1) = (v_2 - v_3) + (v_3 - v_4) + (v_4 - v_2).$$

Las diferencias entre nodos que aparecen en la fórmula anterior pueden sustituirse por los valores que proporcionan la ley de Ohm y los valores de las fuentes. Concluimos entonces que

$$\begin{aligned} R_1 i_1 + R_3(i_2 + i_1) - V_a &= 0, \\ R_2 i_2 - V_b + R_3(i_2 + i_1) &= 0. \end{aligned}$$

Estas ecuaciones pueden ordenarse un poco mejor en las incógnitas i_1 e i_2:

$$\begin{aligned} (R_1 + R_3)i_1 + R_3 i_2 &= V_a, \\ R_3 i_1 + (R_1 + R_3)i_2 &= V_b. \end{aligned}$$

Obtenemos así un sistema lineal de ecuaciones en estas variables, cuya solución nos permitirá conocer los valores de las corrientes.

Ejercicio 1.3. Resolver las ecuaciones para las corrientes de malla con los valores particulares con que trabajamos en el ejercicio 1.2. Determinar los valores de todas las corrientes y los voltajes en el circuito.

1.1.4. Flujos sobre redes

Presentamos en esta sección un modelo relativamente general para los flujos sobre una red. Todas las redes admiten la siguiente descripción: un conjunto de cosas a unir o relacionar, **nodos**, algunos de los cuales se interconectan entre sí, por **aristas** o **arcos**. Los nodos pueden ser uniones en una estructura, computadoras, elementos de un conjunto, esquinas de la ciudad, etcétera. Y las aristas pueden representar barras, cables, relaciones de pertenencia, posibles llamadas telefónicas, rutas, etcétera.

En muchas aplicaciones algo fluye sobre las aristas de la red: electrones, vehículos, petróleo, bienes, información, etcétera, y es de interés determinar cómo se distribuye este flujo. Por ejemplo, en una red eléctrica interconectada, el problema de calcular las corrientes en cada una de sus líneas a partir del conocimiento de las demandas y la generación en cada nodo de la red es básico para su diseño y operación.

También, en sistemas de distribución en los que varios proveedores de energía suministran electricidad a diferentes usuarios hay que decidir cómo repartir el costo de las líneas entre los distintos agentes que operan sobre la red. A diferencia de lo que ocurre en una carretera, en la red eléctrica es imposible distinguir la cantidad enviada por un agente de la enviada por otro, por lo que la determinación del peaje a pagar por cada usuario es más complicada que sobre una red vial.

Un método posible para hacerlo consiste en determinar la proporción de cada línea que cada usuario emplea, lo que hace necesario conocer cómo se distribuye sobre la red el resultado de extraer en el nodo en el que se encuentra el usuario la energía inyectada por un generador ubicado en otro nodo de la red. En lo que sigue estudiaremos el problema más general de determinar el flujo que produce sobre la red cualquier combinación de demandas en sus nodos.

En lo que sigue discutiremos un modelo general para los flujos sobre una red que debe satisfacer ciertas demandas en sus nodos.

En un problema real los números nodos y aristas pueden ser grandes. Ejemplificaremos con una red de cinco nodos y siete aristas, con el esquema de conexiones que se muestra en la figura, pero nuestros argumentos serán completamente generales.

Las demandas en los nodos representan algo que se agrega o se extrae

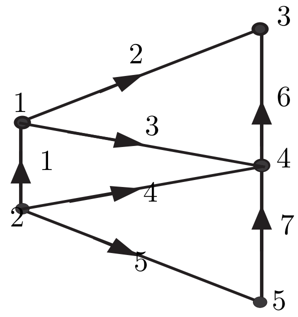


Figura 1.2: Una red con cinco nodos y siete arcos.

de la red: potencia eléctrica generada o consumida, producción o consumo de bienes, inyección o extracción de fluido, etcétera. En nuestra red de cinco nodos hay cinco demandas d_k, $k = 1, \ldots, 5$. Adoptaremos la convención de que las demandas son positivas inyectemos o agreguemos algo en un nodo, y negativas cuando lo extraigamos.

El problema que vamos a tratar admite la siguiente formulación general: dada la red y un vector de demandas $(d_1, d_2, d_3, d_4, d_5)$ determinar los corrientes en cada una de las aristas o arcos. Hemos fijado (arbitrariamente) un sentido positivo en cada arco. Los flujos en la dirección del sentido prefijado tendrán signo positivo, y serán negativas en caso contrario. En resumen, los flujos en las aristas, y las demandas en los nodos tienen signo, según vayan o vengan, entren o salgan.

Ya tenemos la descripción general de la red y de las demandas en cada nodo. Asumiremos que en cada nodo debe satisfacerse la ley de conservación de Kirchoff, que, en este ejemplo, se traduce en las ecuaciones

$$\begin{aligned} i_1 - i_2 - i_3 &= d_1, \\ -i_1 - i_4 - i_5 &= d_2, \\ i_2 + i_6 &= d_3, \\ i_3 + i_4 - i_6 + i_7 &= d_4, \\ i_5 - i_7 &= d_5, \end{aligned} \tag{1.6}$$

donde hemos llamado i_k al flujo en el arco k. Los números d_k representan, tal como mencionamos antes, las demandas en los nodos.

Salvo en situaciones triviales, la Ley de Kirchoff es insuficiente para determinar los flujos sobre la red. Asumiremos un principio adicional, análogo a la Ley de Ohm para las corrientes sobre los circuitos: sobre los nodos podemos medir un *potencial* y los flujos en un arco son proporcionales a la diferencia de potencial entre sus extremos. El potencial puede ser un voltaje en el caso de que consideremos una red de corriente continua, una presión en un problema de mecánica o un precio en una cuestión de economía: introduce en el sistema una fuerza que genera el flujo.

Esta hipótesis acerca de los potenciales se traduce para cada arco en una ecuación de la forma

$$i_k = c_k(\Delta v)_k,$$

donde i_k es la corriente en el arco k, $(\Delta v)_k$ la diferencia de potencial entre los nodos en que el arco comienza y termina, y c_k es la constante de proporcionalidad que vincula, para ese arco, el flujo con la variación del potencial. Si c_k es chico hace falta una gran caída de potencial para producir un flujo grande. Cuando c_k es grande ocurre lo contrario. La constante c_k mide entonces que tanto permite el arco k el flujo a lo largo de él, y nos referiremos a ella como la *conductividad* del arco. En un circuito, la conductividad c_k de un arco es el inverso de su resistencia. Supondremos que las constantes c_k son on características conocidas de la red.

Las diferencias $(\Delta v)_k$ dependen del estado de operación de la red y corresponden a la resta de los potenciales en los extremos del arco k. Al tener en cuenta en que nodo comienza y termina cada arco de la red con la que estamos trabajando obtenemos las siguientes ecuaciones para los arcos:

$$\begin{aligned}
i_1 &= c_1(v_2 - v_1),\\
i_2 &= c_2(v_1 - v_3),\\
i_3 &= c_3(v_1 - v_4),\\
i_4 &= c_4(v_2 - v_4),\\
i_5 &= c_5(v_2 - v_5),\\
i_6 &= c_6(v_4 - v_3),\\
i_7 &= c_7(v_5 - v_4).
\end{aligned} \tag{1.7}$$

Ahora podemos usar las ecuaciones (1.7) para sustituir los flujos en (1.6) y obtener el siguiente sistema de cinco ecuaciones con cinco incógnitas

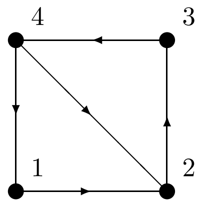
$$\begin{aligned}
(c_1 + c_2 + c_3)v_1 + c_1v_2 + c_2v_3 + c_3v_4 &= d_1\\
c_1v_1 - (c_1 + c_4 + c_5)v_2 + c_4v_4 + c_5v_5 &= d_2\\
c_2v_1 - (c_2 + c_6)v_3 + c_6v_4 &= d_3\\
c_3v_1 + c_4v_2 + c_6v_3 - (c_3 + c_4 + c_6 + c_7)v_4 + c_7v_5 &= d_4\\
c_5v_2 + c_7v_4 - (c_5 + c_7)v_5 &= d_5
\end{aligned} \tag{1.8}$$

Si somos capaces de resolver estas ecuaciones en las variables v_i luego sólo tenemos que sustituir los valores hallados en las ecuaciones (1.7) para calcular las intensidades.

Consideraremos ahora el caso particular de (1.8) en el que todas las conductividades toman el valor 1. Obtenemos el sistema

$$
\begin{array}{rcrcrcrcrcl}
-3v_1 & + & v_2 & + & v_3 & + & v_4 & & & = & d_1\\
v_1 & - & 3v_2 & + & & - & v_4 & + & v_5 & = & d_2\\
v_1 & & & - & 2v_3 & + & v_4 & & & = & d_3\\
v_1 & + & v_2 & + & v_3 & - & 4v_4 & + & v_5 & = & d_4\\
 & & v_2 & & & + & v_4 & - & 2v_5 & = & d_5
\end{array}
\tag{1.9}
$$

Ejercicio 1.4. Plantear las ecuaciones para la red de la figura suponiendo que todas las conductividades son iguales a 1 y que cada nodo i, $i = 1, 2, 3, 4$, recibe una demanda d_i. Resolver las ecuaciones y calcular flujos y potenciales en el caso en que $d_4 = 1$, $d_3 = -1$, $d_1 = d_2 = 0$.



1.1.5. Los esfuerzos en un reticulado.

Estudiaremos ahora el problema de cómo se distribuyen los esfuerzos en un *reticulado*, o estructura de barras sometida a cargas exteriores. Estas cargas producen tensiones en las barras, que deben ser toleradas para que la estructura se mantenga en equilibrio y cumpla la función para la que fue diseñada y construida. Calcular el valor de las tensiones es, entonces, un problema relevante para el diseño de este tipo de estructuras

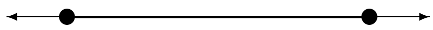
Para tratar el problema trabajaremos con un modelo del reticulado en el que las barras serán completamente rígidas y las consideraremos como segmentos de rectas que confluyen en nodos, donde se articulan varias barras. En la figura del ejemplo 1.1.3, página 16, mostramos un nodo en el que confluyen tres barras y que además soporta una carga externa P. En la figura 1.3, página 19, aparece el esquema de un reticulado con los anclajes que lo mantienen fijo.

Seleccionaremos ahora las variables relevantes para nuestro modelo, y buscaremos cuáles son las relaciones que deben satisfacer.

El equilibrio de una barra

Supondremos que sobre las barras sólo actúan fuerzas aplicadas en sus extremos. Para que cada barra esté en equilibrio las resultantes de estas fuerzas deben tener la dirección de la barra, si no fuera así las fuerzas harían girar a la barra. Además, la resultante de las fuerzas que actúan en un extremo debe ser exactamente la opuesta de la que actúa en el otro. Si esto no ocurriera la fuerza neta que actuaría sobre la barra produciría su desplazamiento.

En resumen, el efecto neto de las fuerzas que actúan sobre la barra en equilibrio es el que se muestra en la figura. Actúan en los extremos fuerzas opuestas, alineadas con la dirección de la barra.

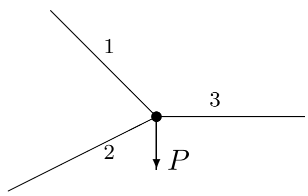


El esfuerzo que realiza la barra queda entonces completamente caracterizado por el módulo de la fuerza que la tracciona (como hemos dibujado en la figura) o comprime. Como la dirección de la fuerza es conocida sólo necesitamos especificar el valor de la tensión T a la que está sometida la barra. Daremos signo positivo a T cuando la barra está traccionada y negativo cuando está comprimida. El valor de T determina entonces el módulo y sentido de las fuerzas que actúan en cada extremo.

Ecuaciones de equilibrio en los nodos

Hasta aquí hemos descrito el equilibrio de cada barra. Para estudiar el equilibrio del conjunto analizaremos el equilibrio de cada nodo. La condición que debe satisfacerse es que todas las fuerzas que actúan sobre un nodo dado, provenientes de las barras que confluyen en el nodo o de cargas externas, estén equilibradas. En un reticulado plano, como el de la figura 1.3, debemos asegurarnos de que tanto las componentes horizontales como las componentes verticales de las fuerzas estén equilibradas. Por lo tanto, tendremos dos ecuaciones para cada nodo. En el próximo ejemplo mostramos cómo son estas ecuaciones para un nodo en el que confluyen tres barras.

Ejemplo 1.1.3. Sobre el nodo de la figura actúan cuatro fuerzas,las de las tres barras y la carga exterior P. La barra 1 forma un ángulo de $\pi/4$ con la vertical, la barra 2 un ángulo de $\pi/6$ con la horizontal y la barra 3 está en posición horizontal.



Cada barra está sometida a una tensión T_i, $i = 1, 2, 3$. La fuerza con que la barra actúa sobre el nodo dependerá de T_i y la orientación de la barra, y puede descomponerse en una componente vertical y una horizontal para plantear las ecuaciones de equilibrio. Tomaremos las orientaciones usuales de los ejes horizontal y vertical, en las que el signo positivo corresponde a lo que apunta hacia la derecha o hacia arriba respectivamente. Entonces, por ejemplo, la componente vertical de la fuerza que ejerce la barra 1 sobre el nodo es

$$T_1 \cos(\pi/4) = T_1/\sqrt{2},$$

en tanto que la componente horizontal es

$$-T_1 \operatorname{sen}(\pi/4) = T_1/\sqrt{2}.$$

Por el principio de acción y reacción, esta fuerza es exactamente opuesta a la fuerza que el nodo trasmite a la barra 1.

Ejercicio 1.5. Hallar en función de T_2, T_3 y las posiciones de las barras las componentes horizontales y verticales de los esfuerzos que las barras 2 y 3 realizan sobre el nodo.

Proyectando los esfuerzos en la dirección de ambos ejes y planteando la condición de equilibrio para cada una de las dos direcciones encontramos las ecuaciones que vinculan las tensiones y las cargas en el nodo. En este ejemplo, la ecuación que corresponde a la dirección horizontal es

$$-\frac{1}{\sqrt{2}}T_1 - \frac{\sqrt{3}}{2}T_2 + T_3 = 0.$$

El factor $-1/\sqrt{2}$ que multiplica a T_1 es igual al coseno del ángulo $3\pi/4$ que la barra 1 forma con la dirección positiva del eje horizontal. El ángulo de la barra 2 es $7\pi/6$, cuyo coseno es el factor $-\sqrt{3}/2$ que aparece multiplicando a T_2. Si el lector lo prefiere puede pensar que la barra 2 forma un ángulo de $\pi/6$ con la dirección negativa del eje vertical, y considerar el factor $-\sqrt{3}/2$ como el producto de $\cos(\pi/6)$ por -1, que da cuenta de que la proyección de la barra sobre la vertical quedó "apuntando hacia atrás". La barra 3 ejerce una fuerza T_3 en el sentido positivo del eje horizontal, y la carga externa P no tiene componente horizontal.

El análisis de las componentes verticales conduce a la segunda ecuación de equilibrio,

$$\frac{1}{\sqrt{2}}T_1 - \frac{1}{2}T_2 - P = 0. \tag{1.10}$$

Los detalles son similares a lo que acabamos de hacer, y los dejamos como ejercicio para el lector.

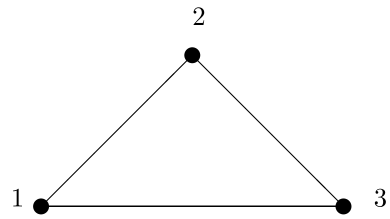
Ejercicio 1.6. Justificar la ecuación (1.10) ♣

Una primera aproximación al problema de determinar las tensiones consiste en escribir todas las ecuaciones de equilibrio de los nodos, y buscar los valores de las tensiones que las satisfacen. En el ejercicio 1.7 iremos recorriendo este camino, resolveremos algunos reticulados, y finalmente encontraremos que esta aproximación tiene limitaciones serias, que obligarán a recurrir a una modelización más elaborada de un sistema de barras para poder alcanzar el objetivo de conocer las tensiones con que la estructura reacciona a las cargas a las que es sometida.

Cargas y reacciones

Antes de plantear el ejercicio conviene hacer un comentario acerca de las fuerzas externas que actuarán sobre nuestra estructura. Las distinguiremos entre *cargas* y *reacciones*. Las cargas son, en general, conocidas. Por ejemplo, el peso de lo que queremos colgar de la estructura. Las reacciones son los esfuerzos que se trasladan al anclaje de la estructura para que no se venga todo abajo. Las reacciones pueden ser calculadas a partir de las cargas, pero también podemos momentáneamente olvidarnos de ellas asumiendo que el anclaje es capaz de soportar cualquier esfuerzo y que, automáticamente, equilibrará cualquier carga.

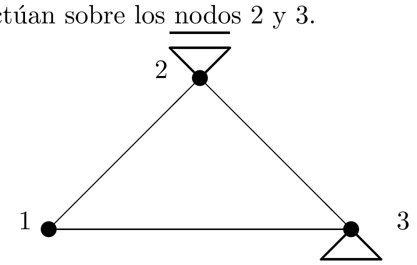
Ejercicio 1.7. En este ejercicio trabajaremos con el sencillo reticulado triangular que aparece en la figura. Lo consideraremos sometido a distintas cargas, y con distintos tipos de anclaje. También analizaremos su comportamiento bajo solicitaciones completamente generales en sus tres nodos.



1. ESFUERZOS EN EL RETICULADO ANCLADO Y CARGADO. Supongamos que el reticulado está sujeto por los nodos 2 y 3, de forma tal que el anclaje en el nodo 3 es capaz de resistir cualquier esfuerzo, y el del nodo 2 puede resistir cualquier esfuerzo vertical.

Con el esquema de anclajes que acabamos de describir, y que se muestra en la figura, se aplica una carga vertical de $10T$ en el nodo 1. Escribir las ecuaciones de equilibrio del nodo 1 y la condición de equilibrio de las componentes horizontales de las fuerzas que actúan sobre el nodo 2. Determinar las tensiones en las barras a partir de estas ecuaciones.

2. REACCIONES EN LOS ANCLAJES. En la parte anterior nuestro interés era calcular las tensiones en las barras. Por eso no planteamos las condiciones de equilibrio del nodo 3, ya que supusimos que su anclaje es capaz de equilibrar cualquier esfuerzo.Tampoco la condición de equilibrio que corresponde a las componentes verticales de las fuerzas que actúan sobre el nodo 2. Ninguna de estas ecuaciones arroja información sobre las tensiones en las barras. Pero sí dan información sobre las reacciones de los anclajes. Determinar entonces las fuerzas externas que actúan sobre los nodos 2 y 3.



3. ESFUERZOS Y REACCIONES PARA OTRO ANCLAJE. Calcular los esfuerzos y las reacciones cuando el reticulado está sujeto de forma tal que el anclaje del nodo 2 es capaz de compensar cualquier esfuerzo, y el del nodo 3 sólo puede hacer fuerza en el sentido vertical (es decir, sabemos que las únicas fuerzas que actúan sobre ese nodo y tienen componente horizontal son las que provienen de las barras).

Sobre el nodo 1 supondremos que actúa una carga C genérica, con componente horizontal H y vertical V.
4. ESFUERZOS EN LAS BARRAS PARA CARGAS CUALESQUIERA. Calcular las tensiones en las barras cuando actúan cargas externas conocidas sobre los nodos. Llamemos H_i y V_i, $i = 1, 2, 3$, a las componentes horizontales y verticales de estas fuerzas. ¿Es posible encontrar tensiones que equilibren los esfuerzos para cualquier valor de las cargas H_i y V_i, $i = 1, 2, 3$? Discutir e interpretar los resultados.

El objetivo de nuestro próximo ejercicio es emplear lo que hemos aprendido acerca de las ecuaciones de equilibrio de un reticulado en un problema algo más complejo. La parte 2 ilustra acerca de las limitaciones del modelo con el que estamos trabajando.

Ejercicio 1.8.

1. Plantear las ecuaciones y hallar las tensiones en el sistema de barras de la figura 1.3, que se supone sometido a las cargas f y g que allí se ilustran. El

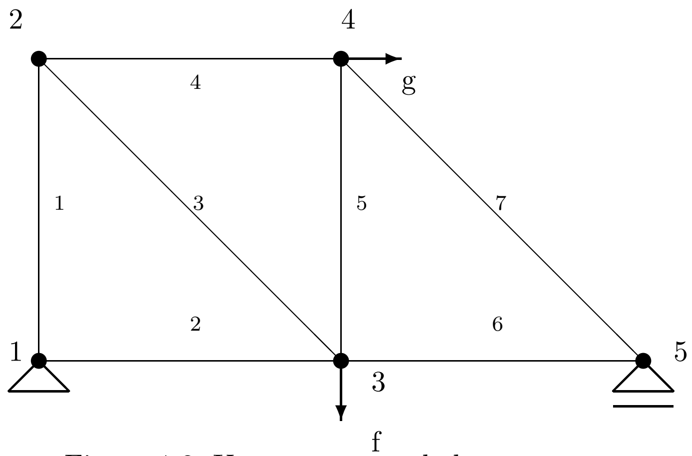


Figura 1.3: Una estructura de barras

anclaje en el nodo 1 es capaz de equilibrar cualquier esfuerzo, y el anclaje en el nodo 5 sólo puede equilibrar cualquier esfuerzo vertical, la componente horizontal de la reacción en ese punto es nula.

2. Se agrega una octava barra entre los nodos marcados 1 y 4. Hallar el sistema que deben satisfacer las tensiones para que el reticulado esté en equilibrio y determinar los valores de las tensiones en cada barra. ¿Cuáles son los valores de las tensiones que equilibran el sistema? Para una carga dada, ¿escogerán las barras una solución en función del estado de ánimo con que se encuentren? ¿Cómo se soluciona este problema de indeterminación en la solución? ♣

Observación 1.1.4. Modelización matemática del reticulado
En el modelo que estamos considerando las barras no se deformarán, independientemente de cuánto las carguemos. Por supuesto, las cosas no son así en la realidad. Pero recién estamos haciendo una primera aproximación. Modelos más elaboradas y precisos tienen en cuenta el hecho de que las barras se deforman cuando están sometidas a una carga, pero aún así constituyen una enorme simplificación del problema.

Recordemos que cualquier modelo constituye una simplificación. Un compromiso entre la descripción ajustada de la realidad y nuestra capacidad de calcular con el modelo. Y su validez dependerá de que permita determinar los valores de las variables que nos interesa conocer con un grado de aproximación suficiente para nuestros propósitos. ¿Cuál es ese grado de aproximación? Imposible saberlo sin saber antes cuál es la situación que queremos analizar y en que contexto lo hacemos. ♠

Observación 1.1.5. Redes, grafos y álgebra lineal
Algunos de los ejemplos tratados tienen que ver con redes: una red vial, un circuito eléctrico, un modelo general para flujos sobre redes, y un reticulado (una red de barras articuladas entre sí).

Sistemas complejos con diversas partes interconectadas entre sí (redes) aparecen en muchas áreas de la ingeniería. El álgebra lineal es una herramienta indispensable en su consideración, por lo que esperamos que constituyan una fuente de motivación y de problemas interesantes para estudiantes que se orientan hacia alguna rama de la ingeniería. Las complejas redes de comunicaciones de nuestra sociedad –la Internet, redes viales, redes cualesquiera de distribución– son ejemplos de esta naturaleza.

Pero señalemos otro, que tiene que ver con posibles interacciones, en un ejemplo tomado de la biología. No es el número de genes de un organismo el que determina su complejidad, sino la red de interconexiones que vincula a los genes. Esta razón mueve en la actualidad a muchos científicos a estudiar redes de genes, y sus estudios tienen como contrapartida los estudios de redes que aparecen en otras disciplinas.

Así, las matemáticas de las redes constituyen una apasionante área de investigación multidisciplinaria, siendo un dominio común de trabajo de matemáticos, físicos, biólogos, epidemiólogos, economistas, por citar algunas de las especialidades involucradas.

Tal como vimos, las redes pueden ser descritas como un conjunto de nodos vinculados por arcos. A su vez, estos arcos pueden ser vistos como parejas de nodos: los dos nodos que el arco une. Este cuadro tan general se presta a una formulación general, abstracta, que englobe todas las posibilidades sin aludir directamente a ninguna. Es decir ¡a una teoría matemática! La teoría correspondiente es la *teoría de grafos*. Quizás interese al lector saber que esta teoría se sirve del álgebra lineal que queremos mostrar en este curso. Ver, por ejemplo, el ejercicio 2.8, en la página 170. El capítulo 3 del texto [G] contiene una introducción a la teoría de grafos. ♠

1.1.6. Otros ejemplos

Cerramos esta sección introductoria con un par de ejercicios en los que mostramos problemas que pueden ser resueltos formulando sistemas de ecuaciones lineales adecuados.

Ejercicio 1.9. Dietas
¡Basta de comida chatarra! Hemos decidido querernos a nosotros mismos y empezar a comer decentemente. ¿Qué nutrientes necesitamos?

Luego de relevar la información al respecto hemos llegado a la conclusión de que necesitamos que lo que comamos cada día nos aporte 3.000 calorías, 60 gramos de proteínas, 1 gramo de calcio y 10 miligramos de hierro[3]. Nuestras costumbres, prejuicios y/o principios nos restringen a una dieta formada por los siguientes 6 alimentos, que contienen cada 100 gramos las cantidades de los nutrientes básicos que aparecen en la tabla:

Alimento	Calorías	Proteínas (g.)	Calcio (mg.)	Hierro (mg.)
Carne vacuna	250	18	10	2,5
Garbanzos	360	20	130	8
Lechuga	16	1,3	30	0,8
Huevos	160	12	60	3
Mayonesa	718	1,1	18	0,5
Leche	65	3,3	120	0,1

Diseñar una dieta adecuada a nuestras necesidades.

Ejercicio 1.10. Transiciones en una población (I)
La población de un país está distribuida en tres grupos: los que viven en ciudades, los que viven en zonas rurales, y los que viven en el exterior del país. Llamemos c, r y e a los porcentajes de pobladores en cada grupo. Cada año, algunos habitantes cambian su situación, según el siguiente esquema:

de la **c**iudad	de la población **r**ural	del **e**xterior
60 % sigue en la ciudad,	40 % se mudan a la ciudad,	10 % vuelven a la ciudad,
10 % pasa a zona rural,	40 % permanece donde está,	10 % a zonas rurales,
30 % emigran al exterior,	20 % parte fuera del país,	80 % sigue en el exterior.

[3]Los requisitos reales para una dieta incluyen otros nutrientes que hemos omitido. No hace falta que los incorporemos al enunciado del ejercicio, porque el problema sería siendo esencialmente el mismo, pero ¡más vale que no nos olvidemos de ellos a la hora de comer!

En este modelo del comportamiento de la población no tenemos en cuenta el efecto de muertes y nacimientos.

1. Si un año comienza con un 40 % de la población en la ciudad, un 30 % en zonas rurales, y un 30 % en el extranjero, ¿cuál será la distribución a fin del año? ¿Cómo comenzó el año anterior?

2. Hallar la expresión general de la transición de los porcentajes (c, r, e) de un año al siguiente.

3. ¿Existe alguna distribución (c, r, e) de la población tal que los porcentajes con los que las personas se distribuyen al final del año sean iguales a los del comienzo del año?

Observación 1.1.6. Un estado como el que se describe en la parte 3 del ejercicio anterior es lo que se llama un *estado estacionario* o de *equilibrio*. Cuando la población se coloca en ese estado, con esa distribución, a pesar de los movimientos de personas, los porcentajes de pobladores en la ciudad, las zonas rurales y en el exterior no cambian de un año otro. Esto ocurre porque los distintos flujos de gente se cancelan unos con otros. En este estado de equilibrio un cierto número de personas continuará abandonando, por ejemplo, la ciudad cada año. Pero un número exactamente igual llegará desde el exterior o el campo para hacer que el porcentaje de habitantes en la ciudad no se modifique. ♠

1.1.7. Para tener presente

- Los sistemas de ecuaciones lineales aparecen en la modelización matemática de una gran variedad de situaciones. Muchas de ellas son de interés para la Ingeniería.

1.2. Presentación general de los sistemas de ecuaciones lineales

Aunque más tarde volveremos sobre ellos, dejaremos ahora de lado los problemas que dan lugar a sistemas de ecuaciones lineales, y nos dedicaremos a estudiar una manera sistemática de resolverlos y a comprender las propiedades de sus soluciones. Nuestro objeto de trabajo será entonces un sistema lineal genérico, cuyos coeficientes e incógnitas pueden ser números cualesquiera. Tampoco pondremos limitación alguna al número de ecuaciones ni al número de incógnitas.

A lo largo de las próximas secciones casi no nos importará de dónde vienen los coeficientes del sistema, ni qué significan sus soluciones, o la ausencia de soluciones. Vamos a desarrollar una teoría matemática abstracta, que no requiere hacer referencia a nada exterior a ella para sostenerse. Aún en este marco trataremos de que los conceptos que se introduzcan tengan significado para el lector, o lo adquieran rápidamente. La mayor parte de ellos aparecerá en el curso de la búsqueda de respuestas a dos preguntas básicas. Dado un sistema de ecuaciones:

- ¿tiene o no tiene soluciones?
- Si tiene soluciones, ¿cuántas tiene? ¿cómo podemos describirlas?

1.2.1. Notación, conjuntos numéricos, soluciones

Lo primero que necesitaremos para trabajar con esta generalidad es una notación adecuada. Emplearemos más de una a lo largo del curso, pero comenzamos por introducir la que sólo consiste en sistematizar los nombres de todas las variables en juego. Escribiremos un *sistema lineal de m ecuaciones con n incógnitas* $x_1, x_2, \ldots, x_n$, en la forma

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2, \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m, \end{array}\right. \tag{1.11}$$

donde los números

$$a_{ij}, \quad i = 1, \ldots, m, \; j = 1, \ldots, n,$$

son un dato que supondremos conocido. Los llamaremos *coeficientes* del sistema. También supondremos conocidos los números

$$b_j, \quad j = 1, \ldots, m,$$

a los que nos referiremos como *términos independientes*). El problema que abordaremos consiste en hallar las *incógnitas*, números

$$x_i, \quad i = 1, \ldots, n,$$

tales que se satisfagan las m ecuaciones en (1.11). Nos referiremos a un sistema de este tamaño como a *un sistema* $m \times n$, donde m hace referencia a la cantidad de ecuaciones, y n a la de incógnitas.

Observación 1.2.1. Los conjuntos numéricos para la teoría.
Una precisión se hace imprescindible: especificar cuál es el conjunto numérico sobre el que trabajaremos. El siguiente ejemplo ilustra esta necesidad.

Ejemplo 1.2.2. Nos planteamos resolver el sistema de ecuaciones

$$\begin{array}{rcl} 2x - y & = & 0, \\ 2x + y & = & 2. \end{array} \tag{1.12}$$

Sumando ambas ecuaciones obtenemos $4x = 2$, de donde concluimos que

$$x = 1/2.$$

Volviendo a cualquiera de las dos ecuaciones originales y sustituyendo x por el valor que acabamos de hallar encontramos que

$$y = 1.$$

Por lo tanto $x = 1/2$, $y = 1$, es la única solución del sistema. ♣

Observemos que el sencillo sistema de nuestro ejemplo no tiene soluciones que sean números naturales, aunque sus coeficientes son naturales. Esto es así porque al resolver los sistemas de ecuaciones lineales es necesario realizar sumas, restas, multiplicaciones y divisiones, y algunas de estas operaciones no son siempre posible dentro del conjunto $\mathbb{N}$ de los números naturales. Trabajaremos entonces en un conjunto numérico donde todas estas operaciones sean posibles, lo que hace necesario especificar que el conjunto de números tenga *estructura de cuerpo.*

Por ejemplo, el conjunto $\mathbb{Q}$ de los números racionales es un cuerpo. Observemos que la solución del sistema (1.12) requiere la consideración de números racionales. También son cuerpos el conjunto $\mathbb{R}$ de los números reales, y el de los complejos $\mathbb{C}$. Los cuerpos $\mathbb{Z}_2$ y $\mathbb{Z}_3$ son ejemplos de cuerpos que sólo tienen un número finito de elementos (2 y 3 respectivamente). Esta breve enumeración no agota la lista de los cuerpos, y todavía quedan muchos más cuerpos sumamente interesantes: los cuerpos $\mathbb{Z}_2$ y $\mathbb{Z}_3$ son casos particulares de una construcción que para cualquier número primo p origina un cuerpo discreto llamado $\mathbb{Z}_p$; existe una infinidad de cuerpos intermedios entre $\mathbb{Q}$ y $\mathbb{R}$. Algunos de estos cuerpos intervienen en la resolución de problemas clásicos como la trisección del ángulo o la cuadratura del círculo; los *bites*, listas de ocho ceros y unos, pueden ser dotados de una estructura de cuerpo. Mucha de la tecnología de las telecomunicaciones actuales emplea esta estructura para procesar la información. .

En este curso trabajaremos habitualmente sobre $\mathbb{R}$, pero aparecerán algunos ejemplos que emplean otros cuerpos. En general, desarrollaremos toda la teoría sobre un cuerpo $\mathbb{K}$ cualquiera, sin hacer referencia a ninguno en particular. Sólo haremos uso de las propiedades de la estructura de cuerpo, no de la naturaleza precisa de los números que forman el cuerpo. Las propiedades de la estructura de cuerpo y alguna información adicional que extiende estos comentarios se recogen en el apéndice A.2. ♣

Una vez especificado que trabajaremos sobre algún cuerpo $\mathbb{K}$, es decir, que los coeficientes de los sistemas lineales serán elementos de un cuerpo $\mathbb{K}$ y que buscaremos soluciones en ese mismo cuerpo, sigamos adelante con nuestra teoría.

Hemos reservado el primer subíndice de los coeficientes a_{ij} para indicar a qué ecuacion pertenecen, y el segundo para indicar a qué variable multiplican. Así, al variar el índice i en a_{ij} vamos cambiando de renglón en (1.11), y al aumentar j nos movemos de izquierda a derecha, recorriendo las columnas en las que aparece alineado todo lo que corresponde a una misma incógnita. Por supuesto, este ordenamiento de los subíndices es absolutamente convencional. Lo adoptaremos porque es estándar, y es coherente con el que se utiliza para las matrices que introduciremos en la página 37.

Observación 1.2.3. El problema de hallar los valores x_i de las n incógnitas puede pensarse como el problema de hallar una única lista

$$X = (x_1, x_2, \ldots, x_n)$$

de n números. En una primera consideración no parece haber una gran ventaja en este enfoque, más allá de pensar en un único objeto en vez de en n, pero estas *listas* de n números, o n*-uplas*, pueden ser vistas como *vectores*, elementos de una estructura algebraica y geométrica que nos ayudará muchísimo en la manipulación y comprensión de los sistemas lineales. Comenzaremos a introducir esta estructura en la observación 1.2.21, página 44, y en la subsección 1.2.6. La emplearemos sistemáticamente a partir de la sección 1.3. ♣

A la luz de la observación precedente, digamos que *una solución del sistema* (1.11) es una lista de n números

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

tales que si se sustituye

$$x_1 = \alpha_1,\ x_2 = \alpha_2, \ldots, x_n = \alpha_n$$

se verifican simultáneamente las m ecuaciones en (1.11).

Ejemplo 1.2.4. Vamos a discutir en este ejemplo tres sencillos sistemas reales de dimensiones 2×2.

1. El sistema
$$\left\{ \begin{array}{rcl} x_1 + x_2 & = & 1, \\ x_1 - x_2 & = & 1, \end{array} \right.$$
tiene una única solución $(x_1, x_2) = (1, 0)$. Es evidente que este par de números es una solución del sistema. Además, es la única, porque si sumamos las dos ecuaciones del sistema obtenemos $2x_1 = 2$, lo que implica $x_1 = 1$. Y al restar la segunda de la primera nos encontramos con $2x_2 = 0$, que sólo puede ser satisfecha por $x_2 = 0$.

2. No puede haber ninguna solución del sistema de ecuaciones
$$\left\{ \begin{array}{rcl} x_1 + x_2 & = & 1, \\ x_1 + x_2 & = & 2. \end{array} \right.$$

3. Pero el sistema
$$\left\{ \begin{array}{rcl} x_1 + x_2 & = & 1, \\ 2x_1 + 2x_2 & = & 2, \end{array} \right.$$
admite infinitas soluciones. Todas las parejas $(x, 1 - x)$, con $x \in \mathbb{R}$, satisfacen ambas ecuaciones. ♣

Ejercicio 1.11. En cada caso, determinar si las listas que se especifican son o no son soluciones de los sistemas dados.

1. Las ternas $(1, -1, 1)$ y $(2 - 3y, y, -1 + 2y)$, donde y puede ser cualquier número real, para el sistema
$$\left\{ \begin{array}{rcr} x_1 + x_2 + x_3 & = & 1, \\ x_1 - x_2 + 2x_3 & = & 0, \\ x_1 - 3x_2 + 3x_3 & = & -1. \end{array} \right.$$

2. Las cuaternas $(1,1,1,0)$, $(1,1,0,1)$ y $(0,0,1,0)$ para el sistema

$$\left\{\begin{array}{rcl} x_1+x_2+x_4 & = & 0, \\ x_1+x_2+x_3 & = & 1, \\ x_1+x_3+x_4 & = & 1, \end{array}\right.$$

   en $\mathbb{Z}_2$.

3. Las duplas $(1-i,1+i)$ y $(1+i,1-i)$, para el sistema con coeficientes complejos

$$\left\{\begin{array}{rcrcr} (i-1)x_1 & + & (1+i)x_2 & = & 4, \\ x_1 & - & x_2 & = & 2i, \end{array}\right.$$

El ejemplo 1.2.4 y el ejercicio 1.11 muestran que un sistema lineal puede tener una única solución, ninguna o muchas. Nuestra tarea al estudiarlos será encontrar y describir la solución de estos sistemas en el siguiente sentido: dado un sistema de ecuaciones llamaremos *conjunto solución* del sistema al conjunto de todas las soluciones de S. **Resolver un sistema es determinar su conjunto solución**.

Ejemplo 1.2.5. Los conjuntos solución de los sistemas del ejemplo 1.2.4 son

1. $\{(1,0)\}$,
2. el conjunto vacío,
3. el conjunto $\{(x,1-x);\ x\in\mathbb{R}\}$,

respectivamente. ♣

Ejemplo 1.2.6. El ejemplo más sencillo de un sistema lineal es el sistema de una ecuación con una incógnita

$$ax=b, \tag{1.13}$$

donde a y b son dos números (elementos de algún cuerpo $\mathbb{K}$) cualesquiera. La letra x indica a la única incógnita del sistema.

En pocos renglones podemos discutir con total generalidad cómo es el conjunto solución.

- Si $a\neq 0$ entonces el sistema tiene una única solución, independientemente del valor de b. Sólo tenemos que multiplicar ambos lados de la ecuación (1.13) por el inverso a^{-1} del número a, para concluir que si x es una solución entonces $x=a^{-1}b$. Este número es una solución de la ecuación, porque
$$a(a^{-1}b)=(aa^{-1})b=1b=b.$$
  La discusión precedente muestra que además es la única.

- Si $a = 0$ tendremos
$$ax = 0x = 0,$$
independientemente del valor de x. Debemos distinguir entonces dos casos:
  - si $b \neq 0$ entonces, para cualquier x, tendremos $0x = 0 \neq b$. Por lo tanto no hay solución del sistema;
  - si $b = 0$ entonces la igualdad en (1.13) siempre se satisface, y cualquier x es solución del sistema. En este caso el sistema tiene muchas soluciones, tantas como elementos tenga el cuerpo $\mathbb{K}$ sobre el que estamos trabajando, porque cualquier número en el cuerpo es solución[4].

Naturalmente, este problema es extremadamente sencillo pero, como acontece muchas veces, ya podemos apreciar en este simple ejemplo algunos comportamientos que reaparecerán en situaciones más generales: problemas de compatibilidad o incompatibilidad asociados a la invertibilidad de los coeficientes del sistema; inexistencia o multiplicidad de soluciones dependiendo de los coeficientes y términos independientes, etcétera. ♣

Llamaremos *homogéneo* a un sistema lineal con m ecuaciones en el que todos los términos independientes b_i, $i = 1, \dots, m$ son iguales a 0. La justificación del término se encuentra en la página 132, en la observación 1.7.2 de la sección 1.7.

Ejercicio 1.12. Mostrar que un sistema lineal homogéneo siempre es compatible, porque tiene al menos una solución en la que todas las incógnitas toman el valor 0.

Conviene introducir aquí una terminología más o menos corriente:

- si un sistema no tiene ninguna solución, o, equivalentemente, el conjunto solución es vacío, diremos que es *incompatible*;
- si existe alguna solución llamaremos al sistema *compatible*. Distinguiremos entonces dos casos.
  - Cuando la solución es única diremos que el sistema es *compatible determinado*;
  - y si admite más de una solución diremos que es *compatible indeterminado*.

[4]Si el cuerpo es infinito –este es el caso para, por ejemplo, $\mathbb{Q}$, $\mathbb{R}$ o $\mathbb{C}$– entonces la ecuación tiene infinitas soluciones. Si el cuerpo es finito –por ejemplo, $\mathbb{Z}_p$, con p un número primo– la solución no es única, pero hay un número finito de ellas. En el caso de $\mathbb{Z}_p$ hay exactamente p.

1.2.2. El método de eliminación de Gauss

El objetivo de esta sección es presentar un método sistemático para la resolución de sistemas lineales de cualquier tamaño. La idea del método es muy simple: **ir reduciendo en cada paso el problema a un problema que tiene una ecuación menos y una incógnita menos, y transformar así, por medio de la eliminación inteligente de incógnitas, el problema original en un problema de sencilla solución**. Este método es conocido como *método de eliminación de Gauss* o *método de eliminación gaussiana*. El nombre alude a que en cada paso vamos eliminando una o más incógnitas, y es un reconocimiento a quien lo introdujo: el matemático Carl Friederich Gauss.

Nota Histórica 1.1. CARL GAUSS (1777-1855)
Carl Gauss nació en el año 1777, en Brunswick, un pueblo que actualmente pertenece a Alemania. A los siete años sorprendió a sus maestros al sumar casi instantáneamente todos los números naturales entre 1 y 100. Lo hizo a partir de la observación de que se trataba de sumar 50 pares de números naturales, dónde cada pareja sumaba 101. En 1795 Gauss abandonó Brunswick para estudiar en la Universidad de Göttingen. Permanecería allí hasta 1798. Al abandonar Göttingen Gauss no había conseguido ningún diploma, pero había descubierto como construir con regla y compás un polígono regular de 17 lados, lo que constituyó el mayor avance en esta área desde el tiempo de las matemáticas griegas.

Este promisorio comienzo de la carrera de Gauus como investigador fue confirmado por su producción futura: en su tesis doctoral demostró el teorema fundamental del álgebra; realizó luego importantes contribuciones en teoría de números; creó el método de los mínimos cuadrados para predecir la posición del asteroide Ceres, descubierto en 1801; desarrolló áreas de la estadística y la teoría del potencial; demostró el llamado *Teorema egregio* de la geometría diferencial, que tiene como una de sus consecuencias el hecho de que no es posible realizar un mapa plano de la superficie de la tierra que respete fielmente las distancias entre pares de puntos cualesquiera; trabajó fecundamente en problemas de electromagnetismo, etcétera. Una serie de contribuciones que hacen a Gauss merecedor de un lugar destacado en la historia de la ciencia, en particular de la Matemática. Gauss falleció en 1855.

Gauss introdujo el método de eliminación que hoy lleva su nombre hacia 1800, con el propósito de resolver problemas de ajuste de datos (por el método de mínimos cuadrados) de observaciones astronómicas. Más tarde aplicaría el método en problemas de geodesia. De hecho, la primera aparición del método en un trabajo publicado ocurrió en un manual de geodesia de la autoría de Wilhelm Jordan, y, curiosamente, el método de eliminación de Gauss fue considerado durante años parte de la geodesia y no de la matemática.

El método de eliminación de Gauss tiene un antecedente en manuscritos chinos del siglo II AC –muy anteriores a Gauss, por cierto– en los que se explicaba como resolver un sistema de tres ecuaciones con tres incógnitas.

En nuestro próximo ejemplo mostramos un sistema lineal de una forma muy particular, que lo hace especialmente sencillo de resolver, porque de una de sus ecuaciones podremos despejar inmediatamente una incógnita. Con esta incógnita conocida podremos despejar otra, y así sucesivamente hasta hallarlas todas.

Luego mostraremos cómo el método creado por Gauss transforma el problema de resolver un sistema lineal cualquiera en un problema en el que sólo hay que resolver un sistema con esta forma tan conveniente, en la que será inmediato saber qué incógnitas pueden despejarse a partir del conocimiento de las demás, y cuáles pueden ser fijadas arbitrariamente.

Ejemplo 1.2.7. Busquemos tres números reales x, y y z que satisfagan

$$\left\{\begin{array}{rcl} x+y+2z &=& 8, \\ 2y+z &=& 8 \\ z &=& 2. \end{array}\right.$$

Éste es un sistema 3×3 muy fácil de resolver. De la última ecuación se despeja

$$z = 2.$$

Con este valor de z sustituimos en al segunda ecuación y tenemos

$$2y + 2 = 8 \Rightarrow 2y = 6 \Rightarrow y = 3.$$

Finalmente con los valores de y y z hallados “subimos” a la primera ecuación y sustituyendo se tiene que

$$x + 3 + 2 \times 2 = 8 \Rightarrow x = 1.$$

En consecuencia, el sistema tiene una única solución, que es

$$x = 1, \quad y = 3, \quad z = 2.$$

Podemos escribir esta solución como la lista de tres números reales $(1, 3, 2)$. El conjunto de soluciones del sistema es $\{(1, 3, 2)\}$.

El procedimiento de despejar una incógnita en una ecuación y sustituirla en otra de arriba se denomina *sustitución hacia arriba*, y es posible por la forma particular de este sistema. ♣

Observación 1.2.8. Notará el lector que en el último ejemplo hemos abandonado momentáneamente la notación (x_1, x_2, x_3) para referirnos a las incógnitas, que pasaron a llamarse (x, y, z). Para problemas con pocas incógnitas es corriente usar las letras x, y o z. Incluso también t, u o v, y de tanto en tanto nos apartaremos de los subíndices y usaremos estos nombres. Es especialmente corriente usar x, y y z en problemas de geometría. Seguiremos esta costumbre en cuando trabajemos con la geometría del espacio de dimensión 3, y en esas secciones del texto abandonaremos a x_1, x_2 y x_3 en beneficio de x, y y z. De todos modos, más allá de que los cambios de notación pueden resultar enojosos, el nombre de las incógnitas es absolutamente irrelevante. ♠

Diremos que un sistema como el del ejemplo 1.2.7, en el que

1. las incógnitas que aparecen en cada ecuación también aparecen en la anterior;

2. en cada ecuación aparecen menos incógnitas que en la anterior,

está *escalerizado*, por la forma de "escalera" que adoptan sus ecuaciones al ser escritas. **Para estos sistemas escalerizados tenemos un método fácil de resolución: calcular las incógnitas que aparecen en las últimas ecuaciones e ir sustituyendo "hacia arriba"**.

Mostraremos ahora con un ejemplo como **el método de eliminación de Gauss permite transformar el problema de calcular las soluciones de un sistema lineal en el problema de resolver un sistema lineal escalerizado**. Por esta razón es también llamado *método de escalerización.*

Ejemplo 1.2.9. Busquemos las soluciones del sistema

$$\left\{\begin{array}{l} x+y+2z=8, \\ x+2y+z=0, \\ 2x+y+z=4. \end{array}\right. \tag{1.14}$$

Explicitaremos los pasos de nuestra resolución.
Paso 1. Comenzamos por "hacer desaparecer la variable x de todas las ecuaciones menos una", usando el siguiente procedimiento: para eliminar la x en la segunda ecuación le restamos la primera; como x aparece en la tercera ecuación con un factor 2 multiplicamos la primera ecuación por 2, y luego la restamos de la tercera. Obtenemos

$$\left\{\begin{array}{rcl} x+y+2z &=& 8, \\ y-z &=& -8, \\ -y-3z &=& -12. \end{array}\right.$$

Observación 1.2.10. Si estuviéramos trabajando con el sistema general de ecuaciones (1.11), restarímos a la i-ésima ecuación, para $i=2,3,\ldots,m$, el resultado de multiplicar la primera por el cociente a_{i1}/a_{11}, entre el coeficiente a_{i1} que aparece multiplicando a x_1 en la ecuación i, y el coeficiente a_{11} que x_1 tiene en la primera ecuación. Si x_1 no aparece en la primera ecuación entonces a_{11} es cero y no podemos hacer este cálculo. En este caso cambiaremos el orden de las ecuaciones, y pondremos en primer lugar una ecuación en que aparezca la variable x_1. ♠

El resultado del paso 1 es que x no aparece en la segunda ni en la tercera ecuación. Si miramos esas dos ecuaciones aisladamente nuestro problema se reduce al sistema 2×2

$$\begin{cases} y - z = -8, \\ -y - 3z = -12. \end{cases}$$

con incógnitas y y z. Una vez resuelto este problema, más pequeño que el original, podremos calcular x usando la primera ecuación. De esta manera el algoritmo de eliminación va reduciendo el tamaño del problema.
PASO 2. Haremos desaparecer la variable y de la tercera ecuación. Simplemente sumamos la segunda ecuación a la tercera, y el sistema se transforma en

$$\begin{cases} x + y + 2z = 8, \\ y - z = -8, \\ -4z = -20. \end{cases}$$

Nuestro sistema ya alcanzó la forma escalerizada que lo vuelve de fácil resolución, por lo que detenemos aquí nuestro algoritmo.
PASO 3. Resolvemos este último sistema. De la tercera ecuación concluimos $z = 5$. Sustituyendo hacia arriba y despejando de la segunda ecuación obtenemos $y = -3$. Finalmente, usamos los valores de hallados en la primera ecuación, para concluir que $x = 1$.
ÚLTIMO PASO. VERIFICACIÓN. Parece que hemos resuelto el problema. Pero, ¿qué nos garantiza que las soluciones halladas son las soluciones del problema original? Podemos verificar que es así, sustituyendo los valores

$$x = 1, \quad y = -3, \quad z = 5,$$

en el sistema:

$$\begin{cases} 1 \times 1 + 1 \times (-3) + 2 \times 5 = 8, \\ 1 \times 1 + 2 \times (-3) + 1 \times 5 = 0, \\ 2 \times 1 + 1 \times (-3) + 1 \times 5 = 4. \end{cases}$$

Vemos que, efectivamente, nuestra presunta solución es realmente solución del sistema.

Estupendo, hemos hallado la solución del sistema. ¡Un momento! Sólo hemos verificado que lo que tenemos es solución. Pero, ¿será realmente **la** solución? ¿No podrá haber otras? ¿Qué asegura que nuestras transformaciones no modifican el conjunto de soluciones del sistema?

Observación 1.2.11. El procedimiento de sustituir las soluciones halladas para ver si las ecuaciones se satisfacen o no permite verificar si hemos encontrado soluciones del sistema. Pero no nos permite saber si hay o no más soluciones. Esta limitación de la verificación no debe hacernos perder de vista que es muy recomendable verificar nuestros cálculos siempre que sea posible hacerlo. Más aún, en algunos casos es imperdonable el error de aportar una solución inexacta cuando es posible verificar su corrección.

Este comentario puede extenderse a muchos aspectos de la vida estudiantil y profesional. Es importante tratar de verificar la exactitud de las afirmaciones que se hacen, cotejar las fuentas, estudiar la confiabilidad de los datos de que se dispone, evaluar si el “software” con el que se pretende abordar una tarea es adecuado para ella, etcétera. ♠♣

Las transformaciones que hemos realizado en cada paso del ejemplo 1.2.9 consisten en repetir una y otra vez la misma operación: sumar a una de las ecuaciones un múltiplo de otra. La gracia del método está en escoger hábilmente el factor por el que se multiplica la ecuación a sumar, de modo de “hacer desaparecer” una incógnita en la ecuación resultante.

Mostraremos a continuación que esta operación no modifica el conjunto de soluciones del sistema: el nuevo sistema que obtenemos es *equivalente* al original, en el sentido de que ambos **tienen exactamente el mismo conjunto solución**. Naturalmente, luego de aplicar una serie de transformaciones que preservan la equivalencia de sistemas el sistema final es equivalente al sistema original: no hemos agregado ni perdido soluciones en todo el proceso.

Proposición 1.1. *Si en un sistema lineal de ecuaciones se sustituye una ecuación por el resultado de sumar a esa misma ecuación un múltiplo de otra, entonces el sistema resultante tiene el mismo conjunto solución que el sistema original.*

Prueba. Supongamos que el sistema original sea

$$(S)\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i, \\ \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n = b_j, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{array}\right.$$

y que sumamos a la j-ésima ecuación el resultado de multiplicar ambos miembros de la ecuación i-ésima por un número β. Obtenemos un nuevo sistema

$$(S')\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i, \\ \vdots \\ a_{j1}x_1 + \ldots + a_{jn}x_n + \beta(a_{i1}x_1 + \ldots + a_{in}x_n) = b_j + \beta b_i, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{array}\right.$$

Para probar que ambos son equivalentes deberemos ver que cualquier solución de S es solución de S' y viceversa.

Veamos primero que cualquier solución de S es solución de S'.

Sea $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ una solución de S. Es claro que $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ satisface todas las ecuaciones de S' salvo, tal vez, la j-ésima, pues son las mismas que las de S. Como $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ debe verificar la i-ésima y j-ésima ecuación de S se tiene que

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n = b_i, \quad a_{j1}\alpha_1 + a_{j2}\alpha_2 + \ldots + a_{jn}\alpha_n = b_j.$$

Multiplicando ambos miembros de la primera igualdad por β y sumando miembro a miembro, se deduce inmediatamente que

$$a_{j1}\alpha_1 + \ldots + a_{jn}\alpha_n + \beta(a_{i1}\alpha_1 + \ldots + a_{in}\alpha_n) = b_j + \beta b_i.$$

Por lo tanto, también la j-ésima ecuación de S' se satisface.

La verificación de que cualquier solución de S' es también solución de S emplea esencialmente el mismo argumento. Igual que antes es claro que una solución $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ de S' debe verificar todas las ecuaciones de S salvo tal vez la j-ésima. Por la j-ésima ecuación en S' sabemos que

$$a_{j1}\alpha_1 + \ldots + a_{jn}\alpha_n + \beta(a_{i1}\alpha_1 + \ldots + a_{in}\alpha_n) = b_j + \beta b_i,$$

en tanto que la i-ésima implica

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n = b_i.$$

Multiplicamos ambos miembros de la i-ésima por β, y restamos de la ecuación que ocupa el lugar j, para concluir

$$a_{j1}\alpha_1 + a_{j2}\alpha_2 + \ldots + a_{jn}\alpha_n = b_j.$$

Esto muestra que $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ es también una solución de S, y la prueba esta concluída. $\square$

La proposición 1.1 es el fundamento teórico del método de eliminación que hemos presentado en el ejemplo 1.2.9: nos asegura que en cada paso hemos respetado la equivalencia entre el sistema de partida y el de llegada. Ahora sí, finalmente estamos seguros de que el conjunto solución del sistema en ese ejemplo es $\{(1, -5, 3)\}$, y no hay soluciones diferentes a $(1, -5, 3)$. ¡Bravo!

Observación 1.2.12. No es esencial para nuestro argumento que las ecuaciones del sistema S sean lineales. La proposición es también cierta para sistemas de ecuaciones no lineales, de la forma $g_i(x_1, x_2, \ldots, x_n) = b_i$, $i = 1, 2, \ldots, m$, donde las expresiones g_i pueden depender de las variables x_j, $j = 1, 2, \ldots, n$, de cualquier manera. Sin embargo, a diferencia de lo que ocurre para sistemas lineales, no puede obtenerse un método sistemático de resolución de sistemas cualesquiera a partir de esta proposición.

En el proceso de eliminación hay veces en que es imprescindible intercambiar ecuaciones. Junto con la transformación de sumar a una ecuación un múltiplo de otra, esto es todo lo que necesitaremos para el método de eliminación de Gauss, o escalerización. Además de estas operaciones recurriremos a la de multiplicar una ecuación por algún número distinto de cero. Por ejemplo, cuando todos los coeficientes de una ecuación son negativos en general simplifica las cosas multiplicar la ecuación por -1. O cuando se quiere conseguir que el primer coeficiente no nulo de una ecuación sea igual a 1. Llamaremos transformaciones elementales a estas operaciones.

En resumen, llamaremos *transformaciones elementales* a las siguientes operaciones:

1. sumar a una ecuación un múltiplo de otra;
2. intercambiar de lugar dos ecuaciones;
3. multiplicar una ecuación por un número distinto de cero,

que pueden efectuarse sobre las ecuaciones de un sistema lineal. Nuestro interés por estas transformaciones proviene de que por medio del método de eliminación nos permiten transformar cualquier sistema lineal en un sistema equivalente de fácil resolución. La equivalencia está asegurada por nuestra siguiente proposición y su corolario.

Proposición 1.2. *Si a un sistema de ecuaciones le aplicamos una operación elemental obtenemos un sistema equivalente al original.*

Corolario 1.3. *Si transformamos un sistema de ecuaciones aplicando una secuencia de transformaciones elementales obtenemos un sistema equivalente al original.*

Parte de la prueba de la proposición que acabamos de enunciar y su corolario está contenida en la proposición 1.1. Dejamos el resto como ejercicio para el lector.

Ejercicio 1.13.

1. Completar las pruebas de la proposición 1.2 y el corolario 1.3.
2. ¿Qué puede ocurrir si en un sistema se multiplica una ecuación por 0?
3. Mostrar que si una ecuación es sustituida por el resultado de multiplicarla por un número distinto de cero y sumarle un múltiplo de otra ecuación cualquiera del sistema, entonces el sistema resultante es equivalente al original.

Observación 1.2.13. Notemos que cuando una ecuación es sustituida por el resultado de sumarle un múltiplo de otra la información contenida en la ecuación eliminada no desaparece: está implícita en la nueva ecuación, porque la ecuación sustituida entró con un coeficiente no nulo en la formación de esta nueva ecuación (en la mayoría de los casos este coeficiente será igual a 1). Esta nueva ecuación, junto con la que se usó en la combinación lineal, permiten recuperar la ecuación eliminada. ♠

1.2.3. Notación matricial

Tal como apuntábamos en la observación 1.2.8, al trabajar con un sistema de ecuaciones la representación de las incógnitas (x, y, z, o $x_1, \dots, x_n$, etctera.) no desempeña ningún papel importante y puede usarse cualquiera. De hecho, para determinar las soluciones de un sistema es suficiente con conocer los coeficientes y el término independiente del mismo. Naturalmente, no alcanza con conocer los valores de estos números, sino que es necesario saber el lugar que ocupan: a qué ecuación pertenecen y a qué incógnita multiplican. Por esto resulta conveniente introducir la noción de *matriz*. Una matriz es un arreglo de números en filas y columnas, que nos permite almacenar y manejar ordenadamente información numérica. En el caso de los sistemas de ecuaciones lineales cada fila corresponderá a una misma ecuación, y cada columna a una misma incógnita.

Las matrices no son sólo útiles para trabajar con sistemas lineales. En la sección 2.1 mostramos varios ejemplos en los que empleamos matrices para ordenar la información en distintos problemas. En cada caso filas y columnas adquieren distintos significados, pero el ordenamiento de los datos en una matriz es común a todos los ejemplos que allí se tratan.

Representaremos una matriz A de m filas y n columnas (o simplemente matriz de dimensiones $m \times n$, o matriz $m \times n$) de entradas

$$a_{ij}, \quad i = 1, 2 \ldots, m, \; j = 1, 2, \ldots, n,$$

con la notación

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

En este texto, salvo mención expresa en sentido contrario, las entradas a_{ij} serán números. Es decir, serán elementos de un cierto cuerpo $\mathbb{K}$.

La formalización[5] de estas nociones de matriz como una arreglo ordenado de números está contenida en nuestra próxima definición.

Definición 1.1 (Matrices $m \times n$ sobre un cuerpo $\mathbb{K}$). *Una **matriz** A de m filas y n columnas sobre el cuerpo $\mathbb{K}$ es una función*

$$A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{K}.$$

Observación 1.2.14. Sobre la definición de las matrices y su notación

El conjunto

$$\{1, \ldots, m\} \times \{1, \ldots, n\}$$

está formado por todas las posibles parejas de números naturales (i, j), con i variando entre 1 y m, y j haciéndolo entre 1 y n. Por lo tanto, al especificar que una matriz $m \times n$ es una función definida sobre ese conjunto y con valores en $\mathbb{K}$ estamos diciendo que una matriz A de dimensiones $m \times n$ queda definida cuando a cada una de estas parejas (i, j) le asignamos un valor $A(i, j)$ en $\mathbb{K}$.

Una manera de hacer esto es escribir explícitamente cada valor, por ejemplo

$$\begin{array}{lll} A(1,1) = 1/3, & A(1,2) = 1/4, & A(1,3) = 1/5, \\ A(2,1) = 1/5, & A(2,2) = 1/6, & A(2,3) = 1/7, \end{array}$$

[5] Casi no recurriremos a esta definición formal para trabajar con las matrices, pero es necesaria para sacarlas de la ambigüedad de expresiones como "ordenamiento de números", "arreglos en filas y columnas"que, aunque son de gran ayuda para trabajar una vez que uno sabe exactamente de qué está hablando, no constituyen una definición sobre la que podamos construir una teoría sólida.

algo que se resume mucho mejor en la notación usual para las matrices

$$A = \begin{pmatrix} 1/3 & 1/4 & 1/5 \\ 1/5 & 1/6 & 1/7 \end{pmatrix}, \tag{1.15}$$

en que la ubicación de cada número dice cuáles son los índices que le corresponden: el primer índice i corresponde a las filas, que se cuentan de arriba a abajo, y el segundo índice j corresponde a la columnas, que se cuentan de izquierda a derecha.

Naturalmente, esta manera de definir una matriz corresponde al hecho conocido de que una función queda definida si damos una lista explícita en la que decimos que elemento del codominio corrresponde a cada elemento del dominio. También se puede especificar una matriz dando una fórmula que permita calcular la entrada $A(i,j)$ en función de i y de j. Por ejemplo

$$A(i,j) = 1/(2i+j), \quad i = 1,2,\ j = 1,2,3,$$

es otra manera de describir la matriz (1.15). El lector puede verificarlo evaluando la fórmula para los seis posibles pares (i,j). Tampoco sorprenderá esto al lector, que habrá visto ya muchísimas funciones definidas por fórmulas que permiten calcular su valor para cada elemento del dominio.

Vale la pena mencionar que no es $A(i,j)$ la notación más corriente en la literatura para designar a las entradas de una matriz. Mucho más frecuente es denominar a una matriz con una letra mayúscula, como A, por ejemplo. Y a sus entradas con una expresión como a_{ij}, la letra minúscula correspondiente acompañada de los subíndices i y j. Esta es la notación que utilizaremos más frecuentemente en este texto. Por supuesto, indicaremos con a_{ij} la entrada $A(i,j)$, que está en la fila i y la columna j.

Sin embargo, a la hora de escribir programas para una computadora se impone una notación como $A(i,j)$ que tiene mucho menos ambigüedad, ya que sólo consiste del nombre de la matriz evaluado en los índices de su dominio. Esta notación es útil para almacenar matrices con nombre nemotécnicos en la memoria de una computadora y para aludir claramente a sus entradas. Por ejemplo, si tenemos que almacenar una matriz de costos es una buena idea llamarla `costos`. Entonces la entrada que ocupa la fila 3 y la columna 4 de la matriz `costos` es simplemente `costos(3,4)`. El lector encontrará que esta notación es de uso corriente en la mayoría del "software" existente.

En algunos casos escribiremos una matriz A con entradas a_{ij} con una notación abreviada, sin desplegar todas sus filas y columnas, como

$$A = (a_{ij})_{j=1,\dots,n}^{i=1,\dots,m}.$$

Con esta notación la matriz (1.15) se escribe

$$A = (1/(2i+j))_{j=1,2,3}^{i=1,2}\,.$$

Esperamos que estas distintas maneras de definir una matriz y las diferentes notaciones existentes no confundan al lector. Luego agregaremos alguna notación más: escribiendo las matrices en términos de sus filas o de sus columnas. Todas estas notaciones serán útiles en algún contexto, y, en realidad sólo reflejan el hecho de que se puede aludir a un mismo objeto de diferente maneras. También cambiaremos de notación de tanto en tanto, porque la representación más conveniente de una matriz dependerá de cuál sea la situación precisa en que estemos trabajando. Por lo tanto, en distintas partes del texto usaremos distintas representaciones y notaciones. ♠

Dos matrices son iguales si y sólo si tienen el mismo tamaño y los mismas entradas en las mismas posiciones. Esto es una consecuencia inmediata de la definición de matrices como funciones, porque dos funciones son iguales si y sólo si tienen el mismo dominio y codominio y toman los mismos valores sobre todos los elementos de su dominio. En particular, dos matrices de distinto tamaño siempre serán distintas, y si A y B son dos matrices que tienen las mismas dimensiones $m \times n$ entonces $A = B$ si y sólo si

$$a_{ij} = b_{ij}, \quad i = 1, \ldots, m,\ j = 1, \ldots, n.$$

Ejemplo 1.2.15. Como ya vimos antes, la matriz

$$(1/(2i+j))_{j=1,3}^{i=1,2}$$

es igual a la matriz (1.15). Si

$$A = \begin{pmatrix} 1 & 2 \\ \sqrt{2} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ \sqrt{2} & 1 \end{pmatrix},$$

entonces $A \neq B$ pues $a_{11} = 1 \neq 2 = b_{11}$. ♣

La introducción de las matrices estuvo motivada por nuestro deseo de almacenar en un único objeto, una matriz, toda la información relevante acerca de un sistema de ecuaciones lineales. Veamos cómo nos ayudan en esta tarea.

Dado un sistema lineal de ecuaciones

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2, \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m, \end{array}\right.$$

llamaremos *matriz del sistema* a la matriz A que tiene dimensión $m \times n$, donde m y n son respectivamente iguales al número de ecuaciones e incógnitas del sistema, y que está formada por sus coeficientes. Esto es:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Llamaremos, *matriz ampliada del sistema*, o simplemente *matriz ampliada* a la matriz $m \times (n+1)$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix},$$

que incorpora los términos independientes. Corrientemente escribiremos esta matriz en la forma

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right),$$

para recordar que estamos considerando la matriz ampliada de un sistema, cuya última columna tiene un significado diferente a las anteriores. También las representaremos en la forma breve $A|B$, donde B representa a la columna, o matriz $m \times 1$ formada por los números b_i, $i = 1, 2, \dots, m$.

Naturalmente, toda la información que necesitamos sobre el sistema lineal está encerrada en la matriz ampliada $A|B$. Veremos más adelante que, en realidad, mucha de la información relevante acerca del sistema sólo depende de la matriz A formada por sus coeficientes.

La definición anterior introduce una notación matricial más compacta para un sistema de ecuaciones, que nos libera de escribir una y otra vez los nombres que hemos escogido para las incógnitas. Veamos un ejemplo.

Ejemplo 1.2.16. Consideremos el sistema

$$\left\{ \begin{aligned} x + y + 2z &= 9, \\ 3x + 6y - 5z &= 0, \\ 2x + 4y - 3z &= 1. \end{aligned} \right.$$

Entonces la matriz del sistema y la matriz ampliada son:

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 6 & -5 \\ 2 & 4 & -3 \end{pmatrix}, \quad A|B = \left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 3 & 6 & -5 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right).$$

Esta notación más compacta simplificará la escritura e implementación del algoritmo de eliminación de Gauss. ♣

Antes de mostrar con un ejemplo la implementación del método de escalerización con la notación matricial vale la pena hacer algunas observaciones.

Observación 1.2.17. Cada fila de la matriz del sistema corresponde a una de las ecuaciones del sistema. Las operaciones del método de escalerización se traducirán entonces en operaciones sobre las filas de la matriz. En particular, las transformaciones elementales serán, en este contexto, las siguentes:

1. sumar a una fila el resultado de multiplicar otra por un número cualquiera;
2. intercambiar de lugar dos filas;
3. multiplicar una fila por un número $\alpha \neq 0$.

Cada incógnita del sistema queda en correspondencia con una de las columnas de la matriz A del sistema. Encontrar una entrada a_{ij} igual a 0 en la matriz A es equivalente a que la incógnita x_j no aparezca en la i-ésima ecuación. ♠

Observación 1.2.18. Con la introducción de las matrices ganamos, para empezar, con la economía en la notación. Pero no es todo. Veremos a lo largo de este curso que las matrices pueden ser consideradas desde varios puntos de vista:

- en muchas situaciones son una forma ordenada de escribir la información disponible;
- pueden ser consideradas como un arreglo de $m \times n$ números;
- como un conjunto ordenado de n columnas;
- como un arreglo de m filas;
- como un arreglo de otras submatrices más pequeñas;
- también como un único objeto, que es además un elemento de una estructura algebraica.

- Finalmente, pueden ser vistas como la definición de una transformación de $\mathbb{K}^n$ en $\mathbb{K}^m$.

Iluminar las relaciones entre ellos será una buena parte del trabajo que tenemos por delante. Encontraremos además que al hacerlo lograremos entender mejor algunas de las aplicaciones del álgebra lineal y del cálculo con matrices. ♠

Ejemplo 1.2.19. La escalerización en notación matricial
Vamos a reescribir los pasos del ejemplo (1.2.9) en notación matricial. Esto permite conseguir bastante economía en la notación. El sistema original queda expresado en la forma:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 8 \\ 1 & 2 & 1 & 0 \\ 2 & 1 & 1 & 4 \end{array}\right).$$

Cada ecuación está ahora representada por una fila de coeficientes en la matriz ampliada del sistema.

Observación 1.2.20. Un poco de jerga.
Llamaremos *pivote* a la primera[6] entrada no nula de cada fila de una matriz. Con esta terminología podemos describir la eliminación de Gauss como un procedimiento que busca, a través de la aplicación reiterada de las transformaciones elementales, dejar un único pivote por columna. En cada paso de la escalerización, una vez escogido un pivote, lo empleamos para eliminar todas los otros pivotes de la misma columna. A partir de ahora, usaremos negrita para resaltar el pivote de cada fila. ♠

Paso 1. Las operaciones que antes hacíamos con las ecuaciones deben hacerse ahora sobre filas de coeficientes. Escogemos el pivote 1 en la primera fila, y restamos entonces esta primera fila de la segunda para hacer aparecer un cero en el primer lugar de la segunda fila (esto es equivalente a hacer desaparecer la variable x en la ecuación correspondiente). Luego multiplicamos la primera fila por 2 y la restamos de la tercera. El resultado es

$$\left(\begin{array}{rrr|r} \mathbf{1} & 1 & 2 & 8 \\ 0 & 1 & -1 & -8 \\ 0 & -1 & -3 & -12 \end{array}\right).$$

Sólo un pivote, el 1 de la primera fila, aparece en la primera columna.

[6]El orden que estamos empleando es el obvio: es el orden en que están numeradas las columnas.

Paso 2. Sumamos la segunda fila a la tercera y llegamos a la matriz escalerizada

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 8 \\ 0 & \mathbf{1} & -1 & -8 \\ 0 & 0 & -4 & -20 \end{array}\right). \tag{1.16}$$

Observemos que con la representación matricial hemos aligerado bastante la notación.

La nueva matriz ampliada (1.16) es una representación del sistema de ecuaciones

$$\left\{\begin{array}{rcr} x+y+2z & = & 8, \\ y-z & = & -8, \\ -4z & = & -20, \end{array}\right.$$

cuya solución podemos calcular directamente. La tercera ecuación implica

$$z=5.$$

Al sustituir en la segunda podemos despejar

$$y=-3.$$

Finalmente, la primera implica $x-3+2\times 5=8$, por lo que

$$x=1.$$

El procedimiento de escalerizar y resolver el sistema escalerizado sustituyendo hacia arriba nos permite hallar el conjunto solución del sistema original. ♣

Llamaremos *matriz escalerizada* a una matriz que cumpla las siguientes condiciones:

1. todas las filas, salvo quizás la primera, comienzan con una sucesión de ceros;

2. cada fila tiene al principio por lo menos un cero más que la fila inmediata superior.

Escalerizar una matriz es llevarla a una forma escalerizada por medio de transformaciones elementales (ver la observación 1.2.17, en la página 41). Si E es una matriz que se obtiene escalerizando otra matriz A, entonces diremos que E es una *forma escalerizada* de A. Diremos que un *sistema está escalerizado* si su matriz ampliada lo está. *Escalerizar un sistema* es encontrar otro sistema escalerizado equivalente. Naturalmente, escalerizar un sistema es equivalente a escalerizar la matriz ampliada del sistema.

Observación 1.2.21. OPERACIONES CON LISTAS DE NÚMEROS.
Notemos que hemos introducido, casi sin darnos cuenta, algunas operaciones algebraicas con las filas de las matrices asociadas a los sistemas lineales.

Por ejemplo en el primer paso del ejemplo 1.2.19 nos referimos al resultado de multiplicar por dos cada uno de los números en la fila $(1, 1, 2, 8)$ para obtener la fila $(2, 2, 4, 16)$ como la operación de *multiplicar por* 2 *la fila* $(1, 1, 2, 8)$. Es decir, interpretamos que

$$2 \times (1, 1, 2, 8) = (2 \times 1, 2 \times 1, 2 \times 2, 2 \times 8) = (2, 2, 4, 16).$$

También nos hemos referido al resultado de sumar entrada a entrada las dos filas $(0, 1, -1, -8)$ y $(0, -1, -3, -12)$ como a la operación de *sumar ambas filas.* En otras palabras

$$(0, 1, -1, -8) + (0, -1, -3, -12) = (0, 0, -4, -20).$$

En general, observemos que todo esto es equivalente a definir las siguientes operaciones para filas (listas) de cuatro números:
SUMA DE DOS LISTAS:

$$(x_1, x_2, x_3, x_4) + (y_1, y_2, y_3, y_4) = (x_1 + y_1, x_2 + y_2, x_3 + y_3, x_4 + y_4).$$

PRODUCTO DE UNA LISTA POR UN NÚMERO:

$$\alpha(x_1, x_2, x_3, x_4) = (\alpha\, x_1, \alpha\, x_2, \alpha\, x_3, \alpha\, x_4).$$

Estas operaciones entre filas, o listas, de números son bastante naturales, y, por supuesto, su definición puede extenderse a listas de longitud cualquiera, tanto filas (listas escritas en forma horizontal) como columnas (escritas en vertical). **La existencia de estas operaciones dota al conjunto $\mathbb{K}^n$ formado por listas de n números en el cuerpo $\mathbb{K}$ de una estructura algebraica**. Por eso nos referiremos a este conjunto como el espacio, o espacio vectorial, $\mathbb{K}^n$. Seguiremos haciendo uso de estas operaciones en esta sección y en el resto del texto, y desarrollaremos más estas ideas en la subsección 1.2.6 y secciones siguientes. Por ahora vale la pena ir subrayando la idea de que **llamaremos *espacio* a un conjunto que está dotado de algún tipo de estructura** . En el caso de $\mathbb{K}^n$ nuestro principal objeto de trabajo será la estructura algebraica que se obtiene a partir de las operaciones de suma y producto por un número. Esta estructura nos será muy útil para describir las soluciones de los sistemas de ecuaciones lineales, algo que iremos mostrando a lo largo de este capítulo. Veremos más adelante que la estructura de $\mathbb{K}^n$ es un caso particular de una estructura general, la de *espacio vectorial*, cuya teoría desarrollaremos en el capítulo dedicado a la teoría abstracta de los espacios vectoriales. ♠

Pondremos en práctica la técnica con un nuevo ejemplo.

Ejemplo 1.2.22. Resolver en $\mathbb{R}$ el sistema:

$$\left\{\begin{array}{rcr} x_1 + 2x_2 + x_3 + x_5 + x_6 & = & 1 \\ -x_1 - 2x_2 + x_3 + x_4 - x_6 & = & 0 \\ x_1 + 2x_2 + x_3 + 2x_4 + 5x_5 + 3x_6 & = & 1 \\ x_1 + 2x_2 + x_3 + x_5 + 3x_6 & = & 3 \\ x_1 + 2x_2 + x_3 + 4x_4 + 9x_5 + 3x_6 & = & -1 \end{array}\right. \tag{1.17}$$

La matriz ampliada del sistema es:

$$\left(\begin{array}{rrrrrr|r} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 & 0 \\ 1 & 2 & 1 & 2 & 5 & 3 & 1 \\ 1 & 2 & 1 & 0 & 1 & 3 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 & -1 \end{array}\right).$$

El proceso comienza fijando la entrada no nula de la primera fila como pivote y utilizándola para lograr ceros en el resto de la primera columna. El resultado es la matriz

$$\left(\begin{array}{cccccc|r} \mathbf{1} & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 4 & 8 & 2 & -2 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 + F_1 \\ \longleftarrow F_3 - F_1 \\ \longleftarrow F_4 - F_1 \\ \longleftarrow F_5 - F_1. \end{array}$$

Hemos anotado al margen de la matriz las transformaciones que en este primer paso de la escalerización fueron hechas a la matriz ampliada del sistema. Creemos que la notación empleada se explica por sí sola.

En el ejemplo 1.2.19, se utilizaba la entrada a_{22} en el segundo paso de la escalerización para conseguir ceros en la segunda columna. Aquí esto no es posible pues la segunda columna ya tiene todas sus entradas, salvo la primera, nulas. En consecuencia, el segundo escalón deberá estar en la tercera columna, donde aparece la entrada no nula a_{23}. Por debajo de a_{23} sólo hay ceros, así que continuamos nuestro algoritmo usando el pivote 2 de la entrada a_{34}. He aquí la matriz que se obtiene, junto con la indicación de las operaciones realizadas:

$$\left(\begin{array}{cccccr|r} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & \mathbf{2} & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & \mathbf{2} & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & -2 & -2 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5 - 2F_3 \end{array}$$

En el tercer paso operaremos sobre la sexta columna:

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{2} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5 + F_4. \end{array}$$

Ya tenemos la forma escalerizada, con pivotes en las columnas primera, tercera, cuarta y sexta. El sistema lineal, equivalente al sistema original en el sentido de que tiene exactamente las mismas soluciones, que corresponde a esta matriz es

$$\begin{array}{ccccccccccccc} x_1 & + & 2x_2 & + & x_3 & & & + & x_5 & + & x_6 & = & 1, \\ & & & & 2x_3 & + & x_4 & + & x_5 & & & = & 1, \\ & & & & & & 2x_4 & + & 4x_5 & + & 2x_6 & = & 0, \\ & & & & & & & & & & 2x_6 & = & 2. \end{array}$$

De la última ecuación resulta $x_6 = 1$. Con esta información vamos a la tercera y concluimos

$$x_4 + 2x_5 = -1. \tag{1.18}$$

Esto no permite determinar x_4 ni x_5, pero podemos dejar expresada una incógnita en términos de la otra. Expresemos x_4, una variable que corresponde a una columna con un pivote en la forma escalerizada, en términos de x_5, como

$$x_4 = -1 - 2x_5.$$

Sustituyendo en la segunda ecuación obtenemos

$$x_3 = \frac{x_5}{2} + 1.$$

Combinando toda esta información con la primera ecuación resulta

$$x_1 + 2x_2 + \frac{3x_5}{2} + 1 = 0. \tag{1.19}$$

Nuevamente escogemos despejar la variable que corresponde al pivote para escribir

$$x_1 = -2x_2 - \frac{3x_5}{2} - 1.$$

Concluimos entonces que todas las soluciones del sistema son de la forma

$$(-1 - 2x_2 - 3x_5/2, x_2, 1 + x_5/2, -1 - 2x_5, x_5, 1),$$

donde x_2 y x_5 son dos parámetros reales que podemos fijar a nuestro antojo. La solución no es única, pero todas las soluciones pueden escribirse en términos de las variables x_2 y x_5: cada pareja de valores reales elegidos para x_2 y x_5 determina una solución del sistema. Por ejemplo, la elección más sencilla de todas es

$$x_2 = x_5 = 0,$$

que da lugar a

$$(-1, 0, 1, -1, 0, 1).$$

Para $x_2 = 1$ y $x_5 = 0$ se obtiene la solución

$$(-3, 1, 1, -1, 0, 1).$$

Si $x_2 = 0$ y $x_5 = 1$ se tiene la solución

$$(-5/2, 0, 3/2, -3, 1, 1)\,.$$

El sistema es compatible e indeterminado, porque tiene más de una solución. Además sabemos como construirlas todas. El proceso de escalerización selecciona a las variables x_2 y x_5 como variables libres, y en términos de x_2 y x_5, cuyos valores pueden ser fijados arbitrariamente, hemos escrito todas las soluciones. El sistema tiene entonces dos grados de libertad. No nos extendemos mucho más sobre esto, porque será el objeto de la sección 1.4. ♣

Observación 1.2.23. Variables libres.
La razón para despejar las variables de las columnas con pivotes (x_4 de la ecuación (1.18), x_1 de (1.19)) es que siempre es posible despejar la variable que está multiplicada por el pivote en términos de las otras porque su coeficiente, el pivote, es no nulo. Despejar las otras variables podría ser imposible. He aquí un ejemplo: si la forma escalerizada de una matriz ampliada es

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{array}\right)$$

entonces podemos despejar $x_3 = -2x_4$, o $x_4 = -x_3/2$ de la última ecuación. En la primera podemos despejar $x_1 = -x_3$, pero es imposible expresar x_2 en términos de las restantes variables.

Vamos a introducir la denominación *variables libres* para las incógnitas que corresponden a columnas sin pivotes. Enfaticemos que **siempre es posible expresar las soluciones del sistema en términos de las variables libres**. Cada una de estas variables corresponde a un grado de libertad dentro del conjunto solución del sistema. Las variables x_2 y x_5 son las variables libres en el ejemplo 1.2.22.

El hecho de que las soluciones de un sistema puedan expresarse en términos de las variables libres que selecciona el procedimiento de eliminación gaussiana no excluye la posibilidad de que las soluciones puedan expresarse también términos de otras variables, que incluyan a algunas de las que corresponden a columnas con pivotes. El próximo ejercicio apunta a mostrar este fenómeno.

Ejercicio 1.14. Mostrar que todas las soluciones del sistema (1.17) del ejemplo 1.2.22 también pueden escribirse en términos de las variables x_1 y x_4, pero que es imposible expresarlas en función de x_1 y x_2. ♠

Observación 1.2.24. Cuando un sistema es compatible es determinado, es decir, tiene una única solución, si y sólo si no quedan variables libres que podamos fijar a nuestro antojo. Esto es equivalente a que el número de pivotes sea igual al número de columnas. Esta observación tiene una importante consecuencia: un sistema lineal de ecuaciones que tiene más incógnitas que ecuaciones no puede ser determinado.

Un sistema no lineal no tiene esta propiedad. Por ejemplo, una única ecuación $(x-1)^2+(y+2)^2=0$ determina completamente los valores de las dos incógnitas x e y como $x=1$, $y=-2$. ♠

Veamos un ejemplo más en el que trataremos un sistema trabajando sobre su representación matricial.

Ejemplo 1.2.25. Resolveremos en $\mathbb{Z}_2$ el sistema

$$\left\{\begin{aligned} x_1+x_2+x_4+x_5&=1,\\ x_2+x_3+x_4&=1,\\ x_1+x_3+x_5&=1,\\ x_2+x_3+x_5&=1. \end{aligned}\right. \tag{1.20}$$

La matriz ampliada del sistema es

$$\left(\begin{array}{ccccc|c} 1&1&0&1&1&1\\ 0&1&1&1&0&1\\ 1&0&1&0&1&1\\ 0&1&1&0&1&1 \end{array}\right).$$

La llevaremos a una forma escalerizada. Comenzamos por sumar la primera fila a la tercera:

$$\left(\begin{array}{ccccc|c} \mathbf{1}&1&0&1&1&1\\ 0&1&1&1&0&1\\ 0&1&1&1&0&0\\ 0&1&1&0&1&1 \end{array}\right).$$

Ahora sumaremos la segunda a la tercera y la cuarta:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & \mathbf{1} & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right).$$

Para obtener la forma escalerizada intercambiamos las filas tercera y cuarta:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

Este sistema es incompatible, porque la cuarta ecuación es equivalente a

$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = 1.$$

Como el miembro de la izquierda vale 0 para cualquier elección de las incógnitas, esta ecuación no puede satisfacerse. El conjunto solución del sistema es el conjunto vacío

Subrayemos que es el sistema el que no tiene solución, ¡pero el problema de hallar las soluciones del sistema sí la tiene! El conjunto solución es el conjunto vacío.

Ejercicio 1.15. Resolver el sistema (1.20), pero cambiando los segundos miembros de todas las ecuaciones del sistema de forma tal que la última columna de la matriz ampliada del nuevo sistema sea $(0, 1, 1, 0)$. ♣

Observación 1.2.26. COMPATIBILIDAD Y PIVOTES
La incompatibilidad del sistema (1.20) del ejemplo 1.2.25 se manifiesta en que la forma escalerizada de la matriz ampliada del sistema tiene más pivotes que la forma escalerizada de la matriz del sistema. Hay al menos una fila en que la matriz del sistema tiene todas sus entradas nulas, pero en la columna que corresponde al término independiente, la que se agrega para formar la matriz ampliada, aparece una entrada b no nula. Esta fila corresponde a una ecuación

$$0 = b \neq 0,$$

que no puede ser satisfecha para ningún valor de las incógnitas. Este hecho es completamente general, tal como el lector podrá apreciar en nuestro próximo ejercicio.

Ejercicio 1.16. Resolver el sistema

$$\left\{\begin{array}{rcl} x_1+2x_2+x_3+x_5+x_6 &=& 1\\ -x_1-2x_2+x_3+x_4-x_6 &=& 0\\ x_1+2x_2+x_3+2x_4+5x_5+3x_6 &=& 1\\ x_1+2x_2+x_3+x_5+3x_6 &=& 3\\ x_1+2x_2+x_3+4x_4+9x_5+3x_6 &=& 1 \end{array}\right.$$

Comparar el resultado con el obtenido para el sistema del ejemplo 1.2.22. ♠

Ejemplo 1.2.27. Resolución del problema de la red vial
En la sección 1.1.2, ver la página 6 y siguientes, consideramos el problema de determinar el flujo del tráfico sobre una red vial. Para resolverlo habíamos formulado un sistema de seis ecuaciones con siete incógnitas que representaban los valores medios de tráfico en distintos puntos de la red (en la tabla de la página 7 se explica el significado de cada variable. En el resto del ejemplo nos referiremos a la información que allí se resume). Este sistema es

$$\begin{aligned} -x_1+x_4&=200,\\ x_2-x_4+x_5&=100,\\ x_3+x_5&=700,\\ x_1-x_6&=100,\\ x_2-x_6+x_7&=600,\\ x_3+x_7&=900, \end{aligned}$$

que puede ser escrito en la forma matricial

$$\left(\begin{array}{rrrrrrr|r} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & \\ \hline -1 & 0 & 0 & 1 & 0 & 0 & 0 & 200\\ 0 & 1 & 0 & -1 & 1 & 0 & 0 & 100\\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 700\\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 100\\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 600\\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 900 \end{array}\right).$$

Nos habíamos planteado las siguientes preguntas: ¿es posible cortar completamente el tráfico en la calle Schiaffino entre Varela y Gambetta y atender la demanda que plantea la circulación de vehículos en la hora pico? Si no es posible, ¿qué medida es conveniente adoptar para minimizar el tráfico por esa calle? Entonces, nuestro objetivo es hallar una solución del sistema en la que ninguna de las incógnitas sea negativa y para la que x_4 tenga el mínimo valor posible. Si consiguiéramos $x_4=0$ sería óptimo. Pero no nos adelantemos a los acontecimientos, veamos qué es lo que nuestro modelo predice.

Para estudiar el sistema procedemos a escalerizar la matriz. Indicaremos sobre el margen derecho de cada matriz cuáles fueron las operaciones hechas en cada paso, sin más comentarios. El pivote de la fila que se usó en ese paso de la eliminación aparece en negrita.

$$\left(\begin{array}{rrrrrrr|r} \mathbf{-1} & 0 & 0 & 1 & 0 & 0 & 0 & 200 \\ 0 & 1 & 0 & -1 & 1 & 0 & 0 & 100 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 700 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 300 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 600 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 900 \end{array}\right) \begin{array}{l} \\ \\ \\ \longleftarrow F_4 + F_1. \\ \\ \\ \end{array}$$

$$\left(\begin{array}{rrrrrrr|r} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 200 \\ 0 & \mathbf{1} & 0 & -1 & 1 & 0 & 0 & 100 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 700 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 300 \\ 0 & 0 & 0 & 1 & -1 & -1 & 1 & 500 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 900 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5 - F_2. \\ \\ \end{array}$$

$$\left(\begin{array}{rrrrrrr|r} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 200 \\ 0 & 1 & 0 & -1 & 1 & 0 & 0 & 100 \\ 0 & 0 & \mathbf{1} & 0 & 1 & 0 & 0 & 700 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 300 \\ 0 & 0 & 0 & 1 & -1 & -1 & 1 & 500 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 200 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \\ \longleftarrow F_6 - F_3. \end{array}$$

$$\left(\begin{array}{rrrrrrr|r} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 200 \\ 0 & 1 & 0 & -1 & 1 & 0 & 0 & 100 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 700 \\ 0 & 0 & 0 & \mathbf{1} & 0 & -1 & 0 & 300 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 200 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 200 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5 - F_4. \\ \\ \end{array}$$

$$\left(\begin{array}{rrrrrrr|r} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 200 \\ 0 & 1 & 0 & -1 & 1 & 0 & 0 & 100 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 700 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 300 \\ 0 & 0 & 0 & 0 & \mathbf{-1} & 0 & 1 & 200 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \\ \longleftarrow F_6 - F_5. \end{array}$$

Ahora que hemos escalerizado el sistema podemos ver que es compatible e indeterminado.

Ejercicio 1.17. Se trata de un sistema de 6 ecuaciones y 7 incógnitas, ¿Podría haber sido compatible determinado? ¿E incompatible?.

Si pasamos de la representación matricial a escribir el sistema de ecuaciones que resulta de la escalerización, obtenemos

$$\begin{aligned} -x_1 + x_4 &= 200, \\ x_2 - x_4 + x_5 &= 100, \\ x_3 + x_5 &= 700, \\ x_4 - x_6 &= 300, \\ -x_5 + x_7 &= 200. \end{aligned}$$

Las soluciones pueden expresarse en términos de las variables libres x_6 y x_7, en la forma

$$\begin{aligned} x_1 &= x_6 + 100, \\ x_2 &= x_6 - x_7 + 600, \\ x_3 &= -x_7 + 900, \\ x_4 &= x_6 + 300, \\ x_5 &= x_7 - 200. \end{aligned}$$

Para minimizar x_4 basta escoger el menor valor posible de x_6. Ya que x_6 no puede ser negativo, tenemos

$$x_4 \geq 300,$$

y el valor mínimo de $x_4 = 300$ se alcanza cuando x_6 es nulo. Así que, para atender la demanda de vehículos entrantes y salientes minimizando el tráfico en la calle Schiaffino entre Varela y Gambetta, debemos cerrar el tráfico en Ghiggia entre Gambetta y Varela. De esta manera forzamos la igualdad $x_6 = 0$. Por último, con $x_6 = 0$ se tiene

$$\begin{aligned} x_1 &= 100, \\ x_2 &= -x_7 + 600, \\ x_3 &= -x_7 + 900, \\ x_4 &= 300, \\ x_5 &= x_7 - 200. \end{aligned} \tag{1.21}$$

De la segunda ecuación se deduce que $x_7 \leq 600$. La última ecuación implica la desigualdad $x_7 \geq 200$. Entonces, cuando tomemos las medidas para minimizar el flujo en la cuadra problemática encontraremos que los valores de la circulaciones son los que nos da la fórmula (1.21), donde x_7 tomará algún valor en el intervalo $[200, 600]$. ♣

En el ejemplo anterior encontramos una situación como la que anunciábamos en la página 9, en la observación 1.1.1: las ecuaciones de conservación en los nodos que se obtienen al aplicar la Ley de Kirchoff son insuficientes para determinar completamente los flujos sobre la red. En nuestro próximo ejercicio podremos apreciar el mismo fenómeno.

Ejercicio 1.18. Hallar todas las soluciones del sistema (1.2) de la página 9.

El próximo ejemplo es de una naturaleza completamente diferente. Muestra un sistema de ecuaciones lineales que aparece al estudiar algunos cubrimientos de una esfera por polígonos, ¡con aplicaciones a nuestro deporte más popular!

Ejemplo 1.2.28. Recubriendo una esfera
En este ejemplo estudiaremos distintas maneras de recubrir una esfera con polígonos. Se trata de un caso particular de un problema aún más general: cubrir una superficie cualquiera con polígonos de algún tipo.

Por ejemplo, el plano puede ser cubierto por hexágonos, con una distribución que recuerda a un panal de abejas. Cada arista es común a dos hexágonos adyacentes, y en cada vértice confluyen tres hexágonos. Si observamos una pelota de fútbol encontraremos un diseño parecido. Sobre la superficie se dibujan algunos pentágonos y hexágonos, distribuidos de forma tal que cada arista es común a dos polígonos vecinos, y tres de ellos se encuentran en cada vértice del cubrimiento.

¿Por qué los fabricantes de pelotas se toman la molestia de cortar pentágonos y hexágonos, en vez de dividir la superficie de la esfera usando solamente hexágonos por medio de una construcción parecida a la división del plano que mencionábamos antes? La respuesta es simple: **tal división es imposible**.

La imposibilidad de esta construcción puede mostrarse a partir de una propiedad topológica de la esfera: su *característica de Euler*. Si cubrimos la esfera por polígonos y contamos los números v de vértices, c de caras y a de aristas que el cubrimiento genera, entonces v, c y a deben satisfacer la relación

$$c + v - a = 2. \tag{1.22}$$

No demostraremos (1.22), que es un resultado topológico profundo pero ajeno a los temas que nos interesa desarrollar en este texto. Nos limitaremos a observar que la relación se satisface en las particiones de la esfera que inducen los poliedros regulares. Cada poliedro regular, por ejemplo, un cubo, permite generar una división de la esfera proyectando sus caras, vértices y aristas sobre una esfera inscripta o circunscripta al poliedros. En el caso del cubo obtenemos un “embaldosamiento” de la esfera con 6 cuadriláteros, 8 vértices y 12 aristas. Examinemos qué ocurre con los cinco poliedros regulares:

	c	v	a	$c+v-a$
tetraedro	4	4	6	2
cubo	6	8	12	2
octaedro	8	6	12	2
dodecaedro	12	20	30	2
isocaedro	20	12	30	2

Apliquemos la relación (1.22) a nuestro hipotético cubrimiento por hexágonos. Supongamos que usamos una cantidad h de hexágonos, que es igual entonces al número de caras en el cubrimiento. Cada hexágono tiene 6 aristas. Tenemos entonces, en principio, $6h$ aristas. Pero al tener en cuenta que cada arista es compartida por dos hexágonos encontramos que

$$a = 6h/2 = 3h.$$

El cálculo para el número de vértices es similar: hay 6 vértices en cada hexágono, y cada vértices es común a tres de ellos. Por lo tanto $v = 6h/3 = 2h$. Entonces

$$c = h, \quad a = 3h, \quad v = 2h,$$

lo que implicaría

$$c + v - a = h + 2h - 3h = 0,$$

que está en contradicción con (1.22). En realidad, es imposible cubrir una esfera con hexágonos, incluso si se permite que en algunos vértices confluyan más de tres hexágonos.

Ejercicio 1.19. Refinar el argumento anterior, para mostrar que es imposible cubrir la esfera empleando sólo hexágonos.

Sin embargo hay cubrimientos de la esfera que combinan pentágonos y hexágonos. Analicemos qué restricciones impone la característica de Euler a un cubrimiento con h hexágonos y p pentágonos, distribuidos de forma tal que en cada vértice confluyen tres polígonos. Cada pentágono aporta 5 vértices, cada hexágono 6, pero alrededor de cada vértices hay tres polígonos. Tenemos entonces

$$v = (5p + 6h)/3,$$

lo que establece la relación

$$3v - 5p - 6h = 0$$

entre los números de vértices, pentágonos y hexágonos. Para las aristas podemos hacer un razonamiento similar, que conduce a

$$2a - 5p - 6h = 0.$$

Por supuesto, el número de caras c es la suma de los números p y h de pentágonos y hexágonos respectivamente. Por lo tanto

$$c - p - h = 0$$

Una cuarta relación está dada por la ecuación (1.22) de la característica de Euler. Escribamos todas las ecuaciones juntas, bajo la forma del sistema

$$\left\{\begin{aligned} 3v - 5p - 6h &= 0, \\ 2a - 5p - 6h &= 0, \\ c - p - h &= 0, \\ c + v - a &= 2. \end{aligned}\right.$$

Por supuesto, es más agradable representar matricialmente el sistema en las incógnitas (c, v, a, p, h), para luego proceder a tratarlo con el método de escalerización. Tenemos

$$\left(\begin{array}{rrrrr|r} 0 & 3 & 0 & -5 & -6 & 0 \\ 0 & 0 & 2 & -5 & -6 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 2 \end{array}\right).$$

En el primer paso ponemos la última ecuación en el primer lugar, la tercera en el segundo, y eliminamos el 1 que queda en la primera columna y la segunda fila. El resultado es

$$\left(\begin{array}{rrrrr|r} 1 & 1 & -1 & 0 & 0 & 2 \\ 0 & -1 & 1 & -1 & -1 & -2 \\ 0 & 3 & 0 & -5 & -6 & 0 \\ 0 & 0 & 2 & -5 & -6 & 0 \end{array}\right).$$

Multiplicamos ahora por 3 a la segunda fila, y la sumamos a la tercera. Luego de cancelar un factor -1 en la tercera fila, e intercambiarla con la cuarta obtenemos

$$\left(\begin{array}{rrrrr|r} 1 & 1 & -1 & 0 & 0 & 2 \\ 0 & -1 & 1 & -1 & -1 & -2 \\ 0 & 0 & 2 & -5 & -6 & 0 \\ 0 & 0 & -3 & 8 & 9 & 6 \end{array}\right).$$

Sumamos a la cuarta fila $3/2$ de la tercera, y multiplicamos el resultado por 2, para conseguir la forma escalerizada

$$\left(\begin{array}{rrrrr|r} 1 & 1 & -1 & 0 & 0 & 2 \\ 0 & -1 & 1 & -1 & -1 & -2 \\ 0 & 0 & 2 & -5 & -6 & 0 \\ 0 & 0 & 0 & 1 & 0 & 12 \end{array}\right).$$

La variable h, que representa el número de hexágonos, quedó libre, ¡pero la variable p queda completamente determinada por la cuarte ecuación $p = 12$! El número de hexágonos puede tomar diversos valores, pero un cubrimiento como el que estamos buscando necesariamente debe tener 12 pentágonos.

Dos ejemplos notables son conocidos para nosotros. El caso en que hay 12 pentágonos pero ningún hexágono corresponde al cubrimiento de una esfera que se obtiene a partir de un dodecaedro. El ejemplo que tiene 12 pentágonos y 20 hexágonos es el cubrimiento que se emplea en la fabricación de las pelotas de fútbol. En esta solución hay 32 caras, 60 vértices y 90 aristas.

Para cerrar este ejemplo comentemos que el sistema de ecuaciones lineales que relaciona los números de caras, vértices, aristas, pentágonos y hexágonos es una **condición necesaria** que debe satisfacer cualquier solución al problema de cubrir la esfera. Pero no todas las soluciones del sistema permiten construir cubrimientos. Por ejemplo, para aquellas en que alguna de las variables tome un valor que no sea un número natural, ni siquiera tiene sentido formularse el problema de buscar un cubrimiento con ese número de caras, vértices o aristas. ♣

Ejercicio 1.20. Expresar en forma matricial y resolver en $\mathbb{R}$ los siguientes sistemas de ecuaciones lineales

1. $\left\{\begin{array}{rcrcl} x & + & 2y & = & 8, \\ 3x & - & y & = & 3. \end{array}\right.$

2. $\left\{\begin{array}{rcrcrcr} -x & + & y & - & z & = & -1, \\ 4x & + & 2y & - & z & = & 5, \\ x & + & & & z & = & 2. \end{array}\right.$

3. $\left\{\begin{array}{rcrcrcl} x & & & + & z & = & 1, \\ & & y & + & z & = & 0, \\ x & + & y & + & z & = & 0. \end{array}\right.$

Verificar que las soluciones halladas son las correctas.

Ejercicio 1.21. Expresar en forma matricial y resolver el sistema de ecuaciones lineales

$$\left\{\begin{array}{rcrcrcl} x & & & + & z & = & 1, \\ & & y & + & z & = & 0, \\ x & + & y & + & z & = & 0, \end{array}\right.$$

en $\mathbb{Z}_2$. Verificar las soluciones halladas.

Ejercicio 1.22. Resolver los siguientes sistemas en $\mathbb{Z}_3$.

1. $\left\{\begin{array}{rcrcrcl} x & & & + & z & = & 1, \\ & & y & + & z & = & 0, \\ x & + & y & + & z & = & 0. \end{array}\right.$

2. $\left\{ \begin{array}{rcrcl} x & + & 2y & = & 1, \\ 2x & + & y & = & 2. \end{array} \right.$

Verificar las soluciones halladas.

Ejercicio 1.23. Resolver en $\mathbb{C}$ los siguientes sistemas de ecuaciones, y verificar las soluciones halladas.

1. $\left\{ \begin{array}{rcrcr} (1+i)x & + & (2-i)y & = & 4i+1, \\ (1+3i)x & + & 4y & = & -2+8i, \end{array} \right.$

2. $\left\{ \begin{array}{rcrcrcr} ix & + & (3+i)y & + & iz & = & -3+10i, \\ x & + & 2y & + & (1-i)z & = & 4+2i, \\ (1-i)x & + & (1+2i)y & + & (2-i)z & = & 4-i, \end{array} \right.$

El objetivo del próximo ejercicio es mostrar cómo pueden escalerizarse simultáneamente varios sistemas que comparten la misma matriz pero tienen distintos términos independientes. En el ejercicio 1.30, página 63, mostraremos como continuar el proceso de escalerización para calcular eficientemente las soluciones. Este problema enfatiza el hecho de que el proceso de escalerización sólo depende de la matriz del sistema, no de su matriz ampliada. El procedimiento que se presenta permitirá calcular, en la sección 2.3, la inversa de esta matriz. El algoritmo no tiene en realidad mayor importancia práctica, porque otros procedimientos lo sustituyen con ventaja. Pero tiene interés enfatizar que el cálculo de la inversa es equivalente a la resolución de sistemas lineales. Sobre estos asuntos, ver el comienzo de la sección 2.3. En particular, la observación 2.3.2, en la página 198.

Ejercicio 1.24. Resolver los siguientes sistemas de ecuaciones:

1. $\left\{ \begin{array}{rcrcrl} 2x & + & 2y & - & z & = 1 \\ & & x & + & 2z & = 0 \\ & - & x & - & y & = 0 \end{array} \right.$

2. $\left\{ \begin{array}{rcrcrl} 2x & + & 2y & - & z & = 0 \\ & & x & + & 2z & = 1 \\ & - & x & - & y & = 0 \end{array} \right.$

3. $\left\{ \begin{array}{rcrcrl} 2x & + & 2y & - & z & = 0 \\ & & x & + & 2z & = 0 \\ & - & x & - & y & = 1 \end{array} \right.$

Sugerencia: escalerizar los tres sistemas al mismo tiempo, observar que la escalerización sólo depende de la matriz del sistema, no del término independiente.

En el siguiente ejercicio nos apartaremos un poco de las rutinas de la eliminación gaussiana para incursionar en un tema muy importante para el cálculo: el problema de entender cómo dependen las soluciones de los sistemas de ecuaciones lineales de sus coeficientes. En particular, de determinar cuál es el efecto de una perturbación en algunos de sus coeficientes. Esta cuestión es importante para la ingeniería porque todos los datos de problemas reales están afectados por algún tipo de error. La representación de datos numéricos en una computadora también introduce errores, porque la enorme mayoría de los números reales no se representa exactamente.

Ejercicio 1.25. Número de condición (I)

1. Resolver el sistema de dos ecuaciones con dos incógnitas $AX = B$, donde

$$A = \begin{pmatrix} 1 & 0{,}999 \\ 0{,}999 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1{,}999 \\ 1{,}999 \end{pmatrix}.$$

2. Considerar el sistema $AX = B_1$, con la misma matriz A, y $B_1 = (1{,}989, 2{,}009)^7$. Dar una estimación del valor de X sin resolver el sistema.
3. Resolver $AX = B_1$. ¿Obtuviste un resultado cercano al que esperabas?
4. Resolver $AX = B_2$, con $B_2 = (2{,}009, 2{,}009)$.
5. Discutir los resultados obtenidos.

El título del ejercicio 1.25 se verá justificado más adelante. De momento sólo digamos que el *número de condición* es una medida de que tan sensibles pueden ser las soluciones de un sistema de ecuaciones lineales a las perturbaciones de sus coeficientes. Este ejercicio es el primero de una serie en la que iremos desarrollando estos conceptos.

En el próximo ejercicio proponemos al lector discutir según un parámetro λ el comportamiento de algunos sistemas de ecuaciones. Nuestra selección no es arbitraria, se trata en realidad de resolver *problemas de valores y vectores propios* que son de fundamental importancia dentro del álgebra lineal y sus aplicaciones. El lector encontrará este tipo de problemas en distintas partes de este mismo texto.

De hecho, la primera parte del ejercicio 1.26 en alguna forma ya ha aparecido en el ejercicio 1.25. Tal vez puedas encontrar la conexión entre ambos problemas. De todos modos, es algo que trabajaremos explícitamente más adelante, en el ejercicio 2.7 de la página 58. También, el ejercicio 1.10 de la página 21 contiene un caso particular del problema que se plantea en la parte 6 del próximo ejercicio, ¿podés reconocerlo?

Ejercicio 1.26. Resolver los siguientes sistemas de ecuaciones, discutiendo según λ la existencia y multiplicidad de soluciones.

[7] Tal como es habitual, debemos considerar que B_1 es una columna. Lo hemos escrito como fila para ahorrar espacio en la hoja.

1. $\left\{\begin{array}{rcrcl} x & + & 0{,}999y & = & \lambda x, \\ 0{,}999x & + & y & = & \lambda y. \end{array}\right.$

2. $\left\{\begin{array}{rcrcl} 2x & + & y & = & \lambda x, \\ & & y & = & \lambda y. \end{array}\right.$

3. $\left\{\begin{array}{rcrcl} x & + & y & = & \lambda x, \\ x & & & = & \lambda y. \end{array}\right.$

4. $\left\{\begin{array}{rcrcl} 3x & - & 4y & = & \lambda x, \\ 2x & - & 3y & = & \lambda y. \end{array}\right.$

5. $\left\{\begin{array}{rrcrcrcl} & & & 3y & + & z & = & \lambda x, \\ & 2x & - & y & - & z & = & \lambda y, \\ - & 2x & - & y & - & z & = & \lambda z. \end{array}\right.$

6. $\left\{\begin{array}{rcrcrcl} 0{,}6x & + & 0{,}4y & + & 0{,}1z & = & \lambda x, \\ 0{,}1x & - & 0{,}4y & + & 0{,}1z & = & \lambda y, \\ 0{,}3x & + & 0{,}2y & + & 0{,}8z & = & \lambda z. \end{array}\right.$

7. $\left\{\begin{array}{rrcrcrcl} - & 2x & + & y & - & z & = & \lambda x, \\ - & x & - & y & & & = & \lambda y, \\ & & & y & - & 3z & = & \lambda z. \end{array}\right.$

En nuestro próximo ejercicio vamos a aplicar lo que hemos aprendido sobre la notación matricial y el tratamiento sistemático de los sistemas de ecuaciones lineales para estudiar las ecuaciones a las que nos condujo el problema de circuitos que planteamos en el ejemplo 1.1.4

Ejercicio 1.27. Flujos en una red de distribución
Al analizar en un caso particular la distribución de las corrientes en la red del ejemplo 1.1.4 encontramos que los voltajes v_i, $i = 1, \ldots, 5$, en los nodos de la red debían satisfacer el sistema de ecuaciones

$$\begin{array}{rcrcrcrcrcl} 3v_1 & - & v_2 & - & v_3 & - & v_4 & & & = & d_1, \\ -v_1 & + & 3v_2 & - & & & v_4 & - & v_5 & = & d_2, \\ -v_1 & + & & & 2v_3 & - & v_4 & & & = & d_3, \\ -v_1 & - & v_2 & - & v_3 & + & 4v_4 & - & v_5 & = & d_4, \\ & - & v_2 & - & & & v_4 & + & 2v_5 & = & d_5, \end{array} \tag{1.23}$$

donde los números d_i, $i = 1, \ldots, 5$ representan las demandas en cada nodo. La matriz ampliada del sistema es

$$\left(\begin{array}{rrrrr|l} 3 & -1 & -1 & -1 & 0 & d_1 \\ -1 & 3 & 0 & 1 & -1 & d_2 \\ -1 & 0 & 2 & -1 & 0 & d_3 \\ -1 & -1 & 1 & 4 & 1 & d_4 \\ 0 & -1 & 0 & 1 & 2 & d_5 \end{array}\right).$$

Utilizar la expresión matricial para resolver el sistema. Interpretar la condición de compatibilidad del sistema y el conjunto de sus soluciones.

A continuación introducimos un ejercicio acerca de una aplicación no trivial del álgebra lineal: la protección de la información contra errores en su almacenamiento y transmisión.

Ejercicio 1.28. CÓDIGOS DE HAMMING (I)
Un procedimiento para proteger información al transmitirla es agregar cierta *redundancia*. A continuación mostramos una manera de hacer esto. Supongamos que tenemos que transmitir una secuencia $X = (x_1, x_2, x_3, x_4)$ formada por ceros y unos. La codificamos entonces como una secuencia Y de siete números, definidos por las fórmulas

$$\begin{aligned} y_1 &= x_1 + x_2 + x_4, \\ y_2 &= x_1 + x_3 + x_4, \\ y_3 &= x_1, \\ y_4 &= x_2 + x_3 + x_4, \\ y_5 &= x_2, \\ y_6 &= x_3, \\ y_7 &= x_4. \end{aligned}$$

La aritmética que se emplea es la de $\mathbb{Z}_2$.

1. Hallar la codificación Y que corresponde a $(0, 1, 0, 1)$. Hallar todas las 7-uplas de ceros y unos que pueden obtenerse por este procedimiento.

2. Si se sabe que los cuatro primeros dígitos (y_1, y_2, y_3, y_4) de una palabra codificada Y son $(1, 0, 1, 1)$, hallar la secuencia X que se codificó.

1.2.4. Matriz escalerizada reducida

Existe una alternativa al procedimiento de calcular la última variable y luego comenzar a sustituir hacia arriba: consiste en utilizar las últimas ecuaciones para ir eliminando variables de las anteriores. El procedimiento es el mismo que el de escalerización, pero se van eliminando variables "hacia arriba". Veamos un ejemplo.

Ejemplo 1.2.29. Volvamos al problema del ejemplo 1.2.19. Luego de escalerizar el sistema habíamos obtenido la matriz

$$\left(\begin{array}{rrr|r} 1 & 1 & 2 & 8 \\ 0 & 1 & -1 & -8 \\ 0 & 0 & -4 & -20 \end{array}\right). \tag{1.24}$$

Busquemos ahora eliminar las entradas que corresponden a la variable z en la primera y segunda ecuación. Dividiremos la última ecuación entre -4. Luego la sumaremos a la segunda. También haremos la operación de multiplicar la última fila por 2 y restarla de la primera. Obtenemos

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & -2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & \mathbf{1} & 5 \end{array}\right).$$

Por último, hacemos aparecer un 0 en el segundo lugar de la primera fila, restando la segunda ecuación a la primera. El resultado final es

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & \mathbf{1} & 0 & -3 \\ 0 & 0 & 1 & 5 \end{array}\right).$$

El sistema asociado a esta matriz es, naturalmente,

$$\left\{\begin{array}{l} x = 1, \\ y = -3, \\ z = 5, \end{array}\right.$$

cuya resolución es inmediata. ♣

Consideraremos ahora un caso un poco más complicado.

Ejemplo 1.2.30. Retomamos el sistema del ejemplo 1.2.22 en el punto en que habíamos escalerizado la matriz del sistema. Teníamos que

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

era la matriz asociada al sistema escalerizado. Usaremos ahora los pivotes para hacer aparecer ceros por encima de ellos. Como paso previo dividimos las filas tercera y cuarta por sus pivotes,

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Luego comenzamos por hacer aparecer ceros sobre el pivote de la última columna, y obtenemos

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \longleftarrow F_1 - F_4, \\ \\ \longleftarrow F_3 - F_4. \\ \\ \\ \end{array}$$

A partir de esta nueva matriz operamos con el pivote de la cuarta columna:

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & \mathbf{1} & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 - F_3, \\ \\ \\ \\ \end{array}$$

Ahora lo hacemos con el de la segunda columna, y dividimos la segunda fila entre 2 para que su pivote quede igual a 1. El resultado final es

$$\left(\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & \frac{3}{2} & 0 & -1 \\ 0 & 0 & \mathbf{1} & 0 & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \longleftarrow F_1 - \frac{1}{2}F_2, \\ \longleftarrow \frac{1}{2}F_2, \\ \\ \\ \\ \end{array} \tag{1.25}$$

donde hemos enfatizado además la "forma escalerizada" que tiene la matriz. El sistema de ecuaciones asociado con esta matriz es

$$\left\{ \begin{array}{rcl} x_1 + 2x_2 + \frac{3}{2}x_5 + & = & -1, \\ x_3 - \frac{1}{2}x_5 & = & 1, \\ x_4 + 2x_5 & = & -1, \\ x_6 & = & 1, \end{array} \right.$$

y ahora es completamente directo despejar x_1, x_3 y x_4 en términos de x_2 y x_5 para reencontrar que

$$\{(-1 - 2x_2 - 3x_5/2, x_2, 1 + x_5/2, -1 - 2x_5, x_5, 1)\ ;\ x_2,\ x_5 \in \mathbb{R}\}$$

es el conjunto solución del sistema. ♣

Vemos en nuestros últimos dos ejemplos que continuar con la escalerización "hacia arriba" a partir de la forma escalerizada de la matriz, nos permite llegar a una nueva matriz, escalerizada, equivalente a la forma escalerizada que ya habíamos obtenido y a la original, tal que

1. todos los pivotes son iguales a 1;
2. las columnas correspondientes a los pivotes tienen todas las entradas nulas, salvo el pivote.

Llamaremos *matriz escalerizada reducida* a una matriz escalerizada que cumple además estas dos condiciones. Notemos que siempre que tengamos una matriz escalerizada reducida podremos, si el sistema es compatible, despejar directamente las variables de las columnas de los pivotes, y expresarlas en función de las entradas en la última columna de la matriz ampliada y de las variables libres.

Ejercicio 1.29. Resolver el sistema (1.9) del ejercicio 1.27 llevando la matriz escalerizada del sistema a su forma reducida.

El objetivo del próximo ejercicio es mostrar cómo cuando se prosigue hasta alcanzar una forma escalerizada reducida la escalerización simultánea de varios sistemas se transforma en un eficiente método para calcular todas las soluciones al mismo tiempo.

Ejercicio 1.30. Resolver simultáneamente los tres sistemas de ecuaciones propuestos en el ejercicio 1.24 llevando la matriz de los sistemas hasta su forma escalerizada reducida.

Observación 1.2.31. La forma escalerizada de una matriz no es única. Pero sí lo es su forma escalerizada reducida. Una demostración de este resultado de unicidad puede hallarse, por ejemplo, en el primer apéndice de [L], en la página A1[8]. ♠

1.2.5. Cualquier sistema puede ser escalerizado

Hasta este momento hemos usado el algoritmo de escalerización, y resuelto unos cuantos sistemas de ecuaciones con él. El lector podría preguntarse si cualquier matriz puede ser llevada a una forma escalerizada. La respuesta es que sí. Y la demostración no pasa de ser un análisis cuidadoso del método. Todo esto está contenido en nuestra siguiente proposición.

Proposición 1.4. *Toda matriz puede ser transformada en una matriz escalerizada mediante una cantidad finita de transformaciones elementales*

[8] El texto [L] es una excelente referencia, con un desarrollo, y muchas aplicaciones. Recomendamos al lector su consulta, también la del sitio web `http://www.laylinalgebra.com/`. Existe una versión en español del texto.

Prueba. Razonaremos por inducción sobre el número de filas de la matriz A. Como primer paso, digamos que es obvio que una matriz que tiene una única fila se puede escalerizar por medio de una cantidad finita de operaciones elementales. En realidad no hay que hacer ninguna: una matriz con una única fila ya está escalerizada. No hay nada que hacer con ella.

Supongamos ahora que la proposición es cierta para matrices con $m-1$ filas, y veremos como implica que se puede escalerizar cualquier matriz con m filas. Sea $A = ((a_{ij}))$ una matriz $m \times n$, con $m \geq 2$, sobre un cuerpo $\mathbb{K}$ cualquiera. Indiquemos por

$$A_i = (a_{i1}, a_{i2}, \ldots, a_{in})$$

su fila i-ésima.

Distingamos tres casos:

1. si $a_{11} \neq 0$ entonces realizamos el primer paso de la escalerización: sustituimos cada fila A_i, para $i = 2, \ldots, m$, por
$$A_i' = A_i - \frac{a_{i1}}{a_{11}} A_1.$$
La elección del multiplicador para A_1 asegura que la primera entrada de cada una de estas filas A_i' es nula;

2. si $a_{11} = 0$, pero la primera columna tiene alguna entrada a_{i1} no nula, entonces escogemos una fila A_i tal que $a_{i1} \neq 0$, e intercambiamos[9] la primera fila con A_i. La aplicación de esta operación elemental de intercambiar dos filas nos coloca en la situación del caso 1. Luego procedemos como en ese caso;

3. si toda la primera columna es nula entonces no hacemos nada.

El resultado del paso anterior es una nueva matriz que tiene el siguiente aspecto

$$\left(\begin{array}{c|ccc|} a_{11} & a_{12} & \ldots & a_{1n} \\ \cline{2-4} 0 & a'_{22} & \ldots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{m2} & \ldots & a'_{mn} \\ \cline{2-4} \end{array}\right).$$

[9]Por razones numéricas y para no amplificar innecesariamente errores de redondeo, al resolver sistemas con ayuda de una computadora resulta muchas veces importante, aún cuando $a_{11} \neq 0$, cambiar la primera fila por la fila que tenga la primera entrada mayor. Este procedimiento se conoce como *pivoteo*. Lo discutimos con cierto detalle en la sección dedicada a la descomposición LU. Acerca del pivoteo por razones de precisión ver la página 225.

Para escalerizar esta matriz basta encontrar una forma escalerizada de la submatriz

$$(a'_{ij})_{j=2,\dots,n}^{i=2,\dots,m},$$

que tiene $m-1$ filas. Por lo tanto esta forma escalerizada puede obtenerse por medio de un número finito de transformaciones elementales.

La proposición que acabamos de demostrar tiene un útil corolario para la resolución de sistemas de ecuaciones lineales.

Corolario 1.5. *Todo sistema lineal es equivalente a uno escalerizado.*

Ejercicio 1.31. Mostrar que cualquier matriz puede ser llevada a una forma escalerizada reducida por medio de una cantidad finita de operaciones elementales.

Observación 1.2.32. Notemos que para llegar a la forma escalerizada (o a la escalerizada reducida) no es necesario usar la transformación elemental que consiste en multiplicar una fila de la matriz o una ecuación de un sistema por un número distinto de cero. Usaremos esta observación en la sección 2.5, dedicada al cálculo de determinantes.

Ejercicio 1.32. Consideremos un sistema lineal de ecuaciones con coeficientes y términos independientes racionales. Mostrar que para todas las soluciones reales del sistema todas las incógnitas toman valores racionales.

La parte 2 del próximo ejercicio contiene un corolario fácil de derivar a partir de la teoría que hemos desarrollado. A pesar de su simplicidad, es un resultado crucial, que emplearemos en la sección 1.6 para establecer el concepto de dimensión.

Ejercicio 1.33. Consideremos el sistema lineal homogéneo con m ecuaciones y n incógnitas.

1. Mostrar que si $m < n$ entonces el sistema tiene soluciones no triviales.
2. Mostrar que si el sistema es determinado y admite como única solución la trivial en que todas las incógnitas toman el valor 0, entonces $m \geq n$.
3. Dar ejemplos de sistemas homogéneos indeterminados en los que m sea mayor que n.

Nos planteamos ahora una generalización a sistemas lineales cualesquiera, no necesariamente homogéneos.

Ejercicio 1.34. Consideremos sistemas lineales cualesquiera, con m ecuaciones y n incógnitas.

1. Para el caso $m > n$, en que el sistema tiene más ecuaciones que incógnitas,

a) dar ejemplos de sistemas incompatibles;

b) dar ejemplos de sistemas compatibles y determinados;

c) dar ejemplos de sistemas compatibles e indeterminados.

2. Repetir la parte anterior para el caso $m = n$, en que el sistema tiene el mismo número de ecuaciones que de incógnitas.

3. Para el caso $m < n$, en que el sistema tiene menos ecuaciones que incógnitas,

   a) dar ejemplos de sistemas incompatibles;

   b) dar ejemplos de sistemas compatibles y determinados;

   c) mostrar que el sistema nunca puede ser compatible y determinado.

1.2.6. Las propiedades de las operaciones en el espacio $\mathbb{K}^n$ de las n-uplas

En la observación 1.2.21 habíamos enfatizado el hecho de que el proceso de eliminación podía ser descrito en términos de operaciones realizadas sobre las filas de la matriz asociada a un sistema de ecuaciones lineales. Cada una de las filas de una matriz $m \times n$ es un lista ordenada de n números, o n-upla. Así, al aplicar el método de escalerizar multiplicamos n-uplas por números y las sumamos entre sí. También, en la observación 1.2.3, página 25, habíamos destacado el hecho de que cada solución de un sistema de ecuaciones lineales puede pensarse como un único objeto, una lista de n números, en vez de ser considerada como el resultado de hallar el valor de n incógnitas diferentes.

Comenzamos a sistematizar ahora el trabajo con las n-uplas construidas a partir de un cuerpo $\mathbb{K}$. Llamaremos $\mathbb{K}^n$ al conjunto formado por todas las n-uplas con entradas en el cuerpo $\mathbb{K}$, o listas ordenadas de n-números. Es decir

$$\mathbb{K}^n = \{(x_1, x_2, \ldots, x_n); x_1, x_2, \ldots, x_n \in \mathbb{K}\}.$$

La estructura algebraica de $\mathbb{K}^n$

Sobre este conjunto definimos dos operaciones.

1. Suma de dos listas. Si $X = (x_1, \ldots, x_n)$ e $Y = (y_1, \ldots, y_n)$ entonces definimos la suma $X + Y$ por la fórmula

$$X + Y = (x_1 + y_1, \ldots, x_n + y_n).$$

   Vemos entonces la suma de dos listas se obtiene sumando en cada posición las entradas de cada lista.

2. PRODUCTO DE UNA LISTA POR UN NÚMERO. Para $X(x_1,\dots,x_n)$ y a un número cualquiera en el cuerpo $\mathbb{K}$ definimos

$$aX = (ax_1,\dots,ax_n).$$

   En otras palabras, el producto de una lista por un número se obtiene multiplicando cada entrada de la lista por el número.

Antes de discutir las propiedades más importantes de estas operaciones hagamos un pequeño ejemplo.

Ejemplo 1.2.33.

1. Comencemos con $n = 3$ y el cuerpo de los números reales. Tenemos, por ejemplo

$$\begin{gathered}(1,e,-1)+(0,1,2) = (1+0, 2+e, -1+2) = (1, 2+e, 1),\\ \sqrt{2}(1,0,\sqrt{2}) = (\sqrt{2}{,}1, \sqrt{2}{,}0, \sqrt{2}.\sqrt{2}) = (\sqrt{2}, 0, 2).\end{gathered}$$

2. Tomemos $X = (1,1,1,0,0,0)$, $Y = (1,0,1,0,1,0)$ en $\mathbb{Z}_2^6$. Entonces

$$X + Y = (0,1,0,0,1,0).$$

   Si queremos ejercitar el producto por un número no tenemos mucho para hacer. Hay sólo dos opciones: multiplicar una lista por 0, lo que da como resultado una lista con 6 ceros. O multiplicarla por 1, lo que produce exactamente la misma lista.

3. Todo funciona más o menos igual en el campo de los números complejos. Por ejemplo, en $\mathbb{C}^2$ tenemos $i(1+i,-i) = (-1+i,1)$. ♣

Destaquemos que ahora nuestras listas de números empiezan a adquirir estructura: hay operaciones algebraicas definidas sobre ellas, y con la ayuda de estas operaciones las hemos estado manipulando para resolver y comprender los problemas relativos a sistemas lineales. La introducción de esta estructura algebraica hace que $\mathbb{K}^n$ deje de ser apenas un conjunto, y por eso nos referiremos al *espacio* $\mathbb{K}^{\mathbf{n}}$, o al *espacio vectorial* $\mathbb{K}^{\mathbf{n}}$. También nos referiremos a los elementos de este espacio como a *vectores* en el espacio $\mathbb{K}^n$, aunque de tanto en tanto, y dependiendo del contexto, volveremos a emplear la expresión "listas de longitud n" para designarlos.

Las operaciones tienen una lista de propiedades que emplearemos repetidamente en el cálculo y que nos interesa destacar porque las recogeremos más tarde en la definición de *espacio vectorial.* Para compactar la notación mantenemos el uso de designar a cada n-upla con una letra mayúscula. Por ejemplo, $X = (x_1, \dots, x_n)$. Comenzamos con las propiedades de la suma. En lo que sigue X, Y y Z simbolizan vectores cualesquiera en $\mathbb{K}^n$.

[S1] CONMUTATIVA: $X + Y = Y + X$.

[S2] ASOCIATIVA: $(X + Y) + Z = X + (Y + Z)$.

[S3] NEUTRO DE LA SUMA: Existe un vector $O \in \mathbb{K}^n$ tal que

$$X + O = O + X = X$$

se satisface para cualquier $X \in \mathbb{K}^n$.

[S4] EXISTENCIA DE OPUESTO: Para cada $X \in \mathbb{R}^n$ existe un opuesto $(-X) \in \mathbb{R}^n$ tal que

$$X + (-X) = O.$$

Estas propiedades son absolutamente directas de verificar. Sólo haremos un par de comentarios sobre la existencia de neutro y opuesto para la suma. Es obvio en este contexto que el neutro para la suma es el vector

$$O = (0, 0, \dots, 0)$$

que esta formado por n ceros, y que el opuesto de

$$X = (x_1, \dots, x_n)$$

no es otro que

$$(-x_1, \dots, -x_n).$$

La formulación que hemos escogido para enunciar las propiedades parece pues un tanto rebuscada, ya que podríamos haber dicho directamente quienes eran el opuesto y el neutro. Confesemos al lector que nuestra elección obedece a la idea de extender toda esta teoría al contexto abstracto de los espacios vectoriales en que los vectores con los que operemos pueden ser muy generales. Por ejemplo, elementos de $\mathbb{K}^n$, funciones, sucesiones, matrices, clases de equivalencia construidas por distintos procedimientos, etcétera. Esta amplitud da una gran aplicabilidad a la teoría de los espacios vectoriales.

En cada ejemplo de espacio vectorial habrá que especificar cuál es el vector O y cómo se construye el opuesto de un vector cualquiera. Pero lo único que será común a todos los casos particulares son las propiedades **[S3]** y **[S4]** que deben ser satisfechas por el neutro y el opuesto.

También son de fácil verificación las propiedades del producto entre los números del cuerpo $\mathbb{K}$, a los que llamaremos *escalares*, y los vectores de $\mathbb{K}^n$. En lo que siguen α y β indicaran escalares cualesquiera. Igual que en el caso de la suma, representaremos con X e Y a vectores cualesquiera en $\mathbb{K}^n$. En la propiedad **[P2]** el 1 simboliza a la unidad multiplicativa del cuerpo $\mathbb{K}$.

[P1] ASOCIATIVA: $(\alpha\beta)X = \alpha(\beta X)$.

[P2] NEUTRO DEL PRODUCTO $1X = X$.

[P3] DISTRIBUTIVA RESPECTO A LA SUMA DE ESCALARES: $(\alpha+\beta)X = \alpha X + \beta X$.

[P4] DISTRIBUTIVA RESPECTO A LA SUMA DE VECTORES: $\alpha(X+Y) = \alpha X + \alpha Y$.

La teoría de los espacios vectoriales está basada en estas ocho propiedades que nos servirán como *axiomas* o *postulados* para la teoría. Los espacios $\mathbb{K}^n$ son, en realidad, un caso particular de esta estructura general con la que trabajaremos más adelante.

Ejercicio 1.35. Completar la verificación de que las operaciones de suma y producto por un escalar que hemos definido en $\mathbb{K}^n$ tienen todas las propiedades que acabamos de enunciar.

Observación 1.2.34. Las operaciones que hemos definido en $\mathbb{K}^n$ se extienden, en forma obvia, a filas y columnas de las matrices. Cada fila o columna no es otra cosa que una lista ordenada de números en $\mathbb{K}$. También podemos, y será útil en alguna oportunidad, considerar un elemento de $\mathbb{K}^n$ como una matriz de una sola columna y n filas (esto es esencialmente lo mismo que pensar la lista de n números como una columna) o como una matriz de una fila y n columnas (o sea, una fila).

Otro punto de vista posible es considerar cada elemento

$$X = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$$

como una función definida X definida sobre el subconjunto

$$\{1, 2, \ldots, n\}$$

formado por los primeros n números naturales, y que toma valores en el cuerpo $\mathbb{K}$. Esta función asocia a 1 el número $X(1)$ que indicamos con la notación x_1, a 2 el número $X(2) = x_2$, y así sucesivamente hasta n, cuyo correspondiente es $X(n) = x_n$.

Si quisiéramos enfatizar este último punto de vista sería más adecuado escribir el vector X con la notación

$$X = (X(1), X(2), \ldots, X(n)),$$

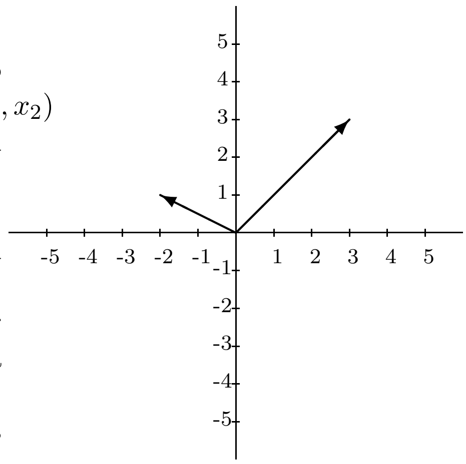
que es corriente en los lenguajes de programación. Esta notación es más sistemática, porque sólo requiere saber el nombre del vector y el lugar de la entrada para interpretar las cosas correctamente. La notación x_1 es más engorrosa de interpretar. Para que tenga sentido hay que saber que debemos identificar la letra mayúscula X con su minúscula y, una convención que un humano rápidamente comprende, pero que en una computadora requiere cierto procesamiento adicional, innecesario.

Este punto de vista es análogo al de considerar las matrices como funciones (ver la definición 1.1 y la observación 1.2.14 al respecto) y será desarrollado con algo más de detalle en el capítulo dedicado a la teoría abstracta de los espacios vectoriales.

Geometría en $\mathbb{R}^n$

Hemos dado el sugerente nombre de *vectores*, que tiene una indudable connotación geométrica, a los elementos de los espacios $\mathbb{K}^n$. Mostraremos ahora cómo se entiende esta geometría en el caso de $\mathbb{R}^2$.

Comencemos por colocar un par de ejes coordenados en el plano, tal como aparecen en la figura. En ese plano coordenado cada pareja de números (x_1, x_2) representa un vector, una instrucción para desplazarse que puede leerse de esta manera: salga desde el origen $(0, 0)$ moviéndose x_1 unidades en la dirección del eje horizontal, y x_2 en la del eje vertical. Si x_1 es positivo hay que viajar hacia la derecha, si es negativo hacia la izquierda. Respectivamente, para x_2 la convención usual es que los valores positivos indican que hay que ir hacia arriba, y los negativos hacia abajo.



Hemos representado los vectores $(3, 3)$ y $(-2, 1)$, siguiendo la convención estándar de dibujarlos como una flecha que comienza en el origen, y tiene su extremo en el punto que se alcanza con nuestro hipotético recorrido.

Los vectores en $\mathbb{R}^2$ pueden multiplicarse por escalares, también sumarse. ¿Cómo se ven estas operaciones en nuestra representación geométrica?

El producto λX de un vector X por un escalar λ alarga o acorta el vector por un factor igual al valor absoluto $|\lambda|$. Si $\lambda > 0$ el vector resultante tiene el mismo sentido que X. Cuando $\lambda < 0$ el sentido se invierte. En ambos casos se conserva la dirección, es decir, la recta que pasa por el origen y sobre la que se "apoya" el vector.

La suma $X+Y$ sigue la "regla del paralelogramo", produce un nuevo vector que se construye sobre la diagonal del paralelogramo que tiene lados X e Y.

Las dos operaciones de suma y producto se ilustran en la figura 1.4 con los vectores $(-2,1)$ y $(3,3)$. Hemos representado los productos

$$1/2 \times (3,3) = (3/2, 3/2), \quad (-1) \times (-2,1) = (2,-1),$$

y la suma

$$(3,3) + (-2,1) = (1,4).$$

El lector podrá construir más de estos ejemplos si así lo desea.

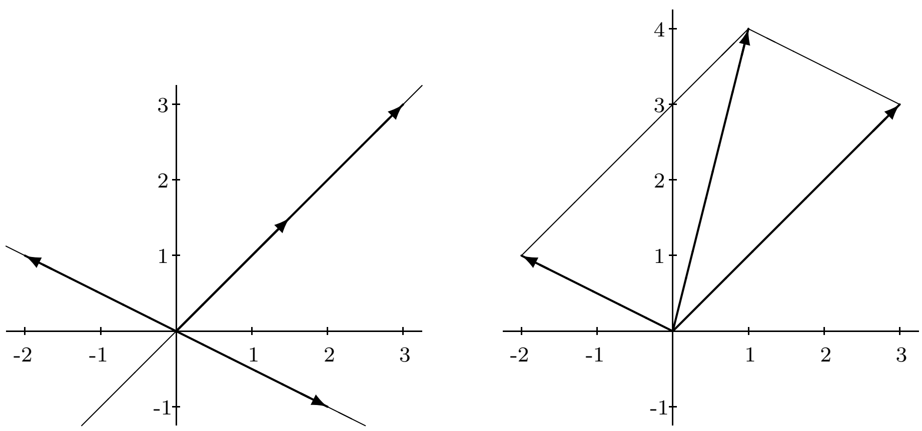


Figura 1.4: Dos productos y una suma en $\mathbb{R}^2$

La extensión de estas consideraciones a $\mathbb{R}^3$ es bastante directa, y se consigue pasando del plano al espacio. La extensión a $\mathbb{R}^n$, con $n > 3$, requiere la capacidad de imaginar las mismas nociones en espacios de dimensión mayor a 3. La geometría funciona esencialmente igual, aunque las representaciones ya no son tan directas, tan cercanas a nuestra experiencia sensorial. En el capítulo 3 extenderemos mucho estas consideraciones geométricas, e introduciremos en nuestra teoría las nociones geométricas básicas de longitud, distancia, y ángulo, que serán definidas con rigor y dejarán de depender de interpretaciones más o menos intuitivas. Daremos definiciones que admiten generalizaciones inmediatas a cualquier espacio $\mathbb{R}^n$.

No debemos preocuparnos demasiado por "no poder dibujar un espacio de dimensión mayor que 3". A decir verdad, tampoco podemos dibujar exactamente un espacio de dimensión 2, ni siquiera uno de dimensión 1. Ningún "triángulo" que dibujemos en una hoja de papel es un verdadero triángulo; ninguna "circunferencia"que consigamos con un compás una auténtica circunferencia. Sólo podemos hacer algunas representaciones que nos resultan útiles. **La clave no está tanto en la representación, sino en nuestra capacidad de interpretarla correctamente**.

Las imágenes geométricas que hemos discutido son útiles para todos los espacios $\mathbb{K}^n$, donde $\mathbb{K}$ es un cuerpo cualquiera. Aunque en el caso de los cuerpos finitos como $\mathbb{Z}_2$ otras representaciones pueden ser preferibles.

1.2.7. Para tener presente

- La teoría de los sistemas de ecuaciones lineales puede desarrollarse sobre cualquier cuerpo, finito o infinito, y abarca aplicaciones continuas y discretas.

- Toda la información relativa a un sistema lineal de ecuaciones puede almacenarse en la matriz ampliada del sistema.

- El algoritmo de eliminación gaussiana proporciona una técnica sistemática y eficiente de resolución de sistemas lineales.

- Cualquier matriz puede ser escalerizada aplicando una cantidad finita de transformaciones elementales.

- Un sistema lineal es incompatible cuando la forma escalerizada de su matriz ampliada tiene mayor número de pivotes que la forma escalerizada de la matriz del sistema.

- Cuando un sistema es compatible, es indeterminado si el número de pivotes en la forma escalerizada de la matriz del sistema es menor que el número de incógnitas. En este caso aparecen variables libres.

- Un sistema que tiene más incógnitas que ecuaciones nunca puede ser determinado.

- Las operaciones con vectores de $\mathbb{K}^n$ dotan a este conjunto de una estructura algebraica que es útil para el estudio de los sistemas lineales de ecuaciones.
- Esta estructura algebraica admite una representación geométrica.

1.3. Espacio de columnas y compatibilidad

En esta sección vamos a cambiar un poco el punto de vista con el que nos enfrentamos a un sistema de ecuaciones lineales de la forma

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1j}x_j + \ldots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2j}x_j + \ldots + a_{2n}x_n & = & b_2, \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mj}x_j + \ldots + a_{mn}x_n & = & b_m, \end{array}\right. \tag{1.26}$$

para concentrarnos en las *columnas* formadas por los coeficientes que multiplican a cada una de las incógnitas

$$x_j, \quad j = 1, 2, \ldots, n.$$

Observemos que cada uno de los números x_j aparece multiplicando a los coeficientes que están en la j-ésima columna de la matriz del sistema. Es decir, cada x_j aparece en una columna

$$\begin{pmatrix} a_{1j}x_j \\ a_{2j}x_j \\ \vdots \\ a_{mj}x_j \end{pmatrix} = x_j \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}. \tag{1.27}$$

En la fórmula (1.27) vemos nuevamente en acción la idea de operar con vectores de $\mathbb{K}^m$ (representados como filas o columnas) que introdujimos en la sección 1.2: ahora la incógnita x_j aparece multiplicando a toda una columna de coeficientes, y la igualdad en (1.27) está justificada por la convención de que multiplicar un número por una columna (fila) es lo mismo que multiplicar cada entrada de la columna (fila) por ese número.

A partir de la observación precedente construiremos una nueva manera de mirar a las soluciones del sistema (1.26), que enfatiza las columnas de la matriz como objetos básicos de trabajo. Nuestro análisis empleará fuertemente la estructura algebraica de $\mathbb{K}^n$, y al desarrollarlo irán emergiendo algunos de los conceptos básicos de esta estructura: *combinaciones lineales*, *subespacios vectoriale*, *conjuntos generadores*, etcétera. Todas estas nociones resultan fundamentales para tratar los sistemas lineales, pero su interés va todavía más lejos y ayudan a comprender otras situaciones. Por lo tanto reaparecerán una y otra vez a lo largo de este texto.

1.3.1. Columnas y combinaciones lineales

En principio, una solución

$$(x_1, x_2, \ldots, x_n)$$

del sistema (1.26) puede ser vista como un conjunto de n números que hace que se satisfagan todas las ecuaciones del sistema. Pero también podemos organizar la información como en la fórmula (1.27), enfatizando que cada x_j multiplica una columna de coeficientes. La suma

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_j \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \tag{1.28}$$

de las n columnas de la matriz multiplicadas por los números x_j es entonces lo mismo que

$$\begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{pmatrix} + \ldots + \begin{pmatrix} a_{1j}x_j \\ a_{2j}x_j \\ \vdots \\ a_{mj}x_j \end{pmatrix} + \ldots + \begin{pmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{pmatrix}. \tag{1.29}$$

Para sumar las columnas en (1.29) debemos hacer m sumas entrada a entrada, con j variando en $1, 2 \ldots, n$. Al hacer esta operación obtenemos la columna

$$\begin{pmatrix} a_{11}x_1 + \ldots + a_{1j}x_j + \ldots + a_{1n}x_n \\ a_{21}x_1 + \ldots + a_{2j}x_j + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mj}x_j + \ldots + a_{mn}x_n \end{pmatrix} \tag{1.30}$$

que contiene todos los primeros miembros de las ecuaciones del sistema. Naturalmente, que el sistema se satisfaga es equivalente a la igualdad

$$\begin{pmatrix} a_{11}x_1 + \ldots + a_{1j}x_j + \ldots + a_{1n}x_n \\ a_{21}x_1 + \ldots + a_{2j}x_j + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mj}x_j + \ldots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

entre la columna (1.30) y la columna

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{1.31}$$

formada por los números b_i, $i = 1, 2, \ldots, m$, de los segundos miembros de las ecuaciones que forman el sistema lineal (1.26).

Podemos ver entonces nuestro sistema de m ecuaciones como una **única** ecuación escrita en términos de vectores de $\mathbb{K}^m$: las columnas de la matriz y la columna de los términos independientes. Si llamamos A_j a la j-ésima columna de la matriz, es decir

$$A_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad j = 1 \ldots, n,$$

podemos representar el sistema (1.26) como

$$x_1 A_1 + x_2 A_2 + \ldots x_j A_j \ldots + x_n A_n = B, \tag{1.32}$$

donde B es la columna (1.31) del término independiente. Notemos que el miembro de la izquierda en (1.32) es lo mismo que (1.28), pero expresado con una notación más breve.

Disponemos ahora de un nuevo punto de vista: **resolver el sistema lineal de ecuaciones (1.26) es buscar una lista de n números**

$$x_1, \ldots, x_n,$$

tales que se satisfaga la igualdad (1.32) entre columnas en $\mathbb{K}^m$.

Ejemplo 1.3.1. Volveremos ahora sobre algunos de los ejemplos de la sección anterior, pero empleando esta vez la notación de columnas.

1. En el ejemplo 1.2.4, página 26, encontramos que el sistema

$$\begin{cases} x_1 + x_2 = 1, \\ x_1 - x_2 = 1, \end{cases}$$

tiene una única solución $(1, 0)$. Al escribir el sistema en la forma vectorial

$$x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

este resultado puede expresarse como

$$1\begin{pmatrix}1\\1\end{pmatrix}+0\begin{pmatrix}1\\-1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}.$$

Analizando en los mismos términos las otras partes de ese ejemplo podemos decir que no existen dos números reales x_1 y x_2 tales que

$$x_1\begin{pmatrix}1\\1\end{pmatrix}+x_2\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}.$$

Y que la igualdad

$$x\begin{pmatrix}1\\2\end{pmatrix}+(1-x)\begin{pmatrix}1\\2\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}$$

se satisface para cualquier $x\in\mathbb{R}$.

2. Encontramos la solución $(1,3,2)$ para el sistema del ejemplo 1.2.7, en la página 30. Algo que ahora podemos escribir del siguiente modo:

$$\begin{pmatrix}1\\0\\0\end{pmatrix}+3\begin{pmatrix}1\\2\\0\end{pmatrix}+2\begin{pmatrix}2\\1\\1\end{pmatrix}=\begin{pmatrix}8\\8\\2\end{pmatrix}$$

3. Análogamente, para el ejemplo 1.2.9 de la página 31 podemos escribir:

$$\begin{pmatrix}1\\1\\2\end{pmatrix}-3\begin{pmatrix}1\\2\\1\end{pmatrix}+5\begin{pmatrix}2\\1\\1\end{pmatrix}=\begin{pmatrix}8\\0\\4\end{pmatrix}$$

4. Al trabajar con el ejemplo 1.2.22 encontramos que la igualdad

$$(-1-2x_2-3x_5/2)\begin{pmatrix}1\\-1\\1\\1\\1\end{pmatrix}+x_2\begin{pmatrix}2\\-2\\2\\2\\2\end{pmatrix}+(1+x_5/2)\begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix}+$$

$$(-1-2x_5)\begin{pmatrix}0\\1\\2\\0\\4\end{pmatrix}+x_5\begin{pmatrix}1\\0\\5\\1\\9\end{pmatrix}+\begin{pmatrix}1\\-1\\3\\3\\3\end{pmatrix}=\begin{pmatrix}1\\0\\1\\3\\-1\end{pmatrix}$$

se satisface para cualquier elección de los números reales x_2 y x_5. ♣

La sencilla operación de multiplicar algunos vectores por coeficientes y luego sumarlos es tan importante que merece un nombre especial y una definición.

Definición 1.2 (Combinaciones lineales). *Si A_j, $j = 1, \ldots, n$, son vectores de $\mathbb{K}^m$, y x_j, $j = 1, \ldots, n$, son números en el cuerpo $\mathbb{K}$, diremos que*

$$x_1 A_1 + x_2 A_2 + \ldots + x_j A_j + \ldots + x_n A_n \tag{1.33}$$

*es la **combinación lineal** de la familia de vectores $(A_1, A_2, \ldots, A_j)$ con **coeficientes** $(x_1, x_2, \ldots, x_j)$.*

*Diremos también que un vector $B \in \mathbb{K}^m$ es **una combinación lineal** de los vectores A_j, $j = 1, \ldots, n$, si existen escalares $x_j \in \mathbb{K}$, $j = 1 \ldots, n$ tales que*

$$B = x_1 A_1 + x_2 A_2 + \ldots + x_j A_j + \ldots + x_n A_n.$$

Observemos entonces que **resolver el sistema lineal de ecuaciones (1.26) es buscar coeficientes $x_1, \ldots, x_n$, tales que B sea una combinación lineal de las columnas de la matriz A del sistema con esos coeficientes**. También es cierto que **el sistema lineal de ecuaciones (1.26) es compatible si y sólo si B es una combinación lineal de las columnas de la matriz A.**

Pasemos ahora a examinar algunos ejemplos de combinaciones lineales. Escribiremos todos los vectores como filas, no como columnas, simplemente para ahorrar algo de espacio en el papel.

Ejemplo 1.3.2. Sea $A = ((1, 2, 1), (2, -2, 2)) \subset \mathbb{R}^3$ entonces $(0, 6, 0)$ es combinación lineal de A con coeficientes 2 y -1 pues

$$(0, 6, 0) = 2(1, 2, 1) + (-1)(2, -2, 2).$$

Pero no es posible escribir $(0, 6, 1)$ como una combinación lineal de estos dos vectores. Verifiquémoslo. Expresar $(0, 6, 1)$ como combinación lineal de $(1, 2, 1)$ y $(2, -2, 2)$ es lo mismo que encontrar dos coeficientes x e y tales que

$$(0, 6, 1) = x(1, 2, 1) + y(2, -2, 2) = (x + 2y, 2x - 2y, x + 2y).$$

Y esta igualdad es equivalente al sistema de ecuaciones

$$\left\{ \begin{array}{rcl} x + 2y &=& 0, \\ 2x - 2y &=& 6, \\ x + 2y &=& 1. \end{array} \right.$$

Al escalerizarlo encontramos

$$\left\{\begin{array}{rcl} x+2y &=& 0, \\ -4y &=& 6, \\ 0 &=& 1, \end{array}\right.$$

por lo que concluimos que el sistema es incompatible y es imposible expresar $(0,6,1)$ como combinación lineal de los dos vectores dados. Observemos que hemos pasado del problema de expresar un vector como combinación lineal de otros, a un sistema de ecuaciones lineales. Tal como comentábamos, ambos problemas son en realidad equivalentes. ♣

Ejemplo 1.3.3. Sea $A = ((1,2,1),(2,-2,2),(1,8,1))$. Combinemos linealmente los vectores de A con coeficientes 3,−1 y 1. Obtenemos

$$3(1,2,1)+(-1)(2,-2,2)+1(1,8,1)=(2,16,2).$$

Y si hacemos la combinación con coeficientes 6, -2 y 0 resulta

$$6(1,2,1)+(-2)(2,-2,2)+0(1,8,1)=(2,16,2).$$

¡Exactamente el mismo vector! Vemos entonces que distintas combinaciones lineales de una misma familia de vectores pueden dar el mismo resultado. Mirando las cosas desde otro punto de vista: hay ejemplos –acabamos de ver uno– en que un vector puede ser expresado en más de una forma como combinación lineal de una familia de vectores dados. Naturalmente, estos ejemplos corresponden a sistemas lineales compatible e indeterminados. ♣

Ejemplo 1.3.4. La m-upla nula $O = (0,\dots,0)$ es combinación lineal de cualquier colección de m-uplas, basta tomar todos los coeficientes nulos para obtenerla. Por ejemplo, si tomamos $n = 3$ y los tres vectores del item anterior, tenemos, obviamente

$$0(1,2,1)+0(2,-2,2)+0(1,8,1)=(0,0,0).$$

Este hecho corresponde a que **un sistema lineal que tiene una columna de ceros en los términos independientes es siempre compatible** porque tiene al menos la solución trivial que consiste en fijar en 0 el valor de todas sus incógnitas. Un resultado que encontramos al resolver el ejercicio 1.12

Pero, **la solución trivial no tiene por qué ser la única**. De hecho, ya sabemos que un sistema que tiene más incógnitas que ecuaciones no puede tener solución única. Establecimos este resultado en la parte $3c$ del ejercicio 1.34, en la página 66.

Por ejemplo, el sistema de ecuaciones lineales reales

$$\left\{\begin{array}{rcrcrcl} x & + & 2y & - & z & = & 0 \\ 2x & + & 4y & - & 3z & = & 0 \end{array}\right.$$

tiene como solución cualquier terna

$$(x, y, z) = (-2\lambda, \lambda, 0),$$

donde $\lambda \in \mathbb{R}$ puede escogerse arbitrariamente. Si fijamos $\lambda = 0$ reencontramos la ecuación trivial, pero $(-2, 1, 0)$, que se obtiene escogiendo $\lambda = 1$, es otra solución del sistema. Naturalmente, este resultado tiene su correlato en términos de combinaciones lineales de las columnas de la matriz del sistema. La igualdad

$$-2\lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 4 \end{pmatrix} + 0 \begin{pmatrix} -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

se satisface para cualquier $\lambda \in \mathbb{R}$.

Como ya mencionamos, **siempre es posible obtener el vector nulo a partir de cualquier familia de vectores de $\mathbb{K}^m$ haciendo una combinación lineal con todos los coeficientes nulos**, pero este ejemplo nos muestra que también **puede ser posible obtener el vector nulo por medio de una combinación lineal no trivial**. ♣

Ejercicio 1.36. Consideremos las 4-uplas

$$X_1 = (1, 2, 2, 1), \quad X_2 = (2, 1, -2, 0), \quad X_3 = (-1, 1, 4, 1).$$

Hallar las combinaciones lineales $aX_1 + bX_2 + cX_3$ para:

1. $a = 0$; $b = 2$; $c = 5$
2. $a = -3$; $b = 2$; $c = 1$
3. $a = 1$; $b = -1$; $c = -1$

Ejercicio 1.37. Consideremos el conjunto

$$A = ((1, 0, 1, -1), (2, 0, 3, 1), (0, 2, 1, 0))$$

formado por tres 4-uplas de números reales. Determinar en cada caso si X puede obtenerse como combinación lineal de los elementos de A. Si la respuesta es afirmativa hallar los respectivos coeficientes.

1. $X = (0, 2, 0, -3)$
2. $X = (5, -2, 0, 0)$
3. $X = (5, -6, 4, 1)$

Observación 1.3.5. Notación de sumatorias
Es posible abreviar la escritura de la suma

$$x_1A_1 + x_2A_2 + \ldots + x_jA_j + \ldots + x_nA_n$$

que representa la combinación lineal de los vectores A_j con coeficientes

$$x_j, \quad j = 1, 2, \ldots, n,$$

con la notación que emplea el signo $\sum$ para las sumas. Escribiremos entonces

$$x_1A_1 + x_2A_2 + \ldots + x_jA_j + \ldots + x_nA_n = \sum_{j=1}^{n} x_jA_j,$$

donde j es un número entero que indica cual es el índice que varía al sumar, y 1 y n son los valores mínimo y máximo que toma j en la suma que estamos considerando. ♠

En la subsección 1.3.3 introduciremos una notación aún más concisa para expresar las combinaciones lineales de vectores en $\mathbb{K}^m$ cuando definamos el producto entre una matriz A y un vector X. Pero antes de pasar a considerar el producto de matrices y vectores discutiremos brevemente la geometría de las combinaciones lineales y los sistemas de ecuaciones.

1.3.2. Interpretación geométrica de las combinaciones lineales

Es sumamente interesante ver una representación geométrica de las combinaciones lineales, que está basada en la interpretación geométrica de los espacios $\mathbb{R}^n$ que discutimos en la sección 1.2.6, a partir de la página 70.

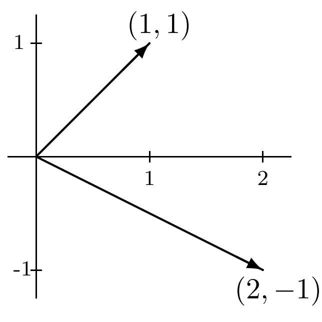
Consideremos, por ejemplo, la combinación lineal

$$x_1A_1 + x_2A_2$$

de dos vectores A_1 y A_2, con coeficientes x_1 y x_2. El vector x_1A_1 es colineal con A_1, pero cambia su longitud en un factor x_1. Además, si x_1 es positivo el vector x_1A_1 tiene el mismo sentido que A_1, y el contrario cuando x_1 es negativo. Una consideración análoga puede hacerse para x_2A_2 en relación A_2. En resumen, $x_1A_1 + x_2A_2$ es la suma de dos vectores, uno alineado con A_1 y el otro alineado con A_2.

Ejemplo 1.3.6. Trabajaremos en este ejemplo con los vectores $(1, 1)$ y $(2, -1)$ en $\mathbb{R}^2$, que se representan en la figura, y su combinación lineal

$$(1, 4) = 3(1, 1) + (-1)(2, -1). \qquad (1.34)$$



Los vectores

$$3(1,1) = (3,3), \quad (-1)(2,-1) = (-2,1)$$

son colineales con $(1,1)$ y $(2,-1)$ respectivamente. Los hemos representado en el primer dibujo de la figura 1.5.

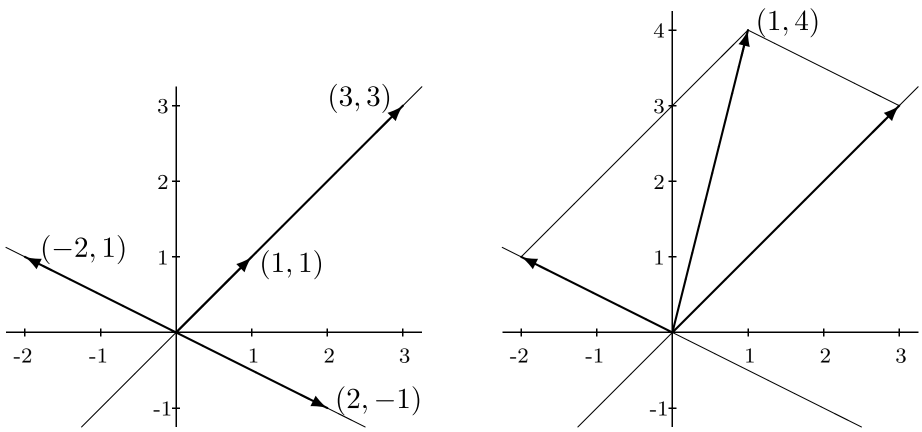


Figura 1.5: El vector $(1,4)$ como combinación lineal de $(1,1)$ y $(2,-1)$

Observemos que $(-2,1)$ tiene sentido contrario a $(2,-1)$, ya que se obtiene a partir de este último vector multiplicando por el coeficiente negativo -1. Hemos representado también rectas que pasan por el origen y que tienen las direcciones de los vectores $(1,1)$ y $(2,-1)$. Los extremos de cualquier vector que se obtenga multiplicando $(1,1)$ por un número caen sobre la línea con la dirección de $(1,1)$. Análogamente, al multiplicar $(2,-1)$ por un escalar obtenemos un nuevo vector en la dirección de $(2,-1)$.

La combinación lineal (1.34) es la suma de $(3,3)$ y $(-2,1)$, que sigue la regla del paralelogramo: el extremo del vector suma es el vértice opuesto a $(0,0)$ en el paralelogramo que tiene a los vectores $(-2,1)$ y $(3,3)$ como dos de sus lados. Esta suma se ilustra en el segundo dibujo de la figura 1.5. ♣

La interpretación geométrica de las combinaciones lineales que acabamos de desarrollar se extiende a una interpretación geométrica de los sistemas de ecuaciones. En nuestro próximo ejemplo desarrollamos este punto de vista, volviendo sobre los vectores con los que trabajamos en el ejemplo 1.3.6, pero presentándolo a través de un sistema de ecuaciones lineales.

Ejemplo 1.3.7. Consideremos el sistema

$$\left\{ \begin{array}{r} x + 2y = 1 \\ x - y = 4 \end{array} \right. \tag{1.35}$$

Podemos escribirlo en forma matricial como

$$\left(\begin{array}{rr|r} 1 & 2 & 1 \\ 1 & -1 & 4 \end{array} \right),$$

una representación que enfatiza el hecho de que estamos buscando una combinación lineal de los vectores $(1,1)$ y $(2,-1)$ que produzca el vector $(1,4)$. Se trata entonces de buscar una descomposición de $(1,4)$ como la suma de un vector colineal con $(1,1)$ y $(2,-1)$, tal como se muestra en la figura 1.6. Por

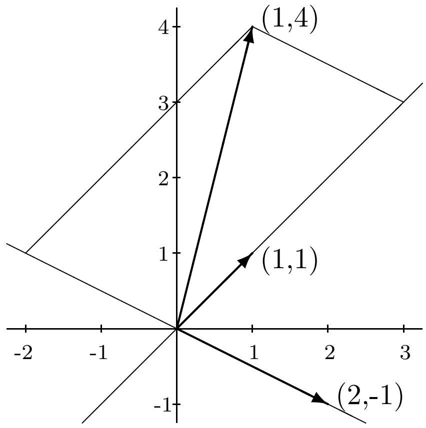


Figura 1.6: Representación geométrica de la solución del sistema (1.35)

supuesto, esta nueva figura es esencialmente la misma que 1.5, pero hemos enfatizado que $(1,4)$ se descompone sobre las líneas en las direcciones de $(1,1)$ y $(2,-1)$. Los vectores en los que se descompone $(1,4)$ son los vectores $(3,3)$ y $(-2,1)$ y los coeficientes necesarios para obtenerlos a partir de $(1,1)$ y $(2,-1)$ son 3 y -1 respectivamente. Naturalmente, la solución del sistema de ecuaciones (1.35) es la pareja $x=3$, $y=-1$. ♣

Interpretación de un sistema de ecuaciones como la intersección de rectas

Existe otra interpretación geométrica de un sistema lineal de ecuaciones, que en el caso de un sistema con dos incógnitas conduce a identificar las soluciones del sistema con los puntos que pertenecen a la intersección de dos rectas.

Para introducir esta nueva interpretación vamos a abandonar por un momento la representación de las parejas (x, y) de números reales como vectores, para considerar que un par de números (x, y) determina un punto en el plano (x, y). En la figura 1.7 ilustramos con la pareja $(2, 1)$ estos dos posibles puntos de vista. En el primer dibujo $(2, 1)$ aparece representada como un vector. En

Figura 1.7: Parejas de $\mathbb{R}^2$ representando vectores o puntos

el segundo como un punto en el plano (x, y). La conexión entre ambas representaciones es la siguiente: el vector $(2, 1)$ es el único vector que tiene su origen en $(0, 0)$ y extremo en el punto $(2, 1)$; el punto $(2, 1)$ es el que se obtiene cuando nos desplazamos desde el origen según el vector $(2, 1)$.

Una ecuación en las variables x e y, como, por ejemplo, la ecuación

$$x + 2y = 1, \tag{1.36}$$

que es la primera ecuación del sistema (1.35), define un cierto subconjunto de puntos (x, y) en $\mathbb{R}^2$. Por ejemplo, el conjunto de los puntos (x, y) que satisfacen (1.36) es la recta que representamos en la figura 1.8. Con estas ideas en

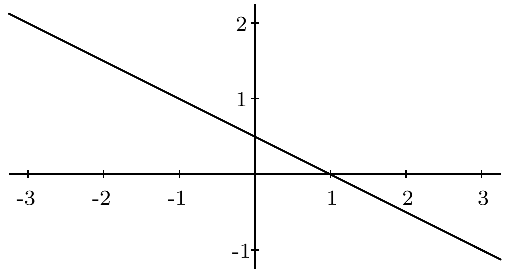


Figura 1.8: La recta de ecuación $x + 2y = 1$

mente volveremos sobre el sistema lineal (1.35).

Ejemplo 1.3.8. Los puntos (x, y) cuyas coordenadas satisfacen el sistema (1.35) están sobre la recta de ecuación

$$x + 2y = 0,$$

y también sobre la recta de ecuación

$$x - y = 4.$$

Por lo tanto, la solución de la ecuación está representada por los puntos que pertenecen a la intersección de ambas rectas. Al representarlas en el plano encontramos la figura 1.9, en la que apreciamos que la intersección de las dos

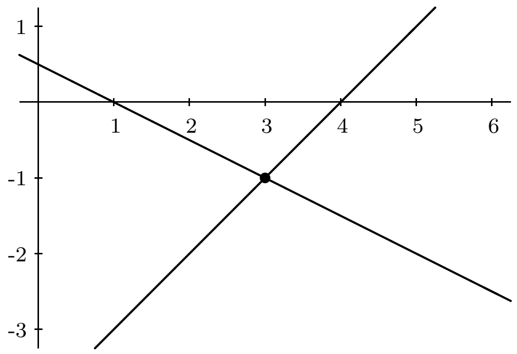


Figura 1.9: La intersección de las rectas de ecuaciones $x + 2y = 1$, $x - y = 4$

rectas es el punto $(3, -1)$. Esto está en concordancia con el resultado que habíamos obtenido al resolver el sistema (1.35). ♣

Ejemplo 1.3.9. Analicemos el sistema indeterminado

$$\left\{ \begin{array}{l} 2x - y = 3, \\ 4x - 2y = 6, \end{array} \right.$$

desde los dos puntos de vista geométricos que hemos estado considerando. Las soluciones de este sistema son todas las parejas $(x, 2x - 3)$, donde x es un número real cualquiera.

En primer lugar observemos que el sistema es equivalente a la combinación lineal

$$x \begin{pmatrix} 2 \\ 4 \end{pmatrix} + y \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

Los tres vectores $(2, 4)$, $(-1, -2)$ y $(3, 6)$ son colineales, y $(3, 6)$ puede obtenerse multiplicando $(2, 4)$ por $3/2$, o $(-1, -2)$ por -3, entre otras muchas posibilidades. La primera corresponde a la solución $(3/2, 0)$, y la segunda a $(0, -3)$.

Por otra parte, las dos ecuaciones

$$2x - y = 3, \quad 4x - y = 6$$

definen exactamente la misma recta, ya que la segunda se obtiene multiplicando a la primera por la constante no nula 2. Por supuesto, la intersección que se obtiene al cortarlas es justamente esa recta. ♣

Ejemplo 1.3.10. Modifiquemos levemente el ejemplo anterior, cambiando el término independiente para considerar el sistema

$$\left\{\begin{array}{l} 2x - y = 3, \\ 4x - 2y = 2, \end{array}\right. \tag{1.37}$$

El vector $(3, 2)$ ya no es colineal con $(2, 4)$ y $(-1, -2)$, y no puede ser generado por una combinación lineal de estos vectores, tal como se muestra en el primer

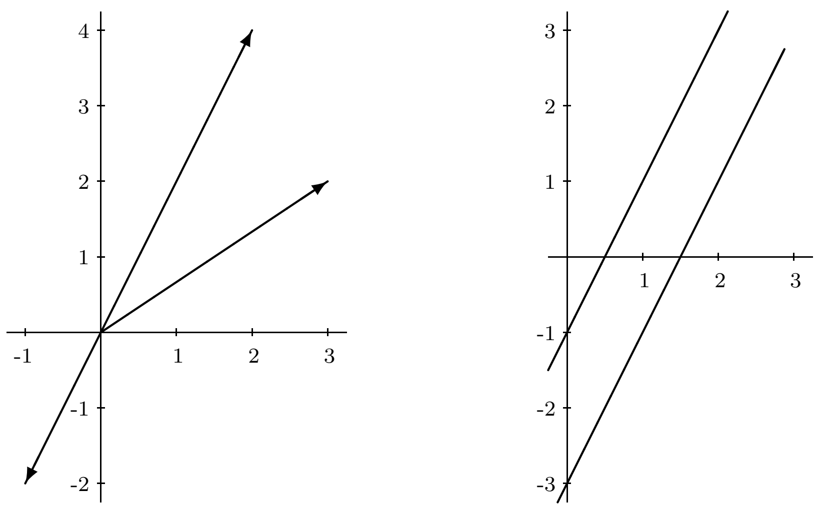


Figura 1.10: Dos representaciones geométricas del sistema incompatible (1.37)

dibujo de la figura 1.10. Las ecuaciones

$$2x - y = 3, \quad 4x - 2y = 2,$$

definen las dos rectas paralelas que están representadas en el segundo dibujo de la figura 1.10. ♣

Ejercicio 1.38. Hacer las representaciones geométricas como combinaciones lineales y como intersecciones de rectas para el sistema

$$\left\{\begin{array}{l} x - 2y = 4, \\ 3x - 2y = 8. \end{array}\right.$$

La interpretación geométrica de los sistemas 2×2 que hemos discutido en esta sección puede extenderse a sistemas con mas incógnitas. En efecto, la noción de combinación lineal ha sido introducida para cualquier cantidad de vectores de cualquier tamaño.

Por otra parte, veremos en el capítulo dedicado a la geometría de $\mathbb{R}^3$ que una ecuación lineal con tres incógnitas representa un plano en el espacio. Las soluciones de un sistema lineal con tres incógnitas aparecerán entonces como intersecciones de planos, que pueden ser vacías, puntos, rectas o planos, dependiendo de las posiciones relativas de los planos que se intersequen.

Al igual que ocurre con otros aspectos geométricos de la teoría que estamos desarrollando, estas consideraciones pueden extenderse a cualquier número n de incógnitas, lo que corresponde a desarrollar una geometría de $\mathbb{R}^n$, tal como apuntábamos en la página 1.2.6.

1.3.3. El producto de matrices y vectores

Presentaremos ahora una definición del producto entre una matriz y un vector que se adecua perfectamente a nuestro trabajo con los sistemas de ecuaciones lineales y a la noción de combinación lineal.

Definición 1.3 (Producto de matrices y vectores). *Sean A una matriz $m \times n$ sobre el cuerpo $\mathbb{K}$ con columnas*

$$A_j, \quad j = 1, 2, \ldots, n,$$

y sea

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

una columna de longitud n. Definimos el producto AX de la matriz A por la columna X como la combinación lineal

$$AX = x_1A_1 + x_2A_2 + \ldots x_jA_j \ldots + x_nA_n$$

de las columnas de A con los coeficientes almacenados en la columna X.

Observación 1.3.11. De momento, para la definición anterior lo único importante es que X sea una lista de longitud n. La restricción de escribirla como una columna quedará justificada más adelante, en la sección 2.2, cuando definamos el producto de matrices en un contexto más general que incluye el caso particular del producto entre una matriz $m \times n$ y una columna con n elementos, que puede ser interpretada como una matriz $n \times 1$

A partir de la definición 1.3 es posible introducir una notación aún más concisa que (1.32) para un sistema lineal de ecuaciones. La voluminosa y amenazadora expresión

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1j} + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2j} + \ldots + a_{2n}x_n & = & b_2 \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mj} + \ldots + a_{mn}x_n & = & b_m \end{array}\right.$$

para el sistema con matriz y término independiente

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

respectivamente, queda condensada en la casi insignificante fórmula

$$AX = B. \tag{1.38}$$

¿Hemos ganado algo con todo este alarde de notación? Poco, al menos ahora es más fácil escribir un sistema lineal... pero su resolución sigue igual de complicada que antes.

Aunque el propósito de simplificar la notación podría justificar por sí mismo introducir tal definición, no hemos definido el producto AX sólo para simplificar nuestra escritura. Este producto tiene además buenas propiedades respecto a la estructura lineal con la que estamos trabajando y múltiples aplicaciones, por lo que aparecerá una y otra vez a lo largo de este curso.

En términos del producto de una matriz por un vector podemos decir que **resolver el sistema lineal de ecuaciones (1.26) es equivalente a encontrar un vector**

$$X = (x_1, \ldots, x_n) \in \mathbb{K}^n$$

tal que $AX = B$. El sistema $AX = B$ es compatible si y sólo si la columna B puede escribirse en la forma AX para algún $X \in \mathbb{K}^n$.

Observación 1.3.12. Componentes y notación con sumatorias
Cuando A es una matriz $m \times n$ la notación AX engloba en una sencilla expresión un vector que tiene m componentes. A veces nos será útil considerar las matrices y vectores como un único objeto, sin detenernos a pensar en cada

una de sus componentes individualmente. En ocasiones haremos lo contrario y trabajaremos con las componentes. Para estos casos es conveniente saber que en la i-ésima entrada del producto AX aparece la suma

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n$$

de las entradas i-ésimas de cada columna, multiplicadas por el coeficiente que multiplique a esa columna. Podemos recurrir al signo de sumatoria para expresar una suma de este tipo:

$$\sum_{j=1}^{n} a_{ij}x_j = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n,$$

con lo que se gana cierta economía en la notación. ♠

Ejemplo 1.3.13. Calculemos el producto AX de

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \\ 1 & 2 \end{pmatrix}, \quad X = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Tenemos que

$$AX = 2\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + (-1)\begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 0 \end{pmatrix}.$$

♣

Ejemplo 1.3.14. Sea

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & -2 & 8 \\ 1 & 2 & 1 \end{pmatrix}.$$

Entonces

$$A\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = A\begin{pmatrix} 6 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 16 \\ 2 \end{pmatrix}.$$

♣

El lector atento habrá notado que en los últimos dos ejemplos hemos presentado nuevamente algunas de las combinaciones lineales de los ejemplos 1.3.2 y 1.3.3, ahora escritas en forma de productos de matrices por columnas.

Ejemplo 1.3.15.

1. Al multiplicar cualquier matriz $m \times n$ por la columna O de n ceros se obtiene una columna de m ceros. Es decir, $AO = O$. Notará el lector que el símbolo O aparece en esta fórmula con dos significados posiblemente diferentes: es una lista de n ceros del lado izquierdo de la igualdad, pero una lista de m ceros del lado derecho.

   Por supuesto, cualquier sistema lineal de la forma $AX = O$, es decir, un sistema en que todas las ecuaciones tienen un 0 en el miembro de la derecha, tiene al menos la solución $X = O$, ya que $AO = O$.

   Con estos comentarios no estamos haciendo más que reformular buena parte del ejemplo 1.3.4, ahora usando la notación matricial para los sistemas y las combinaciones lineales.

2. Si multiplicamos una matriz $m\times n$ por una columna formada por ceros en todas sus entradas salvo la i-ésima, en la que ponemos un 1, el resultado es la i-ésima columna de la matriz. ♣

Ejercicio 1.39. En cada caso, multiplicar la matriz A por el vector X.

1. $A = \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix}$ y $X = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$.
2. $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & -1 & 1 \\ 0 & 3 & 3 \end{pmatrix}$ y $X = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$.
3. $A = \begin{pmatrix} 3 & -1 & 0 \\ 1 & 3 & 5 \\ -1 & 3 & 6 \end{pmatrix}$ y $X = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$.
4. Con la aritmética de $\mathbb{Z}_2$, $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ y $X = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

La siguiente proposición muestra que la operación de multiplicar una matriz por un vector tiene buenas propiedades respecto a las operaciones de suma y de producto por un escalar que están definidas sobre los vectores.

Proposición 1.6. Linealidad en X del producto AX.
Sean A una matriz $m\times n$ sobre el cuerpo $\mathbb{K}$, X e Y dos columnas de longitud n, y a un número cualquiera en $\mathbb{K}$. Entonces

1. $A(X + Y) = AX + AY$;
2. $A(aX) = a(AX)$.

Prueba. Escribamos

$$X = (x_1, \dots, x_n), \quad Y = (y_1, \dots, y_n).$$

Entonces la j-ésima entrada de $X + Y$ es $x_j + y_j$. Recurramos a la expresión con sumatorias del producto $A(X + Y)$. La i-ésima entrada de esta columna es

$$\sum_{j=1}^{n} a_{ij}(x_j + y_j) = \sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{n} a_{ij}y_j,$$

que es la suma de la i-ésima entrada de AX con la i-ésima entrada de AY. Por lo tanto $A(X + Y) = AX + AY$. Dejamos la demostración de la segunda propiedad como un ejercicio para el lector.

Ejercicio 1.40. Completar la demostración de la proposición anterior.

Observación 1.3.16. Matrices y transformaciones lineales.
La operación de multiplicar $X \in \mathbb{K}^m$ por una matriz A de dimensiones $m \times n$ define una correspondencia

$$X \mapsto AX$$

que a cada X le asocia $AX \in \mathbb{K}^m$.

Ejemplo 1.3.17. En el ejercicio 1.10, página 21 describíamos las transiciones dentro de una población. Conociendo los proporciones (c, r, e) de pobladores en la *ciudad*, zonas *rurales* y en el *exterior* en un momento dado, dábamos una serie de reglas que permitían calcular la distribución (c_1, r_1, e_1) un año más tarde. No es difícil verificar que la regla general que regula estas transiciones puede escribirse como el producto de una matriz por un vector, en la forma

$$\begin{pmatrix} c_1 \\ r_1 \\ e_1 \end{pmatrix} = \begin{pmatrix} 0{,}6 & 0{,}4 & 0{,}1 \\ 0{,}1 & 0{,}4 & 0{,}1 \\ 0{,}3 & 0{,}2 & 0{,}8 \end{pmatrix} \begin{pmatrix} c \\ r \\ e \end{pmatrix}.$$

La matriz aparece entonces **transformando** el vector con el estado inicial en el vector con el estado final de la población.

En el próximo ejercicio mostramos como una matriz puede transformar una palabra dada en la palabra que le corresponde en un cierto proceso de codificación.

Ejercicio 1.41. En el ejercicio 1.28, página 60, se dan ciertas reglas de codificación de una lista X de cuatro dígitos en una lista Y de siete dígitos. Expresar estas reglas de codificación en forma matricial. Es decir, hallar una matriz G con entradas en $\mathbb{Z}_2$ tal que el resultado Y de codificar una lista de cuatro dígitos X pueda escribirse en la forma del producto $Y = GX$ entre la matriz G y el vector X. ♣

Esta idea de que las matrices actúan transformando cosas es sumamente fecunda, y volveremos sobre ella a lo largo del curso. En particular, lo haremos en el capitulo 2 destinado a las matrices. De momento, digamos que la proposición 1.6 asegura que esta correspondencia es una *transformación lineal*, en el siguiente sentido:

1. la imagen de la suma de dos vectores X e Y es la suma de sus imágenes;

2. la imagen del resultado de multiplicar un vector por un escalar es el producto de ese mismo escalar por la imagen del vector.

Las transformaciones lineales, a las que dedicaremos un capítulo específico son un objeto de estudio central en el álgebra lineal y de interés en diversas aplicaciones. Existe una conexión muy estrecha entre las matrices y las transformaciones lineales, un tema sobre el que trabajaremos cuando introduzcamos la noción de matriz asociada a una transformación lineal. ♠

1.3.4. Compatibilidad del sistema. Subespacio de columnas

Repasemos las nuevas interpretaciones de la resolución de un sistema de ecuaciones lineales con m ecuaciones y n incógnitas: se trata de buscar una lista X de n coeficientes que produzcan el término independiente B al hacer con estos coeficientes una combinación lineal de las columnas de la matriz A. De acuerdo con la definición del producto de una matriz por un vector, esto es completamente equivalente a que B sea igual al producto AX. Este punto de vista ofrece una caracterización de los vectores B que hacen compatible el sistema: son aquellos vectores B que pueden ser escritos como combinación lineal de las columnas de A, o, equivalentemente, como un producto AX, donde X es algún vector en $\mathbb{K}^n$.

En otras palabras, el sistema es compatible si B está en el conjunto formado por **todas las posibles combinaciones lineales** de las columnas de A. Este conjunto es relevante para nuestra teoría, de modo que nos detendremos un momento a ponerle nombre.

Definición 1.4 (Espacio de columnas). *Si A es una matriz $m \times n$ sobre $\mathbb{K}$ llamaremos **espacio de columnas** de A al subconjunto de $\mathbb{K}^m$ formado por todas las posibles combinaciones lineales de las columnas de la matriz.*

Indicaremos este conjunto con la notación $\mathrm{col}(A)$.

Tal como observábamos, el espacio de columnas coincide con el conjunto de todos los posibles productos AX, con $X \in \mathbb{K}^n$. Es decir

$$\mathrm{col}(A) = \{AX;\ X \in \mathbb{K}^n\}.$$

Como resumen de la discusión que precedió a la definición 1.4 encontramos que, **el sistema** $AX = B$ **es compatible si y sólo si** $B \in \mathrm{col}(A)$.

Observación 1.3.18. La primera parte del ejemplo 1.3.15 nos dice que si A es una matriz $m \times n$ entonces el vector nulo $O \in \mathbb{K}^m$ está en $\mathrm{col}(A)$. La segunda parte del mismo ejemplo implica que todas las columnas de la matriz pertenecen a $\mathrm{col}(A)$. ♠

Ejemplo 1.3.19. El vector $(0,1,1,0) \in \mathbb{Z}_2^4$ pertenece al espacio de columnas de la matriz

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix},$$

pero $(1,1,1,1)$ no. Para justificar estas afirmaciones sólo hay que volver el ejemplo 1.2.25, en la página 48, y reinterpretar los resultados en términos del espacio de columnas de la matriz A. ♣

Ejercicio 1.42. Hallar el espacio de columnas de la matriz G que se halló en el ejercicio 1.41 (como estamos trabajando sobre el cuerpo finito $\mathbb{Z}_2$ este espacio de columnas es un conjunto finito. Tiene dieciséis elementos). Comparar el resultado con el que se obtuvo en la parte 1 del ejercicio 1.28.

Observación 1.3.20. El conjunto $\mathrm{col}(A)$ puede interpretarse también teniendo en cuenta la transformación $X \mapsto AX$ que entre los espacios $\mathbb{K}^n$ y $\mathbb{K}^m$ define la matriz A (ver la observación 1.3.16). Los vectores AX de $\mathbb{K}^m$ son justamente aquellos que son la imagen de algún vector $X \in \mathbb{K}^n$ a través de la transformación que A define. El conjunto $\mathrm{col}(A)$ resulta ser entonces la *imagen* de esta transformación, por lo que es también es usual referirse a él con la notación $\mathrm{im}(A)$. ♠

Es sumamente importante para la caracterización de la compatibilidad del sistema $AX = B$ en términos del conjunto $\mathrm{col}(A)$ el hecho de que este espacio de columnas está dotado de una estructura algebraica que nos permitirá manejarlo y describirlo: es posible sumar vectores en $\mathrm{col}(A)$ y multiplicarlos por números sin salir de $\mathrm{col}(A)$. Ese es el contenido de nuestra próxima proposición.

Proposición 1.7. *Si A es una matriz $m \times n$ sobre $\mathbb{K}$ entonces:*

1. *si Y_1 e Y_2 pertenecen a $\mathrm{col}(A)$, su suma $Y_1 + Y_2$ también pertenece a $\mathrm{col}(A)$;*
2. *si Y pertenece a $\mathrm{col}(A)$ y a es un número cualquiera en $\mathbb{K}$, el producto aY también pertenece a $\mathrm{col}(A)$.*

Prueba: si Y_1 e Y_2 pertenecen a $\mathrm{col}(A)$ entonces existen vectores X_1 y X_2 en $\mathbb{K}^n$ tales que

$$Y_1 = AX_1, \quad Y_2 = AX_2.$$

Sumando estas dos igualdad obtenemos

$$Y_1 + Y_2 = AX_1 + AX_2 = A(X_1 + X_2).$$

Vemos entonces que la suma Y_1+Y_2 es el producto de la matriz A por el vector

$$X_1 + X_2 \in \mathbb{K}^n.$$

Por lo tanto $Y_1 + Y_2$ es una combinación lineal de las columnas de A y está en $\mathrm{col}(A)$. La demostración para el producto aY es similar, y la dejamos como ejercicio para el lector. □

Ejercicio 1.43. Completar la prueba de la proposición 1.7.

Es usual referirse a las propiedades de $\mathrm{col}(A)$ que demostramos en la proposición 1.7 diciendo que $\mathrm{col}(A)$ es cerrado bajo las operaciones de suma de vectores y producto por un escalar. Esto significa que cuando sumamos dos vectores en $\mathrm{col}(A)$, o cuando multiplicamos un vector en $\mathrm{col}(A)$ por un escalar, no nos "salimos" de $\mathrm{col}(A)$ en el sentido de que el resultado de ambas operaciones también pertenece a este conjunto.

Un subconjunto $\mathbb{S}$ no vacío de $\mathbb{K}^m$ que es cerrado bajo la suma y el producto por un escalar es lo que llamaremos un *subespacio vectorial* de $\mathbb{K}^m$.

La designación de *espacio* obedece al hecho de que se puede sumar y multiplicar por escalares dentro de $\mathbb{S}$, es decir, $\mathbb{S}$ **tiene una estructura algebraica**. Nos referimos a $\mathbb{S}$ como *subespacio* porque está contenido dentro del espacio más grande $\mathbb{K}^m$.

Con estos términos, teniendo en cuenta la observación 1.3.18 que asegura que $\mathrm{col}(A)$ no es vacío, el resultado de la proposición 1.7 puede enunciarse diciendo que $\mathrm{col}(A)$ **es un subespacio vectorial de** $\mathbb{K}^m$.

El subespacio $\mathrm{col}(A)$ está *generado* por las columnas de la matriz A en el sentido de que cualquier vector de $\mathrm{col}(A)$ puede escribirse como combinación lineal de las columnas de A. Esta caracterización provee una descripción de $\mathrm{col}(A)$, pero veremos luego que en muchos casos no hace falta utilizar **todas** las columnas de A para generar $\mathrm{col}(A)$: puede ocurrir que un subconjunto más pequeño de las columnas de A sea suficiente para hacerlo, y se consigue así una descripción más breve del espacio $\mathrm{col}(A)$.

En la sección 1.4 utilizaremos la estructura de subespacio de $\mathrm{col}(A)$ y el método de escalerización para dar dos descripciones de este subespacio:

- por medio de un conjunto de ecuaciones lineales que son satisfechas por los elementos de $\text{col}(A)$;
- identificando un subconjunto de las columnas de A que forman un generador óptimo, en el sentido de que si elimina alguna columna de este subconjunto el subconjunto resultante ya no puede generar todo $\text{col}(A)$.

Para avanzar en esta discusión formalizaremos las nociones de subespacio y generador que acabamos de introducir informalmente. Comenzamos por dar una definición precisa de subespacio, y algunos ejemplos.

Definición 1.5. *Un subconjunto $\mathbb{S}$ no vacío de $\mathbb{K}^n$ es un* ***subespacio vectorial*** *de $\mathbb{K}^n$ si tiene las siguientes dos propiedades:*

1. *la suma de dos vectores cualesquiera de $\mathbb{S}$ pertenece a $\mathbb{S}$;*
2. *el producto de un vector de $\mathbb{S}$ por un escalar cualquiera en $\mathbb{K}$ pertenece a $\mathbb{S}$.*

Ejemplo 1.3.21.

1. El espacio de columnas de una matriz A de dimensiones $m \times n$ es un subespacio de $\mathbb{K}^m$. Este conjunto es no vacío, porque todas las columnas de A están en $\text{col}(A)$, y además es cerrado bajo las operaciones de suma y producto por un escalar.
2. Hay dos subconjuntos de $\mathbb{K}^n$ para los que es fácil verificar que se trata de subespacios vectoriales:

   a) El *subespacio trivial* que sólo está formado por O, el vector nulo del espacio $\mathbb{K}^n$. Este conjunto es no vacío, tenemos además que

   $$O + O = O,$$

   y para cualquier $a \in \mathbb{K}$ se satisface

   $$aO = O.$$

   Éste es el subespacio más chico que podemos tener. De hecho, está contenido en cualquier otro subespacio. Dejamos la verificación de esta afirmación como un ejercicio para el lector.

   Ejercicio 1.44. Mostrar que si $\mathbb{S}$ es un subespacio vectorial de $\mathbb{K}^m$ entonces $O \in \mathbb{S}$, por lo que $\{O\} \subset \mathbb{S}$. Sugerencia: utilizar el hecho de que el producto del escalar 0 por cualquier vector de $\mathbb{S}$ debe estar en $\mathbb{S}$.

b) El propio $\mathbb{K}^n$.

3. El subconjunto
$$\mathbb{S} = \{(x,y) \in \mathbb{R}^2;\ x = y\}$$
es un subespacio vectorial de $\mathbb{R}^2$. Justifiquemos esta afirmación. Claramente $\mathbb{S}$ es no vacío. Un elemento genérico de $\mathbb{S}$ es de la forma (x,x), con x un número real cualquiera. Consideremos dos elementos (x_1,x_1), (x_2,x_2) de $\mathbb{S}$, y un escalar a cualquiera. Al sumar tenemos
$$(x_1,x_1)+(x_2,x_2) = (x_1+x_2, x_1+x_2).$$
Al multiplicar uno de ellos, por ejemplo (x_1,x_1) por el escalar a resulta
$$a(x_1,x_1) = (ax_1, ax_1).$$
Ambos resultados tienen la primera componente igual a la segunda, por lo tanto satisfacen la condición de pertenencia a $\mathbb{S}$.

4. El subconjunto
$$\mathbb{S} = \{(x,y) \in \mathbb{R}^2;\ x \geq 0\}$$
no es un subespacio vectorial de $\mathbb{R}^2$. La suma de dos elementos de $\mathbb{S}$ siempre está en $\mathbb{S}$. Pero si $(x,y) \in \mathbb{S}$ y $x > 0$, al multiplicarlo por un número negativo obtenemos una pareja que ya no está en $\mathbb{S}$.

5. El subconjunto
$$\mathbb{S} = \{(x,y) \in \mathbb{R}^2;\ xy = 0\}$$
no es un subespacio vectorial de $\mathbb{R}^2$. Al multiplicar un elemento de $\mathbb{S}$ por cualquier escalar volvemos a obtener un elemento de $\mathbb{S}$. Pero la suma no tiene la propiedad de permanecer dentro de $\mathbb{S}$. Por ejemplo
$$(1,0)+(0,1) = (1,1) \notin \mathbb{S},$$
aunque $(1,0)$ y $(0,1)$ son ambos elementos de $\mathbb{S}$. ♣

El próximo ejercicio completa la discusión que precedió a la definición de subespacio, ya que establece el importante resultado de que **un subespacio vectorial es cerrado bajo la operación de formar combinaciones lineales de sus elementos**.

Ejercicio 1.45. Mostrar que si $\mathbb{S}$ es un subespacio vectorial y

$$X_1, X_2, \ldots, X_n$$

están en $\mathbb{S}$, entonces cualquier combinación lineal

$$\lambda_1 X_1 + \lambda_2 X_2 + \ldots + \lambda_n X_n$$

de esos vectores también está en $\mathbb{S}$.

Ahora introduciremos formalmente la noción de *generador de un subespacio*, y mostraremos generadores de los subespacios del ejemplo 1.3.21.

Definición 1.6 (Generadores). *Sea $\mathbb{S}$ un subespacio vectorial $\mathbb{S}$ de $\mathbb{K}^n$. Diremos que un subconjunto $\mathcal{A} \subset \mathbb{S}$ es un **generador** de $\mathbb{S}$ si cualquier vector de $\mathbb{S}$ puede ser expresado como una combinación lineal de elementos de $\mathcal{A}$.*

Ejemplo 1.3.22.

1. Si A es una matriz entonces las columnas de A son un generador de $\text{col}(A)$. Esto es una consecuencia directa de las definición 1.4 del espacio de columnas y de la definición de generador.

2. Llamemos E_i, $i = 1, \ldots, n$ al vector de $\mathbb{K}^n$ que tiene ceros en todas sus posiciones salvo en la i-ésima, en la que hay un 1. Por ejemplo, si $n = 3$ los tres vectores E_i, $i = 1, 2, 3$ son

$$E_1 = (1, 0, 0), \quad E_2 = (0, 1, 0), \quad E_3 = (0, 0, 1).$$

   Entonces la familia $\mathcal{C} = (E_1, E_2, \ldots, E_n)$ es un generador de $\mathbb{K}^n$. En efecto, si $X = (x_1, x_2, \ldots, x_n)$ es un vector de $\mathbb{K}^n$ entonces

$$X = x_1 E_1 + x_2 E_2 + \ldots + x_n E_n$$

   es una combinación lineal de $\mathcal{C}$.

3. El subespacio

$$\mathbb{S} = \left\{(x, y) \in \mathbb{R}^2;\ x = y\right\}$$

   tiene a la familia formada por el vector $(1, 1)$ como un generador. Porque cualquier vector de $\mathbb{S}$ es de la forma (x, x), con $x \in \mathbb{R}$, y puede escribirse como

$$(x, x) = x(1, 1).$$

   Es decir, como una combinación lineal de $\{(1, 1)\}$ con coeficiente x.

Observación 1.3.23. Generadores:una descripción de un subespacio Un generador de un subespacio $\mathbb{S} \subset \mathbb{K}^n$ es un subconjunto de vectores de $\mathbb{S}$ que nos permite escribir cualquier otro vector en $\mathbb{S}$ haciendo combinaciones lineales de los vectores del generador. En otras palabras, si conocemos un generador **podemos reconstruir todo el subespacio** $\mathbb{S}$ a partir del generador. En muchos ejemplos el generador provee una descripción más o menos económica de $\mathbb{S}$: unos pocos de sus elementos permiten reconstruirlo todo.

Ejemplo 1.3.24. El espacio $\mathbb{Z}_2^n$ tiene 2^n elementos. Pero está generado por las n listas E_i que se contruyen poniendo un 1 en el lugar i y ceros en las restantes posiciones, tal como mostramos en la parte 2 del ejemplo 1.3.22. Vale la pena observar que 2^n es un número muchísimo mayor que n. Veamos, por ejemplo, ¿cuántos dígitos tiene la expresión decimal de 2^{100}?. ♣

Ejemplo 1.3.25. El espacio de columnas de

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

está formado por todas las columnas (x_1, x_2) tales que $x_1 = x_2$. Este espacio tiene un generador obvio: el conjunto

$$\{(1,1),(2,2)\}$$

formado por las dos columnas de A. Pero puede generarse a partir de la primera columna, porque la segunda es un múltiplo de la primera.

Estamos diciendo entonces que el conjunto formado por la columna $(1,1)$ es un generador de $\operatorname{col}(A)$, y no tenemos necesidad de incluir a la segunda columna en un generador para poder construir todo el espacio por combinaciones lineales (en realidad ya habíamos encontrado este resultado en la parte 3 del ejemplo 1.3.22).

Hemos conseguido así un generador más económico que el que está formado por todas las columnas. Observemos que en este ejemplo también podríamos haber usado la segunda columna en vez de la primera. ♣ ♠

Tal como muestra el ejemplo 1.3.25 un mismo espacio puede admitir varios generadores distintos. Volvemos a mostrar este fenómeno en nuestro próximo ejemplo.

Ejemplo 1.3.26. El subespacio

$$\mathbb{S} = \{(x,y,0); x,y \in \mathbb{R}\}$$

de $\mathbb{R}^3$ está generado por $\{(1,0,0),(0,1,0)\}$. En efecto, cualquier vector $(x,y,0)$ en $\mathbb{S}$ admite la expresión

$$(x,y,0) = x(1,0,0) + y(0,1,0),$$

como combinación lineal de $(1,0,0)$ y $(0,1,0)$. Pero $\mathbb{S}$ también está generado por $\{(1,1,0),(1,-1,0)\}$, porque

$$(x,y,0) = \frac{x+y}{2}(1,1,0) + \frac{x-y}{2}(1,-1,0),$$

expresión que muestra que todo vector de $\mathbb{S}$ es también una combinación lineal de $(1,1,0)$ y $(1,-1,0)$.

Vale la pena observar que hay infinitos generadores de $\mathbb{S}$. Dejamos planteado el siguiente ejercicio para el lector.

Ejercicio 1.46. Encontrar un nuevo conjunto de vectores que también sea generador de $\mathbb{S}$. ♣

Estamos recién presentando el comienzo de una teoría que nos permitirá entender los sistemas lineales de ecuaciones. Antes de despedirnos de esta sección adelantamos al lector algunos de los pasos que vendrán.

En la sección 1.4 nos ocuparemos del problema de buscar generadores para el espacio de columnas de una matriz, que sean "óptimos" en un sentido que allí discutiremos. En el correr de esa discusión los sistemas homogéneos determinados, que corresponden a familias que sólo permiten escribir el vector nulo del espacio haciendo una combinación con todos los coeficientes nulos, desempeñarán un importante papel. Formalizaremos esta idea en la sección 1.5, dedicada a la *independencia lineal*.

En la sección 1.8 introduciremos una definición general de los subespacios generados haciendo combinaciones lineales, la definición 1.8.1, que engloba al espacio de columnas de una matriz como un caso particular de subespacio generado.

En el capítulo 4, volveremos a tratar, en $\mathbb{K}^n$ y en otros contextos, estas cuestiones relativas a subespacios, generadores y subespacios generados por medio de combinaciones lineales de una familia de vectores. Lo haremos en el marco de una teoría general que extiende a *espacios vectoriales* cualesquiera los conceptos que ahora estamos presentando.

1.3.5. Para tener presente

- Una combinación lineal de algunos vectores de $\mathbb{K}^m$ es el resultado de multiplicarlos por coeficientes en $\mathbb{K}$, y luego sumar.

- El vector nulo siempre puede obtenerse como combinación lineal de cualquier familia de vectores, simplemente escogiendo la combinación trivial que tiene todos los coeficientes iguales a cero.

  Pero algunas familias permiten formar el vector nulo con combinaciones lineales no triviales. Esta situación corresponde a sistemas lineales indeterminados.

- Las combinaciones lineales tienen significado geométrico.

- Formar el producto AX de una matriz por un vector es equivalente a hacer una combinación lineal de las columnas de A con los coeficientes almacenados en X.

- Un sistema $AX = B$ es compatible si y sólo si B es una combinación lineal de las columna de A. Los coeficientes de la combinación forman el vector X, solución del sistema.

- El espacio de columnas de una matriz está formado por todas las posibles combinaciones lineales de sus columnas.

- Un sistema $AX = B$ es compatible si y sólo si B pertenece al espacio de columnas de la matriz del sistema.

- Los subespacios de $\mathbb{K}^m$ son subconjuntos de $\mathbb{K}^m$ dotados de una estructura algebraica: dentro de ellos se puede sumar y multiplicar por un escalar;

- Los subespacios tienen generadores. Familias que permiten obtener todos los vectores del subespacios haciendo combinaciones lineales.

- Un generador es, en general, mucho más chico que el subespacio que genera. Provee una descripción breve del subespacio.

- La compatibilidad de un sistema de ecuaciones queda caracterizada por un subespacio, que a su vez puede describirse por alguno de sus generadores.

1.4. Descripción de los espacios de columnas

En la sección 1.3 dimos un criterio de compatibilidad para un sistema lineal de la forma $AX = B$ en términos del espacio de columnas $\mathrm{col}(A)$ de A: el sistema es compatible si y sólo sí B está en $\mathrm{col}(A)$.

Ahora nos ocuparemos de dar dos descripciones del espacio $\mathrm{col}(A)$. Ambas se obtendrán por medio del algoritmo de eliminación gaussiana, y la estructura de subespacio vectorial que tiene $\mathrm{col}(A)$ aparecerá como un ingrediente fundamental de ellas. Una primera caracterización de $\mathrm{col}(A)$ estará basada en un conjunto de ecuaciones que lo definen. La otra consistirá en buscar un generador óptimo para ese subespacio.

En el curso de esta segunda discusión aparecerán las importantes nociones de *independencia lineal* y de *base de un subespacio vectorial*. Ambos conceptos serán desarrollados en la sección 1.7. Veremos cómo proceder a través del análisis de un caso: la matriz del sistema de ecuaciones lineales que aparece en el ejemplo 1.2.22.

1.4.1. Ecuaciones para el espacio de columnas

Comenzaremos por obtener un conjunto de ecuaciones que deben ser satisfechas por los vectores que están en el espacio de columnas de una matriz. Este conjunto de ecuaciones puede reducirse a una única ecuación, o a ninguna. Para ir haciendo boca ya vamos planteando un ejercicio al lector.

Ejercicio 1.47. ¿En qué casos no habrá ninguna ecuación que deba ser satisfecha por un vector de $\mathbb{K}^m$ para asegurar su pertenencia al espacio de columnas de una matriz A con m filas y n columnas?

Ejemplo 1.4.1. Tal como anunciábamos en la introducción a esta sección, consideremos la matriz real del sistema del ejemplo 1.2.22. Ésta es la matriz

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix}.$$

Nos preguntamos ahora cuáles son los vectores B de $\mathbb{R}^5$ que hacen que el sistema

$$AX = B$$

sea compatible. Para determinarlos consideraremos una columna genérica

$$B = (b_1, b_2, b_3, b_4, b_5)$$

y analizaremos el comportamiento del sistema.

La técnica para hacerlo será el método de eliminación, y para ello escalerizaremos la matriz ampliada $A|B$ del sistema. Esta matriz es

$$\left(\begin{array}{rrrrrr|c} 1 & 2 & 1 & 0 & 1 & 1 & b_1 \\ -1 & -2 & 1 & 1 & 0 & -1 & b_2 \\ 1 & 2 & 1 & 2 & 5 & 3 & b_3 \\ 1 & 2 & 1 & 0 & 1 & 3 & b_4 \\ 1 & 2 & 1 & 4 & 9 & 3 & b_5 \end{array}\right).$$

Ahora procederemos a escalerizarla. Escribiremos la sucesión de matrices que produce el método de escalerización, sin mayores comentarios. Los pasos son los mismos que hicimos en el ejemplo 1.2.22. Al costado de cada matriz representaremos las operaciones que hemos realizado sobre la matriz del paso anterior:

$$\left(\begin{array}{cccccc|c} \mathbf{1} & 2 & 1 & 0 & 1 & 1 & b_1 \\ 0 & 0 & 2 & 1 & 1 & 0 & b_2+b_1 \\ 0 & 0 & 0 & 2 & 4 & 2 & b_3-b_1 \\ 0 & 0 & 0 & 0 & 0 & 2 & b_4-b_1 \\ 0 & 0 & 0 & 4 & 8 & 2 & b_5-b_1 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2+F_1 \\ \longleftarrow F_3-F_1 \\ \longleftarrow F_4-F_1 \\ \longleftarrow F_5-F_1. \end{array}$$

$$\left(\begin{array}{cccccr|c} 1 & 2 & 1 & 0 & 1 & 1 & b_1 \\ 0 & 0 & 2 & 1 & 1 & 0 & b_2+b_1 \\ 0 & 0 & 0 & \mathbf{2} & 4 & 2 & b_3-b_1 \\ 0 & 0 & 0 & 0 & 0 & 2 & b_4-b_1 \\ 0 & 0 & 0 & 0 & 0 & -2 & b_5-2b_3+b_1 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5-2F_3 \end{array}$$

Llegamos a la forma escalerizada,

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & b_1 \\ 0 & 0 & 2 & 1 & 1 & 0 & b_2+b_1 \\ 0 & 0 & 0 & 2 & 4 & 2 & b_3-b_1 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{2} & b_4-b_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_5+b_4-2b_3 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5+F_4, \end{array} \tag{1.39}$$

¡y la respuesta a nuestra pregunta está ante nuestros ojos!

La forma escalerizada muestra que el sistema $AX = B$ es compatible si y sólo sí

$$b_5 + b_4 - 2b_3 = 0. \tag{1.40}$$

Esta ecuación es la que caracteriza al espacio de columnas de A. Si se satisface entonces la última ecuación del sistema se verifica automáticamente, y podemos encontrar valores de las incógnitas que hagan que se satisfagan la cuatro primeras ecuaciones. Concluimos entonces que

$$\mathrm{col}(A) = \{(b_1, b_2, b_3, b_4, b_5);\ b_5 + b_4 - 2b_3 = 0\}.$$

Esta descripción tiene la ventaja de que dado un vector B cualquiera, sólo tenemos que calcular el primer miembro de (1.40) con las entradas de B para saber si B está o no está en el espacio de columnas, y concluir así sobre la compatibilidad o incompatibilidad del sistema $AX = B$.

Por ejemplo, $B_1 = (4, 0, 12, 6, 18)$ está en $\mathrm{col}(A)$, porque

$$18 + 6 - 2 \times 12 = 0.$$

Si el lector quiere hacer una comprobación adicional puede observar que este vector es la suma de todas las columnas de A excepto la segunda. El sistema

$$AX = B_1$$

es compatible, y

$$X = (1, 0, 1, 1, 1, 1)$$

es una de sus soluciones. Sin embargo

$$B_2 = (1, 1, 0, 1, 1)$$

no está en el espacio de columnas de A, porque al sumar sus entradas cuarta y quinta y restar dos veces la tercera resulta

$$1 + 1 - 2 \times 0 = 2 \neq 0.$$

Naturalmente

$$AX = B_2$$

es un sistema incompatible. ♣

Notemos que lo que hicimos ilustra un método general para hallar ecuaciones que caractericen a la imagen de una matriz: intentamos resolver un sistema genérico $AX = B$, y la condición de compatibilidad aparecerá expresada como una condición sobre los coeficientes de B.

Ejercicio 1.48. Sea la matriz

$$A = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 3 & 2 \\ -1 & 2 & 7 & 5 \\ 2 & -1 & -5 & -4 \end{pmatrix}.$$

1. Determinar si los siguientes vectores de $\mathbb{R}^4$ pertencen al espacio de columnas de A:

   a) $(1,-1,-3,3)$;

   b) $(0,1,0,1)$;

   c) $(0,x,x,-x)$, $x\in\mathbb{R}$;

   d) $(x,0,-x,2x)$, $x\in\mathbb{R}$

2. Hallar ecuaciones que caractericen al espacio de columnas.

Ejercicio 1.49. Hallar un conjunto de ecuaciones que caractericen el espacio de columnas de la matriz real

$$A=\begin{pmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&4\end{pmatrix}.$$

Ejercicio 1.50. Repetir el ejercicio anterior para la matriz G con entradas en $\mathbb{Z}_2$ que se encontró al resolver el ejercicio 1.41, página 91. Expresar estas ecuaciones matricialmente, en la forma $HY=O$, donde H es una matriz que se determinará.

Observación 1.4.2. Si sólo queremos decidir acerca de la compatibilidad o incompatibilidad del sistema para un B dado, sin tratar de entender la estructura del conjunto de los B que hacen compatible el sistema $AX=B$ podemos resumir la discusión en la siguiente observación: el sistema es compatible si y sólo si las formas escalerizadas de A y de la matriz ampliada $A|B$ tienen el mismo número de pivotes (o de escalones, para expresarlo en un lenguaje más gráfico). ♠

Una base para la imagen de A.

Vamos a ver ahora otra caracterización de la imagen de la matriz A del ejemplo 1.4.1. El significado del término “base”será aclarado a lo largo de la discusión, y en la sección 1.5.

La forma escalerizada (1.39) pone en evidencia cuáles son los vectores B que hacen compatible el sistema, y también que para esos vectores B la solución es indeterminada porque hay variables libres, en este caso dos. Sabemos además que podemos elegir como variables libres las que corresponden a columnas que no tienen pivotes: las variables x_2 y x_5. Esto implica que podemos fijar arbitrariamente x_2 y x_5, y luego calcular las restantes variables para obtener una solución del sistema.

Por lo tanto cada columna B en la imagen de A puede ser escrita como combinación lineal de las columnas de A de muchas maneras. En particular, podemos elegir

$$x_2 = x_5 = 0,$$

y conseguir una solución que, en realidad, quedará determinada una vez que hemos fijado estos valores para x_2 y x_5.

Hagamos una observación sencilla pero fundamental: **fijar $x_2 = x_5 = 0$ es lo mismo que eliminar estas variables del sistema**, y también **es lo mismo que eliminar la segunda y quinta columna de la matriz del sistema**. Por lo tanto, eliminar la segunda y quinta columna de la matriz del sistema no altera, en este caso, el conjunto de vectores B para los que el sistema lineal es compatible. Es decir, no se altera el espacio de columnas con esta operación "quirúrgica"sobre A.

Para continuar nuestra discusión escribiremos la matriz A como

$$A = (A_1\ A_2\ A_3\ A_4\ A_5\ A_6),$$

donde A_i, $i = 1, \ldots, n$, indican las columnas de A.

La existencia de variables libres tiene que ver con la las columnas de A tienen cierta estructura de dependencia entre ellas: una vez que tenemos las columnas

$$A_1,\ A_3,\ A_4,\ A_6,$$

las columnas A_2 y A_5 son, en algún sentido, redundantes, ya que **cualquier vector que esté en el espacio de columnas de A puede ser expresado como una combinación lineal de las columnas A_1, A_3, A_4, A_6.** Al eliminar A_2 y A_5 no perdemos absolutamente nada del espacio de columnas. Dicho en otras palabras, sabemos, es un hecho obvio, que el conjunto

$$\{A_1, A_2, A_3, A_4, A_5, A_6\}$$

formado por todas las columnas de A es un generador de $\mathrm{col}(A)$. Pero también

$$\{A_1, A_3, A_4, A_6\}$$

es un generador de $\mathrm{col}(A)$. Es un generador más pequeño que el que está formado por todas las columnas de la matriz. Por eso ofrece una descripción más económica del mismo conjunto.

Observemos que col(A) coincide con el espacio de columnas de la matriz

$$\bar{A} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 2 & 3 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 4 & 3 \end{pmatrix},$$

que hemos construido tachando las columnas 2 y 5 de A. Para escalerizar $\bar{A}$ deberíamos repetir los mismos pasos que usamos para A y obtendríamos

$$\bar{E} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

una matriz sin variables libres, porque cada columna tiene pivotes. Digamos que no hemos repetido todas las cuentas. Sólo nos limitamos a tachar las columnas 2 y 5 en la forma escalerizada, porque la escalerización trata cada columna por separado.

La forma escalerizada $\bar{E}$ pone en evidencia que cuando un sistema lineal con matriz $\bar{A}$ tiene solución, entonces la solución es única. Al eliminar las columnas redundantes, y con ellas las variables libres, hemos perdido la libertad que teníamos para fabricar soluciones del sistema. En resumen, el sistema

$$\bar{A}X = B$$

es compatible para los mismos vectores B que hacen compatible

$$AX = B,$$

y tiene la propiedad adicional de que cuando la solución existe entonces es única. Es decir, $\bar{A}X = B$ es determinado siempre que es compatible.

Podemos expresar esto mismo en el lenguaje de combinaciones lineales, generadores y subespacios: cuando un vector B está en el espacio de columnas de la matriz A (o el de $\bar{A}$, ya que los espacios de columnas coinciden) **hay una única manera de escribirlo como combinación lineal de las columnas A_1, A_3, A_4 y A_6.** Por lo tanto

1. (A_1, A_3, A_4, A_6) es un **generador** de col(A);

2. este generador tiene la propiedad adicional de que cualquier vector de $\text{col}(A)$ admite una *única* expresión como combinación lineal de los elementos en el generador. En particular, **la única combinación lineal de** (A_1, A_3, A_4, A_6) **que es igual al vector nulo** $O \in \mathbb{R}^5$ **es la que tiene todos los coeficientes nulos**.

La segunda propiedad que acabamos de enunciar es lo que llamaremos *indepedencia lineal* de la familia (A_1, A_3, A_4, A_6). Un generador de un subespacio vectorial que es además *linealmente independiente*, en el sentido que acabamos de introducir, es lo que llamaremos una *base* del subespacio. Ambos conceptos son muy importantes, y serán el objeto central de discusión en la sección 1.5 (ver las definiciones 1.7, en la página 112, y 1.8, página 117).

Una base de un subespacio puede ser vista como un generador sin redundancia alguna. Por lo tanto, nos da una forma óptima de describir el subespacio. En nuestro ejemplo hemos conseguido una base eliminando algunos vectores de un generador más grande (el que estaba formado por todas las columnas de la matriz). La siguiente observación muestra la optimalidad de nuestra elección: si seguimos quitando columnas entonces perdemos la capacidad de generar todo $\text{col}(A)$. Los detalles se discuten a continuación.

Observación 1.4.3. El subconjunto de columnas

$$\mathcal{A} = \{A_1, A_3, A_5, A_6\}$$

tiene la siguiente propiedad: ninguna de sus cuatro columnas puede obtenerse como una combinación lineal de las demás. Es decir, si eliminamos una de sus columnas ya no podremos recuperarla a partir de las demás haciendo combinaciones lineales. Vamos a justificar esta afirmación.

Por ejemplo, la columna A_1 no puede expresarse como combinación lineal de

$$\{A_3, A_5, A_6\}.$$

Veámoslo. Sabemos que cualquier columna en el espacio de columnas, o imagen, de A puede ser escrita como una combinación lineal de $\mathcal{A}$ de manera única. Esto es cierto, en particular, para A_1, que admite la expresión obvia

$$A_1 = A_1 + 0A_3 + 0A_5 + 0A_6.$$

Si pudiéramos escribir además

$$A_1 = aA_3 + bA_5 + cA_6 \tag{1.41}$$

para alguna terna de coeficientes a, b y c, obtendríamos inmediatamente la expresión

$$A_1 = 0A_1 + aA_3 + bA_5 + cA_6,$$

que es otra forma de escribir A_1 como combinación lineal de $\mathcal{A}$. Esto es contradictorio, y muestra la imposibilidad de (1.41).

El mismo razonamiento puede hacerse para cualquiera de las otras columnas, e implica que si eliminamos más columnas de A resulta una nueva matriz con un espacio de columnas estrictamente más chico que el de A, porque la columna eliminada no está en el espacio que las restantes generan[10].

Notemos que, cuando hay variables libres, hemos mejorado nuestra descripción del espacio de columnas de A por medio de un generador: al menos somos capaces de darla usando menos columnas que todas las que estaban en la matriz original. Y no podemos mejorar más, porque sabemos que si sacamos más columnas "perdemos un pedazo del espacio de columnas".

Podemos eliminar A_2 y A_5 sin afectar al espacio de columnas de A, porque cualquiera de estas dos columnas puede ser expresada como una combinación lineal de

$$\mathcal{A} = \{A_1, A_3, A_4, A_6\}.$$

Ejercicio 1.51. Expresar A_2 y A_5 como combinación lineal de $\mathcal{A}$. Sugerencia: no hace falta plantear un nuevo sistema de ecuaciones y resolverlo. Para despejar A_2, por ejemplo, es suficiente construir una solución de $AX = O$ con $x_2 = -1$, $x_5 = 0$.

Los resultados discutidas en el correr de esta observación son generales. Ver, al respecto, el ejercicio 1.60, en la página 117. ♠

Ejercicio 1.52. Para cada una de las matrices A eliminar algunas de sus columnas para hallar una matriz $\bar{A}$ tal que

$$\mathrm{col}(A) = \mathrm{col}(\bar{A})$$

y que el sistema $\bar{A}X = B$ tenga solución única para todos los B que lo hacen compatible. En cada caso expresar las columnas eliminadas como una combinación lineal de las columnas de $\bar{A}$, y verificar que ninguna de las columnas de $\bar{A}$ puede escribirse como combinación lineal de las restantes.

1. $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}$.

[10]En realidad perdemos mucho más que la columna eliminada: ninguna de las combinaciones lineales en las que la columna que hemos eliminado aparece con un coeficiente distinto de cero puede escribirse sólo en términos de las columnas que hemos dejado

2. $A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix}$.

3. $A = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$. Trabajar en $\mathbb{Z}_2$.

Ejercicio 1.53.

1. Para la matriz real A y los vectores B_1, B_2 dados por
$$A = \begin{pmatrix} 1 & 1 & -2 & 1 \\ 2 & 1 & -1 & 1 \\ 2 & 2 & -4 & 2 \\ 3 & 2 & -3 & 2 \end{pmatrix}, \quad B_1 = \begin{pmatrix} -1 \\ 2 \\ -2 \\ 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 0 \end{pmatrix}$$
resolver los sistemas $AX = B_i$, $i = 1, 2$. Expresar de dos maneras diferentes B_1 como una combinación lineal de las columnas de A
2. Hallar todos los vectores B que hacen que el sistema $AX = B$ sea compatible.
3. Formar una nueva matriz $\bar{A}$ eliminando algunas columnas de A, de modo que:
   - $\bar{A}X = B$ sea compatible para todos los vectores B hallados en la parte 2, y sólo para esos B;
   - la solución de $\bar{A}X = B$ siempre sea única.
4. Expresar B_1 como combinación lineal de las columnas de $\bar{A}$. Expresar el vector columna $O = (0, 0, 0, 0)$ como combinación lineal de las columnas de A y como combinación lineal de las columnas de $\bar{A}$.
5. Repetir las partes 2 y 3 para las matrices
$$C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & -1 \\ 1 & 1 & 3 & 3 & -3 \\ 0 & 0 & 1 & 1 & -2 \\ 2 & 2 & 4 & 4 & -2 \\ -4 & -4 & -6 & -6 & 0 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$
Para D trabajar sobre el cuerpo $\mathbb{Z}_2$.

Ejercicio 1.54. Para la matriz real
$$\begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix}$$

y la matriz

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

con entradas en $\mathbb{Z}_2$ construir, eliminando algunas de sus columnas, dos posibles matrices que tengan las propiedades que se piden en el ejercicio 1.52 para la submatriz de A que hemos llamado $\bar{A}$.

1.4.2. Para tener presente

- Un sistema $AX = B$ es compatible si y sólo si las formas escalerizadas de A y $A|B$ tienen el mismo número de pivotes.

- Si el procedimiento de escalerización de $A|B$ se lleva a cabo con un vector B genérico, el proceso de escalerización da lugar a un conjunto de ecuaciones que se obtienen a partir de las filas de la forma escalerizada de A que no tienen pivotes, y que caracterizan completamente el subespacio col$A$;

- Si eliminamos de una matriz las columnas que corresponden a variables libres de una forma escalerizada (columnas sin pivotes) se obtiene una nueva matriz con el mismo espacio de columnas que la original.

- Si en la forma escalerizada de una matriz A no aparecen variables libres, entonces cualquier vector B en col(A) admite una única expresión como combinación lineal de las columnas de A.

- Sea E una forma escalerizada de una matriz A. Cualquier vector col(A) admite una única expresión como combinación lineal de las columnas de A cuyas posiciones corresponden a columnas de E que no tienen pivotes.

1.5. Independencia lineal y bases

Cerramos la sección 1.4 identificando un generador particular del espacio de columnas de una matriz A, que tenía la propiedad de que cualquier vector en el espacio de columnas admitía una expresión única como combinación lineal de las columnas en ese generador. En un lenguaje más cercano a los sistemas de ecuaciones podemos describir lo que hicimos diciendo que eliminamos algunas columnas de la matriz A de un sistema $AX = B$ para construir una matriz $\bar{A}$ tal que el sistema $\bar{A}X = B$ es compatible si sólo si $AX = B$ lo es, y que tiene la propiedad adicional de que es determinado cuando es compatible.

En el lenguaje de subespacios vectoriales de la sección 1.4 diremos que

$$\mathrm{col}(A) = \mathrm{col}(\bar{A}),$$

y que el sistema $\bar{A}X = B$ es determinado para cualquier columna $B \in \mathrm{col}(A)$. Esto es equivalente a decir que cualquier vector en $\mathrm{col}(A)$ puede escribirse de una única manera como una combinación lineal de las columnas de $\bar{A}$.

Esta cuestión de *unicidad* de las soluciones para un sistema de ecuaciones lineales, o, equivalentemente, en la descomposición de un vector como combinación lineal de otros, nos llevará a introducir otra de las nociones centrales del álgebra lineal: la noción de *independencia lineal*.

1.5.1. La noción de independencia lineal

Comenzamos con una proposición que asegura que el problema de unicidad de soluciones de un sistema $AX = B$ puede reducirse a la unicidad de soluciones de $AX = O$, donde O indica la columna nula, enteramente formada por ceros.

Proposición 1.8. *El sistema lineal $AX = B$ tiene una única solución para todo $B \in \mathrm{col}(A)$ si y sólo si $AX = O$ tiene como única solución la solución trivial $X = O$.*

Prueba: Recordemos que O está en el espacio de columnas de cualquier matriz, o, equivalentemente, un sistema de la forma $AX = O$ siempre es compatible, porque tiene al menos la solución trivial $X = O$.

Si la solución de $AX = B$ es única para cualquier $B \in \mathrm{col}(A)$ entonces, particularizando esta propiedad para $B = O$, el sistema $AX = O$ sólo puede tener la solución trivial.

Supongamos ahora que $AX = O$ sólo tiene la solución trivial $X = O$, y consideremos dos soluciones X_1 y X_2 de $AX = B$. Se satisfacen entonces las igualdades $AX_1 = B$ y $AX_2 = B$. Restándolas resulta

$$AX_1 - AX_2 = B - B = 0.$$

Como $AX_1 - AX_2 = A(X_1 - X_2)$ concluimos que

$$A(X_1 - X_2) = O,$$

y $X_1 - X_2$ es una solución de $AX = 0$. Entonces $X_1 - X_2 = O$, lo que es equivalente a $X_1 = X_2$. Por lo tanto, el sistema $AX = B$ sólo puede tener una solución. $\square$

Observación 1.5.1. La idea central de la proposición 1.8 y del argumento que la demuestra es que podemos "trasladar" todo el problema al O sumando (o restando) un vector adecuado. Esta idea aparecerá también al describir el conjunto solución de un sistema $AX = B$: todo se reducirá a comprender las soluciones de $AX = O$, y se pasará de un conjunto solución a otro sumando (o restando) una solución particular de $AX = B$.

Naturalmente, la condición de que $AX = O$ tenga solución única puede reformularse en término de combinaciones lineales. Esta caracterización con combinaciones lineales es la que tomaremos como definición de *independencia lineal.* La razón es que luego generalizaremos las nociones de combinación lineal e independencia lineal al contexto más general de los espacios vectoriales cualesquiera, donde ya no escribiremos las combinaciones lineales como el producto de una matriz por un vector.

Definición 1.7 (Independencia lineal). *Diremos que una familia ordenada*

$$(A_1, A_2, \ldots, A_l)$$

de vectores en $\mathbb{K}^m$ es ***linealmente independiente*** *si la única manera de expresar el vector nulo como combinación lineal de la familia es escogiendo todos los coeficientes de la combinación iguales a* 0. *Cuando una familia no es linealmente independiente diremos que es* ***linealmente dependiente****.*

Por supuesto, es válida la siguiente proposición:

Proposición 1.9. *El sistema $AX = O$ es compatible y determinado si y sólo si las columnas de A forman una familia linealmente independiente. El sistema $AX = O$ es compatible e indeterminado si y sólo si las columnas de A forma una familia linealmente dependiente.*

Ejercicio 1.55. Revisar la discusión que precede a la definición 1.7 y completar los detalles de la proposición que acabamos de enunciar.

Ejercicio 1.56. Mostrar que una familia linealmente independiente de vectores de $\mathbb{K}^m$ no puede tener más de m elementos. Sugerencia: considerar que los vectores de la familia como las columnas de la matriz de un sistema homogéneo.

Ejemplo 1.5.2. Volviendo sobre la matriz real

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix},$$

con la que trabajamos a lo largo de la sección 1.4, digamos que las columnas de A forman una familia linealmente dependiente. Pero la familia

$$(A_1, A_3, A_5, A_6)$$

formada por la primera, la tercera, la quinta y la sexta columna de A, es linealmente independiente.

Ejemplo 1.5.3.

- Una familia que contenga al vector nulo es linealmente dependiente. Es muy fácil hacer una combinación lineal de tal familia que tenga algún coeficiente distinto de cero pero que arroje como resultado el vector nulo: sólo hay que multiplicar el vector nulo por cualquier coeficiente no nulo (por ejemplo, por 1), todos los demás vectores por cero y luego sumar.
- Una familia que tenga algún vector repetido es linealmente dependiente. Supongamos, que el vector repetido ocupa los lugares i y j, con $i \neq j$. Nuevamente fabricaremos una combinación lineal de la familia con algunos coeficientes distintos de cero y cuyo resultado es el vector nulo: escogemos todos los coeficientes iguales a cero salvo los de los vectores i-ésimo y j-ésimo, que serán iguales a 1 y -1 respectivamente. ♣

Observación 1.5.4. Otra manera de expresar la condición de que una familia

$$\mathcal{A} = (A_1, A_2, \dots, A_l)$$

sea linealmente independiente es decir que la igualdad

$$x_1 A_1 + x_2 A_2 + \ldots + x_l A_l = O \tag{1.42}$$

implica que

$$x_1 = x_2 = \ldots = x_l = 0. \tag{1.43}$$

Esto es así porque el término de la izquierda en (1.42) es una combinación lineal de los vectores de la familia que resulta ser igual al vector nulo O. La definición de independencia lineal requiere que los coeficientes en esta combinación sean todos nulos, tal como se refleja en (1.43). ♠

En la observación 1.4.3, sobre el final de la sección 1.4 aparecía una propiedad que caracteriza a la independencia lineal: si una familia es linealmente independiente ningún vector puede expresarse como combinación lineal de los restantes. Consideremos una familia $\mathcal{A} = (A_1, A_2, \ldots, A_l)$. Si algúno de sus elementos, por el ejemplo el A_{i_0} puede escribirse como combinación lineal de los restantes existe una expresión del tipo

$$A_{i_0} = x_1A_1 + \ldots + x_{i_0-1}A_{i_0-1} + x_{i_0+1}A_{i_0+1} + x_lA_l,$$

por lo tanto

$$O = x_1A_1 + \ldots + x_{i_0-1}A_{i_0-1} + (-1)A_{i_0} + x_{i_0+1}A_{i_0+1} + x_lA_l,$$

y el miembro de la derecha es una combinación lineal de los elementos de $\mathcal{A}$ que es igual al vector nulo y en la que al menos unos de sus coeficientes es distinto de cero. Por lo tanto $\mathcal{A}$ es linealmente dependiente.

Es cierto también el recíproco: si $\mathcal{A}$ es linealmente dependiente alguno de los vectores en $\mathcal{A}$ puede expresarse como una combinación lineal de los restantes. Veamos por qué esta afirmación es cierta. Al ser $\mathcal{A}$ linealmente dependiente existe una combinación lineal de los elementos de la familia $\mathcal{A}$ que es igual a cero y que tiene algún coeficiente distinto de cero. Supongamos que el coeficiente x_{i_0} del vector A_{i_0} es no nulo, y escribamos

$$O = x_1A_1 + \ldots + x_{i_0-1}A_{i_0-1} + x_{i_0}A_{i_0} + x_{i_0+1}A_{i_0+1} + x_lA_l.$$

De esta fórmula podemos despejar el vector A_{i_0} como

$$A_{i_0} = \lambda_1A_1 + \ldots + \lambda_{i_0-1}A_{i_0-1} + \lambda_{i_0+1}A_{i_0+1} + \lambda_lA_l$$

donde

$$\lambda_i = x_i/x_{i_0}, \quad i = 1, \ldots, l, \ i \neq i_0.$$

Por lo tanto hay uno de los vectores de la familia que puede expresarse como combinación lineal de los restantes.

Resumamos los resultados de esta discusión en la forma de una proposición.

Proposición 1.10. *Una familia $\mathcal{A} = (A_1, A_2, \ldots, A_l)$ de vectores de $\mathbb{K}^m$ es linealmente independiente si y sólo si ninguno de los vectores de $\mathcal{A}$ puede expresarse como combinación lineal de los restantes. Equivalentemente, $\mathcal{A}$ es linealmente dependiente si y sólo si alguno de los vectores en $\mathcal{A}$ puede escribirse como una combinación lineal de los restantes.*

Observación 1.5.5. En el contexto de los sistemas lineales es conveniente recordar que la independencia lineal de la familia formada por las columnas de la matriz de un sistema es equivalente a que el sistema sea determinado siempre que es compatible.

Otra forma interesante de pensar acerca de la independencia lineal es la que nos ofrece la proposición 1.10: una familia linealmente independiente es aquella en la que ninguno de sus vectores puede obtenerse a partir de los demás haciendo combinaciones lineales.

Estas nociones son, quizás, mas naturales que la definición de independencia lineal. Pero la definición, tal como está formulada, es más flexible y fácil de manejar que las caracterizaciones que estamos enfatizando. Por eso es preferible como definición, aunque sea menos evidente.

Vale la pena mencionar que en la historia de las teorías los postulados y las definiciones básicas aparecen tarde, cuando ya se sabe bien cómo funciona todo. Ideas importantes pueden ser recogidas en una definición de una forma un tanto oscura, pero más eficiente para el desarrollo lógico de la teoría, que casi nunca es fiel a su desarrollo histórico. La definición de independencia lineal es un ejemplo de esto. ¿Puede el lector identificar en la Matemática que conoce otras definiciones, postulados, o enunciados para los que sean válidos consideraciones similares? ♠

Ejercicio 1.57. Estudiar la dependencia o independencia lineal de las siguientes familias de vectores en $\mathbb{R}^n$. Para las que son linealmente dependientes lineal estudiar qué vectores se pueden expresar como combinación lineal de los restantes.

1. $((1,3),(-2,-6))$;
2. $((1,3),(1,0))$;
3. $((2,-1),(3,4),(2,-3))$;
4. $((1,2,1),(0,1,1),(1,1,1),(2,4,2))$;
5. $((1,-1,1),(-1,1,3),(-1,1,1),(1,1,-1))$;
6. $((1,1,0,0),(0,0,2,2),(3,2,1,2),(1,2,1,2),(1,-1,0,0),(3,2,1,2))$;
7. $((1,2,3,4),(4,3,2,1),(1,1,1,1),(0,0,0,0))$;
8. $\big((1,2,3,4),(4,3,2,1),(1,1,1,1),(-17,25,44,33),(e,\pi,\sqrt{2},\log 8)\big)$;

9. $((1,1,0,0),(0,0,2,2),(1,2,1,2),(1,-1,0,0),(3,2,1,2))$;

10. $((3,-1,2,-1),(1,2,5,2),(1,-3,-8,-5))$

11. $\big((\lambda,\lambda^2,1),(-1,\lambda,\lambda),(0,2\lambda^2,\lambda^2+1)\big)$. Discutir según λ

Ejercicio 1.58. Estudiar la dependencia o independencia lineal de las siguientes familias de vectores en $\mathbb{Z}_2^n$. Para las que son linealmente dependientes lineal estudiar qué vectores se pueden expresar como combinación lineal de los restantes.

1. $((1,1,0),(0,1,1),(1,1,1))$

2. $((1,1,0),(1,0,1),(0,1,1))$

3. $((1,1),(0,1))$

4. $((1,1),(0,0))$

5. $((0,0,1,0),(0,1,0,1),(0,1,1,1))$

Ejercicio 1.59. Sea $A=(X_1,X_2,X_3)$ un conjunto linealmente independiente. Investigar si también lo es $B=(Y_1,Y_2,Y_3)$, con $Y_1=X_1+X_2$, $Y_2=X_1+X_3$ y $Y_3=X_2+X_3$

Observación 1.5.6. TERMINOLOGÍA Y NIVELES DEL LENGUAJE
Cerramos esta sección con un comentario algunas expresiones relacionadas con la independencia lineal. Subrayemos que **la independencia lineal es una propiedad que sólo puede tener una familia de vectores**. Los vectores por sí solos no son linealmente independientes ni linealmente dependientes. Sólo tiene sentido intentar determinar la independencia o dependencia lineal de una familia de vectores. Sin embargo, es corriente decir, sobre todo en el lenguaje hablado, “los vectores tales y cuales son linealmente independientes (dependientes)”. Se trata de una expresión un tanto informal, que abrevia la expresión formalmente correcta “la familia formada por los vectores tales y cuales son linealmente independientes (dependientes)”. No utilizaremos en el texto esta manera abreviada de expresarnos, pero es casi inevitable recurrir a ella al hablar. ♣

1.5.2. Bases

Ahora que hemos introducido la noción de independencia lineal podemos completar la definición de base de un subespacio vectorial que adelantamos en la sección 1.3, cuando destacamos el papel de los generadores que permiten expresar de una única manera sus combinaciones lineales.

Definición 1.8 (Bases). *Si $\mathbb{S}$ es un subespacio vectorial de $\mathbb{K}^n$, diremos que un subconjunto ordenado*

$$(A_1, A_2, \dots, A_l)$$

de vectores de $\mathbb{S}$ es una ***base*** *de $\mathbb{S}$ si*

1. *es un generador de $\mathbb{S}$;*

2. *es linealmente independiente.*

Esta definición cierra el círculo de ideas que habíamos abierto al tratar la compatibilidad de los sistemas

Retomemos la discusión acerca del sistema $AX = B$, con matriz

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix}, \tag{1.44}$$

en el punto en que la dejamos en la página 106: con la definición que acabamos de introducir diremos que la familia

$$(A_1, A_3, A_5, A_6)$$

formada por la primera, tercera, quinta y sexta columnas de la matriz A **es una base del espacio de columnas de** A.

Los argumentos que presentamos al tratar ese ejemplo son en realidad generales, y permiten concluir en general, para una matriz A cualquiera, que las columnas que corresponden a las posiciones de los pivotes de una forma escalerizada de una matriz A forman una base del espacio de columnas de A. En este sentido, presentamos un ejercicio que muestra que las conclusiones obtenidas en la observación 1.4.3, a partir de la página 107 de la sección 1.4, son generalizables generadores y bases cualesquiera.

Ejercicio 1.60. Consideremos una familia

$$\mathcal{A} = (A_1, \dots, A_n)$$

que es un generador de un subespacio vectorial $\mathbb{S}$. Mostrar que $\mathcal{A}$ es una base de $\mathbb{S}$ si y sólo si tiene la siguiente propiedad: al quitar de la familia $\mathcal{A}$ algún vector necesariamente se obtiene una familia que no puede ser un generador de $\mathbb{S}$.

Tenemos entonces a nuestra disposición un algoritmo para hallar bases del espacio de columnas de una matriz. Lo pondremos a prueba en nuestro próximo ejercicio.

Ejercicio 1.61. Hallar una base del espacio de columnas para cada una de las matrices A dadas a continuación. Para cada una de ellas se da además un vector B. Resolver en cada caso el sistema $AX = B$. Si el sistema es compatible expresar B como la única combinación lineal de la base hallada. Sugerencia: utilizar los resultados obtenidos anteriormente, no es necesario volver a escalerizar los sistemas.

1. Hacerlo para la matriz y el vector reales

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \\ -1 \end{pmatrix}.$$

La matriz es la del ejemplo 1.2.22, página 45, con la que hemos trabajado largamente.

2. Repetir para la matriz A del ejemplo 1.2.25 en la página 48, y el vector B en $\mathbb{Z}_2$, con

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

3. Lo mismo para la matriz A del ejercicio 1.27 en la página 59, y el vector B que la acompaña:

$$A = \begin{pmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 3 & 0 & 1 & -1 \\ -1 & 0 & 2 & -1 & 0 \\ -1 & -1 & 1 & 4 & 1 \\ 0 & -1 & 0 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

Ejercicio 1.62. Sea

$$A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ -1 & 1 & 3 & -4 \\ 2 & 1 & 0 & -1 \end{pmatrix}.$$

Hallar todos los subconjuntos de columnas de la matriz A que son una base del espacio de columnas de A.

Resumimos nuestros resultados en la próxima proposición.

Proposición 1.11. *Sea A una matriz $m \times n$ sobre un cuerpo $\mathbb{K}$ que tiene alguna entrada distinta de 0, y sea E una forma escalerizada de A. Entonces las columnas de la matriz A que están en las posiciones en que la matriz E tiene pivotes forman una base del espacio de columnas de A.*

Prueba: Llamemos $\bar{A}$ y $\bar{E}$ a las matrices que se obtienen eliminando de A y de E las columnas que corresponden a variables libres. Entonces $\bar{E}$ es una forma escalerizada de $\bar{A}$. Un sistema $AX = B$ es compatible si y sólo si $\bar{A}X = B$ es compatible, lo que muestra que $\mathrm{col}(A) = \mathrm{col}(\bar{A})$. Naturalmente, las columnas de $\bar{A}$ generan $\mathrm{col}(\bar{A})$, por lo que generan $\mathrm{col}(A)$.

Como hemos eliminado las columnas que corresponden a variables libres el sistema $\bar{A}X = O$ sólo tiene la solución trivial $X = O$. Entonces las columnas de $\bar{A}$ son linealmente independientes.

En conclusión, las columnas de $\bar{A}$ forman una base de $\mathrm{col}(A)$. □

La noción de base no sólo es interesante en el contexto de los sistemas lineales y los espacios de columnas de las matrices de los sistemas. A continuación vemos algunos otros ejemplos. También trabajaremos ampliamente con bases en el capítulo dedicado a los espacios vectoriales.

Ejemplo 1.5.7. La base canónica de $\mathbb{K}^n$
Para $i = 1, \ldots, n$ llamemos E_i al vector de $\mathbb{K}^n$ que tiene un 1 en la i-ésima posición, y ceros en las restantes. Es decir

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad E_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Cualquier otro vector

$$X = (x_1, x_2, \ldots, x_n)$$

puede expresarse como la combinación lineal

$$X = x_1E_1 + x_2E_2 + \ldots x_nE_n$$

de la familia

$$\mathcal{C} = (E_1, E_2, \ldots E_n),$$

y para cualquier $X \in \mathbb{K}^n$ esta es la única manera de expresar X como combinación lineal de $\mathcal{C}$. Por lo tanto $\mathcal{C}$ es linealmente independiente y es una base de $\mathbb{K}^n$. Es usual llamar a $\mathcal{C}$ la **base canónica de** $\mathbb{K}^n$. ♣

Ejemplo 1.5.8. No hay bases para el subespacio trivial
Llamamos subespacio trivial al subespacio

$$\mathbb{S} = \{O\},$$

que sólo contiene al vector nulo de $\mathbb{K}^n$. Sólo hay dos subconjuntos posibles de $\mathbb{S}$: el vacío y todo $\mathbb{S}$. El vacío no permite generar nada. Y el único subconjunto no vacío, $\mathbb{S}$, es linealmente dependiente: podemos escribir el vector O como

$$O = aO,$$

donde a es cualquier escalar del cuerpo $\mathbb{K}$. En particular, se puede escoger $a \neq 0$. Por lo tanto no hay bases para este subespacio. ♣

La cantidad de vectores que tiene una base de un subespacio mide el tamaño que, como estructura lineal, tiene el subespacio. Es lo que llamaremos la *dimensión* del subespacio. Por ejemplo, $\mathbb{K}^n$ tiene dimensión n. Dedicaremos la sección 1.6 a la formalización de este importante concepto.

1.5.3. Para tener presente

- Un sistema $AX = B$ compatible es determinado si y sólo si el sistema homogéneo $AX = O$ lo es.
- Decimos que una familia de vectores es linealmente independiente si la única manera de expresar el vector nulo como combinación lineal de la familia es escoger todos los coeficientes de la combinación lineal iguales a cero.
- Una familia es linealmente dependiente si no es linealmente independiente.
- Las columnas de una matriz A forman una familia linealmente independiente si y sólo si el sistema homogéneo $AX = O$ es determinado.
- Una familia es linealmente independiente si y sólo si es imposible expresar cualquier vector de la familia como combinación lineal de los restantes.
- Llamaremos base de un subespacio vectorial a un generador del subespacio que es además linealmente independiente.
- Las columnas de una matriz A que corresponden a las posiciones de los pivotes de una forma escalerizada de A forman una base del espacio de columnas de A.

1.6. Dimensión de un subespacio vectorial

La cantidad de vectores que tiene una base de un subespacio mide el tamaño que, como estructura lineal, tiene el subespacio. También puede verse este número como una medida de cuanta información es necesaria para especificar el subespacio por medio de un generador, o de cuantos parámetros independientes son necesarios para recorrer todos los elementos del subespacio. Llamaremos *dimensión* del subespacio a este número. En esta sección haremos rigurosa esta noción de dimensión, y comenzaremos a mostrar su relación con la teoría de los sistemas de ecuaciones lineales.

La noción de dimensión es central dentro del Álgebra Lineal y la Geometría. Ya ha aparecido implícitamente en varias situaciones, y reaparecerá una y otra vez. Veamos algunos ejemplos.

Hemos trabajado escalerizando matrices. En algunos casos encontramos para una misma matriz A más de una posible forma escalerizada. Sin embargo, la cantidad de pivotes, o la cantidad de filas no nulas, que tiene cualquier forma escalerizada de A es siempre la misma. Por ejemplo, podemos apreciar esto para la matriz 5×6

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix} \tag{1.45}$$

del sistema de ecuaciones del ejemplo 1.2.22, página 45. Hemos trabajado mucho con ella, y encontramos al menos dos formas escalerizadas. Por ejmplo, una de ellas aparece sobre el final de ese mismo ejemplo, en la página 46. Otra es la forma escalerizada reducida que fue calculada en el ejemplo 1.2.30, página 61. Ambas tienen cuatro pivotes.

Este fenómeno está asociado con la dimensión de dos subespacios asociados a la matriz: el subespacio de columnas, formado por todas las posibles combinaciones lineales de las columnas; y el subespacio formado por todas las posibles combinaciones lineales de las filas, que será introducido y estudiado en la sección 1.8

Un sistema lineal de ecuaciones puede tener muchas soluciones. Hemos aprendido a escribirlas todas en función de un pequeño conjunto de parámetros: las variables libres. La elección de las variables libres es un tanto arbitraria, y un conjunto de variables libres puede ser sustituido por algunas combinaciones de ellas. Sin embargo el **número** de variables libres, la cantidad de parámetros independientes necesarios para describir todas las soluciones, depende solamente de la matriz A del sistema. No de la elección de las variables. Vimos un ejemplo de esto en el ejercicio 1.14, página 48. Este hecho está relacionado con la dimensión del subespacio formado por todas las soluciones de la ecuación homogénea $AX = O$. Presentaremos este subespacio, llamado el *núcleo de la matriz A*, y discutiremos sus propiedades en la sección 1.7.

Cerremos esta introducción con un ejemplo más geométrico. Los planos son objetos que tienen dimensión 2 en el espacio tridimensional , las rectas tienen dimensión 1. La teoría que estamos elaborando da un sentido preciso a esta afirmación. Discutiremos esto en el capítulo 3, dedicado a la geometría del espacio $\mathbb{R}^3$.

1.6.1. La definición de dimensión

Apuntamos a definir la dimensión de un subespacio vectorial de $\mathbb{K}^n$ como la cantidad de vectores que tiene una base. Comencemos por recordar que un subespacio vectorial puede tener muchas bases. Presentamos a continuación dos ejemplos.

Ejemplo 1.6.1. Volvamos a la matriz del ejercicio 1.62, en la página 118. El lector habrá encontrado ya varias bases de su espacio de columnas. ♣

Ejemplo 1.6.2. Algunas bases de $\mathbb{R}^3$
Para cualquier valor de los parámetros a, b y c la familia

$$((1, a, b), (0, 1, c), (0, 0, 1)) \tag{1.46}$$

es una base de $\mathbb{R}^3$. Para ver que esta familia es generadora y linealmente independiente tenemos que ver que cualquier vector (x, y, z) puede escribirse en la forma

$$(x, y, z) = \lambda_1(1, a, b) + \lambda_2(0, 1, c) + \lambda_3(0, 0, 1),$$

como una combinación lineal de la familia, y que la única combinación que produce el vector nulo es la que tiene coeficientes $\lambda_i = 0$, $i = 1, 2, 3$. Analizar esa combinación lineal es equivalente a estudiar el sistema

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & x \\ a & 1 & 0 & y \\ b & c & 1 & z \end{array}\right),$$

que tiene una forma escalerizada

$$\left(\begin{array}{ccc|l} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y - ax \\ 0 & 0 & 1 & z + (ca - b)x - cy \end{array}\right).$$

De esta forma escalerizada se desprende que el sistema es compatible para cualquier elección de (x, y, z). Haciendo $(x, y, z) = (0, 0, 0)$ resulta que su única solución es la trivial. Por lo tanto (1.46) es una base de $\mathbb{R}^3$ para cualquier elección de los parámetros. ♣

Estos dos ejemplos ilustran el hecho de que un subespacio vectorial tiene, en general, muchas bases. En ambos casos todas las bases tienen exactamente la misma cantidad de vectores, dos y tres respectivamente. Este es un hecho completamente general, que hace posible definir la dimensión de un subespacio vectorial como la cantidad de elementos de una base, porque esta cantidad en realidad no depende de la base escogida. **Es una propiedad del subespacio**.

Nuestra próxima proposición apunta a demostrar esta última afirmación. Expresa que en un subespacio vectorial la cantidad de elementos en un conjunto linealmente independiente no puede superar a la cantidad de elementos en un generador.

Proposición 1.12. *Sea $\mathbb{S}$ un subespacio vectorial de $\mathbb{K}^l$, sea*

$$\mathcal{A} = (v_1, v_2, \ldots, v_m) \subset \mathbb{S}$$

un conjunto que genera a $\mathbb{S}$, y

$$\mathcal{B} = (w_1, w_2 \ldots, w_n) \subset \mathbb{S}$$

un conjunto linealmente independiente. Entonces $n \leq m$.

Observación 1.6.3. Esta proposición puede interpretarse heurísticamente en términos de la cantidad de grados de libertad, o de parámetros que podemos fijar a nuestro antojo, con los que contamos para movernos dentro de un subespacio vectorial $\mathbb{S}$. Cada vector de una familia independiente aporta un grado de libertad, porque no puede reducirse a ninguno de los otros vectores de la familia. Por lo tanto la existencia de una familia linealmente independiente con n vectores de $\mathbb{S}$ asegura que hay al menos n grados de libertad en el subespacio $\mathbb{S}$.

Cada vector de una familia generadora nos da, a lo sumo, un grado de libertad. Podría ocurrir que algún vector del generador fuera redundante, generando cosas que ya pueden obtenerse de otros vectores. La existencia de un generador con m vectores asegura que hay, a lo sumo, m grados de libertad en el espacio.

Estas observaciones sugieren entonces que n no puede superar el valor de m. A continuación demostraremos este hecho. La prueba consiste en una interesante aplicación de la teoría que hemos desarrollado para los sistemas de ecuaciones lineales a partir del algoritmo de escalerización. ♠

Prueba de la proposicion 1.12. La demostración se reduce a una observación sobre el sistema de ecuaciones lineales que se obtiene al estudiar la independencia lineal de la familia $\mathcal{B}$ luego de expresar sus vectores como combinaciones lineales de los vectores del generador $\mathcal{A}$.

Como $\mathcal{A}$ es un conjunto generador de $\mathbb{S}$ cada w_j en $\mathcal{B}$, para $j = 1, 2, \dots, n$, puede escribirse como una combinación lineal de $\mathcal{A}$, con ciertos coeficientes

$$a_{ij}, \quad i = 1, 2, \dots, m,$$

tales que

$$w_j = a_{1j}v_1 + a_{2j}v_2 + \dots + a_{mj}v_m.$$

Formemos ahora una combinación lineal de la familia $\mathcal{B}$ con coeficientes

$$\lambda_j, \quad j = 1, 2, \dots, n,$$

e igualémosla al vector nulo O. Tenemos

$$O = \lambda_1 w_1 + \lambda_2 w_2 \dots \lambda_n w_n. \tag{1.47}$$

Como $\mathcal{B}$ es linealmente independiente esta igualdad sólo puede satisfacerse cuando todos los coeficientes λ_j, $j = 1, \dots, n$ son nulos. Analicemos la igualdad (1.47) introduciendo las expresiones de los vectores de $\mathcal{B}$ como combinaciones lineales de $\mathcal{A}$. Resulta

$$\begin{aligned} O = \lambda_1 \left(a_{11}v_1 + \dots a_{m1}v_m\right) + \lambda_2 \left(a_{12}v_1 + \dots a_{m2}v_m\right) + \dots \\ \dots + \lambda_n \left(a_{1n}v_1 + \dots a_{mn}v_m\right). \end{aligned} \tag{1.48}$$

Podemos reordenar los términos de estas combinaciones lineales, agrupando en los vectores v_j de la familia $\mathcal{A}$, y obtenemos

$$\begin{aligned} O = \left(a_{11}\lambda_1 + \dots + a_{1n}\lambda_n\right) v_1 + \left(a_{21}\lambda_1 + \dots + a_{2n}\lambda_n\right) v_2 + \dots \\ \dots + \left(a_{m1}\lambda_1 + \dots + a_{mn}\lambda_n\right) v_m. \end{aligned} \tag{1.49}$$

Naturalmente, si cada uno de los coeficientes

$$a_{i1}\lambda_1 + \dots + a_{in}\lambda_n, \quad i = 1, 2, \dots, m$$

que aparecen multiplicando a los vectores v_i, $i = 1, 2, \dots, m$, en (1.49) es igual a 0, entonces toda la combinación lineal es igual a O, la igualdad (1.49) se satisface, y también se satisface (1.47).

Por lo tanto, una elección de $(\lambda_1, \lambda_2, \dots, \lambda_n)$ que satisfaga el sistema lineal de ecuaciones

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{1.50}$$

hace que la combinación lineal de $\mathcal{B}$ que aparece en (1.47) sea igual al vector nulo. Como $\mathcal{B}$ es linealmente independiente, la única elección posible de los coeficientes λ_i, $i = 1, 2, \ldots, n$ es tomarlos todos nulos. De modo que el sistema homogéneo (1.48) sólo puede admitir la solución trivial. Esto sólo puede ocurrir si el número m de ecuaciones es mayor o igual que el número n de incógnitas (ver la parte 2 del ejercicio 1.33, en la página 65). De modo que $n \leq m$. □

La proposición 1.12 tiene como consecuencia el siguiente corolario.

Corolario 1.13. *Sea $\mathbb{S}$ un subespacio vectorial de $\mathbb{K}^l$. Si $\mathbb{S}$ tiene una base que tiene n vectores, entonces todas las bases tienen n vectores.*

PRUEBA. Sean $\mathcal{A}$ una base de $\mathbb{S}$, con n vectores respectivamente, y $\mathcal{B}$ otra base de $\mathbb{S}$. Como $\mathcal{A}$ es un generador y $\mathcal{B}$ es linealmente independiente, entonces $\mathcal{B}$ puede contener a lo sumo n vectores. Por lo tanto tiene una cantidad finita m de vectores, con $m \leq n$.

Aplicamos ahora otra vez la proposición 1.12 invirtiendo los papeles de $\mathcal{A}$ y $\mathcal{B}$, usando ahora que $\mathcal{B}$ es un generador y $\mathcal{A}$ es linealmente independiente. Concluimos que $n \leq m$.

Entonces $m = n$. □

Podemos definir ahora la dimensión de un espacio vectorial, sin exponernos a ambigüedades de ningún tipo.

Definición 1.9 (Dimensión de un subespacio vectorial). *Sea $\mathbb{S}$ un subespacio vectorial de $\mathbb{K}^n$, y $\mathcal{B}$ una base de $\mathbb{S}$. Llamaremos* ***dimensión*** *de $\mathbb{S}$ a la cantidad de vectores que tiene la base $\mathcal{B}$. El subespacio vectorial trivial $\{O\}$ no tiene bases. Convendremos en asignarle dimensión 0.*

Indicaremos con la notación $\dim(\mathbb{S})$ la dimensión del subespacio $\mathbb{S}$.

Ejemplo 1.6.4. LA DIMENSIÓN DE $\mathbb{K}^n$
Vimos en el ejemplo 1.5.7, página 119, que el espacio $\mathbb{K}^n$ tiene una base, que hemos dado en llamar base canónica, que tiene n vectores. Por lo tanto la dimensión de $\mathbb{K}^n$ es n. ♣

Ejemplo 1.6.5. La dimensión del espacio de columnas de la matriz

$$\begin{pmatrix} 1 & 0 & -1 & 1 \\ -1 & 1 & 3 & -4 \\ 2 & 1 & 0 & -1 \end{pmatrix}$$

con la que trabajamos en el ejercicio 1.62 es igual a 2. ♠

En general, sabemos como identificar una base del espacio de columnas col(A) de una matriz A cualquiera: tenemos que seleccionar las columnas de A que corresponden a posiciones de pivotes en una forma escalerizada. Por lo tanto una base de col(A) tendrá tantos vectores como pivotes una forma escalerizada. Esta observación constituye la prueba de nuestra próxima proposición.

Proposición 1.14. *Sea A una matriz. La dimensión* $\dim(\mathrm{col}(A))$ *de su espacio de columnas es igual al número de pivotes de una forma escalerizada de A.*

Ejemplo 1.6.6. Como ejemplo de la aplicación de esta proposición, digamos que el espacio de columnas de la matriz (1.45), en la introducción de esta sección, tiene dimensión cuatro. El lector sólo debe examinar alguna de sus formas escalerizadas para verificarlo. ♣

Observación 1.6.7. La dimensión de un subespacio es la cantidad de direcciones independientes que "caben" en él. Por ejemplo, si identificamos los vectores de $\mathbb{R}^3$ con puntos del espacio tridimensional (lo que es esencialmente equivalente a poner un sistema de coordenadas en el espacio) entonces los subespacios de dimensión uno son rectas que pasan por el origen de coordenadas; y los de dimensión 2 son planos que pasan por el origen. El único subespacio de dimensión 3 es todo $\mathbb{R}^3$, que representaría a todo el espacio. Vemos que la noción de dimensión concuerda con nuestra idea intuitiva de que una recta es un objeto que tiene una dimensión, un plano tiene dos, y el espacio tiene tres. Daremos más detalles sobre esto en las secciones del curso dedicadas a la geometría del espacio. También al discutir la teoría general de los espacios vectoriales. ♠

Observación 1.6.8. Comentarios sobre la definición de dimensión Nuestra definición no asegura, en principio, que todo subespacio vectorial de $\mathbb{K}^n$ tenga dimensión. Para asignar la dimensión hace falta contar los vectores de una base. Y todavía no hemos demostrado que todo subespacio de $\mathbb{K}^n$ tenga una base. Este hecho es cierto, y daremos una demostración en el capítulo dedicado a la teoría general de los espacios vectoriales. Es cierto además que todas las bases de subespacios de $\mathbb{K}^n$ tienen una cantidad finita de vectores. Este comentario es de interés teórico, pero no tiene mayor importancia práctica, porque siempre que trabajemos con algún subespacio particular podremos identificar alguna de sus bases.

Las bases de un subespacio no trivial $\mathbb{S} \subset \mathbb{K}^n$ son subconjuntos no vacíos de $\mathbb{S}$, porque el conjunto vacío no puede generar a $\mathbb{S}$. Por lo tanto el único subespacio que puede tener dimensión 0 es el subespacio trivial $\{O\}$. Para los otros subespacios la dimensión será algún número natural mayor o igual que 1, pero no mayor que n (ver el corolario 1.17). ♣

De la definición de dimensión y la proposición 1.12 se desprende que la dimensión de un subespacio vectorial es igual a la máxima cantidad de vectores que puede tener una familia linealmente independiente contenida en el subespacio, y a la mínima cantidad de vectores en un generador del subespacio. Vale además la siguiente proposición.

Proposición 1.15. *Sea $\mathbb{S}$ un subespacio vectorial de $\mathbb{K}^n$ de dimensión $m \geq 1$. Entonces*

1. *Todo subconjunto de $\mathbb{S}$ que sea linealmente independiente y contenga m vectores es una base de $\mathbb{S}$.*
2. *Todo subconjunto de $\mathbb{S}$ que sea un generador de $\mathbb{S}$ y contenga m vectores es una base de $\mathbb{V}$*

La demostración constituye un ejercicio interesante que obliga a revisar las nociones de base, independencia lineal y generadores, de modo que la dejamos para el lector.

Ejercicio 1.63. Demuestre la proposición 1.15.

Ejercicio 1.64. Utilizando la proposición anterior investigar si los siguientes conjuntos son base del espacio dado.

1. $((1,1,1,2),(0,1,0,2),(-1,1,0,1))$ de $\mathbb{R}^4$
2. $((1,0),(0,1),(1,1))$ de $\mathbb{Z}_2^n$
3. $((1,1,1),(0,1,-1),(2,-2,0),(-1,-2,1))$ de $\mathbb{R}^3$

En el mismo círculo de ideas que la proposición 1 está nuestro próximo ejercicio.

Ejercicio 1.65. Sea $\mathbb{S}$ un subespacio vectorial de $\mathbb{K}^n$ y $\mathcal{A}$ un subconjunto ordenado de $\mathbb{S}$. Mostrar que $\mathcal{A}$ es una base de $\mathbb{S}$ si y solo si todo vector de $\mathbb{S}$ puede escribirse de una única manera como una combinación lineal de los vectores de $\mathcal{A}$.

Si una familia linealmente independiente $\mathcal{A}$ contenida en un subespacio $\mathbb{S}$ de dimensión m tiene menos de m vectores entonces no puede generar todo el espacio. Existe entonces un vector $w \in \mathbb{S}$ que no es una combinación lineal de $\mathcal{A}$.

Ejercicio 1.66. Mostrar que la familia que se obtiene agregando a $\mathcal{A}$ el vector w está contenida en $\mathbb{S}$ y es linealmente independiente.

La dimensión de un espacio vectorial es una medida del "tamaño" del espacio, medido en el sentido de la estructura lineal. En este sentido[11] $\mathbb{R}^3$, con sus tres dimensiones, es "más grande" que $\mathbb{R}^2$, que es un espacio de dimensión 2. A continuación presentamos dos resultados que ilustran esta idea.

Proposición 1.16. *Sean* $\mathbb{S}$ *y* $\mathbb{T}$ *dos subespacios vectoriales de* $\mathbb{K}^n$*, de dimensiones* $\dim(\mathbb{S})$ *y* $\dim(\mathbb{T})$ *respectivamente, tales que* $\mathbb{S} \subset \mathbb{T}$*. Entonces*

1. $\dim(\mathbb{S}) \leq \dim(\mathbb{T})$.
2. *Si* $\dim(\mathbb{S}) = \dim(\mathbb{T})$ *entonces* $\mathbb{S} = \mathbb{T}$.

PRUEBA. Comencemos con la primera parte de la proposición. Si $\mathbb{S} = \{O\}$, entonces
$$\dim(\mathbb{S}) = 0 \leq \dim(\mathbb{T}).$$
Si $\mathbb{S} \neq \{O\}$ sabemos que su dimensión es igual a la cantidad de elementos de una base $\mathcal{A}$ de $\mathbb{S}$. Una base de $\mathbb{S}$ es un subconjunto linealmente independiente de $\mathbb{S}$. Como $\mathbb{S} \subset \mathbb{T}$ entonces $\mathcal{A}$ es también un subconjunto linealmente independiente de $\mathbb{T}$, y puede tener, a lo sumo, tantos vectores como $\dim(\mathbb{T})$. Esto prueba $\dim(\mathbb{S}) \leq \dim(\mathbb{T})$.

Para probar la segunda parte basta mostrar que todo vector de $\mathbb{T}$ está en $\mathbb{S}$, porque la inclusión de $\mathbb{S}$ en $\mathbb{T}$ está asegurada por las hipótesis.

Si $\mathbb{T} = \{O\}$ el resultado es obvio.

Si $\mathbb{T} \neq \{O\}$ entonces consideremos una base $\mathcal{A}$ de $\mathbb{S}$. Como $\mathbb{S}$ está contenido en $\mathbb{T}$ la familia $\mathcal{A}$ está contenida en $\mathbb{T}$, es linealmente independiente por ser una base de $\mathbb{S}$, y por la misma razón tiene $\dim(\mathbb{S}) = \dim(\mathbb{T})$ vectores. La parte 1 de la proposición 1.15 implica entonces que es también una base de $\mathbb{T}$.

Por lo tanto, cualquier elemento t de $\mathbb{T}$ puede escribirse como una combinación lineal de $\mathcal{A} \subset \mathbb{S}$, lo que implica que t está en $\mathbb{S}$, ya que $\mathbb{S}$ es un subespacio (ver el ejercicio 1.45, en la página 97). $\square$

Enunciamos a continuación un corolario de la última proposición, que nos será especialmente útil más adelante.

Corolario 1.17. *Sea* $\mathbb{S}$ *un subespacio de* $\mathbb{K}^n$ *que tiene dimensión* m*. Entonces* $m \leq n$*, y* $m = n$ *si y sólo si* $\mathbb{S} = \mathbb{K}^n$.

[11] Sin embargo el tamaño $\mathbb{R}^2$ y $\mathbb{R}^3$ medido con su cardinalidad como conjuntos es el mismo: ambos tiene exactamente la misma cantidad de elementos, porque es posible encontrar una función biyectiva entre $\mathbb{R}^2$ y $\mathbb{R}^3$.

Ejercicio 1.67. Demostrar el corolario.

Ejercicio 1.68. Hallar la dimensión del subespacio $\mathbb{S}$.

1. $\mathbb{S} = \left\{(x, y, z) \in \mathbb{R}^3;\ x - y = 0,\ y + z = 0\right\}.$
2. $\mathbb{S} = \left\{(x_1, x_2) \in \mathbb{Z}_2^2;\ x_1 + x_2\right\}.$

Ejercicio 1.69. Hallar la dimensión de los siguientes subespacios.

1. El subespacio $\mathbb{S} \in \mathbb{R}^3$ generado por $\mathcal{A} = ((1, 1, 1), (0, -1, -2), (1, 2, 3))$
2. El subespacio $\mathbb{S} \in \mathbb{Z}_2^3$ generado por $\mathcal{A} = ((1, 1, 1), (0, 1, 1), (0, 1, 0))$
3. El subespacio $\mathbb{S} \in \mathbb{R}^2$ generado por $\mathcal{A} = ((1, 1), (0, 1), (1, 0))$

Ejercicio 1.70. Sea el subespacio $\mathbb{S} = \left\{(x, y, z, t) \in \mathbb{R}^4;\ x + y = 0,\ z - 2t = 0\right\}.$

1. Hallar una base de $\mathbb{S}$.
2. Agregar vectordes de $\mathbb{R}^4$ a la base hallada hasta obtener una base de $\mathbb{R}^4$.

Ejercicio 1.71. Mostrar que un subespacio $\mathbb{S}$ de dimensión k en $\mathbb{Z}_{\nvDash}{}^{k}$ tiene 2^k vectores. Comparar 2^k (la cantidad de vectores en $\mathbb{S}$) con k (la cantidad de vectores en una base de $\mathbb{S}$) para distintos valores de k.

1.6.2. Para tener presente

- Todas las bases de un subespacio vectorial $\mathbb{S}$ de $\mathbb{K}^n$ tienen la misma cantidad de vectores. La dimensión de $\mathbb{S}$ es igual a la cantidad de vectores que tiene cualquiera de sus bases.

- Un subconjunto linealmente independiente de vectores de un subespacio $\mathbb{S}$ no puede tener más vectores que $\dim(\mathbb{S})$. Si tiene exactamente $\dim(\mathbb{S})$ es una base de $\mathbb{S}$.

- Un generador de un subespacio $\mathbb{S}$ no puede tener menos vectores que $\dim(\mathbb{S})$. Si tiene exactamente $\dim(\mathbb{S})$ es una base de $\mathbb{S}$.

- El único subespacio de $\mathbb{K}^n$ que tiene dimensión n es el propio $\mathbb{K}^n$.

1.7. Determinación de los sistemas lineales y núcleo

En esta sección vamos a considerar con más detalle la unicidad o no unicidad de soluciones de los sistemas de ecuaciones lineales $AX = B$. Describiremos el conjunto solución de un sistema lineal en términos de un nuevo subespacio vectorial de $\mathbb{K}^n$ asociado a la matriz A del sistema: el núcleo de la matriz. Encontraremos nuevamente que las nociones del álgebra lineal que hemos venido introduciendo, como son los conceptos de subespacio, generadores, independencia lineal, y dimensión, son de gran utilidad para comprender el comportamiento de los sistemas lineales.

1.7.1. Sistemas indeterminados: forma general de las soluciones

Vamos a comenzar el análisis del problema de unicidad de soluciones de un sistema de ecuaciones lineales volviendo a uno de nuestros ejemplos favoritos, que nos viene acompañando desde la sección 1.2.

Ejemplo 1.7.1. Retomemos el estudio del sistema real $AX = B$, con matriz y término independiente definidos por

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \\ -1 \end{pmatrix}, \tag{1.51}$$

con el que trabajamos en varios de los ejemplos anteriores y analicemos la unicidad (en este caso más bien no unicidad), de sus soluciones.

Al estudiar el sistema lineal habíamos llegado a la siguiente forma escalerizada reducida para la matriz ampliada del sistema (ver la página 62):

$$\left(\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & \frac{3}{2} & 0 & -1 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{1.52}$$

De aquí dedujimos que las soluciones del sistema son de la forma

$$(-1 - 2x_2 - 3x_5/2, x_2, 1 + x_5/2, -1 - 2x_5, x_5, 1).$$

Vamos a expresar una columna X, solución del sistema, como

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -\frac{3}{2} \\ 0 \\ \frac{1}{2} \\ -2 \\ 1 \\ 0 \end{pmatrix}. \tag{1.53}$$

¿Qué es cada una de las columnas que aparece en el miembro derecho de esta fórmula?

La primera columna, que no aparece multiplicada por las variables libres, es una solución del sistema. Se trata de una solución particular, que es la que resulta de escoger los valores $x_2 = x_5 = 0$ para las variables libres.

Si revisamos las operaciones que hicimos para llegar a la fórmula (1.53) encontraremos que las entradas de esa columna se obtienen de la última columna de la matriz escalerizada reducida, que es justamente la que almacenaba la información acerca del término independiente B.

Miremos ahora las columnas que multiplican x_2 y x_5: éstas no se ven afectadas por el término independiente, porque no dependen de las entradas en la última columna de la forma escalerizada de la matriz ampliada. Si cambiáramos esa columna por una columna de ceros obtendríamos lo mismo. De hecho, los sumandos en (1.53) que contienen a las variables libres x_2 y x_5 son una solución de $AX = O$.

Veamos la justificación de la afirmación que cierra el párrafo anterior. Imaginemos que tenemos que resolver $AX = O$. Buscaríamos entonces una forma escalerizada para la matriz $A|O$, que es la matriz ampliada de este sistema. Las transformaciones elementales necesarias para escalerizar $A|O$ son exactamente las mismas que hicimos para escalerizar $A|B$, ya que sólo dependen de la matriz A del sistema, no de la columna de los términos independientes.

La forma escalerizada reducida que obtendríamos al final del proceso sería una matriz como (1.52), pero con ceros en la última columna. Al calcular las soluciones reencontraríamos entonces una fórmula como (1.53), en la que la primera columna sería una columna de ceros. Es decir, la primera columna no aparecería.

Concluimos entonces que el segundo y tercer sumando de (1.53) contienen la forma general de las soluciones de $AX = O$.

Nuestro análisis ha puesto en evidencia que la solución general del sistema

$$AX = B$$

está compuesta de dos partes:

1. una solución particular del sistema;
2. la expresión de la solución general del sistema $AX = O$. ♣

Mostraremos ahora que las conclusiones a las que llegamos en el ejemplo anterior son completamente generales: si conocemos el conjunto formado por todas las soluciones de

$$AX = O$$

y una solución particular de la ecuación

$$AX = B$$

, entonces podemos recuperar cualquier solución de esta última ecuación.

Observación 1.7.2. Llamaremos *homogéneo* al sistema

$$AX = O$$

porque tiene la propiedad de que el producto aX de un número a cualquiera por una solución X también es solución del sistema (ver la proposición 1.19, en la página 133 de esta misma sección). Recordemos que habíamos introducido este término en la página 28, inmediatamente antes del ejercicio 1.12.

Es usual referirse entonces a $AX = B$, con $B \neq O$, como el *sistema no homogéneo*.

Mostraremos a continuación que las conclusiones a las que llegamos en el ejemplo anterior son completamente generales: si conocemos el conjunto formado por **todas** las soluciones de $AX = O$ y **una** solución particular de $AX = B$, entonces **podemos recuperar cualquier solución** de $AX = B$. Este hecho suele expresarse de la siguiente manera: **la solución general del sistema no homogénea es la suma de una solución particular más la solución general del sistema homogéneo**. Esta observación explica también la notación que usaremos en el enunciado y la demostración de nuestra próxima proposición. ♠

Proposición 1.18. *Sea X_{NH_0} una solución del sistema de ecuaciones lineales*

$$AX = B. \tag{1.54}$$

Entonces la suma de X_{NH_0} con una solución cualquiera X_H del sistema

$$AX = B. \tag{1.55}$$

es otra solución de (1.54)*. Más aún, cualquier solución de* (1.54) *puede expresarse como*

$$X_{NH_0} + X_H,$$

donde X_H es una solución de (1.55)*.*

Prueba: Consideremos la suma de la solución X_{NH_0} con una solución X_H cualquiera de (1.55). Entonces

$$A(X_{NH_0} + X_H) = AX_{NH_0} + AX_H = B + O = B,$$

y $X_{NH_0} + X_H$ satisface (1.54).

Supongamos que tenemos una solución X_1 de (1.54). Entonces la diferencia $X_1 - X_{NH_0}$ satisface

$$A(X_1 - X_{NH_0}) = AX_1 - AX_{NH_0} = B - B = 0.$$

Por lo tanto $X_1 - X_{NH_0}$ es una solución de (1.55), y

$$X_1 = X_{NH_0} + (X_1 - X_{NH_0})$$

es la suma de X_0 más una solución de la ecuación homogénea. □

La consecuencia de esta proposición es que para entender cḿo es el conjunto de todas las soluciones de un sistema lineal compatible $AX = B$ basta entender cómo son las soluciones de $AX = O$. En particular $AX = B$ tiene solución única si y sólo si $AX = O$ tiene solución única, un resultado que ya habíamos probado en la proposición 1.8 de la página 111.

1.7.2. El núcleo de una matriz

Ya sabemos que la información acerca de la estructura del conjunto de soluciones de $AX = B$ está encerrada en el conjunto de soluciones de la ecuación homogénea $AX = O$.

Definición 1.10 (Núcleo de una matriz). *Llamaremos* ***núcleo*** *de la matriz A al conjunto de las soluciones de la ecuación homogénea $AX = O$.*

Indicaremos este conjunto con el símbolo[12] $\ker(A)$.

Si A es una matriz $m \times n$ sobre el cuerpo $\mathbb{K}$ podemos escribir entonces

$$\ker(A) = \{X \in \mathbb{K}^n;\ AX = O\}.$$

La primera noticia interesante es que este conjunto resulta ser también un subespacio vectorial.

Proposición 1.19. *Si A es una matriz $m \times n$ sobre $\mathbb{K}$, entonces el núcleo de A es un subespacio vectorial de $\mathbb{K}^n$.*

[12] Proveniente de la palabra alemana kernel

Prueba: La demostración es bastante sencilla, y descansa en la linealidad de todas los cálculos que hay que hacer para verificar que un vector está en el núcleo. Supongamos que tenemos dos vectores X e Y en el núcleo. Entonces su suma $X+Y$ satisface

$$A(X+Y) = AX + AY = O + O = O,$$

por lo que también está en el núcleo. Para el producto de cualquiera de ellos, por ejemplo X, por un escalar a en el cuerpo tenemos

$$A(aX) = a(AX) = aO = O.$$

Entonces también aX está en el núcleo. □

Como el núcleo es un subespacio podremos aprovechar la estructura lineal de $\mathbb{K}^n$ para describirlo de manera eficaz en términos de ecuaciones lineales o a través de la especificación de una base. Otra vez, toda la información necesaria será obtenida del método de escalerización.

Un conjunto de ecuaciones para el núcleo

El núcleo está definido por la ecuación vectorial $AX = O$, que es equivalente a m ecuaciones escalares. El proceso de escalerización de la matriz A la transforma en una matriz escalerizada (o escalerizada reducida) E, que está asociada a un sistema lineal que tiene el mismo conjunto solución que el sistema original $AX = O$. En otras palabras, el núcleo de A es igual al conjunto de los X que satisface $EX = O$.

El sistema $EX = 0$ tiene tantas ecuaciones escalares como pivotes tenga la matriz escalerizada E, y el número de ecuaciones ya no puede reducirse más. El sistema escalerizado (o puesto en la forma escalerizada reducida) es una forma de caracterizar el núcleo de A por un conjunto de ecuaciones en el que hemos eliminado las ecuaciones redundantes.

Notemos que estas consideraciones también nos dice que

$$\ker(A) = \ker(E).$$

Ejemplo 1.7.3. Consideremos la matriz A del ejemplo 1.7.1 con que comenzamos esta sección (ver la fórmula (1.51). Al llevar la matriz ampliada $A|O$ del sistema $AX = O$ a su forma escalerizada reducida encontramos la matriz

$$\left(\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

que es equivalente a las ecuaciones

$$\left\{\begin{aligned} x_1 + 2x_2 + 3x_5/2 &= 0, \\ x_3 - x_5/2 &= 0, \\ x_4 + 2x_5 &= 0, \\ x_6 &= 0. \end{aligned}\right.$$

Estas ecuaciones caracterizan al núcleo de A, y no pueden simplificarse más. En particular, ninguna de ellas puede obtenerse a partir de las otras haciendo combinaciones lineales. Demostraremos este hecho con toda generalidad en la proposición 1.23, en la página 147 de la sección 1.8. ♣

Una base para el núcleo

También es posible caracterizar el núcleo por medio de una base. Mostraremos esto en el ejemplo con el que hemos venido trabajando.

Ejemplo 1.7.4. En la expresión (1.53) para las soluciones del ejemplo 1.7.1 cada variable libre aparece multiplicando un vector. El que corresponde a x_2 es

$$(-2, 1, 0, 0, 0, 0),$$

y el que está asociado con x_5 es

$$(-3/2, 0, 1/2, -2, 1, 0).$$

Con estos dos vectores podemos generar todas las soluciones de $AX = O$. Además son linealmente independientes porque cada una de las variables libres aparece sólo en el vector que está asociado con ella. En efecto, si planteamos la igualdad

$$x_2(-2, 1, 0, 0, 0, 0) + x_5(-3/2, 0, 1/2, -2, 1, 0) = (0, 0, 0, 0, 0, 0)$$

al mirar la la segunda y quinta componentes obtenemos

$$x_2 = 0, \quad x_5 = 0,$$

respectivamente. De modo que la única combinación que produce el vector nulo es la trivial. La familia formada por estos dos vectores constituye una base para el núcleo de A, que resulta tener dimensión igual a 2. ♣

El procedimiento anterior es completamente general: al expresar las soluciones de $AX = O$ en términos de las variables libres cada variable libre queda asociada a un vector (en realidad este vector es el que se obtiene fijando en uno el valor de la variable que estamos considerando, y en cero todas las demás). Esta familia de vectores es una base para el núcleo, algo que demostraremos en general en la proposición 1.20. Tenemos entonces un procedimiento para hallar una base del núcleo.

El número de variables libres tiene una interpretación clara: es igual a la dimensión del núcleo de la matriz del sistema. Las variables libres pueden cambiar, porque su elección está asociada a la elección de una base para el núcleo, y hay muchas bases posibles. Pero su número esta fijo. Es una propiedad de la matriz del sistema, que no depende de cómo ordenemos los cálculos para expresar las soluciones.

Observación 1.7.5. No es necesario obtener la forma escalerizada reducida de A para construir una base del núcleo. Todo lo que hace falta es expresar las soluciones de $AX = O$ en términos de las variables libres, lo que puede hacerse a partir de la forma escalerizada. En nuestra presentación preferimos llegar a la forma escalerizada reducida, porque ésta hace más evidente que la información que proviene del término independiente aparece separada de las variables libres en la solución general del sistema. ♠

Ejercicio 1.72. Encontrar una base del núcleo de las matrices reales

1. $\begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix};$

2. $\begin{pmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 3 & 0 & 1 & -1 \\ -1 & 0 & 2 & -1 & 0 \\ -1 & -1 & 1 & 4 & 1 \\ 0 & -1 & 0 & 1 & 2 \end{pmatrix};$

3. $\begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 3 & 2 \\ -1 & 2 & 7 & 5 \\ 2 & -1 & -5 & -4 \end{pmatrix}.$

Ejercicio 1.73. Hallar una base del núcleo de la matriz

$$B = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

con entradas en $\mathbb{Z}_2$.

Vale la pena recordar al lector que ya trabajamos con estas matrices en ejercicios anteriores.

En la sección 1.5 logramos describir la imagen de una matriz A por medio de algunas ecuaciones lineales. Al emplear este procedimiento resulta que la imagen de A es el conjunto de soluciones de un sistema de ecuaciones lineal y homogeneo, por lo tanto la imagen de A es igual al núcleo de la matriz de este sistema.

Ejemplo 1.7.6. En la página 103 encontramos que la imagen de la matriz A del ejemplo 1.7.1 es

$$\mathrm{col}(A) = \{(b_1, b_2, b_3, b_4, b_5);\ b_5 + b_4 - 2b_3 = 0\}.$$

Por lo tanto $\mathrm{col}(A)$ es el núcleo de la matriz fila

$$(0\ 0\ -2\ 1\ 1).$$

Esta matriz ya esta (trivialmente) escalerizada y sólo tiene un pivote en la tercera entrada. Las variables b_1, b_2, b_4 y b_5 son variables libres. La solución general de la ecuación es

$$(b_1, b_2, (b_4 + b_5)/2, b_4, b_5), \quad b_1, b_2, b_4, b_5 \in \mathbb{R}$$

y

$$((1,0,0,0,0), (0,1,0,0,0), (0,0,1/2,1,0), (0,0,1/2,0,1))$$

es una base del espacio de columnas de A que se obtiene por el procedimiento de estudiar el núcleo de la matriz del sistema de ecuaciones que caracteriza al espacio $\mathrm{col}(A)$. ♣

Ejercicio 1.74. Para las siguientes matrices reales buscar una base del espacio de columnas con la técnica del ejemplo anterior. Expresar cada columna de dichas matrices como combinación lineal de los elementos de la base hallada.

1. $\begin{pmatrix} 1 & 1 & 5 & -3 \\ -1 & 5 & 2 & -2 \\ -1 & 2 & -\frac{3}{2} & \frac{1}{2} \\ 2 & -4 & 3 & -1 \\ -1 & 5 & 2 & -2 \end{pmatrix}$

2. $\begin{pmatrix} 4 & 15 & 12 & 5 \\ 2 & 8 & 4 & 2 \\ 1 & 5 & -2 & 0 \\ 1 & 4 & 2 & 1 \end{pmatrix}$

Nuestra siguiente proposición demuestra que el procedimiento que hemos encontrado para construir una base del núcleo de una matriz es correcto.

Proposición 1.20. *Consideremos una matriz A de dimensiones $m \times n$ sobre un cuerpo $\mathbb{K}$ y E una forma escalerizada de A. Supongamos que las columnas*

$$i_1, i_2, \ldots, i_k$$

de E no tienen pivotes (es decir, corresponden a variables libres). Llamemos

$$D_l, \quad l = 1, 2, \ldots, k,$$

a la solución de $AX = 0$ que se obtiene asignando el valor 1 a la l-ésima variable libre x_{i_l}, y 0 a las $k-1$ variables libres restantes. Entonces

$$\mathcal{D} = (D_1, D_2, \ldots, D_k)$$

es una base del núcleo de A.

Prueba: Consideremos la forma escalerizada reducida

$$E = (e_{ij})_{i=1,2\ldots,m}^{j=1,2\ldots,n}$$

correspondiente a la matriz A. Las entradas en una solución X de la ecuación

$$AX = O$$

quedan determinadas por las variables libres

$$x_{i_l}, \quad l = 1, \ldots, k.$$

La i-ésima variable x_i se calcula en términos de las variables libres por una expresión de la forma

$$x_i = d_{i1}x_{i_1} + d_{i2}x_{i_2} + \ldots + d_{ik}x_{i_k}.$$

Si x_i es una de las variables que corresponde a columnas con pivotes entonces

$$d_{ij} = -e_{ii_j}, \quad j = 1, 2, \ldots, k,$$

y si x_i es una de las variables libres x_{i_l} entonces

$$d_{ij} = \begin{cases} 1, & j = l, \\ 0, & j \neq l. \end{cases}$$

Formemos con estos números la matriz

$$D = (d_{ij})_{i=1,2...,n}^{j=1,2...,k}.$$

Cada solución X es una combinación lineal de las columnas de la matriz, cuyos coeficientes son los valores que para esa solución toman las variables libres. Las columnas de D son justamente los vectores

$$D_l, \quad l = 1, 2, \ldots, k,$$

lo que muestra que $\mathcal{D}$ es un generador del núcleo de A.

Resta mostrar que $\mathcal{D}$ es linealmente independiente. Consideremos una combinación lineal de los vectores de $\mathcal{D}$ que sea igual al vector nulo. Es decir, escalares

$$\lambda_l, \quad l = 1, 2, \ldots, k,$$

tales que

$$\lambda_1 D_1 + \lambda_2 D_2 + \ldots + \lambda_k D_k = O.$$

Esta es una igualdad entre vectores de $\mathbb{K}^n$. Examinemos las entradas que ocupan los lugares

$$i_l, \quad l = 1, 2, \ldots, k,$$

(que corresponden a las variables libres) en el miembro de los izquierda. Sabemos que

$$d_{i_l j} = 0$$

salvo cuando $j = i_l$, en cuyo caso se tiene

$$d_{i_l i_l} = 1.$$

Por lo tanto en la i_l-ésima entrada de la combinación lineal sólo aparece λ_l. Resulta entonces

$$\lambda_l = 0, \quad l = 1, 2, \ldots, k,$$

, lo que muestra la independencia lineal de la familia $\mathcal{D}$. □

1.7.3. Las dimensiones del núcleo y del espacio de columnas

Tal como se muestra en la proposición 1.20 nuestro procedimiento para hallar una base del núcleo genera un vector de la base por cada una de las variables libres. Por lo tanto, una base del núcleo tendrá tantos vectores como variables libres haya en la forma escalerizada y la dimensión del núcleo será igual al número de variables libres.

Observemos que cuando no hay variables libres el núcleo de A se reduce al subespacio trivial formado por la lista de n ceros, que es la única solución de $AX = O$. La convención de asignar dimensión cero al subespacio trivial hace que la dimensión del núcleo coincida con el número de variables libres también en este caso.

Las variables libres corresponden a columnas que no tienen pivotes. Entonces el número de variables libres es la diferencia $n - p$ entre la cantidad n de variables, o columnas en la matriz del sistema, y el número p de pivotes.

Concluimos que la dimensión del núcleo de A es igual a $n - p$.

Recordemos que podemos fabricar una base del espacio de columnas de una matriz A seleccionando las columnas de A que corresponden a las posiciones de los pivotes de una forma escalerizada de A. Una base de $\mathrm{col}(A)$ tendrá entonces p vectores, tantos como pivotes. Por lo tanto la dimensión del espacio de columnas es p, un resultado que es el contenido de la proposición 1.14, en la página 126 de la sección 1.6.

Recojamos nuestros últimos comentarios en el siguiente teorema acerca de las dimensiones del espacio de columnas y el núcleo.

Teorema 1.1 (Teorema de las dimensiones). *Sea A una matriz $m \times n$ sobre un cuerpo $\mathbb{K}$. Y sea p el número de pivotes de una forma escalerizada de A. Entonces*

1. $\dim(\mathrm{col}(A)) = p$;
2. $\dim(\ker(A)) = n - p$;
3. $\dim(\ker(A)) + \dim(\mathrm{col}(A)) = n$.

Ejercicio 1.75. Verificar el resultado del teorema 1.1 para las matrices de los ejercicios 1.72 y 1.73.

Observación 1.7.7. La última igualdad del teorema 1.1 es sumamente interesante en términos de la interpretación de la matriz A como una transformación. De hecho, es un caso particular de uno de los resultados fundamentales de la teoría de las transformaciones lineales.

Habíamos visto en la observación 1.3.16 que la matriz A define sobre $\mathbb{K}^n$ una función

$$X \mapsto AX$$

que toma valores en $\mathbb{K}^m$. El espacio de columnas de A es la imagen de esta transformación, de modo que se cumple

$$\dim(\ker(A)) + \dim(\mathrm{im}(A)) = n.$$

Es decir, la dimensión n del espacio de salida de la transformación es la suma de la dimensión del núcleo más la de la imagen.

En el núcleo está todo lo que tiene como imagen al vector O de $\mathbb{K}^m$, así que podemos pensar que el núcleo se "comprime" yendo todo al $O \in \mathbb{K}^m$. ¡Hemos perdido un subespacio de la dimensión del núcleo en esto! Pero la transformación es todavía capaz de producir un subespacio de dimensión

$$\dim(\mathrm{im}(A)) = n - \dim(\ker(A))$$

como su imagen.

Esta transformación "respeta las dimensiones": las que no se gastaron en el núcleo reparecen en la imagen. Y la suma es el valor n que es la dimensión del espacio de salida.

1.7.4. Para tener presente

- El núcleo de una matriz A es el conjunto de soluciones de la ecuación homogénea $AX = O$.
- El núcleo de una matriz $m \times n$ es un subespacio vectorial de $\mathbb{K}^n$.
- Todas las soluciones de $AX = B$ se obtienen como suma de una solución particular y soluciones de $AX = O$.
- El núcleo de una matriz es el mismo que el de cualquiera de sus formas escalerizadas.
- La dimensión del núcleo de A es igual a la cantidad de variables libres en cualquiera de sus formas escalerizadas.
- Si A es una matriz $m \times n$ entonces las dimensiones de su núcleo y de su espacio de columnas, o imagen, suman n.

1.8. Espacio de filas y rango

Hemos interpretado la resolución de sistemas lineales en términos de combinaciones lineales de las columnas de la matriz del sistema, y el estudio de los sistemas lineales ha revelado una rica estructura que, a su vez, da información acerca de las soluciones del sistema. Surgieron así el espacio de columnas y el núcleo de una matriz. En particular, el espacio de columnas apareció porque a cada incógnita del sistema de ecuaciones se le asocia, en forma completamente natural, una columna de la matriz del sistema.

Pero cuando operamos con el sistema empleando el método de escalerización no son las columnas las que manipulamos, sino las filas de la matriz. Si el lector vuelve a la observación 1.2.21, en la página 44, recordará que fue a través de estas operaciones con las filas que introdujimos la idea de que las listas de números podían ser consideradas como elementos de una estructura algebraica, los espacios $\mathbb{K}^m$, que es la que sustenta toda la teoría que hemos venido desarrollando a lo largo de este capítulo.

En esta sección consideraremos el espacio generado por las filas de una matriz. Mostraremos que su dimensión coincide con la del espacio de columnas, y llamaremos *rango* de una matriz la dimensión de estos dos subespacios. Veremos que los posibles comportamientos de un sistema lineal $AX = B$ quedan completamente caracterizados por el número de ecuaciones y de incógnitas que tiene, y por el rango de la matriz A del sistema. Es decir, por el tamaño $m \times n$ de la matriz, y su rango.

1.8.1. El espacio de filas de una matriz

Al igual que hicimos para las columnas, podemos considerar el conjunto formado por todas las posibles combinaciones lineales de las filas de A, tal como recogemos en nuestra próxima definición.

Definición 1.11. *Si A es una matriz $m \times n$ sobre $\mathbb{K}$ llamaremos* ***espacio de filas*** *de A al subconjunto de $\mathbb{K}^n$ formado por todas las posibles combinaciones lineales de las filas de la matriz.*

Indicaremos este conjunto con la notación $\text{fil}(A)$.

Observación 1.8.1. SUBESPACIO GENERADO
A esta altura podemos dar un paso más en la teoría y reconocer un procedimiento general de construcción de subespacios a partir de lo hecho con los espacios de filas y de columnas de una matriz. La definición del espacio de filas es el análogo para las filas de la definición 1.4, página 92, del espacio de columnas.

Ambas definiciones son casos particulares de la construcción que consiste en tomar todas las posibles combinaciones lineales de los vectores de un conjunto $\mathcal{A}$ contenido en $\mathbb{K}^n$. Lo que justifica la siguiente definición.

Definición 1.12 (Subespacio generado). *Si $\mathcal{A}$ es un conjunto de vectores en $\mathbb{K}^n$, llamaremos **subespacio generado por** $\mathcal{A}$ al conjunto formado por todas las posibles combinaciones lineales de los vectores en $\mathcal{A}$.*

Ejemplo 1.8.2. Si A es una matriz entonces $\mathrm{col}(A)$ es el subespacio generado por las columnas de A, y $\mathrm{fil}(A)$ al subespacio generado por las filas de A.

Nuestro próximo ejercicio muestra que los subespacios generados por este procedimiento son, efectivamente, subespacios vectoriales de $\mathbb{K}^n$. También que los términos que estamos usando se combinan armoniosamente.

Ejercicio 1.76. Consideremos un subconjunto $\mathcal{A} = \{A_1, A_2, \dots, A_k\} \subset \mathbb{K}^n$.

1. Mostrar que el conjunto de todas las posibles combinaciones lineales de elementos de $\mathcal{A}$ es un subespacio vectorial de $\mathbb{K}^n$.
2. Mostrar que $\mathcal{A}$ es un generador del subespacio generado por $\mathcal{A}$.
3. Mostrar que si $\mathcal{A}$ es un generador de un cierto subespacio $\mathbb{S}$ contenido en $\mathbb{K}^n$, entonces el subespacio generado por $\mathcal{A}$ es justamente $\mathbb{S}$. ♠

Las operaciones necesarias para llevar la matriz A a su forma escalerizada generan nuevas filas que están en el espacio de filas de la matriz original, porque se obtienen haciendo combinaciones lineales con las filas de A.

Más interesante aún es que la escalerización no modifica el espacio de filas de la matriz. Mostraremos esta afirmación verificándola para cada una de las operaciones elementales del proceso de escalerización. Recordemos cuáles eran estas operaciones elementales:

1. sumar a una fila un múltiplo de otra;
2. multiplicar una fila por un escalar no nulo;
3. intercambiar dos filas.

Naturalmente, si aplicar una transformación elemental produce una matriz con el mismo espacio de filas que la matriz original también es cierto que el espacio de filas no cambiará luego de aplicar una cantidad finita de operaciones elementales. Por lo tanto, no cambiará a lo largo del proceso de escalerización.

Proposición 1.21. *Sea A una matriz $m \times n$ sobre $\mathbb{K}$, y B una matriz que se obtiene a partir de A aplicando una de las transformaciones elementales del algoritmo de escalerización. Entonces*

$$\text{fil}(A) = \text{fil}(B).$$

Prueba. Todo lo que tenemos que demostrar es que si una fila puede escribirse como combinación lineal de las filas de A entonces también puede escribirse como combinación lineal de las filas de B, y viceversa. Para ahorrarnos la segunda comprobación hacemos la siguiente observación: si B se obtiene de A por medio de una transformación elemental entonces también es cierto que A se puede obtener de B por una transformación elemental. Para mostrar que esto es así separaremos la prueba en tres casos. Uno para cada transformación elemental. Indicaremos con

$$A_i,\ B_i,\quad i = 1, 2, \dots, m,$$

a las filas de A y B respectivamente.

1. Cuando a una fila se le suma un múltiplo de otra. Supongamos que la matriz B se obtiene sumando a la fila j de A el resultado de multiplicar la fila k, con $k \neq j$ por el número a. Entonces la j-ésima fila de B es
$$B_j = A_j + aA_k,$$
mientras que todas las filas restantes coinciden con la correspondiente fila de A. En particular, $B_k = A_k$. Podemos hacer entonces sobre las filas de B la operación que deshace la acción de sumar aA_k en la fila j: si a la fila j de B le restamos aB_k tenemos
$$B_j - aB_k = A_j + aA_k - aA_k = A_j,$$
y recuperamos la fila A_j. Las restantes filas de A ya coincidían con las filas de B, de modo que hemos conseguido la matriz A haciendo una transformación elemental a la matriz B.

2. Cuando se multiplica una fila por un escalar no nulo. La transformación elemental que hay que aplicar a B para obtener la matriz A consiste en multiplicar la misma fila por el inverso del escalar.

3. Cuando se intercambian dos filas. Podemos deshacer la operación aplicando una vez más el mismo intercambio.

Ahora sólo tenemos que mostrar que si una fila está en el espacio de columnas de B entonces también está en el de A. Consideramos entonces una fila F que es una combinación lineal de las filas de B. Por lo tanto podremos expresar F en la forma

$$F = \lambda_1 B_1 + \ldots + \lambda_i B_i + \ldots \lambda_m B_m,$$

para algún conjunto de escalares λ_i, $i = 1, \ldots, m$. Nuevamente, separamos la demostración en tres casos.

1. Cuando a una fila se le suma un múltiplo de otra. Como antes, supongamos que a la fila A_j le hemos sumado aA_k para obtener B_j. Tomando en cuenta la expresión de las filas de B en términos de las filas de A podemos escribir F en la forma

   $$F = \lambda_1 A_1 + \ldots + \lambda_j (A_j + aA_k) + \ldots \lambda_m A_m,$$

   que es una combinación lineal de las filas de A. Si el lector lo prefiere puede reordenarla un poco, para poner junto todo lo que multiplica a la fila A_k,

   $$F = \lambda_1 A_1 + \ldots + \lambda_j A_j + \ldots (\lambda_k + a\lambda_j) A_k + \ldots \lambda_m A_m,$$

   pero no es esencial hacerlo.

2. Cuando se multiplica una fila por un escalar no nulo. Si hemos multiplicado a la fila j de A por un escalar a, entonces $B_j = aA_j$ y

   $$F = \lambda_1 A_1 + \ldots + (a\lambda_j) A_j + \ldots + \lambda_m A_m.$$

3. Cuando se intercambian dos filas. Las filas de A y de B son exactamente las mismas. Sólo cambia el orden con el que aparecen en la combinación lineal. □

Si aplicamos repetidas veces la proposición anterior (tantas como operaciones elementales sean necesarias para el proceso de escalerización) obtenemos nuestro próximo corolario.

Corolario 1.22. *Si E es una forma escalerizada de la matriz A entonces los espacios de filas de E y A son iguales.*

Observación 1.8.3. Es importante notar que el espacio de columnas cambia a lo largo del proceso de escalerización. Un ejemplo sencillo mostrará esto.

Ejemplo 1.8.4. El espacio de columnas asociado con la matriz real

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

está formado por todas las columnas (x_1, x_2) de $\mathbb{R}^2$ tales que $x_1 = x_2$. El espacio de filas por las filas (x_1, x_2) que satisfacen $x_2 = 2x_1$. La forma escalerizada reducida de A es

$$E = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix},$$

que, naturalmente, tiene el mismo espacio de filas. Pero el espacio de columnas de E está formado por todos los vectores de la forma $(x, 0)$, donde x es un número real cualquiera. ♣♠

Una consecuencia de nuestro último corolario es que las filas no nulas de una forma escalerizada E de una matriz A generan el espacio de filas de A. Esto es cierto porque las filas no nulas de E generan exactamente lo mismo que todas las filas de E, ya que las filas nulas no aportan nada en una combinación lineal: sólo ceros. Y todas las filas de E generan el espacio de filas de E, que coincide con el de A.

Además, las filas no nulas de una matriz escalerizada forman una familia linealmente independiente. Veamos por qué. Sean

$$E_1, E_2, \ldots, E_p,$$

las p filas no nulas, donde p es el número de pivotes de la forma escalerizada. Fabriquemos una combinación lineal de las filas e igualemos al vector nulo. Es decir, consideremos coeficientes λ_i, $i = 1, \ldots, p$ tales que

$$\lambda_1 E_1 + \lambda_2 E_2 + \ldots + \lambda_p E_p = O. \tag{1.56}$$

Ahora examinemos el lugar que corresponde al pivote de E_1. Este pivote está en una fila a la que llamaremos j_1 ¿Qué otras filas aportan algo en este lugar? Ninguna. Sólo E_1 lo hace, porque el pivote e_{ij_1} en E_1 es distinto de 0, y las filas restantes tienen ceros en la columna j_1 de ese pivote. Al igualar a cero esta componente de la combinación lineal obtenemos

$$\lambda_1 e_{ij_1} = 0,$$

lo que implica $\lambda_1 = 0$.

Una vez que $\lambda_1 = 0$ el primer sumando en el término de la izquierda de (1.56) es nulo y esa igualdad se reduce a

$$\lambda_2 E_2 + \ldots + \lambda_p E_p = 0.$$

Podemos repetir el razonamiento que acabamos de hacer, pero basándonos en la columna j_2 en que está el pivote de E_2 para concluir que $\lambda_2 = 0$. Aplicando el argumento p veces mostramos que la igualdad (1.56) implica

$$\lambda_i = 0, \quad i = 1, \ldots, p.$$

Por lo tanto, la familia formada por las filas no nulas de una matriz escalerizada son linealmente independientes.

El lector estará de acuerdo en que hemos demostrado la proposición que enunciaremos a continuación:

Proposición 1.23. *Sean A una matriz y E una forma escalerizada de A. Entonces las filas no nulas de la matriz E forman una base de* $\mathrm{fil}(A)$.

Ejemplo 1.8.5. En este ejemplo vamos a calcular escalerizando bases del espacio de filas de una matriz. Usaremos la matriz del sistema del ejemplo 1.2.22 de la página 45:

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 5 & 3 \\ 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 \end{pmatrix}$$

En ese ejemplo encontramos una forma escalerizada

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

de A. A partir de esta matriz encontramos la base

$$\mathcal{B}_1 = ((1,2,1,0,1,1),(0,0,2,1,1,0),(0,0,0,2,4,2),(0,0,0,0,0,2))$$

del espacio de filas de A.

Más tarde, en el ejemplo 1.2.30 de la página 61, hallamos una forma escalerizada reducida de ese mismo sistema, con matriz

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 3/2 & 0 \\ 0 & 0 & 1 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Las filas no nulas de esta nueva matriz proporcionan otra base de $\text{fil}(A)$ distinta a la anterior:

$$\mathcal{B}_2 = ((1,2,0,0,3/2,0),(0,0,1,0,-1/2,0),(0,0,0,1,2,0),(0,0,0,0,0,1)).$$

Todas las filas de A se pueden escribir de manera única como combinación lineal de los elementos de $\mathcal{B}_1$ o de los elementos de $\mathcal{B}_2$.

Ejercicio 1.77. Escribir la última fila de A como combinación lineal de los vectores en las bases $\mathcal{B}_1$ y $\mathcal{B}_2$

Ejercicio 1.78. Hallar una base del espacio de filas de las siguientes matrices:

1. $\begin{pmatrix} 1 & 0 & -1 & 2 & -2 \\ 2 & 1 & -1 & 3 & -2 \\ 1 & 2 & 1 & 0 & 2 \\ -1 & 3 & 4 & -5 & 8 \end{pmatrix}$
2. $\begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & -1 & 7 & 0 \\ -1 & 1 & -5 & -1 \end{pmatrix}$

Observación 1.8.6. Un adelanto de la descomposición LU Toda la información necesaria para escribir las filas de una matriz A como combinación lineal de las filas de su forma escalerizada se genera durante el proceso de eliminación gaussiana. En la sección 2.4 aprenderemos a almacenar sistemáticamente esta información, en un procedimiento conocido como *descomposición LU* que se traduce en la descomposición de A en dos factores L y U más simples[13]. Este procedimiento es de importancia para la resolución sistemática de sistemas lineales de ecuaciones.

Ejercicio 1.79. Consideremos la matriz real

$$A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -3 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

[13]En realidad U es una forma escalerizada de A. En la sección 2.4 se explica por qué la designamos con la letra U.

Llevar la matriz A a su forma escalerizada, a esta matriz la llamaremos U. Luego hallar los coeficientes que permiten escribir las filas de A como combinación lineal de las filas de U. ¿Qué relación hay entre los coeficientes hallados y los multiplicadores que se usaron durante el proceso de escalerización? ♠

El método de escalerizar para conseguir bases del espacio de filas puede emplearse para obtener una base del espacio de columnas de una matriz A: sólo hay que considerar una nueva matriz que tiene como filas a las columnas de A. A esta matriz se le llama la *matriz traspuesta* de A, y se indica con la notación A^t.

Ejemplo 1.8.7. A continuación presentamos dos ejemplos de matrices traspuestas:

1. Sea $A = \begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & 1 \\ -1 & 2 & 4 \end{pmatrix}$ entonces $A^t = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ 0 & 1 & 4 \end{pmatrix}$

2. Sea $B = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 3 \end{pmatrix}$ entonces $B^t = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \end{pmatrix}$

Ejercicio 1.80. Hallar una base del espacio de columnas de la matriz A del ejemplo 1.8.5 escalerizando su matriz traspuesta. Expresar los elementos de la base

$$((1,0,0,0,0),(0,1,0,0,0),(0,0,1/2,1,0),(0,0,1/2,0,1))$$

del espacio de columnas de A que se obtuvo en el ejemplo 1.7.6 como combinaciónes lineales de esta nueva base.

1.8.2. Rango

Los espacios de filas y columnas de una matriz A tienen exactamente la misma dimensión. Para mostrar este hecho sólo tenemos que revisar lo que hemos aprendido analizando el procedimiento de escalerización. Una vez probado, definiremos el rango como la dimensión común a estos dos subespacios.

Vimos en la proposición 1.23 que si E es una forma escalerizada de A entonces las filas no nulas de E son una base de $\text{fil}(A)$. Recordemos que hay tantas filas no nulas en E como pivotes. La proposición 1.23 tiene entonces el siguiente corolario:

Corolario 1.24. *Sean A una matriz $m \times n$ sobre un cuerpo $\mathbb{K}$, y sea p el número de pivotes de una forma escalerizada de A. Entonces* $\dim(\text{fil}(A)) = p$.

También la dimensión del espacio de columnas es igual al número p de pivotes en una forma escalerizada. Este resultado se recoge en la proposición 1.14, página 126, y fue revisado en la discusión que precede al Teorema de las dimensiones, en la página 140.

Juntando toda la información demostramos el siguiente teorema.

Teorema 1.2 (Teorema del rango). *Sea A una matriz $m \times n$ sobre un cuerpo $\mathbb{K}$. Entonces*

$$\dim(\text{fil}(A)) = \dim(\text{col}(A)).$$

Definición 1.13. *Llamaremos* ***rango*** *de una matriz a la dimensión de su espacio de filas y de su espacio de columnas.*

Ejemplo 1.8.8. El rango de la matriz del ejemplo 1.8.5 es igual a 4. Una forma escalerizada de esta matriz aparece, por ejemplo, en la página 62 y tiene cuatro pivotes.

Ejercicio 1.81. Calcular el rango, hallar una base del espacio de filas y una base del espacio de columnas de la matriz real

$$C = \begin{pmatrix} 1 & 1 & 1 & 0 & 2 & -4 \\ 1 & 1 & 1 & 0 & 2 & -4 \\ 1 & 2 & 3 & 1 & 4 & -6 \\ 1 & 2 & 3 & 1 & 4 & -6 \\ 1 & -1 & -3 & -2 & -2 & 0 \end{pmatrix}.$$

La compatibilidad de un sistema de ecuaciones lineales puede discutirse en términos de esta noción de rango.

Proposición 1.25. Teorema de Roché-Frobenius
Un sistema lineal $AX = B$ es compatible si y sólo si la matriz A y la matriz ampliada $A|B$ tienen el mismo rango.

Ejercicio 1.82. Demostrar el teorema de Roché-Frobenius.

Ejercicio 1.83. Para la matriz real A y los vectores B_1, B_2, dados por

$$A = \begin{pmatrix} 1 & 1 & -2 & 1 \\ 2 & 1 & -1 & 1 \\ 2 & 2 & -4 & 2 \\ 3 & 2 & -3 & 2 \end{pmatrix}, \quad B_1 = \begin{pmatrix} -1 \\ 2 \\ -2 \\ 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 0 \end{pmatrix},$$

calcular el rango de la matriz A y de las matrices ampliadas $A|B_i$, $i = 1, 2$. Decidir cuáles de estas matrices corresponden a sistemas compatibles. Sugerencia: tal vez sea útil recordar que la matriz A y los vectores que la acompñan aparecieron en la parte 1 del ejercicio 1.53, en la página 109.

El próximo ejercicio sólo obliga a una revisión no muy complicada de lo hecho a lo largo de este capítulo, pero nos interesa enunciar el resultado porque haremos uso de el en la sección dedicada al estudio de los determinantes.

Ejercicio 1.84. Mostrar que el rango de una matriz cuadrada $n \times n$ es igual a n si y sólo si todas las entradas

$$e_{ii}, \quad i = 1, \dots, n,$$

que están en la diagonal principal de cualquier forma escalerizada E de A son no nulas.

Para despedirnos del capítulo dedicado específicamente a los sistemas lineales de ecuaciones, prepararemos una especie de resumen final de todo el trabajo acumulado.

Ejercicio 1.85. Sea A una matriz $m \times n$ con rango r. Discutir en términos de los valores de r, m y n todos los posibles comportamientos que puede tener un sistema $AX = B$. Dar un ejemplo de cada posible comportamiento detectado.

1.8.3. Para tener presente

- El espacio de filas de una matriz es el subespacio generado por sus filas.

- Las filas no nulas de una forma escalerizada de A forman una base de su espacio de filas.

- Las dimensiones de los espacios de filas y de columnas coinciden. A la dimensión de estos subespacios se le llama rango de la matriz.

- Los posibles comportamientos de un sistema de ecuaciones $AX = B$ pueden discutirse completamente en términos del tamaño de la matriz A del sistema, y de su rango.

- Si se conoce el rango de una matriz las dimensiones de su núcleo y de su imagen, o espacio de columnas, quedan completamente determinadas.

- **Toda la estructura lineal de los espacios $\mathbb{K}^n$ y sus subespacios, las nociones de combinación lineal, generadores, independencia lineal, bases y dimensión, permiten comprender el comportamiento de los sistemas lineales en términos de unos pocos parámetros: la cantidad de incógnitas en el sistema, la cantidad de ecuaciones, y el rango de la matriz del sistema.**

- **Las cuestiones de compatibilidad y determinación de los sistemas lineales se contestan analizando dos subespacios vectoriales asociados a la matriz del sistema: el espacio de columnas y el núcleo. La dimensión de estos subespacios determina las propiedades del sistema. Conocer el rango de la matriz del sistema es esencialmente equivalente a determinar estas dos dimensiones.**

Capítulo 2

Matrices

En el capítulo 1 hemos introducido las matrices como una herramienta adecuada para almacenar y manipular en forma ordenada los coeficientes de los sistemas de ecuaciones lineales. En este capítulo las matrices ocuparán el centro de la escena y su papel irá mucho más allá que el de ser un artificio útil para representar los sistemas de ecuaciones.

Repasar nuestro trabajo con los sistemas de ecuaciones puede ayudarnos a entender el enfoque general que guiará este capítulo. Para los sistemas de ecuaciones $AX = B$, donde A es una matriz $m \times n$, enfatizamos que se puede ordenar sus coeficientes en filas y columnas e interpretar sus soluciones como vectores de $\mathbb{K}^n$. Fue así que introdujimos la estructura algebraica y geométrica[1] de $\mathbb{K}^n$, y obtuvimos como recompensa una mejor comprensión de las soluciones del sistema y los problemas de compatibilidad y determinación.

Con las matrices recorreremos un camino similar. Veremos cómo estos arreglos de números en filas y columnas pueden ser dotados de una estructura algebraica que permitirá sumarlos, y multiplicarlos entre sí y por escalares. Esta estructura algebraica nos permitirán manipular las matrices de diversas maneras.

A esta altura es posible que el lector se esté preguntando ¿por qué habría yo de mostrar algún interés por la manipulación de matrices, teniendo en el mundo tantas cosas más interesantes sobre las que poner mis manos? Conmovidos por esta hipotética duda de nuestros hipotéticos lectores, hemos incluido en la sección 2.1 unos cuantos ejemplos de problemas en los que toda la información relevante se ordena eficientemente en forma de una matriz, con la que se hace necesario calcular para poder resolver el problema original.

Las operaciones de suma de matrices, de producto de una matriz por un escalar y de producto entre matrices serán objeto de definiciones precisas que aparecerán en la sección 2.2. Los cálculos de la sección 2.1 en realidad anticipan estas definiciones y podrían posponerse para después de la lectura de la sección 2.2, pero nuestra recomendación es que el lector siga el ordenamiento del texto, y se enfrente a los primeros ejemplos guiado por su intuición y la lógica de la situación con la que está trabajando. Creemos que por esta vía las definiciones adquirirán mayor significado.

La sección 2.3 está dedicada al problema de hallar la matriz inversa A^{-1} de una matriz cuadrada A. Hay al menos dos maneras muy naturales de asomarse a esta cuestión:

- en la sección 2.2 veremos que las matrices cuadradas tienen una estructura multiplicativa. En este mundo en el que se puede multiplicar, ¿se

[1] Desarrollaremos más lo que tiene que ver con la geometría en el capítulo 3. Haremos énfasis en el caso particular de $\mathbb{R}^3$.

puede dividir?;

- en las secciones 2.1 y 2.2 se enfatiza la acción de una matriz como una transformación: "las matrices hacen cosas". Nos preguntamos entonces ¿cómo se invierten estas transformaciones? ¿cómo se "deshace" lo que la matriz hace?

En la sección 2.4 volvemos sobre una idea recurrente en este texto: las matrices pueden ser una manera utilísima de almacenar información de muy diversa naturaleza. La descomposición LU de una matriz A guarda para nosotros toda la información que se genera al resolver un sistema $AX = B$ por medio del procedimiento de eliminación gaussiana. Se introduce aquí una idea muy importante para el cálculo: buscar una buena factorización de una matriz es un procedimiento eficiente para resolver gran cantidad de problemas.

Por último, la sección 2.5 está dedicada al estudio de los determinantes de matrices cuadradas. Hemos comprimido todo el material sobre determinantes en una única sección muy extensa, que no es necesario asimilar completamente para seguir adelante. Remitimos al lector a la lectura de la introducción de la sección 2.5, donde encontrará una descripción del material que allí se cubre, y una guía para su lectura.

2.1. Las matrices almacenando y transformando información

En esta sección presentamos una serie de ejemplos y ejercicios en los que algunas matrices recogen, ordenan y transforman información relativa a diversos problemas, en diferentes contextos. Trabajaremos sobre estos problemas, lo que nos llevará a realizar distintas operaciones algebraicas con los números almacenados en las matrices que construiremos. Estos cálculos constituyen, en realidad, un adelanto de las operaciones entre matrices que introduciremos en la sección 2.2, y esperamos que sirvan al lector como una motivación que justifique las definiciones que allí se darán, y como un ejemplo ilustrativo de sus posibles aplicaciones. La lista de actividades que hemos incluido en esta sección no es breve. Esperamos que cada lector identifique las más adecuadas a su gusto e inclinaciones. Tampoco es exhaustiva, en realidad está muy lejos de serlo: muchos ejemplos interesantes han quedado fuera, lo que indica la aplicabilidad de la teoría que desarrollaremos a lo largo de este texto.

2.1.1. Transformaciones definidas por matrices

En esta sección veremos algunos ejemplos de transformaciones que aparecen en diversos contextos, y que son representadas por medio de una matriz.

Ejemplo 2.1.1. Transiciones en una población.
En el ejemplo 1.3.17 del capítulo 1 vimos como las migraciones de la población de un país se describían con una matriz A. Si un año comenzaba con la población distribuida en tres posibles estados: habitantes en la *c*iudad, en zonas *r*urales, o en el *e*xtranjero, según los porcentajes almacenados en el vector $X = (c, r, e)$, la distribución un año más tarde estaba dada por el vector AX, resultado de multiplicar la matriz

$$A = \begin{pmatrix} 0{,}6 & 0{,}4 & 0{,}1 \\ 0{,}1 & 0{,}4 & 0{,}1 \\ 0{,}3 & 0{,}2 & 0{,}8 \end{pmatrix} \tag{2.1}$$

por el vector X. Esta regla define entonces cómo evolucionará el sistema, y debería permitirnos predecir todo el comportamiento futuro. Por lo menos, esto es así mientras el modelo para las transiciones que está implícito en la fórmula (2.1) siga siendo válido. Aparecen entonces preguntas interesantes cuya respuesta podemos buscar en nuestro modelo: ¿cuál será la evolución futura? A la larga ¿se irá todo el mundo al extranjero? ¿se estabilizará la población en algún valor? Si todo el mundo se va de las zonas rurales, ¿cuánto tiempo tenemos que esperar hasta que su población caiga por debajo de un nivel que vuelva insostenible algunas de las actividades económicas de esta zona?, etcétera. Comenzaremos por tratar una pregunta mucho más modesta: si comenzamos con una distribución $(c, r, e)^t$, ¿cuál será el estado de la población dos años más tarde?

Dado que tendremos que tratar con dos años diferentes modifiquemos ligeramente nuestra notación: llamaremos $X_0 = (c, r, e)^t$ al estado inicial, y X_1 al estado un año más tarde. Nuestro objetivo será determinar un nuevo vector X_2 que represente lo que ocurrirá dos años después de nuestro año de partida.

Ya sabemos cómo determinar X_1. Es el resultado $X_1 = AX_0$ de multiplicar la matriz (2.1) por el vector X_0. Al hacer la cuenta obtenemos

$$X_1 = \begin{pmatrix} 0{,}6c + 0{,}4r + 0{,}1e \\ 0{,}1c + 0{,}4r + 0{,}1e \\ 0{,}3c + 0{,}2r + 0{,}8e \end{pmatrix}$$

Arribaremos el estado X_2 al dejar pasar un nuevo año a partir del momento en que alcanzamos el estado X_1, que ahora debe ser tomado como condición inicial de una evolución que nos llevará hasta $X_2 = AX_1$. Teniendo en cuenta cuál es la matriz A y nuestra expresión para X_1 concluimos que

$$X_2 = \begin{pmatrix} 0{,}6\times(0{,}6c+0{,}4r+0{,}1e)+0{,}4\times(0{,}1c+0{,}4r+0{,}1e)+0{,}1\times(0{,}3c+0{,}2r+0{,}8e) \\ 0{,}1\times(0{,}6c+0{,}4r+0{,}1e)+0{,}4\times(0{,}1c+0{,}4r+0{,}1e)+0{,}1\times(0{,}3c+0{,}2r+0{,}8e) \\ 0{,}3\times(0{,}6c+0{,}4r+0{,}1e)+0{,}2\times(0{,}1c+0{,}4r+0{,}1e)+0{,}8\times(0{,}3c+0{,}2r+0{,}8e) \end{pmatrix}.$$

Si dejamos de lado la pereza y operamos un poquito encontramos entonces que

$$X_2 = \begin{pmatrix} 0{,}43c + 0{,}42r + 0{,}18e \\ 0{,}13c + 0{,}22r + 0{,}13e \\ 0{,}44c + 0{,}36r + 0{,}69e \end{pmatrix}.$$

Luego de nuestro buen baño de matrices en el capítulo 1, reconoceremos que esta expresión puede ser escrita como el producto

$$X_2 = \begin{pmatrix} 0{,}43 & 0{,}42 & 0{,}18 \\ 0{,}13 & 0{,}22 & 0{,}13 \\ 0{,}44 & 0{,}36 & 0{,}69 \end{pmatrix} \begin{pmatrix} c \\ r \\ e \end{pmatrix},$$

de la matriz

$$B = \begin{pmatrix} 0{,}43 & 0{,}42 & 0{,}18 \\ 0{,}13 & 0{,}22 & 0{,}13 \\ 0{,}44 & 0{,}36 & 0{,}69 \end{pmatrix}$$

por el vector $X_0 = (c, r, e)^t$ con las condiciones iniciales[2]. Hemos encontrado entonces que la transición $X_0 \mapsto X_2$, que nos permite calcular el estado X_2 que se alcanza dos años más tarde de observar el estado inicial X_0, se resume en la acción de la matriz B que satisface

$$X_2 = BX_0 = A(AX_0)$$

para cualquier condición inicial X_0. Nuevamente ¡la información que andábamos buscando queda naturalmente ordenada en una matriz! Esta matriz B es lo que luego definiremos como el producto AA, o A^2, de la matriz A por sí misma. ♣

Ahora pasamos a un par de ejemplos de naturaleza algo diferente. Se trata de aplicaciones relacionadas con ordenar datos económicos: en el primer caso se trata de la información sobre ventas de determinados productos; en el segundo de las relaciones entre los insumos y los productos en una "cadena de producción".

Ejemplo 2.1.2. Una fábrica de ropa vende cuatro tipo de artículos: polleras, pantalones, camisas y buzos y tiene registros de las cantidades vendidas en los meses de setiembre, octubre, noviembre, diciembre y enero (estos dos últimos especialmente interesantes por las visitas de Papá Noel y los Reyes Magos[3]), para tres sucursales: **C**entro, **P**ocitos y **U**nión. A partir de todos estos datos interesa determinar la facturación total, y también poder obtener subtotales agregados por producto, mes o sucursal.

En la sucursal Centro se venden en setiembre 60 polleras, 130 pantalones, 70 camisas y 150 buzos.En el mes de octubre se venden 100, 200, 120 y 330, en noviembre las ventas son 150, 240, 130 y 510, en diciembre 129, 223, 148 y 947, mientras que en enero las ventas suman 88, 147, 105 y 430 para cada una de las prendas. En Pocitos se venden en setiembre 25, 30, 40 y 38, en octubre 50, 10, 32 y 60, en noviembre 75, 25, 46 y 30, en diciembre 12, 53, 70 y 56, y en enero 62, 14, 31 y 84 polleras, pantalones, camisas y buzos respectivamente. Por último, en la Unión se venden 25, 42, 8 y 19 en setiembre; 15, 24, 13 y 42 en octubre; 13, 47, 50 y 22 en noviembre; 23, 38, 26 y 32 en diciembre y 40, 25, 61 y 45 en enero para cada uno de los artículos antes mencionados.

[2]¡Qué no desespere el lector que no haya reconocido inmediatamente el producto! Tal vez las cosas lleven algo más de tiempo, y cada cuál tiene el suyo propio. En cualquier caso, esperamos que pueda verificar que la afirmación es correcta.

[3]Para no caer en inexactitudes impropias de un texto universitario debemos aclarar que todos ellos son, en realidad, los padres. Aún así, algunos de los redactores de este texto seguimos creyendo en los Reyes Magos

Muchos números, así que hagamos un esfuerzo por ordenarlos. Por ejemplo, los datos provenientes de la sucursal Centro pueden verse mejor si los escribimos así, alineados en filas y columnas:

	Polleras	Pantalones	Camisas	Blusas
Setiembre	60	130	70	150
Octubre	100	200	120	330
Noviembre	150	240	130	510
Diciembre	129	223	148	947
Enero	88	147	105	430

Naturalmente, la tabla que acabamos de armar es una matriz, a la que llamaremos C, de tamaño 5×4. Para cada una de las otras dos sucursales, Pocitos y Unión, podemos resumir la información disponible en las matrices

$$P = \begin{pmatrix} 25 & 30 & 40 & 38 \\ 50 & 10 & 32 & 60 \\ 75 & 25 & 46 & 30 \\ 12 & 53 & 70 & 56 \\ 62 & 14 & 31 & 84 \end{pmatrix}, \quad U = \begin{pmatrix} 25 & 42 & 8 & 19 \\ 15 & 24 & 13 & 42 \\ 13 & 47 & 50 & 22 \\ 23 & 38 & 26 & 32 \\ 40 & 25 & 61 & 45 \end{pmatrix},$$

que ya aparecen escritas en la forma que es convencional para este texto. Las entradas de las matrices P, C y U indican la cantidad de unidades vendidas de cada producto en cada mes de análisis.

Si no queremos los datos desglosados sucursal por sucursal, y sólo nos interesan los totales por mes y producto también podemos resumir la información en una matriz. Sólo tenemos que sumar para cada una de las entradas los datos de las tres sucursales. La operación aparece indicada y ordenada en la forma de la matriz

$$T = \begin{pmatrix} 60+25+25 & 130+30+42 & 70+40+8 & 150+38+19 \\ 100+50+15 & 200+10+24 & 120+32+13 & 330+60+42 \\ 150+75+13 & 240+25+47 & 130+46+50 & 510+30+22 \\ 129+12+23 & 223+53+38 & 148+70+26 & 947+56+32 \\ 88+62+40 & 147+14+25 & 105+31+61 & 430+84+45 \end{pmatrix},$$

que, naturalmente, no es otra cosa que

$$\begin{pmatrix} 110 & 202 & 118 & 207 \\ 165 & 234 & 165 & 432 \\ 238 & 312 & 226 & 562 \\ 164 & 314 & 244 & 1035 \\ 190 & 186 & 197 & 559 \end{pmatrix}.$$

A partir de T y los precios de venta de cada producto podemos determinar la facturación de cada mes. Si sabemos que las polleras cuestan \$250, los pantalones \$450, las camisas \$300 y los buzos \$350 tendremos que multiplicar las columnas de polleras, pantalones, camisas y buzos por las cifras \$250, \$450, \$300 y \$350 respectivamente. Luego sumar las columnas para sacar los totales por mes. Esta operación no es otra cosa que hacer el producto de la matriz T por el vector de precios de los productos, lo que arroja el siguiente vector F de facturaciones:

$$F = \begin{pmatrix} 110 & 202 & 118 & 207 \\ 165 & 234 & 165 & 432 \\ 238 & 312 & 226 & 562 \\ 164 & 314 & 244 & 1035 \\ 190 & 186 & 197 & 559 \end{pmatrix} \begin{pmatrix} 250 \\ 450 \\ 300 \\ 350 \end{pmatrix} = \begin{pmatrix} 226250 \\ 347250 \\ 464400 \\ 617750 \\ 385950 \end{pmatrix}.$$

Otros datos también pueden obtenerse operando sobre la matriz T. Por ejemplo, si nos interesa determinar la cantidad total de unidades vendidas de cada producto sólo tenemos que sumar las filas de T.

En este ejemplo la informació aparece mejor ordenada en forma de algunas matrices. Sumar matrices entrada a entrada tiene significado en este contexto. Digamos que esta es la manera en la que definiremos la suma de matrices cualesquiera. La transformación que permite pasar del vector de precios al que contiene las facturaciones discriminadas mes por mes no es otra cosa que realizar el producto de la matriz T con los totales de ventas. La operación de sumar las filas de T también puede representarse como el producto de la matriz fila $(1\ 1\ 1\ 1)$ por la matriz T, algo que veremos cuando introduzcamos el producto de matrices. ♠

Nuestro próximo ejemplo viene en la forma de un ejercicio en el que se pide calcular las matrices que relacionen distintas cantidades. El producto de matrices está implícito en los cálculos necesarios, tal como el lector podrá comprobar luego de la lectura de la sección 2.2.

Ejercicio 2.1. DE REGRESO A LA ESCUELA: UN PROBLEMA DE "ECONOMÍA DOMÉSTICA" CON REGLAS DE TRES Y MATRICES.
Es sábado, llueve, y la casa de la abuela se llenó de gente. Hay 13 niños y 9 adultos. Es hora de hacer pastelitos de dulce y tortas fritas.

1. Cada niño come 2 pastelitos y una torta frita; cada adulto 2 pasteles y tres tortas. Hallar la matriz que permite calcular las cantidades t y p de tortas y pastelitos en función de las cantidades n y a de niños y adultos presentes.
2. Hallar la matriz que permite calcular los ingredientes necesarios en función de las cantidades de cada producto que se desea preparar[4].

[4] TORTAS FRITAS (salen 12 tortas)

3. Hallar la matriz que permite calcular los ingredientes necesarios en función de la cantidad de personas presentes.
4. Averiguar los precios de cada ingrediente, y calcular el costo total de la merienda.

Antes de despedirnos de esta sección volvemos sobre un ejemplo básico, muy importante, que fue introducido en la página XI de la introducción a este texto, y con el que el lector seguramente trabajará al estudiar el cálculo diferencial para funciones de varias variables.

Ejemplo 2.1.3. MATRIZ JACOBIANA Y DIFERENCIAL
Consideremos una función de $\mathbb{R}^2$ en $\mathbb{R}^2$, que a cada par de números (x, y) le asocia un nuevo par, según el esquema

$$(x, y) \mapsto (f(x, y), g(x, y)).$$

Bajo condiciones muy generales, para valores pequeños Δx y Δy de los incrementos de las variables x e y, los incrementos Δf y Δg de f y g pueden aproximarse muy bien por expresiones lineales en Δx y Δy, que pueden resumirse en el producto de un vector por una matriz.

En el caso de la función definida por

$$f(x, y) = 1 + x^2 + xy, \quad g(x, y) = (x + 2y)^2. \tag{2.2}$$

hicimos en la introducción los cálculos para los incrementos alrededor de $x = 1$, $y = 2$. Ahora que conocemos la definición del producto de una matriz por un vector, podemos expresar los resultados que encontramos allí en la forma

$$\begin{pmatrix} \Delta f \\ \Delta g \end{pmatrix} \approx \begin{pmatrix} 4 & 1 \\ 10 & 20 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}.$$

Cernir $1\frac{1}{2}$ *cucharadita de sal* y 2 *tazas de harina*. Agregar $\frac{1}{4}$ *taza de grasa* blanda y $\frac{1}{2}$ *taza* de agua tibia. Mezclar la masa y amasarla hasta que quede suave. Formar las tortas a mano, tomando un pedacito de masa y estirándola redonda de 10 a 15 centímetros de diámetro. Freír en grasa caliente a 185 grados centígrados. Escurrir y conservar calientes.

PASTELITOS DULCES (salen 28 pastelitos)
Preparar una masa suave con 4 *tazas de harina*, 2 *cucharaditas de sal* y $\frac{1}{4}$ taza de grasa. Mezclar y agregar 2 *huevos batidos* y $\frac{1}{2}$ taza de agua. Amasar hasta que aparezcan pequeñas burbujas en la superficie. Estirar bien finita. Pintar con grasa derretida, espolvorear con harina y doblar. Estirar nuevamente. repetir la operación 4 veces. Cortar la masa en cuadrados de 10 centímetros. Poner un trozo de dulce de membrillo en el centro. Doblar y unir los bordes con agua cerca del relleno, para dejar las orillas sueltas. Freír en grasa abundante, a 184 grados centígrados. Escurrir en papel absorbente.

1 taza de grasa o azúcar $\approx$ 200 gramos; 3 cucharaditas $\approx$ 1 cucharada
16 cucharadas $\approx$ 1 taza = $\frac{1}{4}$ litro; 1 cucharada de dulce de membrillo $\approx$ 10 gramos

La matriz que hemos encontrado es lo que se llama *matriz jacobiana* de la función de varias variables en el punto $(1, 2)$. La transformación lineal que permite calcular la parte principal de los incrementos de f y g en función de los incrementos de las variables independientes es el *diferencial* de la función en ese punto.

Ejercicio 2.2.

1. Calcular la matriz jacobiana de la función (2.2) en cualquier punto (x, y).
2. DERIVADAS PARCIALES. Hallar la derivada de $f(x, y)$ respecto a x, cuando y se trata como una constante. Repetir el cálculo dejando x constante y derivando respecto a x. Luego repetir todo para la g.
3. Comparar con la matriz jacobiana las derivadas calculadas en la parte anterior. Interpretar los resultados. ♣

2.1.2. Transformaciones geométricas

Hemos visto en la sección anterior ejemplos en que distintas matrices representan transformaciones con distintos significados. En esta sección nos interesa destacar que es posible representar transformaciones geomeétricas por medio de matrices. Lo haremos a través de un ejercicio en el que giros, simetrías y proyecciones, y sus composiciones, serán descritos por matrices.

El contexto en el que trabajaremos es el que describimos en la página 84 –ver, en particular, la figura 1.7– cuando discutimos la interpretación geométrica de los sistemas de ecuaciones. Cada par (x_1, x_2) representa las coordenadas de un punto o de un vector respecto a un sistema ortogonal de coordenadas con un origen O –al que le corresponden las coordenadas $(0, 0)$–.

Ejercicio 2.3. TRANSFORMACIONES GEOMÉTRICAS EN EL PLANO.

1. LA MATRIZ DE UN GIRO.
   Para un número θ cualquiera consideramos la matriz
   $$G_\theta = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$
   a) Para $E_1 = (1, 0)^t$ y $E_2 = (0, 1)^t$ calcular GE_1 y GE_2. Interpretar geométricamente el resultado.
   b) Observar que cualquier vector $X = (x_1, x_2)^t$ puede expresarse como
   $$X = x_1E_1 + x_2E_2,$$
   y que al calcular GX usando esta expresión se obtiene
   $$GX = x_1GE_1 + x_2GE_2.$$

Interpretar geométricamente esta observación. Concluir que GX representa el resultado de girar X un ángulo θ en sentido antihorario.

c) Dados dos números θ y ψ, ¿cuál es el resultado Z de calcular $Y = G_\theta E_1$ y luego $Z = G_\psi Y$? ¿Cómo se interpreta esto geométricamente?

d) Comparar el resultado anterior con la acción de la matriz $G_{\theta+\psi}$. ¿Qué famosas fórmulas trigonométricas pueden deducirse de estas manipulaciones?

e) Hallar la matriz $G_{\theta+\psi}$ tal que la igualdad

$$G_{\theta+\psi}X = G_\theta\left(G_\psi X\right)$$

se satisface para todo $X \in \mathbb{R}^2$.

2. La matriz de la simetría respecto a la recta $x_1 = x_2$. Ahora consideramos la matriz

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Calcular el producto SX de S por cualquier vector $X = (x_1, x_2)^t$. En particular hacerlo para $(1,0)^t$, $(0,1)^t$, $(1,1)^t$ y $(1,-1)^t$, e interpretar geométricamente.

3. La matriz de la proyección sobre la recta $x_1 = x_2$. Consideremos

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

a) Calcular el producto SX de S por cualquier vector $X = (x_1, x_2)^t$. En particular hacerlo para $(1,0)^t$, $(0,1)^t$, $(1,1)^t$ y $(1,-1)^t$, e interpretar geométricamente.

b) Observar que cualquier $X = (x_1, x_2)^t$ puede expresarse como

$$X = \frac{x_1 + x_2}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{x_1 - x_2}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Calcular PX usando esta expresión, e interpretar el resultado.

c) Calcular el efecto de aplicar dos veces consecutivas la transformación $X \mapsto PX$ que consiste en multiplicar una columna X por la matriz P. Interpretar geométricamente.

4. Hallar la matriz que representa la composición de la simetría de la parte 2 con un giro de ángulo $\pi/4$ en sentido directo (o antihorario, o el que lleva de $(1,0)$ a $(0,1)$ por el camino más corto). Calcular también la representación matricial de la composición del giro con la simetría.

5. Calcular la composición de la simetría y la proyección de las partes 2 y 3. Componer también la proyección y la simetría.

Esperamos que luego de haber trabajado sobre el ejercicio anterior el lector haya quedado convencido de que es posible describir y operar con transformaciones geométricas por medio de una matriz. Este tipo de conexiones entre el Álgebra Lineal y la Geometría es esencial para ambas, y está en el corazón de muchas aplicaciones tecnológicas interesantes. Por ejemplo, la computación gráfica se sirve directamente de ella, y prácticamente todos los procedimientos eficientes de cálculo emplean profundas imágenes geométricas para ordenar la información numérica disponible.

2.1.3. Vectores propios

Las matrices definen transformaciones. En particular, las matrices cuadradas $n \times n$ definen transformaciones de un espacio vectorial $\mathbb{K}^n$ sobre sí mismo. Estas transformaciones pueden tener cierta complejidad, tal como hemos visto en variados ejemplos. Al multiplicar un vector X de $\mathbb{K}^n$ por una matriz A puede ocurrir cualquier cosa con X. Puede estirarse, acortarse, cambiar de dirección, quedar igual. Pero algunos vectores se transforman de manera especialmente simple. Ya conocemos un primer ejemplo.

Ejemplo 2.1.4. Todos los vectores X que están en el núcleo de una matriz A tiene la propiedad de que AX es igual al vector nulo. El conjunto de vectores con este comportamiento tan particular tiene además mucha estructura: se trata de un subespacio vectorial.

En muchos casos las matrices tienen una rica colección de vectores sobre los que actúan de manera muy simple. El objetivo de esta sección, estructurada en la forma de una serie de ejercicios, es llamar la atención del lector sobre este hecho, presentando los conceptos de *valores y vectores propios*, discutiendo algunas de sus aplicaciones, calculando con ellos, y comenzando a analizar algo de la geometría que encierran.

Ejercicio 2.4. Una matriz cuadrada A, de tamaño $n \times n$ define una transformación que a cada lista X de n números, n-upla o vector de $\mathbb{R}^n$, le asocia el producto AX. Esta acción de la matriz A puede ser bastante compleja, pero a veces ocurre que para algunos vectores particulares todo se vuelve más simple, porque el efecto de multiplicar por A es el mismo que el de multiplicar por un escalar λ. Es decir, $AX = \lambda X$. Si X es un vector no nulo con esta propiedad diremos que X *es un vector propio de* A que está asociado al *valor propio* λ.

1. Sea
$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$
Mostrar que $(1,1)^t$ y $(-1,1)^t$ son vectores propios asociados a los valores propios 4 y 2 respectivamente. Observar que $(1,0)^t$ y $(0,1)^t$ no son vectores propios.

2. Hallar los valores y vectores propios de
$$A = \begin{pmatrix} 5 & -3 \\ 6 & -4 \end{pmatrix}.$$
Sugerencia: la igualdad $AX = \lambda X$ es equivalente a un sistemas de ecuaciones que depende del parámetro λ. Discutir el comportamiento del sistema según el valor de λ, e identificar los valores de λ para los que el sistema tiene soluciones no triviales.

3. Supongamos que X es un vector propio de la matriz A, y que está asociado al valor propio λ. Es decir, se tiene $AX = \lambda X$. Calcular A^2X. Verificar la respuesta calculando directamente A^2 y A^2X para la matriz de la parte 1 de este ejercicio. En general, ¿cuál es el vector A^kX, para $k = 0, 1, 2 \ldots$?

Ejercicio 2.5. Cría de conejos y números de Fibonacci

1. Un hombre colocó una pareja de conejos en un sitio cercado. ¿Cuántos conejos fueron producidos a partir de esta pareja en el plazo de un año, si se supone que cada mes cada pareja fértil de conejos engendra una nueva pareja que se vuelve fértil a partir del mes de vida? Sugerencia: escribir el número x_{n+1} de conejos en el mes $n+1$ en forma recurrente, en términos del número de conejos en los meses x_n y x_{n-1}. Este problema está tomado[5] del texto Liber Abaci, de Leonardo Pisano "Fibonacci"($\approx$ 1170 - $\approx$ 1250)

2. La primera parte del ejercicio puede resolverse a partir de la recurrencia
$$x_{n+1} = x_n + x_{n-1}, \quad n = 1, 2, \ldots. \tag{2.3}$$
Se llama *números de Fibonacci* a los números f_n generados por esa relación, a partir de $f_0 = 0$, $f_1 = 1$. Calcular algunos números de Fibonacci e investigar su crecimiento. Por ejemplo, puede ser útil fabricar gráficos de (n, x_n) y $(n, \log x_n)$, y estudiar como crecen los cocientes x_{n+1}/x_n.

[5] He aquí el texto original:

Quot paria coniculorum in uno anno ex uno pario germinentur.

Quidam posuit unum par cuniculorum in quodam loco, qui erat undique pariete circondatus, ut sciret, quot ex eo paria germinarentur in uno anno: cum natura eorum sit per singulum mensem aliud par germinare; et in secundo mense ab eorum natiuitate germinant. Quia suprascriptum par in primo mense germinat, duplicabis ipsum, erunt paria duo in uno mense. Ex quibus unum, scilicet primum, in secundo mense geminat; et sic sunt in secundo mense paria 3; ex quibus, in uno mense duo pregnantur; et...

Aunque debemos admitir que las reglas según las que se reproducían los conejos de Fibonacci no son demasiado ajustadas a la realidad, en su descargo digamos que este ejercicio constituye el punto de partida hacia un mundo lleno de matemáticas interesantes y entretenidas. Ver, por ejemplo, el sitio `http://www.mscs.dal.ca/Fibonacci/` de "The Fibonacci Association". Ocurre que *Liber Abaci* era un libro de texto que se utilizaba para entrenarse en el uso de la aritmética. En muchos casos los problemas estaban tomados de fuentes anteriores, frecuentemente árabes o hindúes. Como vemos, la tradición de someter a los estudiantes a problemas arbitrarios, sólo salidos de la febril imaginación de los profesores es bastante antigua, ¡y fecunda! ¿Por qué habríamos de abandonarla?

3. Mostrar que las reglas que generan los números de Fibonacci pueden escribirse en forma matricial como

$$F_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad F_{n+1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} F_n, \quad n = 0, 1, 2, \dots,$$

donde

$$F_n = \begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} \qquad n = 1, 2, \dots.$$

En lo que sigue llamaremos A a la matriz 2×2 que acabamos de introducir. Al escribir las cosas de esta manera hemos conseguido que la recurrencia quede expresada sólo en términos del último paso. Veremos también que la expresión matricial nos ayudará en los cálculos.

4. El estudio del crecimiento de los f_n debería haber arrojado como resultado que $\log f_n$ parece crecer linealmente, o que los cocientes f_{n+1}/f_n parecen aproximarse a un cierto número λ. Es decir, $f_{n+1} \approx \lambda f_n$ para n grande y esa constante λ. En términos de la representación matricial esto implica $F_{n+1} \approx \lambda F_n$. Teniendo en cuenta que $F_{n+1} = AF_n$ resulta que $AF_n \approx \lambda F_n$, para n grande. Investigar si existen valores de λ para los que la igualdad $AX = \lambda X$ se satisface exactamente para algún vector X no nulo (es decir, investigar la existencia de valores y vectores propios de A).

5. Hallar las constantes α y β para que

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ \frac{\sqrt{5}-1}{2} \end{pmatrix} + \beta \begin{pmatrix} 1 \\ \frac{\sqrt{5}+1}{2} \end{pmatrix}. \tag{2.4}$$

A partir de esta expresión calcular F_n, $n = 1, 2, \dots$ y hallar una fórmula explícita para los números f_n. Hallar las constantes C y λ que hacen que $\lim_{n\to\infty} f_n/C\lambda^n = 1$.

6. Supongamos que los conejos del problema propuesto en la parte 1 siguieron reproduciéndose desde 1202, fecha en que apareció el texto *Liber Abaci* (ver nota 5, hasta ahora. Calcular el número de conejos que habría en la actualidad si los animalitos nunca murieran y siempre se reprodujeran siguiendo las reglas propuestas por Fibonacci.

Ejercicio 2.6. La solución del ejercicio anterior descansa en que pudimos expresar la condición inicial $(1, 0)$ como combinación lineal de dos vectores propios de la matriz. Con la misma idea se puede resolver el siguiente problema: si la población que se describe en el ejercicio 1.10 y el ejemplo 2.1.1 comienza un cierto año con el 70 % de sus individuos en la ciudad, el 30 % en zonas rurales y ninguno en el exterior ¿cuál será su distribución n años más tarde? ¿Cuál es la distribución en la que tiende a estabilizarse la población? Sugerencia: buscar los vectores propios de la matriz que define la evolución del sistema, y expresar la condición inicial como combinación de vectores propios. Ahora harán falta tres vectores propios.

Mostrar también que, independientmente de cuál sea la distribución inicial, la población tiende a distribuirse en los porcentajes que corresponden al estado estacionario.

El próximo ejercicio retoma el problema tratado en el ejercicio 1.25, de la página 58. Los vectores propios de la matriz del sistema de ese ejercicio nos ayudarán a comprender la reacción de las soluciones del sistema frente a perturbaciones de su término independiente.

Ejercicio 2.7. Número de condición (II)

1. Mostrar que los vectores $(1,1)$ y $(-1,1)$ son vectores propios de la matriz
$$A = \begin{pmatrix} 1 & 0{,}999 \\ 0{,}999 & 1 \end{pmatrix}.$$
Hallar los respectivos valores propios.
2. Calcular las soluciones de
$$AX = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad AX = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$
Sugerencia: no hace falta hacer ningún cálculo, sólo hay que aplicar la parte anterior del ejercicio.
3. Hallar la solución de
$$AX = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$
donde α y β son dos constantes reales cualesquiera.
4. Tener en cuenta la igualdad
$$\begin{pmatrix} 1{,}989 \\ 2{,}009 \end{pmatrix} = 1{,}999 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0{,}01 \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$
y usar las partes anteriores para entender los resultados que se encontraron al resolver el ejercicio 1.25.

Observación 2.1.5. La matriz A del ejercicio 2.7 actúa de manera muy diferente sobre los vectores $(1,1)$ y $(-1,1)$: en la dirección de $(1,1)$ se comporta como una multiplicación por un factor próximo a 2, pero a la dirección de $(-1,1)$ le corresponde un factor de tamaño 1/1.000, que es tres órdenes de magnitud más pequeño.

Este hecho explica el curioso comportamiento de las soluciones del sistema $AX = B$. Para deshacer la acción de A y recuperar el vector X a partir de

B hay que multiplicar por un factor $1/2$ en algunas direcciones, y por 1.000 en otras. Para direcciones intermedias hay que hacer cosas más complicada, por supuesto, porque no en todas partes la matriz actúa como una simple multiplicación por un escalar. En general cambia las direcciones de los vectores a los que se aplica.

El resultado combinado de estas dos escalas tan distintas que A guarda en su interior produce mucha inestabilidad en las soluciones de $AX = B$. El problema $AX = B$ es lo que se llama un problema *mal condicionado*. Definiremos más adelante el *número de condición* de una matriz como el cociente entre la máxima dilatación y la mínima dilatación que produce una matriz al aplicarla a los vectores del espacio. En el caso de A este cociente toma el valor $1.999 \approx 2 \times 10^3$.

Podemos estudiar el sistema $AX = C$, con

$$C = (1{,}989, 2{,}009)$$

considerando que C es una modificación

$$B + \Delta B$$

de

$$B = (1{,}999, 1{,}999)$$

con una perturbación

$$\Delta B = (-0{,}01, 0{,}01).$$

Notemos que el cambio que produce ΔB en las entradas de B es del orden del 0,5 %. Pero la perturbación ΔB está justo en la dirección en la que A multiplica por $1/1.000$, y se ve amplificada por un factor de 1.000 al resolver el sistema, para terminar produciendo un cambio del orden del 900 % en la solución del sistema.

En el capítulo dedicado a la Geometría, luego de haber introducido la noción de *longitud* de un vector, mediremos con ella los errores absolutos y relativos, y estudiaremos cómo el número de condición se relaciona con la propagación de los errores relativos del término independiente B en el sistema $AX = B$ a la solución X del sistema. ♠

2.1.4. Relaciones entre partes

Los ejercicios y ejemplos que cierran esta sección son de naturaleza combinatoria. El primero trata sobre grafos: un modelo para redes de todo tipo. Un *grafo* es un conjunto finito de puntos (*nodos*), en el que se establecen conexiones (*aristas*) entre algunos pares de ellos. Una manera formal de hacer esto es fijar un conjunto $\mathcal{N}$ formado por los nodos, y luego especificar algunas parejas que deben interpretarse como los extremos de una arista. Las parejas pueden estar ordenadas, y establecer así un sentido sobre las aristas. En este caso diremos que el grafo está *orientado*. O ser simplemente un subconjunto de $\mathcal{N}$ formado por dos elementos, entre los que no distinguimos un primero y un segundo. En este caso no estamos privilegiando ninguno de los dos sentidos posibles para una arista sobre el otro. En estos términos, la especificación del grafo orientado del ejercicio 2.8 es

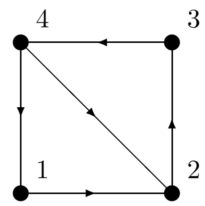
$$\mathcal{N} = \{1, 2, 3, 4\}, \quad \mathcal{A} = \{(1, 2), (2, 3), (3, 4), (4, 1), (4, 2)\},$$

donde $\mathcal{A}$ es el conjunto de las aristas, en este caso orientadas, del grafo. Esta descripción de los nodos y aristas es necesaria para fijar con precisión de qué objeto se está hablando. Sin embargo, parece preferible pensar en un grafo en términos de la representación gráfica de sus nodos y aristas.

En el segundo ejercicio mostraremos cómo codificar la información sobre subconjuntos de un conjunto finito dado. Esta codificación, sumada a la posibilidad de realizar diversas operaciones sobre las matrices, es una herramienta útil para probar muchos resultados combinatorios.

Ejercicio 2.8. Conexiones en grafos

1. El grafo que se muestra en la figura tiene cuatro nodos,



conectados por aristas orientadas. Podemos resumir las conexiones en el grafo por la matriz

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

cuya entrada a_{ij} es igual a la cantidad de arcos que permiten ir del nodo i al nodo j. Calcular la matriz cuya entrada b_{ij} es igual al número de caminos que permiten ir del nodo i al nodo j recorriendo exactamente dos arcos. Calcular la matriz que contiene la información acerca de como ir de un nodo a otro sin recorrer más de dos arcos.

2. Si dejamos que los arcos se recorran en cualquier sentido entonces la matriz de conexiones cambia a
$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$
Repetir la parte anterior para este nuevo esquema de conexiones.

3. Representar gráficamente el grafo orientado cuyas conexiones están dadas por
$$A = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

El siguiente ejercicio trabaja sobre las *matrices de incidencia*, que permiten describir subconjunto de un conjunto finito dado. Sea $X = (x_1, x_2, \cdots, x_n)$ un conjunto ordenado que tiene n elementos, y $\mathcal{A} = (A_1, A_2, \dots, A_m)$ una lista ordenada de m subconjuntos de X. Podemos describir estos subconjuntos por medio de una matriz A, de dimensiones $m \times n$ construida de la siguiente manera: cada fila de A corresponde a uno de los conjuntos A_i, $i = 1, 2, \dots, m$, y tiene un 1 en el lugar j si el elemento x_j está en A_i, y un 0 cuando x_j no está en A_i. Si escribimos las matriz $A = (a_{ij})$ entonces la regla que define A se resume en
$$a_{ij} = \begin{cases} 1 & \text{si} \quad x_j \in A_i, \\ 0 & \text{si} \quad x_j \notin A_i, \end{cases} \qquad i = 1, 2, \dots, m, \; j = 1, 2, \dots, n.$$

Ejercicio 2.9.

1. Hallar la matriz de incidencia que corresponde a la familia de subconjuntos
$$A_1 = \{, @, A, \#\}, \; A_2 = \emptyset, \; A_3 = X, \; A_3 = \{\&, @, \$, \#\}, \; A_1 = \{@, \$, \%, \#\}$$
del conjunto ordenado $X = (\&, @, \$, \%, A, \#)$;

2. Hallar un conjunto y una familia de subconjuntos que correspondan a la matriz
$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

3. Un grafo no orientado puede ser construido de la siguiente manera: damos un conjunto X de nodos, e identificamos cada arista con un subconjunto del conjunto de nodos que tiene exactamente dos elementos. Visto de esta manera, un

grafo no es otra cosa que una lista de parejas dentro del conjunto de nodos. Dar una descripción algebraica de los grafos por medio de una matriz de incidencia. Aplicar esta descripción cuando se considera el grafo del ejercicio 2.8 sin tener en cuenta su orientación.

Para un grafo orientado es posible modificar levemente la construcción de las matrices de incidencia que se describe en el ejercicio 2.8 e incorporar la información acerca de la orientación de las aristas del grafo. Recordemos que cada arista del grafo es un subconjunto del conjunto de nodos que tiene exactamente dos elementos: uno es el nodo en el que la arista empieza; el otro el nodo en que termina. La arista aparece entonces asociada a una fila de la matriz de incidencia que tiene dos entradas iguales a uno (las que están en las columnas que corresponden a los nodos que la arista une), y el resto de las entradas nulas. Si modificamos esta matriz poniendo un $+1$ en el nodo al que llega la arista y un -1 en el nodo del que sale obtenemos una nueva matriz en la que hemos incorporado la información acerca de la orientación en el grafo. En nuestro próximo ejemplo mostramos esta nueva forma de describir un grafo dirigido y algunas de sus aplicaciones.

Ejemplo 2.1.6. Matriz de incidencia arcos-nodos Es posible describir un grafo dirigido $\mathcal{G}$ con m arcos y n nodos con una matriz A que almacena la información acerca de qué nodos conecta cada arco de la siguiente manera: cada fila corresponde a un arco, y cada columna a un nodo. La i-ésima fila tiene sus entradas nulas, salvo un -1 en la posición que corresponde a la columna del nodo inicial del arco i, y un 1 en la del nodo final. Así, la matriz que corresponde al grafo dirigido del ejercicio 2.8 es

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}. \tag{2.5}$$

Observemos que al recorrer las columnas obtenemos toda la información acerca de qué arcos llegan a los nodos o salen de ellos. Cada 1 en la j-ésima columna corresponde a un arco que llega al nodo j, y cada -1 a un arco que sale de ese nodo.

Nuestro próximo ejercicio muestra la relación entre la cantidad de componentes de un grafo y la dimensión del núcleo de su matriz de incidencia. Dejamos los detalles para el lector.

Ejercicio 2.10. Hallar la matriz B de incidencia del grafo orientado con ocho nodos que se muestra en la figura 2.1. Hallar los núcleos de la matriz de incidencia (2.5) y de la matriz B. Hallar las dimensiones de los núcleos. Interpretar los resultados hallados.

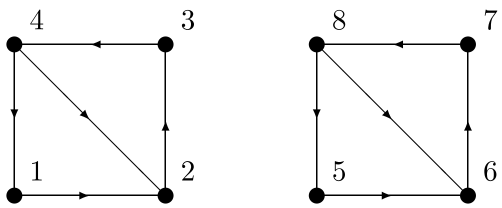


Figura 2.1: Un grafo con ocho nodos

Nuestro próximo ejemplo muestra cómo las relaciones entre los elementos de un circuito quedan codificadas en una matriz: la *matriz de impedancias* del circuito.

Ejemplo 2.1.7. MATRIZ DE IMPEDANCIAS DE UN CIRCUITO
Volvemos aquí sobre el circuito del ejemplo 1.1.3, en la página 8. Cuando hallamos las ecuaciones de este circuito en términos de las corrientes de malla, ver la observación 1.1.2 en la página 10, encontramos el sistema

$$\begin{aligned} (R_1 + R_3)i_1 - R_3 i_2 &= V_a, \\ R_3 i_1 + (R_1 + R_3)i_2 &= -V_b. \end{aligned}$$

La matriz

$$\begin{pmatrix} R_1 + R_3 & -R_3 \\ -R_3 & R_1 + R_3 \end{pmatrix}$$

es lo que se llama *matriz de impedancias del circuito*, y almacena toda la información necesaria para resolver el circuito, salvo la ubicación de las fuentes.

2.2. El álgebra de las matrices

En esta sección definiremos las operaciones de suma de matrices, de producto de un escalar por una matriz y de producto de matrices, que dotan al conjunto de las matrices de una estructura algebraica con la que podremos trabajar. La utilizaremos para tratar una gran variedad de ejemplos y aplicaciones, entre ellos los que vimos en la sección 2.1, en los que, de alguna manera, adelantamos cuáles serían las definiciones de las operaciones con las matrices.

2.2.1. Suma de matrices y producto por un escalar

Comenzaremos por introducir las operaciones de suma entre matrices y producto de una matriz por un escalar. Sumaremos matrices que tengan el mismo tamaño sumándolas entrada a entrada. La definición precisa es la siguiente:

Definición 2.1 (Suma de matrices). *Sean*

$$A = (a_{ij})_{j=1,\dots,n}^{i=1,\dots,m}, \quad B = (b_{ij})_{j=1,\dots,n}^{i=1,\dots,m}$$

dos matrices en $M^{m\times n}(\mathbb{K})$*. Definimos la* ***suma*** *de* A *y* B *como la matriz*

$$A + B = (a_{ij} + b_{ij})_{j=1,\dots,n}^{i=1,\dots,m} \in M^{m\times n}(\mathbb{K}).$$

Esta es la operación que usamos en el ejemplo 2.1.2 al calcular la matriz T con los totales de ventas en todo en todas las sucursales, ver la página 160. También en el ejercicio 2.8 el número de conexiones que pueden establecerse entre dos nodos sin recorrer más de dos arcos puede calcularse sumando la matriz en la que aparece el número de conexiones posibles en las que se recorre exactamente un arco, con la matriz de número de conexiones en las que se recorre exactamente dos arcos.

El producto de una matriz por un escalar se define de manera similar, operando sobre cada entrada.

Definición 2.2 (Producto de números por matrices). *Sean* $\lambda \in \mathbb{K}$ *un número y* $A \in M^{m\times n}(\mathbb{K})$ *una matriz. Definimos el* ***producto*** *de* λ *por* $\mathbf{A}$ *como la matriz*

$$\lambda A = (\lambda a_{ij})_{j=1,\dots,n}^{i=1,\dots,m} \in M^{m\times n}(\mathbb{K}).$$

Observación 2.2.1. Los vectores de $\mathbb{K}^n$ pueden verse como casos particulares de matrices: columnas de dimensiones $n \times 1$, o filas $1 \times n$. La suma de matrices y el producto por escalares que acabamos de definir extienden a matrices cualesquiera las operaciones que ya habíamos definido para vectores.

Esto no es sorprendente, una matriz sobre el cuerpo $\mathbb{K}$, o un vector en $\mathbb{K}^n$ son arreglos de números: en un caso los organizamos en filas y columnas, en el otro en forma de una única fila o columna. Al igual que con los vectores, para sumar dos matrices hemos impuesto la restricción de que tengan el mismo tamaño. ♣

Veamos a continuación algunos ejemplos.

Ejemplo 2.2.2. Sean

$$A = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 1 & 2 & -1 \\ 0 & -1 & \sqrt{2} \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \frac{1}{2} & -1 \\ 2 & 2 & 0 \\ 1 & -2 & 2 \end{pmatrix},$$

entonces

$$A + B = \begin{pmatrix} 1+0 & \frac{1}{2}+\frac{1}{2} & 0-1 \\ 1+2 & 2+2 & -1+0 \\ 0+1 & -1-2 & \sqrt{2}+2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 3 & 4 & -1 \\ 1 & -3 & \sqrt{2}+2 \end{pmatrix}$$

Consideremos ahora la matriz real

$$A = \begin{pmatrix} \sqrt{2} & 1 \\ 0 & -\sqrt{2} \end{pmatrix}.$$

Entonces

$$\sqrt{2}\,A = \begin{pmatrix} \sqrt{2}\times\sqrt{2} & \sqrt{2}\times 1 \\ \sqrt{2}\times 0 & \sqrt{2}\times(-\sqrt{2}) \end{pmatrix} = \begin{pmatrix} 2 & \sqrt{2} \\ 0 & -2 \end{pmatrix}$$

Sean A y B las matrices complejas 3×2

$$A = \begin{pmatrix} 1-i & 2i+1 \\ 1-2i & -3i \\ 2 & 3+i \end{pmatrix}, \quad B = \begin{pmatrix} i & 1-2i \\ 2i & 5+3i \\ 1 & 3-i \end{pmatrix},$$

entonces

$$A + B = \begin{pmatrix} (1-i)+(i) & (2i+1)+(1-2i) \\ (1-2i)+(2i) & (-3i)+(5+3i) \\ (2)+(1) & (3+i)+(3-i) \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 5 \\ 3 & 6 \end{pmatrix}.$$

Veamos, por último, que existen casos en los cuales dos matrices no se pueden sumar empleando la definición 2.1. Consideremos las matrices

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

No está definida la suma $A+B$ porque estas matrices son de diferente tamaño. En este caso no coincide su número de filas. $\square$

Ejercicio 2.11. Sumar las siguientes matrices

1. $\begin{pmatrix} 1 & 2 & -1 \\ -1 & 0 & -3 \end{pmatrix}$ y $\begin{pmatrix} 0 & -2 & -1 \\ 1 & 1 & 2 \end{pmatrix}$
2. $\begin{pmatrix} -2 & 2 \\ 0 & 0 \\ 2 & -2 \end{pmatrix}$ y $\begin{pmatrix} 4 & -3 \\ 1 & -2 \\ 4 & 1 \end{pmatrix}$
3. $\begin{pmatrix} -1 & 3 & -1 \\ 5 & 1 & 0 \\ 4 & 6 & -2 \end{pmatrix}$ y $\begin{pmatrix} 0 & 2 & -3 \\ 4 & 0 & 1 \\ -1 & 1 & -5 \end{pmatrix}$

Observación 2.2.3. Las operación de suma de matrices toma dos matrices A y B en el conjunto $M^{m\times n}(\mathbb{K})$ y produce una tercera matriz $A+B$ en el mismo conjunto. En otras palabras, la suma es una función definida sobre el conjunto de todas las parejas de matrices $m\times n$, que toma valores en el conjunto de las matrices $m\times n$. Algo que representamos con la notación

$$+ : M^{m\times n}(\mathbb{K}) \times M^{m\times n}(\mathbb{K}) \to M^{m\times n}(\mathbb{K}).$$

Análogamente, el producto por un escalar toma un escalar y una matriz y produce una nueva matriz. Es una función

$$\cdot : \mathbb{K} \times M^{m\times n}(\mathbb{K}) \to M^{m\times n}(\mathbb{K}).$$

El producto asocia a cada par (λ, A), donde λ es un escalar y A una matriz, la matriz λA. Observemos que estamos empleando para el producto entre un número y una matriz la misma convención que para el producto entre números: la concatenación del símbolo que representa un escalar con el que representa una matriz deberá interpretarse como la representación del producto entre ambos. ♠

Las operaciones que acabamos de introducir tienen propiedades completamente análogas a la suma y producto por un escalar para n-uplas, o listas de números, a las que hemos dado en llamar vectores a partir de la estructura algebraica que las operaciones de suma y producto por un escalar introducen. Dejamos como ejercicio para el lector completar los detalles. El enunciado preciso de cada propiedad puede deducirse del enunciado correspondiene para n-uplas, ver la página 68.

Ejercicio 2.12.

1. Probar que la operación de suma de matrices tiene las propiedades **conmutativa**, **asociativa**, de **existencia de neutro** y **existencia de opuesto**. Para estas dos últimas habrá que identificar qué matriz actúa como neutro para la suma –a la que llamaremos *matriz nula*–, y determinar cuál matriz es la opuesta de cada matriz dada.

2. Probar que las operaciones de suma entre matrices y de producto entre escalares en el cuerpo $\mathbb{K}$ y matrices tienen las propiedades de **asociatividad respecto al producto de escalares**, de que la unidad en el cuerpo actúa como **unidad para el producto**, y la propiedad **distributiva** respecto a la suma de matrices y respecto a la suma de escalares.

2.2.2. El producto de matrices

En algunos de los ejemplos recogidos en los ejercicios de la sección 2.1 las matrices actúan transformando los objetos a los que se aplican: esto es obvio en los casos del ejercicio 2.3 en que la matriz representa transformaciones geométricas; en el problema 2.1 hay matrices que transforman la información sobre la cantidad de comensales de cada tipo y la comida que hay que preparar, o éste último dato en la descripción detallada de los ingredientes necesarios, y, por último, estos en el precio total de la merienda; en 2.1.1 encontramos una matriz que almacena la información sobre la evolución de la población. Esta matriz transforma el vector con la distribución de la población en el mes k en el vector correspondiente al mes siguiente, $k+1$.

En muchas situaciones en que hay transformaciones operando no nos interesa aplicar una sola de ellas aisladamente, sino que las vamos operando sucesivamente. Por ejemplo, en el ejercicio 2.1 transformábamos primero la información sobre los comensales en información sobre lo que había que preparar, y luego esta información la traducíamos en información sobre ingredientes. Cuando una matriz explica como evoluciona una población mes a mes tenemos que aplicarla varias veces para saber qué ocurre en períodos más largos; etcétera. En ambos casos el efecto de la aplicación sucesiva de dos matrices puede describirse en términos de una nueva matriz. Esta nueva matriz es la que definiremos como el producto de las dos originales. Antes de seguir avanzando en nuestra discusión mostraremos esto en el caso genérico de una matriz B de tamaño 2×3, cuya aplicación sobre un vector de $\mathbb{K}^3$ se ve seguida por la aplicación de una matriz A de tamaño 2×2.

Ejemplo 2.2.4. Consideramos entonces

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Al multiplicar B por una columna $X = (x_1, x_2, x_3)^t$ obtenemos

$$BX = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + b_{13}x_3 \\ b_{21}x_1 + b_{22}x_2 + b_{23}x_3 \end{pmatrix}.$$

El producto de A por BX es

$$A(BX) = \begin{pmatrix} a_{11}\left(b_{11}x_1 + b_{12}x_2 + b_{13}x_3\right) + a_{12}\left(b_{21}x_1 + b_{22}x_2 + b_{23}x_3\right) \\ a_{21}\left(b_{11}x_1 + b_{12}x_2 + b_{13}x_3\right) + a_{22}\left(b_{21}x_1 + b_{22}x_2 + b_{23}x_3\right) \end{pmatrix}$$

que luego de algunas operaciones en las que sacamos de factor común a las entradas x_i, $i = 1, 2, 3$, de la columna X queda en la forma

$$\begin{pmatrix} \left(a_{11}b_{11} + a_{12}b_{21}\right)x_1 + \left(a_{11}b_{12} + a_{12}b_{22}\right)x_2 + \left(a_{11}b_{13} + a_{12}b_{23}\right)x_3 \\ \left(a_{21}b_{11} + a_{22}b_{21}\right)x_1 + \left(a_{21}b_{12} + a_{22}b_{22}\right)x_2 + \left(a_{21}b_{13} + a_{22}b_{23}\right)x_3 \end{pmatrix}.$$

Reconocemos en este último vector el producto de la matriz

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \end{pmatrix} \tag{2.6}$$

por el vector X. La matriz (2.6) es lo que definiremos como el producto AB de las matrices A y B, y recoge el resultado de hacer actuar primero B y luego A. No definiremos el producto BA, porque si aplicamos A a un vector de longitud dos obtenemos un nuevo vector de longitud dos, que no puede ser multiplicado por la matriz B de tamaño 2×3. ♠

En los ejemplos anteriores hemos visto que la acción sucesiva de dos matrices puede describirse por una nueva matriz, cuya acción es equivalente a la acción combinada de las dos matrices originales. Esto no es una peculiaridad de los ejemplos tratados, sino la manifestación de un hecho general que vamos a presentar a continuación.

Consideremos entonces la acción de una matriz B de tamaño $m \times n$ sobre los vectores de $\mathbb{R}^n$. Cada vector $X \in \mathbb{R}^n$ se transforma en un nuevo vector

$$Y = BX \in \mathbb{R}^m$$

al ser multiplicado por B. Si una segunda matriz A $l \times m$, actúa para dar lugar a

$$Z = AY \in \mathbb{R}^l,$$

¿cuál es la transición de X a Z? Observemos que es esencial que A tenga tantas columnas como filas tiene B para que todo esto tenga sentido.

Es fácil dar una primera respuesta a la pregunta que hemos planteado: si $Z = AY$ e $Y = BX$ entonces $Z = A(BX)$. Bien, pero en realidad no estamos diciendo gran cosa, porque esa fórmula, con los paréntesis alrededor de BX, debe leerse "primero apliquemos B a X, y luego apliquemos A al resultado". O sea, sólo expresa en una manera concisa lo que ya sabíamos. Lo que queremos es encontrar una nueva matriz, a la que llamaremos el producto de A y B, y designaremos con AB, tal que su acción sobre X sea la misma que el efecto combinado de A y B. En otras palabras, deseamos que AB sea tal que la igualdad

$$A(BX) = (AB)X \tag{2.7}$$

se satisfaga para cualquier $X \in \mathbb{K}^n$. Ya sabemos que $A(BX)$ tiene longitud l, lo que nos está indicando que la matriz AB debe tener l filas y n columnas.

Hay varias maneras de ordenar los cálculos para extraer la información encerrada en (2.7). Presentamos una que los vuelve relativamente simples, basada en considerar algunos vectores X particulares, especialmente convenientes para calcular, y suficientes para determinar completamente cuál debe ser la definición de AB. Estos son los vectores E_i, $i = 1, 2, \dots, n$, de la base canónica de $\mathbb{K}^n$.

Recordemos que si C es una matriz $p \times n$ cualquiera, entonces el producto CE_i es igual a la i-ésima columna de C. Apliquemos esta observación haciendo

$$X = E_i, \quad i = 1, 2, \dots, n,$$

en (2.7). Obtenemos entonces n igualdades

$$A(BE_i) = (AB)E_i \quad i = 1 \dots n.$$

Cada producto BE_i es igual a B_i, la i-ésima columna de B, y cada producto $(AB)E_i$ es igual a la i-ésima columna de AB, a la que llamaremos $(AB)_i$. Obtenemos entonces

$$A(B_i) = (AB)_i, \quad i = 1, 2, \dots, n. \tag{2.8}$$

Ahora ya tenemos toda la información necesaria para definir el producto de dos matrices cualesquiera.

Definición 2.3 (Producto de matrices). *Si A y B son matrices $l \times m$ y $m \times n$ respectivamente, con entradas en un cuerpo $\mathbb{K}$, entonces el producto AB es la matriz $l \times n$ con entradas en $\mathbb{K}$ tal que para $i = 1, 2, \dots, n$, su i-ésima columna es el producto de la i-ésima columna de B por la matriz A*

Observación 2.2.5. Matrices conformables
En la definición hemos incluido el requisito de que la cantidad de columnas de la matriz A coincida con la cantidad de filas de la matriz B. Esta es una condición indispensable para que exista el producto AB. Cuando se satisface decimos que las matrices A y B son *conformables*. ♣

Observación 2.2.6. Si consideramos a los vectores de $\mathbb{K}^n$ como matrices columna $n \times 1$ entonces el producto AX de una matriz A por un vector X de $\mathbb{K}^n$ resulta ser el caso particular del producto de dos matrices, en el que una de ellas se reduce a una única columna.

Para respetar la definición del producto de matrices, en lo sucesivo siempre consideraremos a los vectores de $\mathbb{K}^n$ como columnas, y cuando aparezcan escritos como filas indicaremos que se trata de la traspuesta del vector.

Ejemplo 2.2.7. El vector X de $\mathbb{R}^2$ que tiene un 1 en su primera entrada y un 2 en la segunda es

$$X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Pero también escribiremos

$$X = (1\ 2)^t,$$

o

$$X = (1, 2)^t,$$

para ahorrar espacio en el papel.

Deberíamos asegurarnos ahora de que la igualdad (2.7) se satisface para todo $X \in \mathbb{K}^n$. De momento, sólo sabemos que nuestra definición asegura que es válida para los vectores de la base canónica. Ese es el contenido de nuestra próxima proposición.

Proposición 2.1. *Sean A y B dos matrices $l \times m$ y $m \times n$ respectivamente, sobre un cuerpo $\mathbb{K}$. Entonces la igualdad*

$$A(BX) = (AB)X \tag{2.9}$$

se satisface para todo $X \in \mathbb{K}$.

La demostración utiliza una de las ideas básicas del álgebra lineal: si sabemos cómo se comporta una operación lineal sobre una base entonces podemos saber cómo se comporta en todo el espacio, extendiendo su acción por medio de la linealidad.

Prueba. Cualquier vector X puede escribirse como combinación lineal de los vectores de la base canónica, en la forma

$$X = \sum_{i=1}^{n} x_i E_i.$$

Usamos ahora la linealidad de la multiplicación por matrices. Primero con la matriz B:

$$BX = B\left(\sum_{i=1}^{n} x_i E_i\right) = \sum_{i=1}^{n} x_i B E_i.$$

Multiplicamos ahora esta expresión por A, y obtenemos

$$A(BX) = A\left(\sum_{i=1}^{n} x_i B E_i\right) = \sum_{i=1}^{n} x_i A(BE_i).$$

Como para los vectores E_i ya sabemos que el resultado que queremos probar es cierto, porque hicimos la definición de AB para que así fuera concluimos

$$A(BX) = \sum_{i=1}^{n} x_i (AB) E_i.$$

Aplicamos una vez más la linealidad del producto por una matriz, ahora para el producto por AB y obtenemos

$$A(BX) = (AB)\left(\sum_{i=1}^{n} x_i E_i\right) = (AB)X.$$

Estas últimas igualdades completan la prueba de la proposición. □

Observación 2.2.8. Dado que $A(BX) = (AB)X$ podemos eliminar los paréntesis en estas fórmulas sin introducir ningún tipo de ambigüedad, e indicar cualquiera de estos dos productos por ABX.

Observación 2.2.9. La definición 2.3 del producto entre matrices puede entenderse en los siguientes términos: a cada matriz $m \times n$ le hemos asociado una transformación de $\mathbb{K}^n$ en $\mathbb{K}^m$ que consiste en multiplicar por la matriz los vectores de $\mathbb{K}^n$. Hemos definido el producto de dos matrices A y B de forma tal que al componer la transformación de multiplicar por B con la transformación de multiplicar por A se obtenga una transformación que consiste en multiplicar por la matriz AB.

Ejemplo 2.2.10. Después de nuestra búsqueda de matrices tales que aplicar primero B a un vector X y luego aplicar A al producto BX, sea equivalente a realizar el producto AB y aplicarlo a dicho vector, estamos listos para calcular el producto de dos matrices A y B. Sean

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 3 & 1 \end{pmatrix} \text{ y } B = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix}$$

Primero debemos multiplicar A por la primera columna de B del siguiente modo:

$$\begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$

Ahora calculamos el producto de la matriz A por la segunda columna de B:

$$\begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 5 \end{pmatrix}.$$

Por último escribimos el producto AB:

$$AB = \begin{pmatrix} 0 & 2 \\ 2 & -4 \\ 1 & 5 \end{pmatrix}.$$

Como las matrices A y B tienen tamaño 3×2 y 2×2 respectivamente, el producto AB es una matriz 3×2. ♠

Ejemplo 2.2.11. En este ejemplo vamos a calcular el producto de dos matrices, pero con coeficientes genéricos. Para hacer este cálculo elegimos dos matrices A y B, de tamaño 3×2 y 2×2 respectivamente, como en el ejemplo 2.2.10. Podemos, entonces, escribir las matrices A y B del siguiente modo:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Si multiplicamos estas matrices según lo que aprendimos en esta sección tenemos que AB es la matriz que llamaremos $C = (c_{ij})$ donde la i-ésima columna de C se obtiene de multiplicar la matriz A por la la i-ésima columna de B. Calculamos primero la columna C_1 de C:

$$C_1 = AB_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} \\ a_{21}b_{11} + a_{22}b_{21} \\ a_{31}b_{11} + a_{32}b_{21} \end{pmatrix}$$

Calculamos ahora la columna C_2 de C:

$$C_2 = AB_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{12} + a_{22}b_{22} \\ a_{31}b_{12} + a_{32}b_{22} \end{pmatrix}$$

Por último debemos "armar" la matriz C juntando las dos columnas:

$$C = (C_1, C_2) = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} \end{pmatrix}$$

Vale la pena adelantar la observación de carácter general de que podemos escribir las entradas de la matriz C empleando el signo de sumatoria. El lector podrá verificar que

$$c_{ij} = \sum_{k=1}^{2} a_{ik}b_{kj}, \quad i = 1, 2, 3 \ j = 1, 2. \tag{2.10}$$

Esta expresión es una forma económica de representar por una única fórmula todas las entradas del producto AB. □

Motivados por el ejemplo 2.2.11, vamos a buscar una expresión para cada entrada c_{ij} de la matriz $C = AB$ que es el producto de dos matrices

$$\begin{aligned} A &= (a_{ik})_{k=1,\dots,n}^{i=1,\dots,m}, \\ B &= (b_{kj})_{j=1,\dots,l}^{k=1,\dots,n}, \end{aligned}$$

sobre un cuerpo $\mathbb{K}$. Entonces $C = AB$ es una matriz $m \times l$, y su entrada c_{ij} es la i-ésima entrada de la j-ésima columna de C, que a su vez es el producto AB_j de la matriz A por B_j, la j-ésima columna de B. Las entradas de B que intervienen en el cálculo de c_{ij} son entonces las

$$b_{kj}, \quad k = 1, \dots, n.$$

En la i-ésima entrada del producto AB_j intervienen las entradas de A que están en la i-ésima fila, es decir, los números

$$a_{ik}, \quad k = 1, \dots, n.$$

En el producto cada entrada a_{ik} multiplica a su correspondiente b_{kj}, y luego se suman, de modo que cada entrada c_{ij} en el producto $C = AB$ es

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \quad i = 1, \dots, m, j = 1, \dots, l.$$

Observemos que es esencialmente la misma expresión que (2.10), pero ahora escrita en una forma genérica que vale para dos matrices A y B conformables cualesquiera sobre el mismo cuerpo $\mathbb{K}$.

El siguiente esquema ayuda a recordar la forma de realizar la operación:

$$
\begin{array}{cc}
 & \begin{pmatrix} b_{11} & \dots & \boxed{b_{1j}} & \dots & b_{1l} \\ b_{21} & \dots & \boxed{b_{2j}} & \dots & b_{2l} \\ \vdots & \dots & \boxed{\vdots} & \dots & \dots \\ b_{m1} & \dots & \boxed{b_{mj}} & \dots & b_{ml} \end{pmatrix} \\
\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \boxed{a_{i1} \quad a_{i2} \quad \dots \quad a_{im}} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} & \begin{pmatrix} \vdots & \vdots & \downarrow & \vdots & \vdots \\ & \longrightarrow & c_{ij} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}
\end{array}
$$

Los ejemplos que acabamos de presentar nos ayudan a ordenar los cálculos a la hora de hacer el producto de dos matrices. Contienen además el siguiente resultado: cada elemento c_{ij} de la matriz $C = AB$ es igual al producto de la i-ésima fila de A por la j-ésima columna de B. En el ejercicio 2.26 veremos como escribir el producto AB en términos de productos de las columnas de A por las filas de B.

Ejercicio 2.13. Ejercitando el cálculo con matrices.

1. Consideremos matrices A y B de dimensión 4×5. Y matrices C, D y E de dimensiones 5×2, 4×2 y 5×4 respectivamente. Todas las matrices tienen sus entradas en el mismo conjunto numérico. Determine cuáles de las siguientes operaciones están definidas:

$$BA, \quad AC + D, \quad AE + B, \quad AB + B, \quad E(A + B), \quad EAC.$$

En caso de estarlo, indique las dimensiones de la matriz resultante.

2. Se consideran las matrices:

$$A = \begin{pmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix}.$$

Realizar las siguientes operaciones: AB, BC, $(AB)C$ y $A(BC)$.

3. Para las matrices

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -3 & 2 & -1 \\ -2 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix},$$

   calcular AB y BA.

4. Dadas las matrices reales

$$A = \begin{pmatrix} 1 & -3 & 2 \\ 2 & 1 & -3 \\ 4 & -3 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 4 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 2 \end{pmatrix},$$
$$C = \begin{pmatrix} 2 & 1 & -1 & -2 \\ 3 & -2 & -1 & -1 \\ 2 & -5 & -1 & 0 \end{pmatrix},$$

   calcular AB y AC.

5. Dadas las matrices

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix},$$

   con entradas en $\mathbb{Z}_2$, calcular AB y BA.

6. ¿Es conmutativo el producto de matrices? Justifique la respuesta.
7. ¿Si $A \neq 0$; $AB = AC \Rightarrow B = C$? Justifique la respuesta.

Ejercicio 2.14. Encontrar ejemplos de matrices reales 2×2 tales que:

1. $A^2 = -I$;
2. $B^2 = O$, $B \neq O$;
3. $CD = -DC$, $(CD \neq O)$;
4. $EF = O$ aunque no tengan ningún elemento cero.

Sugerencia: puede ser útil usar ideas geométricas para buscar ejemplos con las propiedades que se piden (ver el ejercicio 2.3).

Ejercicio 2.15. Determine si las siguientes afirmaciones son verdaderas (indicando por qué) o falsas (dando un contraejemplo):

1. Si la primera y tercera columnas de B son iguales, también lo son la primera y tercera columnas de AB;
2. Si la primera y tercera filas son iguales en B, también lo son en AB;
3. Si la primera y tercera filas son iguales en A, también lo son en AB;

4. Si A y B son matrices $n \times n$ entonces

   a) $(AB)^2 = A^2B^2$;

   b) $(A+B)^2 = A^2 + 2AB + B^2$;

   c) $(A+B)^2 = A^2 + AB + BA + B^2$.

5. Si el producto AB entre las matrices A y B está definido entonces también lo está el producto BA.

Ejercicio 2.16. Mostrar que el espacio de columnas de AB está contenido en el espacio de columnas de A. Dar ejemplos en los que ambos espacios sean iguales, y ejemplos en los que la inclusión sea estricta.

Ejercicio 2.17. Volver a los ejercicios y ejemplos de la sección 2.1 y analizar dónde se utilizaron la suma, multiplicación por un escalar y producto de matrices.

Ejemplo 2.2.12. Formulación matricial de las leyes de Kirchoff y Ohm

En este ejemplo mostraremos como el cálculo matricial aplicado a las matrices de incidencia de una red permiten escribir rápidamente las ecuaciones que gobiernan un flujo sobre la red. Volveremos a tratar la red de la figura 1.2, con la que trabajamos en la sección 1.1.4, página 12. Esta red con arcos dirigidos puede describirse por una matriz de incidencia arcos-nodos, como las que introdujimos en el ejemplo 2.1.6 de la página 172. De la inspección del esquema de conexiones surge que su matriz de incidencia es

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}. \tag{2.11}$$

Observemos que al recorrer las columnas obtenemos toda la información acerca de qué arcos llegan a los nodos o salen de ellos. Cada 1 en la j-ésima columna corresponde a un arco que llega al nodo j, y cada -1 a un arco que sale de ese nodo.

Dedujimos las ecuaciones que gobernaban el flujo sobre la red a partir de las leyes de Kirchoff y de Ohm. Veremos a continuación que ambas leyes, y, en consecuencia, las ecuaciones del circuito, admiten una sencillísima expresión en términos de la matriz de incidencia (2.11). Haremos la deducción para un circuito cualquiera, con m arcos y n nodos. Llamemos

$$\begin{aligned} I &= (i_1, i_2, \ldots, i_m)^t, \\ V &= (v_1, v_2, \ldots, v_n)^t, \\ D &= (d_1, d_2, \ldots, d_n)^t, \end{aligned}$$

respectivamente a los vectores que almacenan en sus entradas los valores de los flujos en los arcos, y de los potenciales y demandas en los nodos. En el producto A^tI estamos multiplicando las filas de A^t (columnas de A) por el vector de flujos. El resultado de multiplicar la j-ésima columna de A por el vector de flujos es sumar todas los flujos que entran o salen del nodo j, con el signo que les corresponda. El resultado debe compensar a la demanda d_j en ese nodo. Plantear el balance entre los flujos y demandas en todos los nodos es equivalente a la ecuación matricial

$$A^tI = D$$

que resulta ser una expresión matricial de las leyes de Kirchoff.

Las diferencias de potenciales entre los nodos iniciales y finales de un mismo arco son las entradas del producto AV entre la matriz de incidencia A y el vector de potenciales V. Si a su vez multiplicamos AV por la matriz diagonal

$$C = \begin{pmatrix} c_1 & 0 & \ldots & 0 & 0 \\ 0 & c_2 & \ldots & 0 & 0 \\ \vdots & \ldots & \ddots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 & c_m \end{pmatrix}$$

que contiene las susceptancias obtenemos el vector de productos de las susceptancias por las diferencias de voltajes, que debe coincidir con el vector de intensidades. En resumen, la ecuación matricial

$$CAV = I$$

expresa la ley de Ohm aplicada a todos los arcos del grafo.

Notemos que las expresiones matriciales $CAV = I$ y $A^tI = D$ de las leyes de Ohm y Kirchoff pueden combinarse en la ecuación

$$A^tCAV = D,$$

que es un sistema de ecuaciones lineales sobre los voltajes en los nodos. Cuando A es la matriz (2.11) del grafo del ejemplo el producto A^tCA es

$$A^tCA = \begin{pmatrix} c_1+c_2+c_3 & -c_1 & -c_2 & -c_3 & 0 \\ -c_1 & c_1+c_4+c_5 & 0 & -c_4 & -c_5 \\ -c_2 & 0 & c_2+c_7 & -c_7 & 0 \\ -c_3 & -c_4 & -c_7 & c_3+_4+c_6+c_7 & -c_6 \\ 0 & -c_5 & 0 & -c_6 & c_5+_6 \end{pmatrix}$$

Vale la pena observar la facilidad con la que obtuvimos las ecuaciones del sistema, y compararlo con el trabajoso procedimiento de sustituir las variables una a una que usamos en la sección 1.1.4 cuando derivamos por primera vez las ecuaciones para los potenciales

Observación 2.2.13. El producto visto por filas
La definición del producto de matrices está basada en las columnas. Cada columna del producto AB es el producto de A por una de las columnas de B. De modo que cada columna de AB es una combinación lineal de las columnas de A con coeficientes almacenados en una columna de B. Vamos a ver ahora algo interesante: las filas de AB son una combinación lineal de las filas de B, cuyos coeficientes están dados por las filas de A.

Comencemos por volver al ejemplo 2.2.11. Las filas de la matriz $C = AB$ de ese ejemplo son

$$(c_{11}\ c_{12}), \quad (c_{21}\ c_{22}), \quad (c_{31}\ c_{32}),$$

donde

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21}, & c_{12} &= a_{11}b_{12} + a_{12}b_{22}, \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21}, & c_{22} &= a_{21}b_{12} + a_{22}b_{22}, \\ c_{31} &= a_{31}b_{11} + a_{32}b_{21}, & c_{32} &= a_{31}b_{12} + a_{32}b_{22}. \end{aligned}$$

La primera fila es entonces

$$(a_{11}b_{11} + a_{12}b_{21}\ \ a_{11}b_{12} + a_{12}b_{22}) = a_{11}\,(b_{11}\ \ b_{12}) + a_{12}\,(b_{21}\ \ b_{22})\,,$$

que es la combinación lineal de las dos filas de B con los coeficientes de la primera fila de A. Análogamente, para la segunda y tercera fila de AB tenemos

$$\begin{aligned} (a_{21}b_{11} + a_{22}b_{21}\ \ a_{21}b_{12} + a_{22}b_{22}) &= a_{21}\,(b_{11}\ \ b_{12}) + a_{22}\,(b_{21}\ \ b_{22})\,, \\ (a_{31}b_{11} + a_{32}b_{21}\ \ a_{31}b_{12} + a_{32}b_{22}) &= a_{31}\,(b_{11}\ \ b_{12}) + a_{32}\,(b_{21}\ \ b_{22})\,. \end{aligned}$$

En esta expresión se ve claramente lo que queríamos: cada una de las filas del producto es una combinación lineal de las filas de B, y los coeficientes que aparecen son las entradas de una de las columnas de la matriz A.

Ejercicio 2.18. Repetir este análisis para las matrices A y B del ejemplo 2.2.10.

Luego de ver qué es lo que ocurre en el caso particular del producto de una matriz 3×2 por una matriz 2×2 trabajaremos con dos matrices conformables

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \dots & b_{1l} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nl} \end{pmatrix}$$

$m \times n$ y $n \times l$ cualesquiera. La entrada que ocupa la fila i y la columna j del producto es

$$\sum_{k=1}^{n} a_{ik} b_{kj}.$$

Por lo tanto, la i-ésima fila del producto es

$$\left(\sum_{k=1}^{n} a_{ik} b_{k1} \ \dots \sum_{k=1}^{n} a_{ik} b_{kj} \dots \sum_{k=1}^{n} a_{ik} b_{kl} \right).$$

Podemos sacar las sumatorias para fuera de la fila, para escribirla en la forma

$$\sum_{k=1}^{n} a_{ik} \left(b_{k1} \ \dots b_{kj} \dots b_{kl} \right),$$

que pone en evidencia que se trata de una combinación lineal de las n filas de B con los coeficientes a_{ik}, $k = 1, \dots, n$, que forman la i-ésima fila de A.

Ejemplo 2.2.14. Analicemos nuevamente el producto AB del ejemplo 2.2.10, pero ahora miremos el producto por filas. Según acabamos de ver, dicho producto es una combinación lineal de las 2 filas de B :

$$F_1 = (0, 2) \quad \text{y} \quad F_2 = (1, -1).$$

La primer fila del producto AB es $(0, 2)$, que es combinación lineal de las dos filas de B con coeficientes 1 y 0. Lo mismo ocurre para las otras dos filas del producto:

$$\begin{array}{lcrcr} (0,2) & = & 1(0,2) + 0(1,-1) & = & 1F_1 + 0F_2 \\ (2,-4) & = & -1(0,2) + 2(1,-1) & = & -1F_1 + 2F_2 \\ (1,5) & = & 3(0,2) + 1(1,-1) & = & 3F_1 + 1F_2 \end{array}$$

Con esto vemos que las tres filas del producto AB son combinación lineal de las filas de B con coeficientes las entradas de A. ♣

Ejercicio 2.19. Transformaciones elementales y matrices A la matriz

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 1 \\ 6 & 5 & 1 \end{pmatrix}$$

se le aplica el proceso de eliminación gaussiana. En el primer paso obtenemos

$$B = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{pmatrix}$$

1. Hallar una matriz M tal que $B = MA$.
2. Hallar una matriz N tal que $A = NB$.

Ejercicio 2.20. Sea A una matriz $m \times n$, y B una matriz que se obtiene intercambiando las filas de A. Mostrar que existe una matriz P de tamaõ $m \times m$ tal que $B = PA$. ♠

Ejercicio 2.21. Sea A una matriz cualquiera, y E una forma escalerizada de A. Mostrar que existen matrices cuadradas M y N tales que $E = MA$ y $A = NE$.

Ejercicio 2.22. Supongamos que al aplicar a la matriz A una serie de operaciones elementales obtenemos una forma escalerizada E. Mostrar que si las mismas operaciones elementales se aplican al producto AB entonces se obtiene la matriz EB.

Nuestro próximo ejercicio tiene que ver con matrices traspuestas. Un concepto relacionado con la trasposición es el de *matriz simétrica*. Decimos que una matriz A es *simétrica* cuando es igual a su traspuesta. Es decir, si $A = A^t$. Notemos que sólo las matrices cuadradas pueden ser simétricas. Llamaremos *antisimétrica* a una matriz A que es igual al opuesto de su traspuesta. O sea, a las que satisfacen $A = -A^t$. Notemos que sólo las matrices cuadradas pueden ser simétricas o antisimétricas.

Observación 2.2.15. Matrices simétricas

La simetría de una matriz es una propiedad muy importante, que muchas matrices que aparecen en las aplicaciones tienen. Por ejemplo, en muchos casos la entrada a_{ij} de una matriz representa la relación que tienen entre sí las partes i y j de una cierta estructura, de cualquier naturaleza. En diversas situaciones, por ejemplo en los grafos no dirigidos, estas relaciones son simétricas y la forma en que la parte i afecta a la j es la misma en la que j afecta a i.

Más interesante todavía es que la trasposición tiene un importante significado geométrico, que se traduce en una rica teoría geométrica para las matrices simétricas. Todavía falta un buen trecho para llegar a ella, pero vale la pena el anuncio, porque adelantos de esta teoría irán apareciendo a lo largo de este texto en algunos ejemplos. ♣

Ejercicio 2.23. El producto y la trasposición.

1. Mostrar que $(A+B)^t = A^t + B^t$ y $(\alpha A)^t = \alpha A^t$.
2. Sean A y B matrices conformables. Mostrar que la traspuesta $(AB)^t$ del producto AB es
$$(AB)^t = B^t A^t.$$
3. Sea A una matriz cualquiera. Mostrar que los productos A^tA y AA^t siempre están definidos y son matrices simétricas.
4. Interpretar el significado de las entradas de A^tA y AA^t cuando A es una matriz de incidencia como las que se describen en el ejercicio 2.9 de la página 171. Explicar que significado tiene en este contexto particular la simetría de A^tA y de AA^t.

Ejercicio 2.24. Traza de una matriz. Sea A una matriz cuadrada. Se define la **traza** $tr(A)$ de la matriz A como la suma de todos los elementos de su diagonal. Entonces, si $A = ((a_{ij}))$ es una matriz $n \times n$ tendremos

$$tr(A) = a_{11} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Probar las siguientes propiedades:

$$tr(A+B) = tr(A) + tr(B), \qquad tr(\alpha A) = \alpha\, tr(A), \qquad tr(AB) = tr(BA).$$

Para el próximo ejercicio es necesario introducir un poco de jerga. Llamaremos *matriz triangular inferior* a una matriz cuadrada en la que todas las entradas que están por encima de la diagonal son ceros. Y *matriz triangular superior* a la que tiene todas las entradas nulas por debajo de la diagonal.

Ejercicio 2.25. Investigar qué productos de matrices triangulares superiores o inferiores son también triangulares.

Hemos visto que las entradas del producto AB son los productos de las filas de A por las columnas de B. En los próximos dos ejercicios analizamos el producto de dos matrices llamando la atención sobre otras submatrices de A y de B.

Ejercicio 2.26. Indiquemos por A_j y B_j, $j = 1, \ldots, n$, a las j columnas y a las j filas de dos matrices A y B conformables. Mostrar que

$$AB = \sum_{j=1}^{n} A_j B_j.$$

Ejercicio 2.27. Producto por bloques
Sean A y B dos matrices sobre un cuerpo $\mathbb{K}$, que se descomponen en bloques, en la forma

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

donde A_{ij} y B_{ij}, para $i, j = 1, 2$, son submatrices de A y B. Investigar qué sentido tiene la igualdad matricial

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

2.2.3. Propiedades del producto

Si A, B y C son tres matrices conformables, entonces el producto tiene la propiedad **asociativa**

$$(AB)C = A(BC).$$

Esta propiedad puede ver como una extensión de las propiedades de la definición de producto de matrices y su demostración queda como ejercicio para el lector.

Ejercicio 2.28. Demostrar la propiedad asociativa del producto de matrices.

También es fácil de demostrar la propiedad asociativa respecto al producto por un escalar.

Ejercicio 2.29. Mostrar que si α es un escalar, y A y B son matrices conformables entonces valen las igualdades $(\alpha A)B = \alpha(AB) = \alpha(AB)$.

Si A y B son matrices de igual tamaño y C es una matriz conformable con A y B entonces vale la **propiedad distributiva a la derecha**

$$C(A + B) = CA + CB.$$

Cuando C es conformable con dos matrices A y B que tienen el mismo tamaño también se verifica la **propiedad distributiva a la izquierda**

$$(A + B)C = AC + BC.$$

Se hace necesario enunciar y demostrar las propiedades distributiva a izquierda y derecha por separado porque puede ocurrir que $C(A + B)$ esté definido y $(A + B)C$ no, o viceversa. Esta demostración también queda como ejercicio. Observar, sin embargo, que la no conmutatividad del producto obliga a probar ambas identidades independientemente, pues en general $C(A+B) \neq (A+B)C$

Ejercicio 2.30. Demostrar las dos propiedades distributivas que acabamos de enunciar.

Quizás la característica más notable del producto de matrices sea que **no se satisface la propiedad conmutativa**. Es decir, cuando de matrices se trata **el orden de los factores altera el producto**.

Un primer ejemplo de esto es el caso de matrices A y B para las cuales el producto AB está definido, pero el producto BA no lo está. Por ejemplo, el producto AB de dos matrices A y B de tamaño 2×2 y 2×3 respectivamente está definido, pero no podemos realizar el producto BA pues las matrices no resultan conformables.

Un segundo ejemplo es el caso de A y B matrices $m \times n$ y $n \times m$ respectivamente, con $m \neq n$. El producto AB es una matriz cuadrada $m \times m$, en tanto que BA es $n \times n$. Es obvio que estas dos matrices no pueden ser iguales porque tienen distinto tamaño.

Lo que ya no es tan obvio es que el producto AB puede diferir de BA incluso incluso en el caso en que A y B sean dos matrices cuadradas $n \times n$.

Ejemplo 2.2.16. Consideremos las dos matrices reales 2×2

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Entonces

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Vemos que los dos productos están definidos, tienen el mismo tamaño, pero $AB \neq BA$. ♣

El ejemplo anterior no debe hacernos creer que el producto AB siempre difiere del BA.

Ejemplo 2.2.17. Cuando

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

Entonces

$$AB = BA \begin{pmatrix} 2 & 0 \\ 2 & 0 \end{pmatrix}.$$

La matriz cuadrada nula O, que tiene ceros en todas sus entradas, es un ejemplo trivial en el que todos los productos AO y OA son iguales a O. Por lo tanto son iguales entre sí. ♣

Cuando el producto AB de dos matrices cuadradas es igual a BA diremos que las matrices *conmutan*. Cuando $AB \neq BA$ diremos que las matrices *no conmutan*.

Observación 2.2.18. Recordemos la discusión que precede a la definición 2.3. Nos dice que el producto de matrices no es más que la manera matricial de representar la composición de dos funciones. Cuando este hecho se tiene en cuenta no resulta para nada sorprendente que el producto de matrices no sea conmutativo, porque La composición de funciones no es conmutativa.

Por ejemplo, si consideramos las sencillas funciones reales

$$f(x) = x^2, \quad g(x) = x + 1,$$

tenemos

$$(g \circ f)(x) = x^2 + 1, \quad (f \circ g)(x) = (x + 1)^2.$$

Esta propiedad no conmutativa de la composición se expresa también en el producto de matrices. ♣

Ejercicio 2.31. Se sabe que A conmuta con B, y B con C. ¿Es cierto que A conmuta con C?

Ejercicio 2.32. Mostrar que el producto de dos matrices simétricas A y B es simétrico si y sólo si A y B conmutan.

El producto de matrices es especialmente interesante cuando se trata de matrices cuadradas $n \times n$, porque origina una nueva matriz que tiene exactamente el mismo tamaño. Tenemos entonces que para cada valor de n el producto de matrices define una operación binaria

$$\cdot : M^{n\times n}(\mathbb{K}) \times M^{n\times n}(\mathbb{K}) \to M^{n\times n}(\mathbb{K}),$$

que tiene además la propiedad distributiva respecto a la operación de suma de matrices. La estructura $(M^{n\times n}(\mathbb{K}), +, \cdot)$ es la de una anillo no conmutativo. Existe además una unidad para el producto. Dejamos al lector la tarea de encontrarla y determinar algunas de sus propiedades.

Ejercicio 2.33. MATRIZ IDENTIDAD $n \times n$

1. Encontrar una matriz I_n, de dimensiones $n \times n$, que cumpla que

   $$AI_n = I_n A = A$$

   para cualquier matriz $A \in M^{n\times n}(\mathbb{K})$. Mostrar que para cada valor de n hay una única matriz con esta propiedad.

2. Mostrar que I_n es la única matriz $n \times n$ que tiene la propiedad de que $I_n X = X$ para todo $X \in \mathbb{K}^n$.

La matriz I_n que se pide hallar en el ejercicio anterior tiene la propiedadad de que no produce ningún cambio en cualquier matriz A que sea multiplicada por ella. Por esa razón llamaremos a I_n la *matriz identidad*, o *matriz identidad* $n \times n$, para enfatizar que para cada valor de n hay una matriz identidad diferente. Usualmente indicaremos a la matriz identidad simplemente con el símbolo I, sin hacer referencia explícita al tamaño de la matriz.

Dado que hay una identidad para el producto, la estructura de las matrices $n \times n$ es algo más rica: se trata de un anillo no conmutativo con unidad.

Ejercicio 2.34. Verificar la afirmación anterior comprobando que se satisfacen todas las propiedades de esa estructura.

Ejercicio 2.35. Si A es una matriz $n \times n$, probar que A conmuta con todas las matrices $n \times n$ si y sólo si A es un múltiplo de la identidad. Es decir, si existe $\lambda \in \mathbb{K}$ tal que $A = \lambda I$.

Ejercicio 2.36. Para $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

1. Hallar todas las matrices 2×2 que conmuten con A.
2. Hallar una matriz 2×2 que no conmute con A.
3. Hallar todas las matrices que conmutan con $U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$

El siguiente ejercicio establece una conexión entre un subconjunto de las matrices 2×2 y el conjunto de los números complejos. La idea es que cada número complejo $z = x + iy$ puede pensarse como una par $(x, y) \in \mathbb{R}^2$. La multiplicación por un número complejo fijo $w = a + ib$ define una transformación

$$z \mapsto wz$$

que envía cada z a su producto por w. Esta transformación tiene un significado geométrico que puede expresarse en términos del módulo y el argumento del complejo w. Veremos que esta transformación también puede escribirse por medio de una matriz, que a su vez puede representarse como el producto de la matriz de un giro y la matriz de una dilatación.

Ejercicio 2.37. Consideremos un número complejo $w = a + ib$.

1. Identificar cada complejo $z = x + iy$ con el par $(x, y) \in \mathbb{R}^2$. Hallar el par (u, v) que corresponde al complejo wz y mostrar que la matriz

$$W = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

representa en $\mathbb{R}^2$ la transformación $(x, y) \mapsto (u, v)$.

2. Mostrar que W puede escribirse como el producto ρG de un número real $\rho \geq 0$ y la matriz G de un giro. Sugerencia: considerar el módulo y el argumento del complejo w.
3. En la parte 1 hemos encontrado una manera de identificar un número complejo $w = a + ib$ con una matriz. Hallar las operaciones sobre las matrices que corresponde a la suma y al producto de dos números complejos w_1 y w_2.
4. Mostrar que todas las posibles matrices W conmutan entre sí. ¿Qué relación tiene esta propiedad con las propiedades algebraicas de los números complejos?
5. Para cada matriz W no nula hallar una matriz W^{-1} tal que $WW^{-1} = WW^{-1} = I$.

2.2.4. Para tener presente

- La suma de dos matrices del mismo tamaño se define como la suma entrada a entrada.
- El producto de un escalar por una matriz se define multiplicando cada entrada por el escalar.
- El producto AB entre dos matrices A y B de tamaño $m \times n$ y $n \times p$ respectivamente se define de modo que la igualdad $A(BX) = (AB)X$ se satisfaga para todo $X \in \mathbb{K}^p$.
- Si el número de columnas de A difiere del número de filas de B el producto AB no está definido.
- La j-ésima columna de AB es igual al producto de A por la j-ésima columna de B.
- La i-ésima fila de AB es igual al producto de la i-ésima fila de A por B.
- El producto de matrices no es, en general, conmutativo. Pero hay parejas de matrices cuadradas que conmutan.
- Las matrices cuadradas, con las operaciones de suma y producto, forman un anillo no conmutativo con unidad.
- En distintos ejemplos, toda esta estructura algebraica adquiere distintos significados.

2.3. La inversa de una matriz

En la sección 2.2 el producto de matrices fue definido de una forma tal que nos permite representar la acción sucesiva de dos matrices por su producto. En esta sección nos preocuparemos por un problema que pertenece al mismo círculo de ideas: deshacer lo hecho por una matriz. Trabajaremos en el contexto de matrices cuadradas $n \times n$.

Volvamos entonces al punto de vista de considerar una matriz A, de tamaño $n \times n$ y entradas en algún cuerpo $\mathbb{K}$, como una manera de definir una transformación

$$X \mapsto Y = AX$$

que a un vector X de K^n le asocia $AX \in \mathbb{K}^m$. ¿Podremos deshacer el efecto de A y determinar X a partir de Y? Esta cuestión es equivalente a la de resolver el sistema de ecuaciones

$$AX = Y.$$

Desde el punto de vista de la estructura algebraica que tienen las matrices $M^{n\times n}(\mathbb{K})$ se trata de buscar la inversa A^{-1} de la matriz A.

Ejemplo 2.3.1. Consideremos el sistema $AX = B$, con matriz

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

y término general $B = (b_1, b_2)^t$. Escalerizamos, y en el primer paso obtenemos

$$A = \left(\begin{array}{rr|l} 1 & 2 & b_1 \\ 0 & -1 & b_2 - 3b_1 \end{array}\right).$$

Eliminamos ahora el 2 de la primera fila y segunda columna, y multiplicamos por -1 la segunda ecuación para obtener

$$A = \left(\begin{array}{rr|l} 1 & 0 & -5b_1 + 2b_2 \\ 0 & 1 & 3b_1 - b_2 \end{array}\right).$$

De aquí se desprende que la solución X es

$$X = \begin{pmatrix} -5b_1 + 2b_2 \\ 3b_1 - b_2 \end{pmatrix}$$

que puede escribirse en forma matricial como

$$X = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} B.$$

Llamamos A^{-1} a la matriz
$$\begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$$
porque es la *inversa* de A, en el sentido de que
$$A^{-1}A = AA^{-1} = I,$$
algo que el lector podrá comprobar directamente, o de que si $AX = B$ entonces $X = A^{-1}B$.

Ejercicio 2.38. Comprobar la igualdad $A^{-1}A = AA^{-1}$. ♣

Observación 2.3.2. Cuando una matriz A tiene inversa entonces la solución de $AX = B$ es $B = A^{-1}X$. En este punto algunos aspectos de la teoría de los sistemas de ecuaciones se separan del cálculo práctico. Si bien desde el punto de vista teórico el producto A^{-1} por X tiene pleno sentido, no es una buena estrategia para resolver $AX = B$ calcular la inversa A^{-1} y luego multiplicar. Ya hemos visto que para resolver $AX = B$ no hace falta calcular A^{-1}. De hecho, hemos resuelto muchos sistemas sin calcular esta inversa. En la sección 2.4 veremos cómo la resolución de sistemas de ecuaciones lineales se traduce a los efectos prácticos en la factorización de la matriz A como el producto LU de dos matrices triangulares. El factor U es una forma escalerizada de A, en tanto que L almacena toda los multiplicadores que se generan a lo largo del proceso de eliminación gaussiana.

Para enfatizar estos comentarios citemos la referencia [TB], un excelente libro de Álgebra Lineal Numérica[6]: *¡Cuando se escribe el producto $X = A^{-1}B$, es importante no permitir que la notación que hace uso de la matriz inversa oculte lo que realmente sucede! Más que pensar en X como el resultado de aplicar A^{-1} a B, debemos entender que es el único vector que satisface $AX = B$. Esto significa que X es el vector de coeficientes de la única expresión de B como una combinación lineal de los vectores en la base formada por las columnas de A.*

Este punto no puede ser demasiado enfatizado, de modo que repetimos:

$A^{-1}B$ es el vector de coeficientes de la expansión de B en la base de columnas de A.

Desde este punto de vista, multiplicar por A^{-1} puede verse como una operación de cambio de coordenadas, una manera de referirse a un vector dando sus coordenadas en una nueva base (la base formadas por las columnas de A). Esta observación será desarrollada en la sección dedicada a los cambios de base, en el capítulo 4. ♠

[6]La traducción es de los autores de este texto

2.3.1. La inversa y el álgebra de matrices

Para una matriz cuadrada $A \in M^{n\times n}(\mathbb{K})$, andamos a la búsqueda de una matriz A^{-1} que pueda deshacer lo que A hace. Esta condición puede expresarse de manera algebraica. Si $Y = AX$ entonces pretendemos que $A^{-1}Y = X$. Esto implica

$$A^{-1}(AX) = \left(A^{-1}A\right) X = X,$$

por lo que el efecto de la matriz producto $A^{-1}A$ sobre cualquier vector $X \in \mathbb{K}^n$ es el de dejarlo incambiado. ¿Cuál es la única matriz que tiene semejante propiedad? Por supuesto, se trata de la matriz identidad I_n, tal como vimos en la parte 2 del ejercicio 2.33 en la página 195. Tenemos ahora una condición algebraica para expresar esta idea de que la matriz A^{-1} es la inversa de A, o deshace lo que A hace. Sobre esta idea construiremos la definición de inversa de una matriz, pero agregaremos además un matiz: pediremos tanto que A^{-1} deshaga el efecto de A, como que A deshaga el de A^{-1}. Ese es el contenido de nuestra próxima definición.

Definición 2.4. *Sea A una matriz $n \times n$ sobre un cuerpo $\mathbb{K}$. Diremos que una matriz A^{-1} es la inversa de A si se satisface*

$$A^{-1}A = AA^{-1} = I_n. \tag{2.12}$$

Observación 2.3.3. La definición de inverso de una matriz es formalmente igual a la condición $a^{-1}a = 1$ que caracteriza el inverso a^{-1} de un número $a \neq 0$ en un cuerpo $\mathbb{K}$ cualquiera. Recordemos que 1 es la unidad para la multiplicación en el cuerpo, de la misma manera que la matriz identidad es la unidad para la multiplicación dentro del conjunto de matrices cuadradas. Pero hay algunas diferencias.

Como el producto de matrices no es, en general, conmutativo hemos incluido en la definición de matriz inversa que los dos posibles productos entre las matrices A^{-1} y A sean iguales a la identidad. En realidad no es esencial tener este cuidado al formular la definición, porque basta con que uno de los productos sea igual a la identidad para que el otro también lo sea. Pero este es un resultado interesante de la teoría, que trataremos algo más adelante.

La otra diferencia notable que encontraremos con el caso escalar es que hay matrices distintas de la matriz nula que no tienen inversa. Ver el ejemplo 2.3.9. En un cuerpo todos los elementos no nulos son invertibles. En el conjunto de las matrices esta afirmación deja de ser cierta. ♣

Observación 2.3.4. La definición de inverso de una matriz se corresponde con la definición de función inversa. La inversa de una función f es una función

f^{-1} tal que las dos composiciones $f \circ f^{-1}$ y $f^{-1} \circ f$ son iguales a la identidad. ♠

Ya vimos en el ejemplo 2.3.1 una matriz 2×2 que tiene inversa. A continuación trataremos un ejemplo de naturaleza geométrica.

Ejemplo 2.3.5. Vimos en la parte 1 del ejercicio 2.3, página 163, que la matriz

$$G_\theta = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

representa en el plano un giro de ángulo θ. ¿Qué movimiento es el inverso de un giro de ángulo θ? Un giro de ángulo $-\theta$.

De acuerdo con esta observación la matriz

$$G_{-\theta} = \begin{pmatrix} \cos(-\theta) & -\operatorname{sen}(-\theta) \\ \operatorname{sen}(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \operatorname{sen}(\theta) \\ -\operatorname{sen}(\theta) & \cos(\theta) \end{pmatrix}$$

debería ser la inversa de G_θ. Dejamos los detalles como un ejercicio para el lector.

Ejercicio 2.39. Verificar que $G_{-\theta}G_\theta = G_\theta G_{-\theta} = I$. ♣

Ejemplo 2.3.6. La matriz

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

no es invertible. En efecto, para cualquier matriz

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

se cumple que

$$AB = \begin{pmatrix} b_{11} & b_{12} \\ 0 & 0 \end{pmatrix}.$$

Este producto nunca puede ser igual a la matriz identidad 2×2, independientemente de que valores tomen b_{11} y b_{12}, porque necesariamente tiene una fila de ceros. ♣

Observación 2.3.7. UNICIDAD DE LA INVERSA.
En nuestra definición de inversa de una matriz nos hemos apresurado en referirnos a *la* inversa. En principio podría ocurrir que, dada una matriz A, muchas matrices A^{-1} satisficieran los requisitos impuestos por la fórmula

(2.12). Pero un sencillo razonamiento muestra que, a lo sumo, hay una. Dada una matriz A consideremos inversas B_1 y B_2, entonces

$$B_1A = AB_1 = I, \quad B_2A = AB_2 = I.$$

La igualdad $AB_1 = I$ implica que

$$B_2(AB_1) = B_2I = B_2.$$

Por lo tanto

$$(B_2A)B_1 = B_2.$$

Como $B_2A = I$ concluimos que

$$B_1 = IB_1 = B_2,$$

por lo que ambas matrices son en realidad la misma y, si la inversa existe, es única. El razonamiento que acabamos de emplear permite mostrar un resultado un poco más fuerte. Como usaremos este resultado más adelante lo enunciaremos bajo la forma de un lema, cuya prueba quedará como ejercicio para el lector.

Lema 2.2. *Sea A una matriz $n \times n$ tal que existen matrices B_1 y B_2 que satisfacen*

$$AB_1 = B_2A = I_n.$$

Mostrar que A es invertible y que que $B_1 = B_2 = A^{-1}$.

Ejercicio 2.40. Demostrar el Lema 2.2. ♠

2.3.2. El cálculo de la inversa.

Discutiremos ahora brevemente el cálculo de la inversa de una matriz cuadrada $n \times n$. Un problema que está muy relacionado con la resolución de sistemas lineales de ecuaciones. El procedimiento de cálculo no es más que la resolución simultánea, ordenada, de n sistemas de ecuaciones lineales. El método ya está implícito en la resolución de los ejercicios 1.24 y 1.30, propuestos en las páginas 57 y 63. La incompatibilidad de alguno de los n sistemas es equivalente a la inexistencia de la inversa, tal como veremos al discutir la teoría correspondiente en la sección 2.3.3. No haremos mucho énfasis en estos algoritmos de cálculo, porque en la práctica no se emplean y el cálculo de inversas se sustituye con ventaja por la descomposición LU que discutiremos en la sección 2.4.

Ejemplo 2.3.8. Consideremos la matriz 2×2

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}.$$

Su inversa, si existe, es una matriz B tal que $AB = I$, donde I es la identidad 2×2. Escribamos B en forma de columnas, como $B = (B_1, B_2)$. Recordemos que la primera columna de AB es igual al producto AB_1 de A por la primera columna de B_1. Si AB es igual a la identidad, entonces AB_1 debe ser igual a la primera columna $E_1 = (1,0)^t$ de la matriz identidad. El mismo razonamiento indica que $AB_2 = E_2 = (0,1)^t$. Por lo tanto, para buscar la inversa B tenemos que resolver los dos sistemas de ecuaciones

$$AB_1 = E_1, \quad AB_2 = E_2.$$

Tenemos una técnica estándar para hacerlo, que es la eliminación gaussiana, que podemos aplicar a ambos sistemas. Pero antes de proceder con cada sistema por separado observemos que en ambos casos tendremos que realizar las mismas operaciones elementales, porque estas dependen de la matriz A, y no del término independiente del sistema de ecuaciones lineales. Entonces podemos realizar **simultáneamente** la escalerización de ambos sistemas. El primero de ellos es

$$\left(\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & 5 & 0 \end{array}\right),$$

y su solución es la primera columna de la inversa. El segundo, que conduce a la segunda columna, es

$$\left(\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & 5 & 1 \end{array}\right).$$

Los representaremos juntos, en la forma

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 5 & 0 & 1 \end{array}\right),$$

que permite aplicar a ambos todos los pasos del procedimiento de escalerización. Primero multiplicamos por 2 la primera fila, y la restamos de la segunda:

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{array}\right).$$

La matriz del sistema ya está escalerizada, por lo que podríamos despejar a partir de aquí. Pero en vez de hacer esto escalerizamos "hacia arriba". Para ello multiplicamos por 3 la segunda fila, y la sumamos a la primera:

$$\left(\begin{array}{rr|rr} 1 & 0 & -5 & 3 \\ 0 & -1 & -2 & 1 \end{array}\right).$$

Para terminar dividimos la segunda fila entre -1, y concluimos

$$\left(\begin{array}{rr|rr} 1 & 0 & -5 & 3 \\ 0 & 1 & 2 & -1 \end{array}\right).$$

La solución de cada uno de los sistema aparece en las columna de la derecha. En la primera columna está la primera columna de la matriz B que estamos buscando, y en la segunda columna la segunda columna de B. O sea, la matriz que se formó a la derecha es justamente la matriz B.

Verificamos, para ver que hemos resuelto correctamente los sistemas de ecuaciones, formando el producto

$$\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Efectivamente, nuestros cálculos fueron correctos y se satisface $AB = I$. Para que B sea la inversa de A también debe verificarse $BA = I$. Hacemos el cálculo de este producto, y obtenemos

$$\begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

de modo que la matriz B que acabamos de calcular es efectivamente la inversa de A. ♣

Veremos luego que sólo hace falta verificar que se satisface $AB = I$ o $BA = I$ para asegurar que B es la inversa de A. Si una de las igualdades se cumple la otra también. Un resultado que parece en principio sorprendente, y que está relacionado con el hecho de que el rango por filas es igual al rango por columnas.

¿Qué ocurre cuando la matriz no tiene inversa? Veamos un ejemplo de esta situación.

Ejemplo 2.3.9. Intentaremos calcular la inversa de

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

El método es el mismo, pero ahora planteamos tres sistemas, cada uno de ellos correspondiente a una de las columnas de la matriz identidad 3×3. Comenzamos por

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 \end{array}\right).$$

En el primer paso restamos la primera fila de la segunda y la tercera. Obtenemos

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{array}\right).$$

Ahora sumamos la segunda fila a la tercera

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 & 1 \end{array}\right).$$

Encontramos entonces que los sistemas son incompatibles, lo que indica que no puede haber una inversa B de la matriz A, porque es imposible satisfacer la ecuación $AB = I$. ♣

Lo que vimos en el ejemplo anterior es completamente general. Cuando la inversa de A no existe entonces alguno de los sistemas $AB_i = E_i$ es incompatible, y el proceso de escalerización se detiene en algún momento.

En los próximos ejercicios proponemos al lector aplicar este algoritmo para la búsqueda de inversas en algunos ejemplos.

Ejercicio 2.41. Calcular la inversa de la matriz

$$\begin{pmatrix} 2 & 2 & -1 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix}$$

Sugerencia: revisar los resultados del ejercicio 1.24 y su continuación, el ejercicio 1.30.

Ejercicio 2.42.

1. Las tres matrices reales siguientes tienen inversas. Calcularlas.

$$A = \begin{pmatrix} 2 & -2 \\ -1 & 3 \end{pmatrix}, \ B = \begin{pmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \ C = \begin{pmatrix} -3 & 2 & -1 & -2 \\ 1 & -1 & 1 & 1 \\ -2 & 1 & -1 & -1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

   Verificar el cálculo.

2. Resolver los siguientes sistemas de ecuaciones lineales usando las matrices inversas calculadas en la parte anterior:

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad B\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}, \quad C\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

3. Calcular la inversa de la matriz

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

sobre $\mathbb{Z}_2$.

4. Hallar la inversa de las siguientes matrices, en las que k y k_i, $i = 1, 2, 3, 4$, indican constantes no nulas:

$$\begin{pmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & k_2 & 0 \\ 0 & k_3 & 0 & 0 \\ k_4 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} k & 0 & 0 & 0 \\ 1 & k & 0 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & k \end{pmatrix}.$$

Ejercicio 2.43. Transformaciones elementales, matrices e inversas

1. Hallar la inversa de la matriz M que se calculó al resolver el ejercicio 2.19, de la página 190. Analizar el efecto que tiene multiplicar a la izquierda por M y por M^{-1} una matriz $3 \times n$ cualquiera.

2. Se considera la matriz

$$M = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -m_{21} & 1 & 0 & \dots & 0 & 0 \\ -m_{31} & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -m_{(n-1)1} & 0 & 0 & \dots & 1 & 0 \\ -m_{n1} & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

Calcular su inversa. Discutir el significado de la multiplicación a la izquierda por M y por M^{-1}.

2.3.3. Existencia de inversa y rango

Consideremos una matriz cuadrada A, de tamaño $n \times n$. En esta sección mostraremos que si existe una matriz B tal que $AB = I$ entonces también se satisface $BA = I$, y B resulta ser la inversa de A. Esto no es para nada obvio, pero resulta relativamente fácil de establecer a partir del concepto de rango.

La igualdad matricial

$$AB = I \tag{2.13}$$

es equivalente a las n igualdades

$$AB_i = E_i, \quad i = 1, \ldots, n,$$

entre columnas de $\mathbb{K}^n$, donde E_i, $i = 1, 2, \ldots, n$ son las n columnas de la matriz identidad $n \times n$. Es decir, se trata de los n vectores en la base canónica de $\mathbb{K}^n$.

Por lo tanto, cuando (2.13) se satisface, todos los vectores de la base canónica de $\mathbb{K}^n$ están en el espacio de columnas de A. Entonces cualquier combinación lineal de estos vectores está en el espacio de columnas $\mathrm{col}(A)$ de A, lo que implica que este espacio es todo $\mathbb{K}^n$. Por lo tanto, a la dimensión de $\mathrm{col}(A)$ es igual a la dimensión de $\mathbb{K}^n$ que es justamente n.

Sabemos que las dimensiones de los espacios de filas y columnas de una matriz coinciden. Este resultado fue establecido en el teorema 1.2, página 150, y condujo a la definición de rango. Esto implica que la dimensión del espacio de filas es n, por lo tanto el espacio de filas es todo $\mathbb{K}^n$.

Por lo tanto, dada cualquier fila $F \in \mathbb{K}^n$ podemos encontrar una combinación lineal de las filas de A que sea igual a F. Esto puede escribirse en forma matricial como $F = CA$, donde C es una fila con los coeficientes de la combinación lineal. En particular, este argumento es cierto para cada una de las filas F_i, de la matriz identidad, por lo que podemos encontrar filas C_i, $i = 1, 2, \ldots, n$, tales que

$$F_i = C_i A, \quad i = 1, 2, \ldots, n.$$

La matriz C formada por las filas C_i satisface $I = CA$. Recordemos que al multiplicar A a la izquierda por C la i-ésima fila del producto CA es el producto de la i-ésima fila C_i de la matriz C por la matriz A. Este producto no es otra cosa que la combinación lineal de las filas de A con los coeficientes almacenados en la fila C_i. Ver, al respecto, la observación 2.2.13 en la que se analiza el producto de matrices desde el punto de vista de las filas.

Hemos encontrado entonces que la existencia de una matriz B tal que $AB = I$ implica que existe una matriz C tal que $CA = I$. Por lo tanto, en virtud del lema 2.2 concluimos que $C = B$, y que ambas matrices son, en definitiva, la inversa A^{-1} de la matriz A.

Ejercicio 2.44. Sea A una matriz cuadrada con entradas en un cuerpo $\mathbb{K}$. Mostrar que si existe una matriz C tal que $CA = I$ entonces C es la inversa de A.

A modo de resumen de los resultados sobre el cálculo de inversas y sistemas de ecuaciones lineales enunciamos el siguiente teorema.

Teorema 2.1. *Sea A una matriz $n \times n$ sobre un cuerpo $\mathbb{K}$. Entonces son equivalentes*

1. *A es invertible;*
2. *existe una matriz $B \in M_{n\times n}(\mathbb{K})$ tal que $AB = I$;*
3. *existe una matriz $C \in M_{n\times n}(\mathbb{K})$ tal que $CA = I$;*
4. *el rango de A es igual a n.*

Dejamos un par de ejercicios al lector:

Ejercicio 2.45. Repasar toda la información acerca de matrices inversas y sistemas lineales, y producir una prueba ordenada del teorema que acabamos de enunciar.

Ejercicio 2.46. Dada una matriz cuadrada A, formular algunas condiciones que sean equivalentes a que A no sea invertible.

El próximo ejercicio agrega algunas equivalencias a la lista del teorema 2.1. Es un ejercicio que nos conduce a un repaso de las propiedades de los sistemas lineales y de los subespacios de $\mathbb{K}^n$ asociados a las matrices.

Ejercicio 2.47. Mostrar que las siguientes condiciones son equivalentes a que una matriz cuadrada A, de tamaño $n \times n$, sea invertible:

1. cualquier sistema $AX = B$ es compatible;
2. el sistema $AX = O$ es determinado;
3. las columnas de A forman una base de $\mathbb{K}^n$;
4. las columnas de A son linealmente independientes;
5. las columnas de A generan $\mathbb{K}^n$;
6. todos los sistemas $AX = E_i$, donde E_i, $i = 1, 2, \dots, n$ son las columnas de la matriz identidad, son compatibles;
7. las filas de A forman una base de $\mathbb{K}^n$;
8. las filas de A son linealmente independientes;
9. las filas de A generan $\mathbb{K}^n$;
10. el espacio $\mathrm{col}(A)$ es todo $\mathbb{K}^n$;
11. el espacio $\ker(A)$ es el subespacio trivial $\{O\}$.

2.3.4. Primeras propiedades de la inversa

Algunas propiedades de las matrices inversas son una consecuencia muy directa de la definición y de la propiedades del cálculo con matrices. Por ejemplo, si una matriz A es invertible, entonces A es la inversa de su inversa.

Ejercicio 2.48. Demostrar la afirmación que acabamos de hacer.

Ejercicio 2.49. Sea A una matriz cuadrada e invertible. Discutir según el valor de λ la invertibilidad de λA.

Es relativamente sencillo conjeturar cuál debe ser la inversa del producto AB de dos matrices invertibles si se sigue el punto de vista de que formar el producto $(AB)X$ es realizar en $\mathbb{K}^n$ una transformación $X \mapsto (AB)X$. Aplicar AB es el resultado de multiplicar primero por B y luego por A. Para deshacer este efecto debemos deshacer primero lo último que hicimos, que fue multiplicar por A, lo que corresponde a multiplicar por A^{-1}. Luego desandamos el camino de multiplicar por B, aplicando B^{-1}. Este procedimiento de multiplicar primero por A^{-1} y luego por B^{-1} es equivalente a multiplicar por el producto $B^{-1}A^{-1}$. Esta matriz debe ser entonces la inversa de AB.

Proposición 2.3. *Sean A y B dos matrices $n \times n$ invertibles entonces el producto AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.*

Prueba. Basta calcular directamente los dos posibles productos entre AB y $B^{-1}A^{-1}$, y usar la propiedad asociativa del producto de matrices. Obtenemos

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1}.$$

Como B^{-1} es la inversa de B entonces $BB^{-1} = I$. Por lo tanto

$$(AB)(B^{-1}A^{-1}) = AIA^{-1} = I.$$

Análogamente

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I.$$

Estas dos igualdades implican el resultado enunciado. □

Ejercicio 2.50. Probar o refutar la siguiente afirmación: si A y B son dos matrices cuadradas e invertibles entonces su suma $A + B$ es invertible.

Proposición 2.4. *Sean A una matriz $n \times n$ invertible. Entonces A^t es invertible y $(A^t)^{-1} = (A^{-1})^t$*

Ejercicio 2.51. Probar la proposición 2.4.

Ejercicio 2.52. ¿Qué propiedades de una matriz real cuadrada, invertible, se conservan en su inversa? Justifique su respuesta.

1. Si A es triangular superior o inferior su inversa también lo es.
2. Si A es simétrica o antisimétrica también lo es su inversa.
3. Si las entradas de A son números enteros también lo son las de su inversa.
4. Si las entradas de A son números racionales también lo son las de su inversa.

Ejercicio 2.53. Mostrar que si A y B son dos matrices invertibles que conmutan entonces sus inversas también conmutan.

Ejercicio 2.54. Mostrar que el producto AB de dos matrices cuadradas A y B es invertible si y sólo si los dos factores A y B son invertibles.

Ejercicio 2.55.

1. Sea A una matriz cuadrada, de dimensiones $n \times n$ y de rango 1. Mostrar que existen dos matrices columna U y V tales que $A = UV^t$.
2. Mostrar que existe una constante real λ tal que $A^2 = \lambda A$.
3. Sea I la matriz identidad $n \times n$, y A una matriz de rango 1. Hallar la condición que debe satisfacer el número λ hallado en la parte 2 para la matriz $I + A$ sea invertible. Mostrar que esa condición es necesaria y suficente.
4. Generalizar los resultados hallados para la inversa de una matriz de la forma $B+A$, donde B es invertible y A tiene rango uno. ¿Cuál es la condición necesaria y suficiente para que $B + A$ sea invertible?

2.3.5. Para tener presente

En lo que sigue A denota una matriz cuadrada $n \times n$.

- Una matriz A tiene inversa si y sólo si su rango es n;
- Una matriz A tiene inversa si y sólo si todo sistema $AX = B$ es compatible;
- Una matriz A tiene inversa si y sólo si el sistema $AX = O$ es determinado;
- Calcular A^{-1} es equivalente a la resolución de n sistemas lineales de ecuaciones.
- Cuando una matriz A es invertible la solución del sistema $AX = B$ es $X = A^{-1}B$. Pero calcular A^{-1} para formar este producto no es la estrategia eficiente para la resolución de ecuaciones lineales.

2.4. Descomposición LU

A lo largo del capítulo 1 hemos hecho un uso extensivo del método de eliminación de Gauss, y aprendido sobre la estructura del espacio de columnas y el espacio de filas de una matriz cualquiera. Ya en este mismo capítulo, en la sección 2.1, mostramos a través de distintos ejemplos como las matrices permiten tratar información de muy variada naturaleza. El objetivo de esta sección es mostrar que el método de Gauss puede representarse en forma matricial, y que esta representación es útil porque permite utilizar una y otra vez la información que el proceso de escalerización genera.

El resultado de nuestro trabajo será una descomposición de una matriz cuadrada A en la forma del producto LU, de una matriz L que sólo tiene entradas no nulas en su diagonal o por debajo de ella, por una matriz U que sólo tiene entradas no nulas en su diagonal o por encima de ella. Las letras L y U aluden a las palabras inglesas *lower* (inferior) y *upper* (superior) respectivamente. La matriz U es simplemente la forma escalerizada de A, en tanto que L almacena toda la información generada durante el proceso de eliminación gaussiana aplicado a A hasta llevarla a la forma escalerizada U. En realidad, **la descomposición LU es la versión matricial de la eliminación gaussiana**.

La descomposición LU es un primer ejemplo de una estrategia de gran importancia en el cálculo con matrices: se busca factorizarlas como un producto de matrices que tienen mayor estructura que la matriz original. Luego se explota la estructura adicional que se ganó con la factorización. Por ejemplo, cuando la descomposición LU se aplica a la resolución de un sistema de ecuaciones $AX = B$, reduce el problema a la resolución de sistemas cuyas matrices están escalerizadas. A lo largo de esta sección iremos desarrollando todas estas ideas.

2.4.1. Registrando los pasos de la eliminación gaussiana

Comencemos por resolver un sistema lineal de ecuaciones

$$AX = B$$

aplicando el que ha sido hasta ahora nuestro método estándar de trabajo, que consiste en recurrir al algoritmo de eliminación gaussiana. Transformamos entonces el sistema en un sistema equivalente

$$EX = C,$$

donde E es una forma escalerizada de A, y el vector C es el que resulta de aplicar a B todos los pasos de la escalerización. Este nuevo sistema es de resolución directa.

Ejemplo 2.4.1. Para fijar ideas en esta discusión supongamos que estamos trabajando con el sistema $AX = B$, con

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ -1 & 1 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}. \tag{2.14}$$

Luego de escalerizar A obtenemos el sistema equivalente $EX = C$, con

$$E = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 7 \end{pmatrix}, \quad C = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \tag{2.15}$$

del que es fácil despejar las componentes

$$x_1 = -1, \quad x_2 = 2, \quad x_3 = 0,$$

del vector incógnita X. ♣

Supongamos ahora que tenemos que resolver un nuevo problema $AX = B_1$, donde A es la matriz que aparece en (2.14). ¿Qué harías? ¿Podrías utilizar la forma escalerizada de la matriz A?

Claro está que no podemos resolver $EX = B_1$ para hallar la solución. La matriz E es la forma escalerizada de A, pero para que esto tuviera sentido tendríamos que aplicar a B_1 todos los pasos de la escalerización. Lamentablemente, a esta altura sólo tenemos la forma escalerizada de A, pero no los pasos de la escalerización. Por lo tanto, si no tuvimos cuidado en registrar lo que hicimos durante la escalerización tendremos que volver a empezar.

Guardemos entonces memoria de lo que vamos haciendo en el proceso de escalerización. Nada nos impide hacer una lista del tipo

> En el primer paso resté de la segunda fila
> cuatro veces la primera.
> Luego sumé un cuarto de la primera a la tercera.
> con la cuarta no hice nada,
> ya había un cero en la primera entrada.
> ...

Como poesía es muy pobre, como algoritmo poco práctico. Una buena manera de almacenar la información de lo que hicimos a lo largo de la eliminación gaussiana es recurrir a una matriz. Examinaremos esta posibilidad recorriendo paso a paso el proceso que nos permitió escalerizar la matriz A del ejemplo 2.4.1. Introduzcamos un superíndice para distinguir a las matrices que van apareciendo a lo largo del proceso de escalerización. Comenzaremos por registrar los multiplicadores que empleamos en cada paso de la eliminación, luego veremos cómo estos multiplicadores intervienen en el cálculo de A a partir de su forma escalerizada. El resultado final será un algoritmo sorprendentemente simple.

Empleando la notación de superíndices escribimos

$$A^{(1)} = A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ -1 & 1 & 4 \end{pmatrix}$$

En el primer paso de la escalerización multiplicamos la primera fila por 2 y la restamos de la segunda. Luego sumamos la primera fila a la tercera, operación que es equivalente multiplicar por -1 y restar. El resultado que obtenemos es la matriz

$$A^{(2)} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 2 & 5 \end{pmatrix}.$$

Tenemos que guardar memoria también de cuáles fueron las operaciones realizadas, para ello necesitamos recordar los multiplicadores $\mathbf{2}$ y $-\mathbf{1}$, junto con la información de dónde los utilizamos. Una forma sencilla de hacer esto –que en realidad adoptamos porque conocemos el resultado final– es conservarlos en la entrada de la matriz A que se eliminó, haciendo aparecer un cero, con la ayuda de cada multiplicador. Escribimos entonces

$$A^{(2)} = \begin{pmatrix} 1 & 1 & 1 \\ \mathbf{2} & -1 & 1 \\ -\mathbf{1} & 2 & 5 \end{pmatrix}.$$

Los números en negrita no deben tenerse en cuenta a los efectos de saber cuál es realmente la matriz $A^{(1)}$. En sus lugares hay que poner ceros. Pero esta notación indica claramente cuáles fueron los dos multiplicadores del primer paso y en qué filas los usamos.

Avanzamos un paso más en la escalerización. Ahora hacemos aparecer un cero en la segunda entrada de la primera fila multiplicando por -2 la segunda fila y restando el resultado de la tercera. Obtenemos así una matriz

$$A^{(3)} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 7 \end{pmatrix},$$

que representamos con nuestra convención de almacenar multiplicadores como

$$A^{(3)} = \begin{pmatrix} 1 & 1 & 1 \\ \mathbf{2} & -1 & 1 \\ -\mathbf{1} & -\mathbf{2} & 7 \end{pmatrix}.$$

Esta matriz $A^{(3)}$ es en realidad la forma escalerizada, a la que hemos decidido llamar U.

Veamos ahora como se han transformado las filas en cada paso. Vamos a designar a la i-ésima fila de cada matriz agregando el subíndice i al nombre de la matriz. Por ejemplo, $A_2^{(3)}$ es la segunda fila de la matriz $A^{(3)}$. Llamaremos m_{ij} a los multiplicadores, indizándolos de acuerdo a la posición en que quedaron almacenados. Así el multiplicador m_{ij} se usó en el j-ésimo paso de la escalerización, en el que pasamos de la matriz $A^{(j)}$ a $A^{(j+1)}$, para producir un cero en la entrada que está en la fila i y la columna j. La operación entre las filas en la que este número intervino fue

$$A_i^{(j+1)} = A_i^{(j)} - m_{ij}A_j^{(j)}. \tag{2.16}$$

Volveremos luego sobre esta fórmula general, pero de momento seamos más modestos, y continuemos con el análisis del ejemplo con el que estábamos trabajando. En el primer paso de la escalerización realizamos las siguientes operaciones:

$$\begin{aligned} A_1^{(2)} &= A_1^{(1)}; \\ A_2^{(2)} &= A_2^{(1)} - 2A_1^{(1)}; \\ A_3^{(2)} &= A_3^{(1)} - (-1)A_1^{(1)}. \end{aligned}$$

Estas fórmulas permiten escribir fácilmente las filas de $A = A^{(1)}$ en términos de las de $A^{(2)}$. Es claro que

$$\begin{aligned} A_1^{(1)} &= A_1^{(2)}; \\ A_2^{(1)} &= A_2^{(2)} + 2A_1^{(2)}; \\ A_3^{(1)} &= A_3^{(2)} - A_1^{(2)}. \end{aligned} \tag{2.17}$$

Es posible hacer una análisis similar del segundo paso:

$$\begin{array}{rcl} A_1^{(3)} & = & A_1^{(2)}; \\ A_2^{(3)} & = & A_2^{(2)}; \\ A_3^{(3)} & = & A_3^{(2)} - (-2)A_2^{(1)}, \end{array}$$

de donde deducimos

$$\begin{array}{rcl} A_1^{(2)} & = & A_1^{(3)}; \\ A_2^{(2)} & = & A_2^{(3)}; \\ A_3^{(2)} & = & A_3^{(3)} - 2A_2^{(3)}. \end{array} \tag{2.18}$$

Si sustituimos las expresiones (2.18) para las filas de la matriz $A^{(2)}$ en las fórmulas (2.17) de las filas de $A^{(1)}$ concluimos

$$\begin{array}{rcl} A_1^{(1)} & = & A_1^{(3)}; \\ A_2^{(1)} & = & 2A_1^{(3)} + A_2^{(3)}; \\ A_3^{(1)} & = & -A_1^{(3)} - 2A_2^{(3)} + A_3^{(3)}. \end{array} \tag{2.19}$$

Si recordamos que $A^{(1)} = A$, llamamos U a la forma escalerizada $A^{(3)}$, y agregamos en (2.19) las filas que no aparecen explícitamente sumándolas con coeficiente 0 obtenemos

$$\begin{array}{rcl} A_1 & = & U_1 + 0U_2 + 0U_3; \\ A_2 & = & 2U_1 + U_2 + 0U_3; \\ A_3 & = & -U_1 - 2U_2 + U_3. \end{array} \tag{2.20}$$

Esta expresión no es otra cosa que la igualdad matricial

$$A = LU,$$

con

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{pmatrix},$$

escrita en términos de las filas de A y U.

Es un cálculo directo verificar la igualdad

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ -1 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 7 \end{pmatrix}.$$

Observación 2.4.2. Nos interesa en este punto destacar cuál es la estructura de L y de U. Comencemos por L:

1. las entradas l_{ij} que están por debajo de la diagonal $(i > j)$ son los multiplicadores de la eliminación gaussiana;
2. la diagonal está formada por unos;
3. por encima de la diagonal sólo hay ceros.

En tanto que la matriz U es la forma escalerizada que se obtuvo por medio del proceso de eliminación gaussiana cuya información quedó almacenada en la matriz L.

Las matrices L y U en la factorización de A son triangulares inferior y superior respectivamente. ♠

Ejemplo 2.4.3. Utilicemos la factorización LU para resolver el sistema $AX = B$, con $B = (0,1,0)^t$. El procedimiento es simple, escribimos el sistema en la forma

$$LUX = B,$$

y a continuación introducimos una variable

$$Y = UX.$$

El sistema original resulta equivalente a

$$\left\{\begin{array}{l} LY = B, \\ UX = Y. \end{array}\right.$$

¡Estupendo! Teníamos un sistema $AX = B$ para resolver y ahora tenemos dos: primero $LY = B$, y una vez calculado Y nos toca hallar X resolviendo $UX = Y$. Parece que hemos retrocedido en vez de avanzar hacia la solución. Sin embargo el retroceso es sólo aparente, porque los dos sistemas con los que ahora tenemos que trabajar ya están en una forma escalerizada, y su resolución es inmediata.

El sistema $LY = B$ es

$$\left(\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 \\ -1 & -2 & 1 & 0 \end{array}\right),$$

para el que determinamos $y_1 = 0$ a partir de la primera ecuación. La segunda nos dice entonces que $y_2 = 1$, y de la tercera concluimos $y_3 = 2$.

Observación 2.4.4. Recordemos que L almacena la información del proceso de escalerización. La solución de $LY = B$ es el mismo vector que se obtendría aplicando a B el proceso de escalerización.

Ejercicio 2.56. Resolver $LY = B$ para $B = (1, 0, 3)$ y comparar el resultado con el vector C en la fórmula (2.15).

En efecto, como $A = LU$ entonces $U = L^{-1}A$, y multiplicar a izquierda por la matriz L^{-1} es lo mismo que realizar el proceso de escalerización. Al resolver el sistema $LY = B$ el resultado es $Y = L^{-1}B$, y el vector Y es justamente el que se obtendría de B si le aplicáramos todos los pasos de la escalerización. Una vez hallado este vector entonces la solución de $AX = B$ es exactamente la misma que la del sistema escalerizado equivalente $UX = Y$

Estas consideraciones en términos de la inversa de L ayudan a entender lo que estamos haciendo, pero digamos que en la práctica no es necesario formar la inversa de L para resolver el sistema. Es más eficiente resolver el sistema triangular $LY = B$ con la técnica de ir despejando las variables y sustituyéndolas en las ecuaciones que aparecen por debajo. ♠

El segundo sistema a resolver es $UX = Y$. Un sistema ya escalerizado, que resulta ser

$$\left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 7 & 2 \end{array}\right)$$

y que tiene solución

$$x_1 = \frac{3}{7}, \quad x_2 = \frac{-5}{7}, \quad \frac{2}{7}.$$

Es decir, tal como anunciábamos, todo el trabajo relativo a la eliminación gaussiana ya está almacenado en la descomposición LU y, una vez dada esta descomposición, los sistemas pueden resolverse directamente despejando las variables en forma sistemática. ♣

Observación 2.4.5. Descomposición LU y espacio de filas.
La posibilidad de escribir a la matriz A como el producto de alguna matriz L por su forma escalerizada U es una consecuencia de que las filas no nulas de U forman una base del espacio de filas de A, tal como se demostró en la proposición 1.23, página 147.

En consecuencia, cada fila de A puede escribirse como una combinación lineal de las filas de U, lo que implica que A puede expresarse como el producto LU, de una matriz L que multiplica U a la izquierda, por U.

También es de esperar que los coeficientes que hay que poner en L tengan que ver con los multiplicadores utilizados durante la eliminación gaussiana. Y un momento de reflexión convencerá al lector de que L debe ser triangular inferior, porque una vez que una fila se ha usado para producir ceros en entradas que están por debajo de ella no se ve modificada en pasos posteriores del proceso de eliminación gaussiana.

Por lo tanto, la forma general de la descomposición LU puede preverse a partir de las propiedades del producto de matrices y del proceso de eliminación de Gauss. Es de destacar la sencilla dependencia en los multiplicadores del proceso de eliminación que tiene L, lo que se traduce en un algoritmo de muy fácil implementación.

Un adelanto de estas ideas, que enfatizan la representación matricial de las operaciones con las filas de la matriz, apareció en el ejercicio 2.19, de la página 190. Proponemos a continuación al lector una extensión de este ejercicio, que aporta otra mirada sobre la descomposición LU.

Ejercicio 2.57. Consideremos la matriz A de la fórmula (2.14), y el proceso de escalerizacion que originó la secuencia $A = A^1$, A^2, $A^3 = U$.

1. Para cada paso de la escalerización hallar matrices M_i y N_i tales que

$$A^{i+1} = M_i A^i, \quad A^i = N_i A^{i+1}, \qquad i = 1, 2.$$

2. Formar los productos $M_2 M_1$ y $N_2 N_1$. ¿Qué relación guardan entre ellos y con la matriz L de la descomposición LU de A?

3. Para una matriz A cualquiera, ¿podrías explicar cómo construir la matriz L de la descomposición LU de A como un producto de matrices que representen operaciones elementales que aparecen durante el proceso de escalerización de A? ♠

Ejercicio 2.58. Investigar cómo extender la descomposición LU al caso de matrices que no sean cuadradas.

Ejercicio 2.59. Para la matriz (2.14) y $B = (1, 2, 3)^t$ resolver el sistema $AX = B$. ♣

Ejercicio 2.60.

1. Hallar la descomposición LU de las siguientes matrices:

$$A = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 2 & -1 \\ -2 & 1 & 3 \\ 4 & 1 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 2 & 0 & -1 \\ -1 & 2 & 0 & 1 \\ -1 & 2 & 2 & -1 \\ 2 & 1 & -2 & -1 \end{pmatrix}.$$

2. Resolver los siguientes sistemas de ecuaciones utilizando la descomposición anterior:

$$AX_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad BX_2 = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}, \quad CX_3 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 3 \end{pmatrix}.$$

Observación 2.4.6. Malas noticias: no siempre existe LU
Hemos calculado descomposiciones LU de unas cuantas matrices. Pero esto no debe hacernos creer que tal descomposición siempre existe. Por ejemplo, la matriz

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

no admite una descomposición LU en dos factores con la estructura que se describe en la observación 2.4.2. Tal factorización sería de la forma

$$\begin{pmatrix} 1 & 0 \\ l_{21} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} \\ l_{21}u_{11} & l_{21}u_{11} + u_{22} \end{pmatrix}.$$

La igualdad con A de este producto implica

$$u_{11} = 0, \quad u_{12} = 0,$$

por lo que

$$l_{21}u_{11} = 0,$$

lo que está en contradicción con que esta entrada debería tomar el valor 1.

Lo que hace fracasar el intento de descomponer A en la forma LU es la necesidad de realizar una operación de intercambio de filas para escalerizar A. Una vez que se hace el intercambio, la matriz resultante admite una descomposición LU, en efecto

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

es una descomposición trivial de la matriz que resulta de intercambiar la primera y segunda fila de A. ♠

En lo que sigue sistematizaremos todo lo aprendido acerca de la obtención de la descomposición LU. Trabajaremos primero en el caso en que ningún intercambio de filas es necesario durante la escalerización de A. Luego trataremos el caso en que el intercambio sí es necesario.

2.4.2. Descomposición *LU* sin intercambio de filas

En esta sección haremos una presentación sistemática de la descomposición *LU* para matrices que pueden ser escalerizadas sin realizar ningún intercambio de filas. Esencialmente se trata de escribir en forma ordenada, general, lo que hicimos a lo largo de la sección anterior.

Nuestro análisis estará basado en una extensión de la fórmula

$$A_i^{(j+1)} = A_i^{(j)} - m_{ij} A_j^{(j)}. \tag{2.21}$$

que describe la operación entre filas en la que entra cada multiplicador m_{ij}. Algunas observaciones nos simplificarán la vida:

1. Cada multiplicador m_{ij} se emplea en el paso j, para $j = 1, 2, \ldots, n-1$, de la eliminación, para eliminar una entrada que está en la fila i restando un múltiplo adecuado de la fila j a la fila i. En esta operación la fila que se resta siempre está por encima de la fila en la que eliminamos una entrada. Tenemos entonces que los multiplicadores están definidos para $i > j$. En particular, la fórmula (2.21), sólo está definida para $i > j$.

2. En el paso j de la eliminación las primeras j filas de la matriz no cambian. Podemos representar este hecho por medio de la fórmula (2.21), conviniendo que $m_{ij} = 0$ si $i \leq j$.

3. Con la notación que estamos usando la matriz $A^{(1)}$ es justamente la matriz A con la que comenzamos la escalerización, en tanto que $A^{(n)}$ es U.

4. Las fila $A_j^{(j)}$ que aparece en (2.21) no cambia en el paso j (en el que la fila se usa para producir cero por debajo de ella) ni en los siguientes. Por lo tanto, la fila $A_j^{(j)}$ es en realidad la j-ésima fila de U.

Este conjunto de observaciones nos permite escribir cada paso de la escalerización en la forma

$$A_i^{(j)} - A_i^{(j+1)} = m_{ij} U_j, \quad j = 1, 2, \ldots, n-1,\ i = 1, 2, \ldots, n, \tag{2.22}$$

donde los números m_{ij} son los multiplicadores de la eliminación gaussiana para $i > j$, y 0 para $i \leq j$. Podemos dar una fórmula explícita para los m_{ij} que no son triviales en términos de las entradas $a_{ik}^{(j)}$ de la matriz $A^{(j)}$:

$$m_{ij} = \frac{a_{ij}^{(j)}}{a_{jj}^{(j)}}, \quad 1 \leq j < i \leq n.$$

Para cada j la fórmula (2.22) nos dice cómo se actualiza la fila i. El efecto acumulado de todos los cambios sobre la fila i se obtiene sumando en el índice j (que indiza los pasos de la eliminación gaussiana) desde $j = 1$ hasta $n-1$. Obtenemos

$$A_i^{(1)} - A_i^{(n)} = \sum_{j=1}^{n-1} m_{ij} U_j, \quad i = 1, 2, \dots, n.$$

Todavía podemos simplificar un poco más las cosas teniendo en cuenta que $A_i^{(1)} = A_i$ y $A_i^{(n)} = U_i$. Entonces

$$A_i = \sum_{j=1}^{n-1} m_{ij} U_j + U_i, \quad i = 1, 2, \dots, n.$$

En esta expresión hemos conseguido escribir las filas de A como combinación lineal de las filas de U. En la sumatoria aparecen todas las filas de U salvo la última, porque la última fila no se utiliza en la eliminación para continuar el proceso por debajo de ella: no hay nada allí. Pero no perdemos nada con sumar hasta $j = n$ e incluirla, porque los coeficientes m_{in} son todos nulos. Haremos esta modificación, para que todas las filas de U entren en la fórmula:

$$A_i = \sum_{j=1}^{n} m_{ij} U_j + U_i, \quad i = 1, 2, \dots, n.$$

El coeficientes m_{ii} es igual a cero, por lo tanto el término $m_{ii} U_i$ que la sumatoria aporta cuando $j = i$ es nulo. Podemos incorporar el término U_i extra a la sumatoria modificando ligeramente los coeficientes, lo que haremos introduciendo coeficientes l_{ij}, para $i, j = 1, 2, \dots, n$, tales que

$$l_{ij} = \begin{cases} m_{ij}, & j < i, \\ 1, & j = i, \\ 0, & j > i. \end{cases}$$

Una vez hecho esto obtenemos

$$A_i = \sum_{j=1}^{n} l_{ij} U_j, \quad i = 1, 2, \dots, n.$$

Esta expresión es justamente el producto matricial

$$A = LU, \quad L = (l_{ij}),$$

expresado a través de sus filas. Observemos que L es una matriz triangular inferior que tiene exactamente la estructura que habímos descrito en la observación 2.4.2.

2.4.3. Descomposición LU con pivoteo: $PA = LU$.

Como ya sabemos, en algunos casos la eliminación gaussiana no puede realizarse sin hacer intercambios de filas, operación a la que es usual referirse con el término *pivoteo*, lo que explica el título de esta sección. Como vimos en la observación 2.4.6 esta circunstancia está relacionada con el hecho de que algunas matrices no tienen una descomposición LU. La novedad que presentamos en esta sección es que para cualquier matriz existe una permutación de filas que produce una matriz que tiene descomposición LU. No es sorprendente: la descomposición LU es la representación matricial del proceso de escalerización, que siempre puede realizarse si se acepta intercambiar filas de vez en cuando.

Estudiaremos ahora cómo incorporar el pivoteo a la descomposición LU. Por ejemplo, para

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 2 & -2 & 4 \\ 1 & -1 & 1 \end{pmatrix}.$$

Naturalmente, no podemos comenzar la escalerización usando el 0 que está en la posición a_{11} como pivote. Pero podemos intercambiar filas, por ejemplo poniendo la tercera en el primer lugar, y a partir de ahí escalerizar. A continuación representamos el primer paso de la escalerización luego del intercambio de filas, con la convención de escribir los multiplicadores en negrita en el lugar que corresponde a las entradas que eliminamos con su ayuda:

$$\begin{pmatrix} 1 & -1 & 1 \\ \mathbf{2} & 0 & 2 \\ \mathbf{0} & 1 & 1 \end{pmatrix}.$$

Por las dudas vayamos recordando cuál fue la permutación que hicimos: cambiamos la fila 1 con la 3, lo que puede representarse haciendo la misma permutación sobre el vector $(1, 2, 3)$ que se transforma en $(3, 2, 1)$. Tampoco podemos hacer el segundo paso sin permutar filas, pero el simple intercambio de las filas 2 y 3 ya nos conduce a una matriz escalerizada. El multiplicador que usamos es, por supuesto, igual a 0. Representamos este último paso como

$$\begin{pmatrix} 1 & -1 & 1 \\ \mathbf{0} & 1 & 1 \\ \mathbf{2} & \mathbf{0} & 2 \end{pmatrix}, \tag{2.23}$$

y registramos la nueva permutación de filas aplicándola a $(3, 2, 1)$ para obtener $(3, 1, 2)$. A partir de la expresión (2.23) podemos fabricar una descomposición

LU, con factores

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Al multiplicar estos factores LU obtenemos

$$LU = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 2 & -2 & 4 \end{pmatrix},$$

matriz en la que reconocemos la filas de A, pero cambiadas de orden. La tercera fila de A es la primera de LU, la primera de A la segunda de LU y la segunda de A reapareció como la tercera en el producto LU. Estas son justamente las permutaciones que hicimos a lo largo de la escalerización. Para verlo basta observar que el vector en el que fuimos recogiendo las permutaciones es justamente el $(3, 1, 2)$, que resume exactamente la misma información que acabamos de presentar acerca de las permutaciones de filas.

El resultado que acabamos de observar es completamente general: si es necesario realizar permutaciones durante el proceso de escalerización de A el algoritmo de escalerización y almacenamiento de los multiplicadores que hemos presentado conduce a una descomposición LU de una matriz que es el resultado de aplicar a A todas las permutaciones.

Vamos a justificar ahora con toda generalidad la afirmación que cierra el párrafo anterior. Consideraremos que A es una matriz $n \times n$ cualquiera. Supongamos que la escalerización requiere algunos intercambios de filas. Si tuviéramos la enorme suerte de que un duende amistoso revelara para nosotros cuáles son las permutaciones necesarias podríamos realizarlas antes de empezar el algoritmo de escalerización, y luego podríamos escalerizar sin intercambiar filas. La información acerca de las permutaciones puede almacenarse en una matriz P, tal que el resultado de aplicar a A las permutaciones es la matriz PA, tal como vimos en el ejercicio 2.20, página 190. Es posible incluso que el duende haya hecho ese ejercicio y ya nos dé la información sobre las permutaciones en la forma de la matriz P.

Como la escalerización de PA no requiere permutación alguna entonces podemos escribir PA como el producto de los factores LU de su descomposición LU, y conseguimos una descomposición

$$PA = LU.$$

Resumimos toda la discusión anterior en una proposición.

Proposición 2.5 (Descomposición $PA = LU$). *Si A es una matriz cuadrada $n \times n$ existe una matriz P de tamaño $n \times n$ cuyas filas son una permutación de las filas de la matriz identidad I_n, y una descomposición de PA como el producto LU de dos matrices L y U de tamaño $n \times n$, con L triangular inferior que tiene unos en su diagonal, y U triangular superior.*

También vale la siguiente proposición acerca de la descomposición LU para matrices invertibles.

Proposición 2.6 (Descomposición LU). *Si A es una matriz cuadrada e invertible y existe una descomposición $A = LU$, donde L y U tiene las propiedades que aparecen en el enunciado de la proposición 2.5, entonces la descomposición es única.*

PRUEBA. Como A es invertible entonces los dos factores L y U también deben serlo. Este es un resultado que encontramos en la sección 2.3, al resolver el ejercicio 2.54, en la página 209.

Si tuviéramos dos factorizaciones diferentes,

$$A = LU = L'U'$$

entonces entonces

$$L'^{-1}L = U'U^{-1}.$$

Pero U y U' son matrices triangulares superiores por lo que U^{-1} es triangular superior y el producto $U'U^{-1}$ es triangular superior. Análogamente L'^{-1} es una matriz triangular inferior, con unos en su diagonal. Por lo tanto $L'^{-1}L$ es triangular inferior y tiene unos en la diagonal. El lector habrá encontrado estos resultados acerca de los productos e inversas de matrices triangulares al resolver los ejercicios 2.25, en la página 191 de la sección 2.2 y 2.52, en la página 209 de la sección 2.3.

De la igualdad

$$L'^{-1}L = U'U^{-1}$$

se deduce que todos los elementos que no estén en la diagonal principal de estas matrices son nulos. Pero como los de la diagonal principal de $L'^{-1}L$ son unos finalmente obtenemos

$$I = L'^{-1}L = U'U^{-1}.$$

Esto implica $L = L'$ y $U = U'$. □

Ejercicio 2.61. Dar dos factorizaciones LU diferentes de una matriz real 3×3 no invertible.

Ejercicio 2.62. Investigar la unicidad de la descomposición $PA = LU$.

A los efectos del uso de la factorización $PA = LU$ digamos que no debemos preocuparnos demasiado si no tenemos un amigo duende interesado en el álgebra matricial: las permutaciones necesarias, almacenadas en la matriz P, son las que se hacen a lo largo del proceso de escalerización, y podemos ir construyendo P durante la eliminación. Por ejemplo, almacenando la información sobre las permutaciones en un vector, o realizando las permutaciones de filas sobre la matriz identidad[7]. En el ejemplo que hemos venido tratando reconstruimos la matriz P de permutaciones inspeccionando el vector $(3, 1, 2)$, que nos dice que hay que poner en P las filas tercera, primera y segunda de la identidad, en ese orden:

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

La descomposición $PA = LU$ resulta, en este caso,

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 2 & -2 & 4 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Ejemplo 2.4.7. Resolvamos ahora el sistema $AX = B$ con $B = (2, -2, -2)^t$ usando la descomposición anterior. Para eso transformamos el sistema en otro equivalente $PAX = PB$, que a su vez es equivalente al sistema $LUX = PB$. Debemos resolver entonces dos sistemas ya escalerizados:

$$\begin{cases} LY = PB \\ UX = Y \end{cases}$$

El primer sistema de ecuaciones es:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}.$$

Resolviendo este sistema tenemos $y_1 = -2$, $y_2 = 2$, $y_3 = 2$. Ahora ya podemos resolver el segundo sistema:

[7] Por esta razón no necesitaremos el duende en el resto del texto, y ya podemos ir recomendándole que busque trabajo en la continuación de la trilogía de “El señor de los anillos”. Otra posibilidad es usar sus poderes para algo distinto que descomponer matrices.

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}.$$

Por tanto obtenemos el resultado: $x_1 = -2,\ x_2 = 1,\ x_3 = 1.$

Ejercicio 2.63.

1. Hallar la descomposición $PA = LU$ de la matriz

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 0 \\ 1 & -2 & 3 \end{pmatrix}.$$

2. Resolver el sistema $AX = \begin{pmatrix} -3 \\ -3 \\ -3 \end{pmatrix}$ usando la descomposición anterior.

2.4.4. Pivoteo por razones de precisión

Hemos visto como modificar la descomposición LU para incorporar a nuestro esquema de cálculo la posibilidad de introducir intercambios de filas. En algunos casos este intercambio es imprescindible porque aparece un cero en el lugar de un pivote, y no puede utilizarse esa fila para proseguir con la eliminación gaussiana. En esta sección mostraremos otra circunstancia en la que el pivoteo es imprescindible: cuando se trabaja con una precisión limitada y aparece un pivote muy pequeño.

Presentamos un ejemplo de un sencillo sistema de ecuaciones lineales. Se trata de resolver $AX = B$, con

$$A = \begin{pmatrix} 1{,}00 & 1{,}10 & 5{,}00 \\ 1{,}10 & 1{,}20 & 2{,}00 \\ 5{,}00 & 2{,}00 & 26{,}00 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Como es usual, representamos el sistema por su matriz ampliada

$$A|B = \left(\begin{array}{ccc|c} 1{,}00 & 1{,}10 & 5{,}00 & 0 \\ 1{,}10 & 1{,}20 & 2{,}00 & 1 \\ 5{,}00 & 2{,}00 & 26{,}00 & 1 \end{array}\right).$$

En nuestro primer paso de la escalerización utilizamos el 1 que está en la primera entrada de la primera fila como *pivote*. Restamos entonces a la segunda fila de A el resultado de multiplicar la primera fila por 1,1, y a la tercera el resultado de multiplicar la primera fila por 5. Obtenemos así la matriz

$$\left(\begin{array}{rrr|r} 1 & 1{,}10 & 5{,}0 & 0 \\ 0 & -0{,}01 & -3{,}5 & 1 \\ 0 & -3{,}50 & 1{,}0 & 1 \end{array}\right).$$

Un paso más en la escalerización, en el que empleamos la entrada $-0{,}01$ como pivote, consiste en restar a la tercera fila el resultado de multiplicar a la segunda por 350. Esto nos lleva a

$$\left(\begin{array}{rrr|r} 1 & 1{,}10 & 5{,}0 & 0 \\ 0 & -0{,}01 & -3{,}5 & 1 \\ 0 & 0{,}00 & 1226{,}0 & -349 \end{array}\right).$$

A partir de la forma escalerizada es muy fácil resolver el sistema. Si llamamos $X = (x_1, x_2, x_3)$ al vector solución entonces

$$x_1 \approx 1{,}8271, \quad x_2 \approx -0{,}3670, \quad x_3 \approx -0{,}2847.$$

Supongamos ahora que implementamos el cálculo en una computadora que sólo tiene dos dígitos en la mantisa.

Observación 2.4.8. PUNTO FLOTANTE, MANTISAS Y EXPONENTES
Estamos haciendo referencia aquí al sistema de *punto flotante* que emplean las computadores y calculadoras para representar los números reales. Utilizan algunos dígitos para la mantisa, dónde se almacenan las cifras más significativas del número que se esta representando, y otros dígitos para un exponente que da idea del tamaño del número.

No es posible representar exactamente todos los números por este procedimiento, y los resultados de prácticamente todos los cálculos se ven afectados por algún error debido a este sistema de almacenamiento.

Para poder simular los resultados de un cálculo hecho en una computadora emplearemos una precisión infinita para los cálculos intermedios, pero los resultados finales serán truncados a la precisión fijada por la longitud de la mantisa con la que estamos trabajando.

La precisión de las máquinas actuales es mucho mayor que la que estamos usando. Nuestra elección está justificada por el deseo de ilustrar con un ejemplo sencillo un problema que afecta a los cálculos reales que se realizan con una precisión mayor, pero que también involucran muchas más variables y muchos más pasos en los que los errores se pueden ir acumulando. ♠

El primer paso en la escalerización del sistema no plantea problema alguno, todos los números resultantes pueden representarse correctamente con sólo dos dígitos significativos. Pero en el segundo y último paso algunas aproximaciones son necesarias, y obtenemos

$$\left(\begin{array}{rrr|r} 1 & 1{,}1 & 5{,}0 & 0 \\ 0 & -0{,}01 & -3{,}5 & 1 \\ 0 & 0 & 1200{,}0 & -350 \end{array}\right).$$

Al resolver a partir de esta forma escalerizada calculamos la solución

$$x_1 \approx -0{,}8333, \quad x_2 \approx 2{,}0833, \quad x_3 \approx -0{,}2917.$$

¡El error es descomunal! Ni una de las cifras de x_1 y x_2 es correcta. ¿Qué ha ocurrido? Errores menores al 5 % en nuestra representación de los números producen resultados disparatados, con errores relativos que superan al 600 %. La respuesta está en el cálculo que hicimos en el segundo paso de nuestra escalerización: multiplicamos la segunda ecuación por 350, la sumamos a la tercera y sustituimos la tercera ecuación por esta combinación lineal. Observemos que la ecuación sustituida, la tercera, fue cambiada por una combinación lineal en que la segunda ecuación aparece con un peso 350 veces mayor. Dicho de otra forma, esta tercera ecuación sólo contribuye con un factor de 1/350 en la combinación, y lo que ella aporta cae fuera de los dígitos que podemos representar. De hecho, buena parte de la información contenida en la tercera ecuación ha desaparecido de la forma escalerizada. Notemos que el sistema de numeración de nuestra hipotética computadora interpretaría

$$-350 \times (-3{,}5) + 1 = 1226 \approx 1200, -350 \times 1 + 1 = -349 \approx -350.$$

En esta aproximación vemos que los unos de los segundos sumandos del los términos de la izquierda, que son la contribución de la tercera de las ecuaciones del sistema, se ven completamente absorbidos por los errores de redondeo. El origen de este problema está en que el pivote $-0{,}01$ es, en valor absoluto, mucho más chico que el coeficiente $-3{,}5$ que queremos “hacer desaparecer.en este paso de la escalerización. Esto es lo que lleva a utilizar un multiplicador grande.

La solución pasa por escoger como pivote para la escalerización el número que tiene el mayor valor absoluto. En este caso usamos el $-3{,}5$ que está en la tercera fila, lo que implementamos por medio de un intercambio de filas. Naturalmente, esta operación no altera las soluciones del sistema, por lo que podemos continuar tranquilamente con nuestro proceso a partir de

$$\left(\begin{array}{rrr|r} 1 & 1{,}10 & 5{,}0 & 0 \\ 0 & -3{,}50 & 1{,}1 & 1 \\ 0 & -0{,}01 & -3{,}5 & 1 \end{array}\right).$$

Con esta permutación conseguimos que el factor que multiplica la segunda fila sea 1/350, y el resultado que, en nuestra aritmética de sólo dos cifras significativas, arroja este nuevo paso en la escalerización es

$$\left(\begin{array}{rrr|r} 1 & 1{,}1 & 5{,}0 & 0 \\ 0 & -3{,}5 & 1{,}1 & 1 \\ 0 & 0{,}0 & -3{,}5 & 1 \end{array}\right).$$

Si escribiéramos con algunas cifras más veríamos alguna diferencia en la tercera fila de la matriz. Una mejor aproximación es

$$\left(\begin{array}{rrr|r} 0 & 0 & -3{,}503 & 0{,}9971 \end{array}\right).$$

Al calcular las soluciones a partir de esta nueva forma escalerizada y con dos cifras en la mantisa obtenemos

$$x_1 \approx 1{,}8, \quad x_2 \approx -0{,}37, \quad x_3 \approx -0{,}28.$$

Este resultado coincide, dentro del margen de precisión que estamos manejando, con la solución correcta. El intercambio de filas, operación a la que nos referiremos con el nombre de *pivoteo*, permitió recuperar un valor cercano al verdadero.

Concluimos entonces que es necesario realizar esta operación de pivoteo para evitar que en la eliminación gaussiana aparezcan multiplicadores muy grandes. En general, digamos que es recomendable escoger en cada paso el pivote con el mayor valor absoluto posible. Observemos que esta operación es inevitable cuando se emplea aritmética exacta y alguno de los pivotes se hace cero. Cuando la aritmética no es exacta los pivotes pequeños deben ser evitados. Por supuesto que “pequeño” o “grande” son términos relativos. Aquí “pequeño” significa de un orden de magnitud mucho menor que las restantes entradas de la matriz que están por debajo del pivote en la misma columna.

2.4.5. Para tener presente

- Toda matriz cuadrada A que pueda escalerizarse sin permutar filas admite una factorización $A = LU$, con L triangular inferior, U triangular superior, y todas las entradas de la diagonal de L iguales a 1.

- Esta factorización es una manera matricial de escribir el resultado de un proceso de escalerización aplicado a A. U es una forma escalerizada de A.
- Conocer la descomposicion LU reduce el sistema $AX = B$ a dos sistemas ya escalerizados.
- Cuando en el proceso de escalerizacion hay que hacer intercambio de filas la factorización es $PA = LU$, donde P es la matriz que hace las permutaciones de filas.
- Cuando se trabaja con una aritmética que no es exacta el intercambio de filas (pivoteo) es necesario para evitar los pivotes pequeños.

2.5. Determinantes

El objetivo de esta sección es presentar la teoría de los determinantes de matrices reales. El determinante es un número que se asocia a una matriz $n\times n$ o, equivalentemente, a una familia de n vectores de longitud n (las n filas o las n columnas de la matriz), y que contiene importante información acerca de la matriz.

Nuestro deseo de hacer una exposición completa ha extendido mucho esta sección, por lo que a continuación incluimos algunos comentarios para ayudar al lector a encontrar su camino a través de todo este material.

Hemos comenzado por introducir explícitamente el determinante para las matrices 2×2, que aparece naturalmente cuando se estudia un sistema lineal con dos ecuaciones y dos incógnitas. En nuestra presentación hemos enfatizado que el determinante es una función multilineal y alternada de las filas (columnas) de las matrices a las que se aplica, lo que permite extender con una gran sencillez la noción de determinante para matrices cuadradas de cualquier tamaño. Para las matrices 3×3 también damos una fórmula explícita. Todo este material se cubre en la subsección 2.5.1. Antes de continuar con la descripción de esta sección digamos al lector que en el resto del texto trabajaremos fundamentalmente con determinantes de matrices 2×2 y 3×3.

Las propiedades de la función determinante que se enfatizan en la sección 2.5.1 implican que es posible calcularlo eficientemente por medio de la eliminación gaussiana. Este material se discute en la subsección 2.5.3. También se presentan allí los importantes teoremas acerca del determinante del producto de matrices y el determinante de la matriz traspuesta (teoremas 2.4 y 2.5 respectivamente). En una primera lectura es posible familiarizarse con los enunciados de estos dos teoremas y aplicarlos, sin entrar en los detalles de sus demostraciones.

La expresión general de los determinantes de matrices $n \times n$ se discute en la subsección 2.5.2, que puede obviarse en una primera lectura. La subsección 2.5.4 contiene algunos resultados clásicos como la regla de Cramer, los desarrollos por filas y columnas, y una fórmula para la inversa de una matriz. No haremos uso de ninguno de estos resultados más adelante, y el lector también puede saltarse esta sección sin perjuicio para el aprovechamiento del resto del texto.

Por último, la sección 2.5.5 contiene la información sobre aspectos geométricos de la teoría de los determinantes: el cálculo de áreas y volúmenes, y la distinción entre las dos posibles orientaciones del plano y el espacio. Su lectura es un requisito previo para la buena comprensión de algunos conceptos que discutiremos en el capítulo dedicado a la geometría.

Antes de lanzarnos a desarrollar la teoría, destacamos cuatro propiedades fundamentales de los determinantes:

1. el cálculo del determinante de una matriz se reduce a escalerizar la matriz. Hay otras maneras de calcularlo, pero ésta es la mas económica desde el punto de vista computacional.

2. el determinante de una matriz cuadrada A es distinto de cero si A es invertible, y nulo cuando A es singular o no invertible.

3. El valor absoluto del determinante de una matriz real A de dimensiones 2×2 es igual al área del paralelogramo cuyos lados están representados por las filas (o columnas) de A. Para una matriz 3×3 vale un resultado análogo: el valor absoluto de su determinante es igual al volumen del paralelepípedo cuyas aristas son las filas o columnas de la matriz A. Este resultado se extiende a dimensiones más altas: es posible definir una noción de *volumen*[8] en el espacio n-dimensional $\mathbb{R}^n$ de forma tal que el volumen del paralelepípedo n dimensional que tiene a las n filas (o columnas) de una matriz A, real y cuadrada, como aristas es igual al valor absoluto del determinante de A.

4. Hay dos maneras de orientar el espacio. Por ejemplo, en el plano podemos girar en el sentido de las agujas del reloj o en el contrario. En el espacio de dimensión 3 los objetos admiten dos orientaciones: la propia o la de su imagen en un espejo. El signo del determinante distingue entre estas dos posibles orientaciones.

En esta sección **todas las matrices serán reales**, aunque no formulemos explícitamente esta condición.

Nota Histórica 2.1. Es interesante señalar que la teoría de los determinantes es más antigua que la de las matrices, aunque el orden del material en este texto parezca sugerir lo contrario.

La primera aparición explícita en Europa del concepto de determinante, aunque no del nombre, es debida al matemático alemán Gottfried Wilhelm Leibnitz (1646-1716), que en una carta a de L'Hôpital (1661-1704) discute las soluciones de un sistema lineal de ecuaciones en términos de un determinante. La teoría de los determinantes se desarrolló a lo largo de los años, y el término que actualmente los designa apareció en un trabajo de 1812 de Agustín Cauchy (1789-1857) que presentaba en forma sistemática mucho de la teoría existente a esa fecha, además de algunas contribuciones originales.

---

[8]Este volumen juega un papel importante en los teoremas de cambio de variable para integrales múltiples, un tema que se trata en los cursos de Cálculo. El problema de integrar en una cierta región se convierte en integrar en otra región, y el cambio de variables es la función que nos lleva de una región a otra. El determinante de una cierta matriz asociada a la función que hace el cambio de variables es quien controla el cambio en la medida de la región original, cuando ésta es transformada en la otra, de modo tal que ambas integrales sean iguales.

La noción de matriz apareció más tarde, y supeditada a la teoría de determinantes: el término *matrix* fue introducido por el matemático inglés Sylvester (1814-1897) como el nombre para un arreglo de números. Su elección del término obedece a que Sylvester veía a la matriz como un "generador" de determinantes. La teoría de las matrices emergió como tal con los trabajos de su compatriota Arthur Cayley (1821-1895), que publicó en 1858 el artículo *Memoir on the Theory of Matrices*.

A lo largo del siglo XX, con la sistematización del Álgebra Lineal, y fundamentalmente a partir de la aparición de las computadoras digitales, la teoría de las matrices fue ganando espacio dentro de la matemática y sus aplicaciones, hasta ocupar el lugar central que tiene hoy en día. Tal como esta breve historia sugiere, la Matemática está evolucionando permanentemente, y cambiando su aspecto y el centro de sus intereses a lo largo del tiempo.

2.5.1. Definición y propiedades fundamentales del determinante

Antes de dar la definición de determinante en el caso general mostraremos cómo este número aparece naturalmente al analizar un sistema de ecuaciones lineales con dos ecuaciones y dos incógnitas. Aplicaremos el método de escalerización al sistema general

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned} \tag{2.24}$$

Lo haremos, naturalmente, con la más económica notación matricial

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array}\right).$$

Si $a_{11} \neq 0$, entonces sustituimos la segunda fila por el resultado de multiplicarla por a_{11} y restarle el producto de a_{21} por la primera fila. Obtenemos

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ 0 & \mathbf{a_{11}a_{22} - a_{21}a_{12}} & a_{11}b_2 - a_{21}b_1 \end{array}\right) \longleftarrow a_{11}F_2 - a_{21}F_1. \tag{2.25}$$

Este sistema es equivalente al original, porque la fila sustituida entró con coeficiente $a_{11} \neq 0$ en la combinación lineal que la sustituyó. Hemos destacado la cantidad

$$a_{11}\,a_{22} - a_{12}\,a_{21} \tag{2.26}$$

porque este número determina el comportamiento del sistema. Si (2.26) es distinto de cero entonces el sistema tiene solución única para cualquier elección de b_1 y b_2. Equivalentemente, cuando esta condición se satisface la matriz del sistema es invertible.

La discusión del párrafo anterior está hecha bajo la hipótesis de que la entrada a_{11} de la matriz del sistema (2.24) es distinta de cero. Completémosla estudiando el caso $a_{11} = 0$. Si $a_{11} = 0$ entonces el sistema se escaleriza intercambiando sus dos filas. El resultado es

$$\left(\begin{array}{cc|c} a_{21} & a_{22} & b_2 \\ 0 & a_{12} & b_1 \end{array}\right).$$

Entonces es necesario y suficiente que los números a_{12} y a_{21}, que ocupan las posiciones de los pivotes, sean no nulos para que el sistema tenga solución única independientemente de los valores que tomen b_1 y b_2, o, equivalentemente, para que la matriz del sistema no sea singular. Por lo tanto, para que esto ocurra debe satisfacerse $0 \neq a_{12}\, a_{21}$, que, bajo la hipótesis $a_{11} = 0$ es equivalente a la no anulación de (2.26).

En resumen, si tenemos una matriz

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

en el número (2.26) está encerrado el grueso de la información acerca del comportamiento de los sistemas de ecuaciones con matriz A. Equivalentemente, conociendo este número sabremos si A es o no es invertible. Llamaremos a este número el *determinante* de la matriz A. Como es un concepto importante lo recogemos en nuestra próxima definición.

Definición 2.5 (Determinante de las matrices 2×2). *Dada una matriz real*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

de dimensiones 2×2 *definimos su* ***determinante*** *como*

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}. \tag{2.27}$$

Para el determinante de una matriz cuadrada A usaremos las notaciones $\det(A)$ o $|A|$. Como un dato curioso, mencionemos que la segunda notación fue introducida por Cayley (ver la nota histórica 2.1).

El determinante tal como se introduce en la definición 2.27 puede entenderse como una función

$$\det : M^{2\times 2}(\mathbb{R}) \to \mathbb{R}$$

que asocia a cada matriz cuadrada A un número, el $\det(A)$, según la regla

$$A \mapsto a_{11}a_{22} - a_{12}a_{21} = \det(A).$$

Esta función tiene un buen número de propiedades interesantes, cuyo tratamiento es el objetivo de esta sección, pero es destacable el hecho de que tres de estas propiedades bastan para caracterizar completamente al determinante, extenderlo a matrices cuadradas de cualquier tamaño, y desarrollar algoritmos de cálculo de determinantes basados en la eliminación gaussiana.

Comenzaremos por enunciar y demostrar estas propiedades fundamentales para el determinante de matrices 2×2.

Proposición 2.7. *El determinante de las matrices* 2×2 *tiene las siguientes tres propiedades:*

1. *es una función* multilineal *de las filas de la matriz: si* λ *y* a_{ij}*,* a'_{ij}*, para* $i, j = 1, 2$*, son números reales cualesquiera, entonces*

   a*) el determinante es lineal respecto a la primera fila, en el sentido de que*

$$\begin{vmatrix} a_{11} + a'_{11} & a_{12} + a'_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a'_{11} & a'_{12} \\ a_{21} & a_{22} \end{vmatrix},$$
$$\begin{vmatrix} \lambda a_{11} & \lambda a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix};$$

   b*) también es lineal respecto a la segunda fila, ya que se satisface,*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} + a'_{21} & a_{22} + a'_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ a'_{21} & a'_{22} \end{vmatrix},$$
$$\begin{vmatrix} a_{11} & a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix};$$

2. *es una función* alternada *de las filas de la matriz, en el sentido de que si se intercambian dos filas el determinante cambia de signo*[9]*:*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = - \begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}$$

3. *el determinante de la matriz identidad es igual a* 1*. Es decir*

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

[9]Es en este punto de la teoría que usamos el hecho de trabajar sobre el cuerpo de los números reales. En $\mathbb{R}$ tenemos que $-1 \neq 1$, a diferencia de lo que ocurre en un cuerpo como $\mathbb{Z}_2$ en que el opuesto de 1 es el propio 1. Para evitar algunas dificultades técnicas en la teoría de los determinantes preferimos entonces restringirnos al contexto de $\mathbb{R}$. Toda la teoría de esta sección, salvo lo que hace referencia a áreas, volúmenes y orientaciones, se extiende sin ninguna modificación al cuerpo $\mathbb{C}$ de los números complejos.

PRUEBA. Comenzamos por probar la primera propiedad en el enunciado de la proposición, es decir, que el determinante es una función multilineal. Veamos primero que depende de manera lineal de la primera fila de la matriz. Tenemos

$$\begin{vmatrix} a_{11} + a'_{11} & a_{12} + a'_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11} + a'_{11})a_{22} - (a_{12} + a'_{12})a_{21}.$$

Distribuyendo los factores a_{22} y a_{21} del miembro de la derecha y ordenando convenientemente resulta

$$\begin{vmatrix} a_{11} + a'_{11} & a_{12} + a'_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}a_{22} - a_{12}a_{21}) + \left(a'_{11}a_{22} - a'_{12}a_{21}\right),$$

donde reconocemos la igualdad

$$\begin{vmatrix} a_{11} + a'_{11} & a_{12} + a'_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a'_{11} & a'_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

La linealidad respecto a la segunda fila se puede probar de manera similar o puede reducirse a la linealidad respecto a la primera fila luego de demostrar que el determinante sólo cambia de signo al intercambiar filas. Dejamos los detalles planteados como el ejercicio 2.64.

Mostremos ahora que el determinante es una función alternada. Es decir, que cambia de signo cuando se intercambian dos filas en la matriz a la que se le calcula el determinante. Calculamos

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix} = a_{21}a_{12} - a_{11}a_{22} = -(a_{11}a_{22} - a_{21}a_{12}) = -\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Para completar la prueba resta calcular el determinante de la matriz identidad 2×2. Tenemos

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \times 1 - 0 \times 0 = 1.$$

La prueba está completa, a menos del siguiente ejercicio.

Ejercicio 2.64. Completar los detalles de la prueba de la proposición 2.7, mostrando que el determinante de las matrices 2×2 depende de manera lineal de la segunda fila de las matrices. □

Observación 2.5.1. La propiedad anterior nos dice que el determinante depende en forma lineal de cada una de las filas de la matriz, pero en general se tiene la desigualdad

$$\det(A + B) \neq \det(A) + \det(B).$$

Ejemplo 2.5.2. Tomemos las matrices

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}.$$

Entonces

$$\det(A) = \det(B) = 2.$$

Pero

$$A + B = \begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix},$$

que tiene un determinante igual a $8 \neq 2 + 2$. ♣

En resumen, **el determinante de la suma de matrices no es igual a la suma de los determinantes**.

Si multiplicamos por 2 la matriz A del ejemplo anterior encontramos

$$2A = \begin{pmatrix} 2 & 4 \\ -2 & 0 \end{pmatrix},$$

y

$$\det(2A) = 8.$$

Vemos que tampoco es cierto que $\det(2A)$ sea igual a $2\det(A)$. En general, **el determinante del producto de una matriz por un número es distinto que el producto del número por el determinante de la matriz**. Sin embargo, para el producto por números puede darse una sencilla regla de cálculo, cuya búsqueda dejamos al lector.

Ejercicio 2.65. Si A es una matriz 2×2 y λ un número. ¿Cómo se calcula $\det(\lambda A)$ en función de λ y $\det(A)$. ♠

De la definición de determinante para matrices 2×2 pueden obtenerse rápidamente algunas importantes propiedades del determinante. Proponemos al lector demostrar dos de ellas en el próximo ejercicio.

Ejercicio 2.66. Sean A y B matrices 2×2. Mostrar que

$$\det(AB) = \det(A)\det(B),$$
$$\det(A) = \det(A^t).$$

Ejercicio 2.67. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

una matriz invertible. Mostrar que

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

El lector interesado en una fórmula para las matrices inversas en el caso general de matrices $n \times n$ puede consultar la página 263.

Antes de enunciar y demostrar la proposicion 2.7 decíamos la algo enigmática frase "estas propiedades bastan para caracterizar completamente al determinante y extenderlo a matrices cuadradas de cualquier tamaño". Comencemos por aclarar en el caso de las matrices 2×2 qué significa eso de que las tres propiedades de la proposición 2.7 caracterizan al determinante: se trata del hecho de que si una función φ asocia un número $\varphi(A)$ a cada matriz A de tamaño 2×2, siguiendo una regla que satisface las propiedades de la proposición 2.7, entonces $\varphi(A)$ es necesariamente el determinante de la matriz A. Enunciamos este resultado en nuestra siguiente proposición.

Proposición 2.8. *Supongamos que a cada matriz* 2×2,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

le asociamos un número real $\varphi(A)$ *de forma tal que la correspondencia* φ *es una función multilineal y alternada de las filas de* A, *y que* $\varphi(I_{2\times 2}) = 1$. *Entonces se satisface que*

$$\varphi(A) = \det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Antes necesitaremos un pequeño lema.

Lema 2.9. *Si* φ *es una función alternada, como en el enunciado de la proposición 2.8, y la matriz* A *tiene dos filas iguales, entonces* $\varphi(A) = 0$.

Observación 2.5.3. El enunciado del lema 2.9 contempla matrices de cualquier tamaño. Notemos que si una matriz 2×2 tiene dos filas iguales entonces las únicas dos filas de la matriz son iguales entre sí. Demostraremos el lema para matrices cualesquiera. ♠

PRUEBA. Supongamos que las filas i y j de la matriz A son iguales, y para cada matriz B llamemos $\tilde{B}$ a la matriz que se obtiene a partir de B intercambiando las filas i y j. Como φ es alternada se satisface, para cualquier matriz B,

$$\varphi(\tilde{B}) = -\varphi(B). \tag{2.28}$$

En particular, esta igualdad es cierta para la matriz A. Por otra parte, como las filas i y j de A coinciden la matriz A no se ve modificada al hacer el intercambio de filas, y tenemos

$$\tilde{A} = A$$

que, obviamente, implica

$$\varphi(\tilde{A}) = \varphi(A).$$

Combinando esta última igualdad con (2.28) aplicada a A concluimos

$$\varphi(A) = -\varphi(A)$$

lo que implica[10] que $\varphi(A) = 0$. $\square$

Mostraremos ahora que las tres propiedades que hemos anotado arriba determinan completamente el la función determinante para matrices 2×2. El cálculo es sencillo y repetitivo, por lo tanto es tedioso. Para tener una notación adecuada a nuestros propósitos escribiremos una matriz 2×2 en la forma

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix},$$

donde A_1 y A_2 son las filas de A.

Si llamamos $E_1 = (1, 0)$ y $E_2 = (0, 1)$ a los vectores de la base canónica de $\mathbb{R}^2$, escritos en forma de filas, tenemos que las filas de A pueden expresarse como combinaciones lineales de estos vectores canónicos en la forma

$$\begin{aligned} A_1 &= a_{11}E_1 + a_{12}E_2, \\ A_2 &= a_{21}E_1 + a_{22}E_2. \end{aligned}$$

A partir de esta manera de escribir la matriz A vamos a trabajar con la evaluación de la función φ. Tenemos

$$\varphi\begin{pmatrix} a_{11}E_1+a_{12}E_2 \\ a_{21}E_1+a_{22}E_2 \end{pmatrix} = a_{11}\varphi\begin{pmatrix} E_1 \\ a_{21}E_1+a_{22}E_2 \end{pmatrix} + a_{12}\varphi\begin{pmatrix} E_2 \\ a_{21}E_1+a_{22}E_2 \end{pmatrix}, \quad (2.29)$$

donde hemos empleado la dependencia lineal de φ respecto a la primera fila de A. Ahora hacemos lo propio con las segundas filas de las matrices que aparecen en el miembro de la derecha de la última igualdad. Comencemos por la primera de ellas. Obtenemos

$$\varphi\begin{pmatrix} E_1 \\ a_{21}E_1 + a_{22}E_2 \end{pmatrix} = a_{21}\varphi\begin{pmatrix} E_1 \\ E_1 \end{pmatrix} + a_{22}\varphi\begin{pmatrix} E_1 \\ E_2 \end{pmatrix}$$

Como φ se anula cuando se aplica sobre una matriz que tiene dos filas iguales esta última fórmula se reduce a

$$\varphi\begin{pmatrix} E_1 \\ a_{21}E_1 + a_{22}E_2 \end{pmatrix} = a_{22}\varphi\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} \quad (2.30)$$

[10] Observemos que esta afirmación es falsa en $\mathbb{Z}_2$, donde 1 es su propio opuesto, pero cierta en $\mathbb{R}$ o $\mathbb{C}$

Trabajaremos ahora con el segundo sumando en el miembro de la derecha de (2.29). Tenemos

$$\varphi\begin{pmatrix} E_2 \\ a_{21}E_1 + a_{22}E_2 \end{pmatrix} = a_{21}\varphi\begin{pmatrix} E_2 \\ E_1 \end{pmatrix} + a_{22}\varphi\begin{pmatrix} E_2 \\ E_2 \end{pmatrix}$$

Podemos simplificar un poco esta última expresión teniendo en cuenta que

$$\varphi\begin{pmatrix} E_2 \\ E_1 \end{pmatrix} = -\varphi\begin{pmatrix} E_1 \\ E_2 \end{pmatrix}, \qquad \varphi\begin{pmatrix} E_2 \\ E_2 \end{pmatrix} = 0.$$

Obtenemos entonces

$$\varphi\begin{pmatrix} E_2 \\ a_{21}E_1 + a_{22}E_2 \end{pmatrix} = -a_{21}\varphi\begin{pmatrix} E_1 \\ E_2 \end{pmatrix}. \tag{2.31}$$

Al sustituir (2.30) y (2.31) en la fórmula (2.29) resulta

$$\varphi(A) = \varphi\begin{pmatrix} a_{11}E_1 + a_{12}E_2 \\ a_{21}E_1 + a_{22}E_2 \end{pmatrix} = (a_{11}a_{22} - a_{12}a_{21})\,\varphi\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} \tag{2.32}$$

Podemos incorporar ahora la condición de normalización $\varphi(I) = 1$, para concluir que

$$\varphi(A) = a_{11}a_{22} - a_{12}a_{21}$$

$\square$

Notemos que la fórmula (2.32) fue obtenida sólo haciendo uso de que φ es una aplicación multilineal y alternada. A partir de ella podemos concluir que cualquier aplicación φ definida sobre las matrices 2×2 que sea multilineal y alternada es de la forma

$$\varphi(A) = \det(A)\varphi(I).$$

Notemos que $\varphi(I)$ es una constante, que no depende de A. Este resultado puede enunciarse como el siguiente corolario de la proposición 2.8.

Corolario 2.10. *Si φ es una aplicación multilineal y alternada definida sobre las matrices 2×2, entonces φ es igual a la función determinante* det *multiplicada por una constante que es igual al valor que φ toma en la matriz identidad.*

Este corolario se extiende a matrices cuadradas de tamaño cualquiera, y es útil para probar algunos resultados de la teoría de determinantes. En particular, para el notable resultado acerca del determinante del producto de matrices (ver el enunciado del teorema 2.4, y la segunda demostración de este teorema en la página 256).

Que haya una única función multilineal y alternada que toma el valor 1 en la matriz identidad no es una peculiaridad de las matrices 2×2. Es cierto para matrices de cualquier tamaño. En el próximo ejercicio proponemos al lector repetir los cálculos de la demostración de la proposición 2.8 para el caso de las matrices 3×3.

Ejercicio 2.68. Supongamos que φ es una función que está definida sobre todas las matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

que φ es multilineal, alternada, y que $\varphi(I_{3\times 3}) = 1$. Entonces

$$\varphi(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Sugerencia: imitar la prueba del caso 2×2.

Los resultados de la proposición 2.8 y el ejercicio 2.68 pueden extenderse para cualquier valor de $n = 1, 2, \ldots$. Para cada n hay una única función definida sobre las matrices cuadradas $n \times n$ que es multilineal, alternada y toma el valor 1 en la matriz identidad $I_{n\times n}$. Esa función es a la que llamaremos *determinante*.

Definición 2.6 (Determinante de las matrices reales $n \times n$). *La función* determinante *es la única función*

$$\det : M^{n\times n}(\mathbb{R}) \to \mathbb{R}$$

que tiene la propiedad de depender de las filas de las matrices de manera multilineal y alternada, y que además satisface la condición $\det(I_{n\times n}) = 1$.

Como ya observamos, la proposición 2.8 asegura que la definición que acabamos de dar implica

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{2.33}$$

¡Nada nuevo bajo el sol!

Si interpretamos el ejercicio 2.68 a la luz de nuestra última definición encontramos que el determinante $\det(A)$ de una matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

se calcula en función de sus entradas por la expresión

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \quad (2.34)$$

Observación 2.5.4. El cálculo del determinante de una matriz 3×3 puede reducirse al cálculo de determinantes 2×2 realizando su *desarrollo* por alguna fila o columna. Los detalles se presentan en la subsección dedicada a estos desarrollos, en la página 261. Sólo es necesario cubrir el material que está comprendido entre el comienzo de esta subsección y el ejercicio 2.88 inclusive para acceder a toda la información que se utilizará en capítulos posteriores. ♠

Ejercicio 2.69. Sabiendo que

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 5,$$

calcular los siguientes determinantes:

$$\begin{vmatrix} d & e & f \\ g & h & i \\ a & b & c \end{vmatrix}, \quad \begin{vmatrix} -a & -b & -c \\ 2d & 2e & 2f \\ -g & -h & -i \end{vmatrix}, \quad \begin{vmatrix} a & b & c \\ d-3a & e-3b & f-3c \\ 2g & 2h & 2i \end{vmatrix}.$$

Las propiedades que caracterizan a la función determinante permiten evaluar fácilmente el determinante de una matriz diagonal. Ese es el contenido de nuestro próximo ejercicio.

Ejercicio 2.70.

1. Mostrar que
$$\begin{vmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{nn} \end{vmatrix} = d_{11}d_{22}\dots d_{nn}.$$

2. Sean A y D dos matrices $n \times n$. A partir del resultado anterior mostrar que si D es una matriz diagonal entonces
$$\det(DA) = \det(D)\det(A).$$

No hemos demostrado completamente el resultado de unicidad necesario para introducir a la función determinante por medio de la definición 2.6. Tampoco lo haremos, pero incluimos a continuación una subsección en la que damos al lector las herramientas necesarias para completar los detalles. En particular, introducimos la noción de *permutación* y toda la notación que le hará falta para poder escribir en el caso general los cálculos que hicimos al demostrar la proposición 2.8 y al resolver el ejercicio 2.68.

2.5.2. Permutaciones y determinantes de matrices cualesquiera.

El material de esta sección no es un requisito previo para secciones posteriores, por lo que puede omitirse en una primera lectura del texto.

En las fórmulas de los determinantes de las matrices A de tamaños 2×2 y 3×3 aparece una suma de productos de las entradas de A en los que hay exactamente un factor de cada fila y de cada columna. Por ejemplo, para las matrices 3×3 hay seis sumandos de la forma

$$\pm a_{1j_1} a_{2j_2} a_{3j_3}$$

donde las ternas de índices (j_1, j_2, j_3) recorren las seis posibles maneras de ordenar los números 1, 2 y 3:

$$(1,2,3), \quad (2,3,1), \quad (3,1,2), \quad (1,3,2), \quad (2,1,3), \quad (3,2,1).$$

Ejercicio 2.71. Hallar todas las posibles maneras de ordenar los números 1 y 2, y ver cómo aparecen en la expresión del determinante de las matrices 2×2.

Cada uno de estos posibles ordenamientos de los números 1 y 2, o de 1, 2 y 3, recibe el nombre de *permutación*. En general, una *permutación* σ de los n primeros números naturales $\{1, 2, \ldots, n\}$ es una función biyectiva

$$\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}.$$

Ejemplo 2.5.5. La permutación σ definida por

$$\sigma(1) = 3, \quad \sigma(2) = 2, \quad \sigma(3) = 1.$$

induce el ordenamiento $(3, 2, 1)$ de los números 1, 2 y 3. ♣

Una permutación queda completamente definida al especificar la lista de valores

$$\sigma(i) = j_i, \quad i = 1, 2, \ldots, n, \tag{2.35}$$

que toma la función. Si pensamos que la permutación σ define un reordenamiento de los primeros n números naturales, entonces j_i es el natural que ocupa el lugar i de la lista. Por esta razón, es usual adoptar la notación

$$\sigma = (j_1 j_2 \ldots j_n)$$

para representar, de una manera mucho más breve, la permutación σ definida por la fórmula (2.35).

Ejemplo 2.5.6. Con esta nueva notación, la permutación del ejemplo 2.5.5 se representa simplemente (321). El símbolo $(12\dots n)$ representa la permutación identidad, que se limita a dejar los números en su orden habitual. ♣

Llamaremos $\mathcal{S}_n$ al conjunto de todas las posibles permutaciones de lo primeros n números naturales $\{1, 2, \dots, n\}$.

Ejemplo 2.5.7. Hay 2 permutaciones en $\mathcal{S}_2$:

$$(12),\ (21).$$

En tanto que en $\mathcal{S}_3$ hay 6:

$$(123),\ (312),\ (231),\ (213),\ (321),\ (132).$$

Ejercicio 2.72. Mostrar que el conjunto $\mathcal{S}_n$ tiene exactamente $n!$ elementos. ♣

Llamaremos *trasposición* a una permutación en la que se intercambian respecto al orden usual las posiciones de exactamente dos números.

Ejemplo 2.5.8. La trasposición $\sigma = (125436)$ en $\mathcal{S}_6$ se obtiene intercambiando el 5 con el 3. Otra forma de expresar esto es decir

$$\sigma(3) = 5, \quad \sigma(5) = 3,$$

en tanto que

$$\sigma(i) = i$$

para todos los valores de i distintos de 3 y de 5. ♣

Un resultado interesante acerca de las trasposiciones es que cualquier permutación puede obtenerse como el resultado de aplicar sucesivamente una cantidad finita de trasposiciones.

Ejemplo 2.5.9. Consideremos la permutación $\sigma = (45213)$. Vamos a ver que puede obtenerse como el resultado de realizar varias trasposiciones al ordenamiento (12345). Comenzamos por intercambiar la posición 1 con la 4, para conseguir (42315). Luego intercambiamos las posiciones 2 y 5, lo que arroja el resultado (45312). Un último cambio entre los lugares 3 y 5 nos permite obtener la trasposición σ.

Ejercicio 2.73. Escribir esta descomposición de σ en trasposiciones en la forma

$$\sigma = \tau_3 \circ \tau_2 \circ \tau_1,$$

donde cada una de las permutaciones τ_i, $i = 1, 2, 3$, es una trasposición. ♣

Naturalmente, una misma permutación puede escribirse de distintas maneras como una composición de trasposiciones. Un ejemplo trivial es el siguiente: si aplicamos dos veces consecutivas la misma trasposición obtenemos la trasposición identidad.

Ejemplo 2.5.10. La permutación (45213) que escribimos en el ejemplo 2.5.9 como la composición de tres trasposiciones admite otras descomposiciones. Por ejemplo, como la cadena de cinco trasposiciones

$$(12345) \stackrel{1\leftrightarrow 2}{\longmapsto} (21345) \stackrel{2\leftrightarrow 3}{\longmapsto} (23145) \stackrel{3\leftrightarrow 4}{\longmapsto} (23415) \stackrel{2\leftrightarrow 5}{\longmapsto} (25413) \stackrel{1\leftrightarrow 3}{\longmapsto} (45213),$$

que hemos representado con una notación que el lector seguramente podrá decodificar. ♣

Hemos visto que una misma permutación σ admite varias descomposiciones en trasposiciones. Ni siquiera el número k_σ de trasposiciones necesarias está definido, pero este número tiene una propiedad notable, que está íntimamente relacionada con el signo $+$ o $-$ con el que aparece cada uno de los sumandos en el cálculo de los determinantes: la paridad del número de trasposiciones necesarias para escribir una permutación sólo depende de la permutación. Si por un procedimiento se ha necesitado un número par (impar) de trasposiciones cualquier otro procedimiento también necesitará un número par (impar) de ellas.

Por lo tanto, el número

$$\operatorname{sig}(\sigma) = (-1)^{k_\sigma}$$

está definido, y sólo depende de la permutación σ. Este número recibe el nombre de *signo* de la permutación. Si k_σ es par, entonces $\operatorname{sig}(\sigma) = 1$ y decimos que σ es *par*. Cuando k_σ es impar, tenemos que $\operatorname{sig}(\sigma) = -1$ y decimos que σ es *impar*.

Ejemplo 2.5.11. Por ejemplo, cualquier representación de la permutación (45213) como un producto de trasposiciones requerirá un número impar de ellas. El signo de esta permutación es -1. ♣

Ejemplo 2.5.12. La permutación (12) es par, en tanto que (21) es impar. Recordemos que en la fórmula

$$a_{11}a_{22} - a_{12}a_{21}$$

para el determinante de las matrices 2×2 cada sumando es de la forma $a_{1j_1}a_{2j_2}$, donde (j_1j_2) es una permutación de (12). El signo con el que cada sumando aparece corresponde al signo de la permutación de los segundos índices en ese sumando.

Un comentario similar vale para la fórmula de los determinantes de las matrices 3×3. Para confirmarlo sólo hay que examinar la fórmula (2.34), y determinar los signos de todas las permutaciones en $\mathcal{S}_3$.

Mostraremos a continuación que las permutaciones

$$(123), \quad (312), \quad (231),$$

son pares. En tanto que

$$(213), \quad (321), \quad (132),$$

son impares.

Es obvio que la permutación identidad (123) es par, porque no requiere ninguna trasposición. El lector podrá verificar rápidamente que las últimas tres permutaciones que aparecen en nuestra lista sólo requieren una trasposición. Con la misma notación que usamos en el ejemplo 2.5.10 escribamos

$$\begin{aligned} (123) &\xmapsto{1\leftrightarrow 3} (321) \xmapsto{2\leftrightarrow 3} (312), \\ (123) &\xmapsto{1\leftrightarrow 3} (321) \xmapsto{1\leftrightarrow 2} (231), \end{aligned}$$

lo que muestra que las permutaciones (312) y (231) son pares, porque admiten una descomposición como el producto de dos trasposiciones. ♣

Con la introducción de las permutaciones y sus signos tenemos todos los ingredientes necesarios para extender las fórmulas del cálculo de determinantes a matrices cuadradas de cualquier tamaño. En el próximo párrafo explicaremos la notación que vamos a emplear.

Sea

$$A = (a_{ij})_{j=1,\dots,n}^{i=1,\dots,n}$$

una matriz cuadrada. Consideremos un producto de n entradas de A tal que hay exactamente un factor en cada fila y en cada columna de A. Un producto de esta clase puede escribirse en la forma

$$a_{1j_1}a_{2j_2}\dots a_{nj_n}, \tag{2.36}$$

donde $(j_1j_2\dots j_n)$ es una permutación σ en $\mathcal{S}_n$. El primer índice en los factores a_{ij} va variando de 1 a n, lo que asegura que hay exactamente un factor en cada fila. La condición de que

$$\sigma = (j_1j_2\dots j_n)$$

sea una permutación asegura que estamos escogiendo exactamente una entrada en cada columna. Al cambiar la permutación cambiamos la forma de elegir, pero no el hecho de que siempre estamos recorriendo todas las columnas con nuestra elección.

Usando que

$$\sigma(i) = j_i, \quad i = 1, 2, \ldots, n,$$

podemos escribir el producto (2.36) en la forma

$$a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{n\sigma(n)}. \tag{2.37}$$

Podemos agregar un signo + o − a estos sumandos, multiplicándolos por el signo de la permutación σ:

$$\operatorname{sig}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{n\sigma(n)}. \tag{2.38}$$

De esta manera, para cada posible permutación σ hemos construido uno de estos productos con signo que aparecen en las fórmulas para los determinantes de matrices 2×2 y 3×3. El lector podrá verificar que para $n = 2$ y $n = 3$ el determinante se obtiene sumando todas las expresiones (2.38) que se obtienen para las posibles permutaciones en $\mathcal{S}_2$ y $\mathcal{S}_3$ respectivamente.

Ejercicio 2.74. Supongamos que φ es una función que está definida sobre todas las matrices reales cuadradas, con n filas y n columnas,

$$A = (a_{ij})_{j=1,\ldots,n}^{i=1,\ldots,n},$$

y que φ es multilineal y alternada en las filas de A. Entonces

$$\varphi(A) = \varphi(I_{n\times n}) \sum_{\sigma\in\mathcal{S}_n} \operatorname{sig}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{n\sigma(n)},$$

donde $I_{n\times n}$ es la matriz identidad $n \times n$.

Una consecuencia inmediata del ejercicio anterior es que hay una única función multilineal y alternada definida sobre las matrices $n \times n$ que toma el valor 1 sobre la matriz identidad. Esta observación extiende a matrices cuadradas cualesquiera los resultados de la proposición 2.8 y del ejercicio 2.68, y permite introducir el determinante a partir de la definición 2.6. También nos da una forma (muy poco práctica, por cierto) para calcular el determinante de una matriz:

$$\det(A) = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sig}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{n\sigma(n)}. \tag{2.39}$$

La fórmula (2.39) no es muy útil como regla de cálculo, salvo en los casos 2×2 y 3×3 en que se reduce a las fórmulas explícitas (2.33) y (2.34) que ya hemos manejado. Pero permite demostrar el importante resultado de que el determinante de una matriz y el de su traspuesta son iguales.

Ejercicio 2.75. Demostrar que la igualdad $\det(A) = \det(A^t)$ se satisface para cualquier matriz cuadrada.

Daremos una prueba diferente de este resultado en la página 257.

2.5.3. Más propiedades. Determinantes y escalerización

En esta sección mostraremos el resultado fundamental de que las operaciones de escalerización no cambian, o, a lo sumo, sólo cambian el signo, el determinante de una matriz. Es una consecuencia bastante directa de la definición de los determinantes como una función que depende de las filas de la matriz en forma multilineal y alternada. A partir de este hecho construiremos un procedimiento de cálculo que consiste en reducir la evaluación del determinante de una matriz cualquiera al cálculo del determinante de una matriz triangular. Las propiedades que vinculan el proceso de escalerización con el cálculo de determinantes nos permitirán también mostrar algunos resultados fundamentales de la teoría de los determinantes.

Haremos un uso muy importante de las propiedades que tiene la función determinante respecto a las operaciones con filas de las matrices. Utilizaremos la siguiente notación: si A_1, A_2, $\ldots A_n$, indican las n filas de una matriz A, escribiremos

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}, \qquad \det(A) = \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{vmatrix}.$$

Comenzamos con dos sencillos lemas. El primero será de utilidad para mostrar la fundamental proposición 2.13. Cualquier función alternada tiene la propiedad consagrada en el lema 2.9 de que se anula cuando se aplica sobre una matriz que tiene dos filas iguales. En particular, como el determinante es alternado concluimos que cuando una matriz tiene dos filas iguales su determinante se anula. Enunciamos este hecho en la forma de un lema sobre la función determinante.

Lema 2.11. *Si dos filas de A son iguales, entonces* $\det(A) = 0$.

También se anula el determinante en presencia de una fila de ceros.

Lema 2.12. *Si A tiene una fila nula, entonces* $\det(A) = 0$.

PRUEBA. Indiquemos con O la fila nula de A. Como $O = O + O$ podemos escribir

$$A = \begin{pmatrix} A_1 \\ \vdots \\ O \\ \vdots \\ A_n \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ O + O \\ \vdots \\ A_n \end{pmatrix}.$$

Ahora usaremos en la última expresión para A la linealidad de la función determinante respecto a la fila en la que están los ceros. Obtenemos

$$\begin{vmatrix} A_1 \\ \vdots \\ O + O \\ \vdots \\ A_n \end{vmatrix} = \begin{vmatrix} A_1 \\ \vdots \\ O \\ \vdots \\ A_n \end{vmatrix} + \begin{vmatrix} A_1 \\ \vdots \\ O \\ \vdots \\ A_n \end{vmatrix}.$$

El determinante que aparece en el miembro de la izquierda de esta igualdad es justamente el determinante de A, escrito haciendo uso del truco $O+O = O$. Los dos sumandos del miembro de la derecha también son iguales al determinante de A. Tenemos entonces

$$\det(A) = \det(A) + \det(A),$$

lo que implica inmediatamente $\det(A) = 0$. □

Pasamos ahora a mostrar una propiedad de la mayor importancia, porque será la base del método de reducir el cálculo de determinantes al proceso de escalerización de Gauss.

Proposición 2.13. *La operación elemental de sustraer de una fila el múltiplo de otra no altera el determinante*

PRUEBA. Esta proposición se prueba usando simplemente la linealidad del determinante y el lema 2.11 que dice que el determinante de una matriz que tiene dos filas iguales es nulo. En la matriz A restemos de la fila A_j el producto del número λ por la fila A_i, y calculemos el determinante de la matriz que resulta. Tenemos

$$\begin{vmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j - \lambda A_i \\ \vdots \\ A_n \end{vmatrix} = \begin{vmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{vmatrix} - \lambda \begin{vmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \\ A_n \end{vmatrix},$$

en virtud de la linealidad del determinante respecto a la j-ésima fila. El primer determinante en el miembro de la derecha es el determinante de A. El segundo determinante es nulo, ya que es el determinante de una matriz que tiene dos filas iguales. Por lo tanto, la matriz que se obtiene a partir de A aplicando la operación elemental de restar a una fila un múltiplo de otra tiene exactamente el mismo determinante que A. □

Recordemos que es posible llevar una matriz A cualquiera a una forma escalerizada aplicando una cantidad finita de veces las siguientes dos operaciones fundamentales (ver la sección 1.2.2 y en particular la página 35):

1. sumar a una fila un múltiplo de otra;
2. intercambiar filas.

La proposición anterior tiene como consecuencia directa el teorema que pasamos a enunciar.

Teorema 2.2. *Si E es una forma escalerizada de A que se obtuvo aplicando las operaciones de sumar múltiplos a una fila e intercambiar filas entonces*

$$\det(A) = \pm \det(E).$$

El signo es $+$ cuando se realiza un número par de intercambios de filas durante la escalerización, y $-$ cuando este número es impar.

Ejemplo 2.5.13. Consideremos la matriz real

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 2 & 1 & 0 & 1 \\ -1 & 0 & 2 & -1 \\ 3 & 0 & -1 & 2 \end{pmatrix}.$$

Buscaremos una forma escalerizada de A, aplicando las operaciones elementales de la eliminación gaussiana. Escribimos los pasos de la escalerización sin mayores explicaciones:

$$\begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & -4 \end{pmatrix} \begin{array}{l} \\ \longleftarrow F_2 - 2F_1, \\ \longleftarrow F_3 + F_1, \\ \longleftarrow F_4 - 3F_1. \end{array}$$

$$E = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -6 \end{pmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow F_4 - 2F_3. \end{array}$$

De momento sólo podemos afirmar que $\det(A) = \det(E)$. O continuar la escalerización hasta obtener una forma escalerizada reducida E_1 de A, que es una matriz diagonal cuyo determinante ya sabemos calcular (ver el ejercicio 2.70):

$$\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -6 \end{pmatrix} \begin{array}{l} \longleftarrow F_1 + F_3, \\ \longleftarrow F_2 - 2F_3, \\ \\ \\ \end{array}$$

$$E_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -6 \end{pmatrix} \begin{array}{l} \longleftarrow F_1 + \frac{1}{2}F_3, \\ \longleftarrow F_2 - \frac{5}{6}F_3, \\ \longleftarrow F_2 + \frac{1}{6}F_3 \\ \\ \end{array}$$

El determinante de la forma escalerizada reducida E_1 coincide con el determinante de A. Por otra parte es fácil de evaluar, porque E_1 es diagonal. Tenemos entonces

$$\det(A_1) = \det(E_1) = -6.$$

Observemos que también el determinante de E tiene el valor -6, porque el paso de E a E_1 se hace por medio de operaciones de eliminación gaussiana que no alteran el determinante de las matrices a las que se aplican.

Subrayemos el hecho de que al pasar de la forma escalerizada a la escalerizada reducida **los pivotes no sufren ninguna modificación**, y aparecen en la diagonal de la forma reducida. Esto indica que no es necesario llegar a la forma reducida, nos basta multiplicar las entradas que están en la diagonal de la forma escalerizada para conocer el determinante de la matriz. Justificaremos completamente esta observación luego, pero haremos uso de ella en nuestro próximo ejemplo. ♣

Ejemplo 2.5.14. Empleemos la técnica de escalerizacion para una nueva matriz real

$$A = \begin{pmatrix} 1 & -\frac{1}{2} & 2 & 3 \\ -1 & \frac{1}{2} & -1 & 1 \\ 2 & 3 & 2 & 4 \\ -2 & 3 & -2 & 1 \end{pmatrix}.$$

Esta matriz requiere un intercambio de filas, tal como se muestra a continuación:

$$\begin{pmatrix} 1 & -\frac{1}{2} & 2 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 4 & -2 & -2 \\ 0 & 2 & 2 & 7 \end{pmatrix} \begin{array}{l} \\ \longleftarrow F_2 + F_1 \\ \longleftarrow F_3 - 2F_1 \\ \longleftarrow F_4 + 2F_1. \end{array}$$

$$\begin{pmatrix} 1 & -\frac{1}{2} & 2 & 3 \\ 0 & 4 & -2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 2 & 2 & 7 \end{pmatrix} \begin{array}{l} \\ F_2 \leftrightarrow F_3 \\ F_2 \leftrightarrow F_3 \\ \\ \end{array}$$

$$\begin{pmatrix} 1 & -\frac{1}{2} & 2 & 3 \\ 0 & 4 & -2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 3 & 8 \end{pmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow F_4 - \frac{1}{2}F_2. \end{array}$$

$$E = \begin{pmatrix} 1 & -\frac{1}{2} & 2 & 3 \\ 0 & 4 & -2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{pmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow F_4 - 3F_3. \end{array}$$

Hemos realizado un intercambio de filas, por lo que

$$\det(A) = -\det(E).$$

El determinante de E se calcula multiplicando las entradas en la diagonal, lo que arroja el resultado -16. Por lo tanto $\det(A) = 16$. ♣

Ahora sabemos que el determinante de una matriz es esencialmente el mismo que el de su forma escalerizada, y tenemos la indicación de que el determinante de una forma escalerizada se calcula multiplicando las entradas de su diagonal.

Proposición 2.14. *El determinante de una forma escalerizada E de una matriz $n \times n$ es igual al producto*

$$e_{11}e_{22}\dots e_{nn}$$

de las entradas en la diagonal principal.

Prueba. Cuando todas las entradas de la diagonal principal son no nulas podemos continuar el proceso de escalerización usando estas entradas como pivotes, hasta conseguir una forma escalerizada reducida E' de la matriz. Este proceso no requiere intercambio de filas. En virtud del teorema 2.2 tenemos

$$\det(E) = \det(E').$$

Por otra parte, durante el pasaje de E a E' las entradas de la diagonal no cambian, por lo que E' tiene la misma diagonal que E. La matriz E' es diagonal, y su determinante es igual al producto de las entradas de su diagonal, que son justamente las de E.

Antes de continuar con la prueba del teorema es conveniente retomar el concepto de rango. El caso en el que todas las entradas en la diagonal principal de E son no nulas corresponde al caso en que el rango de E es igual a n. En efecto, cuando el rango de E es igual a n la forma escalerizada tiene n pivotes, tiene que haber un pivote en cada fila y en cada columna de la matriz, y los pivotes están sobre la diagonal principal. Por lo tanto, cuando el rango de E es igual a n hemos mostrado que su determinante coincide con el producto de las entradas de su diagonal principal. Además el determinante es distinto de cero, porque todos los factores en este producto son no nulos.

Cuando el rango de E es menor que n hay menos pivotes que filas en la forma escalerizada, y aparece una fila de ceros. Por lo tanto, el determinante de E es nulo. Como aparece una fila de ceros hay algún cero en la diagonal principal, de modo que el producto de las entradas en la diagonal también es nulo. □

Antes de emplear la proposición 2.14 para ayudarnos en el cálculo de los determinantes de formas escalerizada, destaquemos que en el curso de su demostración hemos obtenido el siguiente importante resultado.

Corolario 2.15. *Sea A una matriz $n \times n$. Entonces el rango de A es igual a n si y sólo si el determinante de A es distinto de cero.*

Ejercicio 2.76. Calcular el determinante de las siguientes matrices usando escalerización.

$$A_1 = \begin{vmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{vmatrix} \quad A_2 = \begin{vmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 \end{vmatrix} \quad A_3 = \begin{vmatrix} 1 & 1 & 1 & 0 \\ 2 & 2 & 0 & 2 \\ 3 & 0 & 3 & 3 \\ 0 & 4 & 4 & 4 \end{vmatrix}$$

Ejercicio 2.77. Calcular los siguientes determinantes:

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & 1 & 2 \\ 3 & 2 & 3 & 2 \\ -1 & -3 & 0 & 4 \\ 0 & 4 & -1 & -3 \end{vmatrix},$$

$$\begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & -2 & 0 \\ 2 & 3 & -4 & 1 \\ -1 & -2 & 0 & 2 \\ 0 & 2 & 5 & 3 \end{vmatrix}, \quad \begin{vmatrix} 2 & 4 & -5 & 2 & 1 & 5 \\ 0 & -1 & 7 & 3 & -3 & 4 \\ 0 & 0 & 2 & -3 & 9 & 4 \\ 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{vmatrix}.$$

Ejercicio 2.78. Calcular los determinantes

$$d_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 0 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{vmatrix},$$

para $n = 2, 3, \ldots$, donde se supone que la matriz con determinante d_n tiene n filas y n columnas.

Ejercicio 2.79. Modificar el enunciado del teorema 2.2 para incorporar la posibilidad de multiplicar alguna fila por una constante no nula. ¿Cómo afecta esta operación el cálculo de determinantes?

Ejercicio 2.80. Mostrar que el determinante de cualquier matriz cuadrada triangular, superior o inferior, es igual al producto de las entradas en la diagonal. Concluir que si A es una matriz triangular entonces $\det(A) = \det(A^t)$.

En el corolario 2.15 vinculábamos el rango de una matriz con su determinante. Si recordamos que una matriz cuadrada $n \times n$ es invertible si y sólo si su rango es igual a n (ver el teorema 2.1, en la página 207), obtenemos el siguiente teorema.

Teorema 2.3. *Una matriz cuadrada A es invertible si y sólo si* $\det(A) \neq 0$.

A continuación mostraremos en un ejercicio como el sencillo criterio de invertibilidad de una matriz que proporcionan los determinantes nos ayuda a determinar sus vectores propios, sobre los que la matriz actúa simplemente como una multiplicación por una constante λ, que hemos dado en llamar valor propio. Valores y vectores propios contienen importantísima información acerca de la acción de una matriz. Dada una matriz A, la búsqueda de valores y vectores propios consiste en averiguar si existen un escalar λ y un vector no nulo X tales que

$$AX = \lambda X. \tag{2.40}$$

Ejercicio 2.81. Determinantes y valores propios
El objetivo de este ejercicio es mostrar que los determinantes permiten hallar los valores propios asociados a una matriz (en caso de que existan).

1. Mostrar que la la ecuación (2.40) es equivalente a

$$(A - \lambda I)X = 0,$$

   y que λ es un valor propio de A si y sólo la matriz $A - \lambda I$ es no invertible. Con I indicamos la matriz identidad de las mismas dimensiones que A.

2. Mostrar que λ es un valor propio si y sólo si $\det(A - \lambda I) = 0$.

3. Calcular $\det(A - \lambda I)$ para las siguientes matrices A:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0{,}6 & 0{,}4 & 0{,}1 \\ 0{,}1 & 0{,}4 & 0{,}1 \\ 0{,}3 & 0{,}2 & 0{,}8 \end{pmatrix}.$$

   Hallar los valores de λ que anulan el determinante, y, para esos valores, las soluciones no triviales de la ecuación (2.40). Comparar con los resultados obtenidos en los ejercicios (1.26) parte (3) y (1.10).

4. Discutir según θ la existencia en $\mathbb{R}^2$ de vectores propios para la matriz

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

   Interpretar geométricamente el resultado.

Observación 2.5.15. Sobre el cálculo de valores propios
Al resolver el ejercicio 2.81 el lector habrá notado que aprendimos como calcular los valores y vectores propios de una matriz A hallando el determinante $\det(A - \lambda I)$, que resulta ser un polinomio en la variable λ, y luego buscando las raíces de ese polinomio. Este procedimiento es una alternativa al estudio del sistema $AX = \lambda X$, donde λ es un parámetro.

Sin embargo, no se trata de un procedimiento eficiente para la búsqueda de los autovalores de las matrices de mayor porte que suelen aparecer en los cálculos prácticos: evaluar directamente un determinante es costoso, y buscar las raíces de un polinomio cualquiera es un problema muy inestable, demasiado sensible a pequeñas perturbaciones en los valores de los coeficientes de los polinomios. Los métodos eficientes de cálculo para hallar valores propios calculan factorizaciones de las matrices que pongan en evidencia sus valores propios. En su mayoría son procedimientos iterativos, que buscan producir sucesiones de números que converjan rápidamente hacia los autovalores de la matriz.

No nos extenderemos más sobre este punto, que escapa al alcance de este curso, pero recomendamos al lector la consulta de [TB]. En particular de la parte V, dedicada al problemas de autovalores, y a la sección 25, en la página 190 y siguientes, donde se analizan algunos de los inconvenientes que tienen los métodos más obvios para el cálculo de autovalores. ♠

A continuación enunciamos uno de los más importantes resultados del cálculo de determinantes.

Teorema 2.4. *Para dos matrices A y B cuadradas de igual tamaño, el determinante del producto es el producto de los determinantes. Es decir*

$$\det(AB) = \det(A)\det(B).$$

El resultado para matrices 2×2 aparece en el ejercicio 2.66, y puede establecerse por medio de un cálculo directo. En el caso general daremos dos pruebas diferentes. La primera estará basada en el cálculo de determinantes por medio del proceso de escalerización. La segunda en la caracterización de los determinantes como una función multilineal y alternada definida sobre las filas de la matriz.

PRIMERA PRUEBA DEL TEOREMA 2.4. En primer lugar, veamos el caso singular: Si A o B son singulares, entonces $\det(A) = 0$ o $\det(B) = 0$. Por lo tanto

$$\det(A)\det(B) = 0.$$

Si una de las matrices A o B es singular, también el producto AB lo es. Por lo tanto

$$\det(AB) = 0,$$

y queda entonces probada la igualdad $\det(AB) = \det(A)\det(B)$ para este caso particular.

Si A y B son no singulares la demostración puede reducirse al del producto DB de una matriz diagonal D por B, tal como mostramos a continuación.

Partiremos de la matriz A y la reduciremos a su forma escalerizada reducida, que es una matriz diagonal D, mediante una serie de pasos de eliminación gaussiana. Los mismos pasos, aplicados al producto AB, reducen AB al producto DB (ver el ejercicio 2.22, en la página 190). Tendremos entonces, por ser D diagonal (ver el ejercicio 2.70) que

$$\det(DB) = \det(D)\det(B).$$

Por otra parte, la eliminación gaussiana no cambia el determinante, a menos de un posible cambio de signo introducido por la operación de permutar filas. De modo que tenemos

$$\det(DB) = \pm\det(AB), \quad \det(D) = \pm\det(A).$$

En ambas fórmulas va el mismo signo, porque sólo depende de la cantidad de intercambios de filas que hayamos hecho, y realizamos los mismos intercambios sobre AB que sobre A. Concluimos entonces

$$\pm\det(AB) = \pm\det(A)\det(B),$$

de donde se desprende el teorema. □

Segunda prueba del teorema 2.4. Otra demostración posible consiste en considerar la matriz B fija, e introducir la aplicación

$$A \mapsto f(A) = \det(AB).$$

Ejercicio 2.82. Mostrar que la aplicación f depende de manera multilineal y alternada de las filas de A y que $f(I) = \det(B)$. Concluir que $f(A) = \det(A)\det(B)$. Sugerencia: ver el corolario 2.10.

Una vez que hemos completado el ejercicio usamos la definición de f, que asegura que $f(A) = \det(AB)$. □

Es interesante aplicar la fórmula para el determinante de un producto en el caso del producto de una matriz invertible A y su inversa A^{-1}. Tenemos

$$AA^{-1} = (I),$$

por lo que es obvio que

$$\det\left(AA^{-1}\right) = \det(I).$$

El determinante de la identidad es igual a 1, en tanto que el miembro de la izquierda en esta igualdad puede evaluarse con la fórmula para el determinante de un producto. Obtenemos

$$\det(A)\det(A^{-1}) = 1.$$

Un resultado inmediato es el siguiente corolario.

Corolario 2.16. *Si A es invertible, entonces*

$$\det(A^{-1}) = \det(A)^{-1}.$$

Ejercicio 2.83. Demostrar a partir de este último corolario la siguiente proposición: si una matriz A es invertible su determinante es no nulo. ¿Qué consecuencia inmediata se obtiene de este resultado: que si una matriz no es invertible su determinante es cero, o que si su determinante es cero la matriz no es invertible?

El siguiente teorema asegura que el determinante de una matriz coincide con el de su traspuesta. Esto tiene la importante consecuencia de que todos los resultados sobre determinantes que puedan enunciarse en términos de las filas de las matrices, son también válidos en términos de las columnas. Veremos algunos ejemplos y aplicaciones de esto luego de demostrar el teorema.

Teorema 2.5. $\det(A^t) = \det(A)$.

Comencemos por observar que la demostración para el caso de las matrices 2×2 o 3×3 se reduce a una simple verificación a partir de las fórmulas explícitas (2.33) y (2.34). El caso 2×2 aparece en el ejercicio 2.66.

A continuación presentamos una prueba del teorema que está basada en el proceso de escalerización. Más precisamente, en la descomposición $PA = LU$ de una matriz. Necesitaremos el siguiente resultado previo, cuya demostración proponemos como un ejercicio para el lector.

Ejercicio 2.84. Sea P una matriz que se obtiene a partir de la matriz identidad I haciendo una permutación de sus filas. Demostrar que $P^tP = I$, y concluir de este resultado que $\det(P) = \det(P^t) = \pm1$. Dar un ejemplo en el que los determinantes de P y P^t valgan 1, y otro en el que valgan -1.

Prueba del teorema 2.5. Toda matriz A admite una factorización

$$PA = LU, \tag{2.41}$$

donde P es una matriz que se obtiene haciendo una permutación de filas en la matriz identidad, L es triangular inferior y U es triangular superior. Comencemos por observar que el enunciado del teorema es válido para la matriz P (ver el ejercicio 2.84), y también para las matrices L y U que son triangulares (ver el ejercicio 2.80). Sabemos entonces que

$$\det(L) = \det(L^t), \quad \det(U) = \det(U^t), \quad \det(P) = \det(P^t). \tag{2.42}$$

Reduciremos el caso general a una aplicación directa de estas igualdades, utilizando el teorema 2.4 acerca del determinante de un producto de matrices.

Al calcular el determinante de ambos miembros de (2.41) obtenemos

$$\det(P)\det(A) = \det(L)\det(U),$$

de donde podemos despejar, ya que $\det(P) = \pm 1$,

$$\det(A) = \det(L)\det(U)/\det(P). \tag{2.43}$$

Al trasponer la igualdad (2.41) resulta

$$A^tP^t = U^tL^t,$$

donde podemos aplicar otra vez la fórmula para el determinante de un producto. El resultado es

$$\det(A^t)\det(P^t) = \det(U^t)\det(L^t),$$

que nos permite despejar

$$\det(A^t) = \det(U^t)\det(L^t)/\det(P^t). \tag{2.44}$$

La prueba del teorema es ahora prácticamente inmediata. Sólo hay que comparar las expresiones (2.43) y (2.44) de los determinantes de A y A^t, teniendo en cuenta las igualdades en (2.42) □

Una consecuencia inmediata del teorema que acabamos de demostrar es la siguiente proposición.

Proposición 2.17. *Consideremos la función* det *definida sobre las matrices cuadradas $n \times n$. Entonces* det *es una función multilineal y alternada de las columnas de A.*

Ejercicio 2.85. Completar los detalles de la prueba.

Ejercicio 2.86. Sabiendo que

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 5,$$

calcular

$$\begin{vmatrix} a+5c & 3b & c \\ d+5f & 3e & f \\ 2g+10i & 6h & 2i \end{vmatrix}.$$

2.5.4. Otros resultados acerca de los determinantes

Presentamos en esta sección algunos resultados clásicos de la teoría de determinantes. No haremos uso de ellos más adelante, con excepción de los desarrollos de determinantes 3×3 que se presentan a partir de la página 261 y hasta el ejercicio 2.88 inclusive. El resto del a sección puede omitirse en una primera lectura de este texto.

Regla de Cramer

Cuando el determinante $|A|$ de la matriz de un sistema $AX = B$ es no nulo, es posible expresar su solución como un cociente de determinantes. Comenzaremos por verlo en el caso particular de un sistema con dos ecuaciones y dos incógnitas. Para ello volveremos a la página 232, a la discusión previa a la definición de determinante de matrices 2×2.

Si despejamos x_2 de la segunda ecuación en (2.25) obtenemos

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}.$$

Para calcular x_1 podemos sustituir x_2 por su valor, o simplemente multiplicar la primera ecuación del sistema (2.24) por a_{22}, la segunda por $-a_{12}$ y sumar para obtener

$$(a_{11}a_{22} - a_{21}a_{12})\, x_1 = a_{22}b_1 - a_{12}b_2.$$

La condición $|A| \neq 0$ nos permite despejar entonces

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}}.$$

Una rápida inspección de las fórmulas para x_1 y x_2 convencerá al lector de que la solución del sistema es

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

Observemos que en el numerador de la expresión de x_1 aparece una matriz que es la del sistema, con la primera columna sustituida por la columna de los términos independientes. El cálculo de x_2 sigue una regla similar: en el numerador los términos independientes sustituyen a la segunda columna de la matriz del sistema.

Esta expresión de las soluciones de un sistema lineal en términos de determinantes recibe el nombre de *regla de Cramer*, y se puede extender para sistemas compatibles y determinados $n \times n$, con n arbitario, como una aplicación de la proposición 2.17. Proponemos esta tarea al lector

Ejercicio 2.87.

1. Sea A una matriz $n \times n$, que escribiremos en términos de sus columnas como $(A_1, \dots, A_n)$. Sea
$$B = x_1 A_1 + \dots + x_n A_n$$
una combinación lineal de sus columnas. Llamemos $A^{(i)}$ a la matriz que se obtiene sustituyendo la i-ésima columna de A por B. Mostrar que
$$\det(A^{(i)}) = x_i \det(A).$$
Sugerencia: usar la expresión de B, y el hecho de que el determinante depende linealmente de la i-ésima columna de las matrices.

2. Regla de Cramer Sea A una matriz invertible y $X = (x_1, \dots, x_n)$ la solución del sistema $AX = B$. Llamemos $A^{(i)}$ a la matriz que se obtiene sustituyendo la i-ésima columna de A por B. Mostrar que
$$x_i = \frac{\det(A^{(i)})}{\det(A)}.$$

3. Calcular la solución del sistema
$$\begin{aligned} x + y + z &= 2, \\ x - y - z &= 1, \\ x + y - z &= 3, \end{aligned}$$
por medio de la regla de Cramer.

La regla de Cramer debe su nombre al matemático suizo Gabriel Cramer (1704-1752), que la presentó, aunque sin demostración, en un trabajo publicado en 1750. Casos particulares, para sistemas 2×2 y 3×3, eran conocidos con anterioridad.

Este procedimiento de cálculo de las soluciones de un sistema tiene en realidad más importancia histórica que práctica, por lo que hemos escogido presentar las fórmulas de Cramer con toda su generalidad bajo la forma del ejercicio 2.87, en el que aparecen como una aplicación de propiedades básicas de los determinantes. El lector interesado encontrará en la parte 5 del ejercicio 2.93 otra demostración de las reglas de Cramer, basada en las fórmulas para el desarrollo de determinantes. Salvo en el caso 2×2, en que puede ser útil, la regla de Cramer no es un método adecuado para calcular soluciones de sistemas de ecuaciones. Éstas se hallan más eficientemente con el método de eliminación de Gauss, o a partir de una conveniente factorización de la matriz del sistema. Cuando un sistema es compatible e indeterminado, la regla de Cramer no da ninguna información acerca del conjunto solución del sistema.

Desarrollos por filas y columnas

El determinante de una matriz $n \times n$ puede reducirse al cálculo de determinantes de matrices $(n-1) \times (n-1)$ por medio de las fórmulas de *desarrollo por filas o por columnas*. Para poder escribir estas fórmulas es necesario introducir las nociones de *matriz adjunta de un elemento* y de *menor*. Si $A = (a_{ij})$ es una matriz $n \times n$, llamaremos *matriz adjunta* del elemento a_{ij} a la submatriz A_{ij} de la matriz A que se obtiene eliminado la fila i y la columna j de A. Llamaremos *menor* del elemento a_{ij} al determinante $|A_{ij}|$ de la matriz A_{ij}. Veamos estas nociones en un ejemplo.

Ejemplo 2.5.16. Consideremos la matriz

$$A = \begin{pmatrix} 3 & -4 & -3 \\ -1 & 2 & 7 \\ -2 & 4 & 0 \end{pmatrix},$$

la matriz adjunta A_{21} del elemento $a_{21} = -1$, se obtiene eliminando la segunda fila y la primera columna de A, algo que ilustramos gráficamente:

$$\begin{pmatrix} \times & -4 & -3 \\ \times & \times & \times \\ \times & 4 & 0 \end{pmatrix}$$

y que conduce a la matriz adjunta

$$A_{21} = \begin{pmatrix} -4 & -3 \\ 4 & 0 \end{pmatrix}.$$

El menor correspondiente es $|A_{21}| = 12$. ♣

Observemos que las matrices adjuntas de una matriz cuadrada A de tamaño n, son matrices A_{ij} de tamaño $(n-1) \times (n-1)$.

En el próximo ejercicio presentamos al lector las fórmulas de los desarrollos por filas y columnas de los determinantes 3×3. En este caso las verificaciones pueden hacerse directamente, sin recurrir a ningún desarrollo teórico elaborado.

Ejercicio 2.88. Desarrollos de determinantes 3×3

1. Si A es una matriz 3×3 mostrar que se satisface la fórmula

$$|A| = a_{11} |A_{11}| - a_{12} |A_{12}| + a_{13} |A_{13}| ,$$

del desarrollo del determinante $|A|$ por la primera fila.

2. Mostrar que también se puede desarrollar $|A|$ por la primera columna, según la fórmula

$$|A| = a_{11} |A_{11}| - a_{21} |A_{21}| + a_{31} |A_{31}| .$$

3. Calcular el determinante de la matriz del ejemplo 2.5.16 usando las fórmulas de los desarrollos por la primera fila y la primera columna.

4. Hallar las fórmulas para los desarrollos por cualquier fila y cualquier columna. Verificar las fórmulas volviendo a calcular el determinante de la matriz del ejemplo 2.5.16 con ellas.

Aunque prácticamente no las emplearemos en el resto del texto, presentamos a continuación la deducción de las fórmulas generales para el desarrollo de los determinantes, bajo la forma de una serie de ejercicios. En el centro de nuestra presentación está el hecho de que el determinante depende de manera lineal de cada una de las filas y columnas de las matrices.

Ejercicio 2.89.

1. Sea A una matriz $n \times n$ tal que todas las entradas de su primera fila son nulas, salvo la entrada a_{11} que es igual a 1. Mostrar que

$$|A| = |A_{11}|.$$

2. Sea A una matriz $n \times n$ tal que todas las entradas de su i-ésima fila son nulas, salvo la entrada a_{ij} que es igual a 1. Mostrar que

$$|A| = (-1)^{i+j} |A_{ij}|.$$

3. Enunciar y demostrar un resultado análogo para columnas con todas sus entradas nulas salvo una de ellas que toma el valor 1.

Ejercicio 2.90. Sea φ una función $\varphi : \mathbb{R}^n \to \mathbb{R}$ lineal, en el sentido de que se satisface

$$\varphi(X + Y) = \varphi(X) + \varphi(Y), \quad \varphi(\lambda X) = \lambda \varphi(X)$$

para cualquier elección de X e Y en $\mathbb{R}^n$ y de λ en $\mathbb{R}$. Sea $X = (x_1, x_2, \ldots, x_n)$ un vector de $\mathbb{R}^n$ y llamemos E_i, $i = 1, 2, \ldots, n$ a los n vectores de la base canónica de $\mathbb{R}^n$. Probar que

$$\varphi(X) = x_1 \varphi(E_1) + x_2 \varphi(E_2) + \ldots + x_n \varphi(E_n).$$

Ejercicio 2.91. DESARROLLOS POR FILAS Y COLUMNAS

1. Sea A una matriz cuadrada de dimensiones $n \times n$. Demostrar que su determinante satisface, para cada $i = 1, 2, \ldots, n$, la fórmula
$$|A| = (-1)^{i+1} a_{i1} |A_{i1}| + (-1)^{i+2} a_{i2} |A_{i2}| + \ldots + (-1)^{i+n} a_{in} |A_{in}| .$$
del *desarrollo de $|A|$ por la i-ésima fila.*
2. Hallar y demostrar las fórmulas para los desarrollos por columnas.

Ejercicio 2.92. El objetivo de este ejercicio es calcular el determinante, al que llamaremos d_n, de la siguiente matriz $n \times n$:

$$\begin{pmatrix} 5 & 3 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 2 & 5 & 3 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 2 & 5 & 3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2 & 5 & 3 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 2 & 5 \end{pmatrix}.$$

1. Calcular d_1 y d_2.
2. Probar que $d_n = 5d_{n-1} - 6d_{n-2}$, para $n = 3, 4, \ldots$.
3. Hallar d_n. Sugerencia: escribir la recurrencia satisfecha por los números d_n en forma matricial, como una recurrencia sobre un vector $D_n = (d_n, d_{n-1})^t$, y buscar los valores y vectores propios de la matriz que define la recurrencia.

Cofactores y fórmula de la inversa

El lector ya habrá notado que en las fórmulas para los desarrollos del determinante de una matriz A aparecen los menores $|A_{ij}|$ de sus elementos a_{ij} multiplicados por coeficientes de la forma $(-1)^{i+j}$. Las expresiones se simplifican al introducir el *cofactor*

$$c_{ij} = (-1)^{i+j} |A_{ij}| , \quad i, j = 1, 2, \ldots, n,$$

de cada elemento a_{ij}.

Ejercicio 2.93. FÓRMULA DE COFACTORES PARA LA INVERSA
Sea A una matriz $n \times n$.

1. Escribir las fórmulas de los desarrollos de $|A|$ en términos de sus entradas y sus cofactores.

2. Llamamos *matriz de cofactores* de A a la matriz

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix},$$

formada por sus cofactores. Mostrar que

$$AC^t = |A| I_{n\times n}.$$

3. Mostrar que si A es invertible entonces

$$A^{-1} = \frac{C^t}{|A|}.$$

4. Escoger cualquier matriz real invertible $n\times n$ que aparezca en el texto y calcular su inversa con la fórmula de los cofactores. Hacerlo para una matriz 2×2, y para otra matriz con $n > 2$.

5. Utilizar la expresión de la inversa en términos de la matriz de cofactores para dar una nueva demostración de la regla de Cramer.

2.5.5. Determinantes, áreas, volúmenes y orientación

El objetivo de esta subsección es mostrar la vinculación que existe entre el cálculo de determinantes y el cálculo de áreas y volúmenes, y la posibilidad de dotar al plano y el espacio de una orientación. Son resultados de importancia para la Geometría y el Análisis Matemático.

Comenzamos por trabajar con matrices 2×2 y el área de paralelogramos naturalmente asociados con las matrices. Las filas de una matriz 2×2 pueden interpretarse como vectores en el plano. Ésta es una interpretación geométrica que ya hemos introducido en la sección 1.2.6, a partir de la página 70, y manejado en la sección 1.3.2 al discutir el significado de las combinaciones lineales. Si trabajamos con la matriz

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

la primera fila queda identificada con el vector de coordenadas (a_{11}, a_{12}), y la segunda con el vector (a_{21}, a_{22}) tal como se muestra en la figura 2.2.

Mostraremos que el determinante de la matriz es igual al área del paralelogramo que tiene como lados a los vectores (a_{11}, a_{12}) y (a_{21}, a_{22}). Para ello analizaremos desde este punto de vista más geométrico el procedimiento de

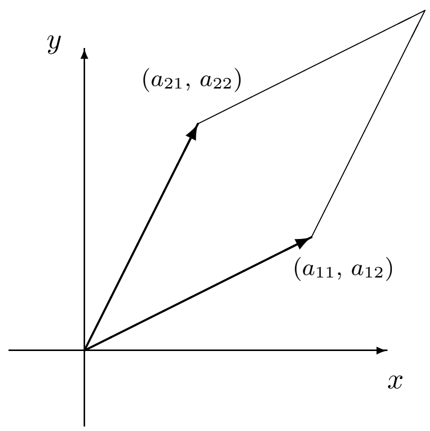


Figura 2.2: Representación geométrica de las filas de una matriz.

eliminación de Gauss. Supongamos que $a_{11} \neq 0$. Dejamos el caso $a_{11} = 0$ como ejercicio para el lector.

¿En qué se traduce en el dibujo de la figura 2.2 la operación de sumar a la segunda fila un múltiplo de la primera para hacer que aparezca un cero en el lugar que antes ocupaba a_{21}? En que desplazamos el extremo del vector (a_{21}, a_{22}) sobre una línea paralela a (a_{11}, a_{12}) hasta que el nuevo vector queda sobre el eje vertical (que es equivalente a la condición de que su primera entrada se anule). El resultado se muestra en la figura 2.3.

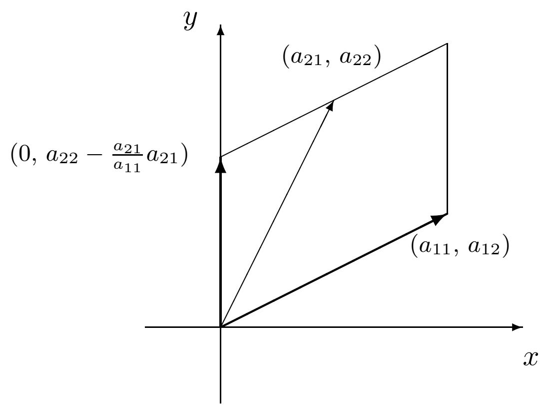


Figura 2.3: Nuevo paralelogramo de base vertical.

El nuevo paralelogramo, de lados

$$(a_{11}, a_{12}), \quad (0, a_{22} - (a_{21}/a_{11})a_{12})$$

tiene la misma área que el paralelogramo original. Ambos comparten la misma base (a_{11}, a_{12}) y la altura referida a esta base no ha cambiado con la operación de desplazar el extremo del segundo vector sobre una línea paralela a (a_{11}, a_{12}). Sin embargo, ahora es fácil calcular el área del paralelogramo en la figura 2.3. Para ello sólo hay que tomar como base el vector vertical y observar que la altura referida a esta base es a_{11}. La vieja fórmula

$$\text{área} = \text{base} \times \text{altura}$$

nos dice que el área es igual a

$$(a_{22} - (a_{21}/a_{11})a_{12}) \times a_{11} = a_{11}a_{22} - a_{21}a_{12} = \det(A).$$

Cómo esta área es igual a la del paralelogramo original encontramos que el área del paralelogramo de lados (a_{11}, a_{12}) y (a_{21}, a_{22}) es igual al determinante de la matriz que tiene estas dos filas. Veamos estos cálculos en acción con algunos ejemplos.

Ejemplo 2.5.17. Consideremos los vectores $(2, 1)$ y $(1, 1)$ en el plano, tal como se muestran en la figura 2.4. Con ellos armamos la matriz

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

El determinante de esta matriz es $\det(A) = 1$, y 1 es el área del paralelogramo de lados $(2, 1)$ y $(1, 1)$.

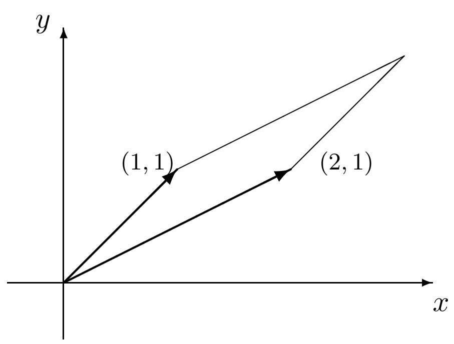


Figura 2.4: Representación geométrica de los vectores $(2, 1)$ y $(1, 1)$.

En la figura 2.5 mostramos los paralelogramos de lados $(1, 0)$ y $(0, 1)$, y de lados $(1, 0)$ y $(c, 1)$. En ambos casos el área es igual a 1.

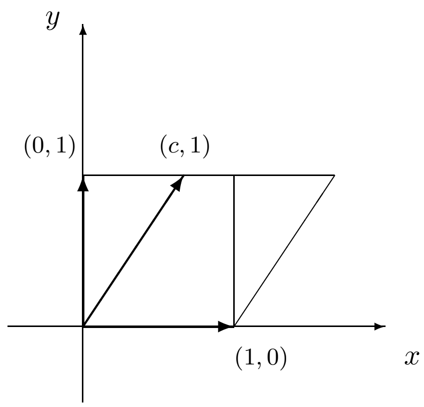


Figura 2.5: Dos paralelogramos en el plano.

El cálculo de los determinantes de las matrices asociadas a estos paralelogramos confirma este dato:

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 1 & 0 \\ c & 1 \end{vmatrix} = 1.$$

Para cerrar este ejemplo hagamos la siguiente observación: cuando consideramos la matriz asociada con los paralelogramos de la figura 2.5 nada nos obliga tomar $(1, 0)$ como primera fila y $(0, 1)$ o $(c, 1)$ como segunda fila. Hagamos el cálculo cambiando el orden de las filas:

$$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1, \quad \begin{vmatrix} c & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

¡Al cambiar las filas el valor del determinante cambió y ahora el "área" arroja un valor negativo! ¿Qué sentido tiene esto? ♣

Vamos a considerar ahora el problema que pone en evidencia la última parte del ejemplo anterior. El determinante mide algo más que el área del paralelogramo cuyos lados son las filas de la matriz, en realidad **el valor del determinante corresponde a un área orientada**, que puede tener signo positivo o negativo. El área en el sentido usual, que es siempre mayor o igual que cero, es en realidad igual al **valor absoluto del determinante**. Discutamos ahora esta cuestión de la orientación y cómo se vincula con el cálculo de áreas y detrminantes.

Orientación

Miremos el plano (x, y). Hay dos posibles sentidos de giro en el plano: el de las agujas del reloj y el contrario a las agujas del reloj. Por ejemplo, para girar el vector $(1, 0)$ hasta el vector $(0, 1)$ hay que realizar un giro de un ángulo $\pi/2$ en el sentido contrario al de las agujas del reloj. Adoptaremos la convención de que este sentido de giro es el *positivo*, por lo que diremos que la pareja $((1, 0), (0, 1))$ está orientada positivamente. Si cambiamos el orden en la pareja entonces encontramos que el giro desde el primer vector al segundo debe hacerse en el sentido de las agujas del reloj, que es el sentido negativo de giro. El determinante detecta este cambio y asigna un valor positivo al área del paralelogramo de lados $((1, 0), (0, 1))$, y negativo para el de lados $((0, 1), (1, 0))$.

En general, si tenemos dos vectores

$$A_1 = (a_{11}, a_{12}), \quad A_2 = (a_{21}, a_{22})$$

no colineales podemos llevar el primero de ellos sobre la dirección del segundo por medio de un giro cuyo ángulo toma valores entre $(0, \pi)$ (por ejemplo, en el caso de los vectores $(1, 0)$ y $(0, 1)$ este ángulo es de $\pi/2$). Si el giro debe hacerse en el sentido positivo, contrario al del movimiento de las agujas del reloj, diremos que la pareja (A_1, A_2) está orientada positivamente. En este caso será

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0.$$

Cuando el giro debe hacerse en el sentido negativo diremos que (A_1, A_2) está orientada negativamente, y el determinante tomará un valor negativo. Cuando A_1 y A_2 son colineales el determinante se anula. Geométricamente esto corresponde al caso en que el paralelogramo degenera en un segmento de recta.

En resumen, vemos que el signo del determinante distingue las dos posibles orientaciones de las parejas de vectores del plano. El valor del determinante corresponde al área orientada del paralelogramo generado por la pareja de vectores que forman las dos filas de la matriz. El valor absoluto del determinante indica el valor del área, prescindiendo de la orientación.

El caso 3×3, volúmenes y orientación

Para las matrices cuadradas, reales, de dimensión 3×3 el determinante representa el volumen de la caja o paralelepípedo cuyas aristas son las filas de la matriz. Comencemos por el caso más simple, el de la matriz identidad

$$I_{3\times 3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

El determinante de esta matriz es igual a 1, lo que corresponde al hecho de que el paralelepípedo de aristas

$$(1,0,0),\ (0,1,0),\ (0,0,1),$$

tiene volumen 1. La relación entre volúmenes y determinantes en el caso general puede establecerse con un argumento basado en la eliminación gaussiana, similar al que usamos para el caso 2×2. Omitiremos los detalles.

También en el caso 3×3 hay una noción de orientación. Así consideraremos que la terna de vectores

$$((1,0,0),\ (0,1,0),\ (0,0,1))$$

está orientada positivamente, en correspondencia con el hecho de que el determinante de la matriz identidad es igual a $1 > 0$. Pero si intercambiamos dos vectores, por ejemplo el primero y el segundo obtenemos

$$((0,1,0),\ (1,0,0),\ (0,0,1)) \tag{2.45}$$

que corresponde a la matriz

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Esta matriz tiene determinante $-1 < 0$, de modo que diremos que la terna (2.45) está negativamente orientada.

Hay entonces dos posibles orientaciones de las ternas de vectores en el espacio, y el determinante discrimina entre ambas. El valor del determinante es, análogamente a lo que ocurre en el caso 2×2 para las áreas, un **volumen orientado**. Estas dos orientaciones corresponden a la orientación de una terna y la de su imagen en un espejo, y sólo tenemos que mirarnos ambas manos para ver un ejemplo bien claro de dos orientaciones diferentes. La existencia de dos posibles orientaciones es relevante para la mecánica, el electromagnetismo y la química orgánica[11]. En nuestro curso esta cuestión reaparecerá cuando estudiemos problemas de geometría en $\mathbb{R}^3$.

Por último digamos que los determinantes permiten extender esta noción de volumen orientado a una dimensión n cualquiera. Si tenemos n vectores en $\mathbb{R}^n$ entonces el volumen del paralelepípedo n-dimensional que estos n vectores generan es igual al valor absoluto del determinante de la matriz que tiene como filas a los n vectores. El signo del determinante corresponde a la orientación de la familia de n vectores. Nuevamente, hay dos orientaciones posibles. No nos extenderemos más en estos comentarios. Pero cerraremos esta sección acerca de volúmenes y orientaciones con un comentario acerca de que todo lo que hemos hecho con las filas de una matriz puede hacerse con sus columnas.

Observación 2.5.18. Como el determinante de una matriz coincide con el de su traspuesta podemos traducir toda la discusión precedente acerca del área del paralelogramo o el volumen del paralelepípedo generado por las filas de una matriz, a una discusión sobre las columnas. En efecto, en el caso 2×2 el determinante es igual al área orientada del paralelogramos cuyos lados son las dos columnas de la matriz. Para verlo sólo basta razonar sobre las filas de la matriz traspuesta. La misma observación vale para el caso 3×3 y el volumen del paralelepípedo generado por las columnas. ♣

2.5.6. Ejercicios adicionales

Ejercicio 2.94. Sea A una matriz $n \times n$. Calcular en función de n y el determinante de A los siguientes determinantes: $\det(2A)$, $\det(-A)$, $\det(A^2)$ y $\det(A^t A)$.

Ejercicio 2.95. Sea A una matriz invertible. Mostrar que

$$\det(A^k) = \det(A)^k, \quad k \in \mathbb{Z}.$$

[11]Para ilustrar la cuestión de las orientaciones en la química recurrimos al siguiente pasaje de *Alicia detrás del espejo*, de Lewis Carrol y a la nota con la que Martin Gardner comenta el fragmento.

¿Te gustaría vivir en la Casa del Espejo, Kitty? No sé si te darán leche allí. Tal vez le leche del Espejo no sea buena de beber.

Gardner comenta *la especulación de Alicia sobre la leche del espejo tiene más importancia de la que Carroll sospechaba. Varios años después de la publicación de* A través del espejo *la estereoquímica descubrió pruebas positivas de que las sustancias orgánicas tienen una disposición asimétrica de los átomos. ...... Los estereósimeros ... forman parejas en espejo. ...... Debido a que la asimilación de alimentos comporta complicadas reacciones químicas entre los alimentos asimétricos y las sustancia asimétricas del cuerpo, a menudo se dan acusadas diferencias en el gusto, el olor y la digestibilidad entre las formas derecha e izquierda de una misma sustancia orgánica.*

Estos textos están tomados de *Alicia anotada*, Lewis Carroll, edición de Martin Gardner, Ediciones Akal, 1999.

Ejercicio 2.96. Consideremos la matriz

$$A = \begin{pmatrix} 4 & 4 & 8 & 8 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Hallar los determinantes de A, A^t, A^{-1} y de la matriz

$$B = \begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 1 & 2 & 2 \\ 4 & 4 & 8 & 8 \end{pmatrix}$$

que se obtiene intercambiando las filas de A.

Ejercicio 2.97. Sea A una matriz $n \times n$ tal que todas sus entradas a_{ij} con $j < i$ son nulas. Hallar el determinante de A en función de n y de sus entradas.

Ejercicio 2.98. Sin desarrollar, probar que

$$\begin{vmatrix} 1 & 7 & 9 & 8 \\ 2 & 1 & 3 & 9 \\ 3 & 2 & 5 & 5 \\ 4 & 8 & 6 & 7 \end{vmatrix}$$

es múltiplo de 31. Sugerencia: observar que

$$\begin{aligned} 1798 &= 1 \times 1000 + 7 \times 100 + 9 \times 10 + 8 &= 31 \times 58, \\ 2139 &= 2 \times 1000 + 1 \times 100 + 3 \times 10 + 9 &= 31 \times 69, \\ 3255 &= 3 \times 1000 + 2 \times 100 + 5 \times 10 + 5 &= 31 \times 105, \\ 4867 &= 4 \times 1000 + 8 \times 100 + 6 \times 10 + 7 &= 31 \times 157. \end{aligned}$$

Ejercicio 2.99. MATRICES ESPECIALES

1. Decimos que una matriz A es *antisimétrica* si $A = -A^t$. Probar que si n es impar y A es una matriz antisimétrica entonces $\det(A) = 0$.

2. Una matriz cuadrada A se llama *nilpotente* si existe algún número natural k tal que $A^k = O$. Mostrar que el determinante de cualquier matriz nilpotente es nulo. Hallar una matriz 2×2 no nula, tal que $A^2 = O$.

3. La matriz A se llama *idempotente* si $A^2 = A$. ¿Qué valores puede tomar el determinante de una matriz idempotente? Construir un ejemplo para cada uno de los valores posibles.

4. ¿Qué valores puede tomar el determinante de una matriz $n \times n$ que satisfaga la igualdad $A^2 = I$? Con I indicamos la matriz identidad $n \times n$. Dar un ejemplo para cada uno de los valores posibles.

5. Se llama *matriz ortogonal*[12] a una matriz A que satisface la igualdad $A^tA = I$. Hallar los posibles valores que puede tomar el determinante de una matriz ortogonal. Construir un ejemplo de matriz ortogonal para cada uno de los posibles valores hallados.

Ejercicio 2.100. Sean A una matriz $n \times n$, B una matriz $n \times m$ y C una matriz $m \times m$. Con O indicaremos una matriz $m \times n$ cualquier dimensión cuyas entradas son todas nulas. Probar que

$$\det \begin{pmatrix} A & B \\ O & C \end{pmatrix} = \det(A)\det(C)$$

Ejercicio 2.101. Sean A, B, C, y D cuatro matrices $n \times n$ que son submatrices de la matriz

$$E = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

de tamaño $2n \times 2n$. Investigar si es cierto o falso que

$$\det(E) = \det(A)\det(D) - \det(C)\det(B).$$

Ejercicio 2.102. Matriz de Vandermonde
Consideremos la matriz

$$V = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}$$

1. Sea $C = (c_0, c_1, \dots, c_{n-1})$ un vector columna en $\mathbb{R}^n$. Mostrar que el producto VC es un vector que contiene las evaluaciones del polinomio

   $$p(x) = c_0 + c_1x + \dots + c_{n-1}x^{n-1}$$

   en los puntos x_i, $i = 1, 2, \dots, n$.
2. Calcular $\det(V)$, y estudiar cuál es la condición para que ese determinante se anule.
3. Investigar la relación entre las dos partes anteriores de este ejercicio.

Ejercicio 2.103. Sean U y V dos vectores de $\mathbb{R}^n$, $n \geq 2$, a los que consideraremos como matrices columna $n \times 1$.

1. Hallar el determinante de la matriz UV^t;
2. Mostrar que

   $$\det(I + UV^t) = 1 + U^tV,$$

   donde I indica a la matriz identidad $n \times n$.

[12]El nombre esta justificado porque cada una de las columnas de la matriz es ortogonal a las demás. La noción de ortogonalidad será introducida en el capítulo dedicado al estudio de la geometría

2.5.7. Para tener presente

Todas las matrices a las que haremos referencia serán matrices cuadradas.

- Si A es 2×2 entonces su determinante $\det(A)$ es
$$a_{11}a_{22} - a_{12}a_{21}.$$
- Si A es 3×3 entonces entonces su determinante $\det(A)$ es
$$a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$
- El determinante es una función definida sobre las matrices $n \times n$. Si todas las filas (o columnas) de una matriz menos una se mantienen constantes, y se considera el determinante como una función de la fila (columna) que se permite variar, entonces el determinante es una función lineal de esa fila (columna).
- Las operación básica de la eliminación gaussiana de sumar a una fila un múltiplo de otra no modifica el determinante de una matriz. Si se suma a una columna un múltiplo de otra el determinante tampoco cambia.
- Si se intercambian dos filas (columnas) de una matriz el determinante cambia de signo.
- El determinante de una matriz puede calcularse escalerizándola.
- Las siguientes tres afirmaciones sobre la matriz A de tamaño $n \times n$ son equivalentes:
  - el determinante de A es distinto de cero;
  - A es invertible;
  - el rango de A es igual a n.
- $\det(AB) = \det(A)\det(B)$;
- El determinante de una matriz es igual al de su traspuesta;
- El valor absoluto del determinante de una matriz 2×2 puede interpretarse como un área; el de una matriz 3×3 como un volumen;
- El signo del determinante indica la orientación de una pareja o de una terna de vectores en el plano o el espacio respectivamente.

- El plano y el espacio tienen dos orientaciones posibles.
- El determinante de una matriz 3×3 puede desarrollarse por la primera fila como

$$|A| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Capítulo 3

Geometría

El propósito de este capítulo es desarrollar la geometría del espacio $\mathbb{R}^3$, introduciendo las nociones de *recta*, *plano*, *distancia*, *ángulo*, *curva* y *superficie* en $\mathbb{R}^3$.

El espacio vectorial $\mathbb{R}^3$ está formado por ternas (x_1, x_2, x_3) o (x, y, z) de números reales. Nos interesa ahora discutir cuál es la interpretación geométrica de estas ternas en el contexto de la teoría que vamos a desarrollar. Volveremos sobre las ideas que presentamos en la página 70 de la sección 1.2, cuando discutimos la interpretación geométrica de los sistemas de ecuaciones, y las trataremos con algo más de profundidad en el contexto de $\mathbb{R}^3$, también de $\mathbb{R}^2$.

Para interpretar geométricamente a $\mathbb{R}^3$ debemos pensar que en el espacio tenemos un sistema de coordenadas con tres ejes ortogonales. Una terna (x, y, z) puede interpretarse como el *vector* que tiene componenentes x, y y z en cada uno de estos ejes. En esta interpretación cada elemento de $\mathbb{R}^3$ representa un vector, que tiene una cierta longitud, dirección y sentido.

La suma $V + W$ de dos elementos V y W de $\mathbb{R}^3$ corresponde a la suma de vectores siguiendo la regla del paralelogramo. Y el producto λV por un escalar λ representa el producto de un vector por un número. Esta operación da lugar a un nuevo vector con la misma dirección que el vector V, una longitud igual a la del vector V multiplicada por el módulo $|\lambda|$ del escalar λ, y el sentido de V o el opuesto, según que λ sea positivo o negativo. Ambas operaciones se ilustran, para el caso de $\mathbb{R}^2$, en la figura 1.4 de la página 71. La interpretación para $\mathbb{R}^3$ es la misma, pero todo transcurre "en el espacio", en vez de hacerlo "en el plano".

Pero también podemos interpretar que una terna (x, y, z) define un *punto* en el espacio. Esto es posible porque al referir la geometría a un sistema de coordenadas estamos introduciendo un punto privilegiado en el espacio: el origen O del sistema de coordenadas, al que le corresponde la terna $(0, 0, 0)$. Una vez fijado este punto privilegiado la correspondencia entre puntos y vectores es muy directa: si nos desplazamos desde O siguiendo el vector (x, y, z) alcanzamos un punto P que tiene coordenadas (x, y, z) en el sistema de coordenadas que estamos empleando.

Ya hemos hecho referencia a la posibilidad de adoptar este nuevo punto de vista en la sección 1.3, al discutir las posibles representaciones geométricas de un sistema de ecuaciones lineales. Las dos representaciones de parejas de números como vectores o puntos del plano se ilustran en la figura 1.7, página 84. La discusión se extiende haciendo solamente algunas modificaciones obvias al caso de ternas de números, o elementos de $\mathbb{R}^3$.

En la teoría que presentaremos utilizaremos ambas interpretaciones, y nos referiremos a los elementos de $\mathbb{R}^3$ como puntos y vectores, según el contexto en el que aparezcan. Desde el punto de vista formal la distinción es innecesaria, porque en todos los casos se tratará de elementos del espacio $\mathbb{R}^3$. Desde el punto de vista de las ideas que se pondrán en juego la distinción es importante, porque ayuda a entender lo que se está haciendo.

En la sección 3.1 introduciremos las nociones de *recta* y *plano* como subconjuntos de $\mathbb{R}^3$. Mostraremos dos maneras de describir estos conjuntos:

- por medio de una representación *paramétrica*, que nos permite recorrer todos los puntos de una recta o de un plano, dejando variar uno o dos parámetros respectivamente;

- como el conjunto solución de sistemas de ecuaciones lineales convenientemente elegidos.

Esta sección extiende y profundiza en el caso particular de $\mathbb{R}^3$ la discusión geométrica contenida en la subsección 1.3.2, página 81, de la sección 1.3.

En la sección 3.2 se define el importante *producto escalar* en $\mathbb{R}^3$. De este producto escalar se derivan las nociones de *ángulo* y *longitud*. Como podremos medir ángulos tendrá sentido hablar de *perpendicularidad* u *ortogonalidad*. En particular, encontraremos que los vectores

$$((1,0,0),(0,1,0),(0,0,1))$$

de la base canónica de $\mathbb{R}^3$ son ortogonales dos a dos, y tienen todos longitud igual a 1, lo que está en correspondencia con nuestra imagen heurística de la teoría.

En la sección 3.3 introducimos el *producto vectorial*, de especial interés para varias aplicaciones físicas. Luego de estas dos secciones avanzamos en las aplicaciones de los productos escalar y vectorial a problemas geométricos, este es el objetivo de la sección 3.4.

Por último, cerramos el capítulo con algunas nociones sobre *curvas* y *superficies* en el espacio, que extienden lo antes hecho para rectas y planos: las curvas y superficies en el espacio pueden describirse por medio de una parametrización que las recorre al dejar variar uno o dos parámetros, respectivamente; o pueden caracterizarse por ecuaciones que deben ser satisfechas por los puntos que pertenecen a ellas. Mostraremos este hecho ejemplificando con esferas, cilindros y conos, además de las rectas y planos. Pero no discutiremos las condiciones de regularidad que aseguran que parametrizaciones y ecuaciones efectivamente representan curvas o superficies en el espacio. Simplemente nos limitaremos a exhibir a través de algunos ejemplos la necesidad de algunas hipótesis adicionales. Los detalles son materia para un curso de geometría diferencial, fuera de los objetivos de este capítulo.

3.1. Rectas y planos en el espacio

El objetivo principal de esta sección es introducir los objetos del espacio $\mathbb{R}^3$ a los que llamaremos *rectas* y *planos*. En nuestro enfoque las rectas y planos serán definidos como algunos subconjuntos particulares de $\mathbb{R}^3$, que quedarán caracterizados por ecuaciones lineales. Comenzaremos por una presentación más bien operativa, dirigida a calcular los elementos que caracterizan a rectas y planos.

Las rectas aparecerán representadas por sistemas lineales de ecuaciones que corresponden al a intersección de dos planos. En la subsección 3.1.3 estudiaremos el rango de las matrices de estos sistemas y luego discutiremos más detalladamente, en la subsección 3.1.4 cómo calcular intersecciones entre rectas y planos.

Las subsecciones 3.1.6 y 3.1.5 escapan al contexto de $\mathbb{R}^3$, porque están dedicadas a mostrar como algunas ideas geométricas pueden usarse en espacios con más dimensiones, o sobre cuerpos discretos.

Antes de seguir adelante digamos que utilizaremos frecuentemente los siguientes términos, corrientes en el tratamiento de la geometría: diremos que una recta r *pasa* por un punto P si $P \in r$. También que un plano π *pasa* por P si $P \in \pi$. Si dos rectas, dos planos, o una recta y un plano, tienen intersección no vacía diremos que *se cortan*. También diremos que dos vectores V y W de $\mathbb{R}^3$ son *colineales* cuando la pareja (V, W) es linealmente dependiente y, por supuesto, *no colineales* cuando la pareja es linealmente independiente. Observemos que si V y W son no colineales entonces ninguno de ellos puede ser el vector nulo. En particular, el vector nulo es colineal con cualquier otro vector del espacio. Cunado V y W son colineales y ambos son distintos del vector nulo, entonces existen números reales λ y $\mu = 1/\lambda$ tales que

$$V = \lambda W, \quad W = \mu V.$$

Ejercicio 3.1. ¿Qué ocurre si alguno de los vectores V o W es nulo?

Ejercicio 3.2. Para cada una de las parejas (V, W) de vectores que aparecen a continuación decidir si son o no son colineales, y hallar λ y μ cuando sea posible.

1. $V = (-\sqrt{2}\pi, \sqrt{2}\pi, -5\pi/\sqrt{2})$, $W = (2e, -2e, 5e)$;
2. $V = (1, 2, 3)$, $W = (3, 2, 1)$;
3. $V = (0, 0, 0)$, $W = (-1, -5, 1)$.

Vectores colineales no nulos definen la misma dirección en el espacio. Para todo lo que haremos en esta sección basta con este enunciado completamente intuitivo, pero la noción de dirección puede hacerse más rigurosa. En este sentido apunta el siguiente ejercicio.

Ejercicio 3.3. La relación $V \equiv W$ si y sólo si V y W son colineales, ¿es una relación de equivalencia en $\mathbb{R}^3$? ¿Y en el conjunto $\mathbb{R}^3 \setminus \{O\}$, que se consigue quitando de $\mathbb{R}^3$ el vector nulo $O = (0,0,0)$? En caso de que se trate de una relación de equivalencia ¿cuál es el significado de las clases de equivalencia?

3.1.1. Rectas

Definiremos las rectas a partir de la idea de que para especificarlas basta dar dos informaciones:

- un punto por donde pasa la recta;
- la dirección de la recta.

Para lo primero daremos un elemento

$$P = (p_1, p_2, p_3) \in \mathbb{R}^3$$

por el que la recta pasa. Para fijar una dirección nos bastará dar un vector no nulo paralelo a esa dirección. Es decir, otro elemento

$$V = (v_1, v_2, v_3) \in \mathbb{R}^3$$

al que ahora le tocará trabajar de vector en esta película[1]. El vector V debe satisfacer la condición de que alguna de sus componentes v_1, v_2 y v_3 sea distinta de cero (el vector nulo $O = (0,0,0)$ no define dirección ninguna).

Cualquier otro punto Q que esté sobre la recta definida por el punto P y el vector V estará caracterizado por la propiedad de que la diferencia $Q - P$ es colineal con V, por lo que existirá algún número real λ tal que

$$Q - P = \lambda V,$$

lo que implica que Q admite una expresión de la forma

$$Q = P + \lambda V.$$

Esta discusión es la que resumimos en nuestra próxima definición.

Definición 3.1 (Recta). *Dados un punto P y un vector no nulo V en $\mathbb{R}^3$ la* recta r que pasa por P y tiene dirección paralela a V *es el conjunto*

$$r = \{P + \lambda V;\ \lambda \in \mathbb{R}\} \subset \mathbb{R}^3.$$

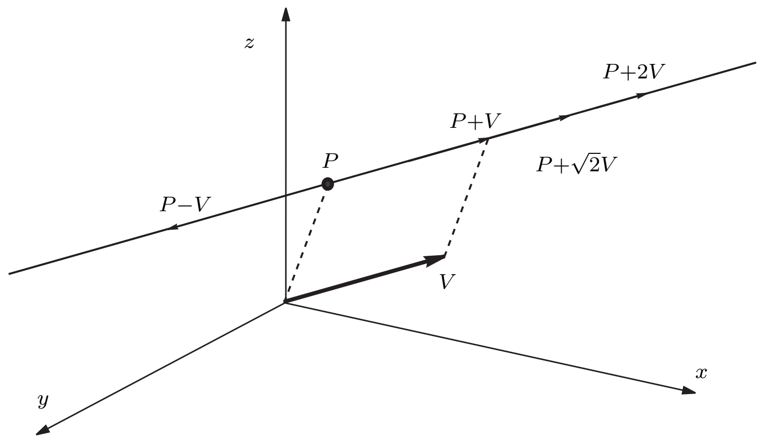


Figura 3.1: Recta que pasa por el punto P y tiene la dirección del vector V.

Para V y r como en la definición que acabamos de dar diremos que V es un *vector director* de la recta r.

La descripción de una recta que pasa por un punto $P = (p_1, p_2, p_3)$ y tiene dirección $V = (v_1, v_2, v_3)$ como el conjunto de puntos de la forma $P + \lambda V$ puede escribirse en coordenadas en la forma de lo que llamaremos *ecuaciones paramétricas* de una recta r. Si un punto $Q = (x, y, z)$ pertenece a r entonces debe satisfacerse

$$Q = P + \lambda V \tag{3.1}$$

para algún valor de λ.

Ecuaciones paramétricas

Escribiendo los vectores coordenada a coordenada, la condición 3.1 se traduce en que x, y, y z deben satisfacer, para algún valor de λ las ecuaciones

$$\left\{\begin{array}{l} x = p_1 + \lambda v_1, \\ y = p_2 + \lambda v_2, \\ z = p_3 + \lambda v_3. \end{array}\right. \tag{3.2}$$

[1]Los distintos papeles justifican los distintos trajes que usan nuestros actores. Perdón, nos referimos a la notación. Al punto lo hemos designado con la P de "punto", y al vector con la V de "vector". En general mantendremos este uso, y usaremos P, Q, R, para elementos de $\mathbb{R}^3$ que deben ser pensados como puntos en el espacio, y U, V, W, para los que hacen de vectores. En otros contextos, por ejemplo al trabajar con sistemas lineales, es más corriente emplear X, Y, Z, para designar a los elementos de $\mathbb{R}^3$. Pero en realidad es todo lo mismo.

Ejemplo 3.1.1. Tomemos $P = (-1, 0, 2)$, $V = (2, -2, 1)$. Entonces la recta r que pasa por P y tiene dirección paralela a V tiene ecuaciones paramétricas

$$\begin{cases} x = -1 + 2\lambda, \\ y = -2\lambda, \\ z = 2 + \lambda, \end{cases} \tag{3.3}$$

Notemos que cuando $\lambda = 0$, obtenemos $(x, y, z) = (-1, 0, 2)$, que es el propio P. Tomando $\lambda = -1$ resulta $(x, y, z) = (-3, 2, 1)$, otro punto en la recta.

Un punto $Q = (x, y, z)$ pertenece a r si las ecuaciones (3.3) se verifican para algún valor de λ. Por ejemplo, intentemos determinar si el punto $(10, 1, 0)$ pertenece a r. Debemos analizar si las tres igualdades

$$\begin{cases} 10 = -1 + 2\lambda, \\ 1 = -2\lambda, \\ 0 = 2 + \lambda. \end{cases}$$

se satisfacen para algún valor de λ. Se trata de un sencillo sistema lineal de tres ecuaciones e incógnita λ, que es, a todas luces, incompatible. La primera ecuación solo puede satisfacerse si $\lambda = 11/2$, en tanto que la segunda requiere $\lambda = -1/2$.

Ejercicio 3.4. Mostrar que el punto $(19, -20, 12)$ está en r. ♣

Observación 3.1.2. En la definición 3.1 la recta r queda caracterizada por uno de sus puntos y un vector que define la dirección de r. Es posible cambiar esta representación y escoger cualquier otro punto de la recta y cualquier otro vector que tenga la misma dirección para representarla.

Ejemplo 3.1.3. El punto $P' = (1, -2, 3)$ pertenece a la recta r del ejemplo 3.1.1, ya que se obtiene haciendo $\lambda = 1$ en (3.3). El vector $V' = (1, -1, 1/2)$ es colineal con V. A partir de P' y V' podemos obtener nuevas ecuaciones paramétricas

$$\begin{cases} x = 1 + \lambda, \\ y = -2 - \lambda, \\ z = 3 + \lambda/2 \end{cases} \tag{3.4}$$

para r. Nuestra intuición geométrica nos dice que las ecuaciones (3.3) y (3.4) describen el mismo conjunto. Se puede confirmar directamente que esto es así, algo que se propone al lector en el próximo ejercicio.

Ejercicio 3.5. Mostrar que cualquier punto $(x, y, z) \in \mathbb{R}^3$ que pueda obtenerse escogiendo adecuadamente un valor de λ en (3.3), también puede obtenerse a partir de (3.4), y viceversa.

Daremos luego un argumento diferente, basado en la teoría de los sistemas de ecuaciones lineales, para mostrar que hemos obtenidos dos parametrizaciones diferentes del mismo conjunto. ♣♠

Ejemplo 3.1.4. RECTAS Y SUBESPACIOS DE DIMENSIÓN 1
Una recta que pasa por el origen $O = (0, 0, 0)$ es un subespacio de dimensión 1 en $\mathbb{R}^3$. Se trata de todos los puntos de la forma

$$Q = O + \lambda V = \lambda V, \quad \lambda \in \mathbb{R}, \tag{3.5}$$

que es justamente el subespacio generado por el vector V que define la dirección de la recta. Como V es un vector no nulo el conjunto que sólo contiene a V es una base de ese subespacio. Claro está que la dimensión de este subespacio es igual a 1.

Recíprocamente, un subespacio de dimensión 1 es una recta que pasa por el origen. Basta tomar una base del subespacio, que estará formada por un único vector V, para concluir que el subespacio contiene a todos los elementos de $\mathbb{R}^3$ que son de la forma (3.5). ♣

Observación 3.1.5. LA PARAMETRIZACIÓN DE UN MOVIMIENTO RECTILÍNEO, Y SU TRAYECTORIA
Comenzamos por recordar el hecho de que una recta r es un subconjunto de $\mathbb{R}^3$. Cada uno de los puntos de r es de la forma $P + \lambda V$ para algún valor real de λ, y al variar λ vamos recorriendo todos los puntos de la recta. Esta manera de describir la recta es lo que llamaremos una *parametrización* de la recta. El parámetro es λ, y la parametrización establece una correspondencia uno a uno entre los números reales y la recta. Podemos enfatizar aún más este punto de vista escribiendo cada punto Q de la recta en la forma

$$Q(\lambda) = P + \lambda V, \quad \lambda \in \mathbb{R}.$$

Como $V \neq 0$ esta expresión define una función

$$\begin{array}{rcccl} Q & : & \mathbb{R} & \rightarrow & \mathbb{R}^3, \\ & & \lambda & \mapsto & P + \lambda V, \end{array}$$

inyectiva, cuya imagen es la recta r.

Quizás la analogía más clara para esta descripción de la recta es la del movimiento con una velocidad uniforme V no nula. Si pensamos que el parámetro λ representa el tiempo t, entonces la fórmula $P + \lambda V$ nos dice cuál es la posición en el instante $t = \lambda$ de un punto móvil que se desplaza con una velocidad constante V y que en el instante $t = 0$ ocupaba (u ocupará, no tenemos por qué pensar que $\lambda > 0$ ni que el tiempo cero está en el pasado) la posición P del espacio. La recta está entonces formada por todos los puntos de $\mathbb{R}^3$ por los que el "punto móvil" pasa cuando λ varía entre $-\infty$ y $+\infty$.

Tal como vimos, una misma recta puede admitir varias parametrizaciones, de la misma forma que una trayectoria dada puede ser recorrida de infinidad de maneras diferentes. En nuestra analogía cinemática, cambiar el punto P que usamos en la parametrización a un nuevo P' equivale a modificar el punto de partida; cambiar el vector V a V' implicar recorrer la recta con una velocidad diferente. Incluso el sentido del recorrido puede cambiar. Esto ocurre si la constante de proporcionalidad entre los vectores V y V' es negativa.

Como cada parametrización origina una terna de ecuaciones paramétricas al ser escrita coordenada a coordenada, dada una recta hay una infinidad de posibles ecuaciones paramétricas.

Ecuaciones reducidas

El problema de saber si un punto $Q = (x, y, z)$ está en la recta r que queda definido por las ecuaciones paramétricas (3.2) puede resolverse para cada punto, buscando si existe o no existe un valor de λ que las satisfaga. Eso fue lo que hicimos en el ejemplo 3.1.1 para la recta r de ecuaciones paramétricas (3.3) y los puntos $(10, 1, 0)$ y $(19, -20, 12)$.

Pero en vez de hacer los cálculos para cada punto particular podemos buscar una condición explícita sobre (x, y, z) que nos permita saber si un punto genérico (x, y, z) está o no está en la recta.

Ejemplo 3.1.6. Determinemos qué condición debe satisfacer $Q = (x, y, z)$ para pertenecer a la recta r de ecuaciones paramétricas (3.3). Notemos que una vez fijados (x, y, z) el problema se reduce a estudiar la compatibilidad del sistema de ecuaciones (3.3) que tiene tres ecuaciones e incógnita λ.

Usamos eliminación gaussiana para escalerizar el sistema, recurriendo a la primera ecuación para eliminar λ de la segunda y la tercera. Luego de sumar a la segunda ecuación la primera, y de multiplicar a la tercera por dos y restarle la primera, obtenemos

$$\begin{cases} x = -1 + 2\lambda, \\ x + y = -1, \\ 2z - x = 5. \end{cases}$$

En las ecuaciones segunda y tercera ya no aparece λ. Son ecuaciones que deben satisfacerse para asegurar la compatibilidad del sistema. Concluimos entonces que la recta r está formada por los puntos (x, y, z) que son solución del sistema de ecuaciones lineales

$$\left\{\begin{array}{l} x+y=-1, \\ -x+2z=5. \end{array}\right. \tag{3.6}$$

Ejercicio 3.6. Usar estas ecuaciones para estudiar la pertenencia a la recta r de los puntos $(10,1,0)$ y $(19,-20-12)$.

En otras palabras la recta r es el conjunto

$$r=\{(x,y,z)\in\mathbb{R}^3;\ x+y+1=0,\ -x+2z-5=0\}.$$

Como resumen de este ejemplo, podemos decir que las ecuaciones (3.3) nos dicen cómo recorrer r. En tanto que (3.6) caracterizan al conjunto r formado por los puntos visitados en nuestro recorrido. ♣

Llamaremos a estas nuevas ecuaciones para r *ecuaciones reducidas* o *implícitas* de la recta. Esto es completamente general y cualquier recta en $\mathbb{R}^3$ puede describirse como el conjunto de puntos (x,y,z) de $\mathbb{R}^3$ que satisfacen un sistema formado por dos ecuaciones lineales independientes. Veremos este hecho en nuestros próximos ejemplos y ejercicios, en los que mostraremos como pasar de ecuaciones paramétricas a reducidas y viceversa.

Ejemplo 3.1.7. Mostraremos que el par de ecuaciones

$$\left\{\begin{array}{l} x-y=3, \\ x+z=-1, \end{array}\right. \tag{3.7}$$

son las ecuaciones reducidas de una recta r.

Comencemos por observar que cualquiera de las coordenadas, x, y o z, puede escogerse como variable libre en el sistema lineal (3.7), y emplearse para expresar todas las soluciones del sistema dejándola variar. Por ejemplo, si decidimos tomar x como variable libre tenemos

$$\left\{\begin{array}{l} y=-3+x, \\ z=-1-x. \end{array}\right. \tag{3.8}$$

Naturalmente, $x=x$, por lo que la expresión paramétrica de las soluciones del sistema es[2].

$$\left\{\begin{array}{l} x=x, \\ y=-3+x, \\ z=-1-x. \end{array}\right. \tag{3.9}$$

[2]Aunque es conveniente usar una notación sistemática, nada obliga a que los parámetros se llamen λ o μ. Si insistimos en llamar λ al parámetro x entonces $x=\lambda$, y las ecuaciones quedan en la forma

$$\left\{\begin{array}{l} x=\lambda, \\ y=-3+\lambda, \\ z=-1-\lambda. \end{array}\right.$$

Reconocemos aquí a la recta que pasa por $P = (0, -3, -1)$ y tiene como vector director a $V = (1, 1, -1)$. ♣

El pasaje de ecuaciones paramétricas a reducidas y viceversa se reduce a dos operaciones ya conocidas para nosotros, con las que hemos trabajado en el capítulo dedicado a los sistemas lineales:

- las ecuaciones paramétricas de una recta pueden interpretarse como un sistema de tres ecuaciones y una incógnita. El sistema es compatible sólo para los puntos que están sobre la recta. **Las ecuaciones reducidas son la condición de compatibilidad de este sistema**. Pueden obtenerse aplicando eliminación gaussiana, al igual que para cualquier sistema lineal. Para un sistema tan sencillo como estos puede no ser necesario recurrir a la escalerización, y la simple eliminación del parámetro o incógnita es suficiente para hallar las ecuaciones reducidas;
- las ecuaciones reducidas son satisfechas por los puntos que están en la recta. Forman un sistema de ecuaciones y la recta es su conjunto solución. **Al expresar las soluciones del sistema en términos de una variable libre encontramos una parametrización de la recta**. La variable libre actúa como parámetro, y al dejarla variar en $\mathbb{R}$ se recorre la recta, o conjunto solución.

Ejemplo 3.1.8. Los comentarios anteriores permiten concluir fácilmente que las dos parametrizaciones (3.3) y (3.4) representan la misma recta. Ambas parametrizaciones son dos maneras diferentes de representar el conjunto solución del sistema (3.6). Los vectores $V = (2, -2, 1)$ y $V' = (1, -1, 1/2)$ están en el núcleo de la matriz

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

del sistema, y son soluciones de la ecuación homogénea

$$\begin{cases} x + y = 0, \\ -x + 2z = 0. \end{cases}$$

Los puntos $P = (-1, 0, 2)$ y $P' = (1, -2, 3)$ son soluciones particulares de (3.6). Como el núcleo de la matriz tiene dimensión 1 cualquier vector del núcleo es de la forma λV, o $\lambda V'$, con $\lambda \in \mathbb{R}$. Reencontramos la expresión $P + \lambda V$, o $P' + \lambda V'$, para los puntos de la recta r al aplicar la proposición 1.18 de la página 132: cualquier solución de (3.6) puede representarse como la suma de una solución particular más una solución cualquiera de la ecuación homogénea asociada. ♣

Hemos visto entonces la expresión de una recta por sus ecuaciones reducidas, que no es otra cosa que volver sobre el hecho de que una recta es el conjunto solución de un sistema de dos ecuaciones lineales independientes con tres incógnitas. Hemos discutido abundantemente en el capítulo 1 dedicado a los sistemas lineales acerca de cómo modificar un sistema de ecuaciones sin cambiar su conjunto solución, por lo que ahora no nos sorprenderá saber que las ecuaciones reducidas de una recta tampoco son únicas. De todos modos, veámoslo en un ejemplo.

Ejemplo 3.1.9. El par de ecuaciones

$$\begin{cases} x + z = 4, \\ 3x + y = 5, \end{cases}$$

define una cierta recta en el espacio, que tiene, por ejemplo, ecuaciones paramétricas

$$\begin{cases} x = x \\ z = 4 - x, \\ y = 5 - 3x. \end{cases}$$

Si escalerizamos el sistema original, eliminando x en la tercera ecuación, obtenemos

$$\begin{cases} x + z = 4, \\ y - 3z = -7. \end{cases}$$

Este par de ecuaciones tienen exactamente el mismo conjunto solución, que es justamente la recta r. De esta manera hemos obtenido un nuevo par de ecuaciones reducidas para r. ♣

En la observación 3.1.16 volveremos sobre esta cuestión de no unicidad de las ecuaciones reducidas, pero adoptaremos allí un punto de vista más geométrico, que consiste en interpretar que cada una de las ecuaciones representa un plano. El sistema formado por las ecuaciones reducidas corresponde a considerar la recta como la intersección de dos planos.

Observación 3.1.10. Es interesante señalar que no hemos introducido las rectas a través de un conjunto de axiomas que éstas deben satisfacer. En este sentido nos apartamos de las presentaciones de la geometría que siguen el estilo que consagró Euclides con su célebre obra *Elementos* y que introducen rectas y puntos como objetos no definidos que satisfacen un conjunto de axiomas, entre ellos axiomas de incidencia del tipo "dados un par de puntos P y Q distintos entre sí, existe una única recta que contiene a P y a Q".

Estamos construyendo una geometría del espacio. Con el mismo tipo de ideas y técnicas se puede hacer geometría en el plano, y es posible recuperar toda la geometría euclidiana del plano definiendo las rectas y construyendo todas las nociones métricas a partir de la noción de producto escalar que introduciremos en la sección 3.2. Esto es así porque las rectas, definidas como en la definición 3.1 satisfacen todos los axiomas de la geometría de Euclides. No nos detendremos a verificar esta afirmación, pero el lector interesado puede trabajar sobre el siguiente ejercicio.

Ejercicio 3.7.

1. Sean P y Q dos puntos distintos de $\mathbb{R}^3$. Mostrar que existe una única recta que pasa por P y Q.
2. Adaptar a $\mathbb{R}^2$ la definición de recta que hemos dado para $\mathbb{R}^3$, y enunciar y demostrar un resultado análogo al de la parte anterior.

En muchas ramas de la matemática es posible introducir los mismos conceptos de diferentes maneras que conducen a construir teorías equivalentes. Tal vez el lector haya visto este hecho en el caso del conjunto de los números reales. Es posible construir los números reales a partir de los racionales, éstos a partir de los enteros y los enteros a partir de los naturales. O declarar que los números reales son un cuerpo ordenado y completo. El resultado de ambas presentaciones es el mismo conjunto numérico. ♠

Ejercicio 3.8. Hallar ecuaciones paramétricas y ecuaciones implícitas (o reducidas) de las siguientes rectas:

1. la que pasa por el punto $P = (1, 2, 5)$, con vector director $V = (2, 1, 3)$;
2. la que pasa por los puntos $A = (4, 3, 0)$ y $B = (1, 0, 1)$.

En el enunciado del próximo ejercicio usamos la expresión *puntos alineados*. Por supuesto, diremos que tres puntos de $\mathbb{R}^3$ están alineados si existe una recta que los contiene a los tres.

Ejercicio 3.9.

1. Averiguar si los puntos $(3, 1, -1)$, $(5, 2, 1)$ y $(5, 0, 0)$ pertenecen a la recta con ecuaciones paramétricas
$$\begin{cases} x = 1 + 2\lambda, \\ y = 2 - \lambda, \\ z = -2 + \lambda. \end{cases}$$
2. Repetir para los puntos $(-1, 0, 0)$, $(0, 1, 1)$ y $(1, -1, 1)$, y la recta que tiene ecuaciones reducidas
$$\begin{cases} x + y - z + 1 = 0, \\ 2x - y + z + 2 = 0. \end{cases}$$

3. Averiguar si los puntos $(1,0,2)$, $(-1,1,1)$ y $(3,-1,1)$ están alineados. Si lo están, encontrar ecuaciones paramétricas y reducidas de la recta que determinan.
4. Repetir para $(1,1,1)$, $(1,0,-1)$ y $(1,2,3)$.

3.1.2. Planos

Introduciremos ahora los *planos* en $\mathbb{R}^3$. Las ideas que manejaremos son muy similares a las de la sección dedicadas a las rectas, pero la diferencia es que los planos serán objetos bidimensionales, que nos permiten movernos en dos direcciones independientes. Por lo tanto, un plano π quedará especificado por un punto $P=(p_1,p_2,p_3)$ que pertenezca a π, y dos vectores

$$U=(u_1,u_2,u_3), \quad V=(v_1,v_2,v_3)$$

no colineales. Cualquier otro punto Q en el plano π estará caracterizado por la

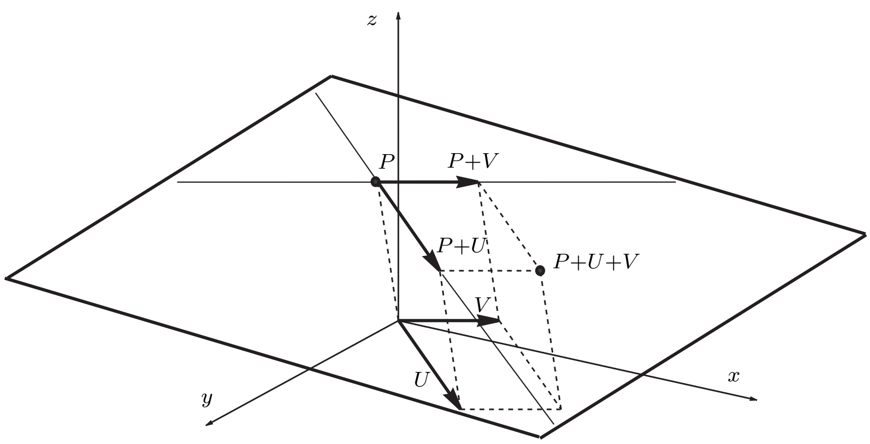


Figura 3.2: Plano determinado por un punto y dos vectores.

propiedad de que la diferencia $Q-P$ puede escribirse como una combinación lineal de los vectores U y V que determinan las direcciones paralelas al plano. Nuestra próxima definición precisa estas nociones. Notemos que es muy similar a la definición 3.1.

Definición 3.2 (Plano). *Dados un punto P y dos vectores U y V no colineales de $\mathbb{R}^3$ el plano π que contiene a P y tiene dirección paralela a los* *vectores U y V es el conjunto*

$$\pi = \{P + \lambda U + \mu V;\ \lambda, \mu \in \mathbb{R}\}$$

Diremos que el par (U, V) es un par de *vectores directores* del plano π.

Esta definición también da lugar a ecuaciones paramétricas para el plano π, de la forma

$$\left\{\begin{array}{l} x = p_1 + \lambda u_1 + \mu v_1, \\ y = p_2 + \lambda u_2 + \mu v_2, \\ z = p_3 + \lambda u_3 + \mu v_3, \end{array}\right. \tag{3.10}$$

donde

$$P = (p_1, p_2, p_3), \quad U = (u_1, u_2, u_3), \quad V = (v_1, v_2, v_3).$$

Al dejar variar en $\mathbb{R}^2$ los parámetros λ y μ se obtienen todos los posibles valores de las coordenadas (x, y, z) de los puntos que están en el plano π.

Ejemplo 3.1.11. El plano π que pasa por el punto $P = (0, -1, 1)$ y tiene a $U = (1, -2, -1)$ y $V = (-1, 3, 1)$ como vectores directores tiene ecuaciones paramétricas

$$\left\{\begin{array}{l} x = \lambda - \mu, \\ y = -1 - 2\lambda + 3\mu, \\ z = 1 - \lambda + \mu. \end{array}\right. \tag{3.11}$$

Consideremos, por ejemplo, el punto $Q = (1, -2, 0)$ y tratemos de determinar si pertenece a π. Esto es equivalente a que Q pueda escribirse en la forma (3.11) para algún valor de λ y μ, y nos lleva a considerar el sistema de ecuaciones

$$\left\{\begin{array}{rcl} 1 &=& \lambda - \mu, \\ -2 &=& -1 - 2\lambda + 3\mu, \\ 0 &=& 1 - \lambda + \mu. \end{array}\right.$$

con incógnitas λ y μ. Su solución es $\lambda = 2$, $\mu = 1$, lo que implica que $Q \in \pi$. Más aún

$$Q = P + 2U + V.$$

Cuando nos planteamos la misma pregunta para $R = (3, -1, -1)$ encontramos que el sistema

$$\left\{\begin{array}{rcl} 3 &=& \lambda - \mu, \\ -1 &=& -1 - 2\lambda + 3\mu, \\ -1 &=& 1 - \lambda + \mu. \end{array}\right.$$

es incompatible. Por lo tanto $R \notin \pi$.

Ejercicio 3.10. Verificar los resultados acerca de los dos sistemas de ecuaciones lineales en λ y μ que aparecen en este ejemplo. ♣

También para los planos la condición de pertenencia de un punto Q de coordenadas (x, y, z) puede expresarse en términos de un conjunto de ecuaciones sobre las coordenadas de Q, que aseguran la compatibilidad del sistema (3.10). En el caso de un plano todo se reduce a una única ecuación lineal, tal como mostramos en nuestro próximo ejemplo

Ejemplo 3.1.12. Un punto $Q = (x, y, z)$ pertenece al plano π del ejemplo 3.1.11 si y sólo si el sistema (3.11) es compatible. Analicemos la condición de compatibilidad recurriendo a la eliminación gaussiana. Para ello representamos el sistema en la forma matricial

$$\left(\begin{array}{rr|c} 1 & -1 & x \\ -2 & 3 & y+1 \\ -1 & 1 & z-1 \end{array}\right).$$

Al sumar dos veces la primera fila de la matriz a la segunda, y la primera a la tercera, el sistema queda escalerizado, en la forma

$$\left(\begin{array}{rr|c} 1 & -1 & x \\ 0 & 1 & 2x+y+1 \\ 0 & 0 & x+z-1 \end{array}\right),$$

y encontramos que

$$x + z = 1 \tag{3.12}$$

es la condición de compatibilidad. Por lo tanto, el plano π es el conjunto

$$\pi = \left\{(x, y, z) \ \in \ \mathbb{R}^3; \ x + z = 1\right\}.$$

La ecuación (3.12) es lo que llamamos ecuación implícita o ecuación reducida del plano π de ecuaciones paramétricas (3.11). ♣

En general, una única ecuación lineal

$$ax + by + cz + d = 0 \tag{3.13}$$

en $\mathbb{R}^3$, que sea no trivial en el sentido de que los tres coeficientes a, b y c no se anulen simultáneamente, define un plano.

Observación 3.1.13. Tal como ocurre con las rectas, las ecuaciones paramétricas

$$Q = P + \lambda U + \mu V$$

de un plano son una expresión de las soluciones de la ecuación reducida (3.13). Vistas de esta manera, el punto P es una solución particular de la ecuación (3.13), y los vectores U y V pertenecen al núcleo de la matriz[3] $(a\ b\ c)$ de la ecuación.

Para el plano π de los ejemplos 3.1.11 y 3.1.12, el punto $(0,-1,1)$ satisface la ecuación (3.12) del plano. La matriz de esta ecuación es

$$A = (1\ 0\ 1).$$

Los vectores

$$U = (1,-2,-1), \quad V = (-1,3,1),$$

que aparecen en las expresiones paramétricas (3.11) están en el núcleo de A, ya que

$$\begin{aligned} 1\times 1 + 0\times(-2) + 1\times(-1) &= 0,\\ 1\times(-1) + 0\times 3 + 1\times 1 &= 0. \end{aligned}$$

Notemos que estos cálculos sencillos permiten verificar que se ha derivado correctamente la ecuación reducida a partir de las parámetricas. ♠

Ejemplo 3.1.14. Consideremos la ecuación

$$2x + 3y - z + 4 = 0.$$

Despejando, por ejemplo, la variable z, obtenemos

$$z = 2x + 3y + 4,$$

lo que es equivalente al juego de ecuaciones paramétricas

$$\begin{cases} x = \lambda,\\ y = \mu,\\ z = 4 + 2\lambda + 3\mu. \end{cases}$$

Esta es la representación paramétrica de un plano que pasa por el punto $(0,0,4)$, que se obtiene haciendo $\lambda = \mu = 0$ en las ecuaciones que acabamos de encontrar, y tiene a los vectores

$$U = (1,0,2), \quad V = (0,1,3),$$

como vectores directores. ♣

[3]Se trata de un sistema con una ecuación y tres incógnitas. La matriz del sistema es 1×3, y se reduce a una única fila con los coeficientes a, b y c con que aparecen las variables, o incógnitas, x, y y z.

Ejemplo 3.1.15. Una ecuación como

$$x = 0$$

define un plano que pasa por el origen, y tiene $(0, 1, 0)$ y $(0, 0, 1)$ como un par de vectores directores. También $y = 0$ y $z = 0$ representan planos en $\mathbb{R}^3$. ♣

Observación 3.1.16. Ahora que sabemos que una ecuación lineal define un plano, tenemos a nuestra disposición una nueva interpretación de las ecuaciones reducidas de las rectas. Cada una de las dos ecuaciones en una pareja de ecuaciones reducidas representa un plano. Los puntos que satisfacen las dos ecuaciones son los que están en la intersección de los planos que ellas definen. Por lo tanto, especificar una recta por medio de sus ecuaciones reducidas es equivalente a representarla como la intersección de dos planos. ♠

Observación 3.1.17. Un plano también queda determinado por tres puntos P, Q y R que no estén alineados. Esta segunda manera de determinar un plano se reduce en realidad a la de un punto y dos vectores no colineales, porque podemos basarnos, por ejemplo, en el punto P y los dos vectores

$$U = Q - P, \quad V = R - P,$$

que no son colineales si P, Q y R no están alineados. ♠

Ejercicio 3.11. Hallar ecuaciones paramétricas y reducidas de los siguientes planos:

1. el que pasa por el punto $(1, 1, 1)$ y tiene a $(2, -1, 1)$ y $(1, 0, -1)$ como vectores directores;
2. el que pasa por los puntos $(1, 1, 1)$, $(2, 2, 3)$ y $(1, 1, -2)$;
3. el que pasa por el punto $(1, 1, 1)$ y contiene a la recta[4]

$$\begin{cases} x + y + z + 2 = 0, \\ x - y - z - 2 = 0. \end{cases}$$

Ejercicio 3.12. Para las dos cuaternas de puntos que aparecen a continuación, averiguar si existe algún plano que las contenga:

1. $(0, 1, 0)$, $(1, 1, 2)$, $(2, 0, 3)$, $(1, -1, 0)$;
2. $(0, -2, -1)$, $(1, 4, 0)$, $(2, 10, 1)$, $(0, 0, 0)$.

En caso afirmativo, hallar una ecuación reducida de ese plano.

[4] En realidad deberíamos decir "la recta cuyas ecuaciones son ...". A esta altura el lector perdonará la metonimia.

Para un plano cualquiera se puede obtener una ecuación reducida expresando la condición de compatibilidad del sistema de ecuaciones paramétricas (3.10) por medio de un determinante. Este procedimiento da una fórmula explícita en términos de los coeficientes de la parametrización.

Ejercicio 3.13.

1. Mostrar que si $U = (u_1, u_2, u_3)$ y $V = (v_1, v_2, v_3)$ son dos vectores no colineales, el sistema de ecuaciones paramétricas (3.10) es compatible si y sólo si el determinante
$$\begin{vmatrix} x - p_1 & y - p_2 & z - p_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$
es nulo.

2. Concluir que el plano de ecuaciones paramétricas (3.10) tiene una ecuación reducida
$$\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}(x - p_1) - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}(y - p_2) + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}(z - p_3) = 0.$$

3. Usar este resultado para hallar una ecuación reducida del plano del ejemplo 3.1.11

Ejercicio 3.14.

1. Hallar ecuaciones reducidas para el plano π de ecuaciones paramétricas
$$\left\{ \begin{array}{l} x = -1 + \lambda - \mu, \\ y = 2 + \lambda + \mu, \\ z = -1 - \lambda - 2\mu. \end{array} \right. \tag{3.14}$$
Hacerlo de dos maneras:
   a) por medio de un determinante, como se discutió en el ejercicio 3.13;
   b) hallando la condición de compatibilidad del sistema formado por las ecuaciones paramétricas.

2. Hallar nuevas ecuaciones paramétricas a partir de las ecuaciones reducidas encontradas en la parte anterior. A partir de cada conjunto de ecuaciones paramétricas identificar un punto en el plano y un par (U_i, V_i), $i = 1, 2$, de vectores directores.

3. Identificar vectores directores (U_3, V_3) a partir de las ecuaciones paramétricas 3.14.

4. Mostrar que cada uno de los vectores U_i y V_i, para $i = 1, 2, 3$, hallado en las partes anteriores puede escribirse de manera única como combinación lineal de cada una de las parejas (U_j, V_j), $j = 1, 2, 3$. Interpretar el resultado.

Nos despedimos de esta sección sobre los planos de $\mathbb{R}^3$ con un ejercicio que vincula las nociones de plano y subespacio.

Ejercicio 3.15. Planos y subespacios de dimensión 2
Mostrar que un subconjunto de $\mathbb{R}^3$ es un plano que pasa por el origen si y sólo si es un subespacio de $\mathbb{R}^3$ que tiene dimensión 2.

3.1.3. Rectas, planos y rango

Las ecuaciones reducidas de las rectas forman un sistema con dos ecuaciones lineales. De acuerdo con lo que hemos aprendido acerca de los planos en el espacio cada una de estas ecuaciones representa un plano. Los puntos que satisfacen ambas ecuaciones a la vez tienen que pertenecer entonces a los dos planos, por lo que las ecuaciones reducidas de las rectas adquieren un nuevo significado: expresan a la recta como la intersección de dos planos. Naturalmente, dada una recta, hay infinitos pares de planos que la tienen como intersección, lo que se refleja en que hay infinitas maneras de representar una recta por medio de ecuaciones reducidas.

Para discutir la representación de una recta por un par de ecuaciones reducidas recurriremos a los conceptos que desarrollamos en el capítulo 1 para estudiar los sistemas de ecuaciones lineales.

Una recta vendrá representada por dos ecuaciones, de la forma

$$\left\{\begin{array}{l} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2. \end{array}\right. \tag{3.15}$$

La matriz del sistema es

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}.$$

Cuando esta matriz tiene rango 2 entonces el sistema es compatible. Tendrá entonces al menos una solución. Llamemos P a esta solución del sistema. Notemos que P es una solución particular de la ecuación no homogénea (3.15).

Además el núcleo de la matriz A tiene dimensión 1. Por lo tanto hay un vector V no nulo que es una base del núcleo de A, y cualquier otro vector en el núcleo es necesariamente de la forma λV, donde λ es algún número real. En la jerga de los sistemas lineales diremos que las soluciones de la ecuación homogénea $AX = 0$ son de la forma λV, con $\lambda \in \mathbb{R}$.

Encontramos así que la solución general del sistema lineal (3.15), que se obtiene sumando a una solución particular P de la ecuación original todas las posibles soluciones λV de la ecuación homogénea asociada, es

$$P + \lambda V, \quad \lambda \in \mathbb{R}.$$

Recuperamos así la expresión paramétrica de la recta.

Desde un punto de vista más geométrico, recordemos que cada una de las ecuaciones en el sistema (3.15) representa un plano. Cuando la matriz A del sistema tiene rango 2 estos planos se cortan en una recta, cuya representación paramétrica se obtiene al resolver el sistema.

Examinemos qué ocurre cuando el rango de la matriz A es igual a 1. Notemos que en este caso la fila (a_1, b_1, c_1) es proporcional a (a_2, b_2, c_2).

Si el sistema (3.15) tiene solución es porque también la matriz ampliada del sistema tiene rango 1. Esto significa que (a_1, b_1, c_1, d_1) es proporcional a (a_2, b_2, c_2, d_2), y en realidad las dos ecuaciones del sistema son esencialmente la misma, porque una se obtiene de la otra multiplicando por una constante. Tenemos entonces que el sistema es equivalente a una **única** ecuación no trivial, que representa a un plano. Resolver el sistema en este caso es equivalente a calcular la intersección de un plano consigo mismo.

Si el sistema (3.15) es incompatible entonces las dos ecuaciones del sistema representan planos que no se cortan. Se trata de las ecuaciones de dos planos paralelos, tal como veremos en la sección 3.4.

Para la ecuación de un plano valen consideraciones similares. La ecuación reducida de un plano es de la forma

$$ax + by + cz = d, \tag{3.16}$$

en la que al menos uno de los coeficientes a, b y c es distinto de cero. No hay inconveniente en pensar en (3.16) como en un sistema lineal con matriz

$$\begin{pmatrix} a & b & c \end{pmatrix},$$

que tiene rango igual a 1, y aplicarle la teoría que hemos desarrollado en el capítulo 1. Este "sistema" es siempre compatible, y el núcleo de su matriz tiene dimensión 2. Una base del núcleo está formada por dos vectores no colineales U y V y la solución general de (3.16) es de la forma

$$P + \lambda U + \mu V, \quad \lambda, \mu \in \mathbb{R},$$

donde P es una solución particular.

Ejercicio 3.16. Considerar cualquier sistema de ecuaciones $AX = B$, donde A es una matriz $m \times 3$, con $m \geq 1$ un número natural cualquiera. Discutir geométricamente cómo puede ser el conjunto solución del sistema. Sugerencia: ordenar la discusión según los posibles valores del rango de A.

También se puede utilizar la noción de rango para analizar la intersección de dos planos π_1 y π_2 cuando vienen expresados, con notación obvia, por ecuaciones paramétricas del tipo

$$\begin{array}{l} Q = P_1 + \lambda_1 U_1 + \mu_1 V_1, \\ Q = P_2 + \lambda_2 U_2 + \mu_2 V_2. \end{array}$$

Un punto Q está en la intersección si admite representaciones paramétricas con las dos ecuaciones. Para los valores de los parámetros λ_1, λ_2, μ_1 y μ_2 con los que se obtiene el punto Q debe satisfacerse

$$P_1 + \lambda_1 U_1 + \mu_1 V_1, = P_2 + \lambda_2 U_2 + \mu_2 V_2,$$

que, naturalmente, implica

$$\lambda_1 U_1 + \mu_1 V_1 - \lambda_2 U_2 - \mu_2 V_2 = P_2 - P_1. \tag{3.17}$$

Este es un sistema lineal con matriz

$$(U_1, V_1, -U_2, -V_2).$$

Ejercicio 3.17.

1. Discutir cómo son las posibles soluciones del sistema lineal (3.17) e interpretar geométricamente las distintas posibilidades. Sugerencia: ordenar la discusión según los posibles valores del rango de la matriz del sistema.
2. Repetir el análisis para la intersección de una recta y un plano expresados en forma paramétrica. Interpretar geométricamente.

3.1.4. Cálculo de intersecciones

Las ecuaciones paramétricas y reducidas de rectas y planos permiten calcular las intersecciones entre rectas, entre planos, o entre rectas y planos. Mostramos a continuación algunos ejemplos, y dejamos otros planteados en forma de ejercicios.

Ejemplo 3.1.18. Consideremos las rectas r y r' de ecuaciones reducidas

$$\left\{ \begin{array}{rcl} x + y - z &=& 1, \\ 2x - y + z &=& 0, \end{array} \right. \qquad \left\{ \begin{array}{rcl} -x + 2y + 2z &=& -1, \\ -x - y + z &=& -1, \end{array} \right. \tag{3.18}$$

respectivamente. Nuestro objetivo es buscar la intersección $r \cap r'$ de ambas rectas, que está formada por los puntos que pertenecen a ambas. Como las ecuaciones reducidas de una recta cualquiera expresan condiciones equivalentes a que un punto (x,y,z) pertenezca a ella, tenemos que un punto (x,y,z) está en $r \cap r'$ si y sólo si satisface a la vez todas las ecuaciones que aparecen en (3.18). Por lo tanto sus coordenadas deben satisfacer el sistema

$$\left\{\begin{aligned} -x+2y+2z &= -1,\\ -x-y+z &= -1,\\ x+y-z &= 1,\\ 2x-y+z &= 0. \end{aligned}\right.$$

Ejercicio 3.18. Mostrar que el sistema tiene como única solución $x = 1/3$, $y = 1/6$, $z = -1/2$.

Luego de resolver el ejercicio, concluimos que la intersección $r \cap r'$ consiste del único punto $(1/3, 1/6, -1/2)$ ♣

Ejemplo 3.1.19. En este ejemplo trataremos la misma intersección del ejemplo anterior, pero ahora expresando una de las rectas en forma paramétrica. Para esto escribamos la recta r' en forma paramétrica, haciendo $y = -\mu$. Obtenemos así las siguientes representaciones para r y r', respectivamente:

$$\left\{\begin{aligned} x+y-z &= 1,\\ 2x-y+z &= 0, \end{aligned}\right. \qquad \left\{\begin{array}{l} x = 1+4\mu,\\ y = -\mu,\\ z = 3\mu. \end{array}\right.$$

Las coordenadas (x,y,z) de cualquier punto en $r \cap r'$ deben satisfacer las ecuaciones reducidas de r, y admitir una representación paramétrica proveniente de las ecuaciones de r'. Por lo tanto, el valor del parámetro μ para un punto de r' que además esté en r debe satisfacer las ecuaciones

$$\left\{\begin{array}{ll} (1+4\mu)+(-\mu)-3\mu & = 1,\\ 2(1+4\mu)-(-\mu)+3\mu & = 0. \end{array}\right.$$

Resolviendo el sistema encontramos $\mu = -1/6$. Sustituyendo este valor en las ecuaciones de r' obtenemos las coordenadas del punto de corte

$$x = 1/3, \quad y = 1/6, \quad z = -1/2,$$

que son justamente las que hallamos en el ejemplo anterior. ¡Menos mal! ♣

Ejemplo 3.1.20. Consideremos los planos π y π' de ecuaciones paramétricas

$$\left\{\begin{array}{l} x = \lambda-\mu,\\ y = -1-2\lambda+3\mu,\\ z = 1-\lambda+\mu, \end{array}\right. \qquad \left\{\begin{array}{l} x = -1+\lambda+\mu,\\ y = 1+\lambda+2\mu,\\ z = 2\lambda-\mu, \end{array}\right.$$

Observar que π es el plano de los ejemplos 3.1.11 y 3.1.12. Hallemos $\pi \cap \pi'$.

Una posibilidad es buscar ecuaciones reducidas de los dos planos, y caracterizar la intersección como el conjunto de puntos que satisface ambas ecuaciones reducidas simultáneamente. Comenzaremos por proceder de esta manera.

En el ejemplo 3.1.12 habíamos encontrado que

$$x + z = 1$$

es una ecuación reducida de π. Encontrar la del plano π' es el objetivo del siguiente ejercicio, que el lector puede obviar en primera instancia, si confía en que el resultado es correcto.

Ejercicio 3.19. Verificar que $5x - 3y - z = -8$ es una ecuación reducida de π'.

Ahora estudiamos el sistema

$$\begin{cases} x + z = 1, \\ 5x - 3y - z = -8. \end{cases}$$

Lo escalerizamos de una manera no muy estándar pero eficiente, sumando a la segunda ecuación la primera para eliminar la tercera variable z. Obtenemos

$$\begin{cases} x + z = 1, \\ 6x - 3y = -7, \end{cases}$$

donde la x puede tomarse como variable libre. Se trata pues de un sistema compatible indeterminado con un grado de libertad. Si llamamos λ a la variable libre, es decir, haciendo $x = \lambda$, resulta que la intersección $\pi \cap \pi'$ admite la descripción paramétrica

$$\begin{cases} x = \lambda, \\ y = \frac{7}{3} + 2\lambda, \\ z = 1 - \lambda, \end{cases}$$

que nos permite identificarla como la recta que pasa por $(0, 7/3, 1)$, y tiene la dirección fijada por el vector $(1, 2, -1)$

No es necesario buscar las ecuaciones reducidas de los planos para calcular la intersección, ya que puede hallarse directamente a partir de las ecuaciones paramétricas. Un punto (x, y, z) está en la intersección si existen λ y ν tales que

$$\begin{cases} x = \lambda - \mu \\ y = -1 - 2\lambda + 3\mu, \\ z = 1 - \lambda + \mu, \end{cases}$$

y existen λ y μ tales que

$$\left\{\begin{array}{l} x=-1+\alpha+\beta, \\ y=1+\alpha+2\beta, \\ z=2\alpha-\beta. \end{array}\right.$$

Pero hay que tener en cuenta un detalle: la primera pareja de valores λ y μ no tiene por qué coincidir con la segunda. Sólo estamos exigiendo que el punto (x,y,z) admita ambas representaciones paramétricas, no que los valores de los parámetros λ y μ para las dos representaciones de (x,y,z) coincidan. Como se trata de variables diferentes, usemos valores diferentes para designarlas. Los puntos (x,y,z) que están en la intersección son aquellos para los que existan valores λ, μ, α y β de los parámetros para los que se satisfaga

$$\left\{\begin{array}{rll} \lambda-\mu & =x= & -1+\alpha+\beta \\ -1-2\lambda+3\mu & =y= & 1+\alpha+2\beta \\ 1-\lambda+\mu & =z= & 2\alpha-\beta \end{array}\right.$$

Naturalmente, si encontramos valores λ, μ, α y β tales que se satisfagan simultáneamente las igualdades

$$\left\{\begin{array}{rcl} \lambda-\mu & = & -1+\alpha+\beta, \\ -1-2\lambda+3\mu & = & 1+\alpha+2\beta, \\ 1-\lambda+\mu & = & 2\alpha-\beta, \end{array}\right. \tag{3.19}$$

habremos encontrado un punto de la intersección. Las coordenadas (x,y,z) del punto se calcular sustituyendo los valores de λ y ν en las ecuaciones paramétricas de π, o α y β en las de π'. Por lo tanto, para hallar la intersección todo lo que hay que hacer es resolver el sistema (3.19) para calcular los valores de los parámetros que corresponden a los puntos en la intersección, y luego recuperar los puntos de la intersección a partir de cualquiera de las ecuaciones paramétricas de los planos.

Reordenamos el sistema y obtenemos

$$\left\{\begin{array}{rcl} \lambda-\mu-\alpha-\beta & = & -1 \\ -2\lambda+3\mu-\alpha-2\beta & = & 2 \\ -\lambda+\mu-2\alpha+\beta & = & -1 \end{array}\right.$$

Escalerizando vemos que $\alpha=2/3$, por tanto α queda determinada y β es variable libre. Volviendo a las ecuaciones paramétricas del plano π' encontramos que la intersección tiene ecuaciones paramétricas

$$\left\{\begin{array}{l} x=-1/3+\beta, \\ y=5/3+2\beta, \\ z=4/3-\beta, \end{array}\right.$$

en las que reconocemos a la recta que pasa por $(-1/3, 4/3, 4/3)$, y tiene la dirección del vector $(1, 2, -1)$. Se trata de una nueva representación paramétrica de la intersección $\pi \cap \pi'$.

Ejercicio 3.20.

1. Completar los cálculos de la resolución del sistema (3.19).
2. Poner los parámetros λ y μ en función de β, y hallar una nueva parametrización de la intersección a partir de las ecuaciones paramétricas para el plano π.
3. Interpretar geométricamente el hecho de que el valor de la variable α haya quedado determinado.

♣

En los ejercicios que proponemos a continuación recurriremos a la figura de referirnos a planos y rectas a través de sus ecuaciones. Creemos que el abuso de lenguaje está justificado por la mayor brevedad de los enunciados. Ver la nota 4 al pie de la página 293. Más adelante en el texto el lector volverá a encontrar este uso.

Ejercicio 3.21. Hallar la intersección de los siguientes planos:

$$2x - 3y + 4z = -2, \qquad \left\{ \begin{array}{l} x = 2 - \lambda + \mu, \\ y = -1 - \lambda + 2\mu, \\ z = -2 - 2\lambda - \mu. \end{array} \right.$$

Ejercicio 3.22. Hallar la intersección del plano y la recta

$$\left\{ \begin{array}{l} x = 2 - \lambda + \mu, \\ y = -1 - \lambda + 2\mu, \\ z = -2 - 2\lambda - \mu. \end{array} \right. \qquad \left\{ \begin{array}{l} x = \alpha, \\ y = 1 - 2\alpha, \\ z = -1 - \alpha. \end{array} \right.$$

Ejercicio 3.23. Se consideran los planos

$$2x + y + z - 2 = 0, \quad \left\{ \begin{array}{l} x = 1 + 2\lambda + 2\mu, \\ y = -3 + \lambda - \mu, \\ z = \lambda + \mu, \end{array} \right.$$

y las rectas

$$\left\{ \begin{array}{r} x + y - 3z = -6, \\ x + 2y - 4z = -8, \end{array} \right. \quad \left\{ \begin{array}{l} x = 3 + \lambda, \\ y = 4 + \lambda, \\ z = 1 - 3\lambda. \end{array} \right.$$

Hallar la intersección de cada una de las dos rectas con cada uno de los dos planos.

Ejercicio 3.24. Perspectiva (I)
Podemos representar una escena tridimensional sobre un plano por medio de la siguiente construcción: el observador se supone ubicado en el punto $O=(0,-1,0)$, y cada punto $P=(x,y,z)$, con $y>0$, se proyecta en un punto que es la intersección de la recta OP con el plano $y=0$. Este nuevo punto tendrá coordenadas $(X,0,Z)$, donde X y Z dependen de las coordenadas (x,y,z) de P. Llamamos

$$\pi(P)=(X,Z),$$

y la correspondencia $P\mapsto\pi(P)$ define entonces una manera de representar el espacio en perspectiva.

1. Hallar las coordenadas (X,Z) de $\pi(P)$ en función de las coordenadas (x,y,z) de P.
2. ¿Cuál es la imagen por π de las rectas del semiespacio $y>0$ que son paralelas al plano $y=0$? En general, ¿cuál es el efecto de π actuando sobre cualquier plano paralelo a $y=0$?
3. Puntos de fuga. Consideremos una recta r que no sea paralela al plano $y=0$. Una vez fijada una de estas rectas hagamos tender al infinito[5] un punto P manteniéndolo sobre r y en el semiespacio $y>0$. Mostrar que cuando P tiende al infinito $\pi(P)$ tiende a un punto del plano que sólo depende de la dirección de r. Llamaremos a este punto el **punto de fuga** correspondiente a esa dirección.
4. Línea del horizonte. Hallar el conjunto formado por los puntos de fuga de las líneas horizontales.

Ejercicio 3.25. Para cada una de las ternas de planos π_1, π_2 y π_3 que se proponen a continuación, hallar la intersección $\pi_1\cap\pi_2\cap\pi_3$ de los tres planos. En caso de que la intersección sea vacía , estudiar las intersecciones dos a dos. Interpretar geométricamente los resultados.

1. $y+z=0$, $2x-y-2z=5$, $3x+3y+2z=7$.
2. $x+2y-z=2$, $2x+y-3z=0$, $-2x-4y+2z=3$.
3. $x-2y+z=5$, $x+z=3$, $x+4y+z=0$.

3.1.5. Extensiones a $\mathbb{K}^m$

Las nociones geométricas que hemos introducido en $\mathbb{R}^3$ están asociadas con la estructura lineal de este espacio, y tienen extensiones a cualquier espacio $\mathbb{K}^m$, en el que disponemos de la misma estructura lineal generada por las operaciones de suma de vectores y producto por un escalar. En realidad, todo puede

[5]Esto es lo mismo que decir que hacer tender a infinito el valor de la coordenada y, manteniéndonos sobre la recta.

extenderse a espacios vectoriales cualesquiera, como los que introduciremos en el capítulo 4.

Comenzamos por introducir la noción de suma de un punto y un conjunto, que permite caracterizar a rectas y planos como la suma de un punto con un subespacio en el que están todos los vectores con las direcciones que corresponden a la recta o al plano que estemos considerando.

Ejercicio 3.26. SUMA DE PUNTO Y CONJUNTO
Podemos definir la suma $P+V$ entre un punto P y un subconjunto cualquiera V de $\mathbb{K}^n$ de la siguiente manera:

$$P+V=\{P+v,\ v\in V\}.$$

Es decir, $P+V$ es el conjunto de puntos que se obtienen sumando a P algún elemento de V.

1. Hallar $P+V$ si $P=(1,3)$ y $V=\{(x,y)\in\mathbb{R}^2;\ 0\leq x\leq 1, 0\leq y\leq 1\}$.
2. Mostrar que la recta de ecuaciones

$$\left\{\begin{aligned} x+y+z&=1,\\ x+z&=3,\end{aligned}\right.$$

   es igual a la suma de la recta de ecuaciones

$$\left\{\begin{aligned} x+y+z&=0,\\ x+z&=0,\end{aligned}\right.$$

   con el punto $(0,-2,3)$.
3. Mostrar que el conjunto de soluciones de cualquier ecuación lineal $AX=B$ es igual a la suma de una solución particular X_0 de esta ecuación, más el núcleo de la matriz A.
4. Interpretar geométricamente esta operación de suma de punto y conjunto. Discutir el caso particular en que el vector nulo de $\mathbb{K}^n$ está en V.

Ejercicio 3.27.

1. Mostrar que $\pi\subset\mathbb{R}^3$ es un plano si y sólo si es un conjunto de la forma $P+V$, donde P es algún punto de $\mathbb{R}^3$ y V es un subespacio vectorial de $\mathbb{R}^3$ que tiene dimensión 2.
2. Expresemos un plano π como suma de un punto y un subespacio de dos formas distintas. Es decir, consideremos

$$\pi=P+V=P'+V',$$

   probar que $V=V'$ y que $P-P'$ está en V. Mostrar que π es igual a la suma de V con cualquier punto en π.

3. Mostrar que la representación de un plano π por medio de un par de ecuaciones paramétricas es equivalente a elegir un punto $P \in \pi$ y una base del subespacio V que permite escribir $\pi = P + V$.
4. Encontrar y demostrar para las rectas del espacio las afirmaciones análogas a las que aparecen en las partes anteriores de este ejercicio.

El ejercicio 3.27 enfatiza el hecho de que planos y rectas en el espacio $\mathbb{R}^3$ tienen la misma estructura algebraica que los conjuntos de soluciones de sistemas de ecuaciones, ver la parte 3 del ejercicio 3.26. Por esta razón podemos representarlos como el conjunto de soluciones de lo que hemos dado en llamar sus ecuaciones reducidas. El objetivo de los ejercicios 3.28 y 3.29 es mostrar cómo esta búsqueda de ecuaciones reducidas y paramétricas puede extenderse a subespacios cualesquiera de cualquier espacio $\mathbb{R}^n$, y también a conjuntos que no son otra cosa que el resultado de trasladar un subespacio. En el ejercicio 3.30 usaremos estas mismas ideas en $\mathbb{Z}_2^7$.

Con este paralelismo entre rectas y planos de $\mathbb{R}^3$ y los subespacios de $\mathbb{K}^n$ esperamos ayudar a construir una cierta intuición geométrica válida para estos espacios, y para la teoría general de espacios vectoriales que desarrollaremos más adelante en el curso. No sorprenderá al lector encontrar que el algoritmo que nos permitirá hacer eficientemente todos los cálculos necesarios es ..., ¿para qué decirlo una vez más?

Ejercicio 3.28.

1. Hallar un sistema de ecuaciones lineales cuyo conjunto solución sea exactamente la imagen de la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 0 & 3 & -1 & 3 \end{pmatrix}.$$

   En lo que sigue llamaremos V a la imagen de A.

2. Determinar una base de V a partir de las ecuaciones halladas. Construir ecuaciones paramétricas para V.
3. Hallar un sistema de ecuaciones lineales cuyo conjunto solución sea exactamente igual a $(1, 1, 3, 3, 6)^t + V$. Repetir para $(-2, 1, 1, 1, 0)^t + V$.
4. Hallar ecuaciones paramétricas para los dos subconjuntos de $\mathbb{R}^5$ que aparecen en la parte anterior.

Ejercicio 3.29.

1. Hallar un sistema de ecuaciones lineales cuyo conjunto solución sea exactamente el conjunto formado por todas las posibles combinaciones lineales de los vectores
$$(1,-1,0,1), \quad (1,1,1,1), \quad (1,0,-1,1)$$
de $\mathbb{R}^4$. Llamemos V a este conjunto, que resulta ser un subespacio vectorial de $\mathbb{R}^4$.
2. Determinar una base de V y construir ecuaciones paramétricas para V.
3. Hallar un sistema de ecuaciones lineales cuyo conjunto solución sea exactamente igual a $(1,1,3,3)^t + V$, y una representación paramétrica del mismo conjunto.

Ejercicio 3.30. CÓDIGOS DE HAMMING (II): MATRICES DE CONTROL
En este ejercicio haremos referencia al ejercicio 1.28 de la sección 1.2.

1. Mostrar que el conjunto de listas $Y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7)^t$ que se obtiene al codificar listas $(x_1, x_2, x_3, x_4)^t$ con las fórmulas del ejercicio 1.28 de la sección 1.2 es el subespacio vectorial de $\mathbb{Z}_2^7$ dado por la imagen de la matriz
$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$
Este subespacio recibe el nombre[6] de **Código de Hamming [7,4]**. Los números 7 y 4 aluden a la longitud de las listas codificadas y de las que se obtienen al codificar. El nombre de Hamming es el de su autor, el matemático norteamericano Richard Hamming (1915-1998) que los introdujo a fines de los años 40.
2. Mostrar que el código de Hamming [7,4] es el conjunto solución de la ecuación lineal $HY = 0$, donde H es la matriz
$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$
A la matriz H se le llama **matriz de control** del código. Observar que hacer la multiplicación HY es una manera sencilla de saber si una lista Y de siete dígitos está o no está en el código. Verificar con la matriz de control algunas de las listas halladas al resolver el ejercicio 1.28 de la sección 1.2 .

[6]En el apéndice A.3.2 hacemos una presentación de las propiedades de este código, y comentamos acerca de su uso para la detección automática de errores. En el apéndice se insiste un poco menos en la estructura lineal del código, para pasar rápidamente a sus propiedades como código corrector, que se presentan a partir de propiedades geométricas basadas en la introducción de la distancia de Hamming.

3. Si B es cualquier columna en $\mathbb{Z}_2^3$, mostrar que el conjunto formado por las listas Y en $\mathbb{Z}_2^7$ que satisfacen la ecuación $HY = B$ puede expresarse como la suma de una cualquiera de estas listas más el Código de Hamming [7,4] (el caso particular en que $B = 0$ es el propio código).

3.1.6. Geometrías discretas, cuadrados latinos y diseño de experimentos

En las subsecciones 3.1.1 a 3.1.4 hemos introducido algunas nociones geométricas en los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$. El objeto de esta subsección es mostrar como el tratamiento algebraico de la geometría permite extenderlas a otros contextos, en particular a conjuntos discretos. Ilustraremos esta posibilidad con un ejemplo en el que resolveremos un interesante problema combinatorio vinculado al diseño de experimentos.

La fórmula

$$y = ax + b, \tag{3.20}$$

donde a y b son dos constantes cualesquiera representa una recta en el plano $\mathbb{R}^2$. Pero la ecuación (3.20) puede formularse también en $\mathbb{K}^2$, donde $\mathbb{K}$ es un cuerpo cualquiera. En particular, para los cuerpos discretos $\mathbb{Z}_p$.

Comencemos por una sencilla observación en $\mathbb{R}^2$: si para un valor fijo de a consideramos todas las posibles rectas (3.20) encontramos que cada punto (x, y) del plano pertenece a una única recta de este tipo. Al variar b generamos una familia de infinitas rectas paralelas cuyos gráficos cubren todo el plano.

La misma construcción puede hacerse sobre cualquier cuerpo $\mathbb{K}$. Ejemplificaremos con $\mathbb{Z}_3 = \{0, 1, 2\}$, pero la construcción es general. El "plano" $\mathbb{Z}_3^2$ tiene sólo 9 puntos, las parejas (x, y) que aparecen en el siguiente cuadro:

$(0,2)$	$(1,2)$	$(2,2)$
$(0,1)$	$(1,1)$	$(2,1)$
$(0,0)$	$(1,0)$	$(2,0)$

.

En este plano hay tres "rectas horizontales", los conjuntos de la forma $y = b$, donde b puede tomar los valores $0, 1, 2$. Estas rectas generan una partición del cuadro en tres filas. También hay tres rectas verticales que particionan el cuadro en columnas.

Más interesante es la partición que generan las rectas $y = x + b$, que tienen "pendiente" igual a 1. Mostramos las tres posibles rectas, con $b = 0, 1, 2$, destacándolas en negrita,

$$\begin{array}{|ccc|}\hline (0,2) & (1,2) & (\mathbf{2},\mathbf{2}) \\ (0,1) & (\mathbf{1},\mathbf{1}) & (2,1) \\ (\mathbf{0},\mathbf{0}) & (1,0) & (2,0) \\ \hline \end{array} \quad \begin{array}{|ccc|}\hline (0,2) & (\mathbf{1},\mathbf{2}) & (2,2) \\ (\mathbf{0},\mathbf{1}) & (1,1) & (2,1) \\ (0,0) & (1,0) & (\mathbf{2},\mathbf{0}) \\ \hline \end{array} \quad \begin{array}{|ccc|}\hline (\mathbf{0},\mathbf{2}) & (1,2) & (2,2) \\ (0,1) & (1,1) & (\mathbf{2},\mathbf{1}) \\ (0,0) & (\mathbf{1},\mathbf{0}) & (2,0) \\ \hline \end{array}$$

$$b = 0,\ y = x; \qquad b = 1,\ y = x + 1; \qquad b = 2,\ y = x + 2.$$

Resulta sugerente el cuadro que se obtiene cuando ponemos en la posición de cada pareja (x, y) del plano $\mathbb{Z}_3^2$ el valor de b que etiqueta a la recta $y = x + b$ a cuyo gráfico pertenece. Obtenemos

$$\begin{array}{|ccc|}\hline 2 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 1 \\ \hline \end{array}. \tag{3.21}$$

Observemos que en este arreglo en cuadro de los tres símbolos 0, 1 y 2, cada uno de ellos aparece exactamente una vez en cada fila y en cada columna.

Un arreglo de n símbolos en un cuadro $n \times n$ de forma tal que cada símbolo aparece exactamente una vez en cada fila y en cada columna es lo que se llama un *cuadrado latino*. Construir cuadrados latinos es un interesante problema combinatorio que tiene aplicaciones en el diseño de experimentos, ver el ejercicio 3.34.

Es posible construir otro cuadrado latino con tres símbolos por medio de la partición del plano $\mathbb{Z}_3^2$ que inducen las rectas de pendiente 2. El resultado es

$$\begin{array}{|ccc|}\hline 2 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ \hline \end{array}, \tag{3.22}$$

tal como podrá verificar el lector. Si recordamos que $2 = -1$ en $\mathbb{Z}_3$, encontramos la explicación del aspecto del cuadro (3.22)

Ejercicio 3.31. Completar los detalles de las construcción del cuadro (3.22)

La construcción de cuadrados latinos con esta técnica puede extenderse a cuerpos finitos $\mathbb{K}$ cualesquiera.

Ejercicio 3.32. Mostrar que si $\mathbb{K}$ es un cuerpo finito con n elementos entonces la construcción precedente puede extenderse en el siguiente sentido: si $a \in \mathbb{K}$ es no nulo, la familia de rectas de ecuación $y = ax + b$ en $\mathbb{K}^2$, donde b toma todos los posibles valores en $\mathbb{K}$, induce una partición de $\mathbb{K}^2$ que da lugar a un cuadrado latino.

Los cuadrados latinos (3.21) y (3.22) guardan entre sí una relación interesante, que se hace evidente al superponerlos:

$$\begin{array}{|ccc|} (2,2) & (1,0) & (0,1) \\ (1,1) & (0,2) & (2,0) \\ (0,0) & (2,1) & (1,2) \end{array}.$$

En este arreglo el primer número de cada pareja es el símbolo que aparece en el cuadro (3.21), y el segundo el que aparece en la misma posición en (3.22). Si examinamos las nueve parejas encontramos que ninguna aparece repetida, y en el cuadro están representadas las 9 posibles parejas formadas por los símbolos 0, 1 y 2. Cuando dos cuadrados latinos tienen esta propiedad al ser superpuestos se dice que son *cuadrados latinos ortogonales.* Nuestro próximo ejercicio extiende el resultado del ejercicio 3.32 para mostrar que la construcción que estamos discutiendo da lugar a cuadrados latinos ortogonales.

Ejercicio 3.33. Sea $\mathbb{K}$ un cuerpo finito. Mostrar que dos cuadrados latinos construidos sobre $\mathbb{K}^2$, con el procedimiento del ejercicio 3.32 a partir de dos familias de rectas

$$y = a_1x + b, \quad y = a_2x + b, \qquad a_1 \neq a_2,$$

son ortogonales.

El siguiente ejercicio muestra la aplicación de los cuadrados latinos al diseño de un experimento.

Ejercicio 3.34. Una revista de automovilismo desea analizar el consumo de cinco coches diferentes, con cinco diferentes tipos de neumáticos y cinco tipos de naftas.

1. Mostrar que hay que hacer 125 ensayos si se desean probar todas las posibles combinaciones de autos, neumáticos y naftas.

2. Mostrar que con sólo 25 ensayos es posible ensayar todas las posibles parejas de autos y naftas, todas las posibles parejas de autos y neumáticos, y todas las posibles parejas de neumáticos y naftas. Hallar un posible diseño experimental que llene estos requisitos. ¿Cuántos diseños de esta naturaleza sos capaz de encontrar?

En esta sección apenas hemos esbozado una pincelada de los temas de diseños combinatorios y geometrías discretas. Quien quiera profundizar un poco más puede consultar las páginas 853–859 del texto [G], también las referencias que allí se citan.

3.1.7. Para tener presente

- Una recta queda determinada al fijar un punto $P \in \mathbb{R}^3$ y un vector no nulo $V \in \mathbb{R}^3$. Distintas elecciones de P y de V pueden determinar la misma recta.
- Un plano queda determinado al fijar un punto $P \in \mathbb{R}^3$ y un par de vectores no nulos y no colineales U y V en $\mathbb{R}^3$. Distintas elecciones de P y de la parajea (U, V) pueden determinar el mismo plano.
- Rectas y planos quedan caracterizados por ecuaciones paramétricas, que especifican cómo recorrerlos. O ecuaciones reducidas, que expresan las condiciones de pertenencia a rectas y planos.
- El pasaje de ecuaciones paramétricas a reducidas puede hacerse usando eliminación gaussiana, y expresando las soluciones en términos de una o dos variables libres, según se trate de una recta o un plano, respectivamente.
- El pasaje de ecuaciones reducidas a paramétricas puede hacerse usando eliminación gaussiana, y obteniendo una o dos ecuaciones de compatibilidad del sistema de ecuaciones paramétricas, según se trate de un plano o una recta.
- El estudio de las intersecciones de rectas y planos se traduce en la formulación de sistemas de ecuaciones lineales. El rango de la matriz de estos sistemas encierra información geométrica relevante.
- La interpretación geométrica de los sistemas de ecuaciones en $\mathbb{R}^3$ puede extenderse de diversas maneras a espacios $\mathbb{K}^n$ más generales, abriendo así nuevos terrenos a la intuición geométrica.

3.2. Producto escalar: distancias y ángulos en $\mathbb{R}^3$

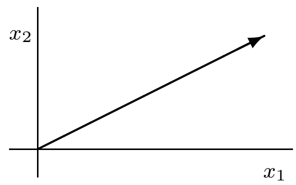
En la sección 3.1 hemos introducido los conceptos de recta y plano a partir de la estructura lineal de $\mathbb{R}^3$. En esta sección vamos a introducir una medida de la longitud de los vectores y de la distancia entre dos puntos o entre dos vectores, y también una noción de ángulo entre vectores. Las medidas de distancias y ángulos que introduciremos modelan correctamente el espacio tal como lo percibimos en las escalas del orden de metros en las que solemos movernos, y en las que el teorema de Pitágoras es válido.

Nuestra presentación nos llevará a introducir el *producto escalar* de dos vectores, para luego definir las nociones de longitud, distancia y ángulo a partir de este producto. Digamos entonces que el producto escalar estará en el centro de nuestro tratamiento de la geometría. Haremos una presentación en el espacio $\mathbb{R}^3$ y el plano $\mathbb{R}^2$, pero todo es fácilmente generalizable a $\mathbb{R}^n$. Sólo hay que desarrollar un poco la imaginación, hasta sentirse cómodo con la idea de un espacio de n dimensiones.

3.2.1. Longitudes, ángulos, producto escalar: discusión preliminar

Comenzaremos nuestra discusión analizando cómo debemos definir la *longitud*, o *módulo*, de los vectores. Para facilitar las cosas discutiremos primero el caso de $\mathbb{R}^2$, para el que podemos representar todos los elementos de la geometría en un dibujo plano de fácil interpretación.

Recordemos que una pareja $X = (x_1, x_2)$ puede ser pensada como un vector, con origen en el origen del sistema de coordenadas $(0,0)$, o como la posición de un punto en el espacio. Concetrémosnos en el primer punto de vista tal como se muestra en la figura.



Los puntos $(0,0)$, $(x_1, 0)$ y (x_1, x_2) son los vértices de un triángulo rectángulo que tiene al vector X como hipotenusa. Los catetos tienen longitudes x_1 y x_2. Si queremos construir una geometría que tenga las propiedades de la geometría euclidiana del plano y sea coherente con la interpretación de que los ejes de abcisas y ordenadas son perpendiculares entre sí, deberíamos definir la longitud $|X|$ del vector X de modo que "el cuadrado de la hipotenusa sea igual a la suma de los cuadrados de los catetos". Es decir, el cuadrado $|X|^2$ debería ser igual a $x_1^2 + x_2^2$, por lo que deberíamos definir

$$|X| = |(x_1, x_2)| = \sqrt{x_1^2 + x_2^2}. \tag{3.23}$$

Insistamos en recordar que **no estamos aplicando el teorema de Pitágoras**. Ni siquiera hemos definido todavía una noción de perpendicularidad que nos habilite a hablar de triángulos rectángulos, catetos e hipotenusas, de modo que **no hay aún ningún teorema de Pitágoras en nuestra teoría**. Lo que estamos haciendo es utilizar ideas geométricas con el fin de **buscar las definiciones que debemos introducir para que haya un teorema de Pitágoras** en nuestro modelo de la geometría basado en los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$. Una vez dadas las definiciones adecuadas enunciaremos y demostraremos una versión adecuada de este famoso teorema (teorema 3.1, en la página 321).

Con esta definición provisional para la longitud de vectores en $\mathbb{R}^2$, podemos pensar que la longitud del vector (x_1, x_2) es justamente la distancia del punto de coordenadas (x_1, x_2) al origen, y extender esta idea a dos puntos cualesquiera del espacio, definiendo la distancia entre X e Y como la longitud del vector diferencia $X - Y$. Si

$$X = (x_1, x_2), \quad Y = (y_1, y_2),$$

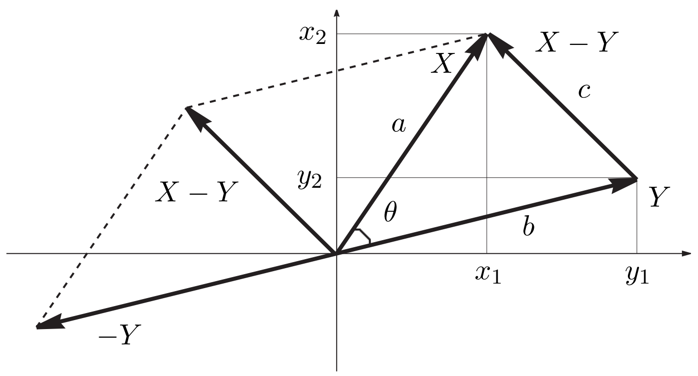
tendremos entonces

$$d\,(X, Y) = |X - Y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, \tag{3.24}$$

donde $d(X, Y)$ representa la distancia de X a Y.

En la fórmula (3.24) está encerrada toda la información necesaria para hacer un tratamiento analítico de la geometría del plano, introduciendo nociones geométricas en $\mathbb{R}^2$.

Analicemos el triángulo de lados X, Y y $X - Y$ que se muestra en la figura.



Hemos llamado a a la longitud $|X|$ de uno de los lados del triángulo, b a $|Y|$, y c a $|X - Y|$. Recordemos el teorema del coseno, una extensión del teorema de Pitágoras que dice

$$c = a^2 + b^2 - ab\cos\theta.$$

Esperamos recuperar toda la geometría clásica en nuestra construcción de la geometría de $\mathbb{R}^2$, de modo que nuestras definiciones deben ser tales que sea válido un teorema del coseno como el que acabamos de evocar. En el triángulo determinado por los vectores X e Y este teorema debería decir

$$|X - Y|^2 = |X|^2 + |Y|^2 - 2|X||Y|\cos\theta. \tag{3.25}$$

El cuadrado $|X - Y|^2$ es relativamente fácil de calcular. Vamos a la fórmula (3.24) y encontramos que

$$|X - Y|^2 = (x_1 - y_1)^2 + (x_2 - y_2)^2.$$

Desarrollando los cuadrados del miembro de la derecha resulta

$$|X - Y|^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 - 2\left(x_1y_1 + x_2y_2\right),$$

en la que reconocemos

$$|X - Y|^2 = |X|^2 + |Y|^2 - 2\left(x_1y_1 + x_2y_2\right).$$

Comparando esta última fórmula con (3.25) reconocemos que, si queremos que en nuestra teoría sea válido el teorema del coseno, la expresión

$$x_1y_1 + x_2y_2 \tag{3.26}$$

debe ser igual a

$$|X||Y|\cos\theta, \tag{3.27}$$

el producto de las longitudes de los vectores por el coseno del ángulo que forman.

La sencilla expresión algebraica (3.26) parece tener un significado geométrico muy importante. Efectivamente es así, y toda la geometría de $\mathbb{R}^2$ está encerrada en el cálculo (3.26), al punto de que lo pondremos en el centro de nuestra teoría.

Recordemos una vez más, a riesgo de hartar al lector, que la noción de ángulo todavía carece de un sentido riguroso en nuestro enfoque algebraico. Sólo nos hemos valido de ella apoyados en la intuición y en nuestra imagen de los elementos de $\mathbb{R}^2$ como coordenadas referidas a un sistema ortogonal. Pero **estas nociones son, por ahora, ajenas a nuestra teoría, porque todavía**

no la hemos definido. Aunque nos estamos acercando a dar una definición. Sin embargo, la expresión (3.26) tiene completo sentido para nosotros, porque

$$x_1,\ x_2,\ y_1,\ y_2$$

son números reales, y podemos multiplicarlos y sumarlos. Tomaremos esta expresión entonces como punto de partida, e iremos construyendo desde esa base las nociones geométricas. Adelantemos que llamaremos *producto escalar* de los vectores X e Y a la expresión (3.26), y la indicaremos con la notación $X \cdot Y$. Tendremos entonces que

$$X \cdot Y = x_1 y_1 + x_2 y_2.$$

Observemos que la longitud de un vector X es igual a la raíz cuadrada del producto escalar $X \cdot X$ de X por sí mismo. En efecto,

$$X \cdot X = x_1^2 + x_2^2 = |X|^2,$$

por lo que

$$|X| = \sqrt{X \cdot X}.$$

Encontramos así que **el módulo o longitud de un vector puede calcularse a partir del producto escalar**. Por lo tanto también las distancias en $\mathbb{R}^2$ son magnitudes que podrán calcularse a partir del producto escalar.

Nuestro análisis nos indica también cómo definir la noción de ángulo entre dos vectores X e Y. Ya sabemos como cálcular $|X|$, $|Y|$ y $X \cdot Y$. Al comparar las expresiones (3.27) y (3.26) concluimos que **deberíamos definir el ángulo θ entre X e Y como el número θ que hace se satisfaga la igualdad**

$$|X||Y| \cos\theta = X \cdot Y.$$

Por lo tanto, despejando θ, encontramos

$$\theta = \arccos\left(\frac{X \cdot Y}{|X||Y|}\right). \tag{3.28}$$

Esta expresión nos acerca a la solución del problema, pero todavía tenemos que trabajar un poco. Notemos que cuando $|X| = 0$ ó $|Y| = 0$ no podemos calcular el cociente que aparece en el miembro de la derecha. Esto no es grave, porque es muy fácil verificar que el único vector que tiene módulo nulo es el vector nulo. Este vector no define ninguna dirección, y no tiene sentido intentar definir el ángulo que forma con cualquier otro vector del espacio. Algo más seria es la cuestión de que ese cociente tiene que tomar valores en el intervalo $[-1,1]$ para que la función arc cos esté definida. Mostraremos luego la importante desigualdad de Cauchy-Schwarz

$$|X \cdot Y| \leq |X||Y|$$

que implica este resultado. Por último, recordemos que arc cos admite múltiples definiciones. Escogeremos la que toma valores en el intervalo $[-\pi, \pi]$.

Abandonaremos ahora la discusión heurística que hemos venido presentando, para pasar a dar las definiciones y propiedades básicas de la teoría. Pero antes dejamos un ejercicio planteado al lector.

Ejercicio 3.35. Estudiar un triángulo de lados

$$X = (x_1, x_2, x_3), \quad Y = (y_1, y_2, y_3), \quad X - Y = (x_1 - y_1, x_2 - y_2, x_3 - y_3),$$

en $\mathbb{R}^3$, y mostrar que la discusión precedente, adaptada a esta nueva situación, sugiere que definamos en $\mathbb{R}^3$ el producto escalar

$$X \cdot Y = x_1y_1 + x_2y_2 + x_3y_3,$$

y que a partir de este producto escalar introduzcamos las nociones de módulo o longitud de un vector, y de ángulo, por medio de las mismas expresiones que encontramos para $\mathbb{R}^2$.

De ahora en adelante desarrollaremos nuestra teoría en $\mathbb{R}^2$ y $\mathbb{R}^3$, poniendo un poco más de énfasis en el caso de $\mathbb{R}^3$. El tratamiento puede ser común, porque sólo depende de las propiedades básicas del producto escalar que se establecen en la proposición 3.1, y que se satisfacen en ambos espacios.

3.2.2. El producto escalar en $\mathbb{R}^3$ y su geometría

Comenzaremos ahora la exposición de la teoría del producto escalar en $\mathbb{R}^3$, poniendo en juego las ideas que acabamos de discutir.

Definición 3.3 (Producto escalar en $\mathbb{R}^3$ y $\mathbb{R}^2$). *Dados dos vectores*

$$X = (x_1, x_2, x_3), \quad Y = (y_1, y_2, y_3),$$

en $\mathbb{R}^3$ definimos su ***producto escalar*** *$X \cdot Y$ por la expresión*

$$X \cdot Y = x_1y_1 + x_2y_2 + x_3y_3. \tag{3.29}$$

Para dos vectores $X = (x_1, x_2)$ e $Y = (y_1, y_2)$ en $\mathbb{R}^2$ definimos

$$X \cdot Y = x_1y_1 + x_2y_2. \tag{3.30}$$

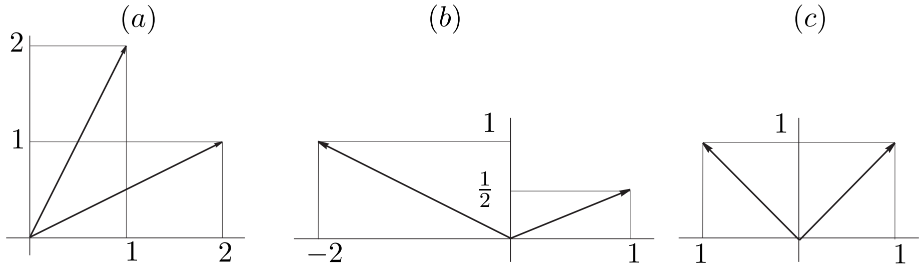


Figura 3.3: Algunas parejas de vectores en $\mathbb{R}^2$

Ejemplo 3.2.1. Calculemos el producto escalar de los vectores que aparecen en la figura 3.3 El primer producto es

$$(1,2) \cdot (2,1) = 1 \cdot 2 + 2 \cdot 1 = 4.$$

El segundo

$$(1,1/2) \cdot (-2,1) = 1 \cdot (-2) + 1/2 \cdot 1 = -3/2,$$

en tanto que el tercero y último arroja el resultado

$$(1,1) \cdot (1,-1) = 1 \cdot 1 + 1 \cdot (-1) = 0.$$

Tal como veremos luego, el hecho de que el producto escalar se anule corresponde a que $(1,1)$ y $(1,-1)$ son perpendiculares. El signo de los otros productos tiene que ver también con el ángulo que hay entre ellos. Ver el ejemplo 3.2.7.

El producto escalar en $\mathbb{R}^3$ tiene las siguientes importantes propiedades.

Proposición 3.1. *Sean X, Y y Z vectores de $\mathbb{R}^3$, y α un número real cualquiera. Entonces el producto escalar satisface las propiedades*

- No negatividad y no degeneración: $X \cdot X \geq 0$ *y* $X \cdot X = 0$ *si y sólo si* X *es el vector nulo;*
- Linealidad respecto a la primera variable:
  - $(X+Y) \cdot Z = X \cdot Z + Y \cdot Z$;
  - $(\alpha X) \cdot Y = \alpha(X \cdot Y)$;
- Simetría: $X \cdot Y = Y \cdot X$.

Ejercicio 3.36.

1. Demostrar la proposición anterior.
2. Demostrar también que el producto escalar depende linealmente de la segunda variable.
3. Mostrar que el producto escalar del vector nulo O con cualquier otro vector del espacio es igual a 0.

Como el producto escalar $X \cdot X$ de un vector por sí mismo es no negativo podemos definir el módulo o longitud de un vector tomando la raíz cuadrada de este producto escalar.

Definición 3.4 (Módulo de un vector). *Dados un vector X en $\mathbb{R}^2$ o $\mathbb{R}^3$ definimos su **módulo** $|X|$ por la expresión*

$$|X| = \sqrt{X \cdot X}. \tag{3.31}$$

Observemos que $|X|^2 = X \cdot X$, y que el único vector de $\mathbb{R}^2$ o $\mathbb{R}^3$ que tiene módulo igual a cero es el vector nulo.

Ejemplo 3.2.2. Podemos calcular el módulo del los vectores del ejemplo 3.2.1. Tenemos

$$(1,2) \cdot (1,2) = 5,$$

de donde concluimos

$$|(1,2)| = \sqrt{5}.$$

Para los demás vectores los cálculos son igual de sencillos, y arrojan los resultados

$$|(2,1)| = |(-2,1)|\sqrt{5}, \quad |(1,1/2)| = \sqrt{5}/2, \quad |(-1,1)| = |(1,1)| = \sqrt{2}.$$

Destaquemos el hecho de que la noción de longitud que hemos introducido está de acuerdo en asignar longitud $\sqrt{2}$ a la diagonal de cuadrado con vértices $(0,0)$, $(1,0)$, $(0,1)$ y $(1,1)$. ♣

Ejercicio 3.37. PRODUCTOS NOTABLES.
Sean X e Y vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$. Demostrar que se satisfacen las igualdades

1. $|X+Y|^2 = |X|^2 + |Y|^2 + 2X \cdot Y$;
2. $(X+Y) \cdot (X-Y) = |X|^2 - |Y|^2$.

Proposición 3.2 (Desigualdad de Cauchy-Schwarz). *Sean X e Y dos vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$. Entonces*

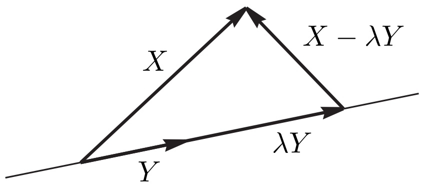
$$|X \cdot Y| \leq |X||Y|. \tag{3.32}$$

Además, la igualdad en (3.32) se alcanza si y sólo si X e Y son colineales.

Prueba. La prueba descansa sobre una idea geométrica muy sencilla: buscar la proyección de X sobre la dirección definida por Y. Es decir, entre todos los posibles vectores λY, con λ real, identificaremos al que hace mínima la longitud

$$|X - \lambda Y|.$$

Esta longitud mínima no puede ser negativa, y es cero sólo si X coincide con su proyección sobre la dirección de Y, lo que requiere que X e Y sean colineales. El resto de la prueba sólo consiste en organizar bien los cálculos para poner



en práctica la idea que acabamos de presentar al lector.

Como la longitud de la diferencia $X-\lambda Y$ involucra tomar una raíz cuadrada, trabajaremos con su cuadrado, por lo que introducimos la función

$$l(\lambda) = |X - \lambda Y|^2,$$

y buscaremos minimizarla. Usando el resultado del ejercicio 3.37 obtenemos

$$l(\lambda) = |X|^2 + |\lambda Y|^2 - 2X \cdot (\lambda Y) = \lambda^2 |Y|^2 - 2\lambda(X \cdot Y) + |X|^2.$$

Se trata de un polinomio en λ, que es de segundo grado cuando $Y \neq O$. Consideremos primero este caso, en que $Y \neq O$. Tenemos que $l(\lambda) \to +\infty$ cuando $\lambda \to \pm\infty$, de modo que la función alcanza su mínimo en algún en el que la su derivada debe anularse. Recordemos además que l siempre es mayor o igual que cero.

Al calcular la derivada encontramos

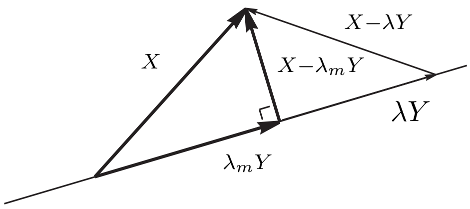
$$l'(\lambda) = 2\lambda |Y|^2 - 2X \cdot Y,$$

que se anula sólo en

$$\lambda_m = \frac{X \cdot Y}{|Y|^2},$$

donde l toma el valor

$$l(\lambda_m) = |X|^2 - \frac{(X \cdot Y)^2}{|Y|^2} \geq 0. \qquad (3.33)$$



De aquí se desprende inmediatamente la desigualdad (3.32) en el caso $|Y| \neq 0$.

Siempre bajo la hipótesis $|Y| \neq 0$ notemos que hay igualdad en (3.32) si y sólo si hay igualdad en (3.33). Por lo tanto, si se satisface la igualdad en (3.32) entonces

$$0 = l(\lambda_m) = |X - \lambda_m Y|,$$

lo que implica $X = \lambda_m Y$.

Por otra parte, si X e Y son colineales y además $Y \neq 0$, se satisface que $X = \lambda Y$ para algún valor de λ. En ese valor de λ la diferencia $Y - \lambda X$ es igual al vector nulo, y la función l se anula. Por lo tanto alcanza su mínimo que es cero, hay igualdad en (3.32), y de allí se deduce la igualdad en (3.32).

Hemos completado entonces la prueba de la proposición cuando $Y \neq 0$. El caso $Y = 0$ es muy sencillo. Si $Y = 0$ tenemos $X \cdot Y = 0$, independientemente de cuál vector sea X, por lo que se satisface la igualdad en (3.32). Además, los vectores X e Y son colineales. □

Observación 3.2.3. Notemos que la desigualdad de Cauchy-Schwarz implica

$$(x_1y_1 + x_2y_2 + x_3y_3)^2 \leq (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2),$$

una desigualdad no trivial para ternas de números. Vale la pena comparar la aparente complejidad de esta expresión, con la sencilla fórmula (3.32). ♠

Ejemplo 3.2.4. Por supuesto, la desigualdad de Cauchy-Schwarz se verifica para todas las parejas de los ejemplos 3.2.1–3.2.2. En efecto, El primer producto es

$$4 = |(1,2) \cdot (2,1)| \leq |(1,2)||(1,2)| = \left(\sqrt{5}\right)^2 = 5;$$

$$\tfrac{3}{2} = |-\tfrac{3}{2}| = \left|\left(1, \tfrac{1}{2}\right) \cdot (-2,1)\right| \leq \left|\left(1, \tfrac{1}{2}\right)\right| |(-2,1)| = \tfrac{\sqrt{5}}{2} \times \sqrt{5} = \tfrac{5}{2};$$

$$0 = |(1,1) \cdot (1,-1)| \leq |(1,1)||(1,-1)| = \left(\sqrt{2}\right)^2 = 2.$$

En todos los casos encontramos una desigualdad, porque ninguna de las parejas está formada por vectores colineales. ♣

Ejercicio 3.38. Para las siguientes parejas de vectores de $\mathbb{R}^2$ y $\mathbb{R}^2$:

1. $(2,1)$ y $(1,3)$;
2. $(1,0,0)$ y $(-1,2,3)$;
3. $(1,-1,2)$ y $(1,1,1)$.

1. Verificar la desigualdad de Cauchy-Schwarz.
2. Hallar λ_m.
3. Interpretar geométricamente los resultados.

Observación 3.2.5. Es importante señalar que ni en la definición de módulo, ni en la demostración de la desigualdad de Cauchy-Schwarz hemos usado para nada la forma explícita del producto escalar. Sólo hemos empleado las propiedades que se enuncian en la proposición 3.1. ♠

Ahora tenemos todos los ingredientes necesarios para definir el ángulo entre dos vectores X e Y no nulos. Sabemos que, bajo esta hipótesis, se tiene $|X| \neq 0$, $|Y| \neq 0$. La desigualdad de Cauchy-Schwarz implica entonces

$$-1 \leq \frac{X \cdot Y}{|X||Y|} \leq 1.$$

Definición 3.5 (Ángulo entre vectores no nulos de $\mathbb{R}^3$). *Dados dos vectores no nulos X e Y en $\mathbb{R}^2$ o $\mathbb{R}^3$ definimos el* ***ángulo*** *θ entre ellos como el único número real $\theta \in [-\pi, \pi]$ tal que*

$$X \cdot Y = |X||Y| \cos \theta.$$

Ejemplo 3.2.6. Consideremos $X = (\sqrt{3}, 1, 0)$ y determinemos los valores que la definición 3.5 arroja para el ángulo con los tres vectores coordenados

$$e_1 = (1,0,0), \quad e_2 = (0,1,0), \quad e_3 = (0,0,1).$$

Comencemos por observar que

$$|X| = \sqrt{3+1+0} = 2.$$

Los vectores e_i, $i = 1,2,3$, tienen módulo 1, por lo que el producto de los módulos es igual a 2 en todos los casos. Los productos escalares dan

$$X \cdot e_1 = \sqrt{3}, \quad X \cdot e_2 = 1, \quad X \cdot e_3 = 0.$$

Los ángulos θ_i entre X y e_i deben satisfacer entonces

$$\cos\theta_1 = \sqrt{3}/2, \quad \cos\theta_2 = 1/2, \quad \cos\theta_3 = 0,$$

de lo que deducimos

$$\theta_1 = \pi/6, \quad \theta_2 = \pi/3, \quad \theta_3 = \pi/2.$$

Estos valores confirman que la definición de ángulo que hemos dado es razonable. ♣

Ejemplo 3.2.7. Podemos calcular los ángulos entre las parejas de vectores de los ejemplos 3.2.1–3.2.2. Tenemos

$$(1,2)\cdot(2,1) = 4, \quad |(1,2)||(1,2)| = 5,$$

de modo que el coseno del ángulo entre $(1,2)$ y $(2,1)$ vale $4/5$. El signo positivo en el producto escalar corresponde a que el ángulo entre los vectores es menor a 90 grados, o $\pi/2$ radianes. El ángulo entre los vectores es cercano a los 37 grados.

El ángulo entre $(1,1/2)$ y $(-2,1)$ tiene un valor cercano a los 127 grados. Un ángulo obtuso, que corresponde al valor negativo $-3/2$ del producto escalar entre ambos vectores.

Los vectores $(1,1)$ y $(-1,1)$ tiene producto escalar nulo. Según nuestra definición les corresponde un ángulo de 90 grados, o $\pi/2$ radianes. ♣

Ejercicio 3.39. Hallar el valor del ángulo entre $(1,1,0)$ y $(1,0,0)$.

Ejercicio 3.40. Calcular el ángulo que forma un vector no nulo X consigo mismo y con $-X$.

En el próximo ejercicio volveremos sobre el significado geométrico que tiene la acción de la matriz

$$G_\alpha = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}.$$

en $\mathbb{R}^2$. En la parte 1 del ejercicio 2.3, página 163, habíamos argumentado en el sentido de que esta matriz representa un giro de ángulo α en el plano. Veremos ahora un par de resultados consistentes con esta imagen.

Ejercicio 3.41. Mostrar que la igualdad

$$|G_\alpha X| = |X|$$

se satisface para todo $X \in \mathbb{R}^2$. Si X no es el vector nulo, mostrar además que el ángulo entre X y $G_\alpha X$ es igual a α.

Ejercicio 3.42. Regla del paralelogramo
Sean X e Y dos vectores cualesquiera. Porbar que

$$\frac{1}{2}\left(|X+Y|^2+|X-Y|^2\right)=|X|^2+|Y|^2.$$

Ejercicio 3.43.

1. Sean X e Y dos vectores cualesquiera. Probar que $X\cdot Y$ está determinado conociendo $|X|$, $|Y|$ y $|X+Y|$. Mostrar que también queda determinado conociendo $|X|$, $|Y|$ y $|X-Y|$.
2. Si $|X|=3$, $|Y|=4$ y $|X+Y|=5$, hallar $X\cdot Y$ y el ángulo entre los vectores X e Y. Repetir para $|X|=3$, $|Y|=4$ y $|X-Y|=5$.
3. Para el caso en que $|X|=|Y|=|X-Y|=1$, hallar $X\cdot Y$ y el ángulo entre los vectores X e Y.

Observemos que cuando el producto escalar entre dos vectores X e Y es nulo, debemos asignar el valor cero al coseno del ángulo θ entre ellos. Por lo tanto $\theta=\pi/2$, lo que conduce a considerar que X e Y son *perpendiculares* u *ortogonales*. A continuación definimos la ortogonalidad entre vectores, de una forma que incluye también al vector nulo.

Definición 3.6. *Dos vectores X e Y se dicen **ortogonales** si $X\cdot Y=0$.*

Esta definición no sólo incluye al vector nulo O, sino que implica además que el vector nulo es ortogonal a todos los vectores del espacio, ver la parte 3 del ejercicio 3.36. Además, el vector nulo es el único que tiene esta propiedad.

Ejercicio 3.44. Mostrar que en $\mathbb{R}^2$ y $\mathbb{R}^3$ se satisface que el vector nulo O es el único vector ortogonal a todos los vectores del espacio.

Ejercicio 3.45. Mostrar que si un vector Y es ortogonal a todos los vectores de una familia

$$(X_1,X_2,\dots X_k),$$

entonces es ortogonal a cualquier combinación lineal de la familia.

Se satisface además el teorema de Pitágoras.

Teorema 3.1 (Pitágoras). *Sean X e Y dos vectores ortogonales de $\mathbb{R}^2$ ó $\mathbb{R}^3$. Entonces*

$$|X+Y|^2=|X|^2+|Y|^2.$$

Prueba. Sólo hay que calcular

$$|X+Y|^2=(X+Y)\cdot(X+Y)=X\cdot X+2X\cdot Y+Y\cdot Y.$$

Como X es ortogonal a Y tenemos $X \cdot Y = 0$. Entonces

$$|X+Y|^2 = X \cdot X + Y \cdot Y = |X|^2 + |Y|^2. \qquad \square$$

La desigualdad de Cauchy-Scwharz permite demostrar la importante desigualdad triangular que satisface el módulo. Resumimos en la próxima proposición las propiedades del módulo.

Proposición 3.3. *Sean X e Y vectores en $\mathbb{R}^2$ o $\mathbb{R}^3$, y α un número real cualquiera. Entonces el módulo $|\cdot|$ satisface las propiedades*

1. No negatividad y no degeneración: *$|X| \geq 0$ y $|X| = 0$ si y sólo si X es el vector nulo del espacio.*
2. Homogeneidad: *$\alpha X = |\alpha||X|$.*
3. Desigualdad triangular: *$|X+Y| \leq |X| + |Y|$.*

Ejercicio 3.46. Demostrar la proposición anterior. Para demostrar la propiedad triangular puede ser útil considerar $|X+Y|^2$ y la desigualdad de Cauchy-Schwarz.

El módulo permite introducir una noción de *distancia* entre los vectores o puntos de $\mathbb{R}^2$ o $\mathbb{R}^3$. Definiremos la distancia de X a Y como la longitud o módulo de la diferencia $X - Y$.

Definición 3.7 (Distancia). *Dados dos vectores X e Y en $\mathbb{R}^2$ o $\mathbb{R}^3$ definimos la **distancia** $d(X,Y)$ entre ambos por la fórmula*

$$d(X,Y) = |X - Y|.$$

Proposición 3.4. *Sean X, Y y Z vectores en $\mathbb{R}^2$ o $\mathbb{R}^3$. Demostrar que la distancia d tiene las siguientes propiedades*

1. No negatividad y no degeneración: *$d(X,Y) \geq 0$ y $d(X,Y) = 0$ si y sólo si $X = Y$.*
2. Simetría: *$d(X,Y) = d(Y,X)$.*
3. Desigualdad triangular: *$d(X,Z) \leq d(X,Y) + d(Y,Z)$.*

Ejercicio 3.47. Demostrar la proposición anterior.

Ejercicio 3.48.

1. Consideremos los vectores $u = (2,-1,-1)/3$, $v = (1,-2,1)$. Hallar $|u|$, $|v|$ y el ángulo entre u e v.

2. Sean u y v dos vectores de $\mathbb{R}^3$. Hallar $|v|$ sabiendo que el ángulo entre u y v es igual a $\pi/4$, $|u| = 3$ y que $u - v$ es perpendicular a u.
3. Hallar $|v|$ y $|u+v|$ sabiendo que el ángulo entre u y v es $\pi/4$, que $|u| = 3$ y que el ángulo entre $u+v$ y u es igual a $\pi/6$.
4. ¿Es cierto que si v es un vector no nulo entonces la igualdad $u \cdot v = w \cdot v$ implica $u = w$? ¿Qué puede decirse de $u - w$?

En el próximo ejercicio usaremos la notación $\overrightarrow{AB}$ para indicar al vector diferencia $B - A$. Tenemos así que $A + \overrightarrow{AB} = B$.

Ejercicio 3.49.

1. Hallar el ángulo que forman dos diagonales en el cubo.
2. Se consideran cuatro puntos coplanares P, Q, R y S. Hallar la condición necesaria y suficiente para que las diagonales del cuadrilátero $PQRS$ sean perpendiculares.
3. Sea $ABCD$ un tetraedro.
   a) Probar que $\overrightarrow{AD} \cdot \overrightarrow{CB} + \overrightarrow{BD} \cdot \overrightarrow{AC} + \overrightarrow{CD} \cdot \overrightarrow{BA} = 0$
   b) Si además $\overrightarrow{AB} \perp \overrightarrow{CD}$ y $\overrightarrow{AD} \perp \overrightarrow{CB}$, probar que $\overrightarrow{AC} \perp \overrightarrow{BD}$.

Antes de cerrar esta sección haremos un par de comentarios sobre otras maneras de representar el producto escalar. En primer lugar digamos que es posible escribir el producto escalar matricialmente, recurriendo a la interpretación de los vectores de $\mathbb{R}^3$ como matrices columna en $M^{3\times 1}(\mathbb{R})$. Si

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

columnas 3×1, entonces el producto

$$X^t Y = (x_1 y_1 + x_2 y_2 + x_3 y_3)$$

es una matriz 1×1, que puede identificarse en forma completamente natural con el número real

$$x_1 y_1 + x_2 y_2 + x_3 y_3 = X \cdot Y.$$

Hechas todas estas identificaciones, tenemos que

$$X \cdot Y = X^t Y,$$

lo que constituye una representación matricial del producto escalar, muy útil para manejar nociones geométricas en el cálculo matricial, o para representar matricialmente operaciones geométricas.

Otra manera corriente de escribir el producto escalar es empleando la notación $\langle\cdot,\cdot\rangle$. Recurriendo a ella escribiríamos

$$X \cdot Y = \langle X, Y\rangle$$

En el próximo ejercicio combinamos los comentarios precedentes, para llamar la atención sobre un resultado de la mayor importancia, a partir del cual es posible derivar un tratamiento geométrico muy interesante del cálculo matricial.

Ejercicio 3.50. Sea A una matriz 3×3. Mostrar que

$$\langle AX, Y\rangle = \langle X, A^tY\rangle.$$

Concluir que la igualdad

$$\langle AX, Y\rangle = \langle X, AY\rangle.$$

se satisface para cualquier pareja de vectores X e Y en $\mathbb{R}^3$ si y sólo si A es una matriz simétrica. Sugerencia: si indicamos con E_i, $i = 1, 2, 3$, a los vectores de la base canónica de $\mathbb{R}^3$, tener en cuenta que

$$a_{ij} = \langle AE_j, E_i\rangle.$$

3.2.3. Proyecciones sobre una recta y versores

En el curso de la demostración de la desigualdad de Cauchy-Schwarz buscamos el vector de la forma λY que hace mínimo el módulo

$$|X - \lambda Y|. \tag{3.34}$$

Lo que hemos hecho entonces es buscar entre todos los vectores colineales con Y el que estaba más cerca de X. Este vector es la *proyección* de X sobre el subespacio generado por Y. Supongamos en lo que sigue que $Y \neq 0$.

Tal como vimos, el valor de λ que hace mínimo (3.34) es

$$\lambda_m = \frac{X \cdot Y}{|Y|^2},$$

por lo que la proyección de X sobre la dirección de Y resulta ser

$$\lambda_m Y = \frac{X \cdot Y}{|Y|^2} Y.$$

Es posible ver más claramente el significado de este vector proyección si lo escribimos en la forma

$$\left(X \cdot \frac{Y}{|Y|}\right) \frac{Y}{|Y|}.$$

¿Qué es ese vector $Y/|Y|$ que aparece en el producto escalar, y también fuera del producto? Se trata de un vector de longitud 1, que tiene la misma dirección y el mismo sentido que Y. Llamaremos e_Y a este vector. Si calculamos $|e_Y|$ obtenemos

$$|e_Y| = \left|\frac{Y}{|Y|}\right| = \frac{1}{|Y|}|Y| = 1.$$

La proyección de X sobre la dirección de Y puede escribirse entonces como

$$(X \cdot e_Y)e_Y.$$

Este vector es justamente la componente de X en la dirección de Y. Observemos que

$$X \cdot e_Y = |X||e_Y|\cos\theta = |X|\cos\theta,$$

donde θ es el ángulo entre X e Y. La longitud de la proyección es

$$|(X \cdot e_Y)e_Y| = |X||\cos\theta|$$

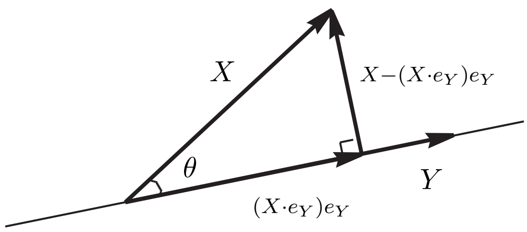


Figura 3.4: Proyección sobre una recta.

Al restar de X su componente $(X \cdot e_Y)e_Y$ en la dirección de Y formamos el vector

$$X - (X \cdot e_Y)e_Y \tag{3.35}$$

que es ortogonal a e_Y. Esto es fácil de verificar haciendo el cálculo del producto interno entre (3.35) y e_Y:

$$(X - (X \cdot e_Y)e_Y) \cdot e_Y = X \cdot e_Y - (X \cdot e_Y)e_Y \cdot e_Y.$$

Como e_Y es un vector de módulo 1 entonces $e_Y \cdot e_Y = 1$ y

$$(X - (X \cdot e_Y)e_Y) \cdot e_Y = X \cdot e_Y - X \cdot e_Y = 0.$$

Se satisface entonces la condición de ortogonalidad que requiere la definición 3.6. Observemos que (3.35) es también ortogonal a Y, ya que Y es colineal con e_Y.

Consideremos el triángulo rectángulo de lados

$$X, \quad (X \cdot e_Y)e_Y, \quad X - (X \cdot e_Y)e_Y.$$

tal como se muestra en la figura 3.4. Su hipotenusa es X, y $(X \cdot e_Y)e_Y$ es la proyección de X sobre la dirección de Y, que resulta ser el cateto adyacente al ángulo θ. La longitud de este cateto es igual al producto $|X| \cos\theta$ de la longitud de la hipotenusa por el coseno del ángulo θ.

Observemos lo sencillo que se vuelve el cálculo de la proyección sobre la dirección de Y una vez que disponemos del vector e_Y, que es colineal con Y y tiene norma 1. Vale la pena distinguir a estos vectores de norma 1, según la siguiente definición.

Definición 3.8. *Diremos que un vector e de $\mathbb{R}^2$ o $\mathbb{R}^3$ es un **vector unitario** o **versor** si el módulo de e es igual a 1.*

Consideremos un vector X y un vector unitario e. Ya sabemos que la proyección $(X \cdot e)e$ de X sobre la dirección de e es el vector que hace mínimo el módulo $|X - \lambda e|$, con λ variando en el conjunto $\mathbb{R}$ de todos los números reales. En el próximo ejercicio proponemos al lector dar una nueva demostración de este hecho, empleando un argumento geométrico basado en el teorema de Pitágoras.

Ejercicio 3.51. Utilizar la igualdad

$$X - \lambda e = (X - (X \cdot e)e) + ((X \cdot e) - \lambda)e$$

para mostrar que

$$|X - \lambda e|^2 \geq |X - (X \cdot e)e|^2,$$

y que la igualdad sólo se alcanza cuando $\lambda = X \cdot e$. Interpretar geométricamente los cálculos.

La proyección de un vector sobre una dirección puede trasladarse para proyectar puntos sobre rectas.

Ejercicio 3.52. Consideremos un punto $Q \in \mathbb{R}^3$ y una recta r de representación paramétrica $P + \lambda V$.

1. Hallar un procedimiento general para determinar el punto Q' de la recta r más cercano a Q. Sugerencia: considerar el versor $V/|V|$;

2. Definimos la distancia $d(Q,r)$ del punto Q a la recta r como
$$d(Q,r) = \text{ínf}\{d(Q,R),\ R \in r\}.$$
Mostrar que $d(Q,r) = d(Q,Q')$.
3. Mostrar que $Q' - Q$ es ortogonal a V.
4. Mostrar que
$$d(Q,r) = |Q - P| \operatorname{sen}\theta,$$
donde θ es el ángulo entre $Q - P$ y V. Interpretar geométricamente este resultado.
5. Hallar el punto de la recta que pasa por $(1,2,3)$ y es paralela a $(1,1,1)$ que está más cerca del origen $O = (0,0,0)$.

3.2.4. Conjuntos ortogonales y ortonormales

La noción de ortogonalidad puede extenderse a familias de vectores.

Definición 3.9. *Diremos que una familia $(X_1, X_2, \cdot, X_n)$ es **ortogonal** si los vectores de la familia son ortogonales dos a dos.*

Una noción muy próxima es la de *ortonormalidad.*

Definición 3.10. *Diremos que una familia $(X_1, X_2, \cdot, X_n)$ es **ortonormal** si es ortogonal, y todos los vectores en ella tienen módulo igual a* 1*.*

Ejemplo 3.2.8. La base canónica
$$((1,0,0),(0,1,0),(0,0,1))$$
de $\mathbb{R}^3$ es una familia ortonormal. Observemos que la noción de perpendicularidad que hemos introducido a través del producto escalar es fiel a la idea intuitiva de pensar el espacio $\mathbb{R}^3$ como el resultado de referir las posiciones en el espacio a un sistema ortogonal de ejes coordenados. ♣

Los vectores de una familia ortonormal son perpendiculares dos a dos, apuntan todos en direcciones diferentes. Esta noción se vincula con la independencia lineal de una familia de vectores, pero es aún más fuerte, tal como se muestra en nuestra siguiente proposición.

Proposición 3.5. *Una familia ortonormal en $\mathbb{R}^2$ o $\mathbb{R}^3$ es linealmente independiente.*

La prueba se reduce a aplicar la definición de independencia lineal, y a aplicar una técnica muy corriente cuando se trabaja con el producto escalar: proyectar la información de que se dispone sobre alguna dirección conveniente, calculando productos escalares con versores convenientemente elegidos.

Prueba. Consideremos una familia ortonormal

$$\mathcal{O} = (e_1, e_2, \ldots, e_n)$$

de versores en $\mathbb{R}^3$, y formemos una combinación lineal que sea igual al vector nulo, escogiendo escalares λ_i, $i = 1, \ldots, n$, de modo que

$$\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_i e_i + \ldots + \lambda_n e_n = O.$$

Al producto escalar de esta combinación con cualquier vector X del espacio, obtenemos, en virtud de la linealidad del producto escalar

$$\lambda_1 e_1 \cdot X + \lambda_2 e_2 \cdot X + \ldots + \lambda_i e_i \cdot X + \ldots + \lambda_n e_n \cdot X = 0.$$

Como la familia $\mathcal{O}$ es ortonormal, si elegimos $X = e_i$ todos los productos $e_j \cdot e_i$ son nulos, salvo el que tiene $j = i$, que vale uno. Concluimos

$$\lambda_i = 0.$$

El razonamiento anterior puede hacerse para cualquier valor de i, lo que implica que todos los coeficientes de la combinación tienen que ser nulos. □

Ejercicio 3.53.

1. ¿Es siempre linealmente independiente una familia ortogonal? ¿Por qué?
2. Mostrar que una familia ortonormal en $\mathbb{R}^3$ puede tener a lo sumo tres vectores, y que si tiene tres vectores es una base de $\mathbb{R}^3$.
3. Enunciar y demostrar una afirmación similar a la de la parte 2 para el caso de $\mathbb{R}^2$.
4. Hallar una base ortonormal de $\mathbb{R}^3$ que contenga al vector $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$.
5. Supongamos que (e_1, e_2, e_3) es una base ortonormal de $\mathbb{R}^3$. Mostrar que para cualquier vector X de $\mathbb{R}^3$ se satisface

$$X = (X \cdot e_1)e_1 + (X \cdot e_2)e_2 + (X \cdot e_3)e_3.$$

6. Hallar la expresión de $(1, 2, 3)$ como combinación lineal de los vectores de la base hallada en la parte 4.

Ejercicio 3.54. Sea (e_1, e_2) una base ortonormal de $\mathbb{R}^2$. Mostrar que si $X \in \mathbb{R}^2$ se expresa como la combinación lineal

$$X = \lambda_1 e_1 + \lambda_2 e_2$$

entonces

$$|X| = \sqrt{\lambda_1^2 + \lambda_2^2}.$$

Enunciar y demostrar un resultado análogo para $\mathbb{R}^3$

Ejercicio 3.55. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ una función tal que

$$|f(X)| = |X|, \quad X \in \mathbb{R}^3.$$

Mostrar que

$$f(X) \cdot f(Y) = X \cdot Y, \quad X, Y \in \mathbb{R}^3.$$

Ejercicio 3.56. Isometrías: transformaciones que conservan longitudes (I)
Sea (e_1, e_2, e_3) una base ortonormal de $\mathbb{R}^3$ y $f : \mathbb{R}^3 \to \mathbb{R}^3$ una función tal que

$$|f(X)| = |X|, \quad X = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

Probar que

$$f(X) = x_1 f(e_1) + x_2 f(e_2) + x_3 f(e_3).$$

Sugerencia: Mostrar que $(f(e_1), f(e_2), f(e_3))$ es una base ortonormal de $\mathbb{R}^3$ y hallar la expresión de $f(X)$ en esta base.

Ahora que tenemos una expresión cuantitativa del tamaño de los vectores de $\mathbb{R}^2$, vamos a volver sobre el problema tratado en los ejercicios 1.25 y 2.7, en las páginas 58 y 168 de este texto. Comenzaremos por estudiar como afecta la acción de la matriz

$$A = \begin{pmatrix} 1 & 0{,}999 \\ 0{,}999 & 1 \end{pmatrix}$$

a la longitud de los vectores $|X|$ a los que se aplica, comparando los módulos de X y de AX.

Ejercicio 3.57. Número de condición (III)

1. Mostrar que los vectores $e_1 = (1/\sqrt{2}, 1/\sqrt{2})$ y $e_2 = (-1/\sqrt{2}, 1/\sqrt{2})$ forman una base ortonormal de $\mathbb{R}^2$.
2. Mostrar que si X se escribe en términos de esta base como $X = \lambda_1 e_1 + \lambda_2 e_2$ entonces

$$AX = 1{,}999\lambda_1 e_1 + 0{,}001\lambda_2 e_2.$$

3. Mostrar que

$$\frac{|AX|^2}{|X|^2} = \frac{1{,}999^2\lambda_1^2 + 0{,}001^2\lambda_2^2}{\lambda_1^2 + \lambda_2^2},$$

y concluir que

$$0{,}001 \leq \frac{|AX|}{|X|} \leq 1{,}999.$$

Observar además que los valores mínimo y máximo, de 0,001 y 1,999, para ese cociente se alcanzan cuando X es colineal con e_1 y e_2 respectivamente.

En el ejemplo del ejercicio anterior encontramos que 0,001 y 1,999 son la mínima y la máxima dilatación que puede producir la matriz A al actuar sobre los vectores de $\mathbb{R}^2$. Tal como vimos en la observación 2.1.5, página 168, estos números son los que controlan lo que ocurre con los errores relativos al resolver el sistema $AX = B$.

Introducimos entonces el número m que mide la más pequeña dilatación que produce A, simplemente examinando todos los cocientes $|AX|/|X|$ para los vectores no nulos del espacio, y tomando el ínfimo. Tenemos

$$m = \text{ínf}\left\{\frac{|AX|}{|X|};\ X \in \mathbb{R}^2,\ X \neq O\right\}. \tag{3.36}$$

De manera análoga, pero tomando un supremo, introducimos el número

$$M = \sup\left\{\frac{|AX|}{|X|};\ X \in \mathbb{R}^2,\ X \neq O\right\}, \tag{3.37}$$

que mide la máxima dilatación que produce A. Observemos que valen las estimaciones

$$m|X| \leq |AX| \leq M|X|, \quad X \in \mathbb{R}^2. \tag{3.38}$$

En el próximo ejercicio vamos a analizar como estos números m y M intervienen en el control de los errores relativos. Consideraremos que al intentar resolver un sistema

$$AX = B,$$

su término independiente B está afectado de error, por lo que en realidad conocemos una perturbación $B+\Delta B$. El vector ΔB es el error absoluto cometido en la determinación de B, la diferencia entre el dato que ponemos en la ecuación, y el valor correcto de B. Es interesante medir este error en relación a B, y es la forma adecuada de medir el error para la mayoría de las aplicaciones. Claro está, por ejemplo, que no es lo mismo cometer un error de algunos milímetros respecto a lo planificado en la construcción de una carretera, que en el análisis de una imagen que se utiliza para preparar una operación al corazón[7]. Introducimos entonces la noción de *error relativo*

$$\delta B = \frac{|\Delta B|}{|B|},$$

comparando el tamaño (módulo) del error absoluto ΔB contra el de B.

Ejercicio 3.58. NÚMERO DE CONDICIÓN (IV)

1. Consideremos una solución X de $AX = B$, y llamemos $X + \Delta X$ a una solución del sistema en el que B se ha visto perturbado a $B + \Delta B$. Es decir, $X + \Delta X$ satisface
$$A(X + \Delta X) = B + \Delta B.$$
El vector ΔX es entonces el error que ΔB introduce en la solución del sistema. Mostrar que $A(\Delta X) = \Delta B$.

2. Concluir que $m|\Delta X| \leq |\Delta B|$. Observar también que $M|X| \geq |B|$. Los números M y m son los que se definen por las fórmulas (3.37) y (3.36).

3. Llamando $\delta X = |\Delta X|/|X|$ al error relativo en X, concluir que se satisface la desigualdad
$$\delta X \leq \frac{M}{m}\delta B.$$
¿Cuál es el sentido de esta desigualdad en el caso en que $m = 0$?

Hemos visto entonces que el cociente M/m controla cómo se propaga el error relativo en los datos a la solución del sistema. Este número es de gran importancia en los cálculos, porque todos los datos de todos los problemas reales están contaminados por algún tipo de error. Además hay que sumar que el procesamiento de los datos suele hacerse en la actualidad en computadoras que no pueden representar exactamente todos los números reales, ya que emplean un sistema de representación de los números que sólo emplea una cantidad finita de dígitos, e introducen errores adicionales.

Este cociente M/m es lo que se llama el *número de condición* de una matriz. Cuando $m \neq 0$ definimos entonces para una matriz A su número de condicion $\kappa(A)$ como

$$\kappa(A) = \frac{\sup\{|AX|/|X|;\ X \neq O\}}{\inf\{|AX|/|X|;\ X \neq O\}}.$$

Por supuesto, vale la acotación

$$\delta X \leq \kappa(A)\delta B$$

[7] En este ejemplo los márgenes de error tolerables no dependen sólo de los órdenes de magnitud de las variables en juego, sino también de lo que está en juego en cada caso.

para el error relativo δX que se comete al resolver $AX = B + \Delta B$ cuando se pretendía resolver $AX = B$. El número de condición $\kappa(A)$ controla la propagación de los errores relativos.

Es posible que el ejemplo que hemos tratado haya dejado en el lector la idea de que el número de condición de una matriz 2×2 es el cociente del máximo valor propio dividido el mínimo valor propio. Aunque esto es cierto para matrices simétricas cuando los valores propios se ordenan según su valor absoluto, no es cierto en general. En el próximo ejercicio mostramos este fenómeno.

Ejercicio 3.59. El objetivo de este ejercicio es determinar el número de condición de las matrices

$$A = \begin{pmatrix} 2 & -n \\ 0 & 1 \end{pmatrix},$$

donde n es un natural cualquiera y compararlo con sus valores propios.

1. Hallar los valores y vectores propios de A.
2. Para calcular el módulo de AX observemos que
$$|AX|^2 = (AX) \cdot (AX) = (AX)^t AX = X^t (A^t A) X = X \cdot (A^t AX).$$
Hallar los valores y vectores propios de A^tA, verificar que existe una base ortonormal (e_1, e_2) de $\mathbb{R}^2$ formada por vectores propios de A^tA y usar esa base para calcular el cociente $|AX|/|X|$ entre el módulo de AX y el de X. Sugerencia: el cuadrado de ese cociente tiene una expresión simple en términos de los coeficientes α_1 y α_2 de la expresion
$$X = \alpha_1 e_1 + \alpha_2 e_2$$
de un vector X cualquiera como combinación lineal de los vectores e_1 y e_2.
3. Hallar el número de condición $\kappa(A)$ de A. Determinar los vectores de $\mathbb{R}^2$ sobre los que A produce las dilataciones máxima y mínima, y sus imágenes al multiplicarlos por A.
4. Calcular $\lim_{n\to\infty} \kappa(A)$. Observar que los valores propios de A no dependen de n. ¿Qué ocurre?

Ejercicio 3.60. Sean A y B dos matrices cuadradas.

1. Mostrar que
$$\kappa(AB) \leq \kappa(A)\kappa(B).$$
2. Dar un ejemplo en el que se alcance la igualdad.
3. Dar un ejemplo en el que $\kappa(AB) = 1$, pero $\kappa(A) = \kappa(B) = 100$.

3.2.5. Producto escalar y ecuaciones de planos y otros objetos

Vimos en la sección 3.1.2 que toda ecuación de la forma

$$ax + by + cz + d = 0, \tag{3.39}$$

con alguno de los coeficientes a, b o c no nulo, representa un plano en $\mathbb{R}^3$. También que cualquier plano puede ser caracterizado por una ecuación reducida como (3.39).

Consideremos un punto particular $P_0 = (x_0, y_0, z_0)$ perteneciente a un plano π representado por la ecuación (3.39). Naturalmente, se satisface

$$ax_0 + by_0 + cz_0 + d = 0, \tag{3.40}$$

Restando (3.40) de (3.39) encontramos que cualquier punto

$$P = (x, y, z) \in \pi$$

debe satisfacer

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

que todavía puede escribirse en la forma más sugerente

$$(a, b, c) \cdot (x - x_0, y - y_0, z - z_0) = 0.$$

Encontramos así que el vector (a, b, c) de los coeficientes de la ecuación es ortogonal a cualquier vector que se obtenga como la diferencia de dos puntos en el plano π. En particular, este vector es ortogonal a los vectores directores del plano. Es corriente llamar N al vector (a, b, c), y decir que es un *vector normal* al plano π. También diremos que es un vector *perpendicular* al plano.

Ejemplo 3.2.9. Si recordamos los ejemplos 3.1.11 y 3.1.12, vemos que dos vectores directores del plano π son ortogonales al vector (a, b, c).

En efecto, la ecuación reducida del plano π es

$$x + z = 1$$

por tanto un vector normal es $N = (1, 0, 1)$. Dos vectores directores del plano son

$$U = (1, -2, -1), \quad V = (-1, 3, 1).$$

Verifiquemos que se cumple que N es ortogonal a U y a V. Para eso debemos calcular $U \cdot N$ y $V \cdot N$ y verificar que ambos productos son nulos:

$$\begin{aligned} U \cdot N &= (1, -2, -1) \cdot (1, 0, 1) = 1 \cdot 1 - 2 \cdot 0 - 1 \cdot 1 = 0 \\ V \cdot N &= (-1, 3, 1) \cdot (1, 0, 1) = -1 \cdot 1 + 3 \cdot 0 + 1 \cdot 1 = 0 \end{aligned}$$

Si recordamos el plano

$$2x + 3y - z + 4 = 0$$

del ejemplo 3.1.14 vemos que un vector normal al plano es $N = (2, 3, -1)$. Este vector es perpendicular a $(1, 0, 2)$ y $(0, 1, 3)$, los dos vectores directores que encontramos en ese ejemplo.

Ejercicio 3.61. Ecuaciones de planos y vectores normales.

1. Hallar la ecuación del plano que pasa por el punto $(-3, 1, 0)$ y es perpendicular a $(-1, 2, 3)$.
2. Expresar el plano π de ecuación $x - y + 3z = 1$ como $(P - P_0) \cdot N = 0$, donde $P = (x, y, z)$ es un punto genérico y N es un vector perpendicular al plano.
3. Hallar dos versores que sean perpendiculares al plano $2x + y = 0$.

El próximo ejercicio vuelve sobre la cuestión que tratamos en el ejercicio 3.13 de la sección 3.1, acerca de hallar una ecuación reducida de un plano a partir de sus ecuaciones paramétricas. En esta oportunidad nuestro tratamiento será más geométrico, y recurriremos al producto escalar para interpretar los resultados.

Ejercicio 3.62. Mostrar que si $U = (u_1, u_2, u_3)$ y $V = (v_1, v_2, v_3)$ son dos vectores directores de un plano π entonces

$$\left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right) \tag{3.41}$$

es un vector normal al plano π. Mostrar, en particular, que este vector es ortogonal a U y a V.

Observación 3.2.10. Producto vectorial
El vector construido a partir de U y V por la fórmula (3.41) es lo que se llama el *producto vectorial* de U y V, y se indica con la notación $U \wedge V$. Este producto vectorial tiene importantes propiedades geométricas, y será nuestro objeto de trabajo en la sección 3.3. ♣

Hemos visto como las ecuaciones de los planos tienen una sencilla interpretación geométrica. Otros objetos geométricos pueden ser caracterizados por ecuaciones construidas a partir del producto escalar. Por ejemplo, una esfera con centro en un punto P_0 y radio $r > 0$ está formada por todos los puntos P del espacio tales que

$$d(P, P_0) = r.$$

Podemos escribir esta ecuación como $|P - P_0| = r$, que es equivalente a la ecuación
$$|P - P_0|^2 = r^2,$$
que es más fácil de manejar.

Ejemplo 3.2.11. Hallemos la ecuación de la esfera de centro $P_0 = (1, -1, 2)$ y radio 3. Para eso tomamos un punto $P = (x, y, z)$ cualquiera y escribamos en coordenadas la condición para que P pertenezca a la esfera:

$$|P - P_0|^2 = (x-1)^2 + (y+1)^2 + (z-2)^2 = r^2 = 3^2$$

Desarrollando y reordenando tenemos:

$$x^2 + y^2 + z^2 - 2x + 2y - 4z = 3$$

que es justamente la ecuación que estábamos buscando. ♣

Ejercicio 3.63.

1. Hallar la ecuación de la esfera de centro $P_0 = (-1, 2, -2)$ y radio 4.
2. Mostrar que $x^2 + y^2 + z^2 + ax + by + cz + d = 0$ representa una esfera si
$$a^2 + b^2 + c^2 + 4d < 0.$$
Hallar el centro y el radio de la esfera.
3. Hallar centro y radio de la siguiente esfera
$$x^2 + y^2 + z^2 + 2x - 4y + 6z - 11 = 0$$

Observación 3.2.12. Es interesante mencionar que podemos escribir las ecuaciones de planos y esferas sin emplear coordenadas.

Las ecuaciones paramétricas de un plano son de la forma

$$P = P_0 + \lambda U + \mu V, \quad \lambda, \mu \in \mathbb{R},$$

donde P_0 es un punto dado del plano, y U y V una pareja de vectores directores. P indica a un punto genérico del plano. Las ecuaciones reducidas de los planos toman la forma
$$N \cdot (P - P_0) = 0,$$
donde P_0 y P tienen el mismo significado que para las ecuaciones paramétricas, y N es un vector normal al plano.

Por último, digamos que las ecuaciones de las esferas son de la forma

$$|P - P_0|^2 = r^2,$$

donde P_0 es el centro, P un punto genérico en la esfera, y r su radio.

Estas expresiones sin coordenadas son muchas veces muy útiles para trabajar. Proponemos al lector un ejercicio en el que puede hacer uso de ellas.

Ejercicio 3.64. Consideremos una esfera E de centro P_0 y radio r, y un punto Q cualquiera en el espacio. Hallar el punto de E que está más cerca de Q, y el que está más lejos. Discutir según Q, y justificar la discusión. ♠

Ejercicio 3.65. Ecuación de la elipse
Consideremos dos números $r > a > 0$, y los puntos $F_1 = (a, 0)$ y $F_2 = (-a, 0)$ de $\mathbb{R}^2$. Mostrar que el conjunto de los puntos $(x, y) \in \mathbb{R}^2$ tales que la suma de su distancia a F_1 y F_2 es igual a $2r$ queda caracterizado por una ecuación de la forma

$$\left(\frac{x}{\alpha}\right)^2 + \left(\frac{y}{\beta}\right)^2 = 1.$$

Hallar α y β en función de a y r. Interpretar geométricamente los parámetros.

3.2.6. Para tener presente

- El objeto básico de la teoría geométrica de $\mathbb{R}^2$ y $\mathbb{R}^3$ es el producto escalar, que puede calcularse a partir de las operaciones de suma y producto de números reales. Las longitudes de los vectores, distancias entre puntos y el ángulos entre vectores se definen luego, a partir del producto escalar.
- El producto escalar permite proyectar con facilidad sobre una dirección o una recta dadas. Es útil expresar los cálculos en términos de vectores unitarios, o versores.
- Una familia ortonormal es linealmente independiente. Una familia ortonormal con tres vectores es una base de $\mathbb{R}^3$, y los coeficientes de la expresión de un vector cualquiera X como combinación lineal en esta base se calculan haciendo productos escalares.
- Las ecuaciones reducidas de los planos tienen una interpretación geométrica en términos de un vector normal al plano que representan.

3.3. Producto vectorial.

En esta sección discutiremos el importante *producto vectorial* en $\mathbb{R}^3$, que fue introducido en la observación 3.2.10, sobre el final de la sección 3.2, al analizar la construcción de un vector normal para un plano a partir de sus vectores directores. A diferencia del producto escalar, el producto vectorial

$$X \wedge Y$$

entre dos vectores X e Y de $\mathbb{R}^3$ es otro vector de $\mathbb{R}^3$.

El producto vectorial tiene un importante significado geométrico, y es la herramienta adecuada para describir diversos fenómenos físicos. En los problemas de mecánica el producto vectorial aparece asociado a los giros y momentos angulares, algo que describiremos luego.

3.3.1. Definición. Primeras propiedades y aplicaciones

Para definir el producto vectorial retomamos la fórmula (3.41).

Definición 3.11 (Producto vectorial). *El producto vectorial $X \wedge Y$ de dos vectores $X = (x_1, x_2, x_3)$ e $Y = (y_1, y_2, y_3)$ de $\mathbb{R}^3$ es*

$$X \wedge Y = \left(\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}, - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \right) \tag{3.42}$$

Podemos desarrollar los tres determinantes, para escribir

$$X \wedge Y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

Los determinantes que aparecen en la definición del producto vectorial, con sus signos, son justamente los que se obtienen al desarrollar por la primera fila un determinante

$$\begin{vmatrix} \circ & \circ & \circ \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}, \tag{3.43}$$

independientemente de qué aparezca en la primera fila, cuyas entradas hemos sustituidos por circulitos.

Basándonos en esta observación, podemos escribir **formalmente** el producto $X \wedge Y$ como el determinante

$$\begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix},$$

donde e_i, $i = 1, 2, 3$, representan a los tres vectores de la base canónica de $\mathbb{R}^3$. Esto evita la necesidad de recordar cuál es la fórmula que define el producto vectorial. Al "calcular" el determinante recuperamos la expresión (3.42), en la que podemos leer las tres componentes del producto vectorial.

Por otra parte, si en la primera fila de la matriz (3.43) aparecen las componentes (z_1, z_2, z_3) de un vector Z el determinante

$$\begin{vmatrix} z_1 & z_2 & z_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

no es otra cosa que el producto escalar $Z \cdot (X \wedge Y)$. Podemos cambiar el orden en el producto escalar, y hacer dos permutaciones de filas en el determinante para escribir la igualdad

$$(X \wedge Y) \cdot Z = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}. \tag{3.44}$$

El producto $(X \wedge Y) \cdot Z$ es lo que se conoce como *producto mixto* de los tres vectores X, Y y Z, y se indica con la notación (X, Y, Z).

Ejercicio 3.66.

1. Mostrar que los tres productos (X, Y, Z), (Z, X, Y) y (Y, Z, X) arrojan el mismo resultado.
2. Mostrar que $(X \wedge Y) \cdot X = (X \wedge Y) \cdot Y = 0$.

Destaquemos el resultado contenido en la parte 2 del ejercicio 3.66: **el producto vectorial $X \wedge Y$ es ortogonal a X y a Y.**

El producto vectorial permite construir fácilmente la ecuación reducida de un plano π a partir de una representación paramétrica de la forma

$$\pi = \{Q \in \mathbb{R}^3;\ Q = P + \lambda U + \mu V,\ \lambda, \mu \in \mathbb{R}\},$$

donde P es un punto cualquiera de $\mathbb{R}^3$, y U y V son dos vectores no colineales. En efecto, el vector $N = U \wedge V$ es un vector normal al plano π, que queda caracterizado como

$$\pi = \{Q \in \mathbb{R}^3;\ (Q - P) \cdot (U \wedge V) = 0\}.$$

Observemos que $(Q - P) \cdot (U \wedge V)$ es el producto triple $(Q - P, U, V)$, que puede escribirse como un determinante. Reencontramos así el resultado del ejercicio 3.13, pero en un contexto en el que la interpretación geométrica es más clara.

A continuación haremos estos cálculos en un ejemplo.

Ejemplo 3.3.1. Sea el plano π de ecuaciones paramétricas

$$\begin{cases} x = 1 + \lambda - \mu, \\ y = -1 + 2\lambda + \mu, \\ z = -2 + \lambda - 2\mu. \end{cases}$$

Calculemos una ecuación reducida del plano. Sabemos que esta ecuación se obtiene como el producto triple $(Q - P, U, V) = 0$ donde

$$Q = (x, y, z), \quad P = (1, -1, -2)$$

es un punto que pertenece al plano,

$$U = (1, 2, 1) \quad \text{y} \quad V = (-1, 1, -2)$$

son dos vectores directores del plano.

Realizamos el producto triple

$$\begin{vmatrix} x-1 & y+1 & z+2 \\ 1 & 2 & 1 \\ -1 & 1 & -2 \end{vmatrix} = 0$$

y obtenemos una ecuación implícita del plano

$$-5x + y + 3z + 12 = 0$$

donde observamos que $N = (-5, 1, 3)$ es ortogonal a U y V y además se verifica fácilmente que $P \in \pi$. ♣

Ejercicio 3.67. Dados los vectores $U = (2, -1, 7)$, $V = (1, 1, -3)$ y $W = (1, -1, 2)$ calcular:

1. $|U|$, $|V|$, $U \cdot V$, $U \wedge V$, $|U \wedge vV|$.
2. $U \wedge (V \wedge W)$ y $(U \wedge V) \wedge W$. Observar que el producto vectorial no es asociativo.
3. (U, V, W), (U, W, V) y (W, U, V).

Ejercicio 3.68. Probar las siguientes identidades, válidas para vectores U, V, W y X cualesquiera en $\mathbb{R}^3$.

1. Fórmula de expulsión. $(U \wedge V) \wedge W = (U \cdot W)\, V - (W \cdot V)\, U$.
2. Identidad de Lagrange. $(U \wedge V) \cdot (W \wedge \bar{X}) = (U \cdot W)(V \cdot X) - (U \cdot X)(V \cdot W)$.

Ejercicio 3.69. Dados dos vectores u y v, hallar todos los vectores w para los que se satisface $u \wedge w = u \wedge v$.

Veamos una nueva aplicación de esta propiedad de ortogonalidad del producto vectorial respecto a cada uno de sus factores.

Ejemplo 3.3.2. La dirección de la intersección de dos planos no paralelos
La intersección de dos planos puede calcularse, por ejemplo, a partir de sus ecuaciones reducidas resolviendo un sistema de la forma

$$\left\{\begin{aligned} ax+by+cz+d&=0,\\ a'x+b'y+c'z+d'&=0. \end{aligned}\right. \tag{3.45}$$

Si sólo deseamos conocer la dirección de la intersección basta observar que esta dirección debe ser normal a los vectores normales a los dos planos, por lo tanto paralela a su producto vectorial. En resumen, el vector

$$(a,b,c)\wedge(a',b',c')$$

es paralelo a la intersección de los planos definidos por las ecuaciones del sistema (3.45).

En la sección 3.4 haremos uso de este tipo de propiedades y construcciones. Ver, por ejemplo, el ejercicio 3.85, en la página 351. ♣

En nuestro próximo ejemplo mostramos cómo el producto vectorial resuelve el problema de describir el estado de movimiento de un cuerpo que está girando.

Ejemplo 3.3.3. La distribución de velocidades de un rígido que gira
Un objeto rígido está girando alrededor de un eje fijo, ¿cuál es la velocidad que en cada instante tiene cada uno de sus puntos?

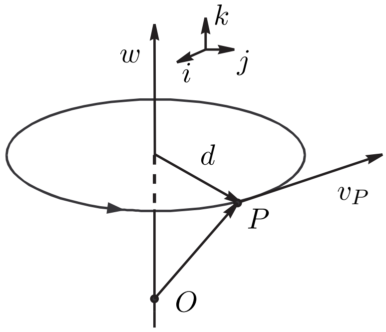
Para fijar ideas supongamos que nuestro objeto gira alrededor del eje Oz, en sentido directo, (antihorario cuando se le observa desde el semieje con $z>0$), con una velocidad angular de ω radianes por unidad de tiempo. Como es usual, determinaremos la posición de un punto P cualquiera en el espacio dando las coordenadas (x,y,z) del vector $P-O$ respecto a una terna ortonormal directa, fija, de vectores. Es usual, en este tipo de problemas, designar a estos vectores como $\vec{i}$, $\vec{j}$, y $\vec{k}$. Mantendremos este uso corriente en los textos de mecánica, y designaremos entonces

$$(\vec{i},\vec{j},\vec{k}),$$

a esta terna de referencia. Al poner coordenadas corresponde a la base $\mathcal{C}=(e_1,e_2,e_3)$ formada por los vectores

$$e_1=(1,0,0),\quad e_2=(0,1,0),\quad e_3=(0,0,1),$$

de $\mathbb{R}^3$. Recordemos que $\mathcal{C}$ es una base ortonormal, por lo que $\mathbb{R}^3$ y el producto escalar permiten modelar correctamente la geometría de la situación que queremos describir.



Todos los puntos del objeto describen circunferencias horizontales alrededor del eje Oz. En cada instante, la velocidad de cada punto es perpendicular al radio vector que une el punto con el eje Oz, tiene un módulo igual al producto de ω por la distancia del punto al eje Oz, y su sentido está determinado por la condición de que el giro ocurre en sentido directo. Por ejemplo, el punto P_x de coordenadas $(x, 0, 0)$ tiene una velocidad igual a $x\omega\vec{j}$. Observemos que $P_x = O + x\vec{i}$. En tanto que el punto $P_y = O + y\vec{j}$, de coordenadas $(0, y, 0)$, tiene velocidad $-y\omega\vec{i}$. Cualquier punto P_z de coordenadas $(0, 0, z)$ tiene velocidad nula.

No es difícil determinar cuál es la velocidad v_P con la que se está desplazando el punto P que ocupa una posición genérica con coordenadas (x, y, z). Su velocidad es horizontal, y perpendicular a $(x, y, 0)$, por lo tanto colineal con $(-y, x, 0)$. El sentido también es el de este vector. Normalicemos $(-y, x, 0)$ para obtener un vector unitario con la dirección y el sentido de la velocidad que queremos calcular. El resultado es

$$e = \frac{(-y, x, 0)}{\sqrt{y^2 + x^2}}.$$

El módulo de la velocidad es $d\omega$, donde $d = \sqrt{x^2 + y^2}$ es igual a la distancia de P al eje O_z. La velocidad $\vec{v}$ es el producto de $d\omega$ por el versor e. Al hacer los cálculos obtenemos

$$\vec{v}_P = \omega(-y\vec{i} + x\vec{j} + 0\vec{k}).$$

Si escribimos $\vec{v}_P$ en coordenadas obtenemos

$$(-\omega y, \omega x, 0),$$

donde reconocemos

$$(-\omega y, \omega x, 0) = (0, 0, \omega) \wedge (x, y, z).$$

El vector

$$(0, 0, \omega) = \omega\vec{k},$$

es la *velocidad angular* del rígido. Allí está codificada la información del eje de giro, definido por la dirección de k; la velocidad de giro, dada por $|\omega|$ y el sentido de giro, definido por el signo de ω. A partir del conocimiento de la velocidad angular, la velocidad de un punto (en nuestro ejemplo usamos que el origen estaba fijo), y la posición de un punto P_0 perteneciente al cuerpo rígido, se puede determinar la velocidad de cualquier otro punto P.

El resultado tiene una característica notable: la velocidad v_P depende linealmente del vector $P-O$, que separa al punto del origen de coordenadas. Esta propiedad de linealidad se obtiene porque $P - O$ entra en el cálculo de v_P a través de un producto vectorial, que depende linealmente de cada uno de sus factores, tal como veremos cuando analicemos las propiedades del producto vectorial. ♣

3.3.2. Propiedades y aplicaciones del producto vectorial

Pasamos ahora a demostrar, a partir de la definición 3.11, una serie de importantes propiedades del producto vectorial. Comenzamos por la antisimetría del producto vectorial: al intercambiar sus factores el producto cambia de signo.

Proposición 3.6. *Para toda pareja de vectores X e Y en $\mathbb{R}^3$ se tiene*

$$X \wedge Y = -Y \wedge X.$$

Prueba. Si $X = (x_1, x_2, x_3)$ e $Y = (y_1, y_2, y_3)$ se tiene

$$\begin{aligned} X \wedge Y &= (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)\,, \\ Y \wedge X &= (y_2x_3 - y_3x_2, y_3x_1 - y_1x_3, y_1x_2 - y_2x_1)\,. \end{aligned}$$

La igualdad $X \wedge Y = -Y \wedge X$ es inmediata. □

A continuación mostramos que el producto vectorial es lineal en cada uno de sus factores.

Proposición 3.7. *El producto vectorial $X \wedge Y$ depende linealmente de cada uno de sus factores en el sentido de que las igualdades*

$$\begin{aligned}(a_1X_1 + a_2X_2) \wedge Y &= a_1X_1 \wedge Y + a_2X_2 \wedge Y,\\ X \wedge (a_1Y_1 + a_2Y_2) &= a_1X \wedge Y_1 + a_2X \wedge Y_2,\end{aligned}$$

se satisfacen para escalares a_1, a_2, y vectores X, X_1, X_2, Y, Y_1, Y_2 cualesquiera.

La demostración se reduce a realizar algunos cálculos sencillos con las componentes. La dejamos como ejercicio para el lector.

Ejercicio 3.70. Demostrar la proposición 3.7. Sugerencia: puede ser útil probar sólo una parte, y usar la proposición 3.6 para probar la otra.

Observación 3.3.4. La proposición 3.7 implica que el producto vectorial es distributivo respecto a la suma de vectores. ♠

Como el producto vectorial es una operación lineal, admite una representación matricial. En el próximo ejercicio mostramos este hecho, que será discutido con mucha mayor generalidad en el capítulo 4.

Ejercicio 3.71. Sea $X = (x_1, x_2, x_3)$ un vector fijo del espacio $\mathbb{R}^3$. Hallar matrices A y B reales 3×3, tales que las igualdades

$$X \wedge Y = AY, \quad Y \wedge X = BY$$

se satisfagan para todo $Y \in \mathbb{R}^3$.

Proposición 3.8.

1. *$X \wedge Y$ es perpendicular a X y a Y;*
2. *el módulo $|X \wedge Y|$ es igual al producto $|X||Y| \operatorname{sen}\theta$, donde θ es el ángulo entre X e Y;*
3. *la terna $(X, Y, X \wedge Y)$ es directa.*

Prueba. La demostración de la parte 1 se reduce a mostrar que los dos productos escalares de $X \wedge Y$ con X e Y son nulos. Este resultado aparece en la parte 2 del ejercicio 3.66.

Calculemos el módulo de $X \wedge Y$, haciendo el producto escalar de este vector por sí mismo, lo que arroja el resultado

$$\begin{aligned}|X \wedge Y|^2 = x_2^2y_3^2 + x_3^2y_2^2 + x_3^2y_1^2 + x_1^2y_3^2 + x_1^2y_2^2 + x_2^2y_1^2 -\\ 2(x_2x_3y_2y_3 + x_1x_3y_1y_3 + x_1x_2y_1y_2).\end{aligned}$$

Los cuadrados que aparecen en los primeros sumandos del miembro de la derecha invitan a completar el producto

$$(x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) = |X|^2|Y|^2.$$

Al hacer aparecer este producto obtenemos

$$\begin{aligned}|X \wedge Y|^2 = {} & |X|^2|Y|^2 - \\ & \left(x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2 + 2x_2x_3y_2y_3 + 2x_1x_3y_1y_3 + 2x_1x_2y_1y_2\right).\end{aligned}$$

En el paréntesis reconocemos

$$(x_1y_1 + x_2y_2 + x_3y_3)^2 = (X \cdot Y)^2,$$

por lo que

$$|X \wedge Y|^2 = |X|^2|Y|^2 - (X \cdot Y)^2.$$

Al tener en cuenta que

$$X \cdot Y = |X||Y| \cos\theta,$$

donde $\theta \in [0, \pi/2]$ es el ángulo entre los vectores X e Y, obtenemos

$$|X \wedge Y|^2 = |X|^2|Y|^2(1 - \cos^2\theta) = |X|^2|Y|^2 \operatorname{sen}^2\theta.$$

La función sen toma valores mayores o iguales que cero en el intervalo $[0, \pi/2]$ que contiene al ángulo θ, entonces

$$|X \wedge Y| = |X||Y| \operatorname{sen}\theta,$$

tal como queríamos probar.

Para mostrar que $(X, Y, X \wedge Y)$ es una terna directa debemos mostrar que el determinante $|X\ Y\ X \wedge Y|$ es mayor o igual que cero. Pero este determinante es igual al producto mixto

$$(X, Y, X \wedge Y) = (X \wedge Y) \cdot (X \wedge Y) = |X \wedge Y|^2 \geq 0. \quad \square$$

Estas propiedades caracterizan completamente al producto vectorial.

Ejercicio 3.72. Dados X e Y no nulos en $\mathbb{R}^3$, mostrar que existe un único vector Z del espacio que tiene la propiedad de que su módulo es igual al producto $|X||Y| \operatorname{sen}\theta$ de los módulos de X e Y y el seno del ángulo θ entre los dos vectores, que es ortogonal a X y a Y, y que hace que el producto triple (X, Y, Z) sea mayor o igual que cero.

Ejercicio 3.73.

1. Mostrar que el producto vectorial $U \wedge V$ es distinto de cero si y sólo si la pareja (U, V) es linealmente independiente;
2. mostrar que cuando (U, V) es linealmente independiente la terna $(U, V, U \wedge V)$ es una base de $\mathbb{R}^3$;
3. mostrar que cuando (U, V) es una pareja ortonormal la terna $(U, V, U \wedge V)$ es una base ortonormal de $\mathbb{R}^3$;
4. mostrar que cuando (U, V) es linealmente independiente un tercer vector W es una combinación lineal de U y V si y sólo si es ortogonal a $U \wedge V$.

Observación 3.3.5. ÁREAS Y VOLÚMENES
Si consideramos el paralelogramo que tiene como lados dos vectores X e Y, encontramos que su área es igual a $|X||Y| \operatorname{sen} \theta$, por lo tanto $|X \wedge Y|$ es igual al área de este paralelogramo.

Si tenemos tres vectores X, Y, Z, podemos considerar el paralelepípedo que los tiene como lados. El volumen V de este paralelepípedo es igual al área de la base por la altura. Supongamos que consideramos como base el paralelogramo de lados X e Y, su área es $|X \wedge Y|$. La altura es igual $|Z| \cos \psi$, donde ψ es el ángulo que forma Z con la dirección perpendicular al plano de X e Y. Este ángulo puede ser igual al ángulo entre Z y $X \wedge Y$, o su suplementario. Entonces

$$V = |X \wedge Y||Z| \cos \psi = \pm(X \wedge Y) \cdot Z = \pm(X, Y, Z),$$

donde aparece un signo de menos cuando $X \wedge Y$ y Z forman un ángulo mayor a $\pi/2$. El volumen coincide con el valor absoluto del producto mixto (X, Y, Z). El producto mixto, que es en definitiva el determinante $|X\ Y\ Z|$, define entonces una noción de volumen orientado: su módulo coincide con el valor del volumen del paralelepípedo de aristas X, Y y Z. Además, es positivo cuando (X, Y, Z) es una terna directa, y negativo cuando la terna no es directa. Esta observación está de acuerdo con la discusión acerca de áreas y volúmenes que presentamos en la sección 2.5.5 en la página 264.

Ejercicio 3.74. Calcular el producto vectorial

$$(a_{11}, a_{12}, 0) \wedge (a_{21}, a_{22}, 0)$$

y usarlo para analizar la interpretación del determinante

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

como un área orientada. ♠

Ejemplo 3.3.6. Distancia de un punto a una recta.
Vimos en la parte 4 del ejercicio 3.52 que la distancia $d(Q,r)$, de un punto Q en el espacio $\mathbb{R}^3$ a la recta $r = \{P + \lambda V;\ \lambda \in \mathbb{R}\}$ que pasa por el punto P y es paralela a la dirección del vector V, puede calcularse por la fórmula $d(Q,r) = |Q-P| \operatorname{sen}\theta$, donde θ es el ángulo que el vector $\vec{PQ} = Q - P$ forma con el vector V que define la dirección de la recta r. El producto vectorial permite calcular fácilmente esta distancia, porque

$$|(Q-P) \wedge V| = |Q-P||V| \operatorname{sen}\theta,$$

por lo tanto

$$d(Q,r) = |Q-P| \operatorname{sen}\theta = \frac{|(Q-P) \wedge V|}{|V|}.$$

Si el vector V tiene módulo igual a 1 entonces la expresión se simplifica porque no es necesario dividir entre $|V|$

Calculemos la distancia del origen $(0,0,0)$ a la recta r de ecuaciones

$$\left\{ \begin{array}{l} x + z = 1, \\ y - z = 2. \end{array} \right. \tag{3.46}$$

Una expresión paramétrica de la recta es

$$\left\{ \begin{array}{l} x = 1 - z, \\ y = 2 + z, \\ z = z, \end{array} \right.$$

de modo que el punto $P = (1,2,1)$ está en r y su dirección está definida por el vector $V = (-1,1,1)$. Calculamos

$$(Q-P) \wedge V = (-1,-2,0) \wedge (-1,1,1) = (-2,1,-3)$$

y obtenemos

$$d(Q,r) = \frac{|(Q-P) \wedge V|}{|V|} = \frac{|(-2,1,-3)|}{|(-1,1,1)|} = \frac{\sqrt{14}}{\sqrt{3}} = \sqrt{\frac{14}{3}}$$

♣

Ejercicio 3.75.

1. Hallar la distancia del origen a la recta de ecuaciones

$$\left\{ \begin{array}{r} x + y + z = 1, \\ x - y = 3. \end{array} \right.$$

2. Hallar la distancia de $(0,0,0)$ al plano $x - y = 3$; y de $(-2,-4,3)$ al plano definido por la ecuación $2x - y + 2z + 3 = 0$.

Ejemplo 3.3.7. Las ecuaciones de los cilindros circulares rectos
Un cilindro circular recto $\mathcal{C}$ de radio r está formado por todos los puntos que están a una distancia r de una recta e dada, a la que llamamos *eje del cilindro*. Entonces, esta superficie es

$$\mathcal{C} = \{Q \in \mathbb{R}^3;\ d(Q,e) = r\}.$$

Vemos entonces que la superficie está caracterizada por la ecuación $d(Q,e) = r$, que puede escribirse fácilmente en coordenadas a partir de una representación paramétrica de la recta e, usando la expresión de la distancia de un punto de a una recta que acabamos de discutir en el ejemplo 3.3.6.

Vamos a calcular la ecuación del cilindro que tiene como eje la recta (3.46), y que contiene al origen $O = (0,0,0)$. La distancia de O al eje es $\sqrt{14/3}$, tal como calculamos en el ejemplo 3.3.6. El vector $V = (-1,1,1)$ tiene módulo $\sqrt{3}$, de modo que están sobre la superficie los puntos (x,y,z) que satisfacen

$$|(x-1, y-2, z-1) \wedge (-1,1,1)| = \sqrt{\frac{14}{3}} \cdot \sqrt{3} = \sqrt{14},$$

que es equivalente a

$$|(x-1, y-2, z-1) \wedge (-1,1,1)|^2 = 14.$$

Al realizar los cálculos encontramos

$$(x-1, y-2, z-1) \wedge (-1,1,1) = (y-z-1, -x-z+2, x+y-3),$$

de donde

$$|(x-1, y-2, z-1) \wedge (-1,1,1)|^2 = (y-z-1)^2 + (-x-z+2)^2 + (x+y-3)^2 = 14$$

implica

$$x^2 + y^2 + z^2 + xy + xz - yz - 5x - 4y - z + 14 = 14$$

finalmente obtenemos

$$x^2 + y^2 + z^2 + xy + xz - yz - 5x - 4y - z = 0. \tag{3.47}$$

Se trata de una ecuación cuadrática que es satisfecha por el origen $(0,0,0)$, porque construimos el cilindro de forma tal que contuviera al origen.

Ejercicio 3.76. Mostrar que un punto (x,y,z) satisface la ecuación (3.47) si y sólo si también

$$(x-\lambda, y+\lambda, z+\lambda)$$

la satisface para todo λ real. Interpretar geométricamente el resultado. ♣

El producto vectorial permite calcular el momento de una fuerza $\vec{F}$ respecto a un punto Q.

Ejemplo 3.3.8. El momento que ejerce una fuerza
Consideremos una fuerza $\vec{F}$ aplicada en un punto P de un cuerpo. Podemos suponer que un punto del cuerpo está sujeto en un punto O, pero no hay ninguna restricción adicional al movimiento[8]. Llamaremos θ al ángulo que forma el vector $\vec{F}$ con $P-O$.

La fuerza $\vec{F}$ tratará de hacer girar el cuerpo en un plano π perpendicular al plano que contiene a la fuerza y al vector $P-O$. Este efecto se describe por la magnitud que llamamos el *momento* de la fuerza respecto al punto O. Se trata de un vector $\vec{M}$ perpendicular a la dirección del plano π (esta dirección perpendicular al plano es la que correspondería al eje de giro) y de un módulo igual al producto de la longitud $|P-O|$ por la componente $|\vec{F}|\operatorname{sen}\theta$ de la fuerza $\vec{F}$ en la dirección del plano π que es perpendicular a $P-O$. El sentido del vector $\vec{M}$ depende del sentido del giro que la fuerza $\vec{F}$ induce al cuerpo, y es convencional definirlo para que la terna $(P-O,\vec{F},\vec{M})$ sea directa.

El vector $\vec{M}$ que tiene todas estas propiedades es justamente el producto vectorial $(P-O)\wedge\vec{F}$. ♣

3.3.3. Para tener presente

- El producto vectorial $X\wedge Y$ entre dos vectores X e Y no colineales es el único vector que satisface

  1. $(X,Y,X\wedge Y)$ es una terna directa;
  2. $X\wedge Y$ es perpendicular a X y a Y;
  3. el módulo de $X\wedge Y$ es igual al producto $|X||Y|\operatorname{sen}\theta$, donde θ es el ángulo entre X e Y.

  Cuando X e Y no son colineales $X\wedge Y=O$. Aunque 1 deja de tener sentido, las condiciones 2 y 3 se siguen satisfaciendo.

- El producto vectorial $X\wedge Y$ es lineal en X e Y, y antisimétrico, en el sentido de que cambia de signo al intercambiar los vectores.

- El producto vectorial tiene muchas aplicaciones geométricas y físicas.

[8]Esta suposición tampoco es esencial, pero ayuda a la visualización de las ideas que buscamos transmitir con este ejemplo

3.4. Rectas y planos: perpendicularidad y paralelismo

En esta sección aplicaremos los productos escalar y vectorial a varias construcciones geométricas en el espacio. Nos concentraremos en las nociones de perpendicularidad u ortogonalidad, y paralelismo. Los ejemplos que trataremos extienden y profundizan nociones que ya hemos tratado antes.

3.4.1. Paralelismo

El paralelismo entre rectas puede caracterizarse con mucha facilidad por medio de sus vectores directores.

Definición 3.12. *Diremos que dos rectas son* ***paralelas*** *cuando sus vectores directores son colineales.*

Vale la pena introducir ahora un término relativamente corriente: cuando dos rectas en el espacio $\mathbb{R}^3$ no son paralelas ni se cortan diremos que *se cruzan.*

Ejercicio 3.77. Dadas las rectas r y r'

$$\left\{\begin{array}{l} x = 1-\lambda, \\ y = -2-\lambda, \\ z = 1-3\lambda. \end{array}\right. \qquad \left\{\begin{array}{l} x-2y = 4, \\ x+z = -1. \end{array}\right.$$

Determinar si son la misma recta. Si no lo son, ¿son paralelas? ¿Se cortan o se cruzan?

Ejercicio 3.78. Consideremos una recta r y un punto Q. Mostrar que por Q pasa una única recta r' que es paralela a r. Mostrar que $r = r'$ si y sólo si $Q \in r$.

Es posible caracterizar el paralelismo de planos en términos de sus vectores directores (ver el ejercicio 3.81), pero daremos una definición en términos de sus vectores normales.

Definición 3.13. *Diremos que dos planos son* ***paralelos*** *cuando sus vectores normales son colineales.*

El significado geométrico de esta definición es evidente. Pero se ve reforzado si volvemos a analizar el problema de la intersección de dos planos, en términos similares a los de la discusión que planteamos en la sección 3.1.4, página 297, cuando estudiamos la representación de una recta como la intersección de dos planos.

Consideremos dos planos planos π_1 y π_2 de ecuaciones

$$a_1x + b_1y + c_1z = d_1, \quad a_2x + b_2y + c_2z = d_2,$$

respectivamente. Los vectores

$$(a_1, b_1, c_1), \quad (a_2, b_2, c_2),$$

son vectores normales a π_1 y π_2. Los puntos que están en la intersección $\pi_1 \cap \pi_2$ corresponden al conjunto solución del sistema

$$\left\{ \begin{array}{l} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2. \end{array} \right. \tag{3.48}$$

La matriz del sistema es

$$A = \left(\begin{array}{ccc} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array} \right).$$

y, según la definición 3.13, tiene rango 1 cuando π_1 y π_2 son paralelos. Hay entonces dos opciones para el sistema (3.48):

1. es incompatible y los planos no se cortan;
2. es compatible. En este caso los dos planos son exactamente el mismo.

En resumen, dos planos paralelos son disjuntos o coinciden. En completo acuerdo con nuestra intuición de la geometría euclidiana de nuestro espacio tridimensional

Ejercicio 3.79. Completar los detalles de la discusión anterior.

Ejercicio 3.80. Dados los planos

$$2x - y - 3z + 2 = 0, \qquad \left\{ \begin{array}{l} x = 2 + \lambda - \mu, \\ y = -2\lambda + 3\mu, \\ z = \mu, \end{array} \right.$$

determinar si son el mismo plano o no. En caso de no serlo, ¿son paralelos?. Hallar su intersección.

Ejercicio 3.81. Mostrar que dos planos son paralelos si y sólo si cada uno de los vectores directores de uno de ellos puede expresarse como combinación lineal de los vectores directores del otro. Sugerencia: si $ax + by + cz = d$ es una ecuación reducida de un plano π entonces los vectores directores de π forman una base del núcleo de la matriz $(a\ b\ c)$.

Así como el ejercicio anterior permite caracterizar el paralelismo entre planos en términos de los vectores directores asociados con representaciones paramétricas, es posible caracterizar el paralelismo de rectas a partir de sus ecuaciones reducidas.

Ejercicio 3.82. Consideremos dos rectas r_1 y r_2 definidas por medio de ecuaciones reducidas. Tendremos dos ecuaciones lineales para cada recta, en total cuatro ecuaciones. Mostrar que las rectas son paralelas si y sólo sí la matriz del sistema formado por estas cuatro ecuaciones lineales tiene rango 2.

Ejercicio 3.83. Mostrar que si dos rectas son coplanares (es decir, hay un plano que las contiene a ambas) entonces son paralelas o tienen un único punto en común.

Dejamos como ejercicio para el lector buscar una o varias definiciones adecuadas para la noción de paralelismo entre una recta y un plano, analizar el sentido geométrico de las posibles definiciones que vengan a su cabeza, mostrar que distintas definiciones razonables son en realidad equivalentes, y buscar, enunciar y demostrar, algún resultado vinculado con esta noción de paralelismo.

Ejercicio 3.84. PARALELISMO DE RECTAS Y PLANOS

1. Formular la parte de la teoría de rectas y planos en el espacio que tiene que ver con el paralelismo de rectas y planos. Enunciar y demostrar los resultados que crea interesante introducir.
2. Sean el plano π y la recta s que se dan a continuación:
$$\begin{cases} x = 2 - \lambda + \mu \\ y = -1 - \lambda + 2\mu \\ z = -2 - 2\lambda - \mu \end{cases} \qquad \begin{cases} x = 1 + \mu \\ y = 1 - \mu \\ z = 0 \end{cases}$$
Decidir si son o no paralelos.

Ejercicio 3.85. Hallar ecuaciones paramétricas de una recta que no corte a ninguno de los planos de ecuaciones
$$x + y + z = 1, \quad x - y = 3$$
y que pasa por el punto $(10, 11, 12)$

Ejercicio 3.86. Hallar ecuaciones reducidas de los planos que satisfacen las condiciones especificadas en cada caso:

1. contiene al punto $(1, 1, 1)$ y es paralelo a las rectas
$$\begin{cases} x = 4 - 5\lambda, \\ y = 5 + 4\lambda, \\ z = -2 + \lambda, \end{cases} \qquad \begin{cases} x + y + z = 0, \\ z = 5. \end{cases}$$
2. Contiene a la intersección de los planos $5x - 2y - z = 3$ y $x + 3y - 2z = -5$, y es paralelo al vector $(7, 9, 17)$.

3.4.2. Perpendicularidad

Las nociones de perpendicularidad pueden reducirse, igual que las de paralelismo, a la perpendicularidad entre vectores.

Diremos que dos rectas son *perpendiculares* si se cortan y sus vectores directores son perpendiculares. Naturalmente, esta definición es independiente de la elección del vector director de cada recta.

Ejercicio 3.87. Mostrar que dos vectores directores U y V de dos rectas r y s son perpendiculares si y sólo si cualquier otra pareja U' y V' de vectores directores de las mismas rectas son perpendiculares.

Diremos que una recta es perpendicular a un plano si su vector director es colineal con el vector normal al plano.

Ejercicio 3.88. ¿Por qué no hemos incluido en la definición de perpendicularidad entre una recta y un plano la condición adicional de que la recta y el plano se corten? Mostrar que la perpendicularidad entre una recta y un plano es una propiedad de la recta y el plano, no de los vectores director y normal escogidos.

Ejercicio 3.89. Probar que las rectas

$$\left\{\begin{array}{l} x+y-3z-1=0, \\ 2x-y-9z-2=0, \end{array}\right. \qquad \left\{\begin{array}{l} 2x+y+2z+5=0, \\ 2x-2y-z+2=0, \end{array}\right.$$

son perpendiculares

Ejercicio 3.90. Probar que las rectas

$$\left\{\begin{array}{l} 2x+y-4z+2=0, \\ 4x-y-5z+4=0, \end{array}\right. \qquad \left\{\begin{array}{l} x=1+2\lambda, \\ y=-2+3\lambda, \\ z=1-6\lambda, \end{array}\right.$$

son perpendiculares.

Ejercicio 3.91. Trazado de perpendiculares.

- Hallar ecuaciones paramétricas y reducidas de la recta que pasa por el punto $(4,4,4)$ y es perpendicular a $2x+y=0$.
- Repetir para la recta que pasa por el origen O y es perpendicular a la recta

$$\left\{\begin{array}{l} 2x+y-4z+2=0, \\ 4x-y-5z+4=0. \end{array}\right.$$

Ejercicio 3.92. Hallar ecuaciones reducidas del plano que contiene a la intersección de los planos $3x-2y+z=3$ y $x-2z=0$, y es perpendicular al plano $x-2y+z=5$.

3.4.3. Proyecciones sobre un plano

Disponer de un vector normal para un plano π proporciona una manera bastante sencilla de determinar cuál es el punto de π que está más próximo a un punto Q cualquiera en el espacio. Hay varias maneras de justificar la construcción. Adoptaremos una que pone en juego ideas geométricas que nos parece interesante destacar.

Proyección sobre un subespacio

Comenzaremos con la proyección de un vector de $\mathbb{R}^3$ sobre un subespacio $\mathbb{S}$ de dimensión 2, que es un plano que pasa por el origen.

Caractericemos a $\mathbb{S}$ por una ecuación de la forma

$$X \cdot N = 0, \tag{3.49}$$

donde N es un versor normal al subespacio. Es decir, tomamos un vector normal N de módulo $|N| = 1$. Esto siempre puede conseguirse dividiendo un vector normal cualquiera por su norma.

El vector $Y \in \mathbb{R}^3$ puede descomponerse en dos componentes. Una componente normal a $\mathbb{S}$, que se obtiene proyectando Y en la dirección de la normal N, y una componente que está en $\mathbb{S}$, de la siguiente manera

$$Y = (Y \cdot N)\, N + (Y - (Y \cdot N)\, N)\,.$$

Es claro que la proyección

$$(Y \cdot N)\, N$$

de Y sobre la dirección de N es un vector colineal con N. Lo que queda de Y luego que restamos la componente según N es perpendicular a N, tal como muestra la igualdad

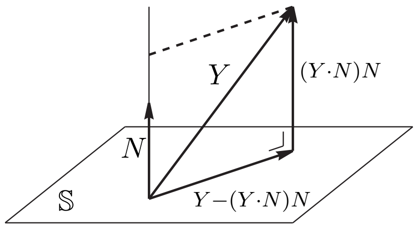
$$(Y - (Y \cdot N)\, N) \cdot N = Y \cdot N - (Y \cdot N)\, N \cdot N = 0,$$

donde hemos usado que $N \cdot N = 1$.

Por lo tanto

$$P_{\mathbb{S}} Y = Y - (Y \cdot N)\, N \tag{3.50}$$

es un vector de $\mathbb{S}$. El vector $P_{\mathbb{S}} Y$ es lo que llamamos la *proyección ortogonal de Y sobre* $\mathbb{S}$. Es el vector de $\mathbb{S}$ más próximo a Y, en el sentido de que si calculamos el módulo $|Y - X|$ de la diferencia entre Y y un vector $X \in \mathbb{S}$ encontraremos que el valor mínimo se alcanza cuando X coincide con $P_{\mathbb{S}} Y$. Mostraremos ahora que esta última afirmación es correcta. El argumento utiliza otra vez ideas de



ortogonalidad y el teorema de Pitágoras, como en el caso de las proyecciones sobre rectas. Para un vector X cualquiera en $\mathbb{S}$ consideramos la diferencia $Y-X$. Sumando y restando el vector $P_{\mathbb{S}}Y$ en esta expresión obtenemos

$$Y-X=(Y-P_{\mathbb{S}}Y)+(P_{\mathbb{S}}Y-X). \tag{3.51}$$

La expresión (3.50) de $P_{\mathbb{S}}Y$ implica

$$Y-P_{\mathbb{S}}Y=(Y\cdot N)\,N$$

El segundo sumando en (3.51) es la diferencia de dos vectores de $\mathbb{S}$, que son ortogonales a N. Por lo tanto es ortogonal a N. Aplicando el teorema de Pitágoras a (3.51) obtenemos

$$|Y-X|^2=|Y-P_{\mathbb{S}}Y|^2+|P_{\mathbb{S}}Y-X|^2\geq|Y-P_{\mathbb{S}}Y|^2.$$

Vemos además que la igualdad se alcanza sólo cuando

$$|P_{\mathbb{S}}Y-X|=0,$$

lo que es equivalente a

$$X=P_{\mathbb{S}}Y.$$

Concluimos entonces que $P_{\mathbb{S}}Y$ es el vector de $\mathbb{S}$ que está más próximo a Y, y que todos los demás vectores de $\mathbb{S}$ están estrictamente más lejos.

Es fácil calcular la distancia de Y al subespacio $\mathbb{S}$. Esta distancia es igual a

$$|Y-P_{\mathbb{S}}Y|=|(Y\cdot N)N|=|Y\cdot N|.$$

Ejemplo 3.4.1. Sean

$$Y=(1,-2,1),\quad \mathbb{S}=\{(x,y,z)\in\mathbb{R}^3:x+y-2z=0\}.$$

Hallaremos la distancia $d(Y,\mathbb{S})$ así como $P_{\mathbb{S}}Y$. Pare ello construiremos un vector N, de longitud uno, ortogonal a $\mathbb{S}$. Sabemos que el vector

$$N'=(1,1,-2),$$

cuyas componentes son los coeficientes de x, y, z en la ecuación que define $\mathbb{S}$, es ortogonal a $\mathbb{S}$. Además

$$|N'| = \sqrt{1+1+4} = \sqrt{6},$$

por lo que normalizamos N con este factor, para obtener

$$N = N'/|N'| = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right).$$

Podemos aplicar ahora la fórmula $P_{\mathbb{S}}Y = Y - (Y \cdot N)N$ para la proyección. Comenzamos por calcular

$$Y \cdot N = (1,-2,1) \cdot \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right) = \frac{1}{\sqrt{6}} - \frac{2}{\sqrt{6}} - \frac{2}{\sqrt{6}} = -\frac{3}{\sqrt{6}}.$$

Entonces,

$$P_{\mathbb{S}}Y = (1,-2,1) - \left(-\frac{3}{\sqrt{6}}\right)\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}\right) = (1,-2,1) + \left(\frac{1}{2}\,, \frac{1}{2}\,, -1\right).$$

En resumen

$$P_{\mathbb{S}}Y = \left(\frac{3}{2}\,, -\frac{3}{2}\,, 0\right)$$

mientras que la distancia de Y a $\mathbb{S}$ es $d(Y,\mathbb{S}) = \frac{\sqrt{6}}{2}$

Ejercicio 3.93. Sean $\mathbb{S} = \{(x,y,z) \in \mathbb{R}^3 : \; -x+3y-2z=0\}$, $P=(0,-1,-2)$ y $Q=(1,0,-1)$ ¿Cuál de los dos puntos está más cerca de $\mathbb{S}$?

Ejercicio 3.94. Sea $\mathbb{S}$ un subespacio de dimensión 2 en $\mathbb{R}^3$, y (e_1, e_2) una base base ortonormal de $\mathbb{S}$. Mostrar que la proyección ortogonal de un vector $X \in \mathbb{R}^3$ sobre $\mathbb{S}$ es

$$P_{\mathbb{S}}X = (X \cdot e_1)e_1 + (X \cdot e_2)e_2.$$

Ejercicio 3.95.

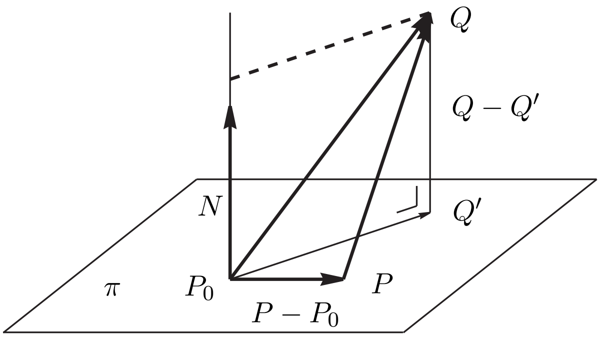
1. Hallar la expresión $P_{\mathbb{S}}X$ de la proyección de un vector genérico $X \in \mathbb{R}^3$ sobre el subespacio $\mathbb{S} = \{(x,y,z) \in \mathbb{R}^3 : \; -x+y-z=0\}$.
2. Mostrar que $P_{\mathbb{S}}X$ puede calcularse como el producto PX de una cierta matriz P que se hallará con X.
3. Mostrar que $P = P^2$ y $P = P^t$. Interpretar geométricamente ambas igualdades matriciales. Sugerencia: volver sobre el ejercicio 3.50, página 324, para analizar el resultado que vincula a P con su traspuesta.

Proyección sobre un plano cualquiera

La construcción anterior puede trasladarse para calcular proyecciones sobre planos cualesquiera, que no tiene por que ser subespacios de $\mathbb{R}^3$. Consideremos un punto $Q \in \mathbb{R}^3$, y un plano π caracterizado por una ecuación de la forma

$$(P - P_0) \cdot N = 0, \tag{3.52}$$

donde P_0 es un punto cualquiera en π y N es un versor normal al plano. Observemos que un punto P está en el plano si y sólo si la diferencia $P - P_0$ está en el subespacio $\mathbb{S}$ caracterizado por la ecuación (3.49).



Nuestro objetivo es determinar un punto Q' en π que sea el más próximo a Q, y la distancia de Q a Q'. Para esto podemos trasladar toda la construcción anterior al punto P_0, observando que la diferencia $Q - P$ entre Q y un punto P cualquiera en π puede escribirse como

$$Q - P = (Q - P_0) - (P - P_0),$$

que es la diferencia entre $Q - P_0$ y un vector $P - P_0$ en S. Por lo tanto $|Q - P|$ se hará mínima para el punto

$$Q' = P_0 + P_S(Q - P_0) = Q - ((Q - P_0) \cdot N)\, N. \tag{3.53}$$

Tendremos además

$$Q - Q' = ((Q - P_0) \cdot N)\, N,$$

que es un vector perpendicular al plano π. La distancia $d(Q, \pi)$ de Q a π es la longitud del vector $Q - Q'$, por lo tanto

$$d(Q, \pi) = |((Q - P_0) \cdot N)\, N|\, .$$

Como el vector N tiene módulo 1, el módulo del vector que aparece en el miembro de la derecha de ésta igualdad es simplemente el valor absoluto del coeficiente $(Q-P_0)\cdot N$ que multiplica a N. Encontramos entoces

$$d(Q,\pi) = |(Q-P_0)\cdot N|\,. \tag{3.54}$$

Esta es una expresión de la distancia de un punto Q a un plano π, que tiene un enorme significado geométrico: la distancia es igual a la longitud de la proyección del vector diferencia entre Q y cualquier punto del plano, sobre una dirección perpendicular al plano. Observemos que el signo de $Q-P_0$ nos indica de qué lado del plano está el punto Q. Cuando es positivo el signo está del lado del plano hacia el que apunta el versor N, cuando es negativo hacia el otro lado. Cuando el producto escalar vale 0 entonces Q está en …¿dónde está Q?.

La expresión (3.54) también puede interpretarse en términos de las ecuaciones reducidas de los planos. Esta manera de mirar las cosas es útil para el cálculo cuando tenemos un plano caracterizado por una única ecuación lineal en $\mathbb{R}^3$, pero es menos geométrica. Notemos que el término de la derecha en (3.54) es el módulo del miembro izquierdo de la ecuación (3.52) del plano π, evaluado en el punto Q. Si estamos trabajando con un plano cuya ecuación en coordenadas es

$$ax+by+cz+d=0,$$

entonces el vector normal es (a,b,c). Para conseguir una ecuación en la que el vector normal tenga módulo 1 dividimos entre el módulo $\sqrt{a^2+b^2+c^2}$ para representar el plano por la ecuación equivalente

$$\frac{ax+by+cz+d}{\sqrt{a^2+b^2+c^2}}=0.$$

Evaluando el miembro derecho en las coordenadas (x,y,z) de un punto cualquiera Q del espacio y tomando el valor absoluto del resultado obtenemos la distancia de Q a π. Es decir, para $Q=(x,y,z)$,

$$d(Q,\pi)=\frac{|ax+by+cz+d|}{\sqrt{a^2+b^2+c^2}}. \tag{3.55}$$

Ejemplo 3.4.2. Dados el punto $Q=(1,0,-1)$ y el plano π

$$2x-y-z+2=0,$$

hallaremos la proyección Q' de Q sobre el plano π, y la distancia $d(Q,\pi)$ del punto al plano. Para calcular la distancia, simplemente aplicamos la fórmula (3.55):

$$d(Q,\pi)=\frac{|2+1+2|}{\sqrt{2^2+(-1)^2+(-1)^2}}=\frac{5}{\sqrt{6}}.$$

Para la proyección, aplicamos la fórmula (3.53). Escogemos en el plano un punto P_0, por ejemplo $P_0 = (-1, 0, 0)$. Tenemos entonces:

$$Q' = (1, 0, -1) - (5/6)(2, -1, -1) = (-4/6, 5/6, -1/6).$$

Como $|N| = \sqrt{6}$, trabajamos con el versor $N/|N|$ en la fórmula (3.53), lo que hace aparecer un factor 6 en el denominador.

Hagamos lo mismo pero ahora con un plano π definido por las ecuaciones paramétricas

$$\begin{cases} x = 1 + \mu, \\ y = -1 + 3\lambda + 2\mu, \\ z = 2 + 2\lambda + \mu. \end{cases}$$

Tomemos $Q = (4, 0, 11)$, y adoptemos como punto P_0 en el plano a $P_0 = (1, -1, 2)$, que es el punto del plano más fácil de calcular a partir de su representación paramétrica. Un vector normal se consigue haciendo el producto vectorial de los vectores directores del plano

$$U = (0, 3, 2), \quad V = (1, 2, 1),$$

que también se obtienen de su representación paramétrica. Construimos un versor normal N normalizando, de modo que

$$N = \frac{U \wedge V}{|U \wedge V|} = \left(-\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, -\frac{3}{\sqrt{14}} \right),$$

es un vector de módulo 1 y normal a π. Podemos calcular la distancia $d(Q, \pi)$ por medio de la fórmula (3.54). Comenzamos evaluando

$$(Q - P_0) \cdot N = (3, 1, 9) \cdot \left(-\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, -\frac{3}{\sqrt{14}} \right) = -2\sqrt{14}.$$

La distancia del punto al plano es entonces

$$d(Q, \pi) = 2\sqrt{14},$$

y la proyección de Q sobre el plano π es el punto

$$Q' = Q + 2\sqrt{14}N = (2, 4, 5).$$

Podemos verificar que los cálculos están bien, porque la diferencia

$$Q - Q' = (4, 0, 11) - (2, 4, 5) = (2, -4, 6)$$

es colineal con $(-1, 2, -3)$, y el punto Q' pertenece al plano, ya que se obtiene escogiendo $\lambda = \mu = 1$. ♣

Ejercicio 3.96. Adaptar los resultados del ejercicio 3.94 para hallar la proyección sobre un plano π dado por ecuaciones paramétricas, cuando los vectores directores del plano forman una familia ortonormal.

Ejercicio 3.97. Dados el punto $Q = (0, -1, 1)$ y el plano π

$$2x - 2y + z - 3 = 0,$$

hallar la distancia del punto al plano, así como la proyección sobre el mismo. Hacerlo por el procedimiento de proyectar sobre la normal, y por el procedimiento del ejercicio 3.96, que pasa por hallar un par de vectores directores que formen una familia ortonormal.

Ejercicio 3.98. Verificar que los siguientes planos son paralelos, y hallar la distancia entre ellos:

1. $x - 2y - 2z = 12$ y $x - 2y - 2z = 6$;
2. $2x - 3y + 6z = 14$ y $4x - 6y + 12z = -21$.

Ejercicio 3.99. En cada caso, hallar las ecuaciones de la recta que satisface las condiciones especificadas:

1. pasa por el punto $(1, 0, 1)$ y es perpendicular al plano $2x + y + 3z - 1 = 0$.
2. Pasa por el punto $(-1, 2, -3)$, se interseca con la recta

   $$P = (1, -1, 3) + \lambda(3, 2, -5),$$

   y es ortogonal a la recta

   $$\left\{\begin{array}{l} x = -1 + 6\lambda, \\ y = -3 - 2\lambda, \\ z = 2 - 3\lambda. \end{array}\right.$$

3. Pasa por el punto $(-4, -5, 3)$ e interseca perpendicularmente a la recta

   $$P = (-1, -3, 2) + \lambda(3, -2, -1).$$

Ejercicio 3.100. En cada caso hallar la ecuación del plano que satisface las condiciones especificadas:

1. Pasa por el punto $(2, -1, 1)$, es perpendicular al plano $2x + 3y - z + 5 = 0$ y paralelo a la recta

   $$\left\{\begin{array}{l} x = 2z + 3, \\ y = z. \end{array}\right.$$

2. Pasa por el punto $B = (1, 0, 1)$ y es paralelo al plano $x + 2y + z + 1 = 0$.
3. Pasa por el punto $C = (1, 1, 1)$, es paralelo al eje Oy y forma un ángulo de $\pi/6$ con el eje Ox (hay dos soluciones posibles).

Ejemplo 3.4.3. Proyección de una recta sobre un plano.
Podemos proyectar toda una recta sobre un plano π, proyectando cada uno de los puntos de la recta sobre el plano. Consideremos una expresión paramétrica de una recta r, de la forma

$$Q = P + \lambda V, \quad \lambda \in \mathbb{R}.$$

Apliquemos a cada punto Q la fórmula (3.53) de la proyección sobre π. Obtenemos puntos

$$Q' = P + \lambda V - ((P + \lambda V - P_0) \cdot N)\, N, \quad \lambda \in \mathbb{R}.$$

Entonces

$$Q' = P - ((P - P_0) \cdot N)\, N + \lambda\, (V - (V \cdot N)N)\,, \quad \lambda \in \mathbb{R}.$$

Si reconocemos que

$$P' = P - ((P - P_0) \cdot N)\, N$$

es la proyección de P sobre π, y que

$$P_S V = (V - (V \cdot N)N)$$

es la proyección de V sobre el subespacio S de las direcciones paralelas al plano π; encontramos que los puntos Q' son de la forma

$$Q' = P' + \lambda P_S V, \quad \lambda \in \mathbb{R}.$$

Esta fórmula es la expresión paramétrica de una recta, salvo en el caso en que $P_S V$ es el vector nulo (la recta es ortogonal a S) .

Calcularemos la proyección en el caso particular de la recta s y el plano π definidos por las ecuaciones

$$\left\{\begin{array}{l} x = 1 + \mu, \\ y = 1 - \mu, \\ z = 0, \end{array}\right. \qquad -x + 2y - 3z - 2 = 0.$$

Comenzamos por identificar los elementos básicos de la recta, un punto P en ella y un vector director V, y la un versor N normal al plano. Encontramos rápidamente, a partir de las ecuaciones,

$$P = (1, 1, 0), \quad V = (1, -1, 0), \quad N = \frac{1}{\sqrt{14}}(-1, 2, -3).$$

Tomaremos $P_0 = (0,1,0)$. Entonces

$$\begin{aligned} P' &= (1,1,0) - \left(\tfrac{1}{14}, -\tfrac{2}{14}, \tfrac{3}{14}\right) = \left(\tfrac{13}{14}, \tfrac{16}{14}, -\tfrac{3}{14}\right), \\ P_S V &= (1,-1,0) - \left(\tfrac{3}{14}, -\tfrac{6}{14}, \tfrac{9}{14}\right) = \left(\tfrac{11}{14}, -\tfrac{8}{14}, -\tfrac{9}{14}\right). \end{aligned}$$

Una representación paramétrica de la recta proyección es

$$\left\{ \begin{array}{l} x = 13/14 + 11\mu, \\ y = 16/14 - 8\mu, \\ 14z = -3/14 - 9\mu. \end{array} \right.$$

El lector notará que hemos multiplicado $P_S V$ por 14, para obtener un nuevo vector director de la recta con una expresión más sencilla. ♣

3.4.4. Normal común y distancia entre rectas

Dadas dos rectas no paralelas en $\mathbb{R}^3$ existe una única recta que es perpendicular a ambas, a la que es usual llamar la *normal común* a las rectas. En esta subsección discutiremos como encontrar esta normal común.

Consideremos dos rectas r y s dada por ecuaciones paramétricas $P + \lambda U$ y $Q + \mu V$ respectivamente, con U y V no colineales. La dirección de $U \wedge V$ es normal a ambas rectas y es, en consecuencia, la dirección de la normal común n a r y a s. Observemos que $U \wedge V$ es no nulo, por la hipótesis de que U y V no son colineales. Para determinar completamente a la normal común n hay que determinar algún punto sobre esta recta. Veremos como hallar los puntos P' y Q' en que n corta a r y s respectivamente.

Observemos que, para algún valor de λ debemos tener

$$P' = P + \lambda U. \tag{3.56}$$

En tanto que para algún valor de μ tendremos

$$Q' = Q + \mu V. \tag{3.57}$$

Además $Q' - P'$ es colineal con $U \wedge V$, por lo que existe un tercer número η tal que

$$Q' - P' = \eta(U \wedge V).$$

Por lo tanto debe satisfacerse

$$(Q + \mu V) - (P + \lambda U) = \eta(U \wedge V),$$

algo que puede reordenarse en la forma

$$\mu V - \lambda U - \eta(U \wedge V) = P - Q. \tag{3.58}$$

Observemos que (3.58) es un sistema lineal de ecuaciones en las incógnitas $(\mu, -\lambda, -\eta)$, cuya matriz tiene a los vectores U, V y $U \wedge V$ como columnas. Por lo tanto, el determinante de la matriz del sistema es el producto mixto

$$(U, V, U \wedge V) = |U \wedge V|^2 \neq 0.$$

Como el determinante es no nulo la matriz es invertible, y (3.58) siempre tiene una única solución.

Ejercicio 3.101. Demostrar que la matriz $(U, V, U \wedge V)$ es invertible siempre que U y V no sean colineales, usando un argumento que no recurra a la teoría de determinantes.

Veremos ahora en un ejemplo como determinar la normal común a dos rectas.

Ejemplo 3.4.4. Hallaremos la normal común a las rectas r y s, de ecuaciones

$$\left\{\begin{array}{l} x = 1 - 2\lambda, \\ y = -3\lambda, \\ z = 2 - \lambda, \end{array}\right. \qquad \left\{\begin{array}{l} x = 2 + \mu, \\ y = 1 - \mu, \\ z = -1 - \mu, \end{array}\right.$$

respectivamente. La recta r tiene a $U = (-2, -3, -1)$ como vector director, en tanto que $V = (1, -1, -1)$ es un vector director de s. A partir de las ecuaciones paramétricas rápidamente identificamos los puntos $P = (1, 0, 2) \in r$, y $Q = (2, 1, -1) \in s$. El vector

$$U \wedge V = (2, -3, 5)$$

es perpendicular a ambas rectas. Buscamos descomponer la diferencia

$$P - Q = (-1, -1, 3)$$

en la suma de un vector paralelo a r, uno paralelo a s y otro perpendicular a ambas rectas. Esto es equivalente a plantearse el sistema (3.58), que, en este ejemplo, toma la forma

$$\mu(1, -1, -1) - \lambda(-2, -3, -1) - \eta(2, -3, 5) = (-1, -1, 3).$$

Igualando componente a componente, obtenemos

$$\left\{\begin{array}{rl} \mu + 2\lambda - 2\eta & = -1 \\ -\mu + 3\lambda + 3\eta & = -1 \\ -\mu + \lambda - 5\eta & = 3 \end{array}\right.$$

La solución del sistema es

$$\mu = -\frac{62}{38}, \quad \lambda = \frac{30}{38}, \quad \eta = -\frac{10}{38},$$

tal como podrá verificar el lector.

Para hallar las ecuaciones paramétricas de la normal común usamos como punto base, por ejemplo,

$$P' = (1,0,2) + \frac{30}{38}(2,-3,5) = \left(\frac{98}{38}\,, -\frac{90}{38}\,, \frac{226}{38}\right).$$

A partir de P' y del vector $U \wedge V$, encontramos las ecuaciones paramétricas

$$\left\{\begin{array}{l} x = 98/38 + 2\mu, \\ y = -90/38 - 3\mu. \\ z = 226/38 + 5\mu, \end{array}\right.$$

para la normal común. ♣

El procedimiento que acabamos de describir no es el único que permite hallar los puntos P' y Q' y la normal común n. De esto trata nuestro próximo ejercicio.

Ejercicio 3.102. Mostrar que el punto Q' puede hallarse trazando el plano que contiene a la recta r y es paralelo a la dirección de $U \wedge V$, y cortándolo con s. Aplicar esta construcción a las rectas r y s del ejemplo anterior.

Los puntos P' y Q' en que la normal común a r y s interseca a estas rectas tienen la propiedad de que minimizan la distancia entre un punto P'' en r y un punto Q'' en s, tal como mostraremos a continuación. Escribamos la diferencia entre P'' y Q'' en la forma

$$P''-Q'' = P''-P'+P'-Q'+Q'-Q'' = (P'-Q')+(P''-P'+Q'-Q''). \quad (3.59)$$

El término $P' - Q'$ es perpendicular a ambas rectas, en tanto que $P'' - P'$ y $Q' - Q''$ tienen las direcciones de las rectas r y s respectivamente. Por lo tanto, los dos sumandos que aparecen agrupados en el término de la derecha de (3.59) son ortogonales. Aplicando el teorema de Pitágoras conseguimos

$$|P'' - Q''|^2 = |P' - Q'|^2 + |P'' - P' + Q' - Q''|^2 \geq |P' - Q'|^2.$$

La igualdad sólo se se obtiene cuando

$$P'' - P' + Q' - Q'' = O, \quad (3.60)$$

donde O es el vector nulo de $\mathbb{R}^3$. Esto muestra que P' y Q' minimizan la distancia entre puntos de las rectas r y s.

Observemos además que $P'' - P'$ y $Q'' - Q'$ están en las direcciones de U y V respectivamente, por lo que existe escalares α y β tales que

$$P'' - P' = \alpha U, \quad Q'' - Q' = \beta V.$$

La igualdad (3.60) se traduce entonces en

$$\alpha U + \beta V = O.$$

Como la pareja (U, V) es linealmente independiente esta última igualdad sólo puede satisfacerse si $\alpha = \beta = 0$. Lo que implica

$$P'' - P' = Q'' - Q' = O.$$

Por lo tanto

$$P'' = P', \quad Q'' = Q',$$

y concluimos que los puntos P' y Q' son los únicos que hacen mínima la distancia entre puntos de r y s.

Es interesante notar que no hace falta hallar los puntos P' y Q' para determinar la distancia entre las rectas r y s. Sólo es necesario proyectar la diferencia $Q' - P'$ en la dirección del vector $U \wedge V$. Como $Q' - P'$ y $U \wedge V$ son colineales

$$(Q' - P') \cdot (U \wedge V) = \pm |Q' - P'||U \wedge V|, \tag{3.61}$$

ya que el ángulo entre ambos vectores es nulo o igual a π, entonces su coseno puede tomar los valores 1 o -1. Usando las expresiones (3.56) y (3.57) encontramos que

$$Q' - P' = Q - P + \mu V - \lambda U,$$

entonces

$$(Q' - P') \cdot (U \wedge V) = (Q - P + \mu V - \lambda U) \cdot (U \wedge V).$$

Desarrollando resulta

$$(Q' - P') \cdot (U \wedge V) = (Q - P) \cdot (U \wedge V) + \mu V \cdot (U \wedge V) - \lambda U \cdot (U \wedge V).$$

Claro está que los dos últimos sumandos son nulos, de modo que

$$(Q' - P') \cdot (U \wedge V) = (Q - P) \cdot (U \wedge V) = (Q - P, U, V) \tag{3.62}$$

y no hace falta determinar Q' y P' para hallar este producto escalar. Combinando (3.61) con (3.62) encontramos la fórmula

$$d(r,s) = |Q' - P'| = \frac{|(Q-P,U,V)|}{|U \wedge V|}, \tag{3.63}$$

que permite calcular la distancia entre las rectas r y s a partir del conocimiento de sus vectores directores U y V, y de puntos P y Q en las rectas.

Ejemplo 3.4.5. Aplicaremos la fórmula (3.63) para calcular la distancia entre las rectas r y s del ejemplo 3.4.4. Recordemos que

$$(Q-P,U,V) = (Q-P) \cdot U \wedge V.$$

En este caso particular, este producto triple toma el valor

$$|(1,1,-3) \cdot (2,-3,5)| = 16.$$

Dividiendo entre $|U \wedge V| = \sqrt{38}$, obtenemos el valor

$$d(r,s) = 16/\sqrt{38}$$

para la distancia entre las rectas r y s. ♣

Ejercicio 3.103. Encontrar la normal común a las rectas r y s de ecuaciones

$$\left\{\begin{array}{l} x = 4 - 4\lambda, \\ y = 5 + 4\lambda, \\ z = -2 + \lambda, \end{array}\right. \qquad \left\{\begin{array}{l} x = 2 + 2\lambda, \\ y = 5 + 2\lambda, \\ z = 1 - \lambda. \end{array}\right.$$

Calcular la distancia entre ambas rectas, y hallar puntos P y Q en r y s respectivamente tales que la distancia entre P y Q sea igual a la distancia entre r y s.

Por supuesto, rectas y planos en el espacio no tienen por qué ser paralelos o perpendiculares. Pueden cortarse formando otros ángulos. El ángulo entre dos rectas que se cortan es el menor ángulo posible entre sus vectores directores. En realidad hay sólo dos posibilidades, un ángulo θ, y su complementario $\pi - \theta$. También podemos definir el ángulo entre dos planos como el menor ángulo posible entre sus normales. De todos modos no llevaremos más lejos esta discusión en el texto.

3.4.5. Ejercicios complementarios

Agregamos aquí algunos ejercicios para que el lector tenga material para trabajar un poco más sobre los conceptos considerados en las secciones anteriores.

Ejercicio 3.104.

1. Hallar las intersecciones dos a dos, y determinar las posiciones relativas de las siguientes rectas[9]:

$$\left\{\begin{array}{l} x-7y=-9, \\ y-z=9, \end{array}\right. \qquad \left\{\begin{array}{l} x+3y+z+2=0, \\ x-y-3z-2=0, \end{array}\right.$$

$$\left\{\begin{array}{l} x=5+2\lambda, \\ y=2-\lambda, \\ z=-7+\lambda, \end{array}\right. \qquad \left\{\begin{array}{l} x=2\lambda+3, \\ y=-\lambda-2, \\ z=\lambda+1. \end{array}\right.$$

2. Demostrar que las ecuaciones

$$\left\{\begin{array}{l} x=1+\lambda, \\ y=2, \\ z=3-2\lambda, \end{array}\right. \qquad \left\{\begin{array}{l} x=2-2\nu, \\ y=2, \\ z=1+4\nu, \end{array}\right.$$

representan la misma recta. Para cada valor de λ hallar el valor de ν que al ser sustituido en las segundas ecuaciones produce el mismo punto que λ al ser sustituido en las primeras. Repetir el cálculo, intercambiando los papeles de λ y ν. ¿Qué relación hay entre las funciones encontradas?

3. Hallar ecuaciones reducidas de la recta de la parte anterior.

4. Hallar ecuaciones paramétricas e implícitas de la recta que pasa por $(-1,2,0)$ y es paralela a la recta de ecuaciones

$$\left\{\begin{array}{l} x=3+\lambda, \\ y=-2\lambda, \\ z=1, \end{array}\right.$$

5. Repetir para la recta que pasa por el punto $(-1,2,0)$ y es paralela a la recta de ecuaciones

$$\left\{\begin{array}{l} 2x+y+z=2, \\ -x+y+z=5, \end{array}\right.$$

[9]Para ser precisos deberíamos referirnos a las rectas definidas por las ecuaciones que se presentan a continuación, pero recurriremos a la figura de referirnos a una recta a través de sus ecuaciones. Ver la nota 4, al pie de la página 293.

Ejercicio 3.105.

1. Hallar ecuaciones paramétricas y reducida del plano que pasa por el punto $(1,1,0)$ y es paralelo al plano

$$\begin{cases} x = 1+\alpha-\beta, \\ y = 1-\alpha+\beta, \\ z = \alpha+\beta. \end{cases}$$

2. Hallar ecuaciones paramétricas y reducida del plano que pasa por el punto $(5,2,0)$ y es paralelo al plano hallado en la parte anterior.

3. Hallar las intersecciones dos a dos, e investigar las posiciones relativas de los planos

$$-x+2y+4z = 17, \qquad \begin{cases} x = 1+4\beta, \\ y = 2-2\alpha, \\ z = -1+\alpha+\beta, \end{cases}$$

$$x+y+z = 3, \qquad \begin{cases} x = 5+2\alpha+4\beta, \\ y = 2-\alpha+2\beta, \\ z = \alpha. \end{cases}$$

Ejercicio 3.106.

1. Determinar si los siguientes pares de rectas son coplanares. Si lo son, hallar el plano que determinan.

$$\begin{cases} x = 1+\lambda, \\ y = 1, \\ z = 1-\lambda, \end{cases} \qquad \begin{cases} x = 3+\lambda, \\ y = -2\lambda, \\ z = 3+7\lambda; \end{cases}$$
$$\begin{cases} x = 1+\lambda, \\ y = 1, \\ z = 1, \end{cases} \qquad \begin{cases} x = 1, \\ y = 1-\lambda, \\ z = 1; \end{cases}$$
$$\begin{cases} x = 1+\lambda, \\ y = 2+\lambda, \\ z = 3+\lambda, \end{cases} \qquad \begin{cases} x = -\lambda, \\ y = -\lambda, \\ z = -\lambda. \end{cases}$$

2. Determinar si las rectas

$$\begin{cases} x = 1+\lambda, \\ y = 5+\lambda, \\ z = 5-\lambda, \end{cases} \qquad \begin{cases} x = 3+2\lambda, \\ y = -1-2\lambda, \\ z = 3+5\lambda, \end{cases} \qquad \begin{cases} x = 1-2\lambda, \\ y = -2+2\lambda, \\ z = 3, \end{cases}$$

son paralelas al plano

$$\begin{cases} x = 5+\alpha-\beta, \\ y = 7-\alpha+\beta, \\ z = \alpha+\beta. \end{cases}$$

3. Probar que la recta
$$\begin{cases} x = 0, \\ y = \lambda, \\ z = \lambda, \end{cases}$$
   a) está incluida en el plano $6x + 4y - 4z = 0$;
   b) es paralela al plano $5x + 2y - 2z = 1$;
   c) corta al plano $y + z = 10$. Hallar el punto de intersección.

Presentamos a continuación una serie de ejercicios que pueden resolverse utilizando una notación vectorial, que no requiere expresar explícitamente las coordenadas de puntos y vectores.

Ejercicio 3.107. Sean π_1 y π_2 dos planos de ecuaciones paramétricas

$$\begin{aligned} P &= P_1 + \lambda U_1 + \mu V_1, \\ P &= P_2 + \lambda U_2 + \mu V_2, \end{aligned}$$

dos planos. Demostrar que π_1 y π_2 son paralelos si y sólo si para todo par de puntos A_1 y B_1 de π_1, y todo punto A_2 de π_2, se cumple que el punto

$$B_2 = A_2 + \overrightarrow{A_1B_1} = B_2$$

pertenece a π_2. Sugerencia: puede ser útil recordar que estos dos planos expresados en forma paramétrica son paralelos si y sólo si U_1 y V_1 pueden expresarse como combinación lineal de U_2 y V_2, y viceversa.

Ejercicio 3.108. Sean A, B, C y D puntos no coplanares del espacio. Consideremos el plano π y la recta r de ecuaciones paramétricas

$$\begin{aligned} P = A + \lambda\overrightarrow{AB} + \mu\overrightarrow{AC} \\ P = D + \sigma\overrightarrow{DB} \end{aligned}$$

Determinar si cada una de las siguientes afirmaciones es verdadera o falsa. En caso de que las afirmaciones involucren un parámetro, determinar su validez en función del parámetro.

1. $A - 3/2\overrightarrow{AC} + 2\overrightarrow{AB} \in \pi$;
2. $C + \mu\overrightarrow{AB} \in \pi$;
3. $D + \sigma\overrightarrow{DA} \in \pi$;
4. La recta $B + \lambda/2\overrightarrow{BD} + \lambda/2\overrightarrow{BA}$ está contenida en π;
5. La recta r' interseca a π;
6. La recta r interseca a r';
7. La recta r es paralela a r'.

Ejercicio 3.109. Sea A un punto y V_1, V_2 y V_3 vectores no coplanares. Se considera el vector

$$W = 2V_1 + V_2 + V_3,$$

y los planos π_1 y π_2 de ecuaciones

$$\begin{aligned} P &= A + \lambda V_1 + \nu V_2, \\ P &= A + \lambda V_3 + \nu w. \end{aligned}$$

respectivamente.

1. Mostrar que:

   a) $A \in \pi_1 \cap \pi_2$;

   b) $Q = A + 2\sqrt{2}V_1 + \sqrt{2}V_2 \in \pi_1 \cap \pi_2$.

2. Hallar una ecuación de la recta $r = \pi_1 \cap \pi_2$.

3. Se consideran un vector cualquiera U no colineal con V_3 y el plano

   $$P = A + \lambda_3 V_3 + \lambda_4 u.$$

   Hallar una ecuación de la recta $r' = \pi_1 \cap \pi_3$.

Ejercicio 3.110. Se consideran las rectas

$$\begin{aligned} r) \quad & P = A + \lambda v, \\ r') \quad & P = B + \lambda w. \end{aligned}$$

Hallar una ecuación del lugar geométrico de los puntos medios de los segmentos PP', con $P \in r$ y $P' \in r'$ y analizar qué representa. Discutir según que v y w sean colineales o no lo sean.

3.4.6. Para tener presente

- Las relaciones de perpendicularidad y paralelismo pueden caracterizarse en términos de los vectores que definen las direcciones de rectas y planos.
- Muchas construcciones geométricas puede realizarse eficientemente reduciendo a los productos escalar y vectorial.
- Las proyecciones sobre subespacios se calculan rápidamente haciendo productos escalares con versores convenientemente elegidos. Trasladando esta construcción, pueden calcularse proyecciones sobre rectas y planos.

3.5. Nociones sobre curvas y superficies

Intuitivamente, una curva en el plano o el espacio es un objeto unidimensional, una superficie en el espacio un objeto bidimensional. Un objeto unidimensional como una curva se recorre variando un parámetro, tal como lo hacíamos con las ecuaciones paramétricas de las rectas. Para recorrer una superficie bidimensional hace falta dejar variar dos parámetros. Vimos esto en el caso de las ecuaciones paramétricas de los planos. Otro ejemplo nos es familiar: para ubicar un punto sobre la superficie de la tierra damos dos coordenas, la latitud y la longitud.

Un plano se describe por una única ecuación lineal que debe ser satisfecha por las coordenadas (x, y, z), su ecuación reducida. Cuando una ecuación que vincula a las variables (x, y, z) no es lineal define una relación entre las coordenadas de los puntos en el espacio, que nos confina a movernos sobre una superficie. Ya vimos esto en el caso de esferas y cilindros. Cuando intersecamos dos superficies conseguimos una curva. Así, una pareja de ecuaciones que deben ser satisfechas por los puntos de $\mathbb{R}^3$ definen, en general, una curva en el espacio.

El espacio $\mathbb{R}^3$ tiene dimensión 3. Hay 3 grados de libertad para moverse en el espacio. Una ecuación impone una restricción que deja sólo dos grados de libertad, dos ecuaciones dejan un único grado de libertad. Esta es otra manera de pensar por qué las superficies quedan definidas por una ecuación y las curvas por dos.

Antes de seguir avanzando digamos que la discusión que acabamos de dar es puramente intuitiva, en general adecuada, pero para que sea completamente cierta hay que agregar diversas condiciones de regularidad sobre las parametrizaciones y las ecuaciones con las que se pretende definir curvas y superficies. Sin estas condiciones bien puede ocurrir que una única ecuación defina una curva, o un punto; o que una curva continua que depende de un único parámetro llene todo el plano o todo el espacio. Ver el ejemplo 3.5.11 y el ejercicio 3.114. En nuestra presentación nos limitaremos a analizar algunos ejemplos, y a realizar algunas construcciones geométricas, pero no nos detendremos sobre estos problemas de regularidad. El lector interesado puede consultar [Do].

3.5.1. Curvas

Una curva es un objeto unidimensional que se recorre al dejar variar un parámetro real. El primer ejemplo que viene a la cabeza es el de las rectas y sus ecuaciones paramétricas.

Ejemplo 3.5.1. RECTAS
Una recta es un caso particular de curva. Si repasamos la definición de recta que pasa por un punto $P = (p_1, p_2, p_3)$ y es colineal con un vector dado $V = (v_1, v_2, v_3)$, encontramos que está formada por todos los puntos Q de la forma

$$P + \lambda V, \lambda \in \mathbb{R}.$$

La expresión $Q(\lambda) = P + \lambda V$ define una función

$$Q : \mathbb{R} \to \mathbb{R}^3,$$

cuyos valores son los puntos sobre la recta. Al variar λ recorremos los puntos de la recta.

La función Q admite una expresión en coordenadas, de la forma

$$Q(\lambda) = (x(\lambda), y(\lambda), z(\lambda)),$$

donde

$$\begin{cases} x(\lambda) = p_1 + \lambda v_1, \\ y(\lambda) = p_2 + \lambda v_2, \\ z(\lambda) = p_3 + \lambda v_3. \end{cases}$$

Es muy sugerente el punto de vista de que el parámetro λ representa el tiempo, cada punto $Q(\lambda)$ la posición en el espacio, en el instante $t = \lambda$ de un móvil, y el conjunto de todos los puntos $Q(\lambda)$ la trayectoria recorrida por el móvil.

Ya vimos también que una recta puede describirse como el conjunto de puntos (x, y, z) que satisfacen un par de ecuaciones lineales, cuya matriz es de rango 2. ♣

Ejemplo 3.5.2. UNA CIRCUNFERENCIA EN EL PLANO (x, y)
Los puntos del plano (x, y) que están a distancia r del origen $(0, 0)$ forman una circunferencia $\mathcal{C}$ de ecuación

$$x^2 + y^2 = r^2.$$

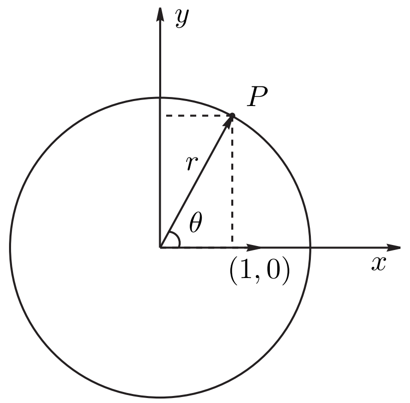
Esta circunferencia puede ser recorrida dejando variar el ángulo θ que el vector (x, y) forma con $(1, 0)$ tal como se muestra en la figura

En función de θ tenemos

$$x = r\cos\theta, \quad y = r \operatorname{sen}\theta.$$

Estas igualdades pueden ser interpretadas como una función

$$P : [0, 2\pi] \to \mathbb{R}^2$$



que a cada valor del ángulo $\theta \in [0, 2\pi]$ le asocia un punto

$$P(\theta) = (x(\theta), y(\theta)) \in \mathbb{R}^2,$$

donde, por supuesto,

$$x(\theta) = r\cos\theta, \quad y(\theta) = r\,\mathrm{sen}\,\theta.$$

La función P es una *parametrización* de la circunferencia $\mathcal{C}$. Su imagen es justamente $\mathcal{C}$.

Naturalmente P no es la única parametrización de $\mathcal{C}$. Por ejemplo, se puede parametrizar la parte de $\mathcal{C}$ que está en el semiplano $y > 0$ por

$$\begin{cases} x = x, \\ y = \sqrt{x^2 + y^2}, \end{cases} \qquad x \in (-r, r).$$

Otras parametrizaciones son posibles. Cada una corresponde a una manera de recorrer la circunferencia, o parte de ella.

Ejemplo 3.5.3. Una circunferencia en el espacio (x, y, z)
Podemos "sumergir" la construcción del ejemplo anterior en el espacio, identificando un punto $(x, y) \in \mathbb{R}^2$ con el punto $(x, y, 0) \in \mathbb{R}^3$, de modo que el plano (x, y) se identifica con el plano de ecuación $z = 0$ en $\mathbb{R}^3$.

La circunferencia $\mathcal{C}$ formada por los puntos del plano $z = 0$ que están a distancia r del origen O queda caracterizada por el par de ecuaciones

$$\begin{cases} x^2 + y^2 = r^2, \\ z = 0, \end{cases}$$

que son satisfechas por todos los puntos (x, y, z) que pertenecen a $= \mathcal{C}$. Observemos que la ecuación $x^2 + y^2 = r^2$ representa a un cilindro de eje O_z y radio r, la ecuación $z = 0$ un plano. El sistema formado por las dos ecuaciones representa la intersección de las dos superficies.

La circunferencia admite otras representaciones con otros pares de ecuaciones, por ejemplo

$$\left\{ \begin{array}{c} x^2 + y^2 + z^2 = r^2, \\ z = 0, \end{array} \right.$$

también representa a $\mathcal{C}$, en este caso como la intersección de la esfera de centro O y radio r con el plano $z = 0$.

También podemos llevar al espacio la parametrización de la circunferencia, como

$$P(\theta) = \left\{ \begin{array}{l} x(\theta) = r\cos\theta, \\ y(\theta) = r\operatorname{sen}\theta, \\ z(\theta) = 0, \end{array} \right. \qquad \theta \in [0, 2\pi]. \tag{3.64}$$

Recorremos los mismos valores de x e y que en el caso plano, pero hemos agregado la especificación de que $z = 0$. ♣

Ejercicio 3.111. ¿Qué ocurre cuando en la parametrización (3.64) se deja variar θ en $(-\infty, +\infty)$?.

Imaginar y dibujar la curva que se obtiene por medio de la parametrización

$$P(\theta) = \left\{ \begin{array}{l} x(\theta) = r\cos\theta, \\ y(\theta) = r\operatorname{sen}\theta, \\ z(\theta) = \theta, \end{array} \right. \qquad \theta \in (-\infty, +\infty).$$

Ejemplo 3.5.4. Parametrización de una elipse Es posible “deformar”la parametrización de una circunferencia para obtener una elipse. Consideremos dos números positivos a y b. Si definimos

$$P : [0, 2\pi] \to \mathbb{R}^2$$

por

$$P(t) = (a\cos t, b\operatorname{sen} t), \tag{3.65}$$

entonces la imagen de esta función es una elipse en el plano, tal como se muestra en la figura 3.5, para el caso $a = 2$, $b = 5$. Escribamos cada punto $P(t)$ en coordenadas, como $(x(t), y(t))$. Entonces

$$x(t)/a = \cos t, \quad y(t)/b = \operatorname{sen} t,$$

por lo que

$$\frac{x^2(t)}{a^2} + \frac{y^2(t)}{b^2} = \cos^2 t + \operatorname{sen}^2 t = 1,$$

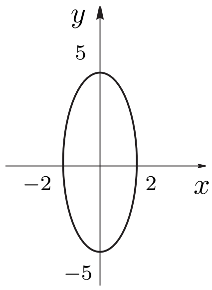


Figura 3.5: Una elipse en el plano.

lo que muestra que los puntos $P(t)$ están sobre la elipse de ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Recíprocamente, si un punto (x, y) satisface esta ecuación entonces $(x/a, y/b)$ es de la forma $(\cos t, \operatorname{sen} t)$ para algún valor de $t \in [0, 2\pi)$. Por lo tanto (x, y) está en la imagen de la parametrización (3.65) ♣

Ejercicio 3.112. Número de condición (V)

1. Sea A una matriz real 2×2. Mostrar que los números

$$\begin{aligned} m &= \text{ínf}\left\{\tfrac{|AX|}{|X|};\ X \in \mathbb{R}^2,\ X \neq O\right\}, \\ M &= \sup\left\{\tfrac{|AX|}{|X|};\ X \in \mathbb{R}^2,\ X \neq O\right\}, \end{aligned}$$

que definen la mínima y la máxima dilatación que produce la matriz A satisfacen

$$\begin{aligned} m &= \text{ínf}\left\{|AX|;\ X \in \mathbb{R}^2,\ |X| = 1\right\}, \\ M &= \sup\left\{|AX|;\ X \in \mathbb{R}^2,\ |X| = 1\right\}. \end{aligned}$$

2. Consideramos la matriz

$$A = \begin{pmatrix} 1 & 0{,}999 \\ 0{,}999 & 1 \end{pmatrix}$$

y la base ortonormal

$$(e_1, e_2) = \left(\left(1/\sqrt{2}, 1/\sqrt{2}\right), \left(-1/\sqrt{2}, 1/\sqrt{2}\right)\right)$$

de $\mathbb{R}^2$. Recordemos que todo $X \in \mathbb{R}^2$ admite una expresión de la forma

$$X = \alpha_1 e_1 + \alpha_2 e_2.$$

Mostrar que $|X| = 1$ si y sólo si $\alpha_1^2 + \alpha_2^2 = 1$. Mostrar que dado un vector

$$Y = \alpha_1 e_1 + \alpha_2 e_2 \in \mathbb{R}^2$$

existe un vector $X \in \mathbb{R}^2$ de módulo 1 tal que $AX = Y$ si y sólo si

$$\frac{\beta_1^2}{1{,}999^2} + \frac{\beta_2^2}{0{,}001^2} = 1.$$

3. Interpretar geoemétricamente los resultados obtenidos.

La primera parte del ejercicio 3.112 nos muestra que para calcular la máxima y la mínima dilatación que produce una matriz A es suficiente examinar su efecto sobre los vectores de módulo 1. En el caso de $\mathbb{R}^2$ estos vectores forman una circunferencia de radio 1 y centro O. El número de condición aparece entonces como una medida de cuánto se deforma esta circunferencia al aplicar A.

Ejemplo 3.5.5. La función

$$P : \mathbb{R} \to \mathbb{R}^3$$

dada por

$$P(t) = (t, t^2, t^3)$$

define una curva en el espacio. En la figura 3.6 mostramos la parte de esta curva que se obtiene cuando t varía en $[-2, 2]$. La imagen de P queda caracterizada por las ecuaciones

$$\left\{ \begin{array}{l} y - x^2 = 0, \\ z - x^3 = 0, \end{array} \right. \tag{3.66}$$

una afirmación cuya verificación dejamos para el lector.

Ejercicio 3.113. Mostrar que un punto (x, y, z) satisface (3.66) si y sólo si es de la forma (t, t^2, t^3) para algún valor de t. ♣

A pesar de que en una primera aproximación podríamos pensar lo contrario, la imagen de una función continua

$$P : I \to \mathbb{R}^2,$$

donde I es un intervalo, no tiene por qué parecerse en nada a algo que merezca llamarse "curva". Veremos este fenómeno en el ejercicio 3.114, en el que construiremos una función P continua, definida sobre $[0, 1]$, cuya imagen cubre completamente el cuadrado

$$[0, 1] \times [0, 1] \subset \mathbb{R}^2.$$

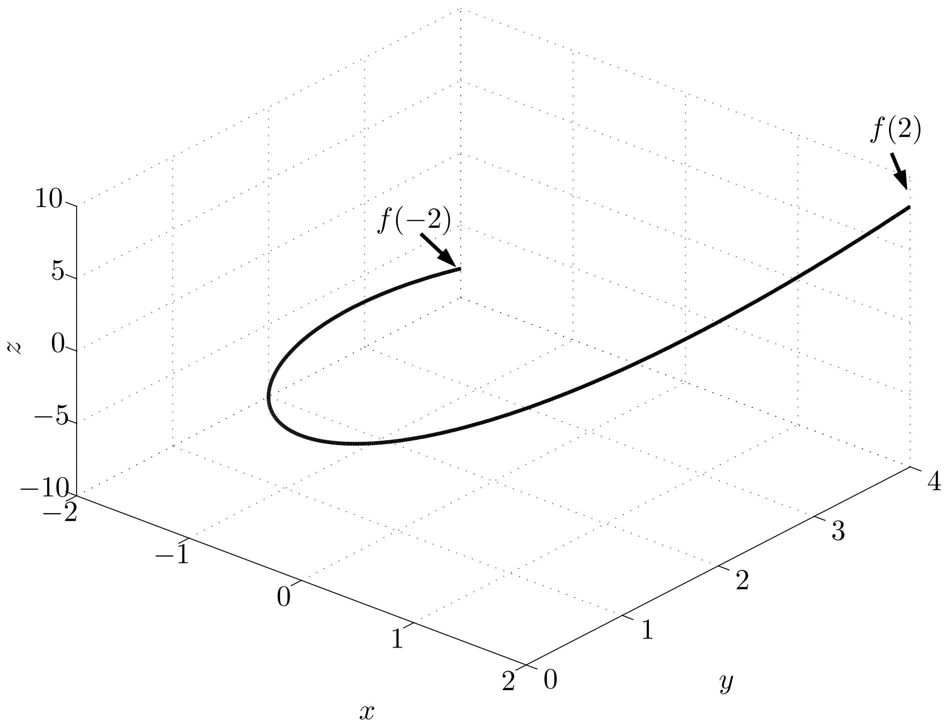


Figura 3.6: Curva en el espacio.

Para ello nos hace falta recordar algunas propiedades de los números reales del intervalo $[0,1]$.

Cualquier número real $t \in [0,1]$ puede representarse como un decimal infinito

$$t = 0{,}t_1t_2t_3t_4t_5\ldots,$$

donde los números t_i, $i = 1,2,\ldots$, toman valores en los dígitos $\{0,1,\ldots,9\}$. Algunos números tienen dos posibles representantes: uno en el que todos los dígitos se hacen 9 a partir de cierto lugar, y otra en el que todos los dígitos son iguales a 9 a partir del lugar siguiente. Convengamos que para estos números tomamos la representación que termina en los 9. Entonces cada $t \in [0,1]$ está asociada a un único desarrollo decimal.

Ejercicio 3.114. Una "curva" continua que llena un cuadrado
Vamos a definir una función $P : [0,1] \to [0,1] \times [0,1]$, de la forma $P(t) = (x(t), y(t))$ por las fórmulas

$$x(0{,}t_1t_2t_3t_4t_5\ldots) = 0,t_1t_3t_5\ldots, \quad y(0{,}t_1t_2t_3t_4t_5\ldots) = 0,t_2t_4t_6\ldots.$$

Mostrar que x e y son funciones continuas de t, y que P es sobreyectiva. ¿Es P inyectiva?

Una de las condiciones de regularidad que debe satisfacer un mapa $P(t)$ definido sobre un intervalo y que toma valores en $\mathbb{R}^3$ para definir realmente una curva es que sus componentes $x(t)$, $y(t)$ y $z(t)$ sean funciones derivables. Entonces $(x'(t), y'(t), z'(t))$ representa un vector tangente a la curva en el punto $P(t) = (x(t), y(t), z(t))$. Lo mismo vale, eliminando la tercera componente, para el caso de curvas planas. No continuaremos analizando estas cuestiones de regularidad, pero proponemos al lector un ejercicio para visualizar curvas y tangentes.

Ejercicio 3.115. Calcular las derivadas de las componentes de todas las parametrizaciones de curvas que aparecen en esta sección. Interpretar geométricamente los resultados obtenidos. ¿Están de acuerdo con tu intuición de cuáles deberían ser los vectores tangentes?.

Hemos visto entonces que una curva en el espacio puede conseguirse "deformando"un intervalo $I \in \mathbb{R}$ y "colocándolo"en $\mathbb{R}^3$, siguiendo una cierta "receta"dada por una parametrización P, que es una función

$$P : I \to \mathbb{R}^3,$$

que envía cada $t \in I$ a

$$P(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3.$$

Para cada valor del parámetro t se obtiene un punto $P(t)$ que está en la curva. Cuando la función toma valores en $\mathbb{R}^2$ obtenemos una curva plana.

También hemos visto que una curva en el espacio puede caracterizarse por un par de ecuaciones reducidas

$$\begin{cases} f(x, y, z) = 0, \\ g(x, y, z) = 0. \end{cases}$$

En general, cada una de estas ecuaciones representa una superficie. La curva aparece entonces como la intersección de dos superficies. Para especificar una curva en el plano basta una ecuación. Una curva plana puede pensarse como un conjunto de puntos en el espacio que satisface la condición adicional $z = 0$. En la próxima sección mostramos algunos ejemplos de superficies. También que la afirmación de que una ecuación describe una superficies debe tomarse con cierto cuidado. Ver el ejemplo 3.5.11.

Ejercicio 3.116.

1. Para $A = (1, 0, 0)$ y $B = (0, 1, 0)$, hallar la ecuación del lugar geométrico $\mathcal{L}$ de los puntos P del espacio tales que $P - A$ y $P - B$ forman un ángulo recto;
2. Probar que $\mathcal{L}$ es una esfera. Hallar su centro y radio.
3. Encontrar las ecuaciones de la circunferencia $\mathcal{C}$, intersección de la esfera $\mathcal{L}$ con el plano $x = 2y$.
4. Hallar el centro C y el radio r de la circunferencia $\mathcal{C}$.

Ejercicio 3.117.

1. Hallar una parametrización de la elipse que es la imagen de la circunferencia de centro $(0, 0)$ y radio 1 por la transformación $(x, y) \mapsto (3x, 2y)$.
2. Un avión vuela con velocidad $(v, 0, 0)$, donde suponemos v expresado en metros por segundo, y su hélice, con aspas de longitud igual a l metros, gira a r revoluciones por segundo. Hallar el punto en que está un extremo de su hélice en el instante t, sabiendo que en el instante $t = 0$ estaba en $(l, 0, 0)$ (en realidad hay dos soluciones posibles. Encontrar ambas).
3. Dar una parametrización de la circunferencia construida en la parte 3 del ejercicio 3.116.
4. Parametrizar la trayectoria de una partícula fijada a un disco de radio r que rueda sin deslizar en el plano (x, y), apoyado sobre la recta $y = 0$. Si x es la abcisa del centro del disco y θ el ángulo que el disco ha girado (medido en sentido directo), la condición de rodadura sin deslizar puede expresarse en la forma $x + r\theta = 0$.

3.5.2. Superficies

Así como una curva es un objeto unidimensional dentro de un espacio ambiente de dos o tres dimensiones (según sea una curva plana o en el espacio), podemos pensar a una superficie en el espacio como un objeto bidimensional.

Podemos caracterizar una superficie por una ecuación, como ya hemos hecho con los planos, esferas y cilindros. Nos interesa ahora enfatizar el punto de vista de que una superficie puede verse como el resultado de tomar un trozo del plano, "deformarlo"y "colocarlo"en $\mathbb{R}^3$. De manera similar a lo que ocurría para las curvas, estas operaciones se describen por medio de una parametrización de la superficie. Se trata de una función

$$P : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3,$$

donde Ω es algún subconjunto de $\mathbb{R}^2$ en el que varían los parámetros (u, v) de la parametrización. Es corriente tomar como Ω un rectángulo finito o infinito en el plano, pero esto no es esencial. La parametrización P asocia a cada punto $(u,v) \in \Omega$ un punto

$$P(u,v) = (x(u,v), y(u,v), z(u,v)) \in \mathbb{R}^3.$$

La forma de la superficie dependerá del dominio de la parametrización y de cuáles sean las funciones $x(u,v)$, $y(u,v)$ y $z(u,v)$. Por supuesto, más de una parametrización P puede dar lugar a la misma superficie, algo que estudiaremos con cierto detalle en el caso de las esferas.

Ejemplo 3.5.6. PARAMETRIZACIÓN DE UN PLANO
Un plano es un caso particular de superficie. El plano que pasa por el punto $P = (p_1, p_2, p_3)$ y tiene a $U = (u_1, u_2, u_3)$ y $V = (v_1, v_2, v_3)$ como vectores directores admite una parametrización

$$Q : \mathbb{R}^2 \to \mathbb{R}^3$$

definida por

$$Q(\lambda, \mu) = P + \lambda U + \mu V.$$

Esta parametrización, cuando se explicitan las coordenadas, es

$$\begin{cases} x(\lambda,\mu) = p_1 + \lambda u_1 + \mu v_1, \\ y(\lambda,\mu) = p_2 + \lambda u_2 + \mu v_2, \\ z(\lambda,\mu) = p_3 + \lambda u_3 + \mu v_3. \end{cases}$$

Se trata simplemente de las ecuaciones paramétricas del plano, escritas de una manera que enfatiza que las coordenadas x, y y z de un punto del plano son funciones de los parámetros λ y μ. ♣

Ejercicio 3.118. PARAMETRIZACIONES DE PLANOS.

1. COORDENADAS POLARES. En muchos problemas es útil describir la posición de un punto (x,y) en el plano por sus **coordenadas polares** r y θ, definidas de forma tal que
$$x = r\cos\theta, \quad y = r\,\mathrm{sen}\,\theta.$$
Mostrar que estas fórmulas definen una parametrización [10] de todo el plano (x,y) excepto la semirrecta $y = 0$, $x \le 0$, desde el "rectángulo"$0 < r < +\infty$, $-\pi/2 < \theta < \pi/2$ del plano (r,θ).

[10] En el sentido de que a cada punto (x,y) de una región le corresponde un único punto (r,θ) de la otra.

2. Hallar la condición que deben satisfacer u y v para que al parametrizar un plano π en la forma
$$P + \lambda\vec{u} + \nu\vec{v}$$
la imagen en el plano π de cualquier segmento del plano (λ, ν) tenga la misma longitud que el segmento original.

Ejemplo 3.5.7. Consideramos ahora la superficie definida por $P : \mathbb{R}^2 \to \mathbb{R}^3$ dada por
$$P(u, v) = (u, v, u^2 - v^2).$$
Una parte de esta superficie se muestra en la figura 3.7. Es facil observar que

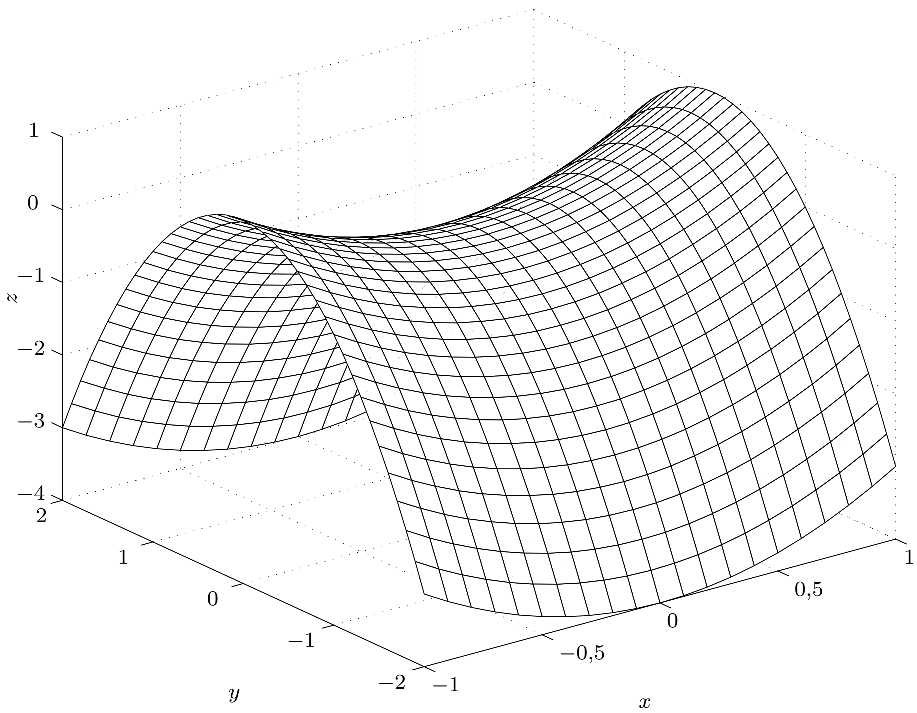


Figura 3.7: Superficie en el espacio.

la superficie está definida por la ecuación $z = x^2 - y^2$.

Ejemplo 3.5.8. Latitud y longitud: una parametrización de la esfera con coordenadas geográficas
Un punto sobre la Tierra queda completamente determinado al dar su latitud ψ y su longitud θ. Este hecho puede traducirse en un conjunto de fórmulas

que dan lugar a una parametrización de la esfera S de centro O y radio 1, cuya ecuación es

$$x^2 + y^2 + z^2 = 1.$$

Dado un punto $P = (x, y, z)$ en la esfera mediremos el ángulo ψ que el vector OP forma con el plano O_{xy}. Este ángulo toma valores entre $-\pi/2$ (en el "Polo Sur"$(0, 0, -1)$) y $\pi/2$ (en el "Polo Norte"$(0, 0, 1)$). La coordenada z es igual a $\operatorname{sen}\psi$, en tanto que la proyección OP' del vector OP sobre el plano O_{xy} tiene una longitud igual a $\cos\psi$.

Mediremos el ángulo θ que esta proyección forma con el semieje positivo del eje O_x (el meridiano de Greenwich y el Ecuador se cortarían en $(1, 0, 0)$). Al proyectar OP' sobre los ejes O_x y O_y encontramos que

$$x = \cos\theta\cos\psi, \quad y = \operatorname{sen}\theta\cos\psi.$$

En resumen, el punto P puede escribirse como $P(\theta, \psi)$ en la forma

$$P(\theta, \psi) = (\cos\theta\cos\psi, \operatorname{sen}\theta\cos\psi, \operatorname{sen}\psi).$$

Este mapa

$$P : [-\pi/2, \pi/2] \times [-\pi, \pi] \to \mathbb{R}^3,$$

tiene como imagen toda la esfera S. Para todos los puntos P de la esfera, salvo un pequeño conjunto excepcional, hay un único valor del par (θ, ψ) tal que $P = P(\theta, \psi)$.

Ejercicio 3.119. Identificar sobre S el conjunto excepcional, y los valores de (θ, ψ) que le corresponden.

Ejercicio 3.120. Dados $r > 0$ y $Q \in \mathbb{R}^3$, construir a partir de la parametrización P de S una parametrización de la esfera de centro Q y radio r. ♣

Ejercicio 3.121. CONOS.

Se llama "cono"o "superficie cónica"de vértice V y directriz $\mathcal{C}$ (donde V es un punto y $\mathcal{C}$ es una curva dada tales que $V \notin \mathcal{C}$) a la superficie engendrada por todas las rectas que pasan por V y cortan a $\mathcal{C}$.

1. Hallar la ecuación del cono $\mathcal{S}_1$ que tiene como vértice al punto $(0, 0, 0)$ y como directriz a la curva de ecuaciones $x^2 - 2z + 1 = 0$, $y = 1$.
2. Hallar la ecuación del cono de vértice $(2, 1, 4)$ y directriz la curva de ecuaciones $x^2 + y^2 + z^2 = 4$, $x = 1$.
3. El eje Oz es el eje de un cono circular que contiene al punto $(3, -4, 7)$ y tiene vértice situado en el origen de coordenadas. Hallar la ecuación de este cono.
4. Hallar la ecuación del cono circular con eje Oy, vértice en el origen de coordenadas, y tal que sus generatrices forman un ángulo de $\pi/3$ con el eje.

5. Hallar la ecuación del cono circular $\mathcal{S}$ sabiendo que su vértice es el origen de coordenadas y que sus generatrices son tangentes a la esfera definida por la ecuación $(x+2)^2+(y-1)^2+(z-3)^2=9$.

Ejercicio 3.122.

1. Hallar la ecuación del cono $\mathcal{S}$ con vértice $(-1,1,1)$ y directriz
$$\begin{cases} x^2+y^2=1, \\ x=z. \end{cases}$$

2. Hallar las ecuaciones de la curva $\mathcal{C}$ que se obtiene cortando el cono $\mathcal{S}$ con el plano $z=x+1$.

3. Hallar las ecuaciones de la proyección ortogonal de la curva $\mathcal{C}$ sobre el plano de ecuación $x+y+2z=0$.

4. Hallar las ecuaciones de la proyección ortogonal de la curva $\mathcal{C}$ sobre $z=0$.

Ejercicio 3.123. Cilindros.

Se llama "cilindro"o "superficie cilíndrica"de directriz $\mathcal{C}$ y generatriz colineal al vector $V=(a,b,c)$ a la superficie engendrada por todas las rectas colineales al vector V que cortan a la curva $\mathcal{C}$.
Hallar las ecuaciones de los cilindros que se especifican:

1. el que tiene directriz de ecuaciones $x^2+y^2=9$, $z=1$, y generatrices paralelas al vector $(2,-3,4)$;

2. con directriz
$$\begin{cases} x^2-y^2=z, \\ x+y+z=0, \end{cases}$$
y generatrices perpendiculares al plano de la directriz.

Ejemplo 3.5.9. Parametrización de un cilindro
El cilindro $C\subset\mathbb{R}^3$ de ecuación
$$x^2+y^2=1$$
puede verse como el conjunto de rectas verticales que se apoya en la circunferencia de ecuaciones
$$\begin{cases} x^2+y^2=1, \\ z=0. \end{cases}$$
A partir de una parametrización de la circunferencia, como la que se obtiene, por ejemplo, tomando $r=1$ en (3.64), podemos agregar un segundo parámetro

v para recorrer cada recta vertical a partir de cada punto en la circunferencia. Encontramos entonces que

$$P(\theta, v) = \begin{cases} x(\theta) = \cos\theta, \\ y(\theta) = \operatorname{sen}\theta, \\ z(\theta) = v, \end{cases} \qquad (\theta, v) \in [0, 2\pi] \times \mathbb{R}. \tag{3.67}$$

es una parametrización del cilindro. ♣

Observación 3.5.10. La parametrización (3.67) del cilindro tiene la propiedad de que no modifica las longitudes de las curvas. Esto es fácil de verificar para los segmentos de rectas horizontales o verticales del plano (θ, v). Los segmentos horizontales se transforman en arcos de circunferencia sobre el cilindro, que tienen exactamente la misma longitud que el segemento original. Los segmentos horizontales se transforman en segmentos verticales, sin sufrir deformación alguna. Es cierto en general, aunque más difícil de verificar[11], que para cualquier curva en el plano (θ, v) su imagen sobre el cilindro por la parametrización (3.67) tiene la misma longitud que la curva original.

La deformación que la parametrización (3.67) hace del rectángulo infinito $[0, 2\pi] \times \mathbb{R}$ del plano (θ, v) corresponde a la acción de arrollar una hoja de papel de ancho 2π para fabricar un cilindro de radio 1: el papel no sufre ningún estiramiento al realizar esta operación.

Es imposible construir una esfera a partir de una hoja de papel, que es rígida. Para fabricar una esfera con una porción de un material plano es necesario estirarlo, introducir deformaciones que cambian las longitudes de las curvas. Esta afirmación es equivalente a la siguiente: es imposible dibujar un mapa plano de la Tierra en el que todos los puntos del mapa estén a la misma distancia que sus correspondientes sobre la superficie terrestre. Para hacer un mapa plano de la Tierra hay que deformar su imagen. Este importantísimo resultado geométrico, que tiene una consecuencia notable para la cartografía, fue demostrado por Gauss, y se conoce con el nombre de *teorema egregio*. ♣

No siempre una ecuación representa una superficie. Mostramos a continuación un ejemplo.

Ejemplo 3.5.11. Cuál es el conjunto formado por los puntos $(x, y, z) \in \mathbb{R}^3$ cuyas coordenadas satisfacen la ecuación

$$x^2 + y^2 = 0?$$

[11]Además, para hacer esta verificación deberíamos definir previamente la noción de *longitud de una curva*.

Esta ecuación es equivalente al par de ecuaciones

$$\begin{cases} x = 0, \\ y = 0, \end{cases}$$

que definen ¡una recta! El eje O_z.

Todavía más, la ecuación

$$x^2 + y^2 + z^2 = a,$$

define la esfera de centro O y radio $\sqrt{a}$ cuando $a > 0$. Pero sólo es satisfecha por el origen O cuando $a = 0$, y no tiene ninguna solución si $a < 0$.

Una ecuación debe satisfacer ciertas condiciones de regularidad para definir una superficie. Estas condiciones son esencialmente equivalentes a que podamos definir un vector normal a la superficie, en el mismo sentido que el vector normal de los planos, en cada punto de $\mathbb{R}^3$ que satisface la ecuación. No discutiremos estos detalles, que el lector interesado puede consultar, por ejemplo, en [Do]. ♣

Ejercicio 3.124. Una perspectiva menos convencional.

En la litografía *Arriba y Abajo* (1947), también en *Cubo de Escalera* (1951), el artista holandés M.C. Escher (1898-1972)[12]. emplea magistralmente una representación de la perspectiva diferente de la que describimos en el ejercicio 9 del práctico 4. En ella, se supone al observador ubicado en el origen $O(0,0,0)$, y para cada punto P se calcula la intersección de la recta OP con el semicilindro

$$\mathcal{C} = \{(x, y, z);\ y > 0,\ y^2 + z^2 = 1\}.$$

Luego este semicilindro se desarrolla sobre un plano la región del plano (x, θ) definida por $-\infty < x < +\infty$, $-\pi/2 < \theta < \pi/2$.

1. Calcular el punto (x, θ) sobre el que se proyecta cada punto (X, Y, Z) del espacio.
2. Hallar las curvas que corresponden a la representación en esta perspectiva de las rectas paralelas a cada uno de los ejes coordenados. ¿Hacia donde fugan las rectas verticales y las rectas paralelas a $(1, 0, 0)$? Sugerencia: escribir x como función de θ para hallar una representación de estas curvas. Esta parte puede hacerse con argumentos analíticos, basados en las fórmulas que definen la perspectiva, o con argumentos de la geometría métrica. Sugerimos buscar ambos caminos.
3. Analizar a la luz de los resultados obtenidos la litografía *Arriba y Abajo*.

[12]Información sobre la vida y obra de Escher se encuentra en, por ejemplo, el sitio web `http://www.mcescher.com/`. Otra posible referencia es *El espejo mágico de M.C. Escher*, Bruno Ernst, Taschen, 1994.

Ejercicio 3.125. Parametrizaciones del cono.

1. Mostrar que al parametrizar un cono como el gráfico de $z = \sqrt{x^2 + y^2}$, es decir, como
$$P(x, y) = (x, y, \sqrt{x^2 + y^2}, \quad (x, y) \in \mathbf{R}^2,$$
la imagen de cualquier circunferencia centrada en el origen tiene la misma longitud que la circunferencia original, pero que la imagen de cualquier segmento que no contenga al origen del plano (x, y), pero esté contenido en una recta que pase por el origen, tiene una longitud mayor que la del segmento original.
2. Gorritos de cumpleaños. Construir la parametrización de la parte del gráfico de $z = \sqrt{x^2 + y^2}$ comprendida en $0 < z < 1$ que corresponde a "plegarün trozo de papel descrito, en coordenadas polares (ver 1 en este mismo ejercicio) por las desigualdades
$$0 < r < \sqrt{2}, \quad -\pi/\sqrt{2} < \theta < \pi\sqrt{2}.$$
En general, hallar el pedazo de papel y la parametrización correspondientes a esta construcción para la región comprendida en $0 < z < 1$ del cono circular recto con vértice en el origen $(0, 0, 0)$, eje en el eje z, y abertura α.

Ejercicio 3.126. Parametrizaciones de la esfera.
En este ejercicio mostramos cuatro maneras de parametrizar partes de la esfera
$$\mathcal{S} = \left\{(x, y, z);\ x^2 + z^2 + y^2 = 1\right\}.$$

1. Coordenadas geográficas: latitud y longitud. Considerar un ángulo θ que mida la **latitud** de cada punto (x, y, z) sobre la esfera. Se tomará como origen de latitud el "ecuador" consistente en la intersección de la esfera con el plano $z = 0$. Introducir un segundo ángulo φ que mida la **longitud** desde la intersección del semiplano $y = 0$, $x > 0$, con la esfera (¿meridiano de Greenwich?). Determinar las coordenadas (x, y, z) de cada punto sobre la esfera en función de su latitud y longitud (θ, φ).
2. Cartografía: la proyección de Mercator. Para fabricar un mapa de la tierra se recurre a la siguiente construcción: para un punto (x, y, z) sobre la esfera se traza desde el origen $(0, 0, 0)$ la semirrecta que contiene a este punto, y luego se busca la intersección de la semirrecta con el cilindro $x^2 + y^2 = 1$. Este cilindro "se corta" por la línea $x = -1$, $y = 0$, y se desenrolla sobre la franja $-\pi < \varphi < \pi$ del plano (φ, η).
   a) Hallar el punto (φ, η) que esta construcción asocia a cada punto (x, y, z) sobre la esfera. Recíprocamente, hallar el correspondiente sobre la esfera de cada punto (φ, η) con $\varphi \in (-\pi, \pi)$ y $\eta \in \mathbb{R}$.
3. Proyección estereográfica desde el polo norte. Fijemos en la esfera dos puntos antipodales, a los que llamaremos N (polo norte) y S (polo sur). El plano del ecuador será entonces el plano perpendicular a la recta NS que pasa por el centro de la esfera. A cada punto $P \neq N$ en la esfera le asociamos el punto del plano del ecuador que es la intersección de este plano con la recta NP.

a) Para la esfera $\mathcal{S}$ escoger $N = (0, 0, 1)$ y construir la correspondencia que a cada punto $P = (x, y, z)$ en $\mathcal{S}$ le asocia las coordenadas (u, v) del punto $(u, v, 0)$ determinado por la construcción que acabamos de describir. Hallar x, y y z en función de (u, v).

b) Repetir la construcción proyectando desde el polo sur. Para cada punto (x, y, z) se obtendrá un punto de coordenadas $(\bar{u}, \bar{v})$. Hallar la expresión de estas nuevas coordenadas $(\bar{u}, \bar{v})$ en función de (u, v).

4. Casquetes de esfera como gráficos de una función. Mostrar que la correspondencia
$$(x, y) \mapsto (x, y, \sqrt{1 - x^2 - y^2}),$$
definida para los puntos (x, y) que satisfacen $x^2 + y^2 < 1$, parametriza una parte de la esfera. ¿Qué parte es ésta? ¿Cuántas parametrizaciones de este tipo son necesarias para cubrir completamente la esfera?

Ejercicio 3.127. Cada butaca de un avión de pasajeros tiene un monitor de 30 centímetros de ancho por 20 de alto. En esta pantalla se proyecta un mapa de la Tierra, construido por medio de una proyección de Mercator. En el punto medio de la pantalla se representa el punto que corresponde a la intersección del Ecuador con el meridiano de Greenwich.

1. Calcular cuáles son las mínima y máxima latitud que la pantalla puede representar si se quiere que el desarrollo del Ecuador ocupe toda la longitud horizontal de la pantalla.
2. Los sistemas de navegación del avión reciben lecturas de latitud y longitud en grados, minutos y segundos. Hallar las fórmulas que es necesario programar para que las computadoras muestren correctamente sobre la pantalla de cada pasajero el lugar en que se encuentra el avión.
3. Ubicar a Montevideo en la pantalla.

3.5.3. Para tener presente

- Las curvas son objetos unidimensionales en el espacio. Típicamente pueden describirse por una parametrización, con un único parámetro real. En este sentido, son la imagen de un intervalo de $\mathbb{R}$ por una función que toma valores en $\mathbb{R}^3$. Cuando las ecuaciones de la parametrización son lineales obtenemos una recta, o parte de ellas. Ecuaciones no lineales producen, en general, otras curvas.

- Las superficies son objetos bidimensionales en el espacio. Típicamente pueden describirse por una parametrización, con dos parámetros reales. En este sentido, son la imagen de una región del plano $\mathbb{R}^2$ por una función que toma valores en $\mathbb{R}^3$. Cuando las ecuaciones de la parametrización son lineales obtenemos un plano, o parte de él. Ecuaciones no lineales producen, en general, otras superficies.

- Una superficie es, en general, el conjunto de puntos cuyas coordenadas satisfacen una ecuación.
- Una curva es, en general, el conjunto de puntos cuyas coordenadas satisfacen un sistema de dos ecuaciones. Cada una de las ecuaciones representa una superficie, y la curva es la intersección de ambas superficies.
- Las consideraciones de los puntos anteriores son esencialmente ciertas, y dan una idea general adecuada. Pero no son completamente ciertas. Para que se satisfagan es necesario incorporar condiciones de regularidad que deben ser satisfechas por las ecuaciones.
- Al igual que para rectas y planos, las ecuaciones implícitas de curvas y superficies son una condición que sólo es satisfecha por los puntos que están en la imagen de las funciones que definen las representaciones paramétricas. Recíprocamente, las ecuaciones paramétricas dan una manera de recorrer las soluciones de las ecuaciones implícitas. Sin embargo, en el caso general de curvas y superficies cualesquiera, a diferencia de lo que ocurre para las rectas y planos que tienen ecuaciones lineales, no hay una manera sistemática de pasar de un tipo de ecuación a otra, porque no hay un algoritmo general de resolución de sistemas de ecuaciones cualesquiera. Las ecuaciones lineales sí tienen un algoritmo general: la eliminación gaussiana.

Capítulo 4

Espacios vectoriales

En muchas ramas de la matemática aparecen conjuntos cuyos elementos pueden ser multiplicados por números y sumados entre sí. Es decir, nos permiten formar *combinaciones lineales*. Ejemplos ya vistos en este curso son los vectores de $\mathbb{K}^m$, y las matrices con entradas en un cuerpo $\mathbb{K}$. Se trata de listas y arreglos de números, que están dotados además de una estructura algebraica definida por las operaciones de suma y producto por un escalar.

Pero estos dos ejemplos no agotan la lista de objetos que pueden sumarse y multiplicarse por un escalar para producir nuevos objetos de la misma clase: en general, las funciones que toman valores en un cuerpo, por ejemplo, las funciones reales, que toman valores en $\mathbb{R}$, o las complejas, que lo hacen en $\mathbb{C}$, tienen esta propiedad. También la tienen muchos de los subconjuntos de funciones que son interesantes en algún sentido. Por ejemplo, las funciones reales de variable real que son continuas, o las que son derivables, o las que son solución de algunas de las ecuaciones de la física. Podemos agregar también a la lista los polinomios, y los vectores tangentes a una superficie en un punto dado de ella, etcétera.

En fin, variedad de objetos diferentes, interesantes por sí mismos y/o por sus aplicaciones –sólo tenemos que detenernos un momento a pensar la gran cantidad de situaciones que pueden ser descritas por medio de una función real de variable real o de una matriz–, admiten estas dos operaciones particularmente simples, con propiedades comunes a situaciones muy diferentes.

Este es el contexto adecuado para formular una teoría matemática que extraiga las propiedades fundamentales, compartidas por toda esta colección de ejemplos, y a partir de ellas produzca conclusiones generales, aplicables en todas las situaciones en que se satisfagan las hipótesis en que las conclusiones fueron deducidas. Es así que las operaciones de suma y producto que están definidas en los ejemplo arriba mencionados se generalizan a los elementos, que llamaremos *vectores*, de una estructura algebraica que por su sencillez y amplitud se utiliza con diversos tipos de objetos y en diversos tipos de aplicaciones: la estructura de *espacio vectorial*.

Desarrollar una teoría abstracta tiene una gran ventaja, porque de una sola vez y con métodos generales se trata una gran diversidad de problemas. Y una desventaja que formulamos un poco caricaturescamente: al hablar de todo parece que ya no hablamos de nada. Problemas que nos eran familiares se han diluido con otros cuya existencia incluso ignoramos en el marco de una teoría que habla de objetos que no requieren de ninguna representación para existir.

Quizás interesa al lector saber que el paso a considerar una estructura abstracta es una dificultad real en el aprendizaje, y ha sido relevado en la literatura sobre la enseñanza del Álgebra Lineal. Por ejemplo, en [C] el autor describe elocuentemente el desconcierto que suele producirse cuando en la clase las rutinas algebraicas dan paso a los conceptos propios del Álgebra Lineal. En [D] el autor apunta a la siguiente dificultad: *cualquier problema lineal al alcance de un estudiante en el primer año de la Universidad puede ser resuelto sin usar la teoría axiomática. El provecho en términos de unificación, generalización y simplificación que trae la teoría formal es solo visible para el experto ... Sin embargo, muchas personas encuentran importante que los estudiantes que comienzan estudios de ciencia y matemática tengan alguna idea acerca de las estructuras axiomáticas, de las cuales la de espacio vectorial es una de las más fundamentales.*

En este texto hemos escogido presentar la teoría axiomática abstracta de los espacios vectoriales, que es el objetivo de este capítulo, tratando de mostrar la potencia, importancia e interés de este enfoque, al tiempo que discutimos con cierto detalle algunas de sus realizaciones concretas. Recomendamos enfáticamente al lector trabajar constantemente sobre ellas e interpretar los resultados en el marco de algún contexto que le sea familiar. Por ejemplo, el de los espacios $\mathbb{K}^n$ con los que ya hemos trabajado. De hecho, tal como veremos, los espacios $\mathbb{K}^n$ son un modelo muy adecuado para cualquier espacio vectorial de dimensión finita.

En la sección 4.1 introducimos la noción de *espacio vectorial*. Se trata de una estructura abstracta, que generaliza a los espacios vectoriales $\mathbb{K}^n$ que nos son conocidos desde el capítulo 1. En la seción 4.2 trabajamos con los *subespacios* de espacios vectoriales cualesquiera. La noción de subespacio es la misma que para los espacios $\mathbb{K}^n$: subconjuntos de un espacios vectorial que son cerrados para las operaciones de suma y de producto por un escalar. Como dentro de un subespacio podremos realizar estas operaciones, resulta que cualquier subespacio vectorial es también un espacio vectorial, y podremos usar en él todos los resultados de la teoría abstracta que desarrollaremos en este capítulo.

La noción de *combinación lineal* es central en la teoría de los espacios vectoriales. A partir de ella se introducen en las secciones 4.3 y 4.4 los importantes conceptos de conjunto generador, independencia y dependencia lineal, base y dimensión, que generalizan los conceptos correspondientes vistos en el contexto de $\mathbb{K}^n$ en el capítulo 1. Una vez definida la dimensión de un espacio vectorial, podemos identificar claramente la estructura que es el principal objeto de estudio en este texto: *los espacios vectoriales de dimensión finita.*

Los vectores de un espacio vectorial de dimensión finita quedan completamente identificados por sus *coordenadas* respecto a una base del espacio. Las coordenadas son el objeto de trabajo de la sección 4.5.

Las coordenadas son un primer ejemplo de *transformación lineal*. A cada vector del espacio le asocian un vector de coordenadas, de forma tal que a la suma de vectores le corresponde la suma de coordenadas, y al producto de un vector por un escalar le corresponde el producto del mismo escalar por el vector de coordenadas. Las transformaciones entre espacios vectoriales que tienen estas propiedades son las *transformaciones lineales*, que se introducen formalmente en la sección 4.6.

Así como cualquier vector puede representarse en coordenadas, la acción de una transformación lineal puede analizarse en términos de coordenadas, y resulta ser equivalente a la multiplicación por una matriz, que llamaremos *matriz asociada* a la transformación.

En la sección 4.7 estudiamos los efectos de las operaciones con transformaciones lineales sobre sus matrices asociadas. En particular, analizamos con detalle lo que ocurre al componer transformaciones lineales. La teoría que resulta permite manejar eficientemente los cambios de base dentro de un espacio vectorial.

Al tomar coordenadas en un $\mathbb{K}$-espacio vectorial de dimensión finita n se introduce una correspondencia lineal biyectiva con el espacio $\mathbb{K}^n$. Esto es lo que se llama un *isomorfismo*, una transformación lineal que identifica completamente la estructura lineal de dos espacios vectoriales diferentes. Estudiaremos estas transformaciones en la sección 4.8, en la que nos ocuparemos de las transformaciones lineales inyectivas y sobreyectivas. Ambas propiedades pueden caracterizarse en términos de subespacios asociados con la transformación: su *imagen* y su *núcleo*. Se obtiene así una teoría que puede verse como una versión abstracta de nuestra respuesta a dos problemas centrales en el estudio de los sistemas lineales: determinar cuándo un sistema es compatible, y cuándo es determinado.

Cerramos el capítulo y el texto con un importante resultado sobre las transformaciones lineales definidas en un espacio vectorial de dimensión finita: el teorema de las dimensiones del núcleo y de la imagen, en la sección 4.9.

4.1. Definición y ejemplos de la estructura de espacio vectorial

El objetivo de esta sección es introducir la noción abstracta de espacio vectorial: una estructura cuyos elementos, vectores, pueden multiplicarse por escalares y sumarse para dar lugar a combinaciones lineales. Generalizaremos así los espacios vectoriales $\mathbb{K}^n$ a un contexto mucho más general que engloba una gran cantidad de situaciones.

Adoptaremos un enfoque axiomático, y nos desentenderemos de cuál es la naturaleza de los vectores para concentrarnos en las relaciones que las operaciones de suma y producto establecen entre ellos y con un conjunto de escalares que tendrá estructura de cuerpo. La caracterización de la estructura estará dada por una serie de axiomas o postulados básicos, que deben ser satisfechos por las operaciones de suma y producto por un escalar. Estos axiomas no son otra cosa que las propiedades de suma de vectores de $\mathbb{K}^n$ o de matrices en $M^{m\times n}(\mathbb{K})$, y del producto de estos objetos por escalares en $\mathbb{K}$, que hemos detallado, respectivamente, en la sección 1.2.6, página 66 y en el ejercicio 2.12, página 177.

Comenzaremos con un ejemplo acerca de un problema de interpolación, cuya solución quedará expresada en términos de combinaciones lineales de unas pocas funciones especialmente simples. Daremos luego un segundo ejemplo: el conjunto formado por las funciones que están definidas sobre un conjunto X cualquiera, y que toman valores en algún cuerpo $\mathbb{K}$. Más tarde mostraremos que ambos ejemplos caen dentro del marco general de la teoría de los espacios vectoriales.

4.1.1. Ejemplos introductorios

La *interpolación* es una operación corriente en la vida cotidiana y en el cálculo científico, que consiste en inferir los valores de una función a partir de unos pocos valores conocidos:

- si la aguja del velocímetro del automóvil está entre dos marcas, mentalmente interpolamos para hacernos una idea de cuál es la velocidad a la que estamos viajando;

- si tenemos datos de una función en algunos puntos podríamos querer conocer valores intermedios, ya sea para calcular con ellos, o para presentar gráficamente los datos;

- la animación por computadora también requiere refinados procesos de interpolación para producir imágenes agradables;

- se emplean técnicas de interpolación para resolver un problema corriente de diseño: hallar una forma funcional para una curva o superficie que ha sido dada en forma gráfica;

- muchos métodos de cálculo numérico realizan, en forma explícita o implícita, algún tipo de interpolación. Por ejemplo, para calcular la integral $\int_a^b f(x)dx$ de una cierta función f sobre un intervalo $[a,b]$ se evalua f en algunos puntos del intervalo y se estima la integral a partir de estas evaluaciones[1]. Los diferentes procedimientos de estimar la integral son equivalentes a tomar una nueva función, digamos ϕ, más simple que la original, que interpole los valores evaluados (por lo tanto es de esperar que constituya una buena aproximación de la función original) y calcular el valor $\int_a^b \phi(x)dx$.

El problema general de interpolación puede plantearse de la siguiente manera: dadas algunas parejas

$$(x_i, y_i), \quad i = 1, 2, \dots, n,$$

de valores, hallar una función f tal que

$$f(x_i) = y_i, \quad i = 1, 2, \dots, n.$$

Naturalmente, el problema así planteado tiene infinitas soluciones, porque podemos escoger f con total libertad. Para plantear el problema correctamente debemos especificar dentro de qué conjunto de funciones vamos a buscar la función f. Es decir, debemos decir con qué tipo de funciones vamos a trabajar.

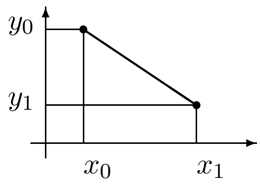
En el próximo ejemplo discutiremos con cierto detalle una técnica específica para realizar la interpolación. En ese ejemplo ilustraremos el procedimiento de interpolación a través del planteo de un problema concreto, y veremos como el espacio de funciones con el que trabajaremos admite una estructura que permite resolver el problema planteado. Esta estructura proviene del hecho de que las funciones pueden sumarse y multiplicarse por números para producir *combinaciones lineales*, tal como ocurre con vectores y matrices.

---

[1]Cualquier texto de cálculo numérico explica estos procedimientos con detalle. Ver, por ejemplo, la página 139 de *Numerical Methods and Software*, de D. Kahaner, C. Moler y S. Nash, Prentice Hall, 1989

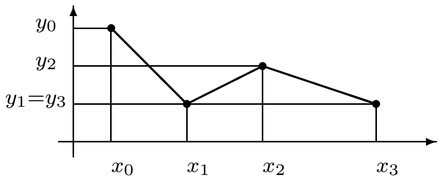
Ejemplo 4.1.1. Interpolación lineal a trozos
Una primera aproximación al problema de interpolación consiste en resolverlo completando los datos inexistentes por funciones lineales.

Así, si conocemos los valores y_0 e y_1 de una función en dos puntos x_0 y x_1, a los que es corriente llamar *nodos*, interpolamos entre x_0 y x_1 por una función lineal, tal como se muestra en la figura.



No hay por qué limitarse a dos puntos, ya que este procedimiento se adapta muy naturalmente al caso en que buscamos interpolar los valores de la función cuando tenemos datos en más puntos.

Si conocemos valores y_0, y_1, $\dots y_n$, de la función en los nodos x_0, x_1, $\dots x_n$, entonces podemos completar entre dos nodos consecutivos por una función lineal. El resultado de aplicar este procedimiento es una función lineal a trozos, cuyo gráfico es del tipo mostrado en la figura, en la que se ilustra esta técnica con 4 nodos.



Las funciones lineales a trozos proveen entonces una posible respuesta a un problema de interpolación: entre dos nodos consecutivos los valores de la función se completan con un interpolante lineal. Vamos a ilustrar este procedimiento trabajando con funciones reales definidas en el intervalo $[-1, 1]$, y con los nodos -1, 0 y 1. Nuestra argumentación es fácilmente generalizable a un intervalo $[a, b]$ cualquiera con nodos

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b,$$

sólo hay que introducir la notación adecuada.

Consideraremos entonces funciones reales continuas definidas sobre $[-1, 1]$, cuyas restricciones a los intervalos $[-1, 0]$ y $[0, 1]$ son funciones lineales. Llamaremos $\mathbb{L}$ al conjunto formado por estas funciones, a las que llamaremos *interpolantes lineales a trozos.*

El conjunto $\mathbb{L}$ tiene dos propiedades notables:

- **la suma de dos funciones f y g de $\mathbb{L}$ es un nuevo elemento $\mathbb{L}$;**
- **el producto de un real por una función de $\mathbb{L}$ también es una función de $\mathbb{L}$.**

O sea, podemos sumar y multiplicar por escalares dentro de $\mathbb{L}$, igual que ocurría en los espacios $\mathbb{K}^n$ o en el conjunto de las matrices sobre un cuerpo $\mathbb{K}$.

En consecuencia, es posible hacer combinaciones lineales de las funciones de $\mathbb{L}$ para obtener nuevas funciones de $\mathbb{L}$.

Cada elemento f de $\mathbb{L}$ queda completamente determinado al fijar los valores

$$y_{-1} = f(-1), \quad y_0 = f(0), \quad y_1 = f(1)$$

que la función toma sobre los nodos. Esto permite, representar cualquier interpolante lineal a trozos a partir de sólo tres funciones muy sencillas: las funciones l_{-1}, l_0 y l_1 cuyos gráficos se muestran en la figura 4.1. La función $l_{-1} \in \mathbb{L}$ queda determinada por las condiciones

$$l_{-1}(-1) = 1, \quad l_{-1}(0) = l_{-1}(1) = 0.$$

En tanto que l_0 y l_1 están, respectivamente, caracterizadas por

$$\begin{array}{ll} l_0(0) = 1, & l_0(-1) = l_0(1) = 0, \\ l_1(1) = 1, & l_1(-1) = l_1(0) = 0. \end{array}$$

Con estas tres funciones podemos "construir" cualquier otra función $g \in \mathbb{L}$,

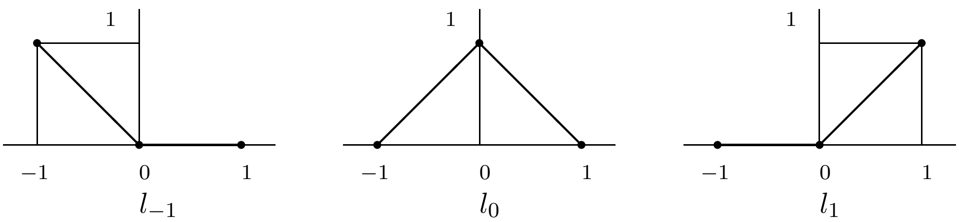


Figura 4.1: Las tres funciones básicas para hacer interpolación lineal a trozos

en el sentido de que g se puede expresar como

$$g(x) = g(-1)l_{-1}(x) + g(0)l_0(x) + g(1)l_1(x).$$

Esta fórmula, permite, por ejemplo calcular fácilmente la integral de g sobre el intervalo $[-1, 1]$ en términos de los valores que g toma en los nodos -1, 0 y 1. Tenemos

$$\int_{-1}^{1} g(x)\,dx = \int_{-1}^{1} [g(-1)l_{-1}(x) + g(0)l_0(x) + g(1)l_1(x)]\,dx.$$

Aplicando la propiedad de linealidad de la integral resulta

$$\int_{-1}^{1} g(x)\,dx = g(-1)\int_{-1}^{1} l_{-1}(x)\,dx + g(0)\int_{-1}^{1} l_0(x)\,dx + g(1)\int_{-1}^{1} l_1(x)\,dx.$$

Las integrales de l_{-1}, l_0 y l_1 son fáciles de evaluar. Al hacerlo y sustituir sus valores en la fórmula anterior, conseguimos la sencilla expresión en función de los valores en los nodos -1, 0 y 1.

$$\int_{-1}^{1} g(x)\,dx = \frac{1}{2}g(-1) + g(0) + \frac{1}{2}g(1).$$

Al igual que ocurre con los espacios $\mathbb{K}^n$, el hecho de que en $\mathbb{L}$ estén definidas operaciones de suma y de producto por un escalar con buenas propiedades dota a este conjunto de estructura. La estructura es la de *espacio vectorial.*

Nos hemos encontrado entonces con una nueva clase de objetos, las funciones lineales a trozos (que resuelven un problema de interpolación), que pueden sumarse y multiplicarse por escalares. En este sentido se comportan como los vectores de $\mathbb{K}^n$, o como las matrices. Además, las operaciones de suma y de producto por un escalar tienen esencialmente las mismas propiedades que en $\mathbb{K}^n$ o las matrices. Las nociones de combinación lineal, independencia lineal, generadores y bases tiene completo sentido en $\mathbb{L}$, y pueden ser definidas exactamente de la misma manera que en $\mathbb{K}^n$. De hecho, veremos luego que $\{l_{-1}, l_0, l_1\}$ es una base de $\mathbb{L}$, en el mismo sentido que encontramos bases de subespacios de $\mathbb{K}^n$ en la sección 1.5.

No nos detendremos ahora en completar los detalles de este ejemplo particular porque centraremos nuestra atención en desarrollar una teoría abstracta, general, que englobe este y otros ejemplos. ♣

Pero antes de dar la definición de espacio vectorial introducimos un nuevo ejemplo que será parte de la teoría general. Este ejemplo tiene una formulación bastante abstracta. Es, en realidad, muy general y mostraremos luego que abarca, entre otros, a nuestros conocidos $\mathbb{K}^n$ y al conjunto de las matrices.

Ejemplo 4.1.2. Las funciones que toman valores en un cuerpo
Sean X un conjunto cualquiera y $\mathbb{K}$ un cuerpo. Consideramos el conjunto $\mathcal{F}$ formado por todas las funciones de X que toman valores en $\mathbb{K}$

$$\mathcal{F} = \{f : X \to \mathbb{K}\}.$$

Definimos en $\mathcal{F}$ las operaciones de suma de funciones y de producto por un escalar, de la manera habitual:

- la operación $+$ de suma de dos funciones f y g queda definida al especificar que la función suma $f+g$ es la que satisface

$$(f+g)(x) = f(x) + g(x), \quad \forall x \in X.$$

- el producto de una función f por un escalar $\alpha \in \mathbb{K}$ a través de la fórmula

$$(\alpha f)(x) = \alpha f(x), \quad \forall x \in X.$$

Estas operaciones tienen las mismas propiedades que las operaciones análogas definidas sobre las n-uplas en $\mathbb{K}$ o las matrices con entradas en $\mathbb{K}$.

Propiedades de la suma. De las propiedades conmutativa y asociativa para la suma en un cuerpo $\mathbb{K}$ podemos deducir las propiedades conmutativa y asociativa de la suma de funciones. Veamos primero la asociativa. Si f, g y h son tres funciones definidas sobre X debemos mostrar que

$$(f+g)+h = f+(g+h).$$

¿Cómo se hace esto? Evaluando las funciones en cada punto y usando que el resultado de evaluar es un número en $\mathbb{K}$. Tenemos entonces, para $x \in X$, que

$$((f+g)+h)(x) = (f+g)(x) + h(x) = f(x) + g(x) + h(x). \tag{4.1}$$

Hasta aquí sólo usamos la definición de suma de funciones. Podemos agrupar los sumandos del miembro de la derecha como nos apetezca, así que juntamos g y h, para hacer aparecer la suma $g+h$. Entonces obtenemos

$$((f+g)+h)(x) = f(x) + (g+h)(x), \quad x \in X$$

Aplicando una vez más la definición de suma, ahora entre f y $g+h$ concluimos

$$((f+g)+h)(x) = (f+(g+h))(x), \quad x \in X$$

lo que es equivalente a (4.1). La propiedad conmutativa se muestra igual. Omitimos los detalles.

Existe un neutro para la suma, porque la función idénticamente nula

$$O(x) = 0, \quad \forall x \in X,$$

satisface

$$f + O = f$$

para cualquier otra función f definida sobre X. También es evidente la existencia de un opuesto para cada función f. El opuesto de f es la función $-f$, que se obtiene multiplicando los valores de f por -1, el opuesto de 1 en el cuerpo $\mathbb{K}$.

Propiedades del producto por un escalar. La asociativa respecto al producto de escalares y las distributivas del producto respecto a la suma de escalares y de funciones se reducen, evaluando en cada $x \in X$, a las propiedades asociativa y distributiva para escalares. El producto $1f$ de la unidad en el cuerpo con una función f cualquiera satisface

$$(1f)(x) = 1f(x) = f(x)$$

para todo $x \in X$, de modo que, efectivamente, $1f = f$.

Veremos luego que estos conjuntos de funciones son un ejemplo de espacio vectorial. ♣

4.1.2. Introducción de la noción abstracta de espacio vectorial

La diversidad de ejemplos en los que los elementos de un conjunto pueden sumarse entre sí y multiplicarse por escalares de forma tal que se satisfacen las propiedades que encontramos en nuestros dos últimos ejemplos justifica hacer una presentación general, abstracta, que englobe a todos estos ejemplos y que desarrolle una teoría válida para todas las situaciones particulares. Además de la generalidad que proporciona el disponer de una notación y métodos generales, aplicables en muchos contextos, esta presentación nos evita enredarnos en los detalles particulares de cada ejemplo. Definamos entonces qué cosa es un *espacio vectorial.*

Definición 4.1. *Un* ***espacio vectorial*** *sobre un cuerpo* $\mathbb{K}$ *consta de cuatro elementos ordenados* $(\mathbb{V}, \mathbb{K}, +, \cdot)$*:*

- *un conjunto* $\mathbb{V}$ *no vacío. Nos referiremos a los elementos de este conjunto como* vectores*;*

- *un cuerpo* $\mathbb{K}$*. Llamaremos* escalares *a los números en* $\mathbb{K}$*;*

- *una operación llamada* suma de vectores*, y denotada con el símbolo* $+$*,*

  $$+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V},$$

  que a cada par de vectores (u, v) *en* $\mathbb{V}$ *le asocia un nuevo vector* $u+v \in \mathbb{V}$*;*

- *una operación llamada* producto de un escalar por un vector*, a la que indicaremos con el símbolo* $\cdot$ *o simplemente escribiendo el escalar junto al vector,*

  $$\cdot : \mathbb{K} \times \mathbb{V} \to \mathbb{V},$$

  que a cada pareja (λ, v)*, con* $\lambda \in \mathbb{K}$ *y* $v \in \mathbb{V}$ *le asocia el producto* $\lambda v \in \mathbb{V}$*.*

Las operaciones de suma y producto por un escalar deben satisfacer, para todo u, v y w en $\mathbb{V}$, y para todo α y β en $\mathbb{K}$, las siguientes propiedades:
Propiedades de la suma.

[S1] Conmutativa: $u + v = v + u$.

[S2] Asociativa: $(u + v) + w = u + (v + w)$.

[S3] Neutro de la suma: *Existe un vector $O \in \mathbb{V}$ tal que*

$$u + O = O + u = u.$$

Llamaremos a O el vector nulo.

[S4] Existencia de opuesto: *existe un vector $(-u) \in \mathbb{V}$ de u, tal que*

$$u + (-u) = O.$$

Llamaremos a $-u$ el opuesto *de u.*

Propiedades del producto.

[P1] Asociativa: $(\alpha\beta)u = \alpha(\beta u)$.

[P2] Neutro del producto $1u = u$.

[P3] Distributiva respecto a la suma de escalares: $(\alpha + \beta)u = \alpha u + \beta u$.

[P4] Distributiva respecto a la suma de vectores: $\alpha(u+v) = \alpha u + \alpha v$.

Ejemplo 4.1.3. El ejemplo más básico.
¡Sí! Los espacios $(\mathbb{K}^n, \mathbb{K}, +, \cdot)$ son espacios vectoriales sobre el cuerpo $\mathbb{K}$. Esto justifica el haberlos denominado "el espacio vectorial $\mathbb{K}^n$" cuando los introdujimos en el capítulo 1.

Con la tranquilidad de espíritu de saber que nuestro principal ejemplo se enmarca dentro de la teoría que estamos construyendo repasemos los principales hitos en la formación de esta teoría y conozcamos a algunos de los protagonistas de la historia.

Nota Histórica 4.1. Muchos ejemplos de métodos lineales en la Matemática, o la Física, pueden rastrearse prácticamente hasta la antigüedad. Pero, hasta el siglo XIX, aparecían cómo métodos aislados. Un mismo autor podía utilizar en dos trabajos diferentes dos ideas "lineales" que hoy reconoceríamos instantáneamente como dos casos particulares de una idea general, sin percatarse de ello.

El proceso de axiomatización de la teoría de espacios vectoriales, hasta su aceptación y difusión generalizada cubre, a grandes rasgos, el período que va de 1870 hasta la década de 1940. En este período emergió vigorosamente dentro de la Matemática el punto de vista de ocuparse de estructuras definidas axiomáticamente. Si bien se trata de una cuestión bastante rutinaria hoy en día, este giro modificó sustancialmente el paisaje de la Matemática, y hasta la concepción de qué es la disciplina.

En 1888 el matemático italiano Giuseppe Peano (1858-1932), explica con claridad en el libro *Cálculo Geométrico* ideas que Grassmann (1809-1877) había presentado en 1844, en su *Teoría Lineal de la Extensión.* La presentación de Peano incluye en un apéndice la novedad de formular axiomáticamente la teoría. Digamos que los axiomas de Peano eran muy similares a las propiedades fundamentales identificadas por Grassman[2].

Sin embargo, la obra de Peano no tuvo gran impacto. La estructura de espacio vectorial fue "redescubierta" varias veces en años sucesivos, por Hermann Weyl (1885-1955), Hans Hahn (1879-1934), Norbert Wiener (1894-1964) y Stefan Banach (1892-1945). Banach, Hahn y Wiener descubrieron la noción de espacio vectorial normado[3] a partir de sus trabajos en el área del Análisis Matemático. La presentación de Banach, en un trabajo de 1922 que resume su tesis doctoral de 1920, fue la que tuvo mayor influencia. Y su publicación motivó la consideración de esta estructura por otros matemáticos.

En los años 40 aparecieron dos textos que exponían la teoría a un nivel adecuado para estudiantes universitarios y que tuvieron amplia difusión: *Álgebra Moderna*, de Garret Birkhoff (1911-1996) y Saunders MacLane (1909-), y *Espacios Vectoriales de Dimensión Finita*, de Paul Halmos (1916-). En Uruguay la teoría comenzó a exponerse en los años 60, en el marco del cuso Análisis Matemático I (antepasado de los actuales Cálculo 1 y Cálculo 2), introducida por José Luis Massera (1915-2002). El principal propósito que en ese curso tenía esta teoría era el de proporcionar fundamentos teóricos claros para la comprensión de la estructura del espacio de soluciones de una ecuación diferencial lineal[4].

El lector interesado en los aspectos históricos de los temas que tratamos en el curso puede consultar los artículos *A general outline of the genesis of vector space theory* de Jean-Luc Dorier, que aparece en las páginas 227–261 del volumen 22, año 1995, de la revista Historia Mathematica, y *The Axiomatization of Linear Algebra: 1875-1940*, de Gregory H. Moore, en las páginas 261–303 del mismo número de la misma revista[5]. Otra interesante fuente de información histórica es el sitio web `http://www-groups.dcs.st-and.ac.uk/~history/`, de la Escuela de Matemática y Estadística de la Universidad de Saint Andrew, que fue consultado varias veces en marzo de 2005 durante la preparación de esta nota histórica.

Observación 4.1.4. Un inofensivo abuso de notación.
Usamos el mismo símbolo $(+)$ para la suma de escalares (suma dentro de $\mathbb{K}$) y para la suma de vectores (suma dentro de $\mathbb{V}$), pero son operaciones diferentes, definidas sobre dominios diferentes. Por ejemplo, si miramos $\mathbb{R}^2$

[2] En este texto hemos hecho un recorrido similar: los axiomas que satisfacen las operaciones en un espacio vectorial son las mismas propiedades fundamentales de la estructura vectorial de los espacios $\mathbb{K}^n$

[3] Un espacio vectorial normado es un espacio vectorial, normado. Es decir, un espacio vectorial, en el que que hay además una noción de longitud de los vectores. La versión abstracta de esta medida de longitud es lo que se llama una *norma*. Por ejemplo, el módulo de los vectores de $\mathbb{R}^2$, o $\mathbb{R}^3$ es un caso particular de una norma. Pero no es esta la única manera de medir la longitud de los vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$, y tampoco es adecuada para medir los vectores de otros espacios vectoriales, por ejemplo, espacios de funciones, que requieren introducir normas apropiadas.

[4] Un ejemplo de este tipo de ecuaciones aparece en el ejemplo 4.2.13, página 421.

[5] El número citado está en la hemeroteca de la comunidad matemática uruguaya, en el Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia", de la Facultad de Ingeniería.

como el espacio vectorial $(\mathbb{R}^2, \mathbb{R}, +, \cdot)$ donde las operaciones son las que fueron introducidas para listas de 2 números reales tenemos como ejemplo de esta situación

$$\begin{array}{ll} 2 + 2{,}5 = 4{,}5 & \text{suma en el cuerpo,} \\ (2,1) + (1,0) = (3,1) & \text{suma de vectores.} \end{array}$$

También para el producto vamos a usar la misma notación (la simple concatenación de los símbolos) en el caso de dos escalares y el de un escalar y un vector, aunque ambas operaciones son diferentes.

Observación 4.1.5. MÁS SOBRE LA NOTACIÓN Y LA JERGA.
Cuando tenemos un espacio vectorial $(\mathbb{V}, \mathbb{K}, +, \cdot)$ es corriente referirnos a "el espacio vectorial $\mathbb{V}$", sin mencionar explícitamente el cuerpo ni hacer referencia a la definición de las operaciones de suma y producto por un escalar. Por ejemplo, frecuentemente diremos "el espacio vectorial $\mathbb{K}^n$" para designar a $(\mathbb{K}^n, \mathbb{K}, +, \cdot)$. No mencionar el cuerpo y a las operaciones de suma y producto implica dar por entendido que son $\mathbb{K}$ y las operaciones usuales.

También es corriente hacer explícito el cuerpo sobre el que se trabaja, diciendo que $\mathbb{V}$ es un $\mathbb{K}$-espacio vectorial, o que $\mathbb{V}$ es un espacio vectorial sobre $\mathbb{K}$, sin mencionar las operaciones. Cuando $\mathbb{K}$ es el cuerpo $\mathbb{R}$ de los números reales es habitual decir que $\mathbb{V}$ es un *espacio vectorial real*. En tanto que diremos que $\mathbb{V}$ es un o *espacio vectorial complejo* cuando $\mathbb{K}$ es el cuerpo $\mathbb{C}$ de los números complejos. ♣

Otro viejo conocido, el conjunto de las matrices $m \times n$ resulta ser también un espacio vectorial.

Ejemplo 4.1.6. EL ESPACIO VECTORIAL DE LAS MATRICES $m \times n$.
Sea $M^{m\times n}(\mathbb{K})$ el conjunto de las matrices $m \times n$ con entradas en el cuerpo $\mathbb{K}$. La cuaterna $(M^{m\times n}(\mathbb{K}), \mathbb{K}, +, \cdot)$, donde las operaciones de suma y producto por un escalar se realizan entrada a entrada, es un $\mathbb{K}$-espacio vectorial. La verificación de las propiedades necesarias se encuentra en la sección 2.2 en la página 174 de este texto.

Observación 4.1.7. UN EJEMPLO BASTANTE UNIVERSAL
Tal vez sea el momento de volver sobre el ejemplo 4.1.2, tratado antes de la definición formal de espacio vectorial. A esta altura huelga decir que $(\mathcal{F}, \mathbb{K}, +, \cdot)$, donde $\mathcal{F}$ es el conjunto de las funciones definidas sobre un conjunto X y que toman valores en un cuerpo $\mathbb{K}$, con las operaciones $+$ y $\cdot$ definidas en ese ejemplo es un espacio vectorial.

Como X puede ser un conjunto cualquiera y $\mathbb{K}$ un cuerpo cualquiera esta construcción contiene muchos ejemplos interesantes.

Si X es un intervalo I cualquiera contenido en $\mathbb{R}$ y $\mathbb{K} = \mathbb{R}$ el conjunto de todas las funciones reales definidas sobre el intervalo I forma un espacio vectorial. También es espacio vectorial el conjunto de las funciones complejas definidas sobre el intervalo I , resultado que surge tomando $\mathbb{K} = \mathbb{C}$.

Pero podemos escoger $X = \mathbb{R}^n$ y $\mathbb{K} = \mathbb{R}$ para mostrar que el conjunto formado por todas las funciones reales con varias variables reales también es un espacio vectorial. Haciendo nuevas variaciones sobre esta construcción podríamos seguir mostrando que muchos conjuntos de funciones que aparecen naturalmente en la matemática y sus aplicaciones tienen esta estructura ubicua de espacio vectorial.

También nuestros familiares espacios $\mathbb{K}^n$ y $M^{m\times n}(\mathbb{K})$ caen dentro de esta construcción. Sólo hay que cambiar un poco la notación para convencerse de esto. Tal como se describía en la observación 1.2.34, página 69, un vector Y de $\mathbb{K}^n$ puede ser pensado como una función definida sobre el conjunto $\{1, 2, \ldots, n\}$ y que toma valores en $\mathbb{K}$. Pero escribimos el valor que la función toma en 1 como y_1, en vez de hacerlo con la notación de $Y(1)$, la habitual cuando se trabaja con funciones.

Con las matrices pasa algo parecido. Una matriz $A \in M^{m\times n}(\mathbb{K})$ es una función definida sobre $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$, pero solemos escribir $A(i, j) = a_{ij}$. Ver, al respecto, la definición 1.1, en la página 37, que introduce formalmente las matrices como funciones.

Ejemplo 4.1.8. El conjunto de las sucesiones reales, en el que la suma de dos sucesiones $\{a_n\}_{n\in\mathbb{N}}$, $\{b_n\}_{n\in\mathbb{N}}$, y el producto de una sucesión por un escalar λ, están definidas por las fórmulas

$$\begin{aligned} \{a_n\}_{n\in\mathbb{N}} + \{b_n\}_{n\in\mathbb{N}} &= \{a_n + b_n\}_{n\in\mathbb{N}}\,, \\ \lambda\,\{a_n\}_{n\in\mathbb{N}} &= \{\lambda a_n\}_{n\in\mathbb{N}}\,, \end{aligned}$$

forman un espacio vectorial real, ya que caen dentro de este marco general. Lo podemos ver tomando $X = \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ en nuestra construcción.

Ejercicio 4.1.

- ¿Cuál es el vector nulo en este espacio de sucesiones?
- Mostrar que el conjunto formado por todas las sucesiones complejas es un espacio vectorial complejo. ♣

El objetivo del siguiente ejercicio es extender la construcción general del ejemplo 4.1.2 para construir nuevos espacios vectoriales: las funciones que toman valores en espacios vectoriales.

Ejercicio 4.2. Sea X un conjunto cualquiera, y $(\mathbb{V}, \mathbb{K}, +, .)$ un espacio vectorial. Consideramos el conjunto $\mathcal{F}$ formado por todas las funciones de X que toman valores en $\mathbb{V}$. Es decir,

$$\mathcal{F} = \{f : \ f : X \to \mathbb{V}\}.$$

Definimos

- SUMA DE DOS FUNCIONES:
$$(f + g)(x) = f(x) + g(x), \quad x \in X;$$
- PRODUCTO DE UNA FUNCIÓN f POR UN ESCALAR λ a través de la fórmula
$$(\lambda f)(x) = \lambda f(x), \quad x \in X.$$

Mostrar que $(\mathcal{F}, \mathbb{K}, +, \cdot)$ es un espacio vectorial. ♠

Vamos ahora a un ejemplo en el que discutimos la estructura de espacio vectorial en el conjunto de los polinomios.

Ejemplo 4.1.9. LOS POLINOMIOS COMO ESPACIO VECTORIAL
Consideremos el conjunto de todos los polinomios con coeficientes reales. Es decir, consideramos las expresiones del tipo

$$\alpha_n x^n + \alpha_{n-1} x^{n-1} + \ldots + \alpha_1 x + \alpha_0, \quad n \in \mathbb{N}, \ \alpha_i \in \mathbb{R}, i = 1, 2, \ldots, n.$$

Cuando $\alpha_n \neq 0$, decimos que el polinomio tiene *grado* n. Al conjunto de todos lo polinomios con coeficientes reales lo notamos como $\mathbb{R}[x]$. En $\mathbb{R}[x]$ están definidas operaciones de suma de polinomios y de producto por un escalar de la siguiente manera.

Para la operación de suma consideremos dos polinomios cualesquiera, de grados m y n,

$$\begin{aligned} p(x) &= \alpha_n x^n + \alpha_{n-1} x^{n-1} + \ldots + \alpha_1 x + \alpha_0 \\ q(x) &= \beta_m x^m + \beta_{m-1} x^{m-1} + \ldots + \beta_1 x + \beta_0. \end{aligned}$$

Si tenemos $m < n$ podemos completar los términos que faltan en $q(x)$ con coeficientes nulos

$$\beta_n = \beta_{n-1} = \ldots = \beta_{m+1} = 0$$

y definimos la suma de p y q por la expresión

$$(p + q)(x) = (\alpha_n + \beta_n) x^n + (\alpha_{n-1} + \beta_{n-1}) x^{n-1} + \ldots + (\alpha_1 + \beta_1) x + \alpha_0 + \beta_0.$$

Definimos el producto de un número $\lambda \in \mathbb{R}$ por el polinomio $p(x)$ como

$$(\lambda p)(x) = \lambda \alpha_n x^n + \lambda \alpha_{n-1} x^{n-1} + \ldots + \lambda \alpha_1 x + \lambda \alpha_0.$$

Con estas operaciones $(\mathbb{R}[x],\ \mathbb{R},\ +,\ \cdot)$ es un $\mathbb{R}$-espacio vectorial. Dado un número natural n podemos restringir las operaciones de suma y producto por un escalar que hemos definido en todo $\mathbb{R}[x]$ al conjunto de los polinomios reales cuyo grado es menor o igual que n, al que designaremos $\mathbb{R}_n[x]$. En efecto, la suma de dos polinomios de grado menor o igual que n es un nuevo polinomio cuyo grado no supera a n, y el producto de un polinomio de grado menor o igual que n por un escalar es también un polinomio de grado menor o igual que n. Por tanto, la suma y el producto por un escalar son operaciones en $\mathbb{R}_n[x]$. Con estas operaciones $(\mathbb{R}_n[x],\ \mathbb{R},\ +, \cdot)$ es también un espacio vectorial.

Observación 4.1.10. En forma análoga podemos definir el $\mathbb{K}$-espacio vectorial $(\mathbb{K}[x],\ \mathbb{K},\ +,\ \cdot)$, formado por los polinomios con coeficientes en $\mathbb{K}$, y el $\mathbb{K}$-espacio vectorial $(\mathbb{K}_n[x],\ \mathbb{K},\ +,\ \cdot)$, que contiene a los elementos de $\mathbb{K}[x]$ cuyo grado no supera a n.

Sin embargo, si consideramos polinomios con un grado fijo, digamos n, este conjunto no forma un espacio vectorial con las operaciones de suma y producto que estamos considerando, porque la suma de dos polinomios de grado n puede dar lugar a un nuevo polinomio cuyo grado es estrictamente menor a n.

Ejercicio 4.3. Completar los detalles del ejemplo, explicando por qué $\mathbb{R}[x]$ y $\mathbb{R}_n[x]$ son espacios vectoriales, y por qué no lo es el conjunto de los polinomios de grado exactamente igual a n. Extender todo lo hecho a polinomios sobre un cuerpo $\mathbb{K}$ cualquiera.

Observación 4.1.11. SUBESPACIOS VECTORIALES
Es obvio que $\mathbb{R}_n[x]$ está incluido en $\mathbb{R}[x]$. Éste es un ejemplo en el que un espacio vectorial está contenido en otro espacio más grande. Expresaremos este hecho diciendo que $\mathbb{R}_n[x]$ es un *subespacio vectorial* de $\mathbb{R}[x]$. En la sección 4.2 desarrollaremos con bastante detalle esta noción de subespacio vectorial. ♠♣

Ejemplo 4.1.12. Un caso particular de espacio $(\mathbb{K}^n, \mathbb{K}, +, \cdot)$ es el que se obtiene cuando se escoge $\mathbb{K} = \mathbb{C}$. Naturalmente, se trata de un $\mathbb{C}$-espacio vectorial, o espacio vectorial complejo. Pero también es posible considerar el espacio vectorial real $(\mathbb{C}^n, \mathbb{R}, +, \cdot)$. Tenemos el mismo el conjunto de vectores, las operaciones definidas por las reglas usuales con las que operamos con listas de números, pero consideramos sólo números reales como el conjunto de escalares. Éste es un nuevo espacio vectorial, un espacio vectorial real, **diferente** del espacio $(\mathbb{C}^n, \mathbb{C}, +, \cdot)$ porque está construido sobre un conjunto diferente de escalares.

Verificar que $(\mathbb{C}^n, \mathbb{R}, +, \cdot)$ es un espacio vectorial se reduce a una simple observación: las propiedades de la suma de vectores siguen satisfaciéndose en esta nueva estructura, porque el cambio del conjunto de escalares no las afecta; las propiedades del producto por un escalar se satisfacen para todo los complejos, lo que asegura que son válidas cuando trabajamos sobre el conjunto de los números reales, ya que $\mathbb{R} \subset \mathbb{C}$.

Ejercicio 4.4.

1. Mostrar que $(\mathbb{R}, \mathbb{Q}, +, .)$ es un espacio vectorial. Las operaciones $+$ y $.$ son la suma y el producto usuales entre números.
2. Consideremos el subconjunto de los números reales
$$\mathbb{Q}\big[\sqrt{2}\big] = \{a + b\sqrt{2};\ a, b \in \mathbb{Q}\}.$$
Mostrar $(\mathbb{Q}\big[\sqrt{2}\big], \mathbb{Q}, +, .)$ es un espacio vectorial[6]. Las operaciones son las usuales.
3. Sea $(\mathbb{V}, \mathbb{K}, +, \cdot)$ un espacio vectorial y $\mathbb{K}'$ un subcuerpo[7] de $\mathbb{K}$. Mostrar que $(\mathbb{V}, \mathbb{K}', +, \cdot)$ es un espacio vectorial. ♣

Luego de esta excursión por las sutilezas de cambiarle el cuerpo a los espacios vectoriales familiares para transformarlos en nuevos objetos, vamos a presentar el espacio vectorial más pequeño que se puede tener: un espacio que tiene al vector nulo como su único elemento.

Ejemplo 4.1.13. El espacio vectorial trivial
Consideremos el conjunto $\{O\}$ formado por el elemento nulo O de un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ cualquiera. La suma está definida en $\{O\}$ porque $O + O = O$, y también el producto por un escalar cualquiera λ porque siempre $\lambda O = O$. Es bastante inmediato verificar que $(\{O\}, \mathbb{K}, +, \cdot)$ satisfacen las propiedades de la estructura de espacio vectorial.

Ejercicio 4.5. Hacer la verificación de que $(\{O\}, \mathbb{K}, +, \cdot)$ es un espacio vectorial.

Estos espacios triviales son los espacios vectoriales más chicos que pueden existir, porque cualquier espacio vectorial debe contener al menos el elemento neutro, es todo lo que hay en los espacios vectoriales triviales que estamos considerando. Observemos además que $\{O\} \subset \mathbb{V}$, este es otro ejemplo de subespacio vectorial.

[6] $\mathbb{Q}\left[\sqrt{2}\right]$ es también un cuerpo. Ver el apéndice A.2.

[7] Un subcuerpo de un cuerpo $\mathbb{K}$ es un subconjunto $\mathbb{K}'$ del cuerpo que es a su vez un cuerpo cuando se consideran en $\mathbb{K}'$ las restricciones a $\mathbb{K}'$ de las operaciones de suma y producto que están definidas en todo $\mathbb{K}$. Por ejemplo, los números racionales son un subcuerpo de los reales. Porque están contenidos en los reales, y porque junto con las operaciones usuales de suma y producto, que no son otra cosa que las operaciones de $\mathbb{R}$ aplicadas al caso particular de los números racionales, forman un cuerpo. Por otra parte, $\mathbb{R}$ es un subcuerpo de los números complejos.

Ejercicio 4.6. Investigar si $(\mathbb{R}^2, \mathbb{R}, +, \cdot)$ es un espacio vectorial en caso de que las operaciones de suma y producto se definan de las siguiente maneras, para todo x_1, x_2, y_1, y_2 y λ reales:

1. $(x_1, y_1) + (x_2, y_2) = (3y_1 + 3y_2, -x_1 - x_2), \quad \lambda(x_1, y_1) = (3\lambda y_1, x_1);$
2. $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \quad \lambda(x_1, y_1) = (\lambda x_1, 0);$
3. $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, 0), \quad \lambda(x_1, y_1) = (\lambda x_1, \lambda y_1);$
4. $(x_1, y_1) + (x_2, y_2) = (x_1, x_2 + y_2), \quad \lambda(x_1, y_1) = (\lambda x_1, \lambda y_1);$
5. $(x_1, y_1) + (x_2, y_2) = (|x_1 + x_2|, |y_1 + y_2|), \quad (x_1, y_1) = (|\lambda x_1|, |\lambda y_1|).$

4.1.3. Propiedades básicas de la estructura de espacio vectorial

Edificaremos la teoría de espacios vectoriales a partir de las **propiedades de la suma y producto por un escalar listadas** en la definición 4.1, junto a las propiedades de la estructura de cuerpo del conjunto de escalares.

En esta sección, mostramos que valen en general algunas propiedades que son más o menos obvias en los ejemplos como $\mathbb{K}^n$, matrices, etcétera, pero que sólo podremos considerar válidas en un espacio vectorial cualquiera una vez que las hayamos deducido de los postulados contenidos en la definición. Subrayemos que el hecho de **que una propiedad se cumpla en un ejemplo particular no asegura su validez en cualquier espacio vectorial**. Sin embargo, **si mostramos alguna propiedad utilizando sólo las propiedades contenidas en la definición 4.1, u otras propiedades obtenidas a partir de la definición, estaremos seguros que vale en todo espacio vectorial**.

Nuestra primera proposición asegura que sólo hay un neutro para la suma. La definición de espacio vectorial solo requiere explícitamente que **exista** un neutro, no que haya solamente uno. Pero las propiedades que aparecen en la definición implican la unicidad de tal neutro.

Proposición 4.1 (Unicidad del neutro). *Sea $(\mathbb{V}, \mathbb{K}, +, \cdot)$ un espacio vectorial. Entonces el elemento neutro de $\mathbb{V}$ es único.*

PRUEBA. Supongamos que O_1 y O_2 son dos elementos de un espacio vectorial que satisfacen la propiedad de ser neutro para la suma. Como O_1 es neutro la suma $O_1 + O_2$ es igual a O_2. Por otra parte, como O_2 es neutro la suma $O_1 + O_2$ es igual a O_1. Es decir

$$O_1 = O_1 + O_2 = O_2,$$

por lo que los dos elementos son en realidad el mismo y el neutro sólo puede ser único. Observe que lo que utilizamos fue la función que se le asigna al neutro de la suma. □

Algo similar ocurre con el opuesto de un vector $v \in \mathbb{V}$: la definición 4.1 no requiere su unicidad, pero ésta se deriva de los axiomas de la teoría.

Proposición 4.2 (Unicidad del opuesto). *Sea* $(\mathbb{V}, \mathbb{K}, +, \cdot)$ *un espacio vectorial. Entonces cada elemento* $v \in \mathbb{V}$ *tiene un único opuesto.*

Prueba. Si $(-v)_1$ y $(-v)_2$ son dos opuestos de algún elemento $v \in \mathbb{V}$ entonces

$$((-v)_1 + v) + (-v)_2 = O + (-v)_2 = (-v)_2.$$

Pero, utilizando la propiedad asociativa de la suma para agrupar de una manera diferente los sumandos en el primer término de esta serie de igualdades:

$$(-v)_1 + (v + (-v)_2) = (-v)_1 + O = (-v)_1.$$

Concluimos entonces $(-v)_1 = (-v)_2$. □

Observación 4.1.14. Un poco de notación
Al saber que el opuesto de un vector es único tiene sentido considerar la resta de un vector: es la suma de su opuesto. Para u y v vectores cualesquiera en un espacio vectorial $\mathbb{V}$ indicaremos la suma $u + (-v)$ de u con el opuesto de v con la notación $u - v$. Obviamente, se satisface para todo v en el espacio que $v - v = O$.

Ejercicio 4.7. Mostrar que $u - (-v) = u + v$. También que $-(u - v) = v - u$.

Sigamos nuestro camino mostrando propiedades que son obvias en prácticamente todos los ejemplos particulares, pero que deben ser demostradas al hacer una presentaciónb axiomática de la teoría. **Las únicas propiedades de los espacios vectoriales que por ahora tenemos a nuestra disposición, son las que se recogen en la definición 4.1, y en las proposiciones 4.1, y 4.2 que acabamos de demostrar**.

Proposición 4.3 (Cero por un vector da cero). *En todo espacio vectorial* $(\mathbb{V}, \mathbb{K}, +, \cdot)$ *se cumple para cualquier* $v \in \mathbb{V}$ *que* $0v = O$.

Prueba. Como $0 = 0 + 0$ tenemos que

$$0v = (0 + 0)v = 0v + 0v,$$

donde usamos la propiedad distributiva respecto a escalares para el producto entre escalares y vectores. Consideremos el opuesto $-(0v)$ del producto $0v$.

Sumando este opuesto a ambos miembros de la igualdad que acabamos de obtener resulta

$$O = 0v - (0v) = 0v + 0v - (0v) = 0v,$$

y así demostramos la proposición. $\square$

Ejercicio 4.8. Demostrar que el producto λO de un escalar $\lambda \in \mathbb{K}$ cualquiera con el vector nulo O es igual al vector nulo.

Para el producto de un escalar por un vector en los espacios vectoriales vale una propiedad que es formalmente parecida a la siguiente propiedad de los números en un cuerpo: si $ab = 0$ entonces $a = 0$ o $b = 0$. La enunciamos en nuestra próxima proposición.

Proposición 4.4. *En un espacio vectorial* $(\mathbb{V}, \mathbb{K}, +, \cdot)$*, sean* $v \in \mathbb{V}$ *y* $\lambda \in \mathbb{K}$. *Entonces* $\lambda v = O$ *si y sólo si* $\lambda = 0$ *o* $v = O$.

Prueba. Comencemos por demostrar que $\lambda v = O$ implica $\lambda = 0$ o $v = O$. Si $\lambda = 0$ no tenemos nada que probar. Consideremos ahora el caso en que $\lambda \neq 0$. Entonces λ tiene un inverso λ^{-1} tal que $\lambda^{-1}\lambda = 1$. Multiplicamos ambos miembros de la igualdad $\lambda v = O$ por λ^{-1} y obtenemos

$$\lambda^{-1}(\lambda v) = \lambda^{-1}O = O.$$

Empleando la propiedad asociativa para el producto por un escalar conseguimos

$$1v = \left(\lambda^{-1}\lambda\right) v = O.$$

Por último, como la unidad 1 del cuerpo actúa como unidad en el espacio vectorial resulta $v = 1v = O$, lo que completa la prueba de $\lambda v = O$ implica $\lambda = 0$ o $v = O$.

Recíprocamente, la proposición 4.3 y el ejercicio 4.8 implican que el producto λv es el vector nulo cuando $\lambda = 0$ o $v = O$. $\square$

Cerramos esta serie de proposiciones, destinadas a enunciar y demostrar algunas propiedades básicas, con el hecho que el opuesto de un vector v es igual a multiplicar el vector por -1, el opuesto de la unidad en el cuerpo de escalares.

Proposición 4.5. *En todo espacio vectorial* $(\mathbb{V}, \mathbb{K}, +, \cdot)$ *se cumple para cualquier* $v \in \mathbb{V}$ *que* $-v = (-1)v$.

Prueba. Como el opuesto $-v$ es único (proposición 4.2) basta ver que $(-1)v$ satisface la propiedad $(-1)v + v = O$ que caracteriza al opuesto de v. Calculamos entonces

$$(-1)v + v = (-1)v + 1v = (1 - 1)v = 0v = O,$$

lo que prueba nuestra afirmación. □

Vale la pena observar que en la demostración de la proposición 4.5 nos hemos empleado pura y exclusivamente los postulados básicos de las estructura de espacio vectorial, y proposiciones que previamente habíamos deducido de esos postulados. Usamos:

- la propiedad de que la unidad del cuerpo actúa como unidad en el espacio vectorial;
- la propiedad distributiva respecto a la suma de escalares;
- la proposición 4.3.

Ejercicio 4.9. Hacer un análisis similar para determinar cuáles fueron las propiedades o proposiciones usadas en las demostraciones de las proposiciones 4.1, 4.2 y 4.3

4.1.4. Para tener presente

- La definición de espacio vectorial es una abstracción, que recoge como postulados básicos de una teoría axiomática las propiedades fundamentales, comunes a una gran variedad de ejemplos.
- Aunque es corriente referirse a un conjunto $\mathbb{V}$ formado por elementos (vectores) que puede sumarse entre sí y multiplicarse por escalares (números en un cuerpo $\mathbb{K}$) como *el espacio vectorial* $\mathbb{V}$, en realidad el espacio vectorial está constituido por toda la estructura $(\mathbb{V}, \mathbb{K}, +, \cdot)$, en la que se especifica cuáles son los vectores, cuáles son los escalares, y cuáles son las relaciones que se establecen entre ellos por las operaciones de suma y producto.
- La noción de espacio vectorial es una noción abstracta. No hace alusión en principio a ninguna realización concreta, pero admite muchas realizaciones distintas.
- Sólo son ciertas para un espacio vectorial general las proposiciones que puedan derivarse, directa o indirectamente, de los axiomas de la teoría.
- Cualquier proposición que sea válida en un espacio vectorial se satisface para cualquier ejemplo concreto de espacio vectorial, y puede emplearse para comprender mejor el ejemplo que se esté tratando. En particular, cuando se discuten las propiedades de un espacio vectorial puede parecer que no se hace referencia a nada, cuando en realidad se está haciendo

referencia en una sola vez a todos los posibles espacios vectoriales. A todos los ejemplos existentes, y a todos los que se puedan introducir en el futuro. Un teorema abstracto referido a la estructura de espacio vectorial ahorra el penoso procedimiento de verificar el teorema una y otra vez en cada caso particular.

- La presentación abstracta de la teoría permite identificar cuáles son las propiedades básicas, importantes, en muchos ejemplos, sin distraer nuestra atención con los detalles de cada ejemplo. En este sentido, la presentación axiomática proporciona una manera de comprender la esencia común de muchos fenómenos aparentemente disímiles.

4.2. Subespacios vectoriales

En esta sección reencontraremos en el contexto general de los espacios vectoriales sobre un cuerpo $\mathbb{K}$ la noción de subespacio vectorial con la que ya hemos trabajado en $\mathbb{K}^n$.

4.2.1. La noción general de subespacio de un espacio vectorial

Cuando estudiamos los sistemas lineales de ecuaciones fueron introducidos varios subconjuntos de $\mathbb{K}^n$ asociados a las matrices, y con "buenas propiedades" respecto a la estructura lineal de $\mathbb{K}^n$. Por ejemplo, los espacios de columnas y de filas de una matriz, y el núcleo. Estos subconjuntos tienen la propiedad de que al sumar dos de sus elementos, o al multiplicar uno por un escalar cualquiera del cuerpo $\mathbb{K}$, se obtiene otro elemento del mismo subconjunto. Esta característica permite trabajar con la estructura lineal **dentro** de cada uno de estos subconjuntos, lo que nos llevó a introducir la noción de *subespacio vectorial* de $\mathbb{K}^n$ (ver la definición 1.5, en la página 95).

Esta noción puede extenderse a un espacio vectorial cualquiera.

Definición 4.2. *Un subconjunto* $\mathbb{S}$ *no vacío de un* $\mathbb{K}$*-espacio vectorial* $\mathbb{V}$ *es un* ***subespacio vectorial de*** $\mathbb{V}$ *si tiene las siguientes dos propiedades:*

1. *la suma de dos vectores cualesquiera de* $\mathbb{S}$ *pertenece a* $\mathbb{S}$*;*
2. *el producto de un vector de* $\mathbb{S}$ *por un escalar cualquiera en* $\mathbb{K}$ *pertenece a* $\mathbb{S}$*.*

Cuando un subconjunto $\mathbb{S}$ no vacío cumple la propiedad 1 decimos que $\mathbb{S}$ es *cerrado frente a la suma.* Análogamente, cuando se satisface 2 decimos que es *cerrado frente a la multiplicación por un escalar* . Eso nos permite resumir la definición 4.2 diciendo que **un subespacio de un espacio vectorial es un subconjunto no vacío y cerrado frente a las operaciones de suma de vectores y producto de escalares por vectores**.

Disponemos ahora de la noción general, bastante flexible, de espacio vectorial que hemos introducido en la definición 4.1. Esta definición engloba todas las situaciones en los que los elementos de un conjunto pueden sumarse entre sí y multiplicarse por escalares, con operaciones que satisfacen las propiedades que se requieren en la definición. En particular, como veremos en el enunciado y la demostración de nuestro próximo teorema, se aplica a los subespacios vectoriales de los espacios vectoriales.

Teorema 4.1. *Sean* $(\mathbb{V}, \mathbb{K}, +, \cdot)$ *un espacio vectorial, y* $\mathbb{S} \subset \mathbb{V}$ *un subespacio vectorial de* $\mathbb{V}$*. Entonces* $(\mathbb{S}, \mathbb{K}, +, \cdot)$ *es un espacio vectorial sobre el cuerpo* $\mathbb{K}$*.*

Observación 4.2.1. Suma y producto en $\mathbb{S}$.
Las operaciones de suma y producto $+$ y $\cdot$ que aparecen en $(\mathbb{S}, \mathbb{K}, +, \cdot)$ son en algún sentido las mismas que teníamos en $\mathbb{V}$, y en algún sentido diferentes.

Son las mismas porque, por ejemplo, para dos vectores v y w de $\mathbb{S}$ la suma $v + w$ en $\mathbb{S}$ arroja como resultado el mismo vector $v + w$ que la suma en $\mathbb{V}$. Es decir, la suma se calcula en $\mathbb{S}$ con el mismo procedimiento que en $\mathbb{V}$.

Ejemplo 4.2.2. Si estamos trabajando en el espacio vectorial $\mathbb{K}^n$ con la suma de vectores que se define sumando coordenada a coordenada, dentro de los subespacios de $\mathbb{K}^n$ también sumaremos coordenada a coordenada. ♣

Una consideración similar vale para el producto por un escalar, que se calcula en $\mathbb{S}$ siguiendo las mismas reglas que se usan en todo $\mathbb{V}$. O sea, desde el punto de vista del cálculo las operaciones $+$ y $\cdot$ son exactamente las mismas que teníamos antes.

Pero las operaciones en algún sentido son diferentes, porque las aplicamos sobre un dominio diferente. Recordemos que la suma $+$ y el producto $\cdot$ en $\mathbb{V}$ son funciones

$$+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}, \quad \cdot : \mathbb{K} \times \mathbb{V} \to \mathbb{V},$$

con dominios $\mathbb{V} \times \mathbb{V}$ y $\mathbb{K} \times \mathbb{V}$ respectivamente, y codominio $\mathbb{V}$. Cuando consideramos la suma $+$ y el producto $\cdot$ en $\mathbb{S}$ estamos en realidad considerando las funciones

$$+ : \mathbb{S} \times \mathbb{S} \to \mathbb{S}, \quad \cdot : \mathbb{K} \times \mathbb{S} \to \mathbb{S},$$

que son una *restricción* a $\mathbb{S}$ de las operaciones en $\mathbb{V}$. Como $\mathbb{S}$ es un subespacio podemos cambiar el codominio $\mathbb{V}$ por $\mathbb{S}$, porque la definición 4.2 requiere que $\mathbb{S}$ sea cerrado frente a la suma de vectores y el producto por un escalar.

Desde este punto de vista más formal las operaciones son en realidad diferentes, y al designarlas con el mismo símbolo $+$ y $\cdot$ estamos cometiendo un abuso de notación. El abuso de notación es cómodo e inofensivo, de modo que lo utilizaremos permanentemente. ♠

Prueba del teorema 4.1. El conjunto $\mathbb{S}$ no es vacío, porque $\mathbb{S}$ es un subespacio. Los comentarios de la observación 4.2.1 muestran que las operaciones $+$ y $\cdot$ están definidas en $\mathbb{S}$.

La suma cumple las propiedades conmutativa y asociativa porque estas se cumplen para todos los vectores de $\mathbb{V}$. En particular, se cumplen para los vectores de $\mathbb{S}$.

Veamos ahora que existe un neutro para la suma en $\mathbb{S}$. Para ello mostraremos que el vector nulo O de $\mathbb{V}$ necesariamente pertenece a $\mathbb{S}$. Como $\mathbb{S}$ es no vacío existe algún $s_0 \in \mathbb{S}$. Por la proposición 4.8 sabemos que

$$O = 0s_0.$$

En consecuencia, O pertenece a $\mathbb{S}$ porque es el producto del escalar 0 por el elemento s_0 de $\mathbb{S}$, y $\mathbb{S}$ es un subespacio vectorial de $\mathbb{V}$. Como O está en $\mathbb{S}$ y satisface

$$O + s = s + O = s$$

para todo $s \in \mathbb{V}$, en particular para todo $s \in \mathbb{S}$, es el neutro para la suma de $\mathbb{S}$.

La demostración de que cada $s \in \mathbb{S}$ tiene opuesto es similar. Sabemos que existe $-s \in \mathbb{V}$, opuesto de $s \in \mathbb{V}$, que cumple $s + (-s) = O$. Veamos que es también opuesto de s en $\mathbb{S}$. Por la proposición 4.5

$$-s = (-1)\, s,$$

por lo tanto pertenece a $\mathbb{S}$ por ser el producto del escalar -1 por el vector $s \in \mathbb{S}$. Entonces $-s \in \mathbb{S}$ y cumple que $s + (-s) = O$, de modo que verifica las condiciones requeridas para ser opuesto de s en $\mathbb{S}$, como queríamos probar.

En cuanto a las propiedades de la función $\cdot$ en $\mathbb{S}$ requeridas en la definición 4.1 se cumplen todas en $\mathbb{S}$, puesto que, por definición, se cumplen en $\mathbb{V}$ que incluye a $\mathbb{S}$. Queda probado entonces que $\mathbb{S}$ es también un espacio vectorial. □

Vale la pena destacar que en el curso de la demostración del teorema 4.1 hemos probado el siguiente corolario.

Corolario 4.6. *Si $\mathbb{S}$ es un subespacio vectorial de $\mathbb{V}$ entonces $O \in \mathbb{S}$.*

El teorema 4.1 nos dice entonces que **un subespacio vectorial es un espacio vectorial que está metido dentro de un subespacio más grande**, una idea que apareció cuando tratamos el ejemplo del espacio $\mathbb{K}_n[x]$, formado por los polinomios de grado menor o igual que n sobre un cuerpo $\mathbb{K}$ (ver la observación 4.1.11, en la página 406—), y en el caso del espacio trivial que contiene únicamente al vector nulo (ver el ejemplo 4.1.13, en la página 407). Vamos a verificar ahora que estos subconjuntos son subespacios vectoriales de los espacios que los contienen.

Ejemplo 4.2.3. Comenzaremos mostrando que para cada $n \in \mathbb{N}$ el conjunto $\mathbb{K}_n[x]$ es un subespacio vectorial de $\mathbb{K}[x]$. Para empezar, notemos que el conjunto de los polinomios de grado menor o igual que n es no vacío. Por ejemplo, los polinomios a_0, donde a_0 es cualquier número en el cuerpo $\mathbb{K}$, pertenecen a $\mathbb{K}_n[x]$ para cualquier valor de n.

La suma de dos polinomios de grado menor o igual que n es un polinomio cuyo grado no supera a n. El grado de la suma es igual a n, salvo en el caso en que los términos de orden n de los sumandos se cancelan, lo que da lugar a un polinomio de grado estrictamente menor que n.

Por último, el producto de un escalar distinto de 0 por un polinomio cualquiera produce un polinomio del mismo grado. El producto de 0 por un polinomio es igual al polinomio nulo, de grado 0. En ambos casos se obtiene un polinomio cuyo grado es menor o igual que el del polinomio original.

Consideremos ahora el caso del subconjunto $\mathbb{S} = \{O\}$, contenido en un $\mathbb{K}$-espacio vectorial $\mathbb{V}$. La definición de $\mathbb{S}$ muestra inmediatamente que es no vacío. Sólo hay que considerar un elemento de $\mathbb{S}$, el O, para determinar si $\mathbb{S}$ es cerrado para la suma y el producto por un escalar. Tenemos

$$O + O = O \in \mathbb{S}, \qquad \lambda O = O \in \mathbb{S}, \ \forall \lambda \in \mathbb{K},$$

lo que completa la demostración de que $\mathbb{S}$ es un subespacio vectorial. ♣

El subespacio vectorial $\{O\}$ es un subespacio obvio de cualquier espacio vectorial $\mathbb{V}$. El espacio $\mathbb{V}$ tiene otro subespacio bastante obvio, que es el propio $\mathbb{V}$. Resumimos estos comentarios en el próximo ejemplo.

Ejemplo 4.2.4. SUBESPACIOS TRIVIALES
En todo espacio vectorial existen dos subespacios que son los llamados *subespacios triviales*. Ellos son

- $\mathbb{S} = \{O\}$, que está formado por un único elemento: el vector nulo del espacio.
- $\mathbb{S} = \mathbb{V}$.

Ejercicio 4.10. Completar el ejemplo, mostrando que $\mathbb{V}$ es un subespacio de $\mathbb{V}$. ♣

Observación 4.2.5. SUBESPACIOS E INFORMACIÓN
Un punto de vista bastante sugerente acerca de la noción de subespacio vectorial es pensar que las condiciones de pertenencia a algún subespacio vectorial $\mathbb{S}$ estrictamente incluido en $\mathbb{V}$ nos ofrecen alguna información adicional[8] sobre

[8] En realidad esta afirmación es cierta para cualquier subconjunto $\mathbb{S}$ estrictamente incluido en $\mathbb{V}$, aunque $\mathbb{S}$ no sea un subespacio vectorial de $\mathbb{V}$, y en muchas situaciones el lector encontrará conjuntos destacados de un espacio vectorial que no son subespacios vectoriales del espacio original. Por ejemplo, la superficie de una esfera en $\mathbb{R}^3$ es uno de estos conjuntos. Pero cuando $\mathbb{S}$ es un subespacio vectorial podemos explotar todas las propiedades de los espacios vectoriales. También, en muchos casos ocurre que las propiedades adicionales, válidas en el subespacio $\mathbb{S}$ se combinan armoniosamente con la propiedades de espacio vectorial de $\mathbb{S}$ para dar lugar a una estructura más rica.

los elementos de $\mathbb{S}$. Veamos esto en un ejemplo, en el que dar cierta información adicional sobre los elementos de un espacio vectorial selecciona un subespacio especialmente interesante.

Podemos considerar el espacio vectorial $\mathcal{F}$ formado por todas las funciones reales de variable real. En la observación 4.1.7 establecimos que $\mathcal{F}$ es un espacio vectorial real. Un subconjunto especialmente interesante de $\mathcal{F}$ es el que está formado por las funciones continuas, al que llamaremos $C(\mathbb{R})$.

Ejemplo 4.2.6. El espacio vectorial de las funciones reales continuas
El conjunto $C(\mathbb{R})$ formado por todas las funciones $f : \mathbb{R} \to \mathbb{R}$ continuas no sólo es un subconjunto del espacio vectorial $\mathcal{F}$. Es además un subespacio vectorial de $\mathcal{F}$. En efecto, $C(\mathbb{R})$ es no vacío, la suma de dos funciones continuas es continua, y el producto de un número real por una función continua define una nueva función continua.

Como $C(\mathbb{R})$ es un subespacio vectorial de $\mathcal{F}$ el teorema 4.1 implica que **el conjunto de las funciones reales de variable real que son continuas forma un espacio vectorial real**. ♣

Para cerrar estos comentarios digamos que el conjunto $\mathcal{F}$ es demasiado grande para casi cualquier propósito. El lector encontrará que el cálculo y el análisis matemático trabajan siempre sobre subespacios vectoriales de $\mathcal{F}$ caracterizados por una o más propiedades adicionales: continuidad, derivabilidad, integrabilidad, acotación, etcétera. ♠

4.2.2. Ejemplos de subespacios vectoriales

El primer ejemplo retoma nociones ya conocidas.

Ejemplo 4.2.7. El núcleo de una matriz $m \times n$ sobre $\mathbb{K}$ es un subespacio vectorial de $\mathbb{K}^n$. Por lo tanto, los elementos de $\mathbb{K}^n$ que satisfacen un conjunto de ecuaciones lineales homogéneas forman un subespacio vectorial.

Por ejemplo, el espacio

$$\mathbb{S} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \ ; \ x_1 + \ldots + x_n = 0\}$$

de las n-úplas de números reales cuyos términos suman cero es un subespacio vectorial, porque es el núcleo de la matriz

$$(1, 1, \ldots, 1),$$

de dimensiones $1 \times n$. Por lo tanto $\mathbb{S}$ es un espacio vectorial.

Ejemplo 4.2.8. El subconjunto $\mathbb{S} \subset \mathbb{K}^n$ formado por las soluciones de una ecuación lineal $AX = B$ no homogénea, donde B es un vector no nulo, no es un subespacio vectorial. Observemos que el vector nulo O de $\mathbb{K}^n$ no es una solución de $AX = B$, por lo tanto no está en $\mathbb{S}$. Por otra parte, el corolario 4.6 asegura que el vector nulo O de $\mathbb{K}^n$ necesariamente está en cualquier subespacio, lo que excluye la posibilidad de que $\mathbb{S}$ sea un subespacio vectorial de $\mathbb{K}^n$.

Las consideraciones anteriores muestran, por ejemplo, que el conjunto

$$\mathbb{S} = \{(z_1, \dots, z_n) \in \mathbb{R}^n \ ; \ z_1 = 5 + 2i\}$$

no es un subespacio vectorial de $\mathbb{C}^n$. Algo que puede mostrarse directamente a partir del hecho de que la suma de dos vectores

$$Z = (z_1, \dots, z_n), \quad W = (w_1, \dots, w_n),$$

en el conjunto $\mathbb{S}$ tiene como primera componente el número complejo

$$z_1 + w_1 = (5 + 2i) + (5 + 2i) = 10 + 4i,$$

que es, a todas luces, distinto de $5 + 2i$. ♣

Ejemplo 4.2.9. El conjunto de las funciones reales de clase C^1, es decir, formado por las funciones reales derivables cuya derivada es una función continua, es un espacio vectorial. Lo designaremos con la notación $C^1(\mathbb{R})$. La verificación es un rápido repaso de las propiedades del cálculo de derivadas.

Es claro que este conjunto es no vacío, ya que están en él, por ejemplo, todas las constantes, los polinomios, la exponencial y las funciones seno y coseno. Si dos funciones f y g tienen derivadas continuas f' y g', su suma $f + g$ tiene derivada $f' + g'$ que también es continua, porque es la suma de las dos funciones continuas f' y g'. Por lo tanto $C^1(\mathbb{R})$ es cerrado frente a la suma. Si $\alpha \in \mathbb{R}$ entonces $(\alpha f)' = \alpha f'$, lo que muestra que $C^1(\mathbb{R})$ es cerrado frente al producto por escalares.

Por lo tanto $C^1(\mathbb{R})$ es un subespacio vectorial del espacio vectorial $C(\mathbb{R})$ formado por todas las funciones continuas. En consecuencia, también es un espacio vectorial real.

♣

Observación 4.2.10. Nuevos espacios dentro de viejos espacios
El procedimiento de mostrar que un conjunto $\mathbb{S}$, en el que están definidas operaciones $+$ de suma y $\cdot$ de producto con los escalares de un cuerpo $\mathbb{K}$, es

un subespacio vectorial de algún $\mathbb{K}$-espacio vectorial conocido es muchas veces el camino más fácil para mostrar que $\mathbb{S}$ es un $\mathbb{K}$-espacio vectorial.

En lugar de verificar todas las condiciones de la definición 4.1, alcanza ver que $\mathbb{S}$ está incluido en $\mathbb{V}$ es cerrado con respecto a la suma y el producto de un escalar por un vector, para luego aplicar el teorema 4.1.

El subespacio "hereda"la estructura vectorial del espacio mayor que lo contiene.

Observemos además que el resultado del teorema 4.1 nos permite aplicar en cualquier subespacio cualquier proposición válida en un espacio vectorial. ♠

En nuestro próximo ejemplo volveremos a utilizar las ideas que destacamos en la observación 4.2.10.

Ejemplo 4.2.11. FUNCIONES LINEALES A TROZOS
Las funciones lineales a trozos que usamos para tratar un problema de interpolación en el ejemplo 4.1.1, página 396, también forman un espacio vectorial. Mostraremos que constituyen un subespacio vectorial del conjunto de todas las funciones reales definidas en $[a, b]$.

Lo haremos en general, para las funciones lineales entre nodos

$$x_i, \quad i = 0, 1, \ldots, n,$$

donde

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

Dos funciones f y g son lineales a trozos si son continuas y en cada intervalo $[x_i, x_{i+1}]$ admiten expresiones de la forma

$$f(x) = a_i x + b_i, \ g(x) = c_i x + d_i, \quad x \in [x_i, x_{i+1}], \ i = 0, 1, \ldots, n-1,$$

donde los números a_i, b_i, c_i y d_i son constantes que sólo dependen de i, es decir, de cuál sea el subintervalo de $[a, b]$ que estemos considerando.

La suma de f y g es una función continua. También el producto αf de un número real α cualquiera por f. En cada intervalo $[x_i, x_{i+1}]$ se satisface

$$\begin{aligned}(f + g)(x) &= (a_i + c_i)x + (b_i + d_i),\\ (\alpha f)(x) &= (\alpha a_i)x + (\alpha b_i),\end{aligned}$$

lo que muestra que $f + g$ y αf admiten expresiones lineales sobre todos esos intervalos.

Por lo tanto el conjunto de funciones lineales a trozos es un subespacio vectorial del espacio de todas las funciones reales definidas sobre $[a, b]$, y, en consecuencia, un espacio vectorial real. También son un subespacio del espacio $C([a, b])$ de las funciones continuas definidas sobre $[a, b]$. Ver el ejemplo 4.2.6.

Ejemplo 4.2.12. Sea $(\mathbb{C}^n, \mathbb{R}, +, \cdot)$ el espacio vectorial de las n-uplas de complejos sobre el cuerpo de los números reales que vimos en el ejemplo 4.1.12. Consideramos el subconjunto $\mathbb{S}$ formado por las n-úplas de números complejos cuyos términos tienen parte real igual a cero. Entonces $\mathbb{S}$ es un subespacio de este espacio vectorial real.

Pero el mismo subconjunto $\mathbb{S}$ no es un subespacio del espacio vectorial complejo $(\mathbb{C}^n, \mathbb{C}, +, \cdot)$. El hecho de que $\mathbb{S}$ sea un subespacio no depende sólo de cuál sea el subconjunto $\mathbb{S}$, sino de toda la estructura que tiene el espacio vectorial que lo contiene. Aquí resulta de fundamental importancia el cuerpo sobre el que se está trabajando.

Ejercicio 4.11. Fundamente las afirmaciones de este ejemplo anterior. ♣

Ejercicio 4.12. Consideremos el espacio vectorial $\mathcal{F}$ formado por todas las funciones reales de variable real. Investigar cuáles de los siguientes subconjuntos de $\mathcal{F}$ son subespacios vectoriales:

1. para un $x_0 \in \mathbb{R}$ dado, el conjunto de las funciones f tales que $f(x_0) = 0$;
2. el conjunto de funciones f para las que existe $x_0 \in \mathbb{R}$ tal que $f(x_0) = 0$.

Ejercicio 4.13. Consideremos el espacio vectorial $\mathbb{V}$ formado por todas las sucesiones reales $\{a_n\}_{n\in\mathbb{N}}$. Investigar cuáles de los siguientes subconjuntos de $\mathcal{F}$ son subespacios vectoriales:

1. para un $n_0 \in \mathbb{N}$ dado, el conjunto de las sucesiones tales que $a_n = 0$ si $n \geq n_0$;
2. el conjunto de las sucesiones para las que existe un número natural n_0 tal que $a_n = 0$ si $n \geq n_0$;
3. el conjunto de las sucesiones para las que existe $lim_{n\to\infty} a_n$;
4. el conjunto de las sucesiones acotadas;
5. para un número $l \in \mathbb{R}$ dado, el conjunto de las sucesiones tales que
$$l = lim_{n\to\infty} a_n.$$
Discutir según l.

Para los conjuntos que sean subespacios, estudiar cuáles son las relaciones de inclusión entre ellos.

Ejercicio 4.14. Se considera el $\mathbb{K}$ espacio vectorial formado por las matrices $n \times n$ con entradas en $\mathbb{K}$. En cada caso, investigar si los subconjuntos de $M^{n\times n}(\mathbb{K})$ son subespacios vectoriales:

1. el conjunto de las matrices simétricas. Es decir, de aquellas matrices que satisfacen $A = A^t$;
2. el conjunto de las matrices antisimétricas; que satisfacen $A = -A^t$;

3. el conjunto de las matrices invertibles;
4. fijado $X \in \mathbb{K}^n$, el conjunto de matrices A tales que $X \in \ker(A)$.

Ejemplo 4.2.13. ECUACIONES DIFERENCIALES LINEALES
Las ecuaciones diferenciales son ecuaciones que relacionan a una función con sus derivadas. Aparecen muy naturalmente cuando se intenta modelar muchos fenómenos de la naturaleza y su estudio constituye una rama importante de la matemática que se sirve de la teoría que estamos desarrollando en este curso. Presentamos aquí un ejemplo de estas ecuaciones.

Sean a y b dos constantes reales. Diremos que una función $x : \mathbb{R} \to \mathbb{R}$ es una solución de la ecuación diferencial

$$x'' + ax' + bx = 0 \tag{4.2}$$

si la igualdad

$$x''(t) + ax'(t) + bx(t) = 0$$

se satisface para todo $t \in \mathbb{R}$.

Se puede pensar, aunque, por supuesto, no es el único enfoque posible, en esta ecuación en términos dinámicos, interpretando que x describe la posición de un partícula que se desplaza bajo la acción de ciertas fuerzas. En cada instante t ocupará cierta posición en el espacio, y la información sobre la trayectoria está almacenada en la función $x(t)$. La ecuación (4.2) es entonces la que gobierna el movimiento de una partícula. Ecuaciones de este tipo se obtienen como un modelo para diversos sistemas físicos, aplicando las leyes de la mecánica en situaciones particulares. Por ejemplo, si a y b son constantes positivas la ecuación (4.2) corresponde al movimiento de una partícula que está oscilando en un medio que opone resistencia a su avance. El conjunto de soluciones de la ecuación contiene todas las posibles trayectorias que nuestro modelo predice para el sistema.

Ejercicio 4.15. Mostrar que el conjunto de soluciones de la ecuación diferencial (4.2) es un espacio vectorial real. ♣

4.2.3. Operaciones con subespacios

Los subespacios de un espacio vectorial $\mathbb{V}$ son subconjuntos contenidos en él. Podemos operar con ellos como con cualquier subconjunto de cualquier conjunto: tiene sentido hablar de la intersección de dos subespacios y de su unión, también del complemento de un subespacio. También podemos operar con los subconjuntos utilizando las operaciones de suma y producto por un escalar para operar con sus elementos.

Intersecciones y uniones

Una pregunta natural en este contexto es si se obtienen nuevos subespacios realizando intersecciones y uniones de conjuntos que son subespacios de $\mathbb{V}$.

Comencemos con la intersección de dos subespacios $\mathbb{S}$ y $\mathbb{T}$ de un espacio $\mathbb{V}$. Tenemos el siguiente resultado:

Proposición 4.7. *Sean $\mathbb{S}$ y $\mathbb{T}$ subespacios de un mismo espacio $\mathbb{V}$. Entonces*

$$\mathbb{U} = \mathbb{S} \cap \mathbb{T}$$

es también un subespacio de $\mathbb{V}$.

PRUEBA. La intersección de $\mathbb{S}$ y $\mathbb{T}$ es un subconjunto de $\mathbb{V}$. El vector nulo O pertenece a $\mathbb{T}$ y a $\mathbb{S}$, por lo tanto pertenece también a su intersección, que resulta ser un subconjunto no vacío de $\mathbb{V}$.

Consideremos dos vectores v y w en $\mathbb{S} \cap \mathbb{T}$. Como ambos pertenecen a $\mathbb{S}$ y $\mathbb{S}$ es un subespacio, entonces su suma $v + w$ también está en $\mathbb{S}$. Análogamente, ambos vectores pertenecen a $\mathbb{T}$, de modo que su suma es un vector de $\mathbb{T}$. Concluimos entonces que la suma $v + w \in \mathbb{S} \cap \mathbb{T}$.

De manera similar se muestra que $\mathbb{S} \cap \mathbb{T}$ es cerrado frente al producto por escalares. Consideremos $\lambda \in \mathbb{K}$ y $v \in \mathbb{S} \cap \mathbb{T}$. Entonces v pertenece a $\mathbb{S}$ y a $\mathbb{T}$. Como $\mathbb{S}$ y $\mathbb{T}$ son subespacios el producto λv también pertenece a $\mathbb{S}$ y a $\mathbb{T}$, por lo tanto a $\mathbb{S} \cap \mathbb{T}$. $\square$

Observemos que un vector v pertenece a la intersección $\mathbb{S} \cap \mathbb{T}$ de dos subespacios $\mathbb{S}$ y $\mathbb{T}$ si satisface simultáneamente las condiciones de pertenencia a ambos. Veamos esto en un ejemplo.

Ejemplo 4.2.14. Consideremos $\mathbb{S} = C(\mathbb{R})$, que es un subespacio del espacio vectorial $\mathcal{F}$ de todas las funciones reales de variable real, tal como vimos en el ejemplo 4.2.6. La condición que debe satisfacer una función f para estar en este subespacio es ser continua.

También el conjunto $\mathbb{T}$ de las funciones reales que se anulan en un punto x_0 dado de $\mathbb{R}$ es un subespacio de $\mathcal{F}$. Éste es un resultado que ya habrá encontrado el lector que haya resuelto el ejercicio 4.12

Las funciones que están en $\mathbb{S} \cap \mathbb{T}$ son las que cumplen a la vez la condición de ser continuas y la de anularse en $x = 0$. Por la proposición 4.7 sabemos que forman un subespacio vectorial de $\mathcal{F}$, por lo tanto un espacio vectorial sobre $\mathbb{R}$. ♣

La proposición 4.7 puede generalizarse a una intersección cualesquiera de subespacios. En el próximo ejercicio pedimos mostrar ese resultado y aplicarlo.

Ejercicio 4.16. INTERSECCIÓN DE UNA COLECCIÓN DE SUBESPACIOS

1. Sea $\{\mathbb{S}_i\}_{i\in I}$ una colección de subespacios de un subespacio vectorial $\mathbb{V}$. Mostrar que la intersección $\mathbb{S} = \cap_{i\in I}\mathbb{S}_i$ de todos los subespacios es un subespacio vectorial.
2. Sean $x_0, \dots, x_n$ números reales. Mostar que el conjunto de las funciones f reales y continuas tales que $f(x_i) = 0$ para $i = 0, 1, \dots, n$ es un espacio vectorial real. En general, si $X \subset \mathbb{R}$ es un subconjunto cualquiera de $\mathbb{R}$, mostrar que las funciones reales y continuas que se anulan en todos los puntos de X forman un espacio vectorial real.

Sin embargo, la unión de subespacios no da lugar a un nuevo subespacio. Esto es fácil de ver en un ejemplo.

Ejemplo 4.2.15. Tomemos $\mathbb{V} = \mathbb{R}^2$ y los subespacios

$$\mathbb{S} = \{(x, y) \in \mathbb{R}^2;\ x = 0\}, \quad \mathbb{T} = \{(x, y) \in \mathbb{R}^2;\ y = 0\}.$$

Entonces $(0, 1) \in \mathbb{S}$ y $(1, 0) \in \mathbb{T}$ pertenecen a la unión $\mathbb{S} \cup \mathbb{T}$, pero su suma

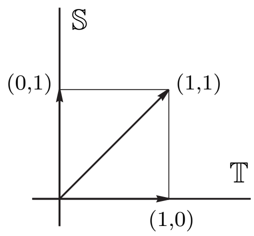


Figura 4.2: Unión de dos subespacios.

$(1, 1)$ no pertenece a $\mathbb{S} \cup \mathbb{T}$, porque no está en $\mathbb{S}$ ni en $\mathbb{T}$. La figura 4.2 ilustra este fenómeno. Por lo tanto $\mathbb{S} \cup \mathbb{T}$ no es cerrado frente a la operación de suma de vectores. ♣

Nuestro próximo ejercicio caracteriza cuándo la unión de subespacios es un subespacio, y muestra que la operación de unión no es demasiado interesante desde el punto de vista de la estructura propia de un espacio vectorial.

Ejercicio 4.17. Dados $\mathbb{S}$ y $\mathbb{T}$, subespacios de $\mathbb{V}$, mostrar que $\mathbb{S} \cup \mathbb{T}$ es un subespacio si y sólo si uno de los subespacios $\mathbb{S}$ o $\mathbb{T}$ está contenido dentro del otro.

Por último, un ejercicio sobre el complemento de un subespacio vectorial.

Ejercicio 4.18. Mostrar que el complemento de un subespacio vectorial $\mathbb{S}$ respecto al espacio vectorial $\mathbb{V}$ que lo contiene jamás es un subespacio vectorial de $\mathbb{V}$.

Sumas de conjuntos y producto por un escalar

El ejemplo 4.2.15 muestra que la unión de dos subespacios $\mathbb{S}$ y $\mathbb{T}$ no es un subespacio porque al sumar un elemento de $\mathbb{S}$ con uno de $\mathbb{T}$ logramos fabricar un vector que no está en $\mathbb{S}$ ni en $\mathbb{T}$. Este ejemplo nos llama la atención sobre la posibilidad de generar nuevos vectores sumando vectores de dos subconjuntos dados de un espacio vectorial $\mathbb{V}$, lo que tiene que ver con el hecho de que las operaciones $+$ y $\cdot$ de $\mathbb{V}$ pueden utilizarse para operar sobre subconjuntos operando sobre todos sus elementos.

Introduciremos entonces las operaciones de *suma de subconjuntos* y de *producto de un subconjunto por un escalar* de la siguiente manera.

Definición 4.3. *Sean $\mathcal{A}$ y $\mathcal{B}$ subconjuntos de un $\mathbb{K}$ espacio vectorial $\mathbb{V}$ y λ un escalar en $\mathbb{K}$. Definimos la* ***suma*** *$\mathcal{A}+\mathcal{B}$ de los subconjuntos $\mathcal{A}$ y $\mathcal{B}$ como el conjunto formado por todas las posibles sumas de elementos de $\mathcal{A}$ con elementos de $\mathcal{B}$. Es decir*

$$\mathcal{A}+\mathcal{B}=\{a+b;\ a\in\mathcal{A},\ b\in\mathcal{B}\}.$$

Definimos el ***producto*** *$\lambda\mathcal{A}$ del escalar λ por $\mathcal{A}$ como el conjunto de todos los posibles productos entre λ y elementos de $\mathcal{A}$. Por lo tanto*

$$\lambda\mathcal{A}=\{\lambda a;\ a\in\mathcal{A}\}.$$

Ejemplo 4.2.16. Consideremos

$$\mathcal{A}=\{(1,2),(3,4)\},\quad \mathcal{B}=\{(1,1),(-1,1)\},$$

en $\mathbb{R}^2$. Entonces la suma de ambos conjuntos es

$$\mathcal{A}+\mathcal{B}=\{(2,3),(0,3),(4,5),(2,5)\}.$$

El producto del escalar π por $\mathcal{A}$ es

$$\pi\mathcal{A}=\{(\pi,2\pi),(3\pi,4\pi)\}.$$

Estos resultados se obtienen aplicando directamente la definición 4.3. ♣

Para la suma de subespacios vale la siguiente proposición.

Proposición 4.8. *Sean $\mathbb{S}$ y $\mathbb{T}$ subespacios de un mismo espacio $\mathbb{V}$. Entonces la suma $\mathbb{S}+\mathbb{T}$ es también un subespacio de $\mathbb{V}$.*

Ejercicio 4.19. Demostrar la proposición 4.8.

Ejemplo 4.2.17. Consideremos $\mathbb{V} = \mathbb{R}^2$, y $\mathbb{S}$ y $\mathbb{T}$ como en el ejemplo 4.2.15. Entonces la suma $\mathbb{S} + \mathbb{T}$ es todo $\mathbb{R}^2$. En efecto, sabemos que $\mathbb{S} + \mathbb{T} \subset \mathbb{R}^2$. Por otra parte, cualquier $(x, y) \in \mathbb{R}^2$ puede descomponerse como la suma

$$(x, y) = (0, y) + (x, 0)$$

de un vector en $(0, y) \in \mathbb{S}$ y $(x, 0) \in \mathbb{T}$, lo que muestra que todo vector de $\mathbb{R}^2$ es un elemento de $\mathbb{S} + \mathbb{T}$. ♣

En el ejemplo anterior hemos descrito un subespacio "grande", como $\mathbb{R}^2$, como la suma de dos subespacios más pequeños. Esta especie de descomposición por medio de la operación de suma de subespacios de un espacio vectorial en subespacios más simples, o que brindan más información sobre los elementos que contienen, es una construcción muy interesante para la teoría del álgebra lineal y para sus aplicaciones. Volveremos sobre ella más adelante. Sin embargo, el producto de un subespacio por un escalar no es demasiado interesante, tal como se muestra en el próximo ejercicio.

Ejercicio 4.20. Mostrar que el producto $\lambda\mathbb{S}$ de un escalar λ por un subespacio $\mathbb{S}$ es igual al subespacio trivial $\{O\}$ o al propio $\mathbb{S}$. Discutir en qué casos se obtiene cada una de estas dos posibilidades.

En el siguiente ejercicio mostramos una aplicación de las operaciones de suma de conjuntos y producto por un escalar.

Ejercicio 4.21. Dados un punto $X \in \mathbb{R}^3$ y un número real $r > 0$ definimos la bola $B_r(X)$ con centro X y radio r como el conjunto de los puntos de $\mathbb{R}^3$ que están a una distancia menor que r de X. O sea

$$B_r(X) = \{Y \in \mathbb{R}^3;\ |Y - X| < r\}.$$

Mostrar que $B_r X = \{X\} + rB_1(O)$.

El significado geométrico del ejercicio anterior es que todas las bolas en $\mathbb{R}^3$ se obtienen dilatando y trasladando una bola de radio unidad centrada en el origen. Proponemos al lector un último ejercicio para jugar un poco con las operaciones que acabamos de definir sobre los subconjuntos de un espacio vectorial.

Ejercicio 4.22. Sea $\mathcal{A}$ un subconjunto cualquiera de un espacio vectorial. Decidir si la igualdad $2\mathcal{A} = \mathcal{A} + \mathcal{A}$ es verdadera o falsa. En caso de que sea falsa decidir si hay alguna relación de inclusión entre estos subconjuntos, y agregar hipótesis sobre $\mathcal{A}$ bajo las cuales la igualdad sea cierta.

Ejercicio 4.23. En cada caso, determinar si S es subespacio vectorial del espacio vectorial dado.

1. Para el espacio vectorial $\mathbb{R}^n$ considerar:
   a) $S = \{(x_1, \ldots, x_n) \in V;\ x_1 \geq 0\}$;
   b) $S = \{(x_1, \ldots, x_n) \in V;\ x_1 = x_2 = \cdots = x_n\}$;
   c) $S = \{(x_1, \ldots, x_n) \in V;\ x_1 = x_2 = \cdots = x_n = 1\}$;
   d) $S = \{(x_1, \ldots, x_n) \in V;\ x_1 + x_2 + \cdots + x_n = 1\}$;
   e) $S = \{(x_1, \ldots, x_n) \in V;\ x_1^2 + x_2^2 + \cdots + x_n^2 = 0\}$.
2. Para el espacio vectorial $\mathbb{R}_n[x]$, formado por los polinomios de grado menor o igual que n y con coeficientes reales, considerar:
   a) $S = \{p \in \mathbb{R}_n[x];\ p(1 - x) = p(1 + x)\ \forall x \in \mathbb{R}\}$;
   b) $S = \{p \in \mathbb{R}_n[x];\ p(\alpha) = 0\}$, donde $\alpha \in \mathbb{R}$ es un valor fijo;
   c) $S = \{p \in \mathbb{R}_n[x];\ p(\alpha) = p'(\alpha) = 0\}$, donde $\alpha \in \mathbb{R}$ es un valor fijo.
3. Para el espacio vectorial $\mathcal{F} = \{f : \mathbb{R} \to \mathbb{R}\}$ formado por las funciones reales de variable real considerar:
   a) $S = \{f \in \mathcal{F};\ f \text{ es continua}\}$;
   b) $S = \{f \in \mathcal{F};\ f \text{ es derivable}\}$;
   c) $S = \{f \in \mathcal{F};\ f(1) = f(0)\}$;
   d) $S = \{f \in \mathcal{F};\ f(x^2) = f(x)^2,\ \forall x \in \mathbb{R}\}$;
   e) $S = \{f \in \mathcal{F};\ f \text{ es par}\}$;
   f) $S = \{f \in \mathcal{F};\ f \text{ es impar}\}$;
4. $V = \mathcal{M}_{n\times m}(\mathbb{R})$
   a) $S = \{A \in V;\ A \text{ es simétrica}\}$;
   b) $S = \{A \in V;\ A \text{ es antisimétrica}\}$;
   c) $S = \{A \in V;\ A \text{ es invertible}\}$.
5. Para el espacio vectorial formado por todas las sucesiones reales, al que indicaremos[9] con la notación $\mathbb{R}^{\mathbb{N}}$, considerar
   a) $S = \{(x_n)_{n\in\mathbb{N}} \in V;\ (x_n)_{n\in\mathbb{N}} \text{ es convergente}\}$;
   b) $S = \{(x_n)_{n\in\mathbb{N}} \in V;\ (x_n)_{n\in\mathbb{N}} \text{ es divergente}\}$;
   c) $S = \{(x_n)_{n\in\mathbb{N}} \in V;\ (x_n)_{n\in\mathbb{N}} \text{ es acotada}\}$.

[9] En general, Y^X es el conjunto de todas las funciones definidas sobre X y que toman valores en Y.

Ejercicio 4.24. Hallar un generador del subespacio S

1. $S = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$
2. $S = \{(a, b, c, d) \in \mathbb{C}^4 : 2ia = b,\ c + d - ib = 0\}$
3. $S = \{p \in \mathbb{R}_3[x]\ ; p(1 - x) = p(1 + x),\ \forall x \in \mathbb{R}\}$
4. $S = \{p \in \mathbb{R}_3[x]\ ; p(0) = 0, \}$
5. $S = \{A \in \mathcal{M}_{3\times3}(\mathbb{R}) : A$ es simétrica $\}$
6. $S = \{A \in \mathcal{M}_{3\times3}(\mathbb{R}) : A$ es antisimétrica $\}$

4.2.4. Para tener presente

- Un subespacio de un espacio vectorial es un subconjunto cerrado para las operaciones de suma y de producto por un escalar.
- Un subespacio "hereda" del espacio que lo contiene toda la estructura lineal.
- El conjunto que está formado únicamente por el vector nulo es un subespacio vectorial, que está contenido en cualquier otro subespacio.
- La intersección de dos subespacios es un nuevo subespacio, pero su unión no.
- La noción adecuada para construir el menor subespacio que contenga a dos subespacios dados es la noción de suma de subespacios. Dos subconjuntos cualesquiera de un espacio vectorial pueden sumarse, pero la suma de dos subespacios tiene la propiedad de ser también un subespacio.

4.3. Combinaciones lineales, conjuntos generadores, e independencia lineal

En esta sección y la siguiente extenderemos al contexto general de los espacios vectoriales cualesquiera algunas de las nociones fundamentales del álgebra lineal. Reaparecerán así los conceptos de *combinación lineal*, *subespacio generado*, *generador de un subespacio*, *independencia y dependencia lineal*, *base* y *dimensión*, que ya fueron introducidos en el marco particular de los espacios $\mathbb{K}^m$. Comenzamos por introducir las combinaciones lineales y a partir de éstas la posibilidad de generar un subespacio realizando combinaciones lineales de una familia dada de vectores, y el problema de la independencia y dependencia lineal de una familia de vectores. En la sección 4.4 discutiremos los conceptos de base y dimensión, y definiremos claramente cuál es la estructura que será nuestro principal objeto de estudio en el resto de este texto: los espacios vectoriales de dimensión finita.

4.3.1. Combinaciones lineales

La idea de *combinación lineal* de una familia de vectores de $\mathbb{K}^n$ fue introducida con precisión en la sección 1.3, a través de la definición 1.2 que aparece en la página 78. La definición para un espacio vectorial cualquiera es esencialmente la misma. Sólo cambia el contexto en que se hace. Introduciremos además ligeras modificaciones en la notación.

Definición 4.4 (Combinaciones lineales). *Sean v_i, $i = 1, \dots, n$, vectores de un $\mathbb{K}$-espacio vectorial $\mathbb{V}$, y λ_i, $i = 1, \dots, n$, números en el cuerpo $\mathbb{K}$. Diremos que*

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_i v_i + \ldots + \lambda_n v_n \tag{4.3}$$

es la ***combinación lineal*** *de la familia de vectores $(v_1, v_2, \dots, v_n)$ con* ***coeficientes*** *$(\lambda_1, \lambda_2, \dots, \lambda_n)$.*

Diremos también que un vector $v \in \mathbb{V}$ es una combinación lineal de los vectores v_i, $i = 1, \dots, n$, si existen escalares $\lambda_i \in \mathbb{K}$, $i = 1 \dots, n$ tales que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_i v_i + \ldots + \lambda_n v_n.$$

La definición que acabamos de dar extiende a la definición 1.2, página 78, en la que introdujimos la combinaciones lineales en el contexto particular de los espacios vectoriales $\mathbb{K}^n$. Por supuesto, abarca otras situaciones, como las que describimos a continuación.

Ejemplo 4.3.1. Consideremos la matriz real, simétrica

$$\begin{pmatrix} \sqrt{2} & \pi \\ \pi & e \end{pmatrix}.$$

Esta matriz puede escribirse como una combinación lineal de las tres sencillas matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

en la forma

$$\begin{pmatrix} \sqrt{2} & \pi \\ \pi & e \end{pmatrix} = \sqrt{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + e\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \pi\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

La familia de coeficientes de esta combinación lineal es $(\sqrt{2}, e, \pi)$. ♣

Ejemplo 4.3.2. PROBLEMAS DE VALORES INICIALES PARA LAS ECUACIONES DIFERENCIALES

Volvemos sobre la ecuación diferencial lineal

$$x'' + ax' + bx = 0 \tag{4.4}$$

que introdujimos en el ejemplo 4.2. Aunque no demostraremos este hecho, digamos que las soluciones de esta ecuación quedan completamente determinadas por la ecuación y condiciones iniciales adicionales sobre el valor de la solución x y su derivada primera x' en algún punto.

En el contexto dinámico que describíamos en el ejemplo 4.2, los valores $x(t_0)$ e $x'(t_0)$ que la función x toma en un instante inicial $t = t_0$ corresponden a la posición y velocidad iniciales de una partícula. Una vez conocidos estos valores, la evolución posterior (¡y la anterior también!) queda determinada por la ecuación diferencial.

En resumen, en el contexto de las ecuaciones diferenciales un problema importante es el de resolver la ecuación **junto** con condiciones iniciales para x y x' en algún tiempo t_0. Es decir, se trata de buscar una función x que satisfaga

$$\left\{ \begin{array}{c} x'' + ax' + bx = 0, \\ x(t_0) = x_0, \ x'(t_0) = v_0, \end{array} \right. \tag{4.5}$$

para algunos valores prefijados t_0, x_0 y v_0.

Ejercicio 4.25. Supongamos conocidas una solución e_1 del problema (4.5) con datos iniciales $e_1(t_0) = 1$, $e_1'(t_0) = 0$, y una solución e_2 con datos iniciales $e_2(t_0) = 0$, $e_2'(t_0) = 1$. Construir a partir de e_1 y e_2 una solución del problema general (4.5), con datos iniciales x_0 y v_0 cualesquiera. ♣

Ejemplo 4.3.3. Cualquier polinomio

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

en el espacio vectorial $\mathbb{K}[x]$ es la combinación lineal de la familia de polinomios

$$\left(x^n, x^{n-1}, \ldots, x, 1\right)$$

con coeficientes

$$(a_n, a_{n-1}, \ldots a_1, a_0)\,.$$

Al realizar la combinación lineal

$$3(x^2 - x + 1) + 4(x^2 + x + 1) - 2(x^2 + 1)$$

de la familia de polinomios

$$\mathcal{A} = \left(x^2 - x + 1, x^2 + x + 1, x^2 + 1\right)$$

se obtiene el polinomio

$$5x^2 + x + 5.$$

Observemos que si hacemos una combinación lineal de $\mathcal{A}$ con coeficientes $(0, 1, 4)$ volvemos a obtener el mismo polinomio.

Ejercicio 4.26. Completar los detalles de los calculos de estas dos combinaciones lineales de la familia $\mathcal{A}$.

Ejemplo 4.3.4. Volvamos al ejemplo 4.1.1, en el que discutimos la interpolación por funciones lineales a trozos de una función g de la que suponíamos conocidos los valores

$$g(-1) = y_{-1}, \quad g(0) = y_0, \quad g(1) = y_1.$$

La solución del problema es la función

$$l = y_{-1} l_{-1} + y_0 l_0 + y_1 l_1,$$

una combinación lineal de las tres funciones especialmente simples l_{-1}, l_0 y l_1.

Muchos problemas relacionados con combinaciones lineales se reducen en la práctica a la consideración de sistemas de ecuaciones lineales apropiados. En el próximo ejemplo tratamos una de estas cuestiones, motivadas por la segunda parte del ejemplo 4.3.3.

Ejemplo 4.3.5. Hemos visto que dos combinaciones lineales diferentes de la familia

$$\mathcal{A} = \left(x^2 - x + 1, x^2 + x + 1, x^2 + 1\right)$$

producen el mismo polinomio $p[x] = 5x^2 + x + 5$.

1. Una pregunta natural en este contexto es ¿cuáles son las combinaciones lineales de $\mathcal{A}$ que producen $p[x]$? Esto es equivalente a determinar los valores λ_i, $i = 1, 2, 3$, tales que

$$5x^2 + x + 5 = \lambda_1(x^2 - x + 1) + \lambda_2(x^2 + x + 1) + \lambda_3(x^2 + 1).$$

   Ordenando en términos del mismo grado obtenemos

$$5x^2 + x + 5 = (\lambda_1 + \lambda_2 + \lambda_3)x^2 + (-\lambda_1 + \lambda_2)x + \lambda_1 + \lambda_2 + \lambda_3.$$

   Encontramos entonces que los coeficientes λ_i producen el polinomio p si y sólo si son una solución del sistema de ecuaciones lineales

$$\left\{\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 5, \\ -\lambda_1 + \lambda_2 &= 1, \\ \lambda_1 + \lambda_2 + \lambda_3 &= 5. \end{aligned}\right.$$

2. Podemos considerar también el problema de determinar cuáles son los polinomios $ax^2 + bx + c$ que pueden expresarse como combinación lineal de $\mathcal{A}$. Un momento de reflexión nos mostrará que son los polinomios tales que el sistema

$$\left\{\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= a, \\ -\lambda_1 + \lambda_2 &= b, \\ \lambda_1 + \lambda_2 + \lambda_3 &= c \end{aligned}\right. \tag{4.6}$$

   es compatible.

Hemos reducido los dos problemas planteados a la consideración de sistemas lineales. A partir del sistema general (4.6) resolveremos ambas cuestiones. El sistema es equivalente al sistema escalerizado

$$\left\{\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= a, \\ -\lambda_1 + \lambda_2 &= b, \\ 0 &= c - a \end{aligned}\right.$$

Por lo tanto es compatible si y sólo si $a = c$, y el conjunto formado por los polinomios que pueden generarse como combinaciones lineales de $\mathcal{A}$ es

$$\{ax^2 + bx + a \quad a, b \in \mathbb{R}\}. \tag{4.7}$$

Cuando tomamos $a = c = 5$, $b = 1$, encontramos que

$$\lambda_2 = 1 + \lambda_1, \quad \lambda_3 = 4 - 2\lambda_1.$$

Concluirmos que

$$\{(\lambda, 1 + \lambda, 4 - 2\lambda) \ \ \lambda \in \mathbb{R}\}$$

es el conjunto de coeficientes que producen el polinomio $5x^2 + x + 5$ al hacer una combinación lineal de la familia $\mathcal{A}$. ♣

Ejercicio 4.27. Determinar cuáles de las funciones reales $\exp(x)$, $\operatorname{sen}^2(x)$ y $\tan(x)$, definidas en $(-\pi/2, \pi/2)$, pueden expresarse como combinación lineal de la familia

$$(\operatorname{sen}(x), \cos(x), \operatorname{sen}(2x), \cos(2x)).$$

Ejercicio 4.28. Escribir el vector v como combinación lineal del conjunto $\mathcal{A}$.

1. $\mathcal{A} = (3x^3 + x, -2x^2 + x - 1, 3x^3 - 2x^2 + 2x - 1)$ y $v = -3x^3 + 4x^2 + x - 2$.
2. $\mathcal{A} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \right)$ y $v = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$

4.3.2. Subespacio generado y conjuntos generadores

Luego de introducir en la sección 1.3 las combinaciones lineales de una familia $\mathcal{A}$ de vectores resultó bastante natural, a partir del análisis de la compatibilidad de los sistemas de ecuaciones, considerar el conjunto de todas las posibles combinaciones lineales de los elementos de una familia $\mathcal{A}$ formada por todas las columnas de una matriz A de dimensiones $m \times n$. Fue así que formulamos en la página 92 la definición 1.4, en la que se introduce el conjunto $\operatorname{col}(A)$, que resulta ser un subespacio de $\mathbb{K}^m$.

Esta construcción puede generalizarse a un espacio vectorial cualquiera, en el sentido de nuestra próxima definición (que es una adaptación de la definición 1.4 al nuevo contexto en que estamos trabajando).

Definición 4.5 (Subespacio generado). *Consideremos un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ y una familia*

$$\mathcal{A} = (v_1, v_2, \ldots, v_n)$$

de vectores de $\mathbb{V}$. Llamaremos ***subespacio generado*** *por $\mathcal{A}$ al subconjunto de $\mathbb{V}$ formado por todas las posibles combinaciones lineales de los elementos de $\mathcal{A}$. Indicaremos este conjunto con la notación $[\mathcal{A}]$.*

Otra manera de expresar esta definición es especificar el subespacio generado $[\mathcal{A}]$ como

$$[\mathcal{A}] = \{v \in \mathbb{V};\ v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n,\ \lambda_i \in \mathbb{K},\ i = 1, 2, \ldots, n\}.$$

Observación 4.3.6. No todos los vectores $v_i \in \mathcal{A}$ tienen que aparecer explícitamente en cada una de las combinaciones lineales que forman $[\mathcal{A}]$. En caso de que algunos de los coeficientes λ_i sean iguales a 0 podemos no escribir los vectores correspondientes a ellos, porque al estar multiplicados por 0 no aportan nada al resultado final.

Naturalmente, es cierto que $[\mathcal{A}]$ es un subespacio vectorial de $\mathbb{V}$, lo que justifica el nombre de subespacio generado. Además, contiene a $\mathcal{A}$.

Proposición 4.9. *Consideremos un* $\mathbb{K}$*-espacio vectorial* $\mathbb{V}$ *y una familia*

$$\mathcal{A} = (v_1, v_2, \ldots, v_n) \subset \mathbb{V}.$$

Entonces $[\mathcal{A}]$ *es un subespacio de* $\mathbb{V}$ *y* $\mathcal{A} \subset [\mathcal{A}]$*.*

Prueba. Cada vector $v_i \in \mathcal{A}$ es una combinación lineal de los vectores de $\mathcal{A}$, ya que puede escribirse como

$$v_i = 0v_1 + 0v_2 + \ldots + 1v_i + \ldots + \lambda_n v_n,$$

donde todos los escalares de la combinación son 0, salvo el que corresponde a v_i que es igual a 1. Esto muestra que $[\mathcal{A}]$ es no vacío y que contiene a $\mathcal{A}$.

Sean ahora $v, w \in [\mathcal{A}]$. Existen escalares

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta n \in \mathbb{K}$$

tales que

$$\begin{aligned} v &= \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n, \\ w &= \beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_n v_n. \end{aligned}$$

Entonces

$$v + w = (\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + \ldots + (\alpha_n + \beta_n)v_n,$$

que es una combinación lineal de elementos de $\mathcal{A}$. El lector notará que hemos hecho uso de las propiedades asociativa, conmutativa y distributiva de la suma para escribir esta igualdad. Hemos probado entonces que $[\mathcal{A}]$ es cerrado respecto a la suma.

Consideremos ahora $\lambda \in \mathbb{K}$ y, por ejemplo, el vector $v \in [\mathcal{A}]$. Entonces

$$\lambda v = \lambda\left(\alpha_1 v_1 + \ldots + \alpha_n v_n\right) = \left(\lambda\alpha_1\right) v_1 + \ldots + \left(\lambda\alpha_n\right) v_n,$$

y encontramos que λv también es una combinación lineal de los vectores de $\mathcal{A}$. O sea, $[\mathcal{A}]$ también es cerrado respecto al producto de un escalar por un vector. Por lo tanto $[\mathcal{A}]$ es un subespacio de $\mathbb{V}$. $\square$

Una vez formulada la noción de subespacio generado vamos a estudiar con cierto detalle el problema de como caracterizar el subespacio generado a partir de un conjunto dado de vectores de un espacio vectorial. Comenzaremos con un ejemplo.

Ejemplo 4.3.7. Consideremos en $M^{2\times 2}(\mathbb{R})$ el conjunto

$$\mathcal{A} = \left(\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \right).$$

Queremos hallar $[A]$, o sea las matrices que se pueden escribir como combinación lineal de estas tres. Trataremos de dar una caracterización fácil de manejar de este conjunto, buscando qué condiciones deben satisfacer las entradas b_{ij} de una matriz

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

para pertenecer a él. La matriz B pertenece a $[\mathcal{A}]$ si y sólo si existen tres números reales α_i, $i = 1, 2, 3$, tales que

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}.$$

Operando con las matrices en el miembro de la derecha encontramos que esta igualdad es equivalente a

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_3 & -\alpha_1 + \alpha_2 \\ 2\alpha_1 + \alpha_2 + 3\alpha_3 & \alpha_1 + \alpha_3 \end{pmatrix},$$

lo que nos conduce a considerar el sistema lineal

$$\left\{ \begin{array}{rrrl} \alpha_1 & & +\alpha_3 & = b_{11}, \\ -\alpha_1 & +\alpha_2 & & = b_{12}, \\ 2\alpha_1 & +\alpha_2 & +3\alpha_3 & = b_{21}, \\ \alpha_1 & & +\alpha_3 & = b_{22}, \end{array} \right.$$

para determinar qué valores de b_{11}, b_{12}, b_{21} y b_{22} hacen el sistema compatible. Escribiendo la matriz ampliada del sistema y luego escalerizando tenemos

$$\left(\begin{array}{rrr|l} 1 & 0 & 1 & b_{11} \\ -1 & 1 & 0 & b_{12} \\ 2 & 1 & 3 & b_{21} \\ 1 & 0 & 1 & b_{22} \end{array} \right) \mapsto \left(\begin{array}{rrr|l} 1 & 0 & 1 & b_{11} \\ 0 & 1 & 1 & b_{11} + b_{12} \\ 0 & 0 & 0 & 3b_{11} + b_{12} - b_{21} \\ 0 & 0 & 0 & b_{11} - b_{22} \end{array} \right).$$

Concluimos que las entradas de la matriza deben cumplir

$$\left\{\begin{array}{rl} 3b_{11}+b_{12}-b_{21} & =0, \\ b_{11}-b_{22} & =0 \end{array}\right.$$

Un resultado que puede escribirse en la forma

$$[\mathcal{A}] = \left[\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}\right] = \left\{\begin{pmatrix} a & b \\ 3a+b & a \end{pmatrix}, \ a,\, b \in \mathbb{R}\right\}.$$

Observación 4.3.8. Es interesante observar en este ejemplo que aunque generamos matrices haciendo combinaciones lineales a partir de una familia que contiene tres matrices el resultado final sólo tiene dos parámetros libres. La razón es que, por ejemplo, la tercera matriz es una combinación lineal de las dos primeras:

$$\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}.$$

Por lo tanto no aporta nada esencialmente nuevo al fabricar el conjunto de todas las posibles combinaciones lineales. En otras palabras, si una matriz es combinación lineal de las tres matrices en $\mathcal{A}$ entonces es combinación lineal de las dos primeras. En la proposicion 4.10 presentamos un resultado general que tiene que ver con este fenómeno. ♠♣

Ejercicio 4.29. Muestre que $[\mathcal{A}]$ es el menor subespacio de $\mathbb{V}$ que contiene a $\mathcal{A}$, en el siguiente sentido: si $\mathbb{S}$ es un subespacio vectorial de $\mathbb{V}$ y $\mathcal{A} \subset \mathbb{S}$, entonces $[\mathcal{A}] \subset \mathbb{S}$. Mostrar también que $[\mathcal{A}]$ es la intersección de todos los subespacios que contienen al conjunto $\mathcal{A}$.

El ejemplo 4.3.7 ilustra, en el contexto de un espacio de matrices, un fenómeno que ya habíamos encontrado al trabajar con espacios de columnas en el capítulo 1: un mismo espacio vectorial puede ser generado por distintas familias de vectores. Cada una de las familias que genera a un espacio es lo que se llama un *generador* del espacio, en el sentido de nuestra siguiente definición, que extiende a la definición 1.6.

Definición 4.6 (Generadores). *Sea $\mathbb{V}$ un espacio vectorial. Diremos que un subconjunto*

$$\mathcal{A} = (v_1, v_2, \dots, v_n) \subset \mathbb{V}$$

*es un **generador** de $\mathbb{V}$, si cualquier vector de $\mathbb{V}$ puede ser expresado como una combinación lineal de los elementos en $\mathcal{A}$.*

Observación 4.3.9.

1. La definición 4.6 puede aplicarse a cualquier subespacio $\mathbb{S}$ de un espacio vectorial cualquiera, porque sabemos que todo subespacio es en sí mismo un espacio vectorial. Al hacerlo obtenemos la siguiente formulación: diremos que $\mathcal{A}$ es un generador de $\mathbb{S}$ si cualquier vector de $\mathbb{S}$ puede ser expresado como una combinación lineal de los elementos de $\mathcal{A}$.

2. El conjunto $\mathcal{A}$ es un generador de $\mathbb{V}$ si y sólo si $[\mathcal{A}] = \mathbb{V}$. Vemos entonces que es posible dar a la definición de generador la siguiente redacción alternativa: $\mathcal{A}$ es un generador de $\mathbb{V}$ si el subespacio generado por $\mathcal{A}$ es $\mathbb{V}$.

3. También es habitual decir que $\mathcal{A}$ *genera* a $\mathbb{V}$, o que los elementos de $\mathcal{A}$ *generan* a $\mathbb{V}$, cuando $\mathcal{A}$ es un generador de $\mathbb{V}$ ♠

El siguiente ejemplo es casi tautológico.

Ejemplo 4.3.10. Toda familia $\mathcal{A}$ de vectores de $\mathbb{V}$ es un generador del subespacio $[\mathcal{A}]$ generado por $\mathcal{A}$. ♣

Ejemplo 4.3.11. Volvamos al ejemplo 4.1.1. El conjunto de funciones

$$(l_{-1}(x),\ l_0(x),\ l_1(x))$$

genera el espacio de las funciones continuas definidas sobre $[-1, 1]$ cuyas restricciones a los intervalos $[-1, 0]$ y $[0, 1]$ son funciones lineales. ♣

Ejemplo 4.3.12. Los polinomios 1, x y x^2 generan el espacio vectorial $\mathbb{K}_2[x]$. En general, los polinomios x^i, $i = 0, 1, \dots, n$ generan $\mathbb{K}_n[x]$. ♣

Nuestra siguiente proposición muestra como se pueden ir eliminando vectores de los generadores de un espacio vectorial que contienen vectores que pueden ser expresados como una combinación lineal de los restantes. Podemos pensar, desde el punto de vista de la estructura lineal del espacio, que estos vectores son redundantes porque si se eliminan pueden recuperarse haciendo combinaciones lineales de los restantes. El procedimiento de eliminar vectores puede repetirse hasta llegar a un generador en el que ningún vector es combinación lineal de los demás. Esto es lo que en la subseccion 4.3.3 llamaremos un generador *linealmente independiente*, o *base* del espacio $\mathbb{V}$.

Proposición 4.10. *Sea* $\mathbb{V}$ *un espacio vectorial y*

$$\mathcal{A} = (v_1, v_2, \dots, v_n) \subset \mathbb{V}$$

un generador de $\mathbb{V}$. Si alguno de los vectores $v_k \in \mathcal{A}$ puede expresarse como combinación lineal de los restantes vectores de $\mathcal{A}$ entonces la familia

$$\mathcal{A}' = (v_1, v_2, \dots, v_{k-1}, v_{k+1}, \dots, v_n)$$

que se obtiene eliminando de $\mathcal{A}$ el vector $\mathbb{V}$, también es un generador de $\mathbb{V}$.

Observación 4.3.13. Si estamos trabajando en un espacio vectorial dado, aunque la familia $\mathcal{A}$ no genere todo el espacio, podemos aplicar el resultado anterior al subespacio vectorial $\mathbb{V} = [\mathcal{A}]$ para concluir que $[\mathcal{A}'] = [\mathcal{A}]$. Es decir, **al eliminar de un conjunto de vectores un vector que es combinación lineal de los restantes obtenemos un subconjunto que genera exactamente el mismo subespacio que genera el conjunto original**.

Prueba de la proposición 4.10. Debemos mostrar que todo vector $v \in \mathbb{V}$ es una combinación lineal de los vectores de $\mathcal{A}'$. Como $\mathcal{A}$ es un generador del espacio existen escalares λ_i, $i = 1, \dots, n$ tales que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_{k-1} v_{k-1} + \lambda_k v_k + \lambda_{k+1} v_{k+1} + \dots + \lambda_n v_n. \tag{4.8}$$

Por otra parte, sabemos que v_k es una combinación lineal de los restantes vectores en $\mathcal{A}$, lo que significa que existen escalares α_i, $i = 1, \dots, k-1, k+1, n$, tales que

$$v_k = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{k-1} v_{k-1} + \alpha_{k+1} v_{k+1} + \dots + \alpha_n v_n. \tag{4.9}$$

Podemos sustituir la expresión (4.9) de v_k en (4.8). Resulta, luego de agrupar los términos en los que un mismo vector aparece multiplicado por escalares,

$$\begin{aligned} v = {} & (\lambda_1 + \lambda_k \alpha_1)\, v_1 + \dots + (\lambda_{k-1} + \lambda_k \alpha_{k-1})\, v_{k-1} \\ & + (\lambda_{k+1} + \lambda_k \alpha_{k+1})\, v_{k+1} + \dots + (\lambda_n + \lambda_k \alpha_n)\, v_n. \end{aligned}$$

O sea, el vector v puede escribirse como una combinación lineal de los vectores de $\mathcal{A}'$. □

Naturalmente, la operación de eliminar un vector de un generador puede repetirse, hasta producir un generador en el que ya ningún vector es combinación lineal de los demás.

Corolario 4.11. *Sea $\mathbb{V}$ un espacio vectorial y*

$$\mathcal{A} = (v_1, v_2, \dots, v_n) \subset \mathbb{V}$$

un generador de $\mathbb{V}$. Entonces existe un subconjunto $\mathcal{A}'$ de $\mathcal{A}$ que también genera a $\mathbb{V}$ y tal que ningún vector de $\mathcal{A}'$ puede expresarse como combinación lineal de los restantes.

Prueba. Encontraremos un generador $\mathcal{A}'$ eliminando sucesivamente vectores de $\mathcal{A}$, hasta llegar a un generador en el que ningún vector es combinación lineal de los restantes.

Si ningún vector de $\mathcal{A}$ es una combinación lineal de los restantes entonces no hay nada que hacer, y basta tomar $\mathcal{A}' = \mathcal{A}$.

En caso contrario, aplicamos la proposición 4.10, eliminamos un vector de $\mathcal{A}$ y obtenemos un nuevo generador $\mathcal{A}^{(1)}$ de $\mathbb{V}$ que tiene un elemento menos. Si ningún vector de $\mathcal{A}^{(1)}$ es una combinación lineal de los restantes tomamos $\mathcal{A}' = \mathcal{A}^{(1)}$.

En caso contrario, volvemos a aplicar la proposición 4.10 para eliminar un vector de $\mathcal{A}^{(1)}$ y obtener un nuevo generador $\mathcal{A}^{(2)}$ de $\mathbb{V}$.

Repetimos el procedimiento hasta obtener un generador de $\mathbb{V}$ con la propiedad deseada, algo que debe ocurrir al cabo de un número finito de pasos porque la familia $\mathcal{A}$ tiene un número finito de vectores. □

Ejercicio 4.30. Con la notación del corolario 4.11, mostrar que si se elimina de $\mathcal{A}'$ algún vector entonces el subconjunto que resulta genera un subespacio que está estrictamente contenido en $\mathbb{V}$.

Vamos a aplicar la construcción sugerida por el corolario estudiando un ejemplo en el espacio de los polinomios $\mathbb{R}_2[x]$.

Ejemplo 4.3.14. La familia de polinomios

$$\mathcal{A} = \left(1 + x^2, 2 + 2x^2, 1 + x + x^3, 3 + x + 2x^2 + x^3\right)$$

genera un cierto subespacio $\mathbb{S} \subset \mathbb{R}_3[x]$. Busquemos eliminar polinomios de uno en uno, hasta conseguir un generador de $\mathbb{S}$ en el que ningún vector sea combinación lineal de los demás. Para intentar determinar si alguno de los polinomios es una combinación lineal de los restantes podríamos usar el procedimiento ingenuo de examinar las cuatro ecuaciones

$$\begin{array}{rcl}
1 + x^2 &=& \lambda_1(2 + 2x^2) + \lambda_2(1 + x + x^3) + \lambda_3(3 + x + 2x^2 + x^3), \\
2 + 2x^2 &=& \lambda_1(1 + x^2) + \lambda_2(1 + x + x^3) + \lambda_3(3 + x + 2x^2 + x^3) \\
1 + x + x^3 &=& \lambda_1(1 + x^2) + \lambda_2(2 + 2x^2) + \lambda_3(3 + x + 2x^2 + x^3), \\
3 + x + 2x^2 + x^3 &=& \lambda_1(1 + x^2) + \lambda_2(2 + 2x^2) + \lambda_3(1 + x + x^3)
\end{array}$$

para determinar cuáles de ellas se satisfacen para alguna terna $(\lambda_1, \lambda_2, \lambda_3)$ de números reales. Podríamos hacerlo, pero sería repetir cuatro veces el mismo trabajo. Para ver esto escribamos las cuatro ecuaciones en la forma

$$\begin{array}{rcl}
0 &=& (-1)(1 + x^2) + \lambda_1(2 + 2x^2) + \lambda_2(1 + x + x^3) + \lambda_3(3 + x + 2x^2 + x^3), \\
0 &=& \lambda_1(1 + x^2) + (-1)(2 + 2x^2)\lambda_2(1 + x + x^3) + \lambda_3(3 + x + 2x^2 + x^3) \\
0 &=& \lambda_1(1 + x^2) + \lambda_2(2 + 2x^2) + (-1)(1 + x + x^3) + \lambda_3(3 + x + 2x^2 + x^3), \\
0 &=& \lambda_1(1 + x^2) + \lambda_2(2 + 2x^2) + \lambda_3(1 + x + x^3) + (-1)(3 + x + 2x^2 + x^3),
\end{array}$$

que pone en evidencia el hecho de que siempre estamos considerando el mismo problema de igualar a cero combinaciones lineales de la familia $\mathcal{A}$, con la restricción de que alguno de los coeficientes tiene que ser igual a -1. Estudiemos pues este problema general que engloba todas las posibles situaciones. Planteamos

$$0 = \lambda_1(1 + x^2) + \lambda_2(2 + 2x^2) + \lambda_3(1 + x + x^3) + \lambda_4(3 + x + 2x^2 + x^3),$$

que es equivalente a

$$0 = (\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4) + (\lambda_3 + \lambda_4)x + (\lambda_1 + 2\lambda_2 + 2\lambda_4)x^2 + (\lambda_3 + \lambda_4)x^3,$$

que es a su vez equivalente al sistema homogéneo de ecuaciones lineales

$$\left\{\begin{aligned} \lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4 &= 0,\\ \lambda_3 + \lambda_4 &= 0,\\ \lambda_1 + 2\lambda_2 + 2\lambda_4 &= 0,\\ \lambda_3 + \lambda_4 &= 0. \end{aligned}\right.$$

Para estudiarlo recurrimos a la representación matricial habitual

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 3 & 0\\ 0 & 0 & 1 & 1 & 0\\ 1 & 2 & 0 & 2 & 0\\ 0 & 0 & 1 & 1 & 0 \end{array}\right).$$

que se reduce a

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 3 & 0\\ 0 & 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

luego del proceso de escalerización. Observemos que λ_2 y λ_4 son variables libres, así que podemos escoger, por ejemplo, $\lambda_4 = -1$ y despejar el cuarto polinomio como combinación lineal de los restantes. Analicemos si luego de eliminar el cuarto vector podemos quitar algún otro. No hace falta que repitamos todos los cálculos, sólo es necesario eliminar la cuarta columna de la matriz del sistema, que proviene de la constante λ_4 correspondiente al vector que hemos quitado en el paso anterior. Obtenemos, luego de la escalerización

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{array}\right)$$

donde λ_2 es una variable libre. El segundo vector puede quitarse. El sistema que correspondería a la familia $\left(1+x^2, 1+x+x^3\right)$, en la que sólo quedan el primer y tercer vector de la familia original sería

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right),$$

que es determinado y sólo admite la solución trivial. Ninguna de las constantes puede escogerse igual a -1, y ninguno de los dos vectores se puede escribir como una combinación lineal del otro. Hemos concluido el proceso de eliminación de vectores.

Por supuesto, no hace falta ir eliminando los vectores de uno en uno. Observemos que lo que hemos hecho es eliminar los vectores que correspondían a variables libres en el sistema de ecuaciones al que hemos reducido nuestro problema, algo que puede hacerse de una vez en cuanto termina la escalerización del sistema. Como veremos en la sección 4.4, hemos reencontrado un procedimiento que empleábamos en el capítulo 1 para construir bases de subespacios. ♣

Ejercicio 4.31. El conjunto $\mathcal{A}$ dado genera un subespacio $\mathbb{S}$. Eliminar elementos de $\mathcal{A}$ hasta conseguir un generador de $\mathbb{S}$ donde ningún elemento sea combinación lineal de los demás.

1. $\left(\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}\right).$
2. $(x^2-1, x^3+2x, x, x^3+x, 2x^2-1)$.

Ejercicio 4.32. Sean $\mathbb{S}$ y $\mathbb{T}$ dos subespacios de un espacios vectorial $\mathbb{V}$, y

$$\mathcal{A} = (v_1, \ldots, v_n), \quad \mathcal{B} = (w_1, \ldots, w_m),$$

generadores de $\mathbb{S}$ y $\mathbb{T}$ respectivamente. Mostrar que la unión $(v_1, \ldots, v_n, w_1, \ldots, w_m)$ de ambos generadores es un generador del subespacio suma $\mathbb{S} + \mathbb{T}$.

4.3.3. Independencia lineal

En esta sección formalizaremos la importante idea de que existen familias de vectores con la propiedad de que ningún vector puede expresarse como una combinación lineal de los anteriores. Hay varias imágenes posibles para describir esta situación: cada vector apunta en una dirección que no puede conseguirse a partir de los restantes vectores; no hay ningún vector redundante, desde el punto de vista de la estructura lineal, en la familia; cada vector contiene información nueva, que no puede obtenerse haciendo combinaciones lineales de los restantes.

Por ese camino extenderemos a espacios vectoriales cualesquiera la noción de *independencia lineal* que en la definición 1.7 introdujimos para familias de vectores contenidas en $\mathbb{K}^n$. Recordemos que en el contexto de estos espacios la noción de independencia lineal apareció a través del estudio de la determinación o indeterminación de los sistemas de ecuaciones lineales de la forma $AX = B$. En especial, al reducir este problema al caso del sistema homogéneo $AX = O$ (ver la discusión que precede a la definición 1.7, en la página 112 y anteriores). Daremos aquí una motivación diferente para introducir la noción de independencia lineal, pero nuestra definición será esencialmente la misma que la que ya teníamos en el contexto de los espacios $\mathbb{K}^n$.

Consideremos entonces una familia $\mathcal{A} = (v_1, v_2, \ldots, v_n)$ contenida en un espacio vectorial $\mathbb{V}$, y supongamos que satisface la condición de que ninguno de sus vectores puede expresarse como una combinación lineal de los restantes. El análisis que propondremos al lector es similar al que hicimos en el ejemplo 4.3.14, cuando buscamos un procedimiento eficiente para decidir si cada uno de los vectores de una familia dada podía ser despejado como combinación lineal de los demás. Observábamos entonces que el i-ésimo vector v_i puede despejarse si somos capaces de expresar el vector nulo O como una combinación lineal de los vectores de la familia $\mathcal{A}$ en la que el coeficiente λ_i es igual a -1, el opuesto de 1. Pero para conseguir esto basta escribir O como una combinación lineal de $\mathcal{A}$ en la que el coeficiente λ_i es distinto de 0. En efecto, si

$$O = \lambda_1 v_1 + \ldots \lambda_i v_i + \ldots + \lambda_n v_n,$$

podemos multiplicar ambos miembros de la igualdad por $-1/\lambda$ para obtener

$$O = \alpha_1 v_1 + \ldots (-1) v_i + \ldots + \alpha_n v_n,$$

donde los coeficientes α_i satisfacen

$$\alpha_j = -\frac{\lambda_j}{\lambda_i}, \quad j = 1 \ldots i-1, i+1 \ldots n.$$

A partir de esta última expresión puede despejarse entonces

$$v_i = \alpha_1 v_1 + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \ldots + \alpha_n v_n.$$

En resumen, para poder expresar algún vector de $\mathcal{A}$ como una combinación lineal de los restantes es suficiente que el vector nulo se pueda escribir como una combinación lineal de la familia en la que algún coeficiente es distinto de cero. Esta condición, que se expresa por medio de una condición algebraica sencilla, es la que usaremos para definir la dependencia e independencia lineal.

Ejercicio 4.33. Mostrar que también es cierto que si algún vector de la familia $\mathcal{A}$ puede escribirse como una combinación lineal de los restantes entonces el vector nulo puede expresarse como una combinación lineal de $\mathcal{A}$ en la que alguno de los coeficientes es distinto de cero.

Definición 4.7 (Independencia lineal). *Diremos que una familia*

$$(v_1, v_2, \ldots, v_n)$$

*de vectores en un espacio vectorial $\mathbb{V}$ es **linealmente independiente** si la única manera de expresar el vector nulo como combinación lineal de la familia es escogiendo todos los coeficientes de la combinación iguales a 0. Cuando una familia no es linealmente independiente diremos que es **linealmente dependiente**.*

La discusión que precede a la definición 4.7 es una demostración de la siguiente proposición.

Proposición 4.12. *Sea $\mathbb{V}$ un espacio vectorial. La familia $\mathcal{A} = (v_1, v_2, \ldots, v_n)$ es linealmente independiente si y sólo si cuando ningún vector de $\mathcal{A}$ es una combinación lineal de los otros. Análogamente, $\mathcal{A}$ es linealmente dependiente si y sólo si existe al menos un vector en $\mathcal{A}$ que es combinación lineal de los demás.*

El siguiente ejercicio propone un refinamiento de la segunda parte de esta proposición, y un par de aplicaciones.

Ejercicio 4.34.

1. Mostrar que si $(v_1, v_2, \ldots, v_n)$ es una familia linealmente dependiente si y sólo si existe un vector de la familia que puede ser expresado como una combinación lineal de los anteriores.

2. Utilizar la parte anterior para mostrar que si $\mathcal{A} = (v_1, v_2, \ldots, v_n)$ es linealmente independiente y $v \notin [\mathcal{A}]$ entonces la familia

$$(v_1, v_2, \ldots, v_n, v)$$

que se obtiene agregando el vector v a $\mathcal{A}$ también es linealmente independiente.

3. Encontrar un polinomio $p[x] \in \mathbb{R}_2[x]$ tal que la familia

$$\left(1 + x + x^2, 2 - x, p[x]\right)$$

sea linealmente independiente. Para el polinomio $p[x]$ hallado, mostrar que no se puede agregar un segundo polinomio $q[x]$ de forma tal que se obtenga una nueva familia también linealmente independiente.

El corolario 4.11 de la proposición 4.10 puede expresarse ahora en la forma de la siguiente proposición.

Proposición 4.13. *Sea* $\mathbb{V}$ *un espacio vectorial y*

$$\mathcal{A} = (v_1, v_2, \ldots, v_n) \subset \mathbb{V}$$

un generador de $\mathbb{V}$*. Entonces existe un subconjunto* $\mathcal{A}'$ *de* $\mathcal{A}$ *que es linealmente independiente y que también genera a* $\mathbb{V}$*.*

Ejemplo 4.3.15. Una familia $\mathcal{A}$ que tiene dos vectores iguales es linealmente dependiente. Basta fabricar una combinación lineal en la que a uno de los vectores repetidos le asignamos el coeficente 1, al otro -1, y a los restantes 0.

Observación 4.3.16. Aparece en este ejemplo el interés de formular la definición de independencia lineal sobre conjuntos ordenados. Por ejemplo, si no tuviéramos en cuenta el orden, el conjunto de las columnas de la matriz

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

se reduce a

$$\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$$

que es linealmente independiente. La definición que dimos detecta que la primera[10] columna y la tercera columna son iguales, y declara que las columnas de A son linealmente dependientes. Observemos que la tercera columna puede escribirse como una combinación lineal obvia de la primera y la segunda. ♠♣

Determinar la independencia o dependencia lineal de un conjunto de vectores se reduce en muchos casos a estudiar un sistema de ecuaciones lineales. Veamos un ejemplo.

[10] Por cierto, es un error, frecuente, pero error al fin, decir o escribir “la *primer* fila”, o “la *tercer* fila”. Primer/a, tercer/a son adjetivos, y deben concordar en género con el sustantivo al que califican.

Ejemplo 4.3.17. En el espacio de matrices $M^{2\times 2}(\mathbb{R})$ consideremos el conjunto

$$\mathcal{A} = \left(\left(\begin{array}{cc} 1 & 1 \\ 1 & 0 \end{array} \right), \left(\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right) \right).$$

Para saber si es linealmente independiente aplicaremos directamente la definición. Formamos una combinación lineal

$$\lambda_1 \left(\begin{array}{cc} 1 & 2 \\ 3 & 0 \end{array} \right) + \lambda_2 \left(\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right) = \left(\begin{array}{cc} \lambda_1 + \lambda_2 & 2\lambda_1 + \lambda_2 \\ 3\lambda_1 + \lambda_2 & \lambda_2 \end{array} \right)$$

con coeficientes genéricos λ_1 y λ_2, e igualamos esta combinación al vector nulo del espacio:

$$\left(\begin{array}{cc} \lambda_1 + \lambda_2 & 2\lambda_1 + \lambda_2 \\ 3\lambda_1 + \lambda_2 & \lambda_2 \end{array} \right) = \left(\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right). \tag{4.10}$$

Si la igualdad 4.10 entre matrices implica que λ_1 y λ_2 deben ser nulos, entonces la familia es linealmente independiente. Si la igualdad puede satisfacerse para alguna elección de estos coeficientes en que alguno de ellos es distinta de cero, entonces la familia es linealmente dependiente. Notemos que 4.10 es equivalente al sistema lineal de ecuaciones

$$\left\{ \begin{array}{rr} \lambda_1 + \lambda_2 = & 0, \\ 2\lambda_1 + \lambda_2 = & 0, \\ 3\lambda_1 + \lambda_2 = & 0, \\ \lambda_2 = & 0. \end{array} \right.$$

No hace falta escalerizarlo para darse cuenta de que es compatible determinado, y sólo admite la solución $\lambda_1 = \lambda_2 = 0$. Por lo tanto, el conjunto $\mathcal{A}$ es linealmente independiente.

Ejemplo 4.3.18. Estudiemos si la familia

$$\mathcal{A} = \left(x^2, 2x+2, 3x^2+x+1 \right) \subset \mathbb{R}[x],$$

es linealmente dependiente o independiente. La igualdad

$$\lambda_1 x^2 + \lambda_2(2x+2) + \lambda_3(3x^2+x+1) = 0$$

es equivalente a

$$(\lambda_1 + 3\lambda_3)x^2 + (2\lambda_2 + \lambda_3)x + 2\lambda_2 + \lambda_3 = 0,$$

que a su vez es equivalente al sistema de ecuaciones

$$\left\{\begin{array}{rcl} \lambda_1 + 3\lambda_3 &=& 0, \\ 2\lambda_2 + \lambda_3 &=& 0, \\ 2\lambda_2 + \lambda_3 &=& 0, \end{array}\right.$$

Las soluciones son de la forma

$$(\lambda_1, \lambda_2, \lambda_3) = (6\lambda, \lambda, -2\lambda), \quad \lambda \in \mathbb{R}.$$

Obviamente, hay soluciones no triviales. Por lo tanto $\mathcal{A}$ es linealmente dependiente.

Escogiendo, por ejemplo, $\lambda = -1$ conseguimos

$$(-6)x^2 + (-1)(2x + 2) + 2(3x^2 + x + 1) = 0,$$

de donde podemos despejar con total facilidad

$$2x + 2 = (-6)x^2 + 2(3x^2 + x + 1).$$

Ejercicio 4.35. Expresar cada uno de los otros dos vectores de la familia como combinación lineal de los restantes. ♣

El ejemplo anterior no debe llevarnos a concluir que en una familia linealmente dependiente siempre se puede despejar cualquier vector como combinación lineal de los demás. El siguiente ejercicio ilustra este fenómeno.

Ejercicio 4.36. Determinar que vectores de la familia linealmente dependiente

$$\mathcal{A} = \left(x^3 + x + 1, x^3 + x^2, -2x^3 + x + 1, x^3 + 3x^2 + x + 1\right) \subset \mathbb{R}[x],$$

pueden expresarse como combinación lineal de los restantes.

Ejemplo 4.3.19. Sea $\mathbb{V}$ un espacio vectorial, y v un vector en $\mathbb{V}$ distinto del vector nulo. La familia (v) que sólo está formada por el vector v es linealmente independiente. Pero la familia (O), formada por el vector nulo, es linealmente dependiente. En general, cualquier familia de vectores que contenga al vector nulo es linealmente dependiente.

Ejercicio 4.37. Completar los detalles de este ejemplo.

La definición de independencia lineal nos dice que una familia linealmente independiente tiene la propiedad de que el vector nulo del espacio tiene una única representación como combinación lineal de los vectores de la familia. Esta propiedad se extiende a cualquier vector del subespacio generado por la familia. La razón es simple. Si algún vector admitiera dos representaciones diferentes, al restarlas encontraríamos una representación del vector nulo con coeficientes no nulos, lo que contradiría la independencia lineal. A continuación enunciamos y demostramos con detalle esta proposición.

Proposición 4.14. *Sea* $\mathbb{V}$ *un espacio vectorial y* $\mathcal{A} = (v_1, \ldots, v_n)$ *un conjunto de vectores de* $\mathbb{V}$. *Entonces* $\mathcal{A}$ *es linealmente independiente si y sólo si todo vector del subespacio generado por* $\mathcal{A}$ *se puede expresar de una única manera como combinación lineal de los vectores de* $\mathcal{A}$.

Prueba. Supongamos que $\mathcal{A}$ es linealmente independiente, y demostremos que todo vector del subespacio $[\mathcal{A}]$ generado por $\mathcal{A}$ se puede expresar de una única manera como combinación lineal de los vectores de $\mathcal{A}$. La definición de $[\mathcal{A}]$ asegura que cualquiera de sus vectores se puede escribir como combinación lineal de vectores de $\mathcal{A}$. Veremos ahora que esa combinación lineal es única.

Supongamos que un vector $v \in [\mathcal{A}]$ admite las expresiones

$$\begin{aligned} v &= \alpha_1 v_1 + \ldots + \alpha_n v_n, \\ v &= \beta_1 v_1 + \ldots + \beta_n v_n, \end{aligned}$$

como combinación lineal de $\mathcal{A}$. Restando ambas igualdades obtenemos

$$O = (\alpha_1 - \beta_1)\, v_1 + \ldots + (\alpha_n - \beta_n)\, v_n,$$

que es una combinación lineal de $\mathcal{A}$, igualada a 0. Por hipótesis $\mathcal{A}$ es linealmente independiente, lo que implica que los coeficientes de esta última combinación lineal deben ser todos nulos. Concluimos que

$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2, \ldots \alpha_n = \beta_n.$$

Por lo tanto las dos formas de escribir v como combinación lineal de $\mathcal{A}$, son en realidad la misma.

Para probar el recíproco sólo hay que observar que O está en el espacio $[\mathcal{A}]$. Nuestras hipótesis implican que hay una única combinación lineal de los vectores de $\mathcal{A}$ que es igual O. Como

$$O = 0v_1 + 0v_2 + \ldots + 0v_n,$$

esta única combinación lineal es necesariamente la que acabamos de exhibir, y que tiene todos los coeficientes nulos. □

Observación 4.3.20. La proposición 4.14 es una generalización de la proposición que asegura que un sistema lineal compatible $AX = B$ tiene solución única si y sólo si el sistema lineal homogéneo $AX = O$ tiene una única solución. Recordemos que los productos AX no son otra cosa que las combinaciones lineales de las columnas de A con los coeficientes almacenados en X ♣

La proposición 4.14 tiene el siguiente importantísimo corolario.

Corolario 4.15. *Si*

$$\mathcal{A} = (v_1, \ldots, v_n)$$

es un generador linealmente independiente de un $\mathbb{K}$*-espacio vectorial* $\mathbb{V}$*, entonces para cada vector* $v \in \mathbb{V}$ *existe en* $\mathbb{K}^n$ *un único vector*

$$(\lambda_1, \ldots, \lambda_n)$$

de coeficientes que permiten expresar a v *como combinación lineal*

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Ejercicio 4.38. Verificar los detalles necesarios para completar la prueba del corolario 4.15

Por supuesto, los generadores que son además linealmente independientes son los que llamaremos *bases* del espacio. Al poner una base en el espacio se establece así una correspondencia uno a uno entre los vectores de $\mathbb{V}$ y las n-uplas $(\lambda_1, \ldots, \lambda_n)$ en el espacio $\mathbb{K}^n$, a las que llamaremos *coordenadas* respecto a la base. Nos ocuparemos de las bases en la sección 4.4, y de las coordenadas en la 4.5.

Ejercicio 4.39. Determinar si el conjunto de vectores A es un generador del espacio vectorial $\mathbb{V}$.

1. $\mathbb{V} = \mathbb{R}^3$, $A = ((1,0,0),(0,1,0))$.
2. $\mathbb{V} = \mathbb{R}^3$, $A = ((0,1,1),(0,0,1),(1,1,1),(1,-1,1))$.
3. $\mathbb{V} = \mathcal{M}_{2\times 2}(\mathbb{R})$, $A = \left(\begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 3 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 3 & 1 \end{pmatrix} \right)$
4. $\mathbb{V} = \mathcal{M}_{2\times 2}(\mathbb{R})$, $A = \left(\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 4 & -1 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} -2 & 5 \\ 6 & 0 \end{pmatrix} \right)$

Ejercicio 4.40. Hallar un generador del subespacio $\mathbb{S}$

1. $\mathbb{S} = \{(x,y,z) \in \mathbb{R}^3 : x + y + z = 0\}$
2. $\mathbb{S} = \{(a,b,c,d) \in \mathbb{C}^4 : 2ia = b,\ c + d - ib = 0\}$
3. $\mathbb{S} = \{p \in \mathbb{R}_3[x]\ ; p(1-x) = p(1+x),\ \forall x \in \mathbb{R}\}$
4. $\mathbb{S} = \{p \in \mathbb{R}_3[x]\ ; p(0) = 0, \}$
5. $\mathbb{S} = \{A \in \mathcal{M}_{3\times 3}(\mathbb{R}) : A \text{ es simétrica }\}$
6. $\mathbb{S} = \{A \in \mathcal{M}_{3\times 3}(\mathbb{R}) : A \text{ es antisimétrica }\}$

Ejercicio 4.41. En los siguientes casos determinar si el conjunto $\mathcal{A}$ es linealmente independiente. Cuando no lo sea encontrar un subconjunto linealmente independiente que permita expresar a los restantes vectores como combinación lineal del subconjunto seleccionado.

1. $\mathcal{A} = ((a, a^2, 1),\ (-1, a, a),\ (0, 2a^2, a^2 + 1))$. Discutir según a.
2. $\mathcal{A} = (p_1, p_2, p_3, p_4) \subset \mathbb{R}_2[x]$, donde
$$p_1(x) = x^2 + 1, \quad p_2(x) = x^2 + x, \quad p_3(x) = x + 2, \quad p_4(x) = x^2 + 3x.$$
3. $\mathcal{A} = (p_1, p_2, p_3) \subset \mathbb{R}_2[x]$, donde
$$p_1(x) = 4x + 3, \quad p_2(x) = x^2 - 1, \quad p_3(x) = ax^2 + 4x + 5.$$
Discutir según $a \in \mathbb{R}$.

Ejercicio 4.42. Sean $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ una matriz,
$$\mathcal{C} = (X_1, X_2, \dots, X_l)$$
un subconjunto de vectores de $\mathbb{K}^n$ y
$$\mathcal{B} = (AX_1, AX_2, \dots, AX_l) \subset \mathbb{R}^m.$$
Discutir si las siguientes afirmaciones son verdaderas o falsas:

1. Si $\mathcal{C}$ es linealmente independiente entonces $\mathcal{B}$ es linealmente independiente.
2. Si $\mathcal{B}$ es linealmente independiente entonces $\mathcal{C}$ es linealmente independiente.

En el caso de que alguna de las afirmaciones sea falsa dar un contraejemplo, y estudiar que hipótesis adicionales sobre A permiten asegurar que la afirmación es verdadera.

Ejercicio 4.43. Sea $(v_1, v_2, \dots, v_n)$ un conjunto linealmente independiente de un espacio vectorial V sobre el cuerpo $\mathbb{K}$. Se considera el vector
$$v = \sum_{i=1}^{n} a_i v_i, \quad \text{con} \quad a_1, a_2, \dots, a_n \in \mathbb{K}.$$

1. Sabiendo que
$$\sum_{i=1}^{n} a_i \neq 1, \quad \sum_{i=1}^{n} \lambda_i (v_i - v) = 0,$$
probar que $\displaystyle\sum_{i=1}^{n} \lambda_i = 0$.
2. Sabiendo que
$$\sum_{i=1}^{n} a_i \neq 1$$
probar que $((v_1 - v), (v_2 - v), \dots, (v_n - v))$ es linealmente independiente.

3. Sabiendo que
$$\sum_{i=1}^{n} a_i = 1$$
probar que $((v_1 - v), (v_2 - v), \dots, (v_n - v))$ es linealmente dependiente.

Existe un análogo para subespacios de esta idea de independencia lineal, que es la noción de *suma directa*. Recordemos que dados dos subespacios $\mathbb{S}$ y $\mathbb{T}$ de un espacio $\mathbb{V}$ el subespacio suma $\mathbb{S} + \mathbb{T}$ está formado por todas las posibles sumas de vectores de $\mathbb{S}$ con vectores de $\mathbb{T}$.

En resumen, $v \in \mathbb{S} + \mathbb{T}$ si y sólo si existen $s \in \mathbb{S}$ y $t \in \mathbb{T}$ tales que $v = s + t$. Diremos además que la *suma es directa* si para cualquier vector v en el subespacio $\mathbb{S} + \mathbb{T}$ la descomposición de v como suma de un elemento $s \in \mathbb{S}$ más un elemento $t \in \mathbb{T}$ es única. Cuando la suma es directa es usual representarla con la notación $\mathbb{S} \oplus \mathbb{T}$.

Ejemplo 4.3.21. En $\mathbb{R}^3$ consideramos los subespacios

$$\begin{aligned} \mathbb{S} &= \{(x_1, x_2, 0)\,,\ x_1, x_2 \in \mathbb{R}\}\,, \\ \mathbb{T} &= \{(0, x_2, x_3)\,,\ x_2, x_3 \in \mathbb{R}\}\,, \\ \mathbb{U} &= \{(0, 0, x_3)\,,\ x_3 \in \mathbb{R}\}\,. \end{aligned}$$

Veamos primero que cualquier vector $X = (x_1, x_2, x_3) \in \mathbb{R}^3$ se puede escribir como suma de un vector $S = (s_1, s_2, 0) \in \mathbb{S}$ y un vector $T = (0, t_2, t_3) \in \mathbb{T}$. Para que esto ocurra, fijados x_1, x_2 y x_3 cualesquiera, tenemos que poder encontrar s_1, s_2, t_1 y t_2 tales que

$$(x_1, x_2, x_3) = (s_1, s_2 + t_2, t_3).$$

Naturalmente, este problema es equivalente a resolver el sistema

$$\left\{ \begin{aligned} s_1 &= x_1, \\ s_2 + t_2 &= x_2, \\ t_3 &= x_3. \end{aligned} \right.$$

Salta a la vista que el sistema es siempre compatible. Los valores de s_1 y t_3 quedan determinados, pero no los de s_2 y t_2, de modo que hay muchas soluciones. Por lo tanto, hay muchas maneras de descomponer X como la suma de un vector de $\mathbb{S}$ y uno de $\mathbb{T}$. Por ejemplo

$$(1, 1, 1) = (1, 1, 0) + (0, 0, 1)\,, \quad (1, 1, 1) = (1, 1/2, 0) + (0, 1/2, 1)\,.$$

Observación 4.3.22. La no unicidad en la descomposición proviene de que hay vectores no nulos en la intersección $\mathbb{S} \cap \mathbb{T}$. Por ejemplo, a partir de la descomposición

$$(1,1,1) = (1,1,0) + (0,0,1)$$

podemos fabricar otras descomposiciones, simplemente sumando y restando un vector $(0,\lambda,0) \in \mathbb{S} \cap \mathbb{T}$. Tenemos

$$(1,1,1) = (1,1,0) + (0,\lambda,0) - (0,\lambda,0) + (0,0,1)\,.$$

Agrupando convenientemente resulta

$$(1,1,1) = (1,1+\lambda,0) + (0,-\lambda,1)\,.$$

Como los vectores $\pm(0,\lambda,0)$ están en $\mathbb{S}$ y $\mathbb{T}$ podemos sumarlos con los vectores S y T de la descomposición sin salir de $\mathbb{S}$ ni de $\mathbb{T}$, lo que da lugar a una nueva descomposición. La unicidad de la descomposición es equivalente a que la intersección sea trivial. Ver el ejercicio 4.47.

Concluimos que

$$\mathbb{R}^3 = \mathbb{S} + \mathbb{T},$$

pero esta suma no es directa, la manera de escribir un vector en $\mathbb{R}^3$ como suma de un elemento en $\mathbb{S}$ y otro en $\mathbb{T}$ no es única.

Veamos ahora que el espacio $\mathbb{R}^3$ puede expresarse como suma directa de $\mathbb{S}$ y $\mathbb{U}$. Para ello consideramos, como hicimos al analizar la suma de $\mathbb{S}$ y $\mathbb{T}$ un vector $X = (x_1, x_2, x_3)$ genérico en $\mathbb{R}^3$, e intentamos expresarlo como la suma

$$(x_1, x_2, x_3) = (s_1, s_2, 0) + (0, 0, u_3) = (s_1, s_2, u_3).$$

de un vector $S = (s_1, s_2, 0)$ en $\mathbb{S}$ y un $U = (0, 0, u_3)$ en $\mathbb{U}$. Es obvio que la igualdad se satisface cuando

$$s_1 = x_1, \quad s_2 = x_2, \quad u_3 = x_3,$$

y que esta es la única posibilidad. Por lo tanto cualquier vector de $\mathbb{R}^3$ se puede expresar de manera única como la suma de un vector en $\mathbb{S}$ y otro en $\mathbb{U}$. Tenemos entonces

$$\mathbb{R}^3 = \mathbb{S} \oplus \mathbb{U}.$$

Ejercicio 4.44. Determinar cuál es el subespacio $\mathbb{T} + \mathbb{U}$.

Para cerrar este ejemplo recomendamos al lector elaborar una representación geométrica de $\mathbb{S}$, $\mathbb{T}$ y $\mathbb{U}$, y de los subespacios que se obtienen haciendo las sumas e intersecciones de estos tres subespacios. También del fenómeno de no unicidad en la descomposición que se analiza en la observación 4.3.22. ♣

Observación 4.3.23. Comentarios sobre la suma directa
Cuando un espacio vectorial $\mathbb{V}$ es la suma directa

$$\mathbb{V} = \mathbb{S} \oplus \mathbb{T}$$

de dos subespacios, tenemos a nuestra disposición una descomposición del espacio $\mathbb{V}$ en dos subespacios más chicos. En este sentido, se trata de una descomposición del espacio en elementos más simples. También podemos pensar que los subespacios $\mathbb{S}$ y $\mathbb{T}$ están determinados por información adicional, que determina a estos subconjuntos de $\mathbb{V}$. Veamos algunos ejemplos.

Ejemplo 4.3.24. Funciones pares e impares
Consideremos por ejemplo $C(\mathbb{R})$, el espacio vectorial real formado por las funciones continuas definidas sobre la recta real $\mathbb{R}$. Recordemos que una función es

- *par* si $f(x) = f(-x)$ para todo $x \in \mathbb{R}$;
- *impar* si $f(-x) = -f(x)$ para todo $x \in \mathbb{R}$.

Una función continua f cualquiera puede descomponerse como suma de una función par y una impar, de la siguiente manera:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{-f(-x) + f(x)}{2}, \quad x \in \mathbb{R}.$$

Esta descomposición induce la descomposición del espacio $C(\mathbb{R})$ como suma directa del subespacio de las funciones pares con el subespacio de las funciones impares. Los detalles se dejan como un ejercicio para el lector.

Ejercicio 4.45. Sean $\mathbb{S}$ y $\mathbb{T}$ los subconjuntos de $C(\mathbb{R})$ formados por las funciones pares e impares respectivamente. Mostrar que ambos son subespacios vectoriales, y que $C(\mathbb{R}) = \mathbb{S} \oplus \mathbb{T}$.

Notemos que en este ejemplo cada uno de los subespacios que aparece en la descomposición está formado por funciones que tienen más estructura que una función cualquiera del espacio, ya que se trata de funciones con algún tipo de simetría. ♣

Ejemplo 4.3.25. Señales: promedios más detalles
En este ejemplo mostraremos que una señal continua o discreta puede descomponerse en la suma de su valor promedio más los detalles que miden en cada punto la diferencia de la señal con su valor promedio. Esta descomposición se realiza como la descomposición del espacio vectorial que representa a la señal como suma directa de dos subespacios más simples.

Comencemos estudiando señales continuas, que pueden modelarse, por ejemplo, por medio de una función $f : [0,1] \to \mathbb{R}$. Consideraremos en este ejemplos funciones continuas. El valor promedio de la función f es

$$\Lambda(f) = \int_0^1 f(x)dx.$$

El apartamiento para cada $x \in [0,1]$ de la función f respecto a este valor medio está medido por la función

$$g(x) = f(x) - \Lambda(f).$$

Observemos que el valor promedio de g es igual a 0. Hemos descompuesto nuestra función original f como

$$f = g + \Lambda, \tag{4.11}$$

donde Λ es una función constante (en la que queda codificado el valor medio de la señal f), y g es una función con valor promedio cero (que almacena los detalles, o fluctuaciones de f que la apartan de su valor medio). Esta descomposición de f es única. En efecto, si g tiene promedio cero, al integrar la igualdad (4.11) sobre el intervalo $[0,1]$ obtenemos

$$\int_0^1 f(x)\,dx = \int_0^1 g(x)\,dx + \int_0^1 \Lambda\,dx = \Lambda,$$

de modo que Λ es necesariamente igual a la integral de f. En consecuencia g queda determinada como $f - \Lambda$.

Los argumentos que hemos presentado muestran que el espacio $C([0,1])$ formado por las funciones continuas definidas sobre $[0,1]$ puede descomponerse como

$$C([0,1]) = \mathbb{S} \oplus \mathbb{T},$$

donde $\mathbb{S}$ es el subespacio que está formado por las funciones constantes, y $\mathbb{T}$ por las funciones que tienen integral nula.

Ejercicio 4.46. Generalizar este análisis al caso discreto, mostrando que $\mathbb{R}^n$ se puede descomponer como suma directa en la forma

$$\mathbb{R}^n = [(1,1,\ldots,1)] \oplus \{(x_1,x_2,\ldots,x_n);\ x_1 + x_2 + \ldots + x_n = 0\}.$$

Para cerrar el ejemplo digamos que este tipo de análisis en el que se separan los valores medios de los detalles puede hacerse a varias escalas, y aparece muchas veces representado frente a nuestros ojos al descargar una imagen en nuestra computadora. Vemos aparecer primero manchas de colores, valores medios en una determinada zona de la imagen, sobre las que poco a poco se van perfilando los detalles. ♣♠

Ejercicio 4.47. Sean $\mathbb{S}$ y $\mathbb{T}$ dos subespacios de un espacio vectorial $\mathbb{V}$. Mostrar que las siguientes afirmaciones son equivalentes:

1. la suma $\mathbb{S}+\mathbb{T}$ es directa;
2. si dos vectores $s \in \mathbb{S}$ y $t \in \mathbb{T}$ satisfacen que su suma $s+t$ es igual al vector nulo del espacio, entonces $s=t=O$;
3. $\mathbb{S} \cap \mathbb{T} = \{O\}$.

Ejercicio 4.48. Sean $\mathbb{S}$ y $\mathbb{T}$ dos subespacios de un espacio vectorial $\mathbb{V}$ tales que su suma es directa. Sean $\mathcal{A} = (v_1, \ldots, v_n)$ y $\mathcal{B} = (w_1, \ldots, w_m)$ conjuntos linealmente independientes contenidos en $\mathbb{S}$ y $\mathbb{T}$ respectivamente. Mostrar que la unión $(v_1, \ldots, v_n, w_1, \ldots, w_m)$ de ambos conjuntos es linealmente independiente.

4.3.4. Subespacio generado e independencia lineal para familias infinitas

En las subsecciones anteriores hemos dado definiciones de subespacio generado, conjunto generador e independencia lineal para subconjuntos finitos de un espacio vectorial $\mathbb{V}$. Vamos a extenderlas ahora a conjuntos cualesquiera de vectores, posiblemente infinitos, reescribiendo las definiciones de una manera que abarque a subconjuntos de $\mathbb{V}$ que tengan cualquier cardinalidad. Sólo se trata de un pequeño ajuste de redacción que requiere una precisión: una combinación lineal de elementos de un espacio vectorial es necesariamente una combinación lineal de una cantidad finita[11] de elementos de $\mathcal{A}$.

Definición 4.8 (Subespacio generado). *Consideremos un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ y una familia $\mathcal{A}$ de vectores de $\mathbb{V}$. Llamaremos **subespacio generado** por $\mathcal{A}$ al subconjunto de $\mathbb{V}$ formado por todas las posibles combinaciones lineales de elementos de $\mathcal{A}$.*

También en este caso emplearemos la notación $[\mathcal{A}]$ para indicar el subespacio generado por $\mathcal{A}$. Al igual que en el caso finito, diremos que un subconjunto $\mathcal{A}$ de un espacio vectorial $\mathbb{V}$ es un *generador* de $\mathbb{V}$ si $\mathbb{V} = [\mathcal{A}]$. Una definición que puede expresarse con un estilo similar al de la definicón 4.6, tal como se hace a continuación.

Definición 4.9 (Generadores). *Sea $\mathbb{V}$ un espacio vectorial. Diremos que un subconjunto $\mathcal{A}$ de $\mathbb{V}$ es un **generador** de $\mathbb{V}$, si cualquier vector de $\mathbb{V}$ puede ser expresado como una combinación lineal de elementos en $\mathcal{A}$.*

[11] Una combinación lineal infinita requiere introducir alguna noción de convergencia en el espacio, en el mismo sentido que la consideración de series en el conjunto de los números reales descansa sobre la idea de límite. Éste es un tema propio del análisis matemático, interesante en sí mismo y por sus aplicaciones, que el lector seguramente abordará en otros cursos.

Por último, extendemos la definición de independencia lineal al caso general.

Definición 4.10 (Independencia lineal). *Diremos que una familia $\mathcal{A}$ de vectores en un espacio vectorial $\mathbb{V}$ es* ***linealmente independiente*** *si cualquier combinación lineal de vectores de $\mathcal{A}$ que sea igual al vector nulo tiene necesariamente todos sus coeficientes iguales a cero. Cuando una familia no es linealmente independiente diremos que es* ***linealmente dependiente****.*

Ejemplo 4.3.26. El conjunto de polinomios

$$\mathcal{A} = \{x^n,\ n = 0, 1, 2, \ldots\} \subset \mathbb{K}[x]$$

es linealmente independiente y un generador de $\mathbb{K}[x]$. Una combinación lineal de l elementos

$$x^{n_1}, x^{n_2}, \ldots, x^{n_l},$$

de $\mathcal{A}$, con coeficientes

$$\lambda_1, \lambda_2, \ldots, \lambda_l,$$

en $\mathbb{K}$, es un polinomio

$$\lambda_1 x^{n_1} + \lambda_2 x^{n_2} + \ldots + \lambda_l x^{n_l}. \tag{4.12}$$

La expresión (4.12) es igual al vector nulo sólo cuando todos los coeficientes λ_i, $i = 1, 2, \ldots, l$ son iguales a 0. Este hecho prueba que $\mathcal{A}$ es linealmente independiente. Por otra parte, cualquier polinomio

$$p[x] = a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0$$

es la combinación lineal de los $n+1$ primeros elementos de $\mathcal{A}$ con coeficietes a_i, $i = 0, 1, \ldots, n$. Por lo tanto $\mathcal{A}$ es un generador de $\mathbb{K}[x]$. ♣

Ejercicio 4.49. Considere en el espacio vectorial $\mathbb{R}_n[x]$ de los polinomios de grado menor o igual a n el subconjunto $\mathcal{A}$ de los polinomios de grado n. ¿Cuál es el subespacio generado por $\mathcal{A}$?

Ejercicio 4.50. Sea $\mathbb{V}$ un espacio vectorial cualquiera, y $\mathbb{S}$ y $\mathbb{T}$ dos subespacios de $\mathbb{V}$. Mostrar que el subespacio suma $\mathbb{S} + \mathbb{T}$ es el subespacio generado por la unión $\mathbb{S} \cup \mathbb{T}$ de $\mathbb{S}$ y $\mathbb{T}$.

Ejercicio 4.51. Mostrar que un conjunto $\mathbb{A}$ es linealmente independiente si y sólo si todos sus subconjuntos son linealmente independientes.

Ejercicio 4.52. Generalizar la proposición 4.14 para familias linealmente independientes cualesquiera, y demostrarla.

4.3.5. Para tener presente

- En un espacio vectorial podemos formar combinaciones lineales de sus vectores.
- A partir de la noción de combinación lineal podemos extender al contexto abstracto de los espacios vectoriales cualesquiera las nociones de conjunto generador, subespacio generado, independencia lineal y dependencia lineal que habíamos introducido previamente en el contexto de los espacios vectoriales $\mathbb{K}^n$.
- Al eliminar de un generador de $\mathbb{V}$ un vector que es combinación lineal de los restantes obtenemos una nueva familia de vectores que sigue siendo un generador de $\mathbb{V}$.
- Un conjunto es linealmente independiente si y sólo si ninguno de sus vectores es una combinación lineal de los restantes.
- Al agregar a un conjunto linealmente independiente un vector que no es combinación lineal de los vectores en el conjunto, obtenemos un nuevo conjunto linealmente independiente.
- Todo generador finito de un espacio vectorial contiene un generador del espacio que es linealmente independiente.
- Si una familia $\mathcal{A}$ es un generador linealmente independiente de $\mathbb{V}$, entonces hay una única manera de escribir cualquier vector de $\mathbb{V}$ como combinación lineal de $\mathcal{A}$.

4.4. Bases y dimensión de un espacio vectorial

En esta sección continuamos con la introducción de las nociones básicas de la estructura de espacio vectorial en el contexto abstracto de espacios vectoriales cualesquiera. Tal como hicimos en la sección anterior, la mayoría de los conceptos y resultados son generalizaciones de lo expuesto en el capítulo 1 en el contexto de los espacios $\mathbb{K}^n$ y sus subespacios, aunque presentamos también algunos resultados que son necesarios para el desarrollo de la teoría abstracta, y que habían sido omitidos al trabajar en $\mathbb{K}$.

4.4.1. Bases

Comenzamos esta sección retomando el resultado del corolario 4.15, en la página 447: un generador linealmente independiente de un espacio vectorial $\mathbb{V}$ permite expresar cada vector del espacio de manera única, y pone cada vector en correspondencia con el conjunto de coeficientes de la combinación lineal que le corresponde. Tal como vimos en la sección 4.15 para los subespacios de $\mathbb{K}^n$, y adelántabamos en la sección anterior para el caso general, llamaremos *bases* del espacio vectorial a estos generadores tan destacados.

Definición 4.11 (Base de un espacio vectorial). *Sea $\mathbb{V}$ un espacio vectorial y*

$$\mathcal{A} = (v_1, v_2, \dots, v_n)$$

un subconjunto de $\mathbb{V}$. Diremos que $\mathcal{A}$ es una base *de $\mathbb{V}$ si*

1. *$\mathcal{A}$ es un generador de $\mathbb{V}$;*

2. *$\mathcal{A}$ es un conjunto linealmente independiente.*

El lector notará que se trata esencialmente de la misma definición de base que dimos en la sección 1.5 para la clase particular de espacios vectoriales que está constituida por todos los posibles subespacios de algún espacio $\mathbb{K}^n$. Ver la definición 1.8, en la página 117.

La segunda definición de esta sección introduce la noción de *coordenadas* de un vector respecto a una base $\mathcal{A} = (v_1, v_2, \dots, v_n)$, a partir del hecho de que la base pone a cada vector v del $\mathbb{K}$-espacio vectorial $\mathbb{V}$ en correspondencia con la n-upla de $\mathbb{K}^n$ formada por los coeficientes de la expresión de v como combinación lineal de $\mathcal{A}$.

Definición 4.12 (Coordenadas). *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, y*

$$\mathcal{A} = (v_1, v_2, \dots, v_n)$$

una base de $\mathbb{V}$. Sea $v \in \mathbb{V}$, entonces existe un único vector

$$(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n \tag{4.13}$$

tal que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

Llamaremos coordenadas *de v en la base $\mathcal{A}$ a este vector de $\mathbb{K}^n$.*

Indicaremos con la notación $(v)_{\mathcal{A}}$, o con $\mathrm{coord}_{\mathcal{A}}(v)$, al vector (4.13) de coordenadas de v respecto a la base $\mathcal{A}$. En la sección 4.5 discutiremos el interés que tiene introducir coordenadas de los vectores, y analizaremos las propiedades de la correspondencia $\mathrm{coord}_{\mathcal{A}}$ que a cada vector de un espacio vectorial le asocia sus coordenadas respecto a una base $\mathcal{A}$ dada del espacio.

Presentaremos ahora algunos ejemplo de bases de distintos espacios vectoriales. El primero ya apareció cuando la noción de base fue introducida en el contexto de los subespacios de los espacios $\mathbb{K}^n$ (ver el ejemplo 1.5.7, en la página 119), pero es tan importante que lo repetimos aquí.

Ejemplo 4.4.1. La base canónica de $\mathbb{K}^n$
Para $i = 1, \dots, n$ llamemos E_i al vector de $\mathbb{K}^n$ que tiene un 1 en la i-ésima posición, y ceros en las $n-1$ restantes. Llamamos *base canónica* de $\mathbb{K}^n$ a la familia

$$\mathcal{C} = \{E_1, E_2, \dots E_n\},$$

formada por estos n vectores. Para cada $X \in \mathbb{K}^n$ el vector $(X)_{\mathcal{C}}$ de sus coordenadas respecto a la base canónica es igual a X. Los detalles están desarrollados en el ejemplo 1.5.7). ♣

Ejemplo 4.4.2. En este ejemplo determinaremos todos los valores de a para los que la familia

$$((1, a), (a, 1))$$

es una base de $\mathbb{R}^2$.
Para eso necesitamos probar dos cosas: primero determinar para qué valores de a la familia genera a $\mathbb{R}^2$, y luego verificar si es una familia linealmente independiente. Veamos ahora para qué valores de a el sistema

$$(x_1, x_2) = \lambda_1(1, a) + \lambda_2(a, 1)$$

es compatible. La matriz de coeficientes del sistema es:

$$\left(\begin{array}{cc|c} 1 & a & x_1 \\ a & 1 & x_2 \end{array} \right).$$

Escalerizamos y tenemos

$$\left(\begin{array}{cc|c} 1 & a & x_1 \\ 0 & a^2-1 & ax_1-x_2 \end{array}\right).$$

Por tanto el sistema es compatible para todo x_1, x_2 solamente cuando $a^2-1 \neq 0$. En ese caso, la familia $((1,a),(a,1))$ genera a $\mathbb{R}^2$.

Cuando $a = 1$ y $a = -1$ la familia no genera a $\mathbb{R}^2$. Vemos que en esos dos casos la familia no generan $\mathbb{R}^2$

$$((1,1),(1,1))\,, \quad ((1,-1),(-1,1))\,.$$

La primera familia genera vectores en los cuales $x_1 = x_2$, y la segunda aquellos en los que $x_1 = -x_2$.

Por último verifiquemos que si $a \neq \pm 1$ la familia es linealmente independiente. Para probarlo debemos escribir al vector nulo como combinación lineal de $(1,a),(a,1)$, es decir $(0,0) = \lambda_1(1,a) + \lambda_2(a,1)$ cuya matriz de coeficientes es

$$\left(\begin{array}{cc|c} 1 & a & 0 \\ a & 1 & 0 \end{array}\right).$$

Escalerizando tenemos

$$\left(\begin{array}{cc|c} 1 & a & 0 \\ 0 & a^2-1 & 0 \end{array}\right).$$

Lo cual permite concluir que $\lambda_1 = \lambda_2 = 0$ por tanto la familia es linealmente independiente siempre que $a \neq \pm 1$.

♣

Por supuesto, el lector puede recorrer todos los ejemplos y ejercicios en los que trabajamos con bases de subespacios vectoriales de $\mathbb{K}^n$ en el capítulo 1.2, por lo que no expondremos aquí otros ejemplos de bases de subespacios vectoriales de $\mathbb{K}^n$. Al igual que $\mathbb{K}^n$, otros espacios corrientes tienen algunas bases bastante evidente, a las que es usual llamar la *base canónica* del espacio en cuestión. Presentamos a continuación las bases canónicas de los espacios de matrices y de polinomios.

Ejemplo 4.4.3. La base canónica de $M^{m\times n}(\mathbb{K})$
Consideremos una matriz cualquiera

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Esta matriz se puede escribir como combinación lineal de las matrices que tienen 1 en una entrada y 0 en todas las demás, es decir

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = a_{11} \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} + \ldots + a_{1n} \begin{pmatrix} 0 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} + \ldots$$
$$+ a_{m1} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 0 \end{pmatrix} + \ldots + a_{mn} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}.$$

Este conjunto de matrices es una base de $M^{m\times n}(\mathbb{K})$, a la que nos referiremos como la base canónica $\mathcal{C}$ de este espacio. En resumen, la base canónica del espacio de las matrices $m \times n$ es

$$\mathcal{C} = \left(\begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \ldots \begin{pmatrix} 0 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \ldots \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 0 \end{pmatrix}, \ldots \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} \right).$$

Observar que las coordenadas de una matriz M en la base $\mathcal{C}$ son

$$(M)_{\mathcal{C}} = (a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, a_{22}, \ldots, a_{2n}, \ldots, a_{m1}, a_{m2}, \ldots, a_{mn}).$$

Se trata simplemente de las filas de la matriz M, copiadas una a continuación de la otra. ♣

Ejemplo 4.4.4. La base canónica de $\mathbb{K}_n[x]$
La base canónica de $\mathbb{K}_n[x]$ es

$$\mathcal{C} = \left(1, x, x^2, \ldots, x^{n-1}, x^n\right).$$

Un polinomio p de $\mathbb{K}_n[x]$ es de la forma

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0.$$

Es claro que p se puede escribir fácilmente como combinación lineal de la base $\mathcal{C}$. En esta base las coordenadas del polinomio p son

$$(p)_{\mathcal{C}} = (a_0, a_1, a_2, \ldots, a_{n-1}, a_n)$$

Observemos que las coordenadas respecto a esta base son los coeficientes que aparecen en cada término del polinomio, pero ordenados en forma creciente según el grado. Este orden es el contrario del que habitualmente se emplea para escribir un polinomio, ya que es usual poner primero los términos de mayor grado. ♣

Ejercicio 4.53. Consideremos el espacio vectorial real

$$(\mathbb{C}^n, \mathbb{R}, +, \cdot),$$

introducido en el ejemplo 4.1.12, página 406, que tiene n-uplas de números complejos como vectores. Verifique que el conjunto

$$\{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)\}$$

no es una base. (Sugerencia: El vector $(i, 0, \ldots, 0)$ donde i es la unidad imaginaria no puede expresarse como una combinación lineal de ese conjunto). Pero

$$\{(1, 0, \ldots, 0), (i, 0, \ldots, 0), (0, 1, \ldots, 0), (0, i, \ldots, 0), \ldots, (0, 0, \ldots, 1), (0, 0, \ldots, i)\}$$

es una base de $(\mathbb{C}^n, \mathbb{R}, +, \cdot)$.

Ejemplo 4.4.5. El espacio trivial no tiene base.
Consideremos el espacio vectorial $\mathbb{V} = \{O\}$ que sólo está formado por el vector nulo O. Este subespacio no contiene ningún subconjunto linealmente independiente, porque el único subconjunto posible contiene necesariamente al vector nulo y es, por lo tanto, linealmente dependiente (ver el ejemplo 4.3.19, en la página 445). Por lo tanto el espacio trivial $\mathbb{V}$ no puede tener una base. ♣

Ejemplo 4.4.6. Hallemos una base del espacio de matrices simétricas 2×2. Recordemos que las matrices simétricas son aquellas que satisfacen $A = A^t$. Una matriz A está en este espacio si se cumple

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$$

Esto nos permite concluir que $a_{12} = a_{21}$ y que a_{11} y a_{22} pueden ser cualquiera. Tenemos entonces que A está en el espacio de las matrices simétricas si es de la forma

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = a_{11} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + a_{12} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + a_{22} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Una base del espacio de matrices simétricas 2×2 es

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Ejercicio 4.54. Hallar una base para el espacio de las matrices simétricas, reales, de dimensiones $n \times n$, donde n indica a un número natural cualquiera. Recordemos que una matriz cuadrada es *antisimétrica* cuando satisface $A^t = -A$. Hallar una base del espacio de las matrices antisimétricas, reales, de dimensiones $n \times n$.

♣

Ejercicio 4.55. Para los siguientes subespacios $\mathbb{S}$ del espacio vectorial $\mathbb{V}$ encuentre una base del mismo.

1. $\mathbb{V} = \mathbb{R}^3$, $\mathbb{S} = \{(x, y, z) \in \mathbb{R}^3;\ x + 2y - z = 0\}$.
2. $\mathbb{V} = \mathbb{R}^3$, $\mathbb{S} = \{(x, y, z) \in \mathbb{R}^3;\ x = -2y,\ z = 0\}$.
3. $\mathbb{V} = \mathbb{C}^4$, $\mathbb{S} = \{(a, b, c, d) \in \mathbb{C}^4;\ 2ia + b = 0,\ 3ic + d - ib = 0\}$.
4. $\mathbb{V} = \mathbb{R}_3[x]$, $\mathbb{S} = \{p \in \mathbb{R}_3[x];\ p(2) = 0\}$.
5. $\mathbb{V} = \mathbb{R}_3[x]$, $\mathbb{S} = \{p \in \mathbb{R}_3[x];\ p(0) = p'(0) = 0\}$.
6. $\mathbb{V} =$ sucesiones de números reales, $\mathbb{S} = \{(a_n)_{n\in\mathbb{N}};\ a_n = 0\ \forall\ n \geq 4\}$.
7. $\mathbb{V} = M_{3\times3}(\mathbb{R})$, $\mathbb{S} = \{A \in M_{3\times3}(\mathbb{R});\ A \text{ es simétrica}\}$.
8. $\mathbb{V} = M_{3\times3}(\mathbb{R})$, $\mathbb{S} = \{A \in M_{3\times3}(\mathbb{R});\ \ A \text{ es antisimétrica}\}$.

Además de identificar bases examinando la forma general de los vectores de un espacio vectorial podemos obtenerlas eliminando vectores de un generador. En ese sentido, podemos expresar la proposición 4.13 en la forma

Proposición 4.16. *Sea* $\mathbb{V}$ *un espacio vectorial y*

$$\mathcal{A} = (v_1, v_2, \ldots, v_n) \subset \mathbb{V}$$

un generador de $\mathbb{V}$*. Entonces existe un subconjunto* $\mathcal{A}'$ *de* $\mathcal{A}$ *que es una base de* $\mathbb{V}$*.*

El último ejercicio de esta subsección ofrece una caracterización de las bases.

Ejercicio 4.56. Sea $\mathbb{V}$ un espacio vectorial y $\mathcal{A}$ un subconjunto ordenado de $\mathbb{V}$. Mostrar que $\mathcal{A}$ es una base de $\mathbb{V}$ si y solo si todo vector de $\mathbb{V}$ puede escribirse de una única manera como una combinación lineal de los vectores de $\mathcal{A}$.

Ejercicio 4.57. Sean $\mathbb{S}$ y $\mathbb{T}$ dos subespacios de un espacio vectorial $\mathbb{V}$ tales que su suma es directa. Sean $\mathcal{A} = (v_1, \ldots, v_n)$ y $\mathcal{B} = (w_1, \ldots, w_m)$ bases de $\mathbb{S}$ y $\mathbb{T}$ respectivamente. Mostrar que la unión $(v_1, \ldots, v_n, w_1, \ldots, w_m)$ de ambas bases es una base de $\mathbb{S} \oplus \mathbb{T}$.

4.4.2. Dimensión

El objetivo de esta subsección es introducir la noción de *dimensión*. Al igual que hicimos en la sección 1.5 definiremos la dimensión como el número de vectores de una base (ver la definición 1.9, en la página 125). Como todos los espacios vectoriales salvo dos tienen muchas bases diferentes hay que asegurarse, igual que hicimos en el caso de $\mathbb{K}^n$ al formular el corolario 1.13, página 125, a la proposición 1.12, que dos bases diferentes tienen la misma cardinalidad. Dejamos esta tarea como ejercicio para el lector.

Ejercicio 4.58. Sea $\mathbb{V}$ un espacio vectorial. Mostrar que si el espacio tiene un generador con m vectores, y una familia linealmente independiente con n vectores, entonces $n \leq m$. Concluir que si una base de $\mathbb{V}$ contiene n vectores, entonces todas las bases tienen n vectores. Sugerencia: estudiar la demostración dada para subespacios de $\mathbb{K}^n$.

Ejercicio 4.59. ¿Cuáles son los espacios vectoriales que no tienen más de una base?

Definición 4.13 (Dimensión de un espacio vectorial). *Sea $\mathbb{V}$ un espacio vectorial. Si existe una base de $\mathbb{V}$ que tiene n elementos entonces llamaremos a n la* ***dimensión*** *de $\mathbb{V}$.*

Igual que hacíamos para los subespacios de $\mathbb{K}^n$, indicaremos con $\dim(\mathbb{V})$ la dimensión del espacio $\mathbb{V}$.

Ejemplo 4.4.7. Según esta definición tenemos que

- $\dim(\mathbb{K}^n) = n$,
- $\dim(M^{m\times n}(\mathbb{K})) = mn$,
- $\dim(\mathbb{K}_n[x]) = n + 1$,
- el espacio real de n-uplas complejas del ejercicio 4.53 tiene dimensión $2n$.

Tal como vimos en el ejemplo 4.4.5, el espacio trivial $\{O\}$ no tiene bases. Daremos una definición especial para este caso tan particular.

Definición 4.14 (Dimensión del espacio vectorial trivial). *La dimensión del espacio vectorial trivial $\mathbb{V} = \{O\}$ es igual a 0.*

Si un espacio vectorial tiene dimensión 0 o igual a un número natural n diremos que es un espacio vectorial de *dimensión finita*. Los espacios vectoriales de dimensión finita constituyen el objeto de estudio del álgebra lineal, y entender sus propiedades es uno de los objetivos centrales de este curso. Como veremos a continuación, no todos los espacios vectoriales tienen dimensión finita.

El espacio vectorial $\mathbb{K}[x]$ contiene a todos los espacios $\mathbb{K}_n[x]$. Por lo tanto contiene conjuntos linealmente independientes con una cantidad arbitrariamente grande de vectores, lo que implica que no puede tener una base con un número finito de vectores. Diremos que este espacio tiene *dimensión infinita*. Discutiremos esta situación con algo más de detalle en la subsección 4.4.3.

Ejercicio 4.60. Mostrar que si un espacio vectorial tiene un generador finito entonces tiene dimensión finita.

Tal como vimos, la dimensión de un espacio vectorial es igual a la máxima cantidad de vectores en una familia linealmente independiente, y a la mínima cantidad de vectores en un generador del espacio. Vale además la siguiente proposición.

Proposición 4.17. *Sea $\mathbb{V}$ un espacio vectorial de dimensión finita $n \geq 1$. Entonces*

1. *Todo subconjunto de $\mathbb{V}$ que sea linealmente independiente y contenga n vectores es una base de $\mathbb{V}$.*
2. *Todo subconjunto de $\mathbb{V}$ que sea un generador de $\mathbb{V}$ y contenga n vectores es una base de $\mathbb{V}$*

La demostración constituye un ejercicio interesante que obliga a revisar las nociones de base, independencia lineal y generadores, de modo que la dejamos para el lector.

Ejercicio 4.61. Demuestre la proposición 4.17.

Observación 4.4.8. La proposición 4.17 proporciona un útil criterio para determinar si un subconjunto dado de un espacio vectorial de dimensión n finita y conocida es una base. Por supuesto, para que tal subconjunto pueda ser una base debe tener n elementos. Si satisface tal criterio aún debemos determinar si es un generador y si es lineamente independiente. Pero la proposición 4.17 nos dice que es suficiente estudiar si satisface **una** de estas dos propiedades. Si una es satisfecha la otra se cumple automáticamente. Si una no lo es, la otra tampoco. Agreguemos que, en general, la más sencilla de estudiar es la independencia lineal.

La proposición 4.17 es una versión abstracta de la propiedad que dice que un sistema lineal $AX = B$ de dimensiones $n \times n$ es compatible para todo B si y sólo si es determinado siempre que es compatible. ♠

Nuestro próximo ejercicio gira en torno a la misma línea de ideas que la proposición 4.17.

Ejercicio 4.62. Demuestre que un subconjunto $\mathcal{A} = (v_1, v_2, \dots, v_n)$ de un espacio vectorial $\mathbb{V}$ es una base si y sólo si es linealmente independiente y todo subconjunto de $\mathbb{V}$ con más de n vectores es linealmente dependiente.

Si una familia linealmente independiente contenida en un espacio de dimensión n tiene menos de n vectores entonces no puede generar todo el espacio. Tal como se muestra en la parte 2 del ejercicio 4.34, si se agrega a la familia un vector que no pueda ser generado por ella se obtiene una nueva familia linealmente independiente. Tal como mostraremos a continuación, este procedimiento puede repetirse hasta conseguir un generador del espacio

Proposición 4.18. *Sean $\mathbb{V}$ un espacio vectorial de dimensión $n \geq 1$, y $\mathcal{A}$ un subconjunto linealmente independiente de $\mathbb{V}$. Entonces existe una base $\mathcal{B}$ de $\mathbb{V}$ que contiene a $\mathcal{A}$.*

Prueba. Sea p la cantidad de vectores en $\mathcal{A}$. Si $p = n$ entonces $\mathcal{A}$ es una base, y tomamos $\mathcal{B} = \mathcal{A}$.

Si $p < n$ entonces $\mathcal{A}$ no genera $\mathbb{V}$ y existe algún vector v de $\mathbb{V}$ que no es combinación lineal de $\mathcal{A}$. Agreguemos este vector v a la familia $\mathcal{A}$, para obtener una familia $\mathcal{A}^{(1)}$ que es linealmente independiente y tiene $p + 1$ vectores. Si $p+1 = n$ entonces $\mathcal{A}^{(1)}$ es una base y tomamos $\mathcal{B} = \mathcal{A}^{(1)}$. Si no lo es repetimos el procedimiento de adjuntar un vector.

Repitiendo el procedimiento una cantidad finita de veces, hasta conseguir un conjunto linealmente independiente con n vectores, llegamos a completar una base del espacio. □

Observación 4.4.9. La proposición anterior sugiere un procedimiento práctico para extender un conjunto linealmente independiente a una base: ir agregando vectores de uno en uno hasta completar la base. Un procedimiento alternativo consiste en adjuntar a la familia $\mathcal{A}$ toda una base $\mathcal{B}$ de $\mathbb{V}$. Se obtiene así un generador linealmente dependiente del espacio, del que se pueden eliminar tantos vectores de $\mathcal{B}$ como sea necesario para obtener una base, aplicando un procedimiento similar al que se explica en la página 440, al final del ejemplo 4.3.14. ♠

Ejercicio 4.63. Dado el subconjunto $\mathcal{A} = ((1, -2, 5, -3), (0, 7, -9, 2))$ de $\mathbb{R}^4$,

1. probar que $\mathcal{A}$ es un conjunto linealmente independiente.
2. ¿Qué condiciones deben cumplir x, y, z, t para que el vector $v = (x, y, z, t)$ esté en el subespacio generado por $\mathcal{A}$?

3. Completar el conjunto $\mathcal{A}$ hasta obtener una base de $\mathbb{R}^4$.

Ejercicio 4.64. Sea $\mathbb{S} = [(1,-3,2),\ (2,-4,1),\ (1,-5,5)] \subset\ \mathbb{R}^3$.

1. Hallar una base de $\mathbb{S}$.
2. Agregar vectores a la base hallada hasta obtener una base de $\mathbb{R}^3$. Emplear los dos procedimientos esbozados en la observación 4.4.9.

Ejercicio 4.65.

1. Consideremos un subconjunto $\mathcal{A} = (v_1, v_2, \dots, v_n)$ de un espacio vectorial $\mathbb{V}$ de dimensión finita n. Diseñar algún procedimiento eficiente para extraer de $\mathcal{A}$ una base de $[\mathcal{A}]$ y luego completarla a una base de $\mathbb{V}$.
2. Aplicar el procedimiento ideado en la parte anterior para el subconjunto $\mathcal{A} = (x^3 + x^2 - 1, x^2, 2, 2x^2 - 4)$.
3. Hacerlo también para $\mathcal{A} = ((1,0,1,1), (-1,1,0,0), (0,1,1,1), (1,3,2,1))$.

La dimensión de un espacio vectorial es una medida del "tamaño" del espacio, medido en el sentido de la estructura lineal. En este sentido[12] $\mathbb{R}^3$, con sus tres dimensiones, es "más grande" que $\mathbb{R}^2$, que es un espacio de dimensión 2. A continuación presentamos dos resultados que ilustran esta idea.

Proposición 4.19. *Sean $\mathbb{V}$ un espacio vectorial de dimensión finita n y $\mathbb{S}$ un subespacio de $\mathbb{V}$. Entonces $\mathbb{S}$ tiene también dimensión finita, que es menor o igual que n.*

Prueba. Si $\mathbb{S} = \{O\}$, entonces $\dim(\mathbb{S}) = 0$ y hemos completado la prueba, pues $0 \leq n$.

Si $\mathbb{S} \neq \{O\}$, alcanza con probar que existe una base de $\mathbb{S}$ con una cantidad finita de elementos p. Una vez que hayamos demostrado este hecho, por la proposición 1.12 resulta que $p \leq n$, porque una base de $\mathbb{S}$ es un conjunto linealmente independiente de $\mathbb{V}$. La proposición 1.12 asegura que la cantidad de vectores en un conjunto linealmente independiente es menor o igual que la cantidad de elementos de un generador.

En este caso, como $\mathbb{S} \neq \{O\}$ existe algún vector $v_1 \neq O$ en $\mathbb{S}$. Si (v_1) genera a $\mathbb{S}$ entonces es una base, y hemos completado la prueba. Si (v_1) no es una base entonces iremos ampliando esta familia hasta obtener una base, tal como explicamos a continuación.

Si (v_1) no genera $\mathbb{S}$ existe un vector $v_2 \in \mathbb{S}$ que no es combinación lineal de v_1. Por lo tanto (v_1, v_2) es linealmente independiente, porque ningún vector

[12]Sin embargo el tamaño $\mathbb{R}^2$ y $\mathbb{R}^3$ medido con su cardinalidad como conjuntos es el mismo: ambos tiene exactamente la misma cantidad de elementos, porque es posible encontrar una función biyectiva entre $\mathbb{R}^2$ y $\mathbb{R}^3$.

en la familia es una combinación lineal de los anteriores. Si (v_1, v_2) genera a $\mathbb{S}$, ya tenemos una base.

En caso contrario, existe un tercer vector $v_3 \in \mathbb{S}$ que no es combinación lineal de los dos primeros. Por lo tanto el conjunto (v_1, v_2, v_3) es linealmente independiente. Si este conjunto es un generador de $\mathbb{S}$ ya hemos conseguido una base de $\mathbb{S}$.

Si no es un generador agregamos un nuevo vector, y repetimos el razonamiento. Este proceso genera en el i-ésimo paso una familia linealmente independiente con i vectores. Por lo tanto debe detenerse antes del paso $n + 1$, porque en un espacio de dimensión n no existe ninguna familia linealmente independiente con $n + 1$ vectores. En el momento en el que la construcción se detiene es porque hemos encontrado una base de $\mathbb{S}$ □

Observación 4.4.10. Vale la pena notar que hemos demostrado que un subespacio de un espacio de dimensión finita **tiene** una base finita, y por lo tanto **tiene** dimensión. No habíamos demostrado antes este resultado, y tampoco es muy útil a los efectos prácticos, porque cuando se trabaja con un subespacio particular, explícito, siempre es relativamente fácil encontrarle una base. Pero es necesario para el desarrollo de la teoría.

Si el lector revisa el enunciado de la proposición 1.16, página 128, acerca de un subespacio $\mathbb{S}$ contenido en un subespacio $\mathbb{T}$, notará que hemos incluido entre las hipótesis el requisito de que $\mathbb{S}$ tuviera dimensión. Esta hipótesis es innecesaria, tal como se desprende de la proposición 4.19. ♠

Ejercicio 4.66. Sea $\mathbb{V}$ un espacio vectorial de dimensión finita, y $\mathbb{S}$ un subespacio de $\mathbb{V}$. Mostrar que si $\dim(\mathbb{S}) = \dim(\mathbb{V})$ entonces $\mathbb{S} = \mathbb{V}$.

Ejercicio 4.67. Sean $\mathbb{S}$ y $\mathbb{T}$ dos subespacios con dimensión finita de un espacio vectorial $\mathbb{V}$, tales que su suma es directa. Mostrar que

$$\dim(\mathbb{S} \oplus \mathbb{T}) = \dim(\mathbb{S}) + \dim(\mathbb{T})$$

En un espacio $\mathbb{V}$ de dimensión finita todo subespacio $\mathbb{S}$ tiene un complemento $\mathbb{T}$ tal que $\mathbb{V} = \mathbb{S} \oplus \mathbb{T}$. Este complemento no es único. Estas propiedades se analizan en nuestro próximo ejercicio.

Ejercicio 4.68. Complementos de subespacios vectoriales

1. Sea $\mathbb{V}$ un espacio vectorial de dimensión finita, y $\mathbb{S}$ uno de sus subespacios. Mostrar que existe un subespacio $\mathbb{T}$ de $\mathbb{V}$ tal que

$$\mathbb{V} = \mathbb{S} \oplus \mathbb{T}.$$

¿Es único este subespacio?

2. Para el subespacio de $\mathbb{R}^3$ definido por la ecuación
$$x + y + z = 0$$
hallar todos los posibles complementos T como en la parte 1. Interpretar geométricamente el resultado.

Cerremos la serie de ejercicios dedicados a la suma, con un cálculo de dimensiones. Observemos como para espacios vectoriales la dimensión se comporta respecto a la operación de suma de subespacios de una manera parecida a la cardinalidad (cantidad de elementos de un conjunto) respecto a la unión de conjuntos finitos.

Ejercicio 4.69. Sean $\mathbb{S}$ y $\mathbb{T}$ dos subespacios con dimensión finita de un espacio vectorial $\mathbb{V}$.

1. Sea $\mathcal{A} = (u_1, \ldots, u_l)$ una base de $\mathbb{S} \cap \mathbb{T}$. Mostrar que existen bases
$$\begin{aligned} \mathcal{B} &= (u_1, \ldots, u_l, v_1, \ldots, v_m) \\ \mathcal{C} &= (u_1, \ldots, u_l, w_1, \ldots, w_n) \end{aligned}$$
de $\mathbb{S}$ y $\mathbb{T}$ respectivamente, que extienden a la base $\mathcal{A}$.

2. Mostrar que $(u_1, \ldots, u_l, v_1, \ldots, v_m, w_1, \ldots, w_n,)$ es una base de $\mathbb{S} + \mathbb{T}$.

3. Concluir que
$$\dim(\mathbb{S} + \mathbb{T}) = \dim(\mathbb{S}) + \dim(\mathbb{T}) - \dim(\mathbb{S} \cap \mathbb{T}).$$

4.4.3. Espacios de dimensión infinita

Hay espacios vectoriales que no pueden tener un generador finito, porque tienen familias linealmente independientes arbitrariamente grandes. Un ejemplo de estos espacios es $\mathbb{K}[x]$, formado por todos los polinomios sobre un cuerpo $\mathbb{K}$. Diremos que estos espacios tienen *dimensión infinita*.

También para estos espacios tiene sentido la noción de base. Al igual que para espacios de dimensión finita, diremos que un generador linealmente independiente es una base del espacio.

Ejemplo 4.4.11. La familia infinita de polinomios $\{1, x, x^2, x^3, \ldots\}$ es una base de $\mathbb{K}[x]$.

Ejercicio 4.70. Hallar una base del espacio vectorial real formado por todas las sucesiones reales $\{a_n\}_{n\in\mathbb{N}}$.

Sin embargo, digamos que en los espacios vectoriales de dimensión infinita es mucho más interesante introducir nociones de convergencia que nos permitan considerar límites de sucesiones en el espacio, y, en particular, "sumas infinitas". Por supuesto, las nociones de convergencia son también importantes en el contexto de los espacios vectoriales de dimensión finita (seguramente el lector tiene cierta familiaridad con problemas de análisis sobre la recta real, que es un espacio vectorial de dimensión 1). De todos modos, este tipo de problemas nos introduce al dominio del análisis matemático, y escapa al alcance de este texto.

4.4.4. Para tener presente

- Todo espacio vectorial que admita un generador finito tiene una base finita, y dimensión finita.
- Un subespacio de un espacio vectorial de dimensión finita tiene dimensión finita, que es menor o igual que la del espacio que lo contiene. Si la dimensión es igual entonces el subespacio coincide con el espacio que lo contiene.
- Todo generador de un espacio vectorial de dimensión finita contiene una base del espacio.
- Todo conjunto linealmente independiente en un espacio vectorial de dimensión finita puede extenderse hasta completar una base del espacio. Esta extensión no es única.
- Las nociones de independencia lineal y generadores pueden extenderse a los espacios de dimensión infinita, y a familias infinitas. Para trabajar en estos espacios la estructura lineal es, en general, insuficiente, y es necesario agregar alguna noción de convergencia. No ahondaremos en esta cuestión en este texto.

4.5. Coordenadas en una base

En esta sección vamos a comenzar a explotar la posibilidad de referir los vectores de un espacio vectorial de dimensión finita n a una base del espacio, para representarlos por medio de vectores de $\mathbb{K}^n$. Observaremos en primer lugar el hecho de que distintas bases dan lugar a distintas representaciones, y algunas son más adecuadas para calcular en una situación concreta que otras. Luego mostraremos que el pasaje a coordenadas representa de manera fiel todas las propiedades de la estructura lineal del espacio, y utilizaremos esta idea en algunos ejemplos. Más adelante desarrollaremos una teoría sistemática de los cambios de base, o cambios lineales de coordenadas. También emplearemos las coordenadas para representar matricialmente las transformaciones lineales entre espacios vectoriales de dimensión finita, y explotaremos la posibilidad de construir distintas representaciones de una misma transformación por medio de cambios de base.

4.5.1. Eligiendo coordenadas adecuadas

En un espacio vectorial $\mathbb{V}$ de dimensión finita, independientemente de cuál sea la naturaleza de los vectores que lo forman, cada vector queda identificado por una lista de $\dim(\mathbb{V})$ números al fijar una base del espacio. En este sentido, las coordenadas proveen entonces una representación sencilla, y fácil de manejar, de los vectores de un espacio vectorial $\mathbb{V}$.

Pero un mismo espacio vectorial tiene muchas bases, y cada base tiene asociada con ella una manera de asignar coordenadas a los vectores. En muchos problemas, escoger una base adecuada para representar los vectores del espacio puede significar una gran simplificación adicional.

Ejemplo 4.5.1. Volvemos sobre el conjunto $\mathbb{L}$ de las funciones reales definidas sobre el intervalo $[-1, 1]$, continuas, lineales entre los nodos -1, 0 y 1, que forman un espacio vectorial real. En ese espacio la familia (l_{-1}, l_0, l_1) (ver la definición de estas funciones en la página 397) es una base que permite representar cualquier función con muy poco esfuerzo.

Las coordenadas de una función $f \in \mathbb{L}$ respecto a esta base son el vector

$$(f(-1), f(0), f(1)),$$

y calcularlas se reduce a hacer evaluaciones de la función.

Naturalmente, las coordenadas de un vector dependen de cuál sea la base escogida, y otras bases podrían requerir cálculos diferentes, más complejos.

Ejercicio 4.71. Mostrar que la familia de funciones $(1, x, |x|)$ también es una base de $\mathbb{L}$, y hallar las coordenadas de una función $f \in \mathbb{L}$ respecto a esta nueva base. En particular, calcular las coordenadas de l_{-1},l_0 y l_1 en esta base.

Las funciones de $\mathbb{L}$ aparecieron como respuesta a un problema de interpolación. En el ejemplo 4.7.6 de la página 517 discutiremos cómo diferentes bases de los espacios de polinomios reales son adecuadas para realizar distintos tipos de interpolación con polinomios.

En nuestro siguiente ejemplo mostramos como distintas bases de $\mathbb{R}^3$ originan distintas ternas de coordenadas para un mismo vector.

Ejemplo 4.5.2. Los espacios $\mathbb{K}^n$ tienen una base canónica $\mathcal{C}$ que tiene la propiedad de que $\operatorname{coord}_{\mathcal{C}}(X) = X$ para todo $X \in \mathbb{K}^n$ (ver el ejemplo 4.4.1, en la página 457).

Pero las coordenadas de un vector de $\mathbb{K}^n$ respecto a otra base son diferentes al vector original. Ilustramos este hecho con un ejemplo en $\mathbb{R}^3$. Consideremos la base

$$\mathcal{B} = ((1,1,1),(1,1,0),(1,0,0))$$

de $\mathbb{R}^3$, y tomemos un vector genérico $X = (x_1, x_2, x_3) \in \mathbb{R}^3$. Como el lector podrá verificar rápidamente, se tiene que

$$X = x_3(1,1,1) + (x_2 - x_3)(1,1,0) + (x_1 - x_2)(1,0,0).$$

Por lo tanto el vector X tiene coordenadas $(x_3, x_2 - x_3, x_1 - x_2)$ en la base $\mathcal{B}$ y coordenadas (x_1, x_2, x_3) en la base canónica de $\mathbb{R}^3$.

Observación 4.5.3. El orden en que aparecen los coeficientes al tomar las coordenadas está determinado por el orden de los vectores de la base. Es por esto que el orden es importante. Esta fue una de las razones que nos llevó a incluir en la definición de base el requerimiento del orden de los vectores. Si cambiamos el orden de los vectores de la base, entonces las coordenadas cambian ya que se modifica el orden en que aparecen las coordenadas en el vector. ♣

Vemos que también en $\mathbb{K}^n$ podemos identificar un vector cualquiera por sus coordenadas en una base dada. En una primera mirada, puede parecer que no tiene interés representar un vector $\mathbb{K}^n$ por su vector de coordenadas, ya que se trata de reemplazar un objeto por otro de la misma naturaleza. Sin embargo, la posibilidad de escoger dentro de un mismo espacio diferentes bases, o sistemas de coordenadas, es esencial para calcular: muchos problemas se vuelven tratables al trabajar con ellos en una base adecuada al problema.

Ejemplo 4.5.4. Consideremos en $\mathbb{R}^3$ el plano de ecuación[13]

$$x_1 + x_2 + x_3 = 1.$$

Busquemos caracterizar el plano por una condición referida a un sistema de coordenadas distinto de la base canónica de $\mathbb{R}^3$, que sea más afín a la geometría del problema. Escogeremos entonces una base (U, V, W) donde W es perpendicular al plano, y los vectores U y V paralelos a él. El vector W tiene que ser paralelo al vector $(1, 1, 1)$. Como es especialmente conveniente trabajar con bases ortonormales (ver, por ejemplo, la parte 5 del ejercicio 3.53, en la página 328) escogemos

$$W = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

El vector $(1, -1, 0)$ es perpendicular a W, y luego de normalizarlo dividiendo entre su longitud $\sqrt{1^2 + 1^2 + 0^2}$ decidimos que

$$U = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)$$

será el primer vector de la base. El segundo vector V puede obtenerse calculando alguno de los productos vectoriales $U \wedge W$ o $W \wedge U$, o a ojo. Preferimos este último procedimiento porque rápidamente notamos que $(1, 1, -2)$ es perpendicular a los dos vectores anteriores. Entonces escogemos

$$V = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right).$$

Si un vector (x_1, x_2, x_3) tiene coordenadas y_1, y_2, y_3 en esta base entonces

$$\begin{aligned} x_1 &= \tfrac{y_1}{\sqrt{2}} + \tfrac{y_2}{\sqrt{6}} + \tfrac{y_3}{\sqrt{3}}, \\ x_2 &= -\tfrac{y_1}{\sqrt{2}} + \tfrac{y_2}{\sqrt{6}} + \tfrac{y_3}{\sqrt{3}}, \\ x_3 &= -\tfrac{2y_2}{\sqrt{6}} + \tfrac{y_3}{\sqrt{3}}. \end{aligned}$$

Por lo tanto

$$x_1 + x_2 + x_3 = \sqrt{3}y_3,$$

y en este sistema de coordenadas el plano queda caracterizado por la ecuación

$$y_3 = 1/\sqrt{3}.$$

Esta ecuación es más simple que la original, porque hemos escogido un referencial que para este problema en particular es más adecuado que la base canónica. ♣

[13]Dejamos por un momento de lado la notación (x, y, z) para los vectores de $\mathbb{R}^3$, y recurrimos a la más sistemática notación con subíndices.

Ejemplo 4.5.5. Conejos y matrices Para la resolución del ejercicio 2.5 es especialmente útil comprender como actúa la matriz

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

cuando multiplica a un vector $X \in \mathbb{R}^2$. Existen vectores para los que la acción de A es particularmente simple:

1. sobre cualquier vector colineal con $v_1 = (1+\sqrt{5}, 2)$ la matriz actúa como una multiplicación por la constante $\lambda_1 = (1+\sqrt{5})/2$. Esto es algo que aprendimos en la parte 4 del ejercicio 2.5, página 166, pero que podemos verificar directamente. Calculamos el producto y obtenemos

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1+\sqrt{5} \\ 2 \end{pmatrix} = \begin{pmatrix} 3+\sqrt{5} \\ 1+\sqrt{5} \end{pmatrix} = \frac{1+\sqrt{5}}{2} \begin{pmatrix} 1+\sqrt{5} \\ 2 \end{pmatrix};$$

2. el efecto de multiplicar por A cualquier vector colineal con

$$v_2 = (1-\sqrt{5}, 2)$$

se reduce al de multiplicar el vector por la constante $\lambda_2 = (1-\sqrt{5})/2$. Omitimos el cálculo detallado, que es muy parecido al que acabamos de hacer.

Es muy fácil mostrar que $\mathcal{B} = (v_1, v_2)$ forma una base de $\mathbb{R}^2$. Por ejemplo, calculando el producto escalar $v_1 \cdot v_2$ para ver que v_1 es perpendicular a v_2, o usando cualquier otro argumento que satisfaga al lector. Por lo tanto, cualquier vector X de $\mathbb{R}^2$ admite una expresión

$$X = y_1 v_1 + y_2 v_2,$$

donde $(y_1, y_2) = \operatorname{coord}_{\mathcal{B}}(X)$. Tenemos entonces que

$$AX = y_1 A v_1 + y_2 A v_2 = \lambda_1 y_1 v_1 + \lambda_2 y_2 v_2,$$

por lo que

$$\operatorname{coord}_{\mathcal{B}}(AX) = (\lambda_1 y_1, \lambda_2 y_2). \tag{4.14}$$

Observemos que la acción de A respecto a las coordenadas en la base $\mathcal{B}$ es especialmente simple: multiplica cada una de las coordenadas por una de las dos constantes λ_i, $i = 1, 2$. Si representamos los vectores de coordenadas como columnas, la igualdad (4.14) puede escribirse en notación matricial como $\mathrm{coord}_{\mathcal{B}}(AX) = D\mathrm{coord}_{\mathcal{B}}(X)$, donde D es la matriz diagonal

$$D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

La acción de A vista a través del sistema de coordenadas $\mathcal{B}$ se reduce a la multiplicación por una matriz diagonal, un objeto mucho más simple. Es muy fácil, por ejemplo, calcular $A^{1000}X$ o entender el comportamiento asintótico de la sucesión A^nX cuando $n \to \infty$ a partir de esta representación. ♣

Ejemplo 4.5.6. El estándar JPEG
Consideremos una imagen en blanco y negro que se representa en la pantalla de una computadora. La imagen está definida por el valor de gris en cada pixel de la pantalla. En principio este valor es discreto, pero aún así podemos considerar que es un número real. El estándar JPEG proporciona una manera eficiente de codificar toda la información necesaria para representar la imagen. A continuación describimos sus características más importantes.

La imagen se divide en bloques de ocho por ocho píxeles. La información que corresponde a cada bloque consiste entonces en 64 números reales, o un vector v de $\mathbb{R}^{64}$. En vez de dar los 64 números directamente (que sería equivalente a dar un mapa de los valores de gris pixel a pixel) JPEG representa las coordenadas de este vector v respecto a una base $\mathcal{B}$ que tiene la propiedad de que la mayoría de las imágenes naturales generan vectores $\mathrm{coord}_{\mathcal{B}}(v)$ que tiene sus componentes muy concentradas en unas pocas posiciones. Esta concentración se aprovecha para conservar sólo unas pocas coordenadas "grandes" y despreciar las "pequeñas" sin perder demasiada información sobre la imagen.

El umbral que separa a las coordenadas "grandes" de las "pequeñas" es más o menos arbitrario, y se fija en función de la calidad que se pretenda obtener, o de cuánto se quiera comprimir la información de la imagen. De hecho, muchos programas ofrecen al usuario la posibilidad de controlar este parámetro.

Además de las consideraciones de tamaño de las distintas componentes del vector que representa la imagen en la base que emplea JPEG, el estándar utiliza información acerca de que tan sensible es el ojo humano a cada una de estas componentes. Digamos que tiende a despreciar más aquellas componentes que el ojo ve peor. Para cerrar estos comentarios digamos que la base que utiliza JPEG está ordenada en función de la "frecuencia" que corresponde a la imagen que representa cada elemento de la base. Las imágenes en las que los tonos de gris varían suavemente están antes en la base que las imágenes que sufren grandes variaciones en los valores de gris entre un píxel y los adyacentes. ♣

Ejemplo 4.5.7. SEÑALES Y RUIDO
Una señal, por ejemplo, de sonido puede ser modelada como un vector de un espacio vectorial adecuado. Las señales suelen tener "superpuestos" (sumados) el sonido que nos interesa y el ruido que no nos interesa. Si somos capaces de representar el vector que representa a la señal en una base donde el sonido deseado y el ruido estén bien "separados" podremos librarnos del enojoso ruido, simplemente sustituyendo por cero las coordenadas que corresponden al ruido. ♣

4.5.2. La transformación $\mathrm{coord}_{\mathcal{B}}$ y su inversa.

En la definición 4.12, página 456 de la sección 4.4, habíamos destacado en el hecho de que una base

$$\mathcal{B} = (v_1, v_2, \dots, v_n)$$

en un espacio vectorial $\mathbb{V}$ de dimensión finita n permite identificar a cada vector del espacio por medio de una lista de n escalares, un vector de $\mathbb{K}^n$, que hemos dado en llamar las coordenadas $(v)_{\mathcal{B}}$ de v en la base $\mathcal{B}$.

Dado un vector $v \in \mathbb{V}$ hay un único vector

$$(v)_{\mathcal{B}} = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n$$

tal que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

Recíprocamente, dado un vector $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n$ existe un único vector

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

tal que

$$(v)_{\mathcal{B}} = (\lambda_1, \lambda_2, \dots, \lambda_n).$$

Ejemplo 4.5.8. Consideremos en $\mathbb{K}_n[x]$ la base "canónica",

$$\mathcal{C} = \left(1, x, x^2, \dots, x^n\right).$$

Las coordenadas de

$$p[x] = a_n x^n + a_{n-1} x^{n-1} + \ldots a_2 x^2 + a_1 x + a_0$$

es la $n+1$-upla

$$(a_0, a_1, a_2, \dots a_{n-1}, a_n).$$

Por ejemplo, el vector $(1, 1, 0, 0, \dots, 0, 0, 1)$, que contiene ceros en todas las posiciones que no aparecen explícitamente, es el vector de coordenadas del polinomio

$$p(x) = x^n + x + 1$$

en la base canónica de $\mathbb{K}_n[x]$. ♣

Ejemplo 4.5.9. Cada matriz

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

en $M^{2\times 3}(\mathbb{K})$ queda automáticamente identificada por el vector

$$(a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}) \in \mathbb{K}^6$$

cuando se trabaja en $M^{2\times 3}(\mathbb{K})$ con la base canónica. ♣

Pero las coordenadas son mucho más que "etiquetas", porque son vectores de $\mathbb{K}^n$, un espacio con estructura vectorial en el que se puede sumar y multiplicar por escalares. Nuestra próxima proposición muestra que las operaciones lineales de suma y producto por un escalar en un $\mathbb{K}$-espacio $\mathbb{V}$ de dimensión n son fielmente reflejadas por las coordenadas en $\mathbb{K}^n$.

Para continuar adelante con nuestra discusión enfatizaremos el hecho de que la asignación de coordenadas en una base

$$\mathcal{B} = (v_1, v_2, \dots, v_n)$$

de un $\mathbb{K}$-espacio vectorial de dimensión n define una función

$$\mathrm{coord}_{\mathcal{B}} : \mathbb{V} \to \mathbb{K}^n.$$

Por supuesto, el vector de las coordenadas $(v)_{\mathcal{B}}$ de un cierto vector v respecto a la base $\mathcal{B}$ es el valor $\mathrm{coord}_{\mathcal{B}}(v)$ que toma la función $\mathrm{coord}_{\mathcal{B}}$ al aplicarla a v.

Sabemos que a vectores distintos del espacio le corresponden coordenadas distintas, porque las coordenadas determinan completamente al vector. En el lenguaje habitual para la teoría de funciones esto se expresa diciendo que $\mathrm{coord}_{\mathcal{B}}$ es inyectiva. Por otra parte, cualquier vector de $\mathbb{K}^n$ es el vector de coordenadas de algún vector $v \in \mathbb{V}$. Por lo tanto $\mathrm{coord}_{\mathcal{B}}$ es sobreyectiva. Una función que es a la vez inyectiva y sobreyectiva es lo que llamamos una función biyectiva[14].

Toda función biyectiva tiene una inversa. En el caso de las coordenadas la aplicación
$$(\lambda_1, \lambda_2, \dots, \lambda_n) \mapsto \lambda_1 v_1 + \lambda_2 v_2 + \dots, \lambda_n v_n,$$
que a cada n-upla de $\mathbb{K}^n$ le asocia un vector de $\mathbb{V}$ es la inversa
$$\mathrm{coord}_{\mathcal{B}}^{-1} : \mathbb{K}^n \to \mathbb{V}$$
de la función $\mathrm{coord}_{\mathcal{B}}$. En la próxima proposición explicitamos las buenas propiedades de $\mathrm{coord}_{\mathcal{B}}$ respecto a la estructura lineal del espacio.

Proposición 4.20. *Sean $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión finita n y $\mathcal{B} = (v_1, v_2, \dots, v_n)$ una base de $\mathbb{V}$. Entonces la correspondencia*
$$\mathrm{coord}_{\mathcal{B}} : \mathbb{V} \to \mathbb{K}^n$$
es biyectiva, y tanto $\mathrm{coord}_{\mathcal{B}}$ como su inversa $\mathrm{coord}_{\mathcal{B}}^{-1}$ son transformaciones lineales, en el sentido de que la transformación $\mathrm{coord}_{\mathcal{B}}$ satisface

1. $\mathrm{coord}_{\mathcal{B}}(w_1 + w_2) = \mathrm{coord}_{\mathcal{B}}(w_1) + \mathrm{coord}_{\mathcal{B}}(w_2)$*, para* $w_1, w_2 \in \mathbb{V}$*;*
2. $\mathrm{coord}_{\mathcal{B}}(\lambda w) = \lambda \mathrm{coord}_{\mathcal{B}}(w)$*, para* $\lambda \in \mathbb{K}$ *y* $w \in \mathbb{V}$*.*

Y la transformación $\mathrm{coord}_{\mathcal{B}}^{-1}$ satisface

1. $\mathrm{coord}_{\mathcal{B}}^{-1}(X + Y) = \mathrm{coord}_{\mathcal{B}}^{-1}(X) + \mathrm{coord}_{\mathcal{B}}^{-1}(Y)$*, para* $X, Y \in \mathbb{K}^n$*;*
2. $\mathrm{coord}_{\mathcal{B}}^{-1}(\lambda X) = \lambda \mathrm{coord}_{\mathcal{B}}^{-1}(X)$*, para* $\lambda \in \mathbb{K}$ *y* $X \in \mathbb{K}^n$*.*

PRUEBA. Los comentarios que preceden al enunciado de la proposición muestran que $\mathrm{coord}_{\mathcal{B}}$ es una transformación biyectiva entre $\mathbb{V}$ y $\mathbb{K}^n$. Sólo resta mostrar la linealidad de esta transformación y de su inversa.

Comenzaremos mostrando la linealidad de $\mathrm{coord}_{\mathcal{B}}^{-1}$, porque para esta aplicación tenemos la fórmula explícita
$$\mathrm{coord}_{\mathcal{B}}^{-1}(x_1, x_2, \dots, x_n) = x_1 v_1 + x_2 v_2 + \dots x_n v_n.$$
Apliquemos $\mathrm{coord}_{\mathcal{B}}^{-1}$ a la suma
$$(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

[14]El lector que necesite repasar el significado de este término puede mirar las discusiones de los conceptos de inyectividad, sobreyectividad y biyectividad en las páginas 527 y siguientes de la sección 4.8.

de

$$X = (x_1, x_2, \dots, x_n), \quad Y = (y_1, y_2, \dots, y_n).$$

Obtenemos

$$\text{coord}_{\mathcal{B}}^{-1}(X + Y) = (x_1 + y_1)v_1 + (x_2 + y_2)v_2 + \dots (x_n + y_n)v_n.$$

Si empleamos en cada sumando la propiedad distributiva del producto entre escalares y vectores respecto a la suma de escalares, y luego reordenamos la suma para agrupar entre sí los términos que provienen de X y los que provienen de Y obtenemos

$$\text{coord}_{\mathcal{B}}^{-1}(X + Y) = (x_1v_1 + x_2v_2 + \dots + x_nv_n) + (y_1v_1 + y_2v_2 + \dots + y_nv_n).$$

Reconociendo en el miembro de la derecha los vectores de $\mathbb{V}$ cuyas coordenadas son X e Y concluimos

$$\text{coord}_{\mathcal{B}}^{-1}(X + Y) = \text{coord}_{\mathcal{B}}^{-1}(X) + \text{coord}_{\mathcal{B}}^{-1}(Y).$$

Análogamente

$$\text{coord}_{\mathcal{B}}^{-1}(\lambda X) = (\lambda x_1)v_1 + \dots + (\lambda x_n)v_n = \lambda(x_1v_1 + \dots + x_nv_n).$$

Tenemos entonces que

$$\text{coord}_{\mathcal{B}}^{-1}(\lambda X) = \lambda \text{coord}_{\mathcal{B}}^{-1}(X).$$

Una vez que hemos probado la linealidad de $\text{coord}_{\mathcal{B}}^{-1}$, ésta implica la linealidad de $\text{coord}_{\mathcal{B}}$, tal como mostraremos a continuación. Consideremos dos vectores v_1 y v_2 en $\mathbb{V}$, tales que

$$\text{coord}_{\mathcal{B}}(v_1) = X, \quad \text{coord}_{\mathcal{B}}(v_2) = Y.$$

Como cada una de las transformaciones $\text{coord}_{\mathcal{B}}$ y $\text{coord}_{\mathcal{B}}^{-1}$ es la inversa de la otra, estas igualdades son equivalentes a

$$v_1 = \text{coord}_{\mathcal{B}}^{-1}(X), \quad v_2 = \text{coord}_{\mathcal{B}}^{-1}(Y).$$

Al sumar obtenemos

$$v_1 + v_2 = \text{coord}_{\mathcal{B}}^{-1}(X) + \text{coord}_{\mathcal{B}}^{-1}(Y),$$

pero la linealidad de $\text{coord}_{\mathcal{B}}^{-1}$ implica

$$v_1 + v_2 = \text{coord}_{\mathcal{B}}^{-1}(X + Y),$$

que es equivalente a

$$\mathrm{coord}_{\mathcal{B}}(v_1 + v_2) = X + Y.$$

La prueba de que

$$\mathrm{coord}_{\mathcal{B}}(\lambda v) = \lambda \mathrm{coord}_{\mathcal{B}}(v)$$

es similar, y queda como un ejercicio para el lector. □

Ejercicio 4.72. Demostrar que

$$\mathrm{coord}_{\mathcal{B}}(\lambda v) = \lambda \mathrm{coord}_{\mathcal{B}}(v)$$

La linealidad de la operación de tomar coordenadas implica que las coordenadas de una combinación lineal de vectores son la combinación lineal de las coordenadas, tal como se expresa el siguiente corolario a la proposición 4.20.

Corolario 4.21. *Sea $\mathcal{B}$ una base de un espacio vectorial de dimensión finita $\mathbb{V}$, y w_i, λ_i, $i = 1, \ldots, m$, m vectores de $\mathbb{V}$ y m escalares, respectivamente. Mostrar que*

$$(\lambda_1 w_1 + \ldots + \lambda_m w_m)_{\mathcal{B}} = \lambda_1 (w_1)_{\mathcal{B}} + \ldots + \lambda_m (w_m)_{\mathcal{B}}$$

Ejercicio 4.73. Demostrar el corolario 4.21

La proposición 4.20 y su corolario implican que es lo mismo hacer operaciones de suma y producto por un escalar, o tomar combinaciones lineales, en un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ de dimensión n, y luego tomar coordenadas para pasar a $\mathbb{K}^n$, que pasar primero a $\mathbb{K}^n$ y hacer allí las operaciones.

En realidad, cuando se calcula explícitamente en un espacio vectorial, prácticamente siempre se hace con las coordenadas de los vectores respecto a alguna base elegida implícita o explícitamente. Este hecho ya ha aparecido, aunque no lo pusimos en evidencia en su momento, en muchos de los ejemplos tratados. Ahora volveremos a uno de ellos, para mostrar cómo nuestros cálculos eran cálculos con coordenadas.

Ejemplo 4.5.10. En el ejemplo 4.3.5 de la página 4.3.5 estudiábamos cuándo un polinomio

$$p(x) = ax^2 + bx + c$$

puede expresarse como combinación lineal de la familia

$$\mathcal{A} = \left(x^2 - x + 1, x^2 + x + 1, x^2 + 1\right),$$

y cuáles son las posibles maneras de escribir la combinación.

Nos planteábamos la igualdad

$$ax^2 + bx + c = \lambda_1(x^2 - x + 1) + \lambda_2(x^2 + x + 1) + \lambda_3(x^2 + 1)$$

y luego de ordenar según potencias crecientes de x concluíamos que era equivalente a las tres igualdades escalares

$$\left\{ \begin{array}{r} \lambda_1 + \lambda_2 + \lambda_3 = a, \\ -\lambda_1 + \lambda_2 = b, \\ \lambda_1 + \lambda_2 + \lambda_3 = c, \end{array} \right.$$

que forman el sistema de ecuaciones (4.6), en la página 431. Naturalmente, este sistema es lo mismo que la única ecuación vectorial

$$\lambda_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Esta ecuación es muy fácil de interpretar en términos de coordenadas: las columnas que aparecen en ella son las coordenadas de

$$x^2 - x + 1, \quad x^2 + x + 1, \quad x^2 + 1, \quad ax^2 + bx + c,$$

respecto a la base canónica

$$\mathcal{C} = \left(1, x, x^2\right)$$

de $\mathbb{R}_2[x]$. En efecto, expresar los polinomios de $\mathbb{R}_2[x]$ en esta base es equivalente a ordenar sus términos en orden creciente según su grado.

Concluimos que un polinomio es combinación lineal de la familia $\mathcal{A}$ si y sólo si sus coordenadas respecto a la base $\mathcal{C}$ son una combinación lineal de las coordenadas de los vectores en $\mathcal{A}$. ♠

El fenómeno que poníamos en evidencia en el ejemplo anterior es completamente general: para todo lo que tiene que ver con la estructura lineal, da lo mismo trabajar en $\mathbb{V}$, que hacerlo en $\mathbb{K}^n$ a través de sus coordenadas. Al pasar a coordenadas todo se reduce a la consideración de sistemas lineales de ecuaciones, para los que podemos emplear toda la teoría desarrollada en el capítulo 1.

En la sección 4.8 daremos una formulación aún más precisa de la equivalencia entre trabajar en un espacio o hacerlo con sus coordenadas, al introducir la noción de *isomorfismo*. De momento vamos a mostrar una proposición que muestra que todos los cálculos que tiene que ver con combinaciones lineales puede hacerse a través de coordenadas.

Proposición 4.22. *Sea $\mathbb{V}$ un espacio vectorial de dimensión finita n, y $\mathcal{B}$ una base de $\mathbb{V}$. Entonces*

1. *un vector $v \in \mathbb{V}$ es una combinación lineal de $(v_1, \ldots, v_m)$ si y sólo si $(v)_{\mathcal{B}}$ es una combinación lineal de $((v_1)_{\mathcal{B}}, \ldots, (v_m)_{\mathcal{B}})$;*

2. *una familia $(v_1, \ldots, v_m)$ es un generador de $\mathbb{V}$ si y sólo si la familia $((v_1)_{\mathcal{B}}, \ldots, (v_m)_{\mathcal{B}})$ es un generador de $\mathbb{K}^n$;*

3. *una familia $(v_1, \ldots, v_m)$ es linealmente independiente si y sólo si la familia $((v_1)_{\mathcal{B}}, \ldots, (v_m)_{\mathcal{B}})$ es linealmente independiente;*

4. *una familia $(v_1, \ldots, v_n)$ es una base de $\mathbb{V}$ si y sólo si $((v_1)_{\mathcal{B}}, \ldots, (v_n)_{\mathcal{B}})$ es una base de $\mathbb{K}^n$;*

Prueba. Sólo vamos a dar una demostración completa de las partes 3 y 4. Para probar cualquiera de las partes la técnica consiste en formular una propiedad en el espacio $\mathbb{V}$ y luego pasar a $\mathbb{K}^n$ tomando coordenadas, o hacer el camino inverso deshaciendo el pasaje a coordenadas. Con esta información el lector podrá completar la demostración.

Supongamos que $\mathcal{A} = (v_1, \ldots, v_m)$ es linealmente independiente. Para mostrar la independencia lineal de $\mathcal{C} = ((v_1)_{\mathcal{B}}, \ldots, (v_m)_{\mathcal{B}})$ consideremos una combinación lineal tal que

$$\lambda_1 (v_1)_{\mathcal{B}} + \ldots + \lambda_m (v_m)_{\mathcal{B}} = O. \tag{4.15}$$

El miembro de la izquierda en esta igualdad es igual a las coordenadas de la combinación lineal de $(v_1, \ldots, v_m)$ con coeficientes λ_i, $i = 1, \ldots, m$. Por lo tanto

$$(\lambda_1 v_1 + \ldots + \lambda_m v_m)_{\mathcal{B}} = O.$$

Sólo el vector nulo de $\mathbb{V}$ tiene el vector nulo de $\mathbb{K}^n$ como vector de coordenadas, entonces

$$\lambda_1 v_1 + \ldots + \lambda_m v_m = O. \tag{4.16}$$

La independencia lineal de $\mathcal{A}$ implica que todos los coeficientes λ_i, $i = 1, \ldots, m$, son nulos. Por lo tanto, una combinación lineal de $\mathcal{C}$ que sea igual al vector nulo tiene todos sus coeficientes nulos. Esto es, por definición, la independencia lineal de $\mathcal{C}$.

Probemos el recíproco, para el que comenzamos por asumir que $\mathcal{C}$ es linealmente independiente. Formamos entonces una combinación lineal de $\mathcal{B}$ que sea igual al vector nulo, como en (4.16). Al tomar coordenadas respecto a la base $\mathcal{C}$ obtenemos una igualdad como (4.15), donde O indica al vector nulo de $\mathbb{K}^n$. Ahora usamos la independencia lineal de $\mathcal{C}$ para concluir que los coeficientes de la combinación lineal en el miembro de la izquierda de (4.16) son todos nulos, lo que implica la independencia lineal de $\mathcal{B}$. Así concluimos la demostración de la parte 3.

Para demostrar la parte 4 sólo hay que notar que una familia es base de $\mathbb{V}$ si y sólo si es un generador de $\mathbb{V}$ y es linealmente independiente. Pero es un generador de $\mathbb{V}$ si y sólo si las coordenadas de los vectores de la familia son un generador de $\mathbb{K}^n$ (parte 1). Y es linealmente independiente si y sólo si la familia de coordenadas es linealmente independiente (parte 3). $\square$

Ejercicio 4.74. Probar las partes 1 y 2 de la proposición 4.22.

Una vez que hemos hecho explícita la posibilidad de calcular usando coordenadas, reformularemos algunos ejemplos usando esta técnica.

Ejemplo 4.5.11. En el ejemplo 4.3.7 buscábamos el subespacio $[\mathcal{A}]$ generado por

$$\mathcal{A} = \left(\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \right),$$

dentro del espacio $M^{2\times 2}(\mathbb{R})$ de las matrices reales 2×2.

Para saber si una matriz

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

pertenece a $[\mathcal{A}]$ formularemos la condición de pertenencia en términos de sus coordenadas respecto a la base canónica. La matriz B está en ese subespacio si y sólo si existen tres números reales α_i, $i = 1, 2, 3$, tales que

$$\alpha_1 \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} b_{11} \\ b_{12} \\ b_{21} \\ b_{22} \end{pmatrix}. \tag{4.17}$$

Por supuesto, esta condición es equivalente a la compatibilidad del sistema

$$\left(\begin{array}{rrr|c} 1 & 0 & 1 & b_{11} \\ -1 & 1 & 0 & b_{12} \\ 2 & 1 & 3 & b_{21} \\ 1 & 0 & 1 & b_{22} \end{array} \right)$$

Las columnas de la matriz del sistema son las coordenadas de los vectores que aparecen en el cálculo. Esta observación permite formular directamente el sistema, sin pasar por la igualdad vectorial (4.17). En el próximo ejemplo procederemos de esta manera, y escribiremos directamente el sistema de ecuaciones lineales que hay que estudiar para resolver el problema. ♣

Ejemplo 4.5.12. En el ejemplo 4.3.14, página 438, discutíamos un procedimiento para ver si alguno de los polinomios en la familia

$$\mathcal{A} = \left(1 + x^2, 2 + 2x^2, 1 + x + x^3, 3 + x + 2x^2 + x^3\right)$$

podía escribirse como combinación lineal de los restantes. El problema es equivalente a determinar si alguno de sus vectores de coordenadas respecto a la base canónica

$$\mathcal{C} = \left(1, x, x^2, x^3\right)$$

puede escribirse como combinación lineal de los demás, y se resuelve estudiando el sistema lineal de ecuaciones

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right).$$

Para esto último, podemos recurrir al método de eliminación gaussiana, tal como se hace en el ejemplo 4.3.14. ♣

Ejemplo 4.5.13. La independencia lineal de la familia

$$\mathcal{A} = \left(x^2, 2x + 2, 3x^2 + x + 1\right) \subset \mathbb{R}_2[x],$$

es equivalente a que el sistema homogéneo

$$\left(\begin{array}{ccc|c} 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 3 & 0 \end{array}\right)$$

sea determinado, y admita la solución trivial como su única solución. Las columnas de la matriz del sistema son las coordenadas de los vectores de $\mathcal{A}$ respecto a la base canónica de $\mathbb{R}_2[x]$. ♣

Ejercicio 4.75.

1. Probar que $\mathcal{B} = \ ((1,1,0),\ (0,1,1),\ (1,0,1))$ es una base de $\mathbb{R}^3$. Encuentre las coordenadas del vector $v = (1,2,3)$ en la base $\mathcal{B}$.

2. Demuestre que $\mathcal{B} = (p_0, p_1, p_2)$ es base de $\mathbb{R}_2[x]$, donde

$$p_i(x) = (x-1)^i, \quad i = 0, 1, 2.$$

Hallar las coordenadas del vector $p(x) = 6x^2 - 3x + 2$ en la base $\mathcal{B}$.

3. Pruebe que

$$\mathcal{B} = \left(\begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 1 & 5 \end{pmatrix}, \begin{pmatrix} 4 & -1 \\ 2 & 2 \end{pmatrix} \right)$$

es una base de $M_{2x2}(\mathbb{R})$. Hallar las coordenadas de

$$A = \begin{pmatrix} 1 & 3 \\ -4 & 6 \end{pmatrix}$$

en la base $\mathcal{B}$.

Ejercicio 4.76. Sea V un espacio vectorial y (v_1, v_2, v_3, v_4) un subconjunto de V. Si sabemos que (v_1, v_2, v_3) es linealmente independiente y que (v_1, v_2, v_4) es un generador de $\mathbb{V}$, ¿cuál es la dimensión de $\mathbb{V}$?.

Ejercicio 4.77. Consideramos el conjunto

$$\mathbb{S} = \{p \in \mathbb{R}_n[x];\ p(x) = p(1-x)\ \forall\ x \in \mathbb{R}\}.$$

1. Probar que $\mathbb{S}$ es un subespacio vectorial de $\mathbb{R}_n[x]$.
2. Consideremos los polinomios

$$p_i(x) = x^i + (1-x)^i, \quad i = 1, 2, \cdots, n.$$

Muestre que los polinomios p_i pertenecen al subespacio $\mathbb{S}$.

3. Probar que $\{p_1, \cdots, p_n\}$ es un generador del subespacio $\mathbb{S}$.
4. Para $n = 4$, calcular $\dim(\mathbb{S})$.

Ejercicio 4.78.

Consideramos los polinomios

$$h(x) = 1 + x, \quad p_\alpha(x) = 1 + \alpha x + x^2, \quad q(x) = 1 + x^2.$$

1. ¿Para qué valores de $\alpha \in \mathbb{R}$ el conjunto $\mathcal{B}_\alpha = (h(t), p_\alpha(t), q(t))$ es una base de $\mathbb{R}_2[x]$?.
2. Para los valores de α tal que $\mathcal{B}_\alpha$ es base, hallar las coordenadas del vector $q(x) = x^2 + x + 2$ en esa base.

4.5.3. Para tener presente

- La elección adecuada de las coordenadas en un espacio puede simplificar mucho cualquier problema que involucre cálculos con los vectores de ese espacio.
- La correspondencia que a cada vector de un $\mathbb{K}$-espacios vectorial de dimensión n le asocia sus coordenadas respecto a una base es una correspondencia biunívoca entre los vectores del espacio y los vectores de $\mathbb{K}^n$. Esta correspondencia y su inversa son lineales, en el sentido de que la imagen de una combinación lineal de vectores es la combinación lineal de las imágenes de los vectores.
- Una familia es linealmente independiente si y sólo si la familia de sus coordenadas respecto a cualquier base es linealmente independiente.
- Un vector es una combinación lineal de una familia si y sólo si sus coordenadas son una combinación lineal de las coordenadas de los vectores de la familia.

4.6. Transformaciones lineales

En esta sección comenzamos a desarrollar una teoría general de las transformaciones lineales entre espacios vectoriales. Se trata de funciones que tienen como dominio y codominio espacios vectoriales sobre un mismo cuerpo, y que "respetan", en un sentido que explicaremos claramente, la estructura lineal del espacio.

En cualquier estructura matemática es natural estudiar las funciones que guardan alguna relación con la estructura. Por ejemplo, en el caso de las funciones reales de variable real las funciones monótonas, que preservan la estructura de orden de la recta real, tienen interesantes propiedades. Otra clase de funciones destacadas son las funciones continuas, que tienen propiedades relacionadas con la topología de $\mathbb{R}$.

Por otra parte, las aplicaciones de la teoría de espacios vectoriales originan funciones que están muy relacionadas con la descripción de diversos problema. Por ejemplo, si trabajamos con funciones derivables tenemos la correspondencia $f \mapsto f'$ que a cada función le asocia su derivada; al calcular los esfuerzos que se producen al cargar un reticulado encontramos que estos dependen de las fuerzas aplicadas; la solución a un problema de interpolación queda determinada en función de los datos del problema; al considerar el problema geométrico de proyectar un vector sobre un plano estamos definiendo una transformación de todos los vectores del espacio, etcétera. Estudiaremos en esta sección las funciones que guardan relación con la estructura lineal del espacio.

4.6.1. Definición, ejemplos y primeras propiedades

Los espacios vectoriales son estructuras en las que se puede sumar y multiplicar por escalares. Las transformaciones lineales son las que preservan estas operaciones básicas, en el sentido de la siguiente definición.

Definición 4.15. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ y*

$$T : \mathbb{V} \to \mathbb{W}$$

*una función con dominio $\mathbb{V}$ y codominio $\mathbb{W}$. Decimos que T es **lineal** si satisface:*

1. *para todo par de vectores u y v en $\mathbb{V}$ se cumple que*

$$T(u+v) = T(u) + T(v);$$

2. para todo $\lambda \in \mathbb{K}$ *y* $v \in \mathbb{V}$ *se cumple*

$$T(\lambda v) = \lambda T(v).$$

Naturalmente, las propiedades de la definición pueden extenderse a combinaciones lineales cualesquiera, en el sentido de la siguiente proposición.

Proposición 4.23. *Sean* $\mathbb{V}$ *y* $\mathbb{W}$ *dos* $\mathbb{K}$*-espacios vectoriales, y* $T : \mathbb{V} \to \mathbb{W}$*. La función* T *es lineal si y sólo si*

$$T(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 T v_1 + \ldots + \lambda_n T v_n$$

se satisface para cualquier elección de n *vectores* $v_1, \ldots, v_n$ *en* $\mathbb{V}$ *y* n *escalares* $\lambda_1, \ldots, \lambda_n$ *en* $\mathbb{K}$.

Prueba. Mostraremos primero el directo. Para probarlo usaremos la definición de transformación lineal.

Realicemos primero un razonamiento inductivo. Sabemos que si T es lineal se cumple que $T(u+v) = Tu + Tv$. Si hacemos $u = v_1$ y $v = v_2 + v_3$ se cumple

$$T(v_1 + v_2 + v_3) = Tv_1 + T(v_2 + v_3) = Tv_1 + Tv_2 + Tv_3.$$

Pensemos ahora en una suma de n vectores, tenemos

$$T(v_1 + v_2 + \ldots + v_n) = Tv_1 + T(v_2 + \ldots + v_n) = Tv_1 + Tv_2 + \ldots + Tv_n.$$

Ahora debemos mostrar que esto se cumple para una combinación lineal cualesquiera. Para probarlo usamos primero la linealidad respecto a la suma para luego usar la linealidad respecto al producto por un escalar. Hacemos este razonamiento para el primer vector v_1 y lo extendemos a n vectores cualesquiera.

$$T(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = T(\lambda_1 v_1) + T(\lambda_2 v_2 + \ldots + \lambda_n v_n) =$$

$$\lambda_1 T v_1 + T(\lambda_2 v_2 + \ldots + \lambda_n v_n) = \ldots = \lambda_1 T v_1 + \lambda_2 T v_2 + \ldots + \lambda_n T v_n$$

El recíproco se reduce a dos casos particulares: basta realizar combinaciones de dos vectores para el caso de la suma, y de un vector para el producto. □

Podemos decir entonces que las transformaciones lineales son justamente las que transforman una combinación lineal de vectores en la combinación lineal de sus imágenes.

Observación 4.6.1. De la demostración de la proposición 4.23 se desprende que para asegurarse de que una transformación T sea lineal es suficiente verificar que la igualdad

$$T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2)$$

se satisface para una combinación lineal arbitraria $\lambda_1 v_1 + \lambda_2 v_2$, de tan sólo dos vectores cualesquiera v_1 y v_2. Esta propiedad implica que se cumplen los requisitos establecidos en la definición de transformación lineal. ♠

Observación 4.6.2. Un poco de jerga y notación
Es usual referirse a las funciones lineales como las que acabamos de definir como *transformaciones lineales*, por eso hemos adoptado este nombre para esta sección. También se usan las expresiones *aplicación lineal*, y *operador*, u *operador lineal*. Estas dos últimas son especialmente comunes en el contexto de los espacios de dimensión infinita.

La mayor parte del tiempo utilizaremos la expresión transformación lineal. Independientemente de cuál sea el nombre que se use, el lector debe tener presente que se trata de funciones, para las cuales tiene pleno sentido considerar su dominio, codominio, imagen, inyectividad, sobreyectividad, biyectividad, y con las que se puede operar componiendo. Haremos todas estas cosas a lo largo de este capítulo. Además, como las transformaciones lineales toman valores en espacios vectoriales, en los que están definidas las operaciones de suma y de producto por un escalar, podremos definir la suma de funciones lineales y su producto por un escalar, tal como se hace con las funciones que toman valores sobre un cuerpo $\mathbb{K}$. Ver el ejemplo 4.1.2, en la página 398.

Respecto a la notación, digamos que para expresar la imagen de un vector v por una transformación lineal T es corriente escribir Tv en vez de $T(v)$, omitiendo los paréntesis que suelen indicar la evaluación de una función. El lector encontrará ambas notaciones en este texto. Preferiremos Tv a $T(v)$ siempre que la omisión de los paréntesis no dé lugar a confusión. ♣

Ejemplo 4.6.3. Multiplicar por una matriz
Una matriz A de tamaño $m \times n$ sobre un cuerpo $\mathbb{K}$ define sobre $\mathbb{K}^n$ una transformación

$$X \mapsto AX$$

que toma valores en $\mathbb{K}^m$. Esta transformación es lineal, tal como vimos en la proposición 1.6 en la página 90.

Hemos aprendido en la sección 4.5 que los espacios $\mathbb{K}^n$ son un modelo para los espacios de dimensión finita, en el sentido de que sus vectores pueden representarse fielmente por sus coordenadas en una base. Cuando analicemos la representación de las transformaciones lineales en coordenadas, veremos que multiplicar por matrices es un modelo para todas las transformaciones lineales entre espacios de dimensión finita. En este sentido, podemos decir que el ejemplo que acabamos de presentar es universal. ♣

Ejemplo 4.6.4. Coordenadas
Dada una base $\mathcal{A}$ en un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ de dimensión n la correspondencia

$$\mathrm{coord}_{\mathcal{A}} : \mathbb{V} \rightarrow \mathbb{K}^n$$

es lineal, tal como se muestra en la proposición 4.20.

Ejemplo 4.6.5. Hacer combinaciones lineales
Fijemos n vectores

$$v_1, v_2, \ldots, v_n,$$

en un $\mathbb{K}$-espacio vectorial. Dados n escalares cualesquiera

$$\lambda_1, \lambda_2, \ldots, \lambda_n,$$

en $\mathbb{K}$ podemos fabricar la combinación lineal

$$\lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Con este procedimiento estamos definiendos una transformación

$$T : \mathbb{K}^n \rightarrow \mathbb{V},$$

que a cada vector

$$\Lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$$

le asocia

$$T\Lambda = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Cuando $\mathbb{V}$ es el espacio $\mathbb{K}^m$ la definición de T se reduce a la multiplicación de Λ por la matriz A cuyas columnas son los vectores $v_1, \ldots, v_n$. En efecto, el producto $A\Lambda$ es el vector de $\mathbb{K}^m$ que se obtiene haciendo una combinación lineal de las columnas de A con los coeficientes de Λ. Ya vimos en este caso que T es lineal.

También hemos analizado el caso particular en el que $(v_1, v_2, \ldots, v_n)$ forma una base de $\mathbb{V}$, en cuyo caso la aplicación T es la inversa a tomar coordenadas respecto a esta base. Ya vimos la linealidad de esta aplicación en la proposición 4.20, página 476.

Dejamos como ejercicio para el lector mostrar la linealidad de T en un caso cualquiera.

Ejercicio 4.79. Completar la demostración de que T es lineal.

En particular, este ejemplo muestra que si fijamos un vector $v \in \mathbb{V}$ la aplicación

$$\lambda \mapsto \lambda v,$$

definida sobre el cuerpo $\mathbb{K}$ es lineal. Recordemos que $\mathbb{K}$ es un $\mathbb{K}$ espacio vectorial, que corresponde al caso particularísimo $n = 1$ de los espacios $\mathbb{K}^n$. ♣

El párrafo final del ejemplo anterior muestra que el producto λv define una aplicación lineal cuando se le considera como una función de λ. El siguiente muestra que también es lineal cuando se considera como función de v.

Ejemplo 4.6.6. Sea $\mathbb{V}$ un $\mathbb{K}$ espacio vectorial, y fijemos un escalar $\lambda \in \mathbb{K}$. La función $T : \mathbb{V} \to \mathbb{V}$ tal que

$$T(v) = \lambda v$$

es una transformación lineal. Efectivamente, tenemos

$$\begin{aligned} T(v_1 + v_2) = \lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 = T(v_1) + T(v_2), \\ T(\alpha v) = \lambda(\alpha v) = \alpha(\lambda v) = \alpha T(v). \end{aligned}$$

para vectores cualesquiera v_1 y v_2 en $\mathbb{V}$, y α un escalar en $\mathbb{K}$. ♣

En los próximo ejemplos consideraremos espacios vectoriales formados por funciones que toman valores en un cuerpo, en particular, en $\mathbb{R}$. En este contexto, la suma de dos vectores, es decir, de dos funciones f y g, es la función $f + g$ caracterizada por

$$(f + g)(x) = f(x) + g(x), \quad x \in X.$$

El producto λf de f por $\lambda \in \mathbb{K}$ es la función definida por

$$(\lambda f)(x) = \lambda f(x), \quad x \in X.$$

Ejemplo 4.6.7. La derivada
Consideramos el espacio vectorial $C^1(\mathbb{R})$, formado por las funciones reales definidas sobre $\mathbb{R}$ que tienen derivada continua, y el espacio $C(\mathbb{R})$ de las funciones reales continuas. Entonces el operador

$$D : C^1(\mathbb{R}) \to C(\mathbb{R})$$

que a cada función f le asocia su derivada $Df = f'$ es lineal. No demostraremos este resultado aquí, ya que se trata de las propiedades de linealidad de la derivada que aparecen en cualquier texto de cálculo. Ver, por ejemplo, Introducción al cálculo y al análisis matemático, Volumen 1, R. Courant, página 187. ♣

El próximo ejemplo emplea las propiedades de linealidad de la integral.

Ejemplo 4.6.8. LA INTEGRAL INDEFINIDA
La integral

$$I(f)(x) = \int_0^x f(y)\,dy$$

define una aplicación lineal

$$I : C(\mathbb{R}) \to C^1(\mathbb{R}),$$

que a cada función continua f le asocia $I(f)$. ♣

Ejercicio 4.80. EVALUACIONES DE FUNCIONES

- Consideremos un punto x en el dominio X de las funciones que forman el espacio $\mathbb{V}$. Mostrar que la aplicación

  $$f \mapsto f(x)$$

  que a cada $f \in \mathbb{V}$ le asocia el valor de su evaluación en x define una transformación lineal de $\mathbb{V}$ en $\mathbb{K}$.
- Si $x_1, x_2, \dots x_n$, son n puntos en x, mostrar que la correspondencia de $\mathbb{V}$ en $\mathbb{K}^n$ definida por

  $$f \mapsto (f(x_1), f(x_2), \dots, f(x_n))$$

  es una transformación lineal.

En los siguientes ejercicios mostramos que dos tipos de transformaciones geométricas especialmente importantes son lineales: las isometrías y las proyecciones ortogonales.

Ejercicio 4.81. ISOMETRÍAS: TRANSFORMACIONES QUE CONSERVAN LONGITUDES (II)

1. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ una función tal que

   $$|f(X)| = |X|, \; X \in \mathbb{R}^3.$$

   Mostrar que f es lineal. Sugerencia: Ver el ejercicio 3.56.
2. ¿Qué se puede decir de una función $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $|f(X) - f(Y)| = |X - Y|$, con $X, Y \in \mathbb{R}^3$?

Ejercicio 4.82. PROYECCIONES ORTOGONALES
Usar la fórmula (3.50), de la página 353, para mostrar que la proyección ortogonal sobre un subespacio de dimensión 2 de $\mathbb{R}^3$ es una transformación lineal. Mostrar también que la proyección ortogonal sobre un subespacio de dimensión 1 es una transformación lineal.

Ejemplo 4.6.9. CAMBIO DE INTERPOLADOR
Distintos procedimientos de interpolación tienen diferentes propiedades. Interesa en algunos casos cambiar de un tipo de interpolador a otro. Una clase útil para problemas de interpolación es la de las funciones lineales a trozos, que es relativamente fácil de manejar, tal como vimos en el ejemplo 4.1.1 en la página 396. Otra clase está formada por polinomios, que son una familia de funciones sencillas con la propiedad adicional de tener derivadas de todos los órdenes.

Consideraremos en este ejemplo el espacio $\mathbb{L}$ formado por las funciones lineales a trozos definidas en el intervalo $[-1,1]$, con nodos -1, 0 y 1. A cada función $l \in \mathbb{L}$ le haremos corresponder el polinomio $p(x) \in \mathbb{R}_3[x]$ cuyo valor coincide con el de l en los puntos -1, $-1/3$, $1/3$ y 1. Es decir, definimos una transformación

$$T : \mathbb{L} \to \mathbb{R}_3[x]$$

asociando a cada $l \in \mathbb{L}$ el único polinomio $p(x)$ de grado menor o igual que 3 que satisface

$$p(-1) = l(-1), \quad p(-1/3) = l(-1/3), \quad p(1/3) = l(1/3), \quad p(1) = l(1).$$

Esta correspondencia T es lineal. La linealidad puede demostrarse a partir de la definición de T que acabamos de dar, algo que dejamos como ejercicio para el lector interesado, pero nosotros pospondremos la consideración de esta cuestión hasta desarrollar algo más la teoría de las transformaciones lineales. Ver el final del ejemplo 4.7.4, a partir de la página 513.

Ejercicio 4.83. Demostrar la linealidad de T. ♣

Agregamos ahora un ejemplo de naturaleza más geométrica.

Ejemplo 4.6.10. Fijemos un vector $Y \in \mathbb{R}^3$, y consideremos la correspondencia

$$X \mapsto X \cdot Y,$$

que a cada $X \in \mathbb{R}^3$ le asocia su producto escalar con Y. Esta correspondencia es lineal, tal como se desprende de las propiedades del producto escalar que demostramos en la proposición 3.1 en la página 315. ♣

Ejemplo 4.6.11. RESTRICCIÓN DE TRANSFORMACIONES LINEALES
Dentro de un espacio vectorial $\mathbb{V}$ podemos encontrar otros espacios vectoriales: sus subespacios. Para cualquier transformación lineal $T : \mathbb{V} \to \mathbb{W}$ podemos limitarnos a considerar su acción dentro de un subespacio $\mathbb{S}$ de $\mathbb{V}$, y definir una transformación de $\mathbb{S}$ a $\mathbb{W}$ transformando a cada vector de $\mathbb{S}$ por medio de

T. La “fórmula”para calcular la transformación es la misma que para T, pero la diferencia con T es que hemos modificado el dominio de la aplicación, que ahora es $\mathbb{S}$.

La aplicación que se construye por este método es lo que se llama la *restricción* de T al subespacio $\mathbb{S}$, y es corriente indicarla por $T|_{\mathbb{S}}$. Tenemos entonces una aplicación

$$T|_{\mathbb{S}} : \mathbb{S} \to \mathbb{W},$$

que es también una transformación lineal, porque las condiciones de linealidad eran satisfechas por T para todos los vectores de $\mathbb{V}$. Por lo tanto se satisfacen también para los vectores de $\mathbb{S}$. Esto es todo lo que necesitamos para verificar la linealidad de $T|_{\mathbb{S}}$, porque la definición de $T|_{\mathbb{S}}$ implica que la igualdad

$$T|_{\mathbb{S}}(s) = T(s) \in \mathbb{W}$$

se satisface para todo $s \in \mathbb{S}$ ♣

Ejemplo 4.6.12. Inversas de transformaciones lineales
Las transformaciones lineales son funciones entre espacios vectoriales. Igual que ocurre con una función cualquiera entre conjuntos cualesquiera, cuando una transformación lineal

$$T : \mathbb{V} \to \mathbb{W}$$

es *biyectiva* tiene una inversa

$$T^{-1} : \mathbb{W} \to \mathbb{V}.$$

Por ejemplo, si A es una matriz invertible en $M^{n\times n}(\mathbb{K})$ entonces la inversa de la aplicación $X \mapsto Y = AX$ en $\mathbb{K}^n$ es $Y \mapsto X = A^{-1}Y$, que consiste en multiplicar a izquierda por la matriz inversa A^{-1}.

En la proposición 4.20, página 476, establecimos la linealidad de la aplicación $\mathrm{coord}_{\mathcal{B}}$ y su inversa. Los argumentos que usamos allí son bastante generales, y una ligera adaptación permite demostrar la siguiente proposición sobre transformaciones lineales invertibles.

Proposición 4.24. *Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal biyectiva. Entonces su inversa $T^{-1} : \mathbb{W} \to \mathbb{V}$ también es lineal.*

Ejercicio 4.84. Demostrar la proposición 4.24. Sugerencia: revisar la prueba de la proposición 4.20. ♣

Cerramos esta sección introductoria con una sencilla propiedad que deben satisfacer todas las transformaciones lineales: deben llevar el vector nulo de su dominio al vector nulo de su codominio. En su enunciado indicamos con $O_{\mathbb{V}}$ y $O_{\mathbb{W}}$ a los vectores nulos de $\mathbb{V}$ y $\mathbb{W}$ respectivamente.

Proposición 4.25. *Sea $T : \mathbb{V} \rightarrow \mathbb{W}$ una transformación lineal. Entonces,*

$$T\left(O_{\mathbb{V}}\right) = O_{\mathbb{W}}.$$

PRUEBA. Hay varios trucos posibles para hacer la demostración. Se puede escribir $O + O = O$, o escoger un vector $v \in \mathbb{V}$ para escribir

$$O_{\mathbb{V}} = 0v.$$

Entonces

$$T\left(O_{\mathbb{V}}\right) = T(0v) = 0T(v) = O_{\mathbb{W}}.$$

La primera igualdad está justificada por la linealidad de T, que permite "sacar fuera de T" la constante $0 \in \mathbb{K}$. La segunda igualdad es cierta porque el producto de 0 por cualquier vector de $\mathbb{W}$ es igual a $O_{\mathbb{W}}$. □

Ejemplo 4.6.13. Sea $\mathbb{V}$ un espacio vectorial y $v_0 \neq 0$ un vector fijo de $\mathbb{V}$. La función $T : \mathbb{V} \rightarrow \mathbb{V}$ tal que

$$Tv = v + v_0$$

no es una transformación lineal, pues

$$T\left(v_1 + v_2\right) = v_1 + v_2 + v_0 \neq v_1 + v_0 + v_2 + v_0 = Tv_1 + Tv_2.$$

♣

4.6.2. Matriz asociada

Vamos a considerar ahora un hecho de la mayor importancia para el cálculo con transformaciones lineales: así como los vectores de un espacio vectorial de dimesión finita pueden representarse por medio de sus coordenadas respecto a una base, la acción de las transformaciones lineales puede expresarse en coordenadas. Cuando los vectores se escriben en coordenadas los cálculos se reducen a la consideración de sistemas lineales. Cuando las transformaciones lineales se expresan en coordenadas su acción queda representada por matrices. Desarrollar esta última frase es el principal objetivo de esta sección.

Trabajaremos con una transformación lineal

$$T : \mathbb{V} \rightarrow \mathbb{W}$$

definida entre dos $\mathbb{K}$-espacios vectoriales de dimensión finita $\mathbb{V}$ y $\mathbb{W}$, de dimensiones n y m respectivamente. Fijemos entonces una base

$$\mathcal{A} = (v_1, v_2, \ldots, v_n)$$

del espacio $\mathbb{V}$, y una base

$$\mathcal{B} = (w_1, w_2, \ldots, w_m)$$

del espacio $\mathbb{W}$.

A continuación buscaremos comprender cómo se ve la acción de T en estas coordenadas, representando en las bases $\mathcal{A}$ y $\mathcal{B}$ todos los vectores que aparezcan en el cálculo. Consideremos entonces un vector $v \in \mathbb{V}$, con coordenadas

$$(v)_{\mathcal{A}} = (\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Por lo tanto v puede escribirse como una combinación lineal de los vectores de la base $\mathcal{A}$ con los λ_i, $i = 1, \ldots, n$ como coeficientes. Es decir

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

Como T es una transformación lineal resulta

$$Tv = \lambda_1 Tv_1 + \lambda_2 Tv_2 + \ldots + \lambda_n Tv_n.$$

Buscamos ahora la expresión de Tv en coordenadas referidas a la base $\mathcal{B}$ de $\mathbb{W}$. Como tomar las coordenadas de un vector también es una operación lineal encontramos que

$$(Tv)_{\mathcal{B}} = \lambda_1 (Tv_1)_{\mathcal{B}} + \lambda_2 (Tv_2)_{\mathcal{B}} + \ldots + \lambda_n (Tv_n)_{\mathcal{B}}. \tag{4.18}$$

Es decir, las coordenadas del vector Tv, transformado de v por T, en la base $\mathcal{B}$ se obtienen como una combinación lineal de

$$(Tv_1)_{\mathcal{B}},\ (Tv_2)_{\mathcal{B}}, \ldots, (Tv_n)_{\mathcal{B}}, \tag{4.19}$$

en la que los coeficientes son las coordenadas

$$(\lambda_1, \ldots, \lambda_n) \tag{4.20}$$

que corresponden a v en la base $\mathcal{A}$.

Observemos que los vectores de coordenadas que aparecen en (4.19) son vectores de $\mathbb{K}^m$. Por lo tanto la combinación lineal de la fórmula (4.18) puede representarse[15] como **el producto de la matriz**

$$_{\mathcal{B}}T_{\mathcal{A}} = \left((Tv_1)_{\mathcal{B}}, (Tv_2)_{\mathcal{B}}, \ldots, (Tv_n)_{\mathcal{B}}\right), \tag{4.21}$$

[15]Esto no es más que una consecuencia de nuestra definición del producto de una matriz A por un vector X. Fabricamos la definición para que el producto AX fuera una manera concisa de expresar la combinación lineal de las columnas de A con los coeficientes almacenados en X. Ver la página 87.

que tiene como columnas a los vectores (4.19), por el vector columna (4.20) que contiene a las coordenadas de v respecto a la base $\mathcal{A}$.

A la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$ que aparece en la fórmula (4.21) la llamaremos *matriz asociada con la transformación T en las bases $\mathcal{A}$ y $\mathcal{B}$*. La notación escogida expresa cuál es la transformación con la que estamos trabajando, y en que bases del dominio y codominio se representa su acción.

Al tener presente que

$$(\lambda_1, \lambda_2, \dots, \lambda_n) = (v)_{\mathcal{A}}$$

toda la discusión precedente se resume en la fórmula

$$(Tv)_{\mathcal{B}} =_{\mathcal{B}} (T)_{\mathcal{A}} (v)_{\mathcal{A}}, \tag{4.22}$$

o en la frase **las coordenadas de Tv en la base $\mathcal{B}$ son el producto de la matriz asociada con T en las bases $\mathcal{A}$ y $\mathcal{B}$ por el vector de coordenadas de v en la base $\mathcal{A}$.**

Es decir, vista en coordenadas respecto a las bases $\mathcal{A}$ y $\mathcal{B}$, la acción de T es equivalente a la multiplicación por una matriz, la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$.

Antes de seguir avanzando vamos a verificar que la matriz asociada a una transformación lineal respecto a un par de bases dadas es única.

Ejercicio 4.85. En este ejercicio consideraremos dos $\mathbb{K}$-espacios vectoriales $\mathbb{V}$ y $\mathbb{W}$ con bases $\mathcal{A}$ y $\mathcal{B}$, y dimensiones n y m, respectivamente, y una transformación lineal $T : \mathbb{V} \to \mathbb{W}$.

1. Mostrar que si dos matrices A y B sobre $\mathbb{K}$ tienen la propiedad de que

   $$(Tv)_{\mathcal{B}} = A(v)_{\mathcal{A}} = B(v)_{\mathcal{A}}, \quad v \in \mathbb{V},$$

   entonces ambas matrices tienen dimensiones $m \times n$ y

   $$AX = BX, \quad X \in \mathbb{K}^n. \tag{4.23}$$

2. Mostrar que si dos matrices $m \times n$ satisfacen (4.23) entonces son iguales.
3. Concluir que sólo puede haber una matriz ${}_{\mathcal{B}}(T)_{\mathcal{A}}$ tal que se satisfaga (4.22).

Nuestro primer ejemplo muestra que todas las transformaciones lineales definidas entre espacios $\mathbb{K}^n$ se reducen a la multiplicación por matrices.

Ejemplo 4.6.14. TRANSFORMACIONES LINEALES ENTRE ESPACIOS $\mathbb{K}^n$
Ya vimos que una matriz $A \in M^{m\times n}(\mathbb{K})$ define una transformación lineal

$$\begin{array}{rcl} \mathbb{K}^n & \to & \mathbb{K}^m, \\ X & \mapsto & AX, \end{array}$$

que consiste en multiplicar por A. Como en los espacios $\mathbb{K}^n$ los vectores coinciden con sus coordenadas respecto las bases canónicas $\mathcal{C}_n$ la matriz A es justamente la matriz asociada con esta transformación en las bases canónicas $\mathcal{C}_m$ y $\mathcal{C}_n$ de $\mathbb{K}^m$ y $\mathbb{K}^n$. En efecto,

$$(AX)_{\mathcal{C}_m} = AX = A(X)_{\mathcal{C}_n}.$$

Otra manera de ver esto es analizar A columna a columna: la primera columna de A es justamente el producto de A por el primer vector de la base canónica de $\mathcal{C}_n$, y coincide con la expresión de este producto en la base canónica de $\mathcal{C}_m$ Por lo tanto, como es obvio, la matriz A actúa transformando coordenadas, que es justamente la propiedad que caracteriza a la matriz asociada.

En realidad, cualquier transformación lineal

$$T : \mathbb{K}^n \to \mathbb{K}^m$$

consiste en multiplicar los vectores de $\mathbb{K}^n$ por una matriz de tamaño $m \times n$. Si consideramos la matriz asociada con T en las bases canónicas tenemos

$$(TX)_{\mathcal{C}_m} = {}_{\mathcal{C}_m}(T)_{\mathcal{C}_n}(X)_{\mathcal{C}_n}.$$

Al tener en cuenta que TX y X coinciden con sus coordenadas en estas bases encontramos

$$TX = {}_{\mathcal{C}_m}(T)_{\mathcal{C}_n}X,$$

lo que muestra que la acción de T se reduce a multiplicar por una matriz. ♣

Nuestro próximo ejemplo sale del marco de los espacios $\mathbb{K}^n$.

Ejemplo 4.6.15. La derivada en un espacio de polinomios
Consideremos la transformación lineal

$$\begin{aligned} D : \mathbb{R}_2[x] &\to \mathbb{R}_2[x], \\ p(x) &\mapsto p'(x), \end{aligned}$$

que a cada polinomio le asocia su derivada.

Podemos escribir explícitamente la acción de D:

$$ax^2 + bx + c \mapsto 2ax + b.$$

Si escribimos esto mismo en coordenadas respecto a la base canónica

$$\mathcal{C} = (1, x, x^2)$$

tenemos

$$(c, b, a) \mapsto (b, 2a, 0),$$

que admite una representación matricial como

$$\begin{pmatrix} b \\ 2a \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c \\ b \\ a \end{pmatrix}.$$

La matriz en esta última fórmula es justamente ${}_{\mathcal{C}}D_{\mathcal{C}}$.

Tal como sugiere la discusión en la que apareció la matriz asociada, no es necesario calcular explícitamente el efecto de D sobre todos los vectores del espacio. Es suficiente calcular para los vectores de la base del dominio (en este caso la base canónica $\mathcal{C}$), representar las imágenes en la base del codominio (en esta caso el dominio y el codominio son el mismo espacio, y hemos escogido usar la misma base $\mathcal{C}$ como base del codominio, aunque tenemos libertad de tomar cualquier otra base). Las columnas de coordenadas que resulten del cálculo serán las columnas de ${}_{\mathcal{C}}D_{\mathcal{C}}$. A continuación seguimos este procedimiento.

Hallamos entonces las imágenes de los vectores de la base $\mathcal{C}$, y los representamos en la base $\mathcal{C}$, a la que también hemos escogido como base del codominio en este caso particular:

$$\begin{array}{l} D(1) = 0 = 0 \times 1 + 0 \times x + 0 \times x^2, \\ D(x) = 1 = 1 \times 1 + 0 \times x + 0 \times x^2, \\ D(x^2) = 2x = 0 \times 1 + 2 \times x + 0 \times x^2. \end{array}$$

Los vectores de coordenadas de los transformados de los vectores de la base se leen en estas expresiones, y son, respectivamente, las columnas

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}.$$

Sólo resta formar la matriz que tiene estos vectores como columnas, en el mismo orden en el que aparecen en la base los vectores correspondientes a cada columna. Reencontramos

$${}_{\mathcal{C}}D_{\mathcal{C}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Este segundo procedimiento, de calcular sólo los transformados de los vectores de la base, es, en general, más eficiente para el cálculo de la matriz asociada.

Ejemplo 4.6.16. Consideremos la transformación lineal $T : \mathbb{K}_2[t] \to \mathbb{R}^2$ tal que al polinomio $p(t) = at^2 + bt + c$ le hace corresponder

$$Tp = (p(-1), p(1)).$$

Sean $p_2(t) = t^2$, $p_1(t) = t$, $p_0(t) = 1$; las bases $\mathcal{A} = (p_2, p_1, p_0)$ de $\mathbb{K}_2[t]$ y $\mathcal{B} = ((1,1),(1,-1))$ de $\mathbb{R}^2$. Para hallar la matriz asociada a T en dichas bases:

En primer lugar hallamos las imágenes de los vectores de la base $\mathcal{A}$

$$Tp_2 = (1,1), \quad Tp_1 = (-1,1), \quad Tp_0 = (1,1).$$

Luego calculamos las coordenadas de estos en la base $\mathcal{B}$

$$Tp_2 = (1,1) = 1(1,1) + 0(1,-1) \Rightarrow (Tp_2)_{\mathcal{B}} = (1,0),$$

$$Tp_1 = (-1,1) = 0(1,1) - 1(1,-1) \Rightarrow (Tp_1)_{\mathcal{B}} = (0,-1),$$

$$Tp_0 = (1,1) = 1(1,1) + 0(1,-1) \Rightarrow (Tp_0)_{\mathcal{B}} = (1,0).$$

Luego

$${}_{\mathcal{B}}T_{\mathcal{A}} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Ya vimos que cualquier transformación lineal T entre los espacios $\mathbb{K}^n$ es la acción de multiplicar por una matriz, que coincide con la matriz asociada con T en las bases canónicas. Pero se puede representar a T con otras matrices, empleando otras bases. A continuación mostramos dos ejemplos de este tipo. En el ejemplo 4.6.18 se consigue una enorme simplificación en la expresión de T al expresar su acción en una base diferente a la canónica.

Ejemplo 4.6.17. Consideremos la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que

$$T(x,y) = (4x - 2y, 2x + y, x + y) \quad (x,y) \in \mathbb{R}^2,$$

y las bases canónicas

$$\mathcal{C}_2 = ((1,0),(0,1))$$

de $\mathbb{R}^2$ y

$$\mathcal{C}_3 = ((1,0,0),(0,1,0),(0,0,1))$$

de $\mathbb{R}^3$ respectivamente.

Para hallar la matriz asociada a T en dichas bases en primer lugar hallamos las imágenes de los vectores de la base $\mathcal{C}_2$:

$$T(1,0) = (4,2,1), \quad T(0,1) = (-2,1,1).$$

Luego calculamos las coordenadas de estos en la base $\mathcal{C}_3$. Hay poco para hacer, porque las coordenadas de un vector de $\mathbb{R}^n$ respecto a la base canónica $\mathcal{C}_n$ coinciden con el vector. Por lo tanto

$$\begin{aligned}(T(1,0))_{\mathcal{C}_3} &= (4,2,1),\\ (T(0,1))_{\mathcal{C}_3} &= (-2,1,1).\end{aligned}$$

donde hemos representado los vectores coordenadas como filas, en vez de columnas. Por tanto obtenemos

$$_{\mathcal{C}_3}T_{\mathcal{C}_2} = \begin{pmatrix} 4 & -2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

El lector notará que las entradas en cada columna son justamente los coeficientes con que aparecen x e y en cada una de las entradas del vector $T(x,y,z)$. ♣

Ejemplo 4.6.18. En el ejemplo 4.5.5 expresamos la acción de la matriz

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

en coordenadas respecto a la base

$$\mathcal{B} = \left((1+\sqrt{5},2),(1-\sqrt{5},2)\right),$$

y encontramos que

$$(AX)_{\mathcal{B}} = D(X)_{\mathcal{B}},$$

donde D es la matriz diagonal

$$D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

Vemos entonces que la matriz D es la matriz asociada en la base $\mathcal{B}$ con la transformación lineal T que define la multiplicación por A. Podemos pensar que A es la matriz asociada con T en la base canónica. El cambio a la base $\mathcal{B}$ permite entender mucho mejor la estructura de T, es decir, el efecto de la matriz A sobre los vectores de $\mathbb{R}^2$. ♣

Ejemplo 4.6.19. Interpolación y polinomios de Lagrange
La transformación

$$T : \mathbb{L} \to \mathbb{R}_3[x]$$

que hemos definido en el ejemplo 4.6.9 de esta misma sección admite una sencilla descripción matricial cuando consideramos bases en ambos espacios.

En el espacio $\mathbb{L}$ trabajaremos con la base

$$\mathcal{A} = (l_{-1}, l_0, l_1)\,.$$

Para los polinomios tenemos una base canónica, pero no es muy adecuada para tratar problemas de interpolación. Es preferible recurrir a una construcción similar a la que nos llevó a la base $\mathcal{A}$ y considerar para cada nodo

$$x_j, \quad j = 1, 2, 3, 4,$$

un polinomio p_{x_j} que tome el valor 1 en ese nodo, y 0 en los restantes. En nuestro caso los nodos son

$$-1,\ -1/3,\ 1/3,\ 1.$$

Comencemos por p_{-1}. Para que se anule en $-1/3$, $1/3$ y 1 el polinomio debe tener como factores los polinomios

$$x + 1/3, \quad x - 1/3, \quad x - 1.$$

Como p_{-1} es un polinomio de grado menor o igual a tres tendremos

$$p_{-1} = c(x + 1/3)(x - 1/3)(x - 1),$$

donde c es una constante no nula. Elegimos la constante c para que

$$p_{-1}(-1) = 1,$$

entonces

$$p_{-1} = \frac{(x + 1/3)(x - 1/3)(x - 1)}{(-2/3)(-4/3)(-2)} = -\frac{9}{16}(x + 1/3)(x - 1/3)(x - 1).$$

Repitiendo la construcción para los demas nodos encontramos que

$$\begin{aligned}
p_{-1} &= -\tfrac{9}{16}(x + 1/3)(x - 1/3)(x - 1),\\
p_{-1/3} &= \tfrac{27}{16}(x + 1)(x - 1/3)(x - 1),\\
p_{1/3} &= -\tfrac{27}{16}(x + 1)(x + 1/3)(x - 1),\\
p_{1} &= \tfrac{9}{16}(x + 1)(x + 1/3)(x - 1/3).
\end{aligned} \tag{4.24}$$

Estos cuatro polinomios forman una base de $\mathbb{R}_3[x]$, a la que llamaremos $\mathcal{L}$. Es usual referirse a esta base como los polinomios de Lagrange, de ahí la $\mathcal{L}$ en su nombre.

El polinomio que T asocia con cada función $l \in \mathbb{L}$ es simplemente

$$Tl(x) = l(-1)p_{-1}(x) + l(-1/3)p_{-1/3}(x) + l(1/3)p_{1/3}(x) + l(1)p_1(x).$$

Esta fórmula permite probar con poco trabajo la linealidad de T, una cuestión que quedó pendiente en el ejemplo 4.6.9, pero aún así seguiremos posponiendo la prueba, para hacerla más adelante con herramientas más potentes que nos evitarán prácticamente todos los cálculos. Dado que hemos expresado Tl como una combinación lineal de los elementos de la base $\mathcal{L}$ encontramos

$$(Tl)_{\mathcal{L}} = (l(-1), l(-1/3), l(1/3), l(1)).$$

Esta información es la que precisamos para calcular ${}_{\mathcal{L}}T_{\mathcal{A}}$. Ahora sólo tenemos que calcular este vector de coordenadas para los tres vectores de la base $\mathcal{A}$, y ponerlos como columnas de una matriz. La matriz que estamos buscando es

$${}_{\mathcal{L}}T_{\mathcal{A}} = \begin{pmatrix} l_{-1}(-1) & l_0(-1) & l_1(-1) \\ l_{-1}(-1/3) & l_0(-1/3) & l_1(-1/3) \\ l_{-1}(1/3) & l_0(1/3) & l_1(1/3) \\ l_{-1}(1) & l_0(1) & l_1(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 2/3 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 0 & 1 \end{pmatrix}$$

Veremos más adelante, en el ejemplo 4.7.6, como calcular la matriz ${}_{\mathcal{C}}T_{\mathcal{A}}$ que nos da las coordenadas de Tl en la base canónica. Es decir, los coeficientes de Tl. Claro está que podríamos hacer esta cálculo desarrollando los polinomios de la base $\mathcal{L}$ y agrupando los términos del mismo grado, pero un poco de álgebra lineal y cálculo con matrices nos permitirá ordenar los cálculos más eficientemente. ♣

Ejercicio 4.86.

1. Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$T(x, y, z) = (3x + 2y - 4z, x - 5y + 3z).$$

Hallar ${}_{\mathcal{A}}T_{\mathcal{B}}$ en los siguientes casos:

a) $\mathcal{B}$ y $\mathcal{A}$ son las respectivas bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$;

b) $\mathcal{B} = ((1,1,1),(1,1,0),(1,0,0))$ y $\mathcal{A}$ la base canónica de $\mathbb{R}^2$;

c) $\mathcal{B} = ((1,1,1),(1,1,0),(1,0,0))$ y $\mathcal{A} = ((1,3),(2,5))$.

2. Sea $T : \mathbb{R}_2[x] \to \mathbb{R}^4$ tal que

$$T(a + bx + cx^2) = (2a + 3b - 8c, a + b + c, 4a - 5c, 6b).$$

Hallar ${}_\mathcal{A}T_\mathcal{B}$ en los siguientes casos:

   a) $\mathcal{B}$ y $\mathcal{A}$ son las respectivas bases canónicas de $\mathbb{R}_2[x]$ y $\mathbb{R}^4$;
   b) $\mathcal{B} = \left(1, x - 1, (x - 1)^2\right)$ y $\mathcal{A}$ es la base canónica de $\mathbb{R}^4$.

Ejercicio 4.87. Dado $u_0 \in \mathbb{R}^3$ fijo, con $|u_0| = 1$, se definen transformaciones P y Q de $\mathbb{R}^3$ en $\mathbb{R}^3$ por las fórmulas

$$\begin{aligned} P(v) &= (v \cdot u_0)\, u_0, \\ Q(v) &= v - (v \cdot u_0)\, u_0. \end{aligned}$$

1. Probar que P y Q son transformaciones lineales y describirlas geométricamente.
2. Para las transformaciones P y Q y un vector $u_0 = (\alpha, \beta, \gamma)$, hallar las matrices asociadas con P y Q respecto a
   a) una base ortonormal que incluya al vector u_0;
   b) la base canónica de $\mathbb{R}^3$.

Ejercicio 4.88. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

$${}_\mathcal{A}T_\mathcal{B} = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

donde $\mathcal{B} = ((1, -1, 1), (1, 1, 1), (1, 0, 0))$ y $\mathcal{A} = ((1, 2, 0), (2, 1, 0), (1, 1, 1))$.

1. ¿T está bien definida?, es decir, ¿hay una única transformación lineal que verifique las condiciones dadas? Justifique su respuesta.
2. En caso afirmativo, hallar $T(x, y, z)$.

Ejercicio 4.89. Sea $T : \mathbb{R}_2[x] \to \mathbb{R}^3$ lineal tal que

$${}_\mathcal{U}T_\mathcal{E} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix}$$

donde

$$\mathcal{E} = \left(1, x + 1, (x + 1)^2\right), \quad \mathcal{U} = ((1, 1, 0), (1, 2, 3), (3, 2, 1)).$$

Hallar $T(x^2 + x - 1)$.

Ejercicio 4.90. Consideremos una matriz $A \in M^{2\times 2}(\mathbb{R})$ y la transformación

$$T : M^{2\times 2}(\mathbb{R}) \to M^{2\times 2}(\mathbb{R})$$

definida por

$$T(B) = AB.$$

1. Probar que T es lineal.
2. ¿Existen bases en $M^{2\times 2}(\mathbb{R})$ tal que la matriz asociada en dichas bases sea justamente A?. Justifique la respuesta.
3. Para
$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$
hallar la matriz asociada a T en la base canónica de $M^{2\times 2}(\mathbb{R})$.

Ejercicio 4.91. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$
$$T(x, y, z) = (2x - y + z, 3y - z, x + 2y + z),$$
Sea $\mathbb{S}$ el subespacio de $\mathbb{R}^3$ generado por
$$\mathcal{A} = ((1, 0, 1), (-1, 1, 2)).$$
Hallar la matriz asociada de la restricción de T a $\mathbb{S}$, $T|_{\mathbb{S}}$, de la base $\mathcal{A}$ a la base $\mathcal{C}$ canónica de $\mathbb{R}^3$.

Ejercicio 4.92. Sea $M : \mathbb{R}^3 \to \mathbb{R}^3$ una transformación lineal tal que
$$M(x, y, z) = 1/3(3y, -x + 2y + 2z, 4x - 2y + z).$$

1. Sea el subespacio $\mathbb{S} = \{(x, y, z) : x + y + z = 0\}$. Probar que si $X \in \mathbb{S}$ entonces $MX \in \mathbb{S}$.
2. Hallar la matriz asociada a
$$M|_{\mathbb{S}} : \mathbb{S} \to \mathbb{S}$$
en la base $((1, -1, 0), (1, 1, -2))$ de $\mathbb{S}$.
3. Hallar $M(1, 1, 1)$ y la matriz asociada con M en la base
$$((1, -1, 0), (1, 1, -2), (1, 1, 1))$$
de $\mathbb{R}^3$.

Observación 4.6.20. Notemos que en la matriz asociada con una transformación T está toda la información sobre T. Por otra parte, para hallar esta matriz sólo es necesario calcular los transformados de los vectores de una base del dominio. Parece entonces que T queda determinada por su comportamiento sobre una base. Esto es completamente cierto, y es una consecuencia de la linealidad de la transformación T.

Proposición 4.26. *Sean* $\mathbb{V}$ *y* $\mathbb{W}$ *dos* $\mathbb{K}$*-espacios vectoriales. Supongamos que* $\mathbb{V}$ *tiene dimensión finita* n *y que*

$$\mathcal{B} = (v_1, v_2, \dots, v_n)$$

es una base de $\mathbb{V}$*. Consideremos* n *vectores*

$$(w_1, w_2, \dots, w_n)$$

en $\mathbb{W}$*. Entonces existe una única transformación lineal*

$$T : \mathbb{V} \to \mathbb{W}$$

tal que

$$Tv_i = w_i, \quad i = 1, 2, \dots, n.$$

Ejercicio 4.93. Demostrar la proposición 4.26

Ejercicio 4.94.

1. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ una transformación lineal tal que

   $$T(1,0,0) = (2,1,0),\ T(1,1,0) = (-1,2,3),\ T(1,1,1) = (0,0,1).$$

   ¿Hay una única transformación lineal que verifique las condiciones dadas? Justifique su respuesta. Si es afirmativa entonces hallar $T(3,2,1)$.

2. Determinar si existe alguna transformación $T : \mathbb{R}_2[x] \to \mathbb{R}^2$ que satisfaga

   $$T(1) = (1,0),\ T(1+x) = (1,1),\ T(1+x+x^2) = (0,0),\ T(3+2x+x^2) = (2,1).$$

   En caso de que exista alguna hallarlas todas.

Ejercicio 4.95. En los siguientes casos hallar la forma general de las transformaciones lineales que cumplen las condiciones dadas, y determinar cuántas hay:

1. $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

   $$T(1,1,-1) = (2,1,0), \quad T(1,2,1) = (-1,2,3), \quad T(1,0,-3) = (0,0,1);$$

2. $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

   $$T(1,1,-1) = (2,1,0), \quad T(1,2,1) = (-1,2,3), \quad T(1,0,-3) = (5,0,-3);$$

3. $T : \mathbb{Z}_2^3 \to \mathbb{Z}_2^2$ tal que

   $$T(1,1,0) = (1,0),\ T(0,1,1) = (0,1),\ T(1,0,1) = (1,1);$$

4. $T : M^{2\times 2}(\mathbb{R}) \to \mathbb{R}^2$ tal que

$$T(M_1) = (1,-1), \quad T(M_2) = (1,1), \quad T(M_3) = (1,1),$$
$$T(M_4) = (3,1), \quad T(M_5) = (1,-3).$$

donde

$$M_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix},$$
$$M_4 = \begin{pmatrix} 3 & 3 \\ 1 & 2 \end{pmatrix}, \ M_5 = \begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix}.$$

4.6.3. Matriz de cambio de base

Todo espacio vectorial $\mathbb{V}$ de dimensión finita n sobre un cuerpo $\mathbb{K}$ tiene muchas bases, respecto a las cuales podemos referir los vectores del espacio tomando coordenadas. Al cambiar de base cambian las coordenadas, de una manera que está completamente determinada por las bases en juego. Ya hemos visto que introducir una base adecuada, o, equivalentemente, coordenadas adecuadas, es una herramienta útil para tratar diversos problemas, por lo que nos interesa entender cómo operan estos cambios de coordenadas. La teoría de las matrices asociadas a las transformaciones lineales resuelve este problema de una manera especialmente simple y elegante.

Dada una base $\mathcal{A}$ del espacio $\mathbb{V}$ a cada vector $v \in \mathbb{V}$ queda asociado su vector de coordenadas

$$(v)_{\mathcal{A}} \in \mathbb{K}^n.$$

Nos preguntamos entonces ¿cómo cambia la representación si seleccionamos otra base?. Es decir si consideramos otra base $\mathcal{B}$ del espacio $\mathbb{V}$, **¿cómo se relaciona $(v)_{\mathcal{B}}$ con $(v)_{\mathcal{A}}$?**

La respuesta pasa por la observación de que la transformación identidad

$$I : \mathbb{V} \to \mathbb{V}$$

que satisface

$$Iv = v,$$

para todo elemento de $\mathbb{V}$ es una transformación lineal que puede ser expresada por su matriz en las bases $\mathcal{A}$ y $\mathcal{B}$. Al hacerlo encontramos

$$(Iv)_{\mathcal{B}} =_{\mathcal{B}} (I)_{\mathcal{A}}(v)_{\mathcal{A}}.$$

Al tener en cuenta que $Iv = v$ resulta entonces

$$(v)_{\mathcal{B}} =_{\mathcal{B}} (I)_{\mathcal{A}}(v)_{\mathcal{A}},$$

lo que muestra que el cambio de coordenadas puede describirse completamente por la matriz ${}_{\mathcal{B}}(I)_{\mathcal{A}}$, a la que llamaremos *matriz de cambio de base*, de la base ("vieja") $\mathcal{A}$ a la base ("nueva") $\mathcal{B}$.

Al recordar cómo se construye la matriz asociada con una transformación encontramos que las columnas de la matriz de cambio de base de $\mathcal{A}$ a $\mathcal{B}$ son las coordenadas de los vectores de la base $\mathcal{A}$ respecto a los de la base $\mathcal{B}$.

Ejemplo 4.6.21. Si $(A_1, A_2, \ldots, A_n)$ es una base $\mathcal{B}$ de $\mathbb{K}^n$, y $\mathcal{C}$ es la base canónica, entonces ${}_{\mathcal{C}}(I)_{\mathcal{B}}$ es la matriz

$$A = (A_1, A_2, \ldots, A_n)$$

que tiene como columnas a los vectores de la base $\mathcal{B}$. Esta afirmación está justificada porque las columnas A_i, $i = 1, \ldots, n$ coinciden con sus coordenadas en la base canónica.

Otra manera de mostrar que la afirmación es correcta es observar que el vector de $\mathbb{K}^n$ que tiene coordenadas $\Lambda = (\lambda_1, \ldots, \lambda_n)$ en la base $\mathcal{B}$ es justamente la combinación lineal de los vectores de $\mathcal{B}$ con estos coeficientes. Esta combinación es exactamente el producto de la matriz A por el vector Λ, de acuerdo a nuestra definición del producto de una matriz por un vector. ♠

Ejemplo 4.6.22. La matriz de cambio de base de una base a la misma es la matriz identidad $n \times n$. ♣

Ejercicio 4.96. Hallar las dos matrices de cambio de base entre la base canónica de $\mathbb{R}^3$ y la base

$$((1,1,1),(1,0,1),(-1,0,1)).$$

Verificar que una matriz de cambio de base es la inversa de la otra.

Ejercicio 4.97.

1. Sea $\mathcal{B} = (A_1, A_2, A_3)$ una base ortonormal de $\mathbb{R}^3$. Llamemos A a la matriz que tiene a los vectores de $\mathcal{B}$ como columnas. Mostrar que

$$ {}_{\mathcal{B}}(I)_{\mathcal{C}} = A^t. $$

2. Consideremos la base

$$\mathcal{B} = \left(\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right), \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\right)$$

en $\mathbb{R}^3$. Hallar las matrices de cambio de base ${}_{\mathcal{B}}(I)_{\mathcal{C}}$ e ${}_{\mathcal{C}}(I)_{\mathcal{B}}$

3. Verificar que una matriz de cambio de base es la inversa de la otra.

Observación 4.6.23. En general, si $\mathcal{A}$ y $\mathcal{B}$ son dos bases de un mismo espacio vectorial de dimensión finita, las matrices ${}_{\mathcal{A}}(I)_{\mathcal{B}}$ ${}_{\mathcal{B}}(I)_{\mathcal{A}}$ son una inversa de la otra. No es sorprendente: cada una de ellas debe deshacer la acción de la otra.

Tal vez el lector quiera demostrar ahora esta afirmación, pero pospondremos su consideración hasta la sección 4.7, donde de la derivaremos de resultados generales acerca de la composición de transformaciones lineales y su efecto sobre las matrices asociadas. ♠

Ejemplo 4.6.24. Hallemos las matrices de cambio de base entre la base canónica

$$\mathcal{C} = \left(1, x, x^2, x^3\right)$$

de $\mathbb{R}_3[x]$ y la base

$$\mathcal{L} = \left(p_{-1}, p_{-1/3}, p_{1/3}, p_1\right)$$

del ejemplo 4.6.19, formada por los polinomios de Lagrange

$$\begin{aligned}
p_{-1} &= -\tfrac{9}{16}(x+1/3)(x-1/3)(x-1),\\
p_{-1/3} &= \tfrac{27}{16}(x+1)(x-1/3)(x-1),\\
p_{1/3} &= -\tfrac{27}{16}(x+1)(x+1/3)(x-1),\\
p_{1} &= \tfrac{9}{16}(x+1)(x+1/3)(x-1/3).
\end{aligned} \tag{4.25}$$

que permiten resolver fácilmente el problema de interpolación polinómica con nodos -1, $-1/3$, $1/3$ y 1.

Comenzamos por calcular ${}_{\mathcal{L}}(I)_{\mathcal{C}}$. Para esto debemos escribir los polinomios de la base canónica como combinaciones lineales de los polinomios en $\mathcal{L}$. Esto es sencillo, porque la base $\mathcal{L}$ fue construida para que los coeficientes de la expresión de un polinomio cualquiera en esta base se puedan hallar evaluando el polinomio en los nodos correspondientes a los vectores de $\mathcal{L}$. El resultado es

$$\begin{aligned}
1 &= 1p_{-1}(x) + 1p_{-1/3}(x) + 1p_{1/3}(x) + 1p_1(x),\\
x &= -1p_{-1}(x) - 1/3p_{-1/3}(x) + 1/3p_{1/3}(x) + 1p_1(x),\\
x^2 &= 1p_{-1}(x) + 1/9p_{-1/3}(x) + 1/9p_{1/3}(x) + 1p_1(x),\\
x^3 &= -1p_{-1}(x) - 1/27p_{-1/3}(x) + 1/27p_{1/3}(x) + 1p_1(x).
\end{aligned}$$

Ahora podemos escribir las coordenadas de la base $\mathcal{C}$ en la base $\mathcal{L}$, tal como se muestra a continuación,

$$\begin{aligned}
(1)_{\mathcal{L}} = (1,1,1,1), &\quad (x)_{\mathcal{L}} = (-1,-1/3,1/3,1),\\
(x^2)_{\mathcal{L}} = (1,1/9,1/9,1), &\quad (x^3)_{\mathcal{L}} = (-1,-1/27,1/27,1).
\end{aligned}$$

Con estas coordenadas “armamos” la matriz de cambio de base, “colgándolas” como columnas de la matriz

$$
{}_{\mathcal{L}}(I)_{\mathcal{C}} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & -\frac{1}{3} & \frac{1}{9} & -\frac{1}{27} \\ 1 & \frac{1}{3} & \frac{1}{9} & \frac{1}{27} \\ 1 & 1 & 1 & 1 \end{pmatrix}.
$$

Para hallar la matriz de cambio de base ${}_{\mathcal{C}}(I)_{\mathcal{L}}$ basta expresar los polinomios de Lagrange como combinaciones lineales de los polinomios en la base canónica, lo que es equivalente a desarrollar los polinomios en $\mathcal{L}$ y agrupar los términos según su grado. Observemos que en todos los polinomios hay un factor de la forma

$$
(x+a)(x-a) = x^2 - a^2,
$$

lo que simplifica los cálculos. Obtenemos

$$
\begin{aligned}
p_{-1} &= \tfrac{1}{16}\left(-9x^3 + 9x^2 + x - 1\right), \\
p_{-1/3} &= \tfrac{1}{16}\left(27x^3 - 9x^2 - 27x + 9\right), \\
p_{1/3} &= \tfrac{1}{16}\left(-27x^3 - 9x^2 + 27x + 9\right)), \\
p_{1} &= \tfrac{1}{16}\left(9x^3 + 9x^2 - x - 1\right).
\end{aligned} \tag{4.26}
$$

Con los coeficientes del polinomio armaremos la matriz ${}_{\mathcal{C}}(I)_{\mathcal{L}}$. Al copiar los coeficiente en la matriz hay que tener cuidado con el orden, porque en la expresión usual de un polinomio los términos aparecen ordenados de mayor a menor grado, lo que corresponde al orden inverso al que tienen los polinomios en la base $\mathcal{C}$. El resultado es

$$
{}_{\mathcal{C}}(I)_{\mathcal{L}} = \frac{1}{16}\begin{pmatrix} -1 & 9 & 9 & -1 \\ 1 & -27 & 27 & -1 \\ 9 & -9 & -9 & 9 \\ -9 & 27 & -27 & 9 \end{pmatrix}.
$$

Es muy fácil verificar que ${}_{\mathcal{L}}(I)_{\mathcal{C}}$ es la inversa de ${}_{\mathcal{C}}(I)_{\mathcal{L}}$, formando uno cualquiera de los dos posibles productos de estas matrices. Dejamos los cálculos para el lector.

Ejercicio 4.98. Verificar que ${}_{\mathcal{L}}(I)_{\mathcal{C}}$ es la inversa de ${}_{\mathcal{C}}(I)_{\mathcal{L}}$. ♣

4.6.4. Para tener presente

- La imagen de una combinación lineal de vectores por una transformación lineal, es una combinación lineal de las imágenes de los vectores, con los mismos coeficientes de la combinación original.

- Cuando una transformación lineal se expresa en términos de coordenadas se representa como una matriz.

- El cambio de coordenadas entre dos bases distintas se expresa por una matriz, que es la matriz asociada con la transformación identidad en esas bases.

4.7. Operaciones con transformaciones lineales

En esta sección consideraremos la interacción de distintas transformaciones lineales. La composición de transformaciones ocupará un lugar destacado, por dos razones:

- es habitual en el cálculo que varias transformaciones lineales se apliquen una tras otra, para producir una nueva transformación lineal;
- los cambios de base pueden considerarse como un caso particular de transformación lineal. Su aplicación para representar una transformación lineal en una base conveniente puede entenderse en el marco de una teoría general para las composiciones de transformaciones lineales.

Sobre el final de la sección discutiremos la posibilidad de sumar transformaciones lineales, y de multiplicarlas por escalares.

4.7.1. Composición de transformaciones lineales

Comencemos por recordar qué cosa es la composición de dos funciones

$$f : X \to Y, \quad g : Y \to Z.$$

Se trata de la función

$$g \circ f : X \to Z,$$

definida por

$$(g \circ f)(x) = g\left(f(x)\right).$$

La composición $g \circ f$ de las dos funciones es simplemente la función que resulta de aplicar primero f y luego g.

Comenzamos discutiendo la composición de transformaciones lineales a través de un ejemplo que tiene que ver con el cálculo aproximado de la integral de una función.

Ejemplo 4.7.1. Supongamos que nos interesa evaluar la integral de una función continua f definida sobre el intervalo $[-1, 1]$. En general f no tendrá una primitiva que se pueda expresar en términos de funciones elementales, y tendremos que recurrir a algún procedimiento numérico. Implementaremos aquí un procedimiento que esbozábamos al comienzo de la sección 4.1: evaluaremos la función en los puntos -1, 0 y 1, a partir de estos valores construiremos un interpolante lineal a trozos, y calcularemos la integral del interpolante. Tomaremos el valor de esta integral como una aproximación del valor de la integral de f.

Observación 4.7.2. El número así hallado puede ser una aproximación muy grosera al verdadero valor de la integral, pero el procedimiento constituye la base para un procedimiento general de cálculo, que permite obtener la integral con tanta precisión como se desee. Ver el ejercicio 4.99. ♠

Repasemos las operaciones necesarias:

1. comenzamos por evaluar f en los tres nodos que usaremos en el cálculo, para obtener la terna
$$(f(-1), f(0), f(1)).$$
Esta operación define una aplicación lineal
$$E : C([-1,1]) \to \mathbb{R}^3;$$
2. a partir de estos valores construimos el interpolador lineal a trozos que toma valores $(f(-1)$, $f(0)$ y $f(1)$ en los nodos -1, 0 y 1 respectivamente. Esta operación define una nueva transformación lineal
$$L : \mathbb{R}^3 \mapsto \mathbb{L},$$
donde $\mathbb{L}$ es el espacio de funciones lineales a trozos definidas sobre $[-1,1]$ que introdujimos en el ejemplo 4.1.1 de la sección 4.1;
3. por último, integrar una función de $\mathbb{L}$ en $[-1,1]$ define una aplicación lineal
$$I : \mathbb{L} \to \mathbb{R}$$
que a cada función $l \in \mathbb{L}$ le asocia el valor
$$I(l) = \frac{1}{2}l(-1) + l(0) + \frac{1}{2}l(1).$$

De la aplicación sucesiva de estas tres operaciones resulta que asociamos a f el valor aproximado
$$I_a(f) = \frac{1}{2}f(-1) + f(0) + \frac{1}{2}f(1).$$
Observemos que $I_a(f)$ es una transformación lineal
$$I_a : C([-1,1]) \to \mathbb{R},$$
que consiste en la composición
$$I_a = I \circ L \circ E$$
de las tres operaciones de evaluar, interpolar linealmente e integrar.

Ejercicio 4.99. Método de los trapecios
Para alcanzar una mayor precisión en el cálculo de la integral se puede usar este mismo método de interpolación lineal e interpolación colocando más nodos en el intervalo $[a, b]$ en el que se pretende integrar una función $f \in C([a, b])$.

Repetir el análisis para una cantidad arbitraria de nodos

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b,$$

en un intervalo $[a, b]$. ♣

Ejemplo 4.7.3. Composición de transformaciones lineales entre espacios $\mathbb{K}^m$
Supongamos que tenemos dos transformaciones lineales

$$S : \mathbb{K}^n \to \mathbb{K}^m, \quad T : \mathbb{K}^m \to \mathbb{K}^l.$$

Sabemos que cada una de estas transformaciones consiste en la multiplicación por una matriz. Por lo tanto, existen matrices $A \in M^{m\times n}(\mathbb{K})$ y $B \in M^{l\times m}(\mathbb{K})$ tales que

$$S(X) = AX,\ X \in \mathbb{K}^n; \quad T(Y) = BY,\ Y \in \mathbb{K}^m.$$

Aplicar sucesivamente T y S es lo mismo que multiplicar por las matrices B y A que codifican su acción, por lo tanto

$$(T \circ S)(X) = B(AX) = (BA)X.$$

Vemos entonces que la transformación $T \circ S$ es también lineal, y su acción es equivalente a la de multiplicar por el producto BA de las matrices asociadas con T y S.

Como resumen, vemos que la composición de transformaciones lineales entre espacios $\mathbb{K}^n$ es una nueva transformación lineal, y la matriz que la representa es el producto de las matrices de las transformaciones que intervienen en la composición. Este resultado es completamente general, tal como veremos en esta misma sección.

Cerramos este ejemplo con un ejercicio que servirá al lector para revisar la noción de composición, la representación matricial de las transformaciones lineales, y los resultados que aquí se tratan.

Ejercicio 4.100. Para las transformaciones lineales

$$S : \mathbb{R}^3 \to \mathbb{R}^2, \quad T : \mathbb{R}^2 \to \mathbb{R}^2,$$

tales que

$$S(x, y, z) = (2x, y + z), \quad T(x, y) = (y, x),$$

hallar la composición $T \circ S$, y las expresiones matriciales de T, S y $T \circ S$. ¿Es posible formar la composición $S \circ T$? ♣

Comenzamos por mostrar que la composición de transformaciones lineales es una transformación lineal.

Proposición 4.27. *Sean $S : \mathbb{U} \to \mathbb{V}$ y $T : \mathbb{V} \to \mathbb{W}$ dos transformaciones lineales entre los $\mathbb{K}$-espacios vectoriales $\mathbb{U}$, $\mathbb{V}$, y $\mathbb{W}$, entonces $T \circ S : \mathbb{U} \to \mathbb{W}$ es una transformación lineal.*

Prueba. Para cualquier pareja de escalares α_1, α_2, en $\mathbb{K}$, y de vectores u_1, u_2, en el espacio $\mathbb{U}$ consideremos

$$(T \circ S)(\alpha_1 u_1 + \alpha_2 u_2) = T\left(S(\alpha_1 u_1 + \alpha_2 u_2)\right).$$

Esta igualdad está justificada porque aplicar $T \circ S$ es exactamente lo mismo que aplicar primero S y luego T. Usando la linealidad de la aplicación S obtenemos

$$(T \circ S)(\alpha_1 u_1 + \alpha_2 u_2) = T\left(\alpha_1 S u_1 + \alpha_2 S u_2\right).$$

Por la linealidad de T resulta

$$(T \circ S)(\alpha_1 u_1 + \alpha_2 u_2) = T\left(\alpha_1 T(S u_1) + \alpha_2 T(S u_2)\right).$$

Reconocemos la composición $T \circ S$ en la aplicación sucesiva de T y S, y concluimos que para vectores u_1, u_2, y escalares α_1, α_2, cualesquiera se tiene

$$(T \circ S)(\alpha_1 u_1 + \alpha_2 u_2) = T\left(\alpha_1 (T \circ S) u_1 + \alpha_2 (T \circ S) u_2\right),$$

que es equivalente a la linealidad de $T \circ S$. □

Ejemplo 4.7.4. Comenzaremos por volver sobre el problema de interpolación polinómica que tratamos en el ejemplo 4.6.19, página 499. Supongamos que queremos encontrar un polinomio real de grado menor o igual que 3 que tome valores prefijados

$$(y_{-1}, y_{-1/3}, y_{1/3}, y_1) \tag{4.27}$$

en los nodos -1, $-1/3$, $1/3$ y 1, del intervalo $[-1, 1]$. La solución única al problema puede expresarse como una combinación lineal de los polinomios de Lagrange, en la forma

$$p(x) = y_{-1} p_{-1}(x) + y_{-1/3} p_{-1/3}(x) + y_{1/3} p_{1/3}(x) + y_1 p_1(x). \tag{4.28}$$

Ya sabemos que formar combinaciones lineales define una transformación lineal, de modo que la correspondencia, a la que llamaremos

$$P : \mathbb{R}^4 \to \mathbb{R}_3[x],$$

que a cada vector de $\mathbb{R}^4$ como en (4.27) le asocia la combinación lineal (4.28) define una transformación lineal. La forma explícita de los polinomios que intervienen en la combinación lineal se encuentra en la fórmula (4.24), página 500. No es importante para el resultado de linealidad que acabamos de establecer, pero sí lo es para saber que (4.28) es la solución del problema de interpolación.

Supongamos ahora que queremos interpolar los valores que una función continua $f \in C([-1,1])$ toma en los nodos -1, $-1/3$, $1/3$ y 1 por un polinomio en $\mathbb{R}_3[x]$. Este procedimiento produce una aplicación

$$\begin{aligned} C([-1,1]) &\to \mathbb{R}_3[x], \\ f &\mapsto p(x), \end{aligned}$$

que a cada función f le asocia el polinomio de interpolación que le corresponde. Esta aplicación es lineal, porque consiste en la composición de la aplicación lineal

$$f \mapsto (f(-1), f(-1/3), f(1/3), f(1))$$

que consiste en tomar evaluaciones de f, con la aplicación lineal P que construye el polinomio interpolador a partir de los valores en los nodos.

Este mismo razonamiento puede aplicarse a la transformación de "cambio de interpolador" que introdujimos en el ejemplo 4.6.9, página 491. Consideramos allí una transformación

$$T : \mathbb{L} \to \mathbb{R}_3[x]$$

que a cada función l, en el espacio $\mathbb{L}$ formado por las funciones continuas en $[-1,1]$, lineales entre los nodos -1, 0 y 1, le asocia el único polinomio $p(x)$ de grado menor o igual que 3 que satisface

$$p(-1) = l(-1), \quad p(-1/3) = l(-1/3), \quad p(1/3) = l(1/3), \quad p(1) = l(1).$$

Esta transformación T es también la composición de la evaluación de l en los nodos -1, $-1/3$, $1/3$ y 1, con la transformación lineal que resuelve el problema de interpolación polinómica. Por lo tanto T es lineal. ♠

Ahora vamos a trabajar en el contexto de espacios de dimensión finita, y a analizar cómo se refleja la composición en las matrices asociadas. Consideremos tres $\mathbb{K}$-espacios vectoriales $\mathbb{U}$, $\mathbb{V}$ y $\mathbb{W}$, y transformaciones lineales

$$S : \mathbb{U} \to \mathbb{V}, \quad T : \mathbb{V} \to \mathbb{W},$$

tal como en la proposición 4.27. Agreguemos además la hipótesis de que los espacios tienen dimensión finita, y que $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{C}$ son bases de $\mathbb{U}$, $\mathbb{V}$ y $\mathbb{W}$ respectivamente. Supondremos conocidas las matrices ${}_{\mathcal{C}}(T)_{\mathcal{B}}$ y ${}_{\mathcal{B}}(S)_{\mathcal{A}}$ y buscaremos determinar a partir de ellas la matriz ${}_{\mathcal{C}}(T \circ S)_{\mathcal{A}}$.

El cálculo se reduce a la aplicación sucesiva de S y T, expresando todo en coordenadas. Dado $u \in \mathbb{U}$ tenemos

$$(Su)_{\mathcal{B}} = {}_{\mathcal{B}}(S)_{\mathcal{A}}\,(u)_{\mathcal{A}},$$

Podemos aplicar T al vector Su. El resultado es

$$T(Su) = (T \circ S)u.$$

Naturalmente

$$((T \circ S)u)_{\mathcal{C}} = (T(Su))_{\mathcal{C}}.$$

Las coordenadas de la imagen por T de cualquier vector v en $\mathbb{V}$ pueden calcularse aplicando la matriz ${}_{\mathcal{C}}(T)_{\mathcal{B}}$ a $(v)_{\mathcal{B}}$. En particular, podemos aplicar este resultado a Su, para obtener

$$(T(Su))_{\mathcal{C}} = {}_{\mathcal{C}}(T)_{\mathcal{B}}\,(Su)_{\mathcal{B}}.$$

Concluimos

$$((T \circ S)u)_{\mathcal{C}} = {}_{\mathcal{C}}(T)_{\mathcal{B}}\,{}_{\mathcal{B}}(S)_{\mathcal{A}}\,(u)_{\mathcal{A}}.$$

Por lo tanto, las coordenadas de la imagen de u por la composición $T \circ S$ se obtienen multiplicando sucesivamente por las matrices ${}_{\mathcal{C}}(T)_{\mathcal{B}}$ y ${}_{\mathcal{B}}(S)_{\mathcal{A}}$. La definición del producto de matrices es tal que la aplicación sucesiva de estas dos matrices es igual a multiplicar por su producto. Entonces el producto

$${}_{\mathcal{C}}(T)_{\mathcal{B}}\,{}_{\mathcal{B}}(S)_{\mathcal{A}}$$

es la matriz que permite calcular las coordenadas $((T \circ S)u)_{\mathcal{C}}$ a partir de $(u)_{\mathcal{A}}$.

El resultado de esta discusión es nuestra siguiente proposición acerca de la matriz asociada a una composición.

Proposición 4.28. *Consideremos tres $\mathbb{K}$-espacios vectoriales $\mathbb{U}$, $\mathbb{V}$ y $\mathbb{W}$ de dimensión finita, con bases $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{C}$ respectivamente, y dos transformaciones lineales*

$$S : \mathbb{U} \to \mathbb{V}, \quad T : \mathbb{V} \to \mathbb{W}.$$

Entonces

$${}_{\mathcal{C}}(T \circ S)_{\mathcal{A}} = {}_{\mathcal{C}}(T)_{\mathcal{B}}\,{}_{\mathcal{B}}(S)_{\mathcal{A}}. \tag{4.29}$$

En la demostración de la proposición que acabamos de enunciar esta implícito el uso del siguiente resultado del calculo matricial, que dejamos al lector como un ejercicio de repaso.

Ejercicio 4.101. Sean A y B dos matrices en $M^{m\times n}(\mathbb{K})$. Mostrar que la igualdad

$$AX = BX$$

se satisface para todo $X \in \mathbb{K}^n$ si y sólo si $A = B$.

4.7.2. Cambios de base

La teoría que estamos desarrollando permite demostrar con sencillez importantes resultados acerca de cómo se modifica la matriz de una transformación lineal al cambiar de base en el espacio.

Comenzaremos por dar una demostración del resultado, previsible, por cierto, que dice que la matriz de cambio de base de $\mathcal{A}$ a $\mathcal{B}$ es la inversa de la matriz de cambio de $\mathcal{B}$ a $\mathcal{A}$. Ya vimos un adelanto de este resultado en los ejercicios 4.96 y 4.97 de la página 506. Por otra parte, una matriz tiene que deshacer el cambio que la otra hace, por lo que es natural esperar que una sea la inversa de la otra.

Proposición 4.29. *Sean $\mathcal{A}$ y $\mathcal{B}$ dos bases de un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ de dimensión finita n. Entonces la matriz de cambio de base de $\mathcal{A}$ a $\mathcal{B}$ es la inversa de la matriz de cambio de $\mathcal{B}$ a $\mathcal{A}$.*

Prueba. Consideremos la transformación lineal identidad, que compuesta con si misma vuelve a dar la transformación identidad:

$$I_{\mathbb{V}} \circ I_{\mathbb{V}} = I_{\mathbb{V}}.$$

Aplicando la proposición 4.28 obtenemos

$$I_n = {}_{\mathcal{A}}(I_{\mathbb{V}})_{\mathcal{A}} = {}_{\mathcal{A}}(I_{\mathbb{V}})_{\mathcal{B}}\,{}_{\mathcal{B}}(I_{\mathbb{V}})_{\mathcal{A}},$$

lo que implica que

$${}_{\mathcal{A}}(I_{\mathbb{V}})_{\mathcal{B}} = {}_{\mathcal{B}}(I_{\mathbb{V}})_{\mathcal{A}}^{-1}. \quad \square \tag{4.30}$$

Observación 4.7.5. Aunque no es necesario hacerlo para saber que se satisface (4.30), también podemos verificar directamente, a partir de la proposición 4.28 que

$$I_n {}_{\mathcal{B}}(I_{\mathbb{V}})_{\mathcal{A}}\,{}_{\mathcal{A}}(I_{\mathbb{V}})_{\mathcal{B}}.$$

Ejercicio 4.102. Completar los detalles de esta observación. ¿Por qué basta asegurarse de que uno sólo de los dos posibles productos entre ${}_{\mathcal{B}}(I_{\mathbb{V}})_{\mathcal{A}}$ y ${}_{\mathcal{A}}(I_{\mathbb{V}})_{\mathcal{B}}$ es igual a la matriz identidad para saber que una matriz es la inversa de la otra? ♠

Recordemos que una matriz cuadrada A, de tamaño $n \times n$ es *ortogonal* si $A^t A = I_n$.

Ejercicio 4.103.

1. Sea $\mathcal{O}$ una base ortonormal de $\mathbb{R}^3$ y designemos con $\mathcal{C}$ a la base canónica.

    a) Mostrar que ${}_{\mathcal{C}}I_{\mathcal{O}}$ es una matriz ortogonal.

    b) Hallar ${}_{\mathcal{O}}I_{\mathcal{C}}$.

2. Sea $\mathcal{B}$ una base de $\mathbb{R}^3$. Mostrar que $\mathcal{B}$ es ortonormal si y sólo si ${}_{\mathcal{C}}I_{\mathcal{B}}$ es ortogonal.

3. ¿Qué puede afirmarse acerca de las matrices de cambio de base entre bases ortonormales cualesquiera de $\mathbb{R}^3$?

Ejemplo 4.7.6. En este ejemplo completaremos los cálculos relativos a la transformación T entre el espacio $\mathbb{L}$ de las funciones lineales a trozos, con nodos -1, 0 y 1, y $\mathbb{R}_3[x]$, con la que trabajamos en los ejemplos 4.6.19 y 4.7.4.

La acción de T consiste en asociar a cada función $l \in \mathbb{L}$ un interpolador polinómico Tl que coincide con l en los nodos -1, $-1/3$, $1/3$ y 1. Habíamos escogido representar las funciones de $\mathbb{L}$ en la base $\mathcal{A}$, que tiene la propiedad de que

$$l_{\mathcal{A}} = (l(-1), l(0), l(1)).$$

Es decir, el vector de coordenadas de una función l respecto a esta base coincide con el vector de sus evaluaciones en los nodos -1, 0 y 1.

Habíamos calculado la matriz ${}_{\mathcal{L}}(T)_{\mathcal{A}}$, asociada a T en las bases $\mathcal{A}$ y $\mathcal{L}$. Esta matriz permite hallar

$$(Tl)_{\mathcal{L}} = {}_{\mathcal{L}}(T)_{\mathcal{A}} l_{\mathcal{A}}.$$

Si quisiéramos expresar el polinomio Tl en la forma usual, referida a la base canónica, necesitaríamos las coordenadas

$$(Tl)_{\mathcal{C}} = {}_{\mathcal{C}}(I)_{\mathcal{L}}(Tl)_{\mathcal{L}}.$$

Naturalmente, combinando estas dos últimas fórmulas, obtenemos

$$(Tl)_{\mathcal{C}} = {}_{\mathcal{C}}(I)_{\mathcal{L}}\, {}_{\mathcal{L}}(T)_{\mathcal{A}}\, l_{\mathcal{A}}.$$

La conclusión es que las coordenadas de Tl respecto a la base $\mathcal{C}$ se obtienen multiplicando $l_{\mathcal{A}}$ por el producto

$$_{\mathcal{C}}(I)_{\mathcal{L}}\,{}_{\mathcal{L}}(T)_{\mathcal{A}}$$

de la matriz de cambio de base de $\mathcal{L}$ a $\mathcal{C}$ y la matriz asociada a T en las bases $\mathcal{A}$ y $\mathcal{L}$. Por lo tanto, este producto es justamente la matriz $_{\mathcal{C}}(T)_{\mathcal{A}}$, asociada con T en las bases $\mathcal{A}$ y $\mathcal{C}$. La matriz $_{\mathcal{C}}(I)_{\mathcal{L}}$ fue calculada en el ejemplo 4.6.24. página 4.6.24, y es igual a

$$_{\mathcal{C}}(I)_{\mathcal{L}} = \frac{1}{16}\begin{pmatrix} -1 & 9 & 9 & -1 \\ 1 & -27 & 27 & -1 \\ 9 & -9 & -9 & 9 \\ -9 & 27 & -27 & 9 \end{pmatrix}.$$

Concluimos entonces que

$$_{\mathcal{C}}(T)_{\mathcal{A}} = \frac{1}{16}\begin{pmatrix} -1 & 9 & 9 & -1 \\ 1 & -27 & 27 & -1 \\ 9 & -9 & -9 & 9 \\ -9 & 27 & -27 & 9 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 2/3 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 1 & 6 & 1 \\ -4 & 0 & 4 \\ 3 & -6 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Naturalmente,

$$(Tl)_{\mathcal{C}} = \frac{1}{8}\begin{pmatrix} 1 & 6 & 1 \\ -4 & 0 & 4 \\ 3 & -6 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l(-1) \\ l(0) \\ l(1) \end{pmatrix}.$$

Si en vez de expresar todo en coordenadas, escribimos el polinomio que corresponde a l, encontramos

$$\begin{aligned}(Tl)[x] = \tfrac{1}{8}\big((3l(-1) - 6l(0) + 3l(1))\,x^2 + (-4l(-1) + 4l(1))\,x+ \\ l(-1) + 6l(0) + l(1)\big).\end{aligned}$$

A pesar de que habíamos considerado, en principio, polinomios de tercer grado para asegurarnos de poder ajustar los valores en cuatro nodos, la forma especial de las funciones lineales a trozos hace que en el ajuste sólo intervengan los polinomios de segundo grado, en el subespacio $\mathbb{R}_2[x]$, que está estrictamente contenido en $\mathbb{R}_3[x]$.

Observemos entonces que no todos los polinomios de $\mathbb{R}_3[x]$ pueden obtenerse por este procedimiento. Esto es un caso particular de una propiedad general de las transformaciones lineales que implica, entre otras cosas, que no puede existir una aplicación lineal sobreyectiva que tenga como dominio un espacio de dimensión 3, y como codominio un espacio de dimensión 4. Discutiremos este fenómeno en las próximas secciones. ♣

El ejemplo 4.7.6 nos muestra la construcción de una matriz asociada a una transformación en unas bases, a partir de una matriz referida a otras bases. El procedimiento es completamente general, y lo discutiremos a continuación.

Supongamos que tenemos una transformación lineal T, con dominio en un espacio vectorial $\mathbb{V}$ y codominio en $\mathbb{W}$, y conocemos la matriz ${}_{\mathcal{B}}(T)_{\mathcal{A}}$, referida a las bases $\mathcal{A}$ y $\mathcal{B}$ de $\mathbb{V}$ y $\mathbb{W}$. Puestos en esta situación, podría interesarnos representar la acción de T en otras bases diferentes. En el ejemplo 4.7.6 queríamos volver a la base canónica de $\mathbb{R}_3[x]$, porque en las coordenadas respecto a esa base podemos leer directamente los coeficientes de los polinomios. En otros ejemplos la elección de una nueva base puede estar motivada por consideraciones geométricas. También podemos plantearnos el interesantísimo problema de buscar las bases que dan la representación más simple de T, y que permiten entender mejor su acción, una cuestión de la mayor importancia práctica.

Supongamos entonces que tenemos otro par de bases $\mathcal{A}'$ y $\mathcal{B}'$ en $\mathbb{V}$ y $\mathbb{W}$ respectivamente. Nuestro interés es hallar la matriz ${}_{\mathcal{B}'}(T)_{\mathcal{A}'}$, o, equivalentemente, aprender a calcular $(Tv)_{\mathcal{B}'}$ a partir de $v_{\mathcal{A}'}$. Sabemos hacer los cálculos en las bases $\mathcal{A}$ y $\mathcal{B}$, porque conocemos ${}_{\mathcal{B}}(T)_{\mathcal{A}}$, entonces procedemos de la siguiente manera:

1. comenzamos por transformar el vector de coordenadas $v_{\mathcal{A}'}$ en
$$v_{\mathcal{A}} = {}_{\mathcal{A}}\left(I_{\mathbb{V}}\right)_{\mathcal{A}'} v_{\mathcal{A}'};$$
2. tenemos las herramientas para calcular
$$(Tv)_{\mathcal{B}} = {}_{\mathcal{B}}(T)_{\mathcal{A}}\, v_{\mathcal{A}} = {}_{\mathcal{B}}(T)_{\mathcal{A}}\, {}_{\mathcal{A}}\left(I_{\mathbb{V}}\right)_{\mathcal{A}'} v_{\mathcal{A}'};$$
3. pero en realidad nos interesa
$$(Tv)_{\mathcal{B}'} = {}_{\mathcal{B}'}\left(I_{\mathbb{W}}\right)_{\mathcal{B}} (Tv)_{\mathcal{B}} = {}_{\mathcal{B}'}\left(I_{\mathbb{W}}\right)_{\mathcal{B}}\, {}_{\mathcal{B}}(T)_{\mathcal{A}}\, {}_{\mathcal{A}}\left(I_{\mathbb{V}}\right)_{\mathcal{A}'} v_{\mathcal{A}'}.$$

En resumen, para cambiar de la matriz asociada con una transformación lineal en unas bases a su matriz en otras bases sólo hay que multiplicar por las matrices de cambio de base apropiadas. El resultado de esta discusión se recoge en la siguiente proposición, que probamos con un argumento ligeramente diferente al que nos condujo a formularla. De todos modos, digamos que en la discusión preliminar hemos prácticamente completado una demostración alternativa.

Proposición 4.30. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales de dimensión finita sobre un mismo cuerpo $\mathbb{K}$ y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Consideremos dos bases $\mathcal{A}$, y $\mathcal{A}'$ del espacio $\mathbb{V}$, y bases $\mathcal{B}$ y $\mathcal{B}'$ de $\mathbb{W}$. Entonces*
$${}_{\mathcal{B}'}(T)_{\mathcal{A}'} = {}_{\mathcal{B}'}\left(I_{\mathbb{W}}\right)_{\mathcal{B}}\, {}_{\mathcal{B}}(T)_{\mathcal{A}}\, {}_{\mathcal{A}}\left(I_{\mathbb{V}}\right)_{\mathcal{A}'} \tag{4.31}$$

donde las matrices ${}_{\mathcal{B}'}(I_{\mathbb{W}})_{\mathcal{B}}$ y ${}_{\mathcal{A}}(I_{\mathbb{V}})_{\mathcal{A}'}$ son las matrices de cambio de base entre las bases $\mathcal{B}$ y $\mathcal{B}'$, y entre $\mathcal{A}'$ y $\mathcal{A}$, en los espacios $\mathbb{W}$ y $\mathbb{V}$ respectivamente.

Prueba. La prueba se reduce a escribir la transformación T en la forma

$$T = I_{\mathbb{W}} \circ T \circ I_{\mathbb{V}}, \tag{4.32}$$

para hacer aparecer las matrices de cambio de base como expresiones de las transformaciones identidad en ciertas bases.

Al usar (4.32) para expresar T en las bases ${}_{\mathcal{A}'}$ y ${}_{\mathcal{B}'}$ obtenemos

$${}_{\mathcal{B}'}(T)_{\mathcal{A}'} = {}_{\mathcal{B}'}\left(I_{\mathbb{W}} \circ T \circ I_{\mathbb{V}}\right)_{\mathcal{A}'}.$$

Al emplear la proposición 4.28 para escribir la matriz asociada de la composición que aparece en el miembro de la derecha como un producto de matrices asociadas se obtiene (4.31). Sólo hay que tomar la base $\mathcal{A}$ como base del espacio $\mathbb{V}$ considerado como dominio de T y codominio de $I_{\mathbb{V}}$, y la base $\mathcal{B}$ como base del espacio $\mathbb{W}$ considerado como codominio de T y dominio de $I_{\mathbb{W}}$. $\square$

Un caso particular especialmente interesante se obtiene cuando $\mathbb{V} = \mathbb{W}$, $\mathcal{A} = \mathcal{B}$ y $\mathcal{A}' = \mathcal{B}'$.

Corolario 4.31. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión finita, T una transformación lineal de $\mathbb{V}$ en sí mismo y $\mathcal{A}$, $\mathcal{A}'$, dos bases del espacio $\mathbb{V}$. Entonces*

$${}_{\mathcal{A}'}(T)_{\mathcal{A}'} = {}_{\mathcal{A}'}(I)_{\mathcal{A}}\,{}_{\mathcal{A}}(T)_{\mathcal{A}}\,{}_{\mathcal{A}}(I)_{\mathcal{A}'} = {}_{\mathcal{A}'}(I)_{\mathcal{A}}\,{}_{\mathcal{A}}(T)_{\mathcal{A}}\,{}_{\mathcal{A}'}(I)^{-1}_{\mathcal{A}}. \tag{4.33}$$

Ejercicio 4.104.

1. Sea $T : \mathbb{R}_3[x] \to \mathbb{R}^3$ una transformación lineal definida por la matriz

$${}_{\mathcal{B}}T_{\mathcal{A}} = \begin{pmatrix} -1 & 1 & 0 & 2 \\ 2 & 2 & 1 & 3 \\ 0 & -1 & 1 & 2 \end{pmatrix},$$

donde

$$\mathcal{A} = (x^3 - 1, x^2 + 1, x + 2, -1), \quad \mathcal{B} = ((1, -1, 0), (0, 1, 1), (0, 0, 1)).$$

Hallar ${}_{\mathcal{C}}T_{\mathcal{D}}$ donde $\mathcal{D} = (1 + x, x^2 - 1, x + x^3, -1)$ y $\mathcal{C}$ es la base canónica de $\mathbb{R}^3$.

2. Sea la transformación lineal $T : \mathbb{R}_2[x] \to \mathbb{R}^2$ definida por la matriz

$${}_{\mathcal{C}}T_{\mathcal{D}} = \begin{pmatrix} 1 & 2 & 0 \\ -5 & 1 & -4 \end{pmatrix},$$

donde $\mathcal{D}$ y $\mathcal{C}$ son las bases canónicas de $\mathbb{R}_2[x]$ y $\mathbb{R}^2$ respectivamente. Hallar la matriz asociada a T en las bases

$$\mathcal{A} = (x^2 - x + 1, x^2 + x, x^2), \quad \mathcal{B} = ((1, 1), (1, -1)).$$

Ejercicio 4.105. Se considera la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$T(x, y) = \frac{1}{5}(3x + 4y, 4x - 3y).$$

1. Hallar la matriz ${}_{\mathcal{A}}T_{\mathcal{A}}$, siendo

$$\mathcal{A} = ((2, 1)(-1, 2)).$$

2. ¿Cuál es el significado geométrico de la transformación T?

El ejercicio anterior nos muestra que la representación T en una base adecuada puede ser de gran ayuda para comprender su acción. En los próximo ejercicios explotaremos la misma idea, para construir representaciones matriciales de algunas transformaciones geomtŕicas, y para interpretar geométricamente algunas transformaciones definidas por matrices. Recordemos que todas las transformaciones que preservan la longitud de los vectores de $\mathbb{R}^3$ son transformaciones lineales (ver el ejercicio 4.81, en la pǵina 4.81), por lo que operaciones como rotaciones y simetrías, que no alteran la longitud de los vectores a los que se aplican, deben caer en el marco de la teoría lineal que estamos desarrollando.

Ejercicio 4.106. Simetrías

1. Consideremos un plano π que pasa por el origen O de $\mathbb{R}^3$ y tiene versor normal N. Mostrar que el simétrico SX de cualquier vector X respecto al plano π es

$$SX = X - 2(X \cdot N)N. \tag{4.34}$$

2. Hallar la matriz M asociada en la base canónica de una simetría respecto al plano $x + 2y - 3z = 0$. Hacerlo de dos maneras:

   a) empleando la fórmula (4.34);

   b) calculando primero la matriz asociada respecto a la base

   $$((-2, 1, 0), (0, 3, 2), (1, 2, -3))$$

   y haciendo luego los cambios de base necesarios.

3. Hallar la expresión $S(x, y, z)$ del vector simétrico de $(x, y, z) \in \mathbb{R}^3$.

4. Calcular el producto M^2.

Ejercicio 4.107. Giros

Consideremos los giros G_+ y G_-, de ángulo $\pi/2$ alrededor de la recta que pasa por el origen y tiene dirección $(1, 1, 1)$. Definimos G_+ y G_- por la condición de que las ternas

$$(X, G_+X, (1, 1, 1)), \quad (X, G_-X, (1, 1, 1))$$

sean directa e inversa respectivamente.

1. Hallar una base ortonormal $\mathcal{B}$ de $\mathbb{R}^3$ tal que las matrices asociadas con G_+ y G_- respecto a esa base sean, respectivamente,

$$_{\mathcal{B}}(G_+)_{\mathcal{B}} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad _{\mathcal{B}}(G_-)_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2. Hallar las expresiones $G_+(x,y,z)$ y $G_-(x,y,z)$ para cualquier (x,y,z) en $\mathbb{R}^3$.
3. Hallar la expresión de la composición $G_+ \circ G_-$.

Las proyecciones son transformaciones que no conservan la longitud de los vectores, pero también son lineales, tal como vimos en el ejercicio 4.82, página 4.82.

Ejercicio 4.108. Consideremos $P : \mathbb{R}^3 \to \mathbb{R}^3$ la proyección ortogonal sobre el plano $2x - y + z = 0$. Hallar la proyección P para cualquier $X \in \mathbb{R}^3$ de dos maneras diferentes:

1. Hallando la matriz asociada a P en la base canónica utilizando lo aprendido en el capítulo de Geometría (ver la fórmula 3.50 en la página 353).
2. Hallando la matriz asociada a P respecto a una base ortonormal convenientemente elegida, y luego calculando la matriz asociada en la base canónica de $\mathbb{R}^3$.

4.7.3. Suma y producto por un escalar

Las transformaciones que tienen como codominio un espacio vectorial pueden sumarse y multiplicarse por escalares, porque estan operaciones son posibles para todos los valores que toman las funciones. Esta construcción es completamente análoga a la que se usa para definir la suma $f+g$ de, por ejemplo, funciones reales, o el producto λf de una función real f por un número real λ. Introducimos a continuación las definiciones necesarias.

Definición 4.16 (Suma de funciones y producto por un escalar). *Sea X un conjunto cualquiera, $\mathbb{W}$ un $\mathbb{K}$-espacio vectorial, y*

$$f : X \to \mathbb{W}, \quad g : X \to \mathbb{W},$$

*dos funciones cualesquiera. Se define la **suma de** f **y** g como la función*

$$f + g : X \to \mathbb{W}$$

tal que

$$(f+g)(x) = f(x) + g(x) \quad x \in X.$$

Si λ es un elemento del cuerpo $\mathbb{K}$ definimos ***el producto de*** λ ***por*** f *como la función*

$$\lambda f : X \to \mathbb{W}$$

tal

$$(\lambda f)(x) = \lambda f(x), \quad x \in X.$$

Las funciones definidas sobre un conjunto cualquiera X, que toman valores en un $\mathbb{K}$-espacio vectorial $\mathbb{W}$ son, con estas operaciones naturales de suma y producto por un escalar, un nuevo $\mathbb{K}$-espacio vectorial. Este resultado ya se presentó al lector en el ejercicio 4.2.

Cuando el dominio X es otro $\mathbb{K}$-espacio vectorial $\mathbb{V}$ tiene especial interés fijar nuestra atención sobre el conjunto de las transformaciones lineales de $\mathbb{V}$ en $\mathbb{W}$, porque guardan una relación privilegiada con la estructura lineal de los espacios. Llamaremos $\mathbb{L}(\mathbb{V}, \mathbb{W})$ a este conjunto de transformaciones. Es decir

$$\mathbb{L}(\mathbb{V}, \mathbb{W}) = \{T : \mathbb{V} \to \mathbb{W};\ T \text{ lineal}\}.$$

Este conjunto es cerrado bajo las operaciones de suma y producto por escalares, tal como surge de las proposiciones 4.32 y 4.33, que enunciamos a continuación.

Proposición 4.32. *Si $S : \mathbb{V} \to \mathbb{W}$ y $T : \mathbb{V} \to \mathbb{W}$ son transformaciones lineales, entonces $S + T$ es una transformación lineal.*

PRUEBA. Sean v_1 y v_2 dos vectores en $\mathbb{V}$, y α_1 y α_2 escalares. Entonces, por la definición de suma $S + T$ tenemos que

$$(S + T)(\alpha_1 v_1 + \alpha_2 v_2) = S(\alpha_1 v_1 + \alpha_2 v_2) + T(\alpha_1 v_1 + \alpha_2 v_2).$$

La linealidad de S y de T permite transformar la igualdad en

$$(S + T)(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 S v_1 + \alpha_2 S v_2 + \alpha_1 T v_1 + \alpha_2 T v_2.$$

Por supuesto, las propiedades de la suma y el producto por un escalar en el espacio vectorial $\mathbb{W}$ nos permiten reordenar los términos y extraer α_1 y α_2 como factores comunes a los términos en que aparecen. Obtenemos

$$(S + T)(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 (S v_1 + T v_1) + \alpha_2 (S v_2 + T v_2).$$

Al reconocer la suma $S + T$ actuando sobre v_1 y v_2 concluimos

$$(S + T)(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 (S + T) v_1 + \alpha_2 (S + T) v_2,$$

que es justamente la expresión de la linealidad de $T + S$. □

Proposición 4.33. *Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal y λ un escalar entonces λT es una transformación lineal.*

Dejamos la demostración como un ejercicio para el lector.

Ejercicio 4.109. Demostrar la proposición 4.33.

Ejemplo 4.7.7. Consideremos las aplicación $S : \mathbb{R}^3 \to \mathbb{R}^2$ tal que

$$S(x, y, z) = (x + y + z, x - z),$$

y $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$T(x, y, z) = (x - y, y + z).$$

Entonces la suma de S y T es la transformación $S + T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por

$$\begin{aligned}(S + T)(x, y, z) = S(x, y, z) + T(x, y, z) = \\ (x + y + z, x - z) + (x - y, y + z) = (2x + z, x + y).\end{aligned}$$

El producto $5T$ es

$$(5T)(x, y, z) = 5(x - y, y + z) = (5x - 5y, 5y + 5z).$$

Como era de esperar, reconocemos expresiones lineales en las fórmulas que describen la acción de $S + T$ y de $5T$. ♣

Ejercicio 4.110. Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales, mostrar que $\mathbb{L}(\mathbb{V}, \mathbb{W})$, con las operaciones de suma y producto por un escalar definidas en esta sección, es un $\mathbb{K}$-espacio vectorial.

Las operaciones de suma de aplicaciones lineales, y de producto por un escalar, se reflejan sobre las matrices asociadas.

Proposición 4.34. *Consideremos dos $\mathbb{K}$-espacios vectoriales de dimensión finita $\mathbb{V}$ y $\mathbb{W}$, dos transformaciones lineales*

$$T : \mathbb{V} \to \mathbb{W}, \quad S : \mathbb{V} \to \mathbb{W},$$

y un escalar λ en $\mathbb{K}$. Consideremos bases $\mathcal{A}$ y $\mathcal{B}$ bases en $\mathbb{V}$ y $\mathbb{W}$ respectivamente. Entonces

$$\begin{aligned}{}_{\mathcal{B}}(T + S)_{\mathcal{A}} &= {}_{\mathcal{B}}(T)_{\mathcal{A}} + {}_{\mathcal{B}}(S)_{\mathcal{A}}, \\ {}_{\mathcal{B}}(\lambda T)_{\mathcal{A}} &= \lambda\, {}_{\mathcal{B}}(T)_{\mathcal{A}}.\end{aligned}$$

Prueba. Consideremos un vector $v \in \mathbb{V}$. La demostración pasa por analizar la expresión en coordenadas de $(T+S)v$. Tenemos

$$(T+S)v = Tv + Sv,$$

entonces, tomando coordenadas respecto a la base $\mathcal{B}$ en ambos miembros de la igualdad, y usando la linealidad de la aplicación $\mathrm{coord}_{\mathcal{B}}$ encontramos

$$((T+S)v)_{\mathcal{B}} = (Tv)_{\mathcal{B}} + (Sv)_{\mathcal{B}},$$

Las coordenadas de los transformados de v que aparecen en el miembro de la derecha se pueden expresar usando las matrices asociadas, obtenemos

$$((T+S)v)_{\mathcal{B}} = {}_{\mathcal{B}}(T)_{\mathcal{A}}v_{\mathcal{A}} + {}_{\mathcal{B}}(S)_{\mathcal{A}}v_{\mathcal{A}}.$$

El producto de matrices tiene la propiedad distributiva respecto a la suma, lo que nos permite escribir la igualdad anterior en la forma

$$((T+S)v)_{\mathcal{B}} = \left({}_{\mathcal{B}}(T)_{\mathcal{A}} + {}_{\mathcal{B}}(S)_{\mathcal{A}}\right) v_{\mathcal{A}}.$$

Vemos entonces que las coordenadas de $(T+S)v$ en la base $\mathcal{B}$ se obtienen multiplicando la matriz

$${}_{\mathcal{B}}(T)_{\mathcal{A}} + {}_{\mathcal{B}}(S)_{\mathcal{A}}$$

por $v_{\mathcal{A}}$, lo que implica que esta matriz es la matriz asociada con $T+S$ en las bases $\mathcal{A}$ y $\mathcal{B}$.

La demostración de la segunda parte de la proposición, para el producto λT, es similar, y la dejamos como ejercicio para el lector.

Ejercicio 4.111. Completar la demostración de la proposición 4.34. □

Ejercicio 4.112. Sean P y Q las proyecciones ortogonales sobre el plano de $\mathbb{R}^3$ definido por la ecuación

$$x + 2y + 2z = 0,$$

y sobre el subespacio generado por $(1,2,2)$ respectivamente.

1. Hallar las matrices ${}_{\mathcal{A}}P_{\mathcal{A}}$ y ${}_{\mathcal{A}}Q_{\mathcal{A}}$, siendo

$$\mathcal{A} = \left((1/3, 2/3, 2/3), (-2/\sqrt{5}, 1/\sqrt{5}, 0), (-2/3\sqrt{5}, -4/3\sqrt{5}, 5/3\sqrt{5})\right).$$

2. Hallar las expresiones $P(x,y,z)$ y $Q(x,y,z)$ de la proyección para cualquier $(x,y,z) \in \mathbb{R}^3$.
3. Hallar la transformación lineal $P+Q$, e interpretar geométricamente el resultado.

4. Hallar las transformaciones P^2 y Q^2, e interpretar geométricamente el resultado.

Ejercicio 4.113. Sea N un versor de $\mathbb{R}^{\nvDash}$.

1. Mostrar que la aplicación P definida por

$$PX = (X \cdot N)N$$

   es lineal. ¿Cuál es su significado geométrico?

2. ¿Cuál es su significado geométrico de la transformación $Q = I - P$? Como es habitual, I indica la transformación identidad en $\mathbb{R}^3$.

3. ¿ Y cuál es el de $S = I - 2P$?

Ejercicio 4.114. Consideremos dos $\mathbb{K}$-espacios vectoriales $\mathbb{V}$ y $\mathbb{W}$, con ba ses $\mathcal{A}$ y $\mathcal{B}$, y dimensiones n y m respectivamente. Mostrar que la aplicación

$$M : \mathbb{L}(\mathbb{V}, \mathbb{W}) \to M^{m\times n}$$

que a cada transformación T le asocia

$$M(T) = {}_{\mathcal{B}}(T)_{\mathcal{A}}$$

es una transformación lineal.

4.7.4. Para tener presente

- La composición de transformaciones lineales es una nueva transformación lineal.

- La matriz asociada con una composición es el producto de las matrices asociadas[16]. Esta relación es una consecuencia de la definición que dimos para el producto de las matrices, que fue introducido para comportarse bien respecto a la composición.

- La teoría de cambios de base puede manejarse en el marco de la teoría de las matrices asociadas con las transformaciones lineales: un cambio de base es la representación de la transformación identidad en dos bases diferentes.

- Una combinación lineal de transformaciones lineales es una nueva transformación lineal. El conjunto de todas las transformaciones lineales entre dos $\mathbb{K}$-espacios vectoriales es un nuevo espacio vectorial.

[16]Hay que elegir las bases de manera coherente para que esto sea cierto, ver los detalles en el cuerpo de esta misma sección.

4.8. Inyectividad, sobreyectividad, e isomorfismos

En esta sección estudiaremos la inyectividad y sobreyectividad de las transformaciones lineales. El análisis extiende, en algún sentido, al que hicimos en el capítulo 1 a la hora de tratar la compatibilidad y la determinación de los sistemas lineales. En particular, encontraremos que tanto la inyectividad como la sobreyectividad de una transformación lineal pueden caracterizarse por medio de algunos subespacios asociados a la transformación (su núcleo e imagen) o estudiando su efecto sobre combinaciones lineales. Es decir, empleando la estructura lineal.

Consideraremos el importante caso de las transformaciones lineales que son a la vez inyectivas y sobreyectivas, a las que llamaremos *isomorfismos*. Si entre dos espacios hay un isomorfismo entonces, aunque su aspecto pueda ser muy diferente, son indistinguibles para todo lo que tiene que ver con su estructura lineal.

4.8.1. Transformaciones inyectivas y biyectivas

Las nociones de inyectividad y sobreyectividad corresponden al dominio de la teoría de funciones, y se aplican a funciones de cualquier naturaleza. No sólo a las transformaciones lineales.

Recordemos que una función

$$f : X \to Y$$

es *inyectiva* si nunca envía dos elementos diferentes de su dominio X al mismo elemento de su dominio Y. Algo que puede expresarse diciendo que la igualdad

$$f(x_1) = f(x_2)$$

implica

$$x_1 = x_2.$$

Diremos que f es *sobreyectiva* cuando todo elemento y en Y es la imagen por f de algún elemento de x. Es decir, para cada $y \in Y$ existe $x \in X$ tal que

$$f(x) = y.$$

Naturalmente, el elemento x con esta propiedad no tiene por qué ser único, aunque lo será cuando f sea además inyectiva.

La sobreyectividad puede caracterizarse por medio del conjunto *imagen* de f, al que es usual indicar con la notación $f(X)$. Este conjunto es un subconjunto de Y, y está formado por las imágenes de los puntos de X. Es decir

$$f(X) = \{f(x);\ x \in X\}$$

Ejemplo 4.8.1. Si $f : \mathbb{R} \to \mathbb{R}$ está definida por $f(x) = x^2$, entonces su imagen $f(\mathbb{R})$ es el intervalo $[0, +\infty) \subset \mathbb{R}$. Esta función no es sobreyectiva, porque ningún número negativo está en su imagen. Ni inyectiva, porque $f(x) = f(-x)$ para todo $x \in \mathbb{R}$. ♣

Ejemplo 4.8.2. Una nueva visita a los sistemas lineales
A continuación vamos a analizar la inyectividad y la sobreyectividad en el contexto de las transformaciones lineales. Comenzaremos por recordar lo que sabemos en el caso de aplicaciones lineales entre espacios $\mathbb{K}^n$, que se reducen a la multiplicación por matrices. Estudiaremos entonces una aplicación lineal

$$T : \mathbb{K}^n \to \mathbb{K}^m,$$

definida por

$$X \mapsto AX, \tag{4.35}$$

donde A es una matriz $m \times n$ con entradas en un cuerpo $\mathbb{K}$. En este contexto, las afirmaciones acerca de la inyectividad y la sobreyectividad de T como en (4.35) se traducen inmediatamente en afirmaciones acerca de la compatibilidad y determinación de sistemas de ecuaciones lineales. Ya hemos estudiado este problema en el capítulo 1.

Que la aplicación T sea sobreyectiva es equivalente a que el sistema lineal

$$AX = B \tag{4.36}$$

sea compatible para todo B. El subespacio $\operatorname{im}(A)$ de $\mathbb{K}^m$ que hemos dado en llamar la imagen de A, es justamente la imagen de la aplicación T, porque está formado por el conjunto de todos los posibles valores AX. La sobreyectividad de T queda caracterizada por este subespacio de su codominio $\mathbb{K}^m$, ya que T es sobreyectiva si y sólo si

$$\operatorname{im}(A) = \mathbb{K}^m.$$

La transformación T es inyectiva si y sólo si para cada B en su imagen, el subespacio $\operatorname{im}(A)$ hay un único X tal que $AX = B$. Esta condición equivale a pedir que el sistema (4.36) sea determinado siempre que es compatible. Ya sabemos que esto es equivalente a que el núcleo $\ker(A)$ de la matriz A sólo contenga al vector nulo O. Notemos que $\ker(A)$ es justamente el conjunto de vectores de $\mathbb{K}^n$ cuya imagen por T es el vector nulo de $\mathbb{K}^m$. Se trata de un subespacio en el dominio $\mathbb{K}^n$ de la aplicación T. En resumen, T es inyectiva si y sólo si se satisface

$$\ker(A) = \{O\}.$$

Vemos entonces que la inyectividad y la sobreyectividad de cualquier aplicación lineal T definida entre espacios $\mathbb{K}^n$ están caracterizadas por subespacios asociados con T. ♣

El análisis que hemos hecho para las transformaciones lineales entre espacios $\mathbb{K}^n$ puede repetirse para transformaciones lineales T cualesquiera. Comenzamos por introducir el *núcleo* y la *imagen* de T, que resultarán ser subespacios de su dominio y codominio respectivamente.

Definición 4.17. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal.*

- *Llamaremos **núcleo** de T al subconjunto de $\mathbb{V}$ formado por los vectores cuya imagen por T es el vector nulo $O_{\mathbb{W}}$ de $\mathbb{W}$. Indicaremos este conjunto con la notación* $\ker(T)$*. Entonces*

$$\ker(T) = \{v \in \mathbb{V};\ Tv = O_{\mathbb{W}}\}.$$

- *Llamaremos **imagen** de T al subconjunto de $\mathbb{W}$ formado por los vectores que son la imagen por T de algún vector de $\mathbb{V}$. Indicaremos este conjunto con la notación* $\operatorname{im}(T)$*. Entonces*

$$\operatorname{im}(T) = \{Tv;\ v \in \mathbb{V}\}.$$

Observación 4.8.3. Por supuesto, $\operatorname{im}(T)$ no es otra cosa que el conjunto imagen de T en el sentido en que se puede definir el conjunto imagen de cualquier función. Por esta razón también es corriente usar la notación $T(\mathbb{V})$ para designarlo.

Otra manera de expresar la caracterización de la imagen de T es decir que un vector $w \in \mathbb{W}$ pertenece a $\operatorname{im}(T)$ si y sólo si existe algún vector $v \in \mathbb{V}$ tal que $w = Tv$.

La novedad en el contexto de los espacios vectoriales es que este conjunto tiene una estructura lineal, ya que que se trata de un subespacio en el codominio $\mathbb{W}$ de T (ver la proposición 4.35) Esta estructura lineal proviene de la estructura lineal de los espacios $\mathbb{V}$ y $\mathbb{W}$, y de la aplicación T. ♠

Ejemplo 4.8.4. Calculemos el núcleo y la imagen de la transformación lineal

$$T : \mathbb{K}_2[t] \to \mathbb{R}^2$$

del ejemplo 4.6.16 tal que al polinomio $p(x) = ax^2+bx+c$ le hace corresponder

$$Tp = (p(-1), p(1)).$$

Calculemos primero el núcleo de T. Queremos encontrar todos aquellos vectores tales que

$$Tp = O,$$

por tanto buscamos los polinomios que cumplen

$$Tp = (p(-1), p(1)) = (a - b + c, a + b + c) = (0, 0).$$

Ahora debemos resolver un sistema homegéneo de dos ecuaciones

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right).$$

Escalerizando el sistema obtenemos

$$b = 0 \quad c = -a.$$

Por tanto el núcleo de T es

$$\ker(T) = \{p(x) \in \mathbb{K}_2[x] : \ p(x) = ax^2 - a\}.$$

Debemos calcular ahora la imagen de la transformación T. Una forma de hacerlo es investigar para qué valores (α, β) con $\alpha, \ \beta \in \mathbb{R}$ el sistema

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & \alpha \\ 1 & 1 & 1 & \beta \end{array}\right)$$

es compatible. El sistema es claramente compatible para todo $(\alpha, \beta) \in \mathbb{R}^2$, por tanto

$$\operatorname{im}(T) = \left\{(\alpha, \beta) \ \in \mathbb{R}^2\right\}.$$

Vale la pena observar que el sistema lineal que estudiamos para determinar el núcleo de la transformación T es el caso particular $\alpha = \beta = 0$ del sistema que apareció al determinar su imagen. ♣

Ejemplo 4.8.5. Busquemos ahora el núcleo y la imagen del la aplicación $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que

$$T(x, y) = (4x - 2y, 2x + y, x + y) \qquad (x, y) \in \mathbb{R}^2,$$

del ejemplo 4.6.17.

Esta transformación consiste en multiplicar por la matriz

$$A = \begin{pmatrix} 4 & -2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix},$$

que es justamente la matriz asociada con T en las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$.

Un vector $X \in \mathbb{R}^2$ está en el núcleo de T si y sólo si $AX = O$, y un $Y \in \mathbb{R}^3$ está en la imagen si y sólo si es de la forma AX, para algún $X \in \mathbb{R}^2$. Vemos entonces que el núcleo y la imagen de T son justamente el núcleo y la imagen de la matriz A, y podemos calcularlos con las técnicas aprendidas en el capítulo 1.

Al escalerizar la matriz A obtenemos

$$\begin{pmatrix} 4 & -2 \\ 0 & 2 \\ 0 & 0 \end{pmatrix},$$

una forma escalerizada sin variables libres. Esto indica que el núcleo de la matriz es trivial, y que la colección de todas las columnas de la matriz forma una base del espacio de columnas.

Dejamos para el lector el siguiente ejercicio.

Ejercicio 4.115. Hallar un sistema lineal homogéneo que es satisfacho por todos los vectores en la imagen de T, y sólo por ellos. ♣

El núcleo y la imagen de una transformación lineal son subespacios de su dominio y codominio respectivamente, tal como se establece en la próxima proposición.

Proposición 4.35. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ y*

$$T : \mathbb{V} \to \mathbb{W}$$

una transformación lineal. Entonces el núcleo $\ker(T)$ es un subespacio vectorial de $\mathbb{V}$, y la imagen $\operatorname{im}(T)$ es un subespacio vectorial de $\mathbb{W}$.

El lector podrá demostrar esta proposición generalizando lo hecho para matrices, pero la deduciremos a partir de algunas consideraciones más generales acerca de imágenes y preimágenes de conjuntos, y su relación con la estructura lineal cuando se trabaja con subespacios de espacios vectoriales y transformaciones lineales.

Para una función cualquiera

$$f : X \to Y$$

podemos definir la *imagen* $f(S)$ de un subconjunto $S \subset X$ como

$$f(S) = \{f(s);\ s \in S\}.$$

Este conjunto no es otra cosa que el conjunto formado por todas las imágenes de elementos de S.

Ejemplo 4.8.6. Si $f : \mathbb{R} \to \mathbb{R}$ está definida por $f(x) = x^2$, entonces la imagen $f([0,2])$ del intervalo $[0,2]$ es el intervalo $[0,4]$. ♣

Ejemplo 4.8.7. La imagen $f(X)$ de una aplicación es justamente la imagen de todo su dominio X. Para una aplicación lineal $T : \mathbb{V} \to \mathbb{W}$ vale la misma afirmación, la imagen $\operatorname{im}(T)$ es justamente $T(\mathbb{V})$. ♣

Ejemplo 4.8.8. Hallemos la imagen de la familia

$$\mathcal{A} = ((1,1,-1),(-1,-1,2),(1,1,3))$$

por la aplicación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que

$$T(x,y,z) = (x - 2y + z, -y + z).$$

Para eso calculamos el transformado de cada vector de la familia $\mathcal{A}$

$$T(1,1,-1) = (-2,-2) \quad T(-1,-1,2) = (3,3) \quad T(1,1,3) = (2,2).$$

Por tanto, la imagen de $\mathcal{A}$ por la aplicación T es la familia

$$f(\mathcal{A}) = ((-2,-2),(3,3),(2,2)).$$

Observemos que cuando trabajemos con familias ordenadas de vectores consideraremos también la imagen como una familia ordenada, con el orden que se "hereda" de la familia inicial.

También para $f : X \to Y$ podemos definir la *preimagen* $f^{-1}(U)$ de un subconjunto $U \subset Y$ como

$$f^{-1}(U) = \{x \in X;\ f(x) \in U\}.$$

Este conjunto no es otra cosa que el conjunto formado por todos los elementos de X cuyas imágenes caen en U.

Ejemplo 4.8.9. Si $f : \mathbb{R} \to \mathbb{R}$ está definida por $f(x) = x^2$, entonces la preimagen $f^{-1}([-1,4])$ del intervalo $[-1,4]$ es el intervalo $[-2,2]$. ♣

Ejemplo 4.8.10. Hemos definido el núcleo de una aplicación lineal $T : \mathbb{V} \to \mathbb{W}$ como el conjunto de todos los vectores v de $\mathbb{V}$ tales que $Tv = O$. Esta definición puede formularse en términos de la preimagen del subespacio trivial $\{O\} \subset \mathbb{W}$, ya que $Tv = O$ si y sólo si $Tv \in \{O\}$. Por lo tanto

$$\ker(T) = \{v \in V;\ Tv \in \{O\}\},$$

lo que pone en evidencia que el núcleo de T no es otra cosa que la preimagen por T del subespacio $\{O\}$ en $\mathbb{W}$. Por supuesto, sólo se trata de una nueva manera de decir que en el núcleo de T están todos los vectores que T envía al vector nulo de $\mathbb{W}$. ♣

Proposición 4.36. *Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal.*

1. *Si $\mathbb{S}$ es un subespacio de $\mathbb{V}$, entonces $T(\mathbb{S})$ es un subespacio de $\mathbb{W}$.*

2. *Si $\mathbb{U}$ es un subespacio de $\mathbb{W}$, entonces $T^{-1}(\mathbb{U})$ es un subespacio de $\mathbb{V}$.*

Ejercicio 4.116. Demostrar la proposición 4.36. Deducir de ella la proposición 4.35 acerca de la estructura lineal de los conjuntos $\mathrm{im}(T)$ y $\ker(T)$.

La inyectividad de una aplicación lineal T puede caracterizarse por su núcleo. Esto es una generalización de lo que ya sabemos para transformaciones que consisten en multiplicar por matrices.

La inyectividad implica que a vectores distintos le corresponden vectores distintos. Cuando se trata de transformaciones lineales puede formularse simplemente en términos de "direcciones distintas" en el espacio. Recordemos la idea intuitiva de que los vectores de una familia linealmente independiente apuntan en direcciones "esencialmente diferentes". Una transformación lineal inyectiva es la que tiene la propiedad de que transforma vectores que apuntan en direcciones esencialmente diferentes, en vectores que también apuntan en direcciones esencialmente diferentes. El enunciado preciso de esta propiedad se recoge en la siguiente proposición.

Proposición 4.37. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Entonces las siguientes afirmaciones son equivalentes:*

1. *T es inyectiva;*

2. $\ker(T) = \{O_{\mathbb{V}}\}$.

3. *para todo conjunto linealmente independiente*

$$\mathcal{A} = (v_1, v_2, \dots, v_n) \subset V$$

se cumple que $T(\mathcal{A}) \subset \mathcal{W}$ también es linealmente independiente en $\mathbb{W}$.

Prueba. Comenzaremos por mostrar la equivalencia entre las condiciones 1 y 3.

Como T es lineal

$$TO_{\mathbb{V}} = O_{\mathbb{W}}.$$

Si T es inyectiva entonces no puede haber otro vector de $\mathbb{V}$ cuya imagen sea $O_{\mathbb{W}}$, por lo que $\ker(T)$ sólo contiene a $O_{\mathbb{V}}$.

Supongamos ahora que

$$\ker(T) = \{O_{\mathbb{V}}\}.$$

Si para una pareja de vectores v_1 y v_2 en $\mathbb{V}$ se satisface

$$Tv_1 = Tv_2$$

entonces

$$Tv_1 - Tv_2 = O_{\mathbb{W}}.$$

La linealidad de T implica

$$T(v_1 - v_2) = O_{\mathbb{W}},$$

por lo que

$$v_1 - v_2 \in \ker(T) = \{O_{\mathbb{V}}\}.$$

Esta inclusión implica

$$v_1 - v_2 = O_{\mathbb{V}},$$

de donde inmediatamente concluimos que necesariamente

$$v_1 = v_2.$$

Por lo tanto la condición 2 implica la 1.

Mostraremos a continuación la equivalencia entre las condiciones 2 y 3. Comencemos suponiendo que se satisface 2, y consideremos una familia linealmente independiente

$$\mathcal{A} = (v_1, v_2, \ldots, v_n) \subset V.$$

Para mostrar la independencia lineal de

$$T(\mathcal{A}) = (Tv_1, Tv_2, \ldots, Tv_n)$$

recurriremos directamente a la definición. Formemos una combinación lineal de los transformados de los vectores en $\mathcal{A}$ con coeficientes

$$\lambda_i, \quad i = 1, \ldots, n,$$

tales que

$$\lambda_1 Tv_1 + \ldots + \lambda_n Tv_n = O_{\mathbb{W}}. \tag{4.37}$$

Como T es lineal tenemos que

$$T(\lambda_1 v_1 + \ldots + \lambda_n v_n) = 0_{\mathcal{W}}.$$

Por lo tanto, el vector

$$\lambda_1 v_1 + \ldots + \lambda_n v_n$$

está en el núcleo de T, y tiene que ser el vector nulo de $\mathbb{V}$, que es el único elemento de $\ker(T)$. Entonces

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = O_{\mathbb{V}}$$

y la independencia lineal de $\mathcal{A}$ implica

$$\lambda_1 = \ldots = \lambda_n = O.$$

Por lo tanto, los coeficientes de cualquier combinación lineal como en (4.37) deben ser nulos. Esto es, por definición, la independencia lineal de la familia $T(\mathcal{A})$.

Resta mostrar que 3 implica 2. Para ello consideraremos un vector no nulo $v \in \mathbb{V}$, y mostraremos que su imagen es no nula. Si $v \neq O_{\mathbb{V}}$ entonces la pequeña familia

$$(v)$$

es linealmente independiente, y su transformada por T, que se reduce a

$$(Tv)$$

también lo es. Por lo tanto Tv es no nulo, porque la familia $(O_{\mathbb{W}})$ es linealmente dependiente. $\square$

Ejercicio 4.117. Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal, y $\mathbb{S}$ un subespacio de $\mathbb{V}$. Mostrar que la restricción

$$T|_{\mathbb{S}} : \mathbb{S} \to \mathbb{W}$$

de T a $\mathbb{S}$ es inyectiva si y sólo si

$$\mathbb{S} \cap \ker(T) = \{O_{\mathbb{V}}\}.$$

La sobreyectividad de una transformción lineal puede caracterizarse en términos del subespacio $\operatorname{im}(T)$ y de generadores. Este resultado va en la misma línea que la proposición 4.37, pero es menos sutil.

Proposición 4.38. *Son equivalentes las siguientes afirmaciones:*

1. *T es sobreyectiva;*
2. $\operatorname{im}(T) = \mathbb{W}$*;*
3. *para todo conjunto $\mathcal{A}$ que es un generador de $\mathbb{V}$ se cumple que $T(\mathcal{A})$ es un generador de $\mathbb{W}$;*

Prueba. Que la condición 1 implica 2, es obvio. Es prácticamente la definición de sobreyectividad.

Probemos ahora que 2 implica 3. Para ello consideremos un generador

$$\mathcal{A} = (v_1, \dots, v_n)$$

del espacio $\mathbb{V}$ y un vector w cualquier en $\mathbb{W}$. Sabemos que $w \in \operatorname{im}(T)$, por lo que existe $v \in \mathbb{V}$ tal que $T(v) = w$. Este vector v se puede escribir como combinación lineal de los elementos del generador $\mathcal{A}$, en la forma

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Al aplicar T a ambos miembros de esta igualdad, empleando la linealidad, obtenemos

$$w = Tv = \lambda_1 T v_1 + \dots + \lambda_n T v_n.$$

Por lo tanto el vector w puede expresarse como una combinación de $T(\mathcal{A})$. Como w es un vector cualquiera en $\mathbb{W}$ concluimos que $T(\mathcal{A})$ es un generador de $\mathbb{W}$.

Para mostrar que que 3 implica 1 consideremos nuevamente un generador

$$\mathcal{A} = (v_1, \dots, v_n)$$

del espacio $\mathbb{V}$ y un vector w cualquier en $\mathbb{W}$. Sabemos, por la hipótesis 3, que $T(\mathcal{A})$ es un generador de $\mathbb{W}$. Por lo tanto existen escalares λ_i, $i = 1, \dots, n$, tales que

$$w = \lambda_1 T v_1 + \dots + \lambda_n T v_n.$$

Naturalmente, el miembro de la derecha es igual a

$$T(\lambda_1 v_1 + \dots \lambda_n v_n),$$

de donde concluimos que w es el transformado por T del vector

$$\lambda_1 v_1 + \dots \lambda_n v_n \in \mathbb{V}.$$

Como w es un vector cualquiera en $\mathbb{W}$ hemos mostrado que T es sobreyectiva. □

El siguiente ejercicio contiene un resultado que es un refinamiento de alguna de las conclusiones de la proposición anterior.

Ejercicio 4.118.

1. Mostrar que si existe un generador de $\mathbb{V}$ tal que su imagen por T es un generador de $\mathbb{W}$ entonces T es sobreyectiva.
2. Investigar si es suficiente que exista algún conjunto linealmente independiente en $\mathbb{V}$ tal que su imagen por T sea linealmente independiente para asegurar que T es inyectiva.

4.8.2. Isomorfismos entre espacios vectoriales.

En la sección anterior hemos considerado las propiedades de inyectividad y sobreyectividad en el contexto de transformaciones lineales. Consideraremos ahora las transformaciones *biyectivas*. Recordemos que, en general, diremos que una función

$$f : X \to Y$$

es biyectiva si es inyectiva y sobreyectiva. Cuando una función es biyectiva encontramos que para cada $y \in Y$ existe un único $x \in X$ tal que $f(x) = y$. Podemos definir entonces la función inversa

$$f^{-1} : Y \to X,$$

que a cada $y \in Y$ le asocia ese x. Naturalmente, las funciones f y f^{-1} son cada una inversa de la otra, en el sentido de que la aplicación de una de ellas deshace el efecto de la otra. Podemos expresar esto en término de composiciones como

$$f^{-1} \circ f = I_X, \quad f \circ f^{-1} = I_Y,$$

donde las aplicaciones I_X e I_Y son las identidades en X e Y respectivamente, caracterizadas por

$$I_X(x) = x,\ x \in X; \quad I_Y(y) = y,\ y \in Y.$$

Recordemos también que la inversa de una transformación lineal biyectiva es también una transformación lineal. Ver la proposición 4.24, en la página 492. Una transformación lineal biyectiva entre dos espacios permite identificarlos completamente desde el punto de vista de su estructura lineal, tal como ocurre con un $\mathbb{K}$-espacio $\mathbb{V}$ de dimensión finita n y el espacio $\mathbb{K}^n$ por medio de las coordenadas referidas a cualquier base de $\mathbb{V}$. En esta sección desarrollaremos la teoría correspondiente.

Definición 4.18. *Sean* $\mathbb{V}$*,* $\mathbb{W}$ *espacios vectoriales sobre un mismo cuerpo* $\mathbb{K}$*. Diremos que una transformación lineal*

$$T : \mathbb{V} \to \mathbb{W}$$

biyectiva es un ***isomorfismo****.*

Cuando existe un isomorfismo de $\mathbb{V}$ *en* $\mathbb{W}$*, diremos que* $\mathbb{V}$ *y* $\mathbb{W}$ *son* ***isomorfos****.*

Ejemplo 4.8.11. Todo $\mathbb{K}$-espacio vectorial $\mathbb{V}$ de dimensión n es isomorfo a $\mathbb{K}^n$. Cualquier base $\mathcal{B}$ del espacio $\mathbb{V}$ sirve para fabricar un isomorfismo entre ambos espacios, porque

$$\mathrm{coord}_{\mathcal{B}} : \mathbb{V} \to \mathbb{K}^n$$

es una transformación lineal y biyectiva.

Ejemplo 4.8.12. El ejemplo anterior puede extenderse para mostrar que cualquier pareja de $\mathbb{K}$-espacios vectoriales $\mathbb{V}$ y $\mathbb{W}$ de dimensión n son isomorfos. El isomorfismo se construye a través de $\mathbb{K}^n$, fijando bases $\mathcal{A}$ y $\mathcal{B}$ en los espacios $\mathbb{V}$ y $\mathbb{W}$. El mapa

$$\mathrm{coord}_{\mathcal{B}}^{-1} \circ \mathrm{coord}_{\mathcal{A}} : \mathbb{V} \to \mathbb{W},$$

es una transformación lineal y biyectiva entre $\mathbb{V}$ y $\mathbb{W}$.

Ejercicio 4.119. ¿Qué vector de $\mathbb{W}$ es la imagen de cada uno de los vectores de la base $\mathcal{A}$ por este isomorfismo? ♣

Ejercicio 4.120. En este ejercicio consideraremos espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$. Probar que

1. todo espacio vectorial es isomorfo a sí mismo;
2. si $\mathbb{V}$ es isomorfo a $\mathbb{W}$ entonces $\mathbb{W}$ es isomorfo a $\mathbb{V}$;
3. si $\mathbb{U}$ es isomorfo a $\mathbb{V}$ y $\mathbb{V}$ es isomorfo a $\mathbb{W}$, entonces $\mathbb{U}$ es isomorfo a $\mathbb{W}$.

Nuestra próxima proposición combina las caracterizaciones de inyectividad y sobreyectividad de la sección anterior, para producir una caracterización de los isomorfismos entre espacios de dimensión finita. Excluiremos de su enunciado el caso de los espacios triviales que están formados sólo por su vector nulo. Ver el ejercicio 4.122.

Proposición 4.39. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita, y*

$$T : \mathbb{V} \to \mathbb{W}$$

una transformación lineal. Las siguientes afirmaciones son equivalentes :

1. *T es biyectiva.*
2. *Para toda base $\mathcal{B}$ del espacio $\mathbb{V}$ se cumple que $T(\mathcal{B})$ es una base de $\mathbb{W}$.*
3. *El núcleo de T es trivial, y su imagen es todo $\mathbb{W}$.*

Prueba. Es inmediato que 1 y 3 son equivalentes, porque es una consecuencia directa de las proposiciones 4.37 y 4.38.

Mostraremos a continuación la equivalencia entre 1 y 2, comenzando por el hecho de que la condición 1 implica la 2.

Consideremos una base $\mathcal{B}$ de $\mathbb{V}$. La familia $\mathcal{B}$ es entonces linealmente independiente, y generadora de $\mathbb{V}$. Como T es inyectiva, aplicando la proposición 4.37 concluimos que $T(\mathcal{B}$ es linealmente independiente. Como T es sobreyectiva, concluimos que $T(\mathcal{B})$ es un generador de $\mathbb{W}$, por la proposición 4.38. Por lo tanto $T(\mathcal{B})$ es una base de $\mathbb{W}$.

Probemos por último que (2) implica (1). Consideremos una base

$$\mathcal{B} = (v_1, \dots, v_n)$$

de $\mathbb{V}$. Como la imagen de cualquier base de $\mathbb{V}$ es una base de $\mathbb{W}$, esta base $\mathcal{B}$ es un generador de $\mathbb{V}$ cuya imagen por T genera $\mathbb{W}$. Aplicando el resultado del ejercicio 4.118 concluimos que T es sobreyectiva.

Para probar que T es inyectiva mostraremos que su núcleo es trivial. Consideremos un vector cualquiera $v \in \ker(T) \subseteq \mathbb{V}$. Podemos expresarlo como una combinación lineal

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n,$$

de los vectores en $\mathcal{B}$. Aplicando T y usando la linealidad obtenemos

$$Tv = \alpha_1 T v_1 + \ldots + \alpha_n T v_n.$$

Pero como $v \in \ker(T)$, se tiene que $Tv = O$, lo que implica

$$\alpha_1 T v_1 + \ldots + \alpha_n T v_n = O$$

La familia

$$(Tv_1, \dots, Tv_n) = T(\mathcal{B})$$

es linealmente independiente, por ser una base de $\mathbb{W}$, lo que implica que necesariamente

$$\alpha_1 = \ldots = \alpha_n = 0.$$

Estos números son las coordenadas de v en la base $\mathcal{B}$, por lo tanto $v = O$. □

Parte de los resultados contenidos en la proposición anterior pueden refinarse, tal como se sugiere en el enunciado de nuestro próximo ejercicio.

Ejercicio 4.121. Mostrar que si existe una base $\mathcal{B}$ del espacio $\mathbb{V}$ tal que $T(\mathcal{B})$ es una base de $\mathbb{W}$ entonces T es un isomorfismo.

Ejercicio 4.122. . Mostrar que si dos espacios $\mathbb{V}$ y $\mathbb{W}$ son isomorfos, y uno de ellos es el espacio trivial que sólo está formado por el vector nulo, entonces el otro también es un espacio trivial que sólo contiene al vector nulo.

Una consecuencia de la proposición 4.39 es que dos espacios de dimensión finita isomorfos tienen igual dimensión, porque el isomorfismo transforma una base de uno cualquiera de ellos en una base del otro. Cuando no hay bases esta observación también es cierta, en virtud del ejercicio 4.122. Al tener en cuenta además la información que surge del ejemplo 4.8.12 obtenemos el siguiente teorema acerca de los isomorfismos entre $\mathbb{K}$-espacios de dimensión finita.

Teorema 4.2. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos espacios vectoriales de dimensión finita sobre un mismo cuerpo $\mathbb{K}$. Entonces $\mathbb{V}$ y $\mathbb{W}$ son isomorfos si y sólo si tienen la misma dimensión.*

Ejercicio 4.123. Mostrar que el espacio de transformaciones lineales $\mathbb{L}(\mathbb{V}, \mathbb{W})$ entre dos espacios de dimensión finita es isomorfo al de matrices de tamaño $\dim \mathbb{W} \times \dim \mathbb{V}$, construyendo un isomorfismo explícito entre ambos. Concluir que

$$\dim(\mathbb{L}(\mathbb{V}, \mathbb{W})) = \dim(\mathbb{V}) \dim(\mathbb{W}).$$

4.8.3. Coordenadas, núcleo e imagen

Estudiaremos aquí las relaciones entre el núcleo y la imagen de una transformación lineal y el núcleo y la imagen de la matriz asociada a dicha transformación lineal respecto a bases en su dominio y codominio. El principal resultado de la sección es que al tomar coordenadas en el dominio el núcleo de la transformación queda identificado con el núcleo de la matriz asociada, y al hacer lo propio en el codominio, las coordenadas establecen un isomorfismo entre la imagen de la transformación y la de su matriz asociada. Damos a continuación un enunciado preciso de estas afirmaciones.

Proposición 4.40. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$, con dimensiones n y m respectivamente. Sean $\mathcal{A}$ y $\mathcal{B}$ bases de $\mathbb{V}$ y $\mathbb{W}$ respectivamente, y*

$$T : \mathbb{V} \to \mathbb{W}$$

una transformación lineal. Con ${}_{\mathcal{B}}T_{\mathcal{A}}$ indicamos la matriz asociada con T en las bases A y B. Entonces

1. *Un vector $v \in \mathbb{V}$ pertenece a $\ker(T)$ si y sólo si sus coordenadas $(v)_{\mathcal{A}}$ pertenecen al núcleo de la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$.*

2. *Un vector $w \in \mathbb{W}$ pertenece a $\operatorname{im}(T)$ si y sólo si sus coordenadas $(w)_{\mathcal{B}}$ pertenecen a la imagen, o espacio de columnas, de la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$.*

3. La aplicación

$$\text{coord}_{\mathcal{A}} : \ker(T) \to \ker\left({}_{\mathcal{B}}T_{\mathcal{A}}\right) \tag{4.38}$$

es un isomorfismo.

4. $\dim(\ker(T)) = \dim\left(\ker\left({}_{\mathcal{B}}T_{\mathcal{A}}\right)\right)$*.*

5. La aplicación

$$\text{coord}_{\mathcal{B}} : \text{im}(T) \to \text{im}\left({}_{\mathcal{B}}T_{\mathcal{A}}\right)$$

es un isomorfismo.

6. $\dim(\text{im}(T)) = \dim\left(\text{im}\left({}_{\mathcal{B}}T_{\mathcal{A}}\right)\right)$*.*

Prueba. Iremos probando las distintas partes en el mismo orden con el que aparecen en el enunciado.

La parte 1 se sigue de una cadena de equivalencias relativamente fácil de seguir. La incluimos a continuación, sin palabras:

$$v \in \mathbb{V} \Leftrightarrow Tv = O \Leftrightarrow (Tv)_{\mathcal{B}} = O \Leftrightarrow_{\mathcal{B}} (T)_{\mathcal{A}}(v)_{\mathcal{A}} = O \Leftrightarrow (v)_{\mathcal{A}} \in \ker\left({}_{\mathcal{B}}(T)_{\mathcal{A}}\right).$$

Para la parte 2 comencemos por observar que la igualdad

$$Tv = w \tag{4.39}$$

es equivalente a

$${}_{\mathcal{B}}(T)_{\mathcal{A}}(v)_{\mathcal{A}} = (w)_{\mathcal{B}}. \tag{4.40}$$

Supongamos que un vector $w \in \mathbb{W}$ pertenece a $\text{im}(T)$, entonces se satisface (4.39), por lo tanto también (4.40), lo que implica

$$(w)_{\mathcal{B}} \in \text{im}\left({}_{\mathcal{B}}(T)_{\mathcal{A}}\right).$$

Recíprocamente, supongamos que se satisface esta última inclusión. Entonces existe $X \in \mathbb{K}^n$ tal que

$${}_{\mathcal{B}}(T)_{\mathcal{A}}X = (w)_{\mathcal{B}}.$$

Pero X es necesariamente el vector de coordenadas de algún vector $v \in \mathbb{V}$. Para ese vector se satisface (4.40), que implica (4.39). Por lo tanto w es la imagen de v por T, y está en la imagen de T.

Para mostrar la parte 3 comencemos por observar que cualquier vector X en $\ker\left({}_{\mathcal{B}}T_{\mathcal{A}}\right)$ es el vector de coordenadas de algún vector v en $\mathbb{V}$. Por lo tanto

$$X = (v)_{\mathcal{A}}.$$

La parte 1 implica que $v \in \ker(T)$, por lo tanto el mapa (4.38) es sobreyectivo. La inyectividad y la linealidad son directas, porque se trata de la restricción de la transformación lineal e inyectiva $\mathrm{coord}_{\mathcal{A}}$ a un subespacio. Como (4.38) es una transformación lineal biyectiva, es un isomorfismo entre su dominio y su codominio.

La parte 4 es una consecuencia directa de la parte 3 y el teorema 4.2, acerca de las dimensiones de espacios vectoriales isomorfos.

Ejercicio 4.124. Demostrar las partes 5 y 6 de la proposición 4.40. $\square$

Ejemplo 4.8.13. Haremos ahora un ejemplo donde trabajaremos con bases que no son las canónicas. Consideramos una transformación lineal

$$T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^3 \quad \text{tal que} \quad {}_{\mathcal{B}}T_{\mathcal{A}} = \begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 \\ 1 & 1 & 3 & -3 \end{pmatrix}$$

donde

$$\mathcal{A} = \left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right),$$

$$\mathcal{B} = ((1,1,1),(0,1,1),(0,0,1)).$$

Queremos hallar una base del núcleo de T. Primero calculamos una base del núcleo de la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$. Sea $(x,y,z,t) \in \ker({}_{\mathcal{B}}T_{\mathcal{A}})$. Entonces debemos resolver el sistema

$$\left(\begin{array}{cccc|c} 1 & -1 & 1 & 1 & 0 \\ 1 & 0 & 2 & -1 & 0 \\ 1 & 1 & 3 & -3 & 0 \end{array} \right)$$

Escalerizando obtenemos

$$\left(\begin{array}{cccc|c} 1 & -1 & 1 & 1 & 0 \\ 0 & 1 & 1 & -2 & 0 \\ 0 & 2 & 2 & -4 & 0 \end{array} \right) \Longrightarrow \left(\begin{array}{cccc|c} 1 & -1 & 1 & 1 & 0 \\ 0 & 1 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right),$$

por tanto

$$\begin{cases} x - y + z + t = 0 \\ y + z + t = 0 \end{cases} \Longrightarrow \begin{cases} x = -2z - 2t \\ y = -z - t \end{cases}$$

Así los vectores del núcleo de la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$ son de la forma

$$(-2z - 2t, z + t, z, t) = z(-2,1,1,0) + t(-2,1,0,1).$$

Entonces
$$((-2,1,1,0),(-2,1,0,1))$$
es una base del $\ker({}_{\mathcal{B}}T_{\mathcal{A}})$.

Luego debemos aplicar el inverso de coordenadas, es decir obtenemos una base del núcleo de T haciendo
$$\operatorname{coord}_{\mathcal{A}}^{-1}(-2,1,1,0), \quad \operatorname{coord}_{\mathcal{A}}^{-1}(-2,1,0,1)$$
por tanto
$$\left(\begin{pmatrix} -1 & -3 \\ -3 & -2 \end{pmatrix}, \begin{pmatrix} -1 & -2 \\ -3 & -2 \end{pmatrix} \right)$$
es una base del núcleo de T.

Para terminar con este ejemplo calculemos la imagen de T. Sabemos que una base de $\operatorname{im}({}_{\mathcal{B}}T_{\mathcal{A}})$ es
$$((1,1,1),(-1,0,1)),$$
las primeras dos columnas de la matriz. Ahora, deshaciendo coordenadas en la base $\mathcal{B}$ obtenemos una base de la imagen de T
$$(\operatorname{coord}_{\mathcal{B}}^{-1}(1,1,1), \operatorname{coord}_{\mathcal{B}}^{-1}(-1,0,1)) = ((1,2,3),(-1,-1,0)).$$
♣

Ejemplo 4.8.14. Hallemos una base del núcleo y una base de la imagen de la transformación
$$T: M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R}) \text{ tal que } {}_{\mathcal{B}}T_{\mathcal{A}} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 1 & 1 \\ 3 & 2 & 2 & -3 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$
donde
$$\mathcal{A} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right),$$
$$\mathcal{B} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right).$$
En primer lugar hallemos el núcleo de la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$. Para eso consideramos un vector $X = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. Ese vector está en el núcleo de ${}_{\mathcal{B}}T_{\mathcal{A}}$ si

se verifica que ${}_{\mathcal{B}}T_{\mathcal{A}} \cdot X = O$, lo que es equivalente a resolver el sistema cuya matriz ampliada es

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ -1 & 1 & 1 & 1 & 0 \\ 3 & 2 & 2 & -3 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right).$$

Escalerizando obtenemos

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right) \Longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Por tanto los vectores del núcleo de la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$ son los que verifican

$$\left\{ \begin{array}{l} x_1 - x_4 = 0 \\ x_2 + x_3 = 0 \end{array} \right. ,$$

es decir $\ker({}_{\mathcal{B}}T_{\mathcal{A}}) = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 - x_4 = 0, x_2 + x_3 = 0\}$. Los vectores del núcleo son de la forma $(x_4, -x_3, x_3, x_4)$ con $x_3, x_4 \in \mathbb{R}$, lo que nos da una base de $\ker({}_{\mathcal{B}}T_{\mathcal{A}})$:

$$((1, 0, 0, 1), (0, -1, 1, 0)).$$

Ahora sabemos que estos vectores son las coordenadas en la base $\mathcal{A}$ de una base del núcleo de T, deshaciendo coordenadas encontramos que

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right)$$

es una base de $\ker(T)$.

Como ya tenemos ${}_{\mathcal{B}}T_{\mathcal{A}}$ escalerizada, sabemos que una base de la imagen está formada por las dos primeras columnas de la matriz, esto nos permite concluir que una base de la imagen es

$$((1, -1, 3, 0), (0, 1, 2, 1)).$$

Ahora, si deshacemos coordenadas en la base $\mathcal{B}$ obtenemos una base de $\mathrm{im}(T)$:

$$\left(\begin{pmatrix} 4 & -1 \\ -1 & -2 \end{pmatrix}, \begin{pmatrix} 2 & 2 \\ 0 & -2 \end{pmatrix} \right)$$

♣

Ejercicio 4.125.

1. Se considera la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que
$$T(x, y, z) = (x + 2y, y - z, x + 2z).$$
   a) Verificar que el vector $v = (3, 0, 3)$ pertenece a $\operatorname{im}(T)$.
   b) Verificar que el vector $v = (2, -1, -1)$ pertenece a $\ker(T)$.
2. Se considera la transformación lineal $T : M^{2\times 2}(\mathbb{R}) \to M^{2\times 2}(\mathbb{R})$ tal que
$$T(M) = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} M.$$
   a) Verificar que
$$M = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$
   pertenece a $\ker(T)$.
   b) Verificar que
$$M = \begin{pmatrix} 1 & 0 \\ -2 & 0 \end{pmatrix}$$
   pertenece a $\operatorname{im}(T)$.

Ejercicio 4.126. Use la definición para hallar el núcleo y la imagen de las siguientes transformaciones lineales:

1. $T : \mathbb{R}_3[x] \to \mathbb{R}^2$ tal que $T(p) = (p(1) + p(-1), p(0))$;
2. $T : M^{2\times 2}(\mathbb{R}) \to \mathbb{R}$ tal que $T(A) = \operatorname{traza}(A)$;
3. $T : \mathbb{C}^2 \to \mathbb{C}^2$ tal que $T(z, w) = (iz + w, -2z + 2iw)$.

Ejercicio 4.127. Hallar una base del núcleo y la imagen de las siguientes transformaciones lineales:

1. $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que
$${}_B(T)_B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$
   donde $B = ((1, 1, 0), (0, 2, 0), (2, 0, -1))$.
2. $T : \mathbb{R}_2[x] \to M^{2\times 2}(\mathbb{R})$ tal que
$${}_B(T)_C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 1 & 2 \end{pmatrix},$$
   donde
$$B = \left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right)$$
   y C es la base canónica de $\mathbb{R}_2[x]$.

Ejercicio 4.128. Hallar el núcleo y la imagen de las transformaciones P y Q con las que trabajamos en los ejercicios 4.87.

Ejercicio 4.129.

1. Sea $P : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la transformación definida por la matriz

$$ {}_{\mathcal{C}}P_{\mathcal{C}} = 1/2 \begin{pmatrix} 1 & -1 & -1 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{pmatrix}, $$

donde $\mathcal{C}$ es la base canónica de $\mathbb{R}^3$.

a) Hallar núcleo e imagen de la transformación lineal.

b) Verificar que $P = P^2$.

c) Interpretar geométricamente.

2. Sea $P : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la transformación definida por la matriz

$$ {}_{\mathcal{C}}P_{\mathcal{C}} = 1/6 \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & 1 \\ -2 & 1 & 5 \end{pmatrix}, $$

donde $\mathcal{C}$ es la base canónica de $\mathbb{R}^3$.

a) Hallar núcleo e imagen de la transformación lineal.

b) Verificar que $P = P^2$.

c) Probar que $PX \cdot Y = X \cdot PY$.

d) Interpretar geométricamente.

Ejercicio 4.130. Se consideran las transformaciones $T : \mathbb{R}^3 \rightarrow M^{2\times 2}(\mathbb{R})$ tal que

$$ T(x, y, z) = \begin{pmatrix} x - y & y \\ y & y - z \end{pmatrix}, $$

y

$$ S : M^{2\times 2}(\mathbb{R}) \rightarrow \mathbb{R}_3[x] $$

definida por

$$ S(A)(x) = (1, x)A \begin{pmatrix} x \\ x^2 \end{pmatrix}. $$

Verificar que T y S son lineales, y hallar el núcleo y la imagen de T, S, y $S \circ T$.

Ejercicio 4.131. Para las transformaciones P, Q y S que aparecen en el ejercicio 4.113, página 4.113, calcular el núcleo y la imagen de P, Q, S, $I - S$ e $I + S$.

Ejercicio 4.132.

1. Decidir si las dos afirmaciones siguientes son verdaderas o falsas, dando una demostración o un contraejemplo según corresponda.

   a) Si $T : \mathbb{V} \to \mathbb{W}$, lineal, es tal que existe un conjunto $A = (v_1, ..., v_r) \subset \mathbb{V}$ linealmente independiente que cumple que $T(A) = (Tv_1, ..., Tv_r)$ es también linealmente independiente, entonces es inyectiva.

   b) Si $T : \mathbb{V} \to \mathbb{W}$, lineal es tal que existe una base $B = (v_1, ..., v_n)$ de $\mathbb{V}$, que cumple que $T(B) = (Tv_1, ..., Tv_n)$ es también linealmente independiente, entonces es inyectiva.

2. Pruebe utilizando lo visto en este ejercicio que $T : \mathbb{R}^2 \to \mathbb{R}^3$, definida por

$$T(x, y) = (x - y, x + 2y, y)$$

   es inyectiva.

Ejercicio 4.133. Consideremos el espacio

$$\mathbb{V} = [f_1, f_2, f_3],$$

donde

$$f_1(t) = \operatorname{sen}^2(t), \ f_2(t) = \cos^2(t), \ f_3(t) = 1, \quad t \in \mathbb{R},$$

y la transformación lineal

$$T : \mathbb{R}^3 \to \mathbb{V}$$

tal que

$$T(a, b, c) = af_1 + bf_2 + cf_3.$$

Hallar $\ker(T)$ e $\operatorname{im}(T)$, y probar que que T no es inyectiva, pero sí sobreyectiva.

Ejercicio 4.134.

1. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales con $\dim(\mathbb{V}) < \dim(\mathbb{W})$; $T : \mathbb{V} \to \mathbb{W}$ y $S : \mathbb{W} \to \mathbb{V}$, transformaciones lineales.

   a) Probar que T no es sobreyectiva.

   b) Probar que S no es inyectiva.

2. Sean U, $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales; $T : \mathbb{V} \to \mathbb{W}$ y $S : \mathbb{W} \to U$ transformaciones lineales.

   a) Probar que $\ker(T) \subset \ker(S \circ T)$

   b) Si $S \circ T$ es inyectiva, probar que T es inyectiva.

   c) Probar que $\operatorname{im}(S \circ T) \subset \operatorname{im}(S)$

   d) Si $S \circ T$ es sobreyectiva, probar que S es sobreyectiva.

3. Sean A una matriz $n \times m$ y B una matriz $m \times n$, con $n < m$. Probar que BA no es invertible.

Ejercicio 4.135. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que

$$_A(T)_B = \begin{pmatrix} 1 & 4 \\ 2 & 1 \end{pmatrix},$$

donde

$$B = ((1,1),(1,0)), \quad A = ((1,2),(2,-1)).$$

Probar que T es invertible y hallar una matriz asociada a T^{-1}. Especificar las bases correspondientes.

Ejercicio 4.136.

1. Sea $\mathbb{V}$ un espacios vectorial y $T : \mathbb{V} \to \mathbb{V}$ una transformación lineal.

   a) Probar que T es inyectiva si y sólo si T^n es inyectiva para todo $n \in \mathbb{N}$.

   b) Probar que T es sobreyectiva si y sólo si T^n es sobreyectiva para todo $n \in \mathbb{N}$.

2. Sea A una matriz $n \times n$.

   a) Si el rango de A es igual a n, probar que el rango de todas sus potencias A^q, con $q \in \mathbb{N}$, es igual a n.

   b) ¿Es cierto en general que el rango de las potencias de A coincide con el rango de A?

Ejercicio 4.137.

1. Sea $\mathbb{V}$ un espacio vectorial de dimensión finita, y $T : \mathbb{V} \to \mathbb{V}$ una transformación lineal no inyectiva. Probar que existen transformaciones lineales $S_i : \mathbb{V} \to \mathbb{V}$, $i = 1, 2$, no nulas, tales que $T \circ S_1 = S_2 \circ T = O$

2. Para la transformación

$$T(x, y, z) = (x + y, x + z, y - z),$$

   definida en $\mathbb{R}^3$, hallar S_1 y S_2 como en la parte 1.

3. Sea A una matriz $n \times n$ no invertible. Probar que existen una matrices B y C, no nulas, de las mismas dimensiones, tales que $CA = AB = 0$.

Ejercicio 4.138. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios de dimensión finita, y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Probar:

1. si T es inyectiva entonces existe una aplicación lineal $S : \mathbb{W} \to \mathbb{V}$ tal que $S \circ T = Id$.

2. Para la transformación $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$T(x, y, z) = (x + y + z, x + y, y + z, x + z)$$

hallar una aplicación S como se especifica en la parte 1.

3. si T es sobreyectiva entonces existe una aplicación lineal $S : \mathbb{W} \to \mathbb{V}$ tal que $T \circ S = Id$.

4. Para la transformación $T : \mathbb{R}^4 \to \mathbb{R}^3$ definida por

$$T(x, y, z, t) = (x + y + t, x + y + z, x + z + t)$$

hallar una aplicación S como se especifica en la parte 3.

4.8.4. Para tener presente

- La sobreyectividad e inyectividad de las transformaciones lineales quedan caracterizadas por dos subespacios: el núcleo y la imagen de la transformación.

- Los resultados acerca de transformaciones lineales sobreyectivas e inyectivas pueden verse como versiones abstractas, generales, de los resultados acerca de la compatibilidad y determinación de los sistemas de ecuaciones lineales.

- Las imágenes y la preimágenes de subespacios por transformaciones lineales son subespacios. Las transformaciones respetan la estructura lineal.

- Dos espacios isomorfos son indistinguibles desde el punto de vista de su estructura lineal. En particular, el espacio $\mathbb{K}^n$ es un modelo para todos los $\mathbb{K}$-espacios de dimensión n.

- Una vez fijada una base en un espacio vectorial, la transformación que a cada vector le asocia las coordenadas respecto a esa base es un isomorfismo. Este isomorfismo provee una representación fiel de todos los vectores del espacio, y permite calcular con las coordenadas en vez de los vectores. En particular, recordemos que

  - una familia es linealmente independiente si y sólo si la familia de sus coordenadas respecto a cualquier base es linealmente independiente;
  - un vector es una combinación lineal de una familia si y sólo si sus coordenadas son una combinación lineal de las coordenadas de los vectores de la familia.

4.9. El teorema de las dimensiones

El objetivo de esta sección es demostrar el teorema de las dimensiones para transformaciones lineales. Es un resultado central de la teoría de transformaciones lineales entre espacios de dimensión finita, que generaliza el teorema de las dimensiones para matrices, teorema 1.1, página 140, que asegura que la suma de las dimensiones del espacio de columnas, o imagen, de una matriz, y de su núcleo, es igual a la cantidad de columnas de la matriz. Cuando consideramos la matriz como la representación de una transformación lineal definida sobre $\mathbb{K}^n$ entonces el número de columnas es igual a la dimensión del dominio.

Una versión abstracta de este resultado, válida para cualquier espacio de dimensión finita, es el principal objetivo de esta sección.

4.9.1. Enunciado y demostración del teorema

El enunciado general del teorema es el siguiente.

Teorema 4.3 (Teorema de las dimensiones). *Sean $\mathbb{V}$ y $\mathbb{W}$ dos espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$, y*

$$T : \mathbb{V} \to \mathbb{W}$$

una transformación lineal. Supongamos que $\mathbb{V}$ es un espacio de dimensión finita. Entonces

$$\dim(\mathbb{V}) = \dim(\ker(T)) + \dim(\operatorname{im}(T)). \tag{4.41}$$

Vamos a dar varias demostraciones del teorema. Algunas de ellas excluyen los casos triviales en que $\mathbb{V}$ es es el espacio trivial $\{O\}$, que sólo consiste del vector nulo. Dejamos este caso como ejercicio para el lector.

Ejercicio 4.139. Mostrar que el teorema de las dimensiones es cierto en el caso en que el dominio $\mathbb{V}$ de T sea el espacio trivial $\{O\}$.

La primera de ellas reduce el teorema 4.3 al teorema de las dimensiones para matrices, por la vía de considerar la matriz asociada con T respecto a bases de $\mathbb{V}$ y $\mathbb{W}$. Presentaremos la demostración agregando la hipótesis adicional, innecesaria, de que $\mathbb{W}$ también tenga dimensión finita, pero propondremos al lector el ejercicio de completar el argumento para conseguir una demostración del teorema 4.3 en toda su generalidad.

PRIMERA PRUEBA DEL TEOREMA 4.3. Supongamos que la dimensión de $\mathbb{V}$ es n, y que también $\mathbb{W}$ tiene dimensión finita m. Consideremos bases $\mathcal{A}$ y $\mathcal{B}$ de $\mathbb{V}$

y $\mathbb{W}$ respectivamente. Entonces T tendrá una matriz $m \times n$ asociada en estas bases, a la que llamaremos ${}_{\mathcal{A}}T_{\mathcal{B}}$.

El teorema de las dimensiones para matrices nos dice

$$n = \dim\left(\ker\left({}_{\mathcal{A}}T_{\mathcal{B}}\right)\right) + \dim\left(\operatorname{im}\left({}_{\mathcal{A}}T_{\mathcal{B}}\right)\right). \tag{4.42}$$

Pero sabemos, por la proposición 4.40, que

$$\begin{aligned} \dim\left(\ker(T)\right) &= \dim\left(\ker\left({}_{\mathcal{B}}T_{\mathcal{A}}\right)\right), \\ \dim\left(\operatorname{im}(T)\right) &= \dim\left(\operatorname{im}\left({}_{\mathcal{B}}T_{\mathcal{A}}\right)\right). \end{aligned}$$

Al emplear estas fórmulas en (4.42) para sustituir las dimensiones de los subespacios asociados con la matriz ${}_{\mathcal{B}}T_{\mathcal{A}}$ por las dimensiones de los subespacios asociados con T, obtenemos (4.41), porque n es justamente la dimensión de $\mathbb{V}$.

Este argumento completa la demostración en el caso en que $\mathbb{W}$ tiene dimensión finita.

Ejercicio 4.140.

1. Mostrar que $\operatorname{im}(T)$ es un espacio de dimensión finita, incluso en el caso en que $\mathbb{W}$ es un espacio de dimensión infinita;
2. Considerar como codominio de T a $\operatorname{im}(T) \subset \mathbb{W}$, y aplicar el resultado para transformaciones lineales entre espacios de dimensión finita para completar la demostración del teorema 4.3. $\square$

Segunda prueba del teorema 4.3. Nos independizaremos en esta prueba del cálculo con coordenadas. La idea de la prueba consiste en tener en cuenta que sobre $\ker(T)$ la acción de T es trivial, por lo que buscaremos estudiar a T sobre el complemento de $\ker(T)$. Consideraremos el complemento en el sentido de la estructura lineal. Es decir, consideremos cualquier subespacio $\mathbb{S} \subset \mathbb{V}$ tal que

$$\mathbb{V} = \mathbb{S} \oplus \ker(T).$$

La existencia de al menos un subespacio que satisface esta condición está asegurada por la parte 1 del ejercicio 4.68.

La restricción de T a $\mathbb{S}$ es inyectiva. En efecto, si $s \in \mathbb{S}$ y $Ts = O$, entonces s también pertenece a $\ker(T)$, por lo tanto

$$s \in \mathbb{S} \cap \ker(T),$$

lo que implica $s = O$ porque la suma de los dos subespacios es directa.

Al ser la restricción de T a $\mathbb{S}$ inyectiva, la aplicación

$$T : \mathbb{S} \to T(\mathbb{S}), \tag{4.43}$$

en la que hemos considerado a $\mathbb{S}$ como dominio y a su imagen $T(\mathbb{S})$ como codominio es inyectiva. También es sobreyectiva, porque la definición de $T(\mathbb{S})$ asegura que sólo están en este subespacio los vectores de $\mathbb{W}$ que son la i-magen por T de algún vector de $\mathbb{S}$. En conclusión, la transformación (4.43) es biyectiva, por lo tanto un isomorfismo entre su dominio y codominio, de donde concluimos

$$\dim(\mathbb{S}) = \dim\left(T(\mathbb{S})\right). \tag{4.44}$$

Mostraremos ahora que $T(\mathbb{S})$ es en realidad toda la imagen de T. La razón es que el núcleo $\ker(T)$ no aporta nada para la construcción de la imagen. Veamos un argumento riguroso.

Es claro que los vectores de $T(\mathbb{S})$ están en $\mathrm{im}(T)$, porque son la imagen por T de vectores de $\mathbb{S} \subset \mathbb{V}$. Por otra parte, si $w \in \mathrm{im}(T)$ entonces existe $v \in \mathbb{V}$ tal que $Tv = w$. El vector v puede descomponerse como la suma

$$v = s + u,$$

con $s \in \mathbb{S}$ y $u \in \ker(T)$. Por lo tanto

$$w = Tv = T(s+u) = Ts + Tu = Ts + O = Ts \in T(\mathbb{S}).$$

En conclusión, como la imagen de T y $T(\mathbb{S})$ son el mismo espacio, concluimos, en virtud de (4.44), que

$$\dim(\mathbb{S}) = \dim(\mathrm{im}(T)).$$

Como el espacio $\mathbb{V}$ es suma directa de $\mathbb{S}$ y $\ker(T)$ tenemos además

$$\dim(\mathbb{V}) = \dim(\ker(T)) + \dim(\mathbb{S}).$$

Las dos últimas igualdades implican (4.41). □

En el correr de esta prueba hemos utilizado un caso particular del resultado que enunciamos en nuestro próximo ejercicio.

Ejercicio 4.141. Sea T una transformación lineal, y $\mathcal{A}$ y $\mathcal{B}$ dos conjuntos cualesquiera en su dominio. Mostrar que

$$T(\mathcal{A} + \mathcal{B}) = T(\mathcal{A}) + T(\mathcal{B}).$$

Nuestra última prueba del teorema usa ideas que, en algún sentido, están implícitas en la demostración anterior, pero las emplea sin recurrir a la teoría de la suma directa de subespacios, ni a la consideración de T como un isomorfismo definido en un complemento de $\ker(T)$.

La prueba consiste en tomar una base de $\ker(T)$ y extenderla a una base de todo el espacio $\mathbb{V}$. La imagen de esta base por T genera imagen de T, pero los vectores que están en $\ker(T)$ sólo aportan el vector nulo, y las imágenes de los vectores agregados son linealmente independientes. Por lo tanto forman una base de $\operatorname{im}(T)$. A continuación exponemos los detalles. Luego de completada la demostración estudiaremos su relación con la segunda prueba del teorema 4.3.

Tercera prueba del teorema 4.3. Los casos en que $\ker(T) = \{O\}$ y $\ker(T) = \mathbb{V}$ son especialmente sencillos, aunque un tanto excepcionales para el razonamiento que vamos a presentar, de modo que los dejamos como ejercicio para el lector. En cualquier otro caso $\ker(T)$ será un subespacio no trivial de $\mathbb{V}$, cuya dimensión d satisface $1 \leq d < n$.

Consideremos una base

$$\mathcal{B} = (e_1, \ldots, e_d)$$

de $\ker(T)$ y extendámosla a una base

$$\mathcal{A} = (e_1, \ldots, e_d, e_{d+1}, \ldots, e_n)$$

de $\mathbb{V}$. Los transformados

$$(Te_1, \ldots, Te_d, Te_{d+1}, \ldots, Te_n)$$

de los vectores de la base forman un generador de $\operatorname{im}(T)$. Los primeros d vectores de este generador son nulos, porque e_i, $i = 1, \ldots, d$, están en el núcleo de T. Por lo tanto

$$\mathcal{T} = (Te_{d+1}, \ldots, Te_n)$$

es un generador de $\operatorname{im}(T)$.

Mostraremos ahora que $\mathcal{T}$ es linealmente independiente, recurriendo a la definición de independencia lineal. Consideremos entonces una combinación lineal de $\mathcal{T}$, con constantes λ_i, $i = d + 1, \ldots, n$, tales que

$$\lambda_{d+1} Te_{d+1} + \ldots + \lambda_n Te_n = O. \tag{4.45}$$

Como T es lineal tenemos que

$$T(\lambda_{d+1} e_{d+1} + \ldots + \lambda_n e_n) = O,$$

lo que implica que

$$\lambda_{d+1} e_{d+1} + \ldots + \lambda_n e_n \in \ker(T).$$

Por lo tanto, este vector puede escribirse como una combinación de los vectores en la base $\mathcal{B}$ de $\ker(T)$. Existen constantes λ_i, $i = 1, \dots, d$, tales que

$$\lambda_{d+1}e_{d+1} + \ldots + \lambda_n e_n = \lambda_1 e_1 + \ldots + \lambda_d e_d.$$

Podemos reordenar esta igualdad, en la forma más familiar

$$\lambda_1 e_1 + \ldots + \lambda_d e_d - \lambda_{d+1}e_{d+1} - \ldots - \lambda_n e_n = O$$

La independencia lineal de la base $\mathcal{A}$ asegura que todos los coeficientes de esta combinación lineal son nulos. En particular, los coeficientes λ_i, $i = d+1, \dots, n$ que aparecen en la combinación lineal (4.45) son necesariamente nulos, lo que implica la independencia lineal de $\mathcal{T}$.

En resumen $\mathcal{T}$ es una base de $\operatorname{im}(T)$, y este subespacio tiene dimensión igual a $n - d$. La igualdad (4.41) es ahora obvia.

Ejercicio 4.142. Completar esta tercera demostración, estudiando los casos en que $\ker(T)$ es uno de los dos subespacios triviales de $\mathbb{V}$ □

Observación 4.9.1. Las dos últimas demostraciones tienen unos cuantos puntos en común. Los vectores e_i, $i = d + 1, \dots, n$, de la tercera demostrácion generan un subespacio $\mathbb{S}$ que es un complemento de $\ker(T)$, en el sentido de la segunda demostración.

El hecho de que T restringida a cualquier complemento $\mathbb{S}$ de $\ker(T)$ sea inyectiva implica que transforma un subconjunto linealmente independiente de $\mathbb{S}$ en un subconjunto linealmente independiente de $\operatorname{im}(T)$.

La demostración de que Te_i, $i = d + 1, \dots, n$, es un generador de $\operatorname{im}(T)$ recurre a a la misma idea que la demostración de que $T(\mathbb{S})$ es toda la imagen de T: en ambos casos se utiliza que cualquier cosa que esté en el núcleo va a parar al vector nulo O de $\mathbb{W}$ cuando se aplica T.

Vemos entonces que estas dos demostraciones usan ideas parecidas. La segunda recurre a una mayor elaboración teórica previa, porque usa las nociones de suma directa de subespacios, complementos e isomorfismos, y produce una prueba bastante breve, cuya idea central puede resumirse en la frase: cualquier complemento del $\ker(T)$ es isomorfo a $\operatorname{im}(T)$.

La tercera demostración recurre más directamente a las definiciones de combinación lineal, independencia lineal, generador. Su idea central es algo más larga de resumir: al extender una base de $\ker(T)$ a una base de $\mathbb{V}$, los transformados de los vectores que se agregaron en la extensión forman una base de $\operatorname{im}(T)$.

Ambas pruebas del teorema 4.3 descansan sobre el importante hecho de que cualquier conjunto linealmente independiente en $\mathbb{V}$ puede extenderse a una base del espacio $\mathbb{V}$. Este resultado se utiliza para extender una base de un subespacio cualquiera a una base de todo el espacio, y para mostrar que cualquier subespacio de $\mathbb{V}$ tiene un subespacios complementario en el sentido de la suma directa de subespacios. ♠

4.9.2. Consecuencias del teorema de las dimensiones

El teorema de las dimensiones da bastante información acerca de las transformaciones lineales entre dos $\mathbb{K}$-espacios de dimensión finita. Enunciamos a continuación una proposición que recoge unos cuantos resultados que se derivan directamente de él.

Proposición 4.41. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita, con dimensiones n y m respectivamente. En lo que sigue, T siempre designará una transformación lineal $T : \mathbb{V} \to \mathbb{W}$.*

1. *Si $n > m$ entonces T no puede ser inyectiva;*
2. *si $n < m$ entonces T no puede ser sobreyectiva;*
3. *si $n = m$ entonces son equivalentes:*
   a) *T es inyectiva;*
   b) *T es sobreyectiva;*
   c) *T es un isomorfismo.*

Solo demostraremos la parte 3 de la proposición. El resto quedará como ejercicio para el lector.

Trabajemos bajo la hipótesis $n = m$. Todo se reduce a traducir afirmaciones sobre inyectividad y sobreyectividad en subespacios, y a contar dimensiones.

Comenzaremos por probar que las condiciones $3a$ y $3b$ son equivalentes entre sí, y que cualquiera de ellas implica $3c$. La transformación T es inyectiva si y sólo si $\ker(T) = O$, y esta condición es equivalente a que la dimensión del núcleo de T sea nula. El teorema de las dimensiones implica

$$\dim(\text{im}(T)) = n - \dim(\ker(T)),$$

por lo que la dimensión del núcleo es nula si y sólo si la de la imagen es $n = m$. Pero la dimensión de la imagen es m si y sólo si la imagen es todo $\mathbb{W}$, lo que es equivalente a que T sea sobreyectiva. Hemos concluido entonces que si T es inyectiva es sobreyectiva, y que si es sobreyectiva es inyectiva. Por lo tanto, en ambos casos es biyectiva y resulta ser un isomorfismo.

Por otra parte, si T satisface $3c$ es biyectiva, y por lo tanto inyectiva y sobreyectiva.

Ejercicio 4.143. Completar la demostración de la proposición 4.41.

Corolario 4.42. *Si $\mathbb{V}$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, y*

$$T : \mathbb{V} \to \mathbb{V}$$

es una transformación lineal, entonces T es inyectiva si y sólo si es sobreyectiva, y, en tal caso, es un isomorfismo de $\mathbb{V}$ en sí mismo.

Ejemplo 4.9.2. La conclusión del corolario no se aplica a espacios de dimensión infinita. Mostramos dos aplicaciones del espacio de polinomios $\mathbb{R}_n[x]$ en sí mismo, que muestran este hecho. El primero es el operador de diferenciación

$$D : p(x) \mapsto p'(x),$$

que a cada polinomio le asocia su derivada. No es inyectivo, porque envía todos los polinomios constantes al polinomio nulo. En otras palabras, el núcleo de D es igual a las constantes. Pero es sobreyectivo. Para mostrarlo consideremos un polinomio genérico

$$p(x) = a_0 + a_1 x_1 + \dots a_n x^n. \tag{4.46}$$

Tenemos que

$$p(x) = (Dq)(x),$$

con

$$q(x) = a_0 x + \frac{a_1}{2} x + \dots \frac{a_n}{n+1} x^{n+1}. \tag{4.47}$$

Por otra parte, la transformación que a cada polinomio $p(x)$ como en (4.46) le asocia la integral indefinida del polinomio que aparece en la fórmula (4.47) es una transformación lineal inyectiva, pero no sobreyectiva.

Ejercicio 4.144. Verificar esta última afirmación. ♠

Ejercicio 4.145. Sea $\mathbb{V}$ un espacio de dimensión finita, y $T : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Mostrar que

$$\mathbb{V} = \operatorname{im}(T) \oplus \ker(T)$$

si y sólo si

$$\operatorname{im}(T) \cap \ker(T) = \{O\}.$$

Dar un ejemplo de una transformación lineal para la que se satisfagan ambas condiciones, y una para la que no se satisfagan.

Ejercicio 4.146. Sea $T : M^{2\times 2}(\mathbb{R}) \to M^{2\times 2}(\mathbb{R})$ tal que

$$T(X) = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} X$$

1. Hallar $\ker(T)$ e $\operatorname{im}(T)$.
2. Verificar que $\dim(M^{2\times 2}(\mathbb{R})) = \dim(\ker(T)) + \dim(\operatorname{im}(T))$.

Ejercicio 4.147.

1. Sean $T : \mathbb{V} \to \mathbb{R}$ y $S : \mathbb{V} \to \mathbb{R}$ transformaciones lineales. Sabiendo que $\dim(\mathbb{V}) = n$,
   a) probar que $\dim(\ker(T)) = n$ ó $n - 1$.
   b) probar que $\ker(T) = \ker(S)$ si y sólo si existe un número real $\alpha \neq 0$ tal que $T = \alpha S$.
2. Sea $T : \mathbb{R}^3 \to \mathbb{R}$ tal que $T(x, y, z) = x + y + z$. Hallar $S : \mathbb{R}^3 \to \mathbb{R}$, lineal, sabiendo que $\ker(T) = \ker(S)$ y $S(1, 0, 0) = 2$.

Ejercicio 4.148. Sean $T : \mathbb{V} \to \mathbb{W}$ y $S : \mathbb{V} \to \mathbb{W}$ dos transformaciones lineales tales que $\operatorname{im}(T) = \operatorname{im}(S)$.

1. Probar que $\dim(\ker(T)) = \dim(\ker(S))$.
2. ¿Es cierto que $T = S$? ¿Y que $\ker(T) = \ker(S)$? Justifique su respuesta.

Ejercicio 4.149. Sea $T : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Probar que si

$$\ker(T) = \ker(T^2),$$

entonces

$$\ker(T) \oplus \operatorname{im}(T) = \mathbb{V}.$$

En este texto comenzamos el recorrido hasta la introducción de los espacios vectoriales analizando la estructura lineal que emerge al considerar los sistemas de ecuaciones lineales. La teoría de los espacios vectoriales de dimensión finita es una extensión de la teoría de los sistemas de ecuaciones lineales, incluso puede verse como una versión abstracta de ella. Conviene entonces volver al punto de partida, y ver qué nos dicen nuestros resultados cuando los aplicamos en el contexto de los sistemas lineales de ecuaciones.

Ejercicio 4.150. Reformular todos los resultados de la proposición 4.41 y su corolario en la forma de afirmaciones sobre sistemas de ecuaciones lineales $AX = B$.

4.9.3. Para tener presente

- El teorema de las dimensiones es un resultado central acerca de las transformaciones lineales definidas en espacios vectoriales de dimensión finita, en el que se resumen muchas de las propiedades de los espacios vectoriales y las transformaciones lineales.

- El teorema de las dimensiones generaliza el resultado que afirma que la suma de la dimensión del núcleo de una matriz, más la dimensión de su imagen, es igual al número de columnas.
- No puede existir una transformación lineal inyectiva que tenga como codominio un espacio vectorial de menor dimensión que su dominio.
- No puede existir una transformación lineal sobreyectiva que tenga como codominio un espacio vectorial de mayor dimensión que su dominio.
- Una interpretación posible del teorema de las dimensiones es que cualquier transformación lineal establece un isomorfismo entre un complemento cualquiera de su núcleo, y su imagen.

Epílogo

Ahora que hemos dejado al lector mas o menos en el mismo sitio en el que todo empezó, nos despedimos de él, y le agradecemos su compañía.
Es hora de dejar descansar juntos los dos volúmenes que forman este texto, y apreciar el hermoso conjunto **tricolor** que componen, en un merecido homenaje al **glorioso** y **decano**.
¡Chau, hasta la vista! ¡Buena suerte!

Apéndice A

A.1. El alfabeto griego

En los textos de Matemática las letras griegas forman parte de la notación habitual. No reconocerlas, o designarlas incorrectamente, tiende a generar una confusión que dificulta la comunicación oral o escrita. Incluimos a continuación un alfabeto griego[1].

α	A	alfa
β	B	beta
γ	Γ	gama ó gamma
δ	Δ	delta
ϵ	E	épsilon
ζ	Z	dseda
η	H	eta
θ	Θ	zeta
ι	I	iota
κ	K	cappa ó kappa
λ	Λ	lambda
μ	M	mi
ν	N	ni
ξ	Ξ	xi
o	O	ómicron
π	Π	pi
ρ	P	ro
σ	Σ	sigma
τ	T	tau
υ	Υ	ípsilon
φ	Φ	fi
χ	X	ji
ψ	Ψ	psi
ω	Ω	omega

[1] Tal como aparecen en la versión accesible a través de Internel del diccionario de la Real Academia Española, en el sitio `http://www.rae.es`. La consulta fue realizada el 17 de febrero de 2005

A.2. Cuerpos

A lo largo del texto trabajaremos con conjuntos numéricos que tienen las mismas propiedades algebraicas que el conjunto $\mathbb{Q}$ de los números racionales, y el conjunto $\mathbb{R}$ de los números reales. Estas propiedades corresponden a la estructura algebraica llamada *cuerpo*, cuya definición aparece a continuación.

Definición A.1. *La terna* $(\mathbb{K}, +, \times)$ *formada por un conjunto* $\mathbb{K}$ *sobre el que están definidas dos operaciones binarias*

$$+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \quad \times : \mathbb{K} \times \mathbb{K} \to \mathbb{K},$$

a las que llamaremos suma y producto respectivamente, es un ***cuerpo*** *si la suma tiene las siguientes propiedades*

- CONMUTATIVA*:* $x + y = y + x \ \forall x, y$ *en* $\mathbb{K}$
- ASOCIATIVA*:* $x + (y + z) = (x + y) + z \ \forall x, y, z$ *en* $\mathbb{K}$
- EXISTENCIA DE NEUTRO*:* $x + 0 = 0 + x = x \ \forall x$ *en* $\mathbb{K}$
- EXISTENCIA DE OPUESTO*:* $x + (-x) = (-x) + x = 0 \ \forall x$ *en* $\mathbb{K}$*,*

el producto las propiedades

- CONMUTATIVA*:* $x \times y = y \times x \ \forall x, y$ *en* $\mathbb{K}$
- ASOCIATIVA*:* $x \times (y \times z) = (x \times y) \times z \ \forall x, y, z$ *en* $\mathbb{K}$
- EXISTENCIA DE NEUTRO*:* $x \times 1 = 1 \times x = x \ \forall x$
- EXISTENCIA DE INVERSO PARA LOS ELEMENTOS NO NULOS*:* $\forall x \neq 0$ *en* $\mathbb{K}$ *existe* x^{-1} *tal que* $x \times x^{-1} = x^{-1} \times x = 1$

y la suma es DISTRIBUTIVA *frente al producto, en el sentido de que la igualdad*

$$(x + y) \times z = x \times z + y \times z$$

se satisface para $x, y, z \in \mathbb{K}$ *cualesquiera.*

Tenemos así que los conjuntos $\mathbb{Q}$ y $\mathbb{R}$, con las operaciones usuales de suma y producto, son cuerpos. El cuerpo $\mathbb{R}$ extiende al cuerpo $\mathbb{Q}$, en el sentido de que $\mathbb{Q} \subset \mathbb{R}$, y que las operaciones de suma y producto en $\mathbb{Q}$ no son más que las operaciones correspondientes de $\mathbb{R}$ restringidas a $\mathbb{Q}$. La diferencia esencial entre $\mathbb{Q}$ y $\mathbb{R}$ es que este segundo cuerpo es *completo* en el sentido que se discute en las páginas 118 a 121 de "Introducción al cálculo y al análisis matemático" de Courant - John. También se puede consultar "Calculus" volumen 1 de Tom Apóstol en las páginas 31 y 32.

Hay muchos cuerpos comprendidos entre $\mathbb{Q}$ y $\mathbb{R}$. A continuación mostramos uno de ellos.

Ejemplo A.2.1. El cuerpo $\mathbb{Q}[\sqrt{2}]$

El cuerpo $\mathbb{Q}[\sqrt{2}]$ está formado por todos los números de la forma

$$a + b\sqrt{2}, \quad a, b \in \mathbb{Q}, \tag{A.1}$$

donde las operaciones de suma y producto actúan de la siguiente manera:

$$\begin{gathered}(a_1 + b_1\sqrt{2}) + (a_2 + b_2\sqrt{2}) = (a_1 + a_2) + (b_1 + b_2)\sqrt{2},\\ (a_1 + b_1\sqrt{2}) \times (a_2 + b_2\sqrt{2}) = (a_1a_2 + 2b_1b_2) + (a_1b_2 + b_1a_2)\sqrt{2}.\end{gathered} \tag{A.2}$$

Hay varias maneras de pensar este conjunto $\mathbb{Q}[\sqrt{2}]$:

- si suponemos conocido el conjunto $\mathbb{R}$ de todos los números reales entonces se trata del subconjunto de $\mathbb{R}$ formado por los números de la forma (A.1);
- si no conocemos a los números reales y sólo tenemos los racionales a nuestra disposición el número $\sqrt{2}$ no tiene sentido, porque no existe ningún racional q tal que $q^2 = 2$. Podemos entonces introducir $\sqrt{2}$ como un símbolo que tiene la propiedad de que $\sqrt{2} \times \sqrt{2} = 2$, y definir las sumas y productos de expresiones del tipo $a + b\sqrt{2}$ por las fórmulas (A.2);
- una manera alternativa de hacer esta construcción es considerar parejas (a, b) de números racionales, y definir sobre estas parejas las operaciones

  $$\begin{gathered}(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2),\\ (a_1, b_1) \times (a_2, b_2) = (a_1a_2 + 2b_1b_2, a_1b_2 + b_1a_2).\end{gathered}$$

  Observemos que al operar con parejas que tienen ceros en la segunda componente los cálculos se reducen a operar con las primeras componentes con las reglas de suma y producto en $\mathbb{Q}$. Por lo tanto, las parejas con segunda componente 0 son esencialmente lo mismo que $\mathbb{Q}$ y podemos identificar cada $q \in \mathbb{Q}$ con la pareja $(q, 0)$. Por otra parte

  $$(0, 1) \times (0, 1) = (2, 0),$$

  por lo que $(0, 1)$ se comporta como una raíz cuadrada de 2.

Observación A.2.2. Notemos que no hace falta extender $\mathbb{Q}$ al conjunto de todos los números reales para tener una raíz cuadrada de 2 y resolver la ecuación $x^2 = 2$. Podemos adjuntarle a $\mathbb{Q}$ el número $\sqrt{2}$ por el procedimiento que acabamos de mostrar. El interés de pasar del cuerpo $\mathbb{Q}$ al cuerpo $\mathbb{R}$ es que $\mathbb{R}$ es completo.

En lo que sigue consideraremos a los números reales como un conjunto conocido, y trabajaremos con $\mathbb{Q}[\sqrt{2}]$ tratándolo como un subconjunto de $\mathbb{R}$.

Ejercicio A.1. Mostrar que si $a + b\sqrt{2} \neq 0$ entonces su inverso es

$$\frac{a - b\sqrt{2}}{a^2 + 2b^2}$$

que también es un elemento de $\mathbb{Q}[\sqrt{2}]$. Mostrar que $Q[\sqrt{2}]$, con las operaciones de suma y producto usuales, es un cuerpo. ♣

A.2.1. Los números complejos

Así como la ecuación $x^2 = 2$ no tiene ninguna solución en $\mathbb{Q}$, también es cierto que $x^2 = -1$ no tiene ninguna solución real. Vamos a obtener una interesantísima extensión del cuerpo de los números reales agregándoles una raíz cuadrada de -1, un símbolo i que tiene la propiedad de que

$$i^2 = -1.$$

Consideramos todas las posibles combinaciones

$$x + yi, \quad x, y \in \mathbb{R}, \tag{A.3}$$

y sobre ellas definimos las operaciones

$$\begin{aligned} (x_1 + y_1 i) + (x_2 + y_2 i) &= (x_1 + x_2) + (y_1 + y_2)i, \\ (x_1 + y_1 i) \times (x_2 + y_2 i) &= (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i. \end{aligned} \tag{A.4}$$

El conjunto de todas las combinaciones (A.3) es lo que se llama el conjunto $\mathbb{C}$ de los números complejos.

Si se desea, podemos pensar los números complejos $x + yi$ como parejas (x, y) de números reales, sobre las que hemos definido las operaciones

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2), \\ (x_1, y_1) \times (x_2, y_2) &= (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2). \end{aligned}$$

Este punto de vista es consistente con identificar cada número complejo con un punto del plano, o con un vector de $\mathbb{R}^2$.

Ejemplo A.2.3. Ahora vamos a calcular algunas operaciones con complejos usando la definición anterior.

- $(5-2i)+(-3+i)=(5-3)+(-2+1)i=2-i$,
- $(2+2i)\times(3+2i)=(2\cdot 3-2\cdot 2)+(2\cdot 2+2\cdot 3)i=2+10i$,
- $(3+2i)^2=(3+2i)\times(3+2i)=(3\cdot 3-2\cdot 2)+(3\cdot 2+3\cdot 2)i=5+12i$.

Es relativamente fácil determinar cuál tiene que ser el inverso de una pareja $a+ib$ no nula. El inverso, por ahora desconocido, debe ser un complejo $x+iy$ que satisfaga

$$1=(a+bi)\times(x+yi)=(ax-by)+(bx+ay)i,$$

por lo tanto x y y son una solución del sistema de ecuaciones lineales

$$\begin{aligned} ax-by &= 1,\\ bx+ay &= 0. \end{aligned}$$

Si $a\neq 0$ el sistema se escaleriza rápidamente, multiplicando la segunda ecuación por a, la primera por $-b$ y sumando. Obtenemos

$$\begin{aligned} ax-by &= 1,\\ (a^2+b^2)y &= -b. \end{aligned}$$

Despejamos

$$y=\frac{-b}{a^2+b^2},$$

y luego calculamos x sustituyendo el valor de y en la primera ecuación y despejando

$$x=\frac{a}{a^2+b^2}.$$

Si a es nulo entonces $y=-1/b$, $x=0$, que es la forma que toman las expresiones que acabamos de encontrar para x e y cuando $a=0$. Nuestros cálculos pueden resumirse en que el inverso de $a+ib$ es

$$(a+ib)^{-1}=\frac{a-ib}{a^2+b^2}.$$

Ejercicio A.2. Verificar que $\mathbb{C}$, con las operaciones de suma y producto que acabamos de definir, es un cuerpo.

Para un complejo $x + iy$ llamamos a x su *parte real*, y a y su *parte imaginaria*. Podemos identificar a los complejos con parte imaginaria nula con los números reales.

También podemos obtener una representación geométrica de los complejos identificando $x+yi$ con la pareja (x, y) de números reales, y esta pareja como la coordenadas de un punto del plano respecto a un sistema de ejes ortogonal. En esta representación los complejos de parte imaginaria nula quedan identificados con el eje horizontal, que puede a su vez pensarse como la recta real.

Es útil considerar el *conjugado* de cada número complejo $z = x + yi$, que se define como el complejo

$$\bar{z} = x - yi.$$

Si multiplicamos $z\bar{z}$ tenemos

$$z\bar{z} = (x + yi)(x - yi) = x^2 + y^2$$

Pasar de un complejo a su conjugado es equivalente a simetrizar respecto al eje real. El número $x^2 + y^2$ es igual al cuadrado de la distancia del punto de coordenadas (x, y) al origen. Su raíz cuadrada es lo que se llama el *módulo* del número complejo $z = x + yi$, y se indica con la notación

$$|z| = \sqrt{x^2 + y^2}.$$

Observemos que vale

$$z\bar{z} = |z|^2,$$

y también

$$z^{-1} = \frac{\bar{z}}{z^2},$$

lo que da una fórmula sencilla para recordar la expresión del inverso de un número complejo.

Hemos considerado hasta ahora los cuerpos $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$. No son estos los únicos cuerpos interesantes. De hecho, hay otros cuerpos mucho más atractivos y muchas personas prestan más atención a otros cuerpos que a los que acabamos de mostrar. Lo de atractivo es muy subjetivo, por supuesto. Por ejemplo, quienes trabajan en el procesamiento digital de la información suelen encontrar mucho más excitantes los cuerpos discretos, que sólo tienen una cantidad finita de “números”. Otras personas muestran una nítida preferencia por el cuerpo real, y otras por el cuerpo complejo. Cuestión de gustos e inclinaciones.

En cualquier caso, nos ocuparemos de discutir algunos cuerpos discretos en nuestra próxima sección.

A.2.2. Aritmética en $\mathbb{Z}$ módulo un entero p y cuerpos $\mathbb{Z}_p$

El cuerpo $\mathbb{Z}_2$

El primer cuerpo discreto que mostraremos es extremadamente simple. Es el cuerpo $\mathbb{Z}_2$ que está formado por los dos símbolos o números 0 y 1, es decir

$$\mathbb{Z}_2 = \{0, 1\},$$

con operaciones de suma y producto definidas por las tablas

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

Es fácil verificar por la simple inspección de las tablas que el 0 es el neutro para la suma, y el 1 la identidad para el producto. Se podría demostrar que $\mathbb{Z}_2$ es un cuerpo estudiando exhaustivamente todas las posibles combinaciones de operaciones para ver que se satisfacen las propiedades requeridas. No lo haremos porque presentaremos más adelante argumentos más generales y poderosos.

Notemos que los números 0 y 1 pueden representar los dos posibles valores de un bit, o cada uno de los valores *verdadero* y *falso*, o *encendido* y *apagado*. Es por eso que la estructura algebraica de $\mathbb{Z}_2$ tiene interés para las ciencias de la computación, la lógica, la ingeniería eléctrica, etcétera.

La aritmética de $\mathbb{Z}_2$ también reproduce aquella letanía escolar de

par más par, par
par más impar, impar
impar más impar, par

par por par, par
par por impar, impar
impar por impar, impar[2].

De hecho, si interpretamos 0 y 1 como *par* e *impar* respectivamente, las tablas de las operaciones de $\mathbb{Z}_2$ se leen

+	par	impar
par	par	impar
impar	impar	par

×	par	impar
par	par	par
impar	par	impar

[2]Claro está que una vez examinados la suma y el producto de un par y un impar no era necesario hacerlo para un impar y un par. En aquellas épocas escolares en que las matrices todavía no habían aparecido en el horizonte, nadie dudaba de que *el orden de los factores no altera el producto*.

Esto no es casual. El resto que arroja un número par al ser dividido entre 2 es 0, y un número impar arroja resto 1. Todo número natural, también todo número entero queda clasificado según su divisibilidad entre 2 en dos clases:

- los números pares, que son múltiplos de 2 más 0;
- y los números impares, que son múltiplos de 2 más 1;

Si uno opera con dos números enteros y sólo retiene del resultado el resto que resulta de dividir entre 2 se recupera la aritmética de $\mathbb{Z}_2$. Esta aritmética puede ser vista como una aritmética sobre dos clases de números:

- la clase de los números pares, que difieren de 0 en un múltiplo de 2;
- la de los números impares, que difieren de 1 en un múltiplo de 2.

Aritmética módulo un número entero

La construcción con la que cerrábamos el párrafo anterior no exige tomar como punto de partida el número 2 para identificar entre sí a todos los pares y a todos los impares. Puede hacerse usando cualquier número natural n, identificando entre sí a todos los enteros que difieren en un múltiplo de n, y definiendo, a partir de las operaciones usuales de suma y producto en $\mathbb{Z}$ una aritmética sobre las clases que así resultan. Estas clases se identifican naturalmente con los posibles restos de dividir entre n. A continuación describimos esta construcción y enunciamos sus principales propiedades.

Definición A.2. *Consideremos un número natural $n > 1$ Para $a, b \in \mathbb{Z}$ decimos que a es* ***congruente*** *con b* ***módulo n****, y escribimos*

$$a \equiv b \pmod{n}$$

si $a - b$ es un múltiplo de n.

Vale la siguiente proposición.

Proposición A.1. *La congruencia módulo n es una relación de equivalencia sobre $\mathbb{Z}$.*

Ejercicio A.3. Demostrar la proposición anterior.

Como una relación de equivalencia sobre un conjunto induce una partición de éste, para $n \geq 2$, la congruencia módulo n divide a $\mathbb{Z}$ en las n clases de equivalencia

$$\begin{array}{l}
[0] = \{\ldots, -2n, -n, 0, n, 2n, \ldots\} = \{0 + nx / x \in \mathbb{Z}\} \\
[1] = \{\ldots, -2n+1, -n+1, 1, n+1, 2n+1, \ldots\} = \{1 + nx / x \in \mathbb{Z}\} \\
[2] = \{\ldots, -2n+2, -n+2, 2, n+2, 2n+2, \ldots\} = \{2 + nx / x \in \mathbb{Z}\} \\
\vdots \\
[\text{n-1}] = \{\ldots, -n-1, -1, n-1, 2n-1, 3n-1, \ldots\} = \{n - 1 + nx / x \in \mathbb{Z}\}
\end{array}$$

Indicamos con $[i]$ a la clase de equivalencia de i. Notemos que en el caso $n = 2$ se obtiene la partición de $\mathbb{Z}$ en pares e impares.

Esta partición en clases de equivalencia tiene muy buen comportamiento respecto a las operaciones de suma y producto en $\mathbb{Z}$, en el siguiente sentido:

Proposición A.2. *Si $a \equiv b \pmod{n}$ y $c \equiv d \pmod{n}$ entonces*

$$a + c \equiv b + d \pmod{n}, \qquad ac \equiv bd \pmod{n}.$$

Ejercicio A.4. Demostrar esta proposición.

Podemos entonces definir operaciones de suma y producto entre las clases operando sobre representantes de las clases con las reglas de suma y producto en $\mathbb{Z}$. La clase de equivalencia del resultado no dependerá de los representantes elegidos.

Usamos la notación $\mathbb{Z}_n$ para denotar a las n clases que resultan de aplicar en $\mathbb{Z}$ la relación de congruencia módulo n, es decir

$$\mathbb{Z}_n = \{[0], [1], [2], \ldots, [n-1]\}.$$

La suma $+$ y el producto $\times$ en $\mathbb{Z}_n$ quedan definidos por las reglas

$$[a] + [b] = [a + b] \qquad [a] \times [b] = [a \times b]$$

Como es usual, muchas veces no utilizaremos el símbolo $\times$ para indicar el producto, que será indicado por la simple yuxtaposición ab de los símbolos a y b que se desea multiplicar.

Calcular en $\mathbb{Z}_n$ es sencillo. Sólo hay que seguir las reglas usuales, pero volver a 0 cada vez que se alcanza el valor de un múltiplo de n. Por ejemplo si $n = 7$, entonces $[2] + [6] = [2 + 6] = [8] = [1]$, y $[2] \cdot [6] = [12] = [5]$

Observemos que cada clase de equivalencia en $\mathbb{Z}_n$ tiene un representante canónico: el único elemento de la clase que está entre 0 e $i - 1$. Por esa razón también podemos representar a $\mathbb{Z}_n$ en la forma

$$\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\},$$

o podemos pensar Z_n como el conjunto de los números $0, 1, \ldots, n-1$ sobre los cuáles se opera con la suma y el producto módulo n. Esta representación es útil para calcular y trabajar, en tanto que la representación como clases de equivalencia es mucho más adecuada para demostrar las propiedades de la aritmética en $\mathbb{Z}_n$.

Teorema A.1. *Para cualquier número natural $n > 1$, la suma en $\mathbb{Z}_n$ tiene las propiedades conmutativa, asociativa, de existencia de neutro y de existencia de opuesto. El producto tiene las propiedades conmutativa, asociativa, y de existencia de unidad. Además la suma es distributiva respecto al producto.*

Observación A.2.4. Un conjunto $\mathbb{A}$ que tiene operaciones de suma y producto con las propiedades que se listan en el teorema A.1 es lo que se llama un *anillo conmutativo con unidad*. Podemos decir entonces que $\mathbb{Z}_n$ *es un anillo conmutativo con unidad*.

Digamos que un *anillo* es un conjunto que tiene operaciones de suma y producto con las mismas propiedades que aparecen en el enunciado del teorema A.1, salvo la propiedad conmutativa para el producto y la existencia de unidad para el producto. Por ejemplo, el conjunto de las matrices cuadradas $n \times n$ es un anillo con unidad (ver la página PONER REFERENCIA), pero no es conmutativo. El conjunto de los números pares forma un anillo conmutativo, pero no tiene unidad.

La demostración del teorema A.1 se basa en escribir las operaciones a partir de representantes de las clases, tal como se hace en la definición de las operaciones, y usar las propiedades de la aritmética en $\mathbb{Z}$. Lo dejamos como ejercicio para el lector.

Ejercicio A.5. Demostrar el teorema A.1.

Ya hemos visto las tablas de operaciones para $\mathbb{Z}_2$. En las tablas siguientes se describen explícitamente las operaciones para $\mathbb{Z}_3$, $\mathbb{Z}_4$ y $\mathbb{Z}_5$. La construcción es sencilla. Por ejemplo el producto 3×4 en $\mathbb{Z}_5$ se obtiene calculando primero $3 \times 4 = 12$, y luego teniendo en cuenta que $12 \equiv 2 \pmod{5}$. Los restantes productos y sumas se calculan con el mismo procedimiento.

La tablas de suma y producto en $\mathbb{Z}_3$ son

$$
\begin{array}{c|ccc}
+ & 0 & 1 & 2 \\
\hline
0 & 0 & 1 & 2 \\
1 & 1 & 2 & 0 \\
2 & 2 & 0 & 1
\end{array}
,\qquad
\begin{array}{c|ccc}
\times & 0 & 1 & 2 \\
\hline
0 & 0 & 0 & 0 \\
1 & 0 & 1 & 2 \\
2 & 0 & 2 & 1
\end{array}
.
$$

En $\mathbb{Z}_{\ntrianglerighteq}$ son

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

,

×	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

,

y en $\mathbb{Z}_5$ son

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

,

×	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

.

Ejercicio A.6. Construir las tablas para la aritmética en $\mathbb{Z}_6$ y $\mathbb{Z}_{10}$.

Los cuerpos $\mathbb{Z}_p$

A un anillo conmutativo con unidad sólo le falta una propiedad para ser un cuerpo. Que todo elemento no nulo (distinto del neutro para la suma) tenga un inverso para la multiplicación. Un rápido examen de las tablas de multiplicar de $\mathbb{Z}_2$, $\mathbb{Z}_3$ y $\mathbb{Z}_5$ muestra que todo elemento distinto de cero tiene su inverso en la multiplicación, de modo que $\mathbb{Z}_2$, $\mathbb{Z}_3$ y $\mathbb{Z}_5$ son cuerpos. Si indicamos con a^{-1} el inverso de un elemento a de un cuerpo, encontramos, por ejemplo, en $\mathbb{Z}_5$ que

$$1^{-1} = 1, \quad 2^{-1} = 3, \quad 3^{-1} = 2, \quad 4^{-1} = 4.$$

Sin embargo, esto no ocurre para $\mathbb{Z}_4$. No hay ningún elemento que multiplicado por 2 arroje el resultado 1. El 2 en $\mathbb{Z}_4$ es lo que se llama un *divisor de cero*: un número distinto de cero que al ser multiplicado por otro número distintos de cero arroja el resultado cero. En efecto, se tiene que

$$2 \times 2 = 0 \pmod{4}.$$

En los cuerpos no hay divisores de cero.

Ejercicio A.7. Mostrar que en un cuerpo no puede haber divisores de cero, ya que la condición $ab = 0$ implica que $a = 0$ o $b = 0$.

El problema con $\mathbb{Z}_4$ es que 2 divide a 4, y por eso es un divisor de 0. En general, cuando m divide a n entonces m es un divisor de 0 en $\mathbb{Z}_n$ y $\mathbb{Z}_n$ no puede ser un cuerpo. Este problema no aparece cuando n es primo, y se tiene el siguiente resultado general:

Teorema A.2. $\mathbb{Z}_n$ *es un cuerpo si y sólo si* n *es primo.*

La demostración de este teorema es algo más compleja que los ejercicios anteriores. Recomendamos al lector la consulta de la página 719 de "Matemáticas discretas y combinatorias"de Grimaldi en su tercera edición o la página 271 de "Notas de Álgebra 1"de Enzo Gentile, Eudeba, tercera edición.

Es corriente cambiar ligeramente la notación a $\mathbb{Z}_p$ para referirse a los cuerpos que se obtienen con p primo.

A.2.3. Una aritmética para los "bytes": $\mathbb{F}_{2^8}$

Hemos encontrado toda una familia de cuerpos discretos con p elementos, donde p puede ser un número primo cualquiera. Entre ellos está el cuerpo $\mathbb{Z}_2$, formado por los símbolos 0 y 1, que provee una estructura algebraica adecuada al manejo de los bits de información.

Muchas veces los bits aparecen agrupados en bytes, listas de, por ejemplo, 8 bits. Hay $2^8 = 256$ de estas listas. Mostraremos ahora como se puede dar estructura de cuerpo a este conjunto de listas de ocho ceros y unos, definiendo una suma y un producto sobre las listas.

Haremos la construcción detallada para un caso más pequeño, el conjunto

$$\{(0,0),(1,0),(0,1),(1,1)\},$$

y luego indicaremos cómo se hace en el caso general de listas de cualquier longitud, en particular, de longitud 8.

Sumar listas es sencillo: basta sumarlas entrada a entrada, con la aritmética de $\mathbb{Z}_2$. Observemos que esta es la misma noción de suma que empleamos en el espacio vectorial $\mathbb{Z}_2^2$

Para multiplicarlas identificaremos cada pareja

$$(a_0, a_1)$$

de números en $\mathbb{Z}_2$ con el polinomio

$$a_0 + a_1 x$$

que tiene a_0 y a_1 como coeficientes.

Ahora ya sabemos multiplicar dos parejas: simplemente multipliquemos los polinomios que les corresponden. El resultado es un polinomio que puede ser de segundo grado. Y aquí entra en juego una idea similar a la de la construcción de la aritmética módulo un entero: conservemos del polinomio producto solamente el resto de dividir entre un polinomio de segundo grado. Para que la construcción tenga buenas propiedades este nuevo polinomio debe ser "primo" en el siguiente sentido: no debe admitir una factorización como el producto de dos polinomios de grado menor. Por ejemplo, x^2 no nos sirve, porque es el producto de x por sí mismo.

Observemos que si definimos la suma de parejas como la suma de sus polinomios asociados módulo un polinomio irreducible de segundo grado seguimos teniendo la suma original, que consiste en sumar las parejas entrada a entrada. Necesitamos esta propiedad para conseguir la propiedad distributiva de la suma frente al producto.

Es relativamente sencillo encontrar los polinomios irreducibles de grado 2 sobre $\mathbb{Z}_2$, simplemente examinándolos todos. Al fin y al cabo hay sólo 4 posibilidades. Además el término independiente tiene que ser 1, porque si fuera 0 podríamos factorizar x. Sólo nos quedan $x^2 + 1$ y $x^2 + x + 1$. Observemos que

$$x^2 + 1 = (x + 1)(x + 1),$$

ya que al operar en $\mathbb{Z}_2$ al sumar $x+x$ obtenemos 0 como coeficiente del término de primer grado. ¿Será irreducible $p(x) = x^2+x+1$? Ya sabemos que no puede tener el factor x. El otro polinomio de grado 1 es $x + 1$, y tampoco puede ser un factor de $p(x)$, porque $p(1) = 1 + 1 + 1 = 1$. Entonces 1 no es una raíz de $p(x)$ y $x + 1$ no puede ser factor de $p(x)$.

Con el polinomio irreducible a mano calculemos la tabla de multiplicar:

$$(a_0 + a_1x)(b_0 + b_1x) = a_0b_0 + (a_0b_1 + a_1b_0)x + a_1b_1x^2.$$

Hacemos la división del producto entre $x^2 + x + 1$ y obtenemos

$$\begin{aligned}&a_0b_0 + (a_0b_1 + a_1b_0)x + a_1b_1x^2 =\\&\qquad a_1b_1(1 + x + x^2) + a_0b_0 + a_1b_1 + (a_0b_1 + a_1(b_0 + b_1))x.\end{aligned}$$

Vemos entonces que el producto entre dos parejas debe definirse por la fórmula

$$(a_0, a_1) \times (b_0, b_1) = (a_0b_0 + a_1b_1, a_0b_1 + a_1(b_0 + b_1))$$

que da lugar a la tabla

×	00	10	01	11
00	00	00	00	00
10	00	10	01	11
01	00	01	11	10
11	00	11	10	01

Hemos representado cada pareja escribiendo sus dos símbolos uno junto al otro. La multiplicación puede parecer un poco engorrosa, pero se simplifica bastante al observar que

$$(1,0) = (0,1)^0,\ (0,1) = (0,1)^1,\ (1,1) = (0,1)^2.$$

La tabla de suma es, por supuesto

+	00	10	01	11
00	00	10	01	11
10	10	00	11	01
01	01	11	00	10
11	11	01	10	00

El cuerpo que hemos construido por este procedimiento se designa con la notación $\mathbb{F}_{2^2}$, o $\mathbb{GF}(2^2)$. La F proviene del término inglés *field*. La $\mathbb{G}$ en $\mathbb{GF}$ alude a Evaristo Galois (1811-1832), el matemático francés que introdujo este tipo de cuerpos finitos. No demostraremos que se trata de un cuerpo, pero no es demasiado complicado hacerlo a partir de las propiedades del producto y suma de polinomios.

La construcción de cuerpos con 2^k elementos sigue un patrón similar, pero empleando un polinomio irreducible de grado k. Para el caso $k = 8$ puede usarse

$$p(x) = x^8 + x^4 + x^3 + x^2 + 1$$

en la construcción del cuerpo $\mathbb{F}_{2^8}$.

Estos temas se pueden profundizar en el capítulo 17 de "Matemáticas discreta y combinatoria" de Ralph Grimaldi, a partir de la página 835 (en la tercera edición).

A.3. Trabajos

En esta sección presentamos algunas propuestas de trabajos que combinan temas de Geometría y/o Álgebra Lineal con otros temas de Matemática, y en los que se apunta a resolver problemas de otras áreas de la ciencia.

A.3.1. Reacciones en un reticulado hiperestático

En este trabajo consideraremos la estructura reticulada, plana, que se representa en la figura A.1. La estructura está anclada en dos apoyos fijos ubicados en los nodos 1 y 5, que son capaces de ejercer reacciones horizontales y verticales para mantenerla fija. Nuestro objetivo será determinar los valores de

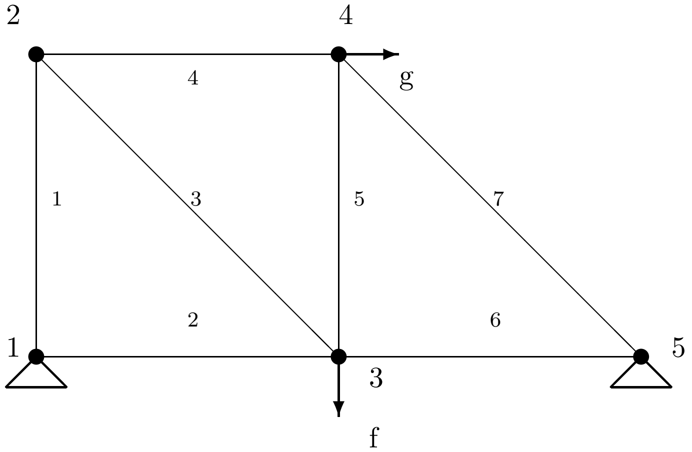


Figura A.1: Una estructura de barras firmemente anclada

las tensiones T_i, $i = 1, \ldots, 7$ en las barras, y de las reacciones en los anclajes, cuando la estructura debe equilibrar cargas f y g en los nodos 3 y 4 como se muestra en la figura.

1. Ecuaciones de equilibrio. Plantear las ecuaciones de equilibrio en los nodos, que vinculan a los valores de las tensiones en las barras y de las reacciones en los anclajes.

2. Un sistema hiperestático. Se llama *isostática* a una estructura cuando los esfuerzos en cada parte de ella pueden calcularse a partir de las ecuaciones de equilibrio, e *hiperestática* cuando las ecuaciones de equilibrio no determinan completamente a la solución. Mostrar que la estructura que se muestra en la figura A.1 es hiperestática[3].

3. Energía de deformación. En esta parte vamos a incorporar al modelo que estamos empleando para estudiar el equilibrio de una estructura de barras el concepto de *energía de deformación*[4]. Supondremos que cada barra tiene asociada una constante positiva k_i, $i = 1, \ldots, 7$, tal que al ser sometida a una

[3]En la parte 2 del ejercicio ejercicio.del.reticulado.sislin0, página 19, consideramos la misma estructura de barras, pero anclada de modo tal que no podía haber reacciones horizontales en el nodo 5. En ese caso el problema era isostático.

[4]Es un concepto similar al de energía de deformación almacenada en un resorte

tensión T_i sufre una variación Δl_i en su longitud, que satisface

$$T_i = k_i \Delta l_i.$$

En estas condiciones, la energía de deformación almacenada en la barra i es

$$E_i = \frac{k_i}{2} (\Delta l_i)^2 = \frac{T_i^2}{2k_i},$$

y la energía total en el sistema de barras es

$$E = \sum_{i=1}^{7} E_i.$$

Incorporaremos a nuestro modelo el siguiente *principio de mínima energía*: **entre todos las posibles determinaciones de las tensiones T_i en las barras, el sistema adopta aquella que hace mínimo el valor de la energía de deformación E.**

Supongamos que el valor de la constante k_i es un valor conocido k para todas las barras, salvo la 3 y las 7 para las que $k_3 = k_7 = k/\sqrt{2}$. Aplicar el principio de mínima energía para calcular los valores de las tensiones y reacciones en el sistema de la figura A.1.

Observación A.3.1. En el nuevo modelo las barras ya no son rígidas. De todos modos, despreciaremos las variaciones en la geometría del reticulado que se producen por los cambios de longitud en las barras para escribir las ecuaciones que relacionan las tensiones en las distintas barras.

A.3.2. Códigos de Hamming

El objetivo de este trabajo es presentar un ejemplo de *código corrector de errores*. Se trata de una estrategia para proteger la información digital contra fallos en su transmisión y almacenamiento. La idea básica consiste en agregar cierta redundancia artificial en la información que permita detectar y corregir los errores. Por ejemplo, una forma un tanto primitiva de alcanzar este propósito consiste en almacenar o transmitir la misma información muchas veces: las repeticiones son redundantes, no agregan nada a la porción original de información, pero permiten detectar y corregir eventuales errores. Los códigos correctores permiten hacer lo mismo pero de manera más eficiente.
1. Codificando. La propuesta de Hamming para codificar una lista

$$X = (x_1, x_2, x_3, x_4)^t$$

de cuatro bits (o sea, donde los caracteres x_i, $i = 1, \dots, 4$ pueden ser ceros o unos) es generar a partir de ella la lista

$$Y = (x_1 + x_2 + x_4, x_1 + x_3 + x_4, x_1, x_2 + x_3 + x_4, x_2, x_3, x_4)^t.$$

Las sumas en esta expresión, al igual que todas las operaciones de suma y producto en este trabajo, deben entenderse en el sentido de $\mathbb{Z}_2$. Representar el esquema de codificación $X \mapsto Y$ por medio de una matriz G tal que $Y = GX$. Observar que en las posiciones 1, 2 y 4 de la lista Y aparecen caracteres redundantes, que no agregan ninguna informacion a los almacenados en las posiciones 3, 5, 6 y 7 porque pueden ser calculados a partir de ellos.

Llamaremos *código de Hamming [7,4,3]* y designaremos con la letra $\mathcal{H}$ al conjunto de todas las posibles listas Y que pueden obtenerse por este procedimiento de codificación. Para usar una jerga más afín a la empleada en la teoría de códigos llamaremos *palabras* a cada una de las lista Y en el código $\mathcal{H}$.

Mostrar que el código de Hamming es un subespacio vectorial de $\mathbb{Z}_2^7$, hallar su dimensión y su cardinalidad. Hacer una lista de todas las palabras del código.

2. Matriz de control. Mostrar que $\mathcal{H}$ es el núcleo de la matriz

$$C = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix},$$

a la que llamaremos *matriz de control* del código.

3. Distancia de Hamming. Dadas dos listas X e Y en $\mathbb{Z}_2^7$ definiremos la *distancia de Hamming* $d(X,Y)$ entre ambas como el número de posiciones en que la lista X difiere de la lista Y. Es decir,

$$d(X,Y) = \#\{j \,:\, 1 \leq j \leq n \,,\, x_j \neq y_j\},$$

donde # indica la cardinalidad o número de elementos de un conjunto. Mostrar que la distancia de Hamming tiene las siguientes propiedades[5]:

1. $d(X,Y) \geq 0$ y $d(X,Y) = 0$ si y sólo si $X = Y$, $X, Y \in \mathbb{Z}_2^7$;
2. $d(X,Y) = d(Y,X)$, $X, Y \in \mathbb{Z}_2^7$;
3. $d(X,Z) \leq d(X,Y) + d(X,Z)$, $X, Y, Z \in \mathbb{Z}_2^7$.

Mostrar además que $d(X,Y) = d(X+Y,0)$, $X, Y \in \mathbb{Z}_2^7$.

4. Distancia al código y entre palabras del código. Mostrar que $Y \in \mathbb{Z}_2^7$ no pertenece al código de Hamming $\mathcal{H}$ existe una pálabra Y' en el código tal que $d(Y,Y') = 1$. Sugerencia: observar que en la matriz de control C están todas las posibles columnas no nulas de $\mathbb{Z}_2^3$.

Mostrar que la distancia entre dos palabras distintas de $\mathcal{H}$ nunca es menor a 3. Sugerencia: usar que $d(X,Y) = d(X+Y,0)$ y examinar la lista de palabras en $\mathcal{H}$.

Concluir que para cada palabra Y que no está en el código hay una única palabra del código que es la más cercana a Y.

5. Corrección de errores con el Código de Hamming. La transmisión de una palabra de $X = (x_1, x_2, x_3, x_4)$ se implementa de la siguiente manera: se códifica X en una palabra $Y = GX$ y se transmite Y. Esta palabra se transmite y el receptor recibe Y_r. Si $Y_r \in \mathcal{H}$ el receptor busca la palabra X_r tal que $Y_r = GX_r$, e interpreta que se transmitió X_r. Si un error adultera el mensaje y resulta que $Y_r \notin \mathcal{H}$ entonces el receptor busca la palabra Y_r' del código que está más cerca de Y_r, e interpreta que se transmitió la palabra X_r tal que $Y_r' = GX_r$.

Mostrar que este procedimiento permite decodificar correctamente todos los mensajes que se transmiten sin ningún error, y también aquellos en los que un error modifica sólo una posición de la palabra Y durante la transmisión.

Recibiendo una transmisión codificada con este criterio nos llegan los mensajes $(0,1,1,0,1,1,0)$ y $(1,1,1,0,1,1,0)$. Hallar en cada caso la palabra $X = (x_1, x_2, x_3, x_4)$ que debe interpretarse que fue codificada para su transmisión.

[5]Es importante notar que estas tres propiedades son compartidas con la distancia usual entre puntos del plano o del espacio. Estas tres propiedades caracterizan a lo que, en general, se llama una función distancia, y a partir de ellas es posible formular una teoría general de los espacios en los que existe una medida de la distancia entre puntos.

A.3.3. Oscilaciones de una masa sujeta a un resorte

En este trabajo estudiaremos cómo se mueve una masa $m > 0$ que está sujeta a un resorte de constante elástica $k > 0$. Para describir la posición de la masa m pondremos coordenadas con origen en la configuración de reposo del sistema, en la que el resorte tiene su longitud natural. Con estas coordenadas describiremos la posición de la masa m en cada instante t, que estará dada por una función real x que indica la posición $x(t)$ que la masa m ocupa al tiempo t.

Cuando la masa m se desplaza a una posición x entonces el resorte ejerce una fuerza cuya componente en la dirección del eje x es

$$F = -kx. \tag{A.5}$$

Por otra parte, si $x(t)$ es la posición de m entonces la derivada de x respecto al tiempo representa su velocidad $v(t)$. Indicaremos esta derivada con la notación $\dot{x}(t)$. La derivada de la velocidad es la aceleración $a(t)$. Es decir, la aceleración $a(t)$ es igual a la derivada segunda de x respecto al tiempo t, a la que indicaremos con $\ddot{x}(t)$, de modo que

$$a(t) = \ddot{x}(t). \tag{A.6}$$

El movimiento de la masa m está gobernado por la Ley de Newton

$$\text{FUERZA} = \text{MASA} \times \text{ACELERACIÓN}.$$

Cuando los valores de la fuerza y la aceleración se escriben en términos de $x(t)$ usando las ecuaciones (A.5) y (A.6) encontramos que esta función debe satisfacer para todos los valores de t la igualdad

$$m\ddot{x}(t) + kx(t) = 0, \tag{A.7}$$

a la que llamaremos *ecuación del resorte*. El objetivo de este trabajo es mostrar que la ecuación (A.7), junto con las condiciones iniciales de posición $x_0 = x(0)$ y velocidad $v_0 = \dot{x}(0)$ de la masa m, determina completamente el movimiento $x(t)$.

1. Dos soluciones[6] particulares. Hallar el valor de ω que hace que las funciones $\operatorname{sen}(\omega t)$ y $\cos(\omega t)$ satisfagan la ecuación del resorte. Mostrar que son soluciones periódicas de la ecuaci'on y hallar su período. Hallar las condiciones

[6] La expresión “solución” tiene aquí el sentido usual: diremos que una función $x(t)$ es solución de la ecuación del resorte si la ecuación (A.7) se satisface para todo t.

iniciales, en $t = 0$, de posición y velocidad para la masa m que corresponden a cada una de estas soluciones.

2. Estructura vectorial de las soluciones. Mostrar que el conjunto de las funciones $x : \mathbb{R} \to \mathbb{R}$ que tienen derivada segunda continua y satisfacen la ecuación del resorte forman un espacio vectorial. Llamaremos $\mathcal{R}$ a este conjunto.

3. Soluciones con condiciones iniciales cualesquiera. Dadas una posición inicial x_0 y una velocidad inicial v_0 cualesquiera hallar una solución $x(t)$ de la ecuación del resorte que satisfaga las condiciones iniciales

$$x(0) = x_0, \quad \dot{x}(0) = v_0.$$

4. Unicidad de soluciones: reducción al caso trivial. Mostrar que el problema de resolver la ecuación del resorte con condiciones iniciales

$$x(0) = x_0, \quad \dot{x}(0) = v_0$$

tiene solución única si y sólo si el problema de resolver la ecuación del resorte con condiciones iniciales

$$x(0) = 0, \quad \dot{x}(0) = 0$$

tiene solución única.

5. Unicidad de soluciones: energía. Mostrar que si $x(t)$ es una solución de la ecuación del resorte entonces la energía

$$E(t) = \tfrac{1}{2}\left(kx^2(t) + m\dot{x}^2(t)\right)$$

es constante. Deducir que la única solución de la ecuación del resorte con datos iniciales $x(0) = 0$, $\dot{x}(0) = 0$, es la solución trivial tal que $x(t) = 0$ para cualquier valor de t.

Concluir que el problema de hallar una función $x : \mathbb{R} \mapsto \mathbb{R}$, con derivada segunda continua, que satisfaga

$$\begin{cases} m\ddot{x}(t) + kx(t) = 0, \\ x(0) = x_0, \\ \dot{x}(0) = v_0, \end{cases}$$

tiene solución única para cualquier par de números x_0 y v_0.

5. Dimensión del espacio de soluciones de la ecuación del resorte. Hallar una base del espacio $\mathcal{R}$ de las soluciones de la ecuación del resorte, y su dimensión. Justificar.

Bibliografía

[C] David Carlson, *Teching Lineal Algebra: Must the Fog Always Roll In*, College Mathematics Journal, volumen 24 (1993), páginas 29–40. El artículo fue reproducido en el volumen en Resources for Teaching Linear Algebra, editado por David Carlson y otros, MAA Notes, volumen 42, Mathematical Association of America, 1997.

[Do] Manfredo Do Carmo, *Geometría diferencial de curvas y superficies*, Alianza, 1976.

[D] Jean Luc Dorier, *Teaching Linear Algebra at University*, Actas del Congreso Internacional de Matemáticos, Pequín, 2002, volumen III, páginas 875–884.

[F] Pablo Fernández, *El secreto de Google y el Álgebra Lineal*, Boletín de la Sociedad Española de Matemática Aplicada, número 30 (2004), páginas 115–141. Disponible en `http://imerl.fing.edu.uy/GAL1/` por cortesía del autor.

[G] Ralph P. Grimaldi, *Matemáticas Discreta y Combinatoria*, Addison-Wesley Iberoamericana, 1997.

[H] Eugenio Hernández, *Álgebra y geometría*, Addison-Wesley/Universidad Autónoma de Madrid, segunda edición, 1987.

[I] Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia", *Geometría y Álgebra Lineal 1*, Oficina de Publicaciones, Centro de Estudiantes de Ingeniería, 2000.

[L] David C. Lay, *Linear Algebra and its Applications*, Addison-Wesley, segunda edición, 2000.

[M] Roberto Markarian y Nelson Möller, *Como cuantificar la importancia individual en una estructura de enlaces: Google-PageRank*, PreMat 2004/78. Disponible en `http:/premat.fing.edu.uy/`

[S] Gilbert Strang, *Introduction to Lineal Algebra*, Wellesley-Cambridge Press, segunda edición, 1998.

[TB] Lloyd N. Trefethen, David Bau III, *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics, 1997.

[T] Alan Tucker, *The Growing Importance of Linear Algebra in Undergraduate Mathematics*, en Resources for Teaching Linear Algebra, editado por David Carlson y otros, MAA Notes, volumen 42, Mathematical Association of America, 1997.

Índice alfabético

N

O

P