Geometría y Algebra Lineal I

Instituto de Matemática y Estadística

"Prof. Rafael Laguardia"

Facultad de Ingeniería Universidad de la República

Índice general

PREFACIO

Estas notas son una reelaboración, corregida y muy modificada de las correspondientes partes de las notas de 1991, en cuya redacción participaron un gran número de docentes del Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia"(IMERL) de la Facultad de Ingeniería de la Universidad de la República.

Varios capítulos fueron redactados nuevamente, teniendo en mente transformarlos a la brevedad en partes de un libro sobre el tema.

Los principales defectos de las notas derivan de la impericia de los firmantes y de la celeridad con que fue armada esta versión (no) final.

El Algebra Lineal constituye hoy una rama básica de la matemática con aplicaciones dentro y fuera de esta. No se han incluido en esta edición, sobre todo por razones de tiempo ejemplos de tales aplicaciones. Sin embargo algunas serán abordadas en clase, además en la bibliografía se indican algunas referencias donde el lector interesado puede encontrar varios ejemplos interesantes.

La carencia de ejercicios no debe extrañar dado que éstos son editados en fascículos aparte.

La estructura general del libro -que no será necesariamente la del curso- es la siguiente. Se incluye un Capítulo Cero, en que se desarrollan los conceptos de Relaciones, Funciones, etc., con diversos ejemplos. Su lectura facilitará la comprensión del lenguaje del curso, y dará una referencia general sobre definiciones y resultados básicos. Fue extraído de *Notas de Álgebra*, de Alfredo Jones, editadas por el IME de la Universidad de San Pablo.

En el Capítulo Uno se estudian los Sistemas de Ecuaciones Lineales, se desarrolla su presentación matricial y se describe el algoritmo de Gauss o de escalerización como herramienta principal para su resolución.

En el Capítulo Dos se estudian las principales propiedades de las matrices y sus operaciones. Se presenta un ejemplo clave del curso: el espacio de n-uplas, $\mathbb{R}^n$ y culmina con el estudio del rango de una matriz, concepto directamente relacionado con las soluciones de los sistemas lineales y con la invertibilidad de las matrices cuadradas.

El Capítulo Tres estudia el Determinante de una matriz cuadrada se prueban algunos resultados clásicos como el teorema de Cramer y el teorema de Binet-Cauchy sobre el determinante de un producto de matrices.

Los Capítulos Cuatro y Cinco (que podrían ser los primeros de las notas) están dedicados a la introducción de un segundo ejemplo básico: el de los vectores libres del espacio. Estos se utilizan para dar una exposición de los primeros elementos de la Geometría Analítica del Espacio.

El Capitulo Seis está dedicado a los Espacios Vectoriales. Es el capítulo decididamente nuevo para todos los estudiantes, y el centro conceptual del curso, en el se introduce de manera axiomática una estructura abstracta que se inspira y generaliza varios de los ejemplos tratados en los capítulos previos.

El Capítulo Seis está dedicado a un tipo particular de funciones entre espacios vectoriales: las Transformaciones Lineales. Estas se caracterizan por preservar la estructura de los espacios vectoriales y son en cierto sentido el modelo mas sencillo de una función. Se estudia aquí con detalle su relación con las matrices y se pone de manifiesto el contenido geométrico de muchas propiedades algebraicas de estas últimas.

Un comentario final es pertinente para preparar al estudiante a aprovechar tanto el curso como estas notas. Aquí cada capítulo interviene de manera fundamental en los siguientes. La gran mayoría de los tópicos abordados en los primeras clases tiene un rol fundamental en el resto del curso. Difícilmente será posible avanzar de un capítulo al siguiente sin haber comprendido el primero.

En la presente edición colaboró particularmente José Díaz.

Lecturas complementarias recomendadas. De cualquiera de los libros que se indican a continuación, hay diversas otras ediciones; en particular en sus lenguas originales:

A.G. Kurosch: *Curso de Álgebra Superior*, Mir-Nimusa.

E. Lages Lima: *Álgebra Linear*, IMPA.

> Un libro excelente, escrito por un experimentado matemático autor de numerosos textos, cubre un amplio tópico de temas con rigor y profundidad pero no incluye los capítulos de geometría. Además sigue un orden algo distinto del de nuestras notas, los capítulos de matrices y sistemas de ecuaciones aparecen luego de tratar espacios vectoriales. Es ampliamente recomendado para ampliar los dos últimos capítulos de nuestro curso.

P.R. Halmos: *Espacios Vectoriales de dimensión finita*, CECSA.

> Una obra clásica sobre el tema, realizada por un destacado matemático, tampoco aborda los temas iniciales de nuestro curso. Su enfoque sobre la teoría de espacios vectoriales esta pensada para quien desea luego profundizar en el estudio de espacios de dimensión infinita.

E. Hernández *Álgebra y Geometría* , Adisson–Wesley.

> Este libro cubre todos los temas de nuestro curso incluyendo los capítulos de geometría, en un orden similar. Es escueto en explicaciones y las pruebas a veces resultan oscuras. No contiene aplicaciones.

R. Hill*Álgebra Lineal Elemental con Aplicaciones*, Prentice Hall.

> Este libro es como su nombre lo indica, tal vez más elemental que los otros sin embargo es claro abarca casi todos los temas del curso y sobre todo tiene un número grande de aplicaciones interesantes a la ingeniería y otras disciplinas. Incluye

una introducción al Matlab y tiene numerosos ejercicios, incluyendo algunos proyectos para trabajar en computadora.

K. Koffman & R. Kunze: *Álgebra Lineal*, Prentice-Hall.

Otro excelente libro, que no cubre los capítulos de geometría, es muy claro y riguroso, en algunos momentos aborda temas (determinantes)con mucha generalidad lo cual puede dificultar la comprensión en una lectura inicial

G. Nakos & D. Joyner *Algebra Lineal con Aplicaciones*, Thomson.

De nivel similar al libro de Hill tiene también un gran número de ejemplos y aplicaciones y algunas notas históricas interesantes. Tal vez su mayor virtud es el gran número de ejemplos para trabajar con computadora. Se incluyen proyectos para trabajar con Matlab, Maple y Mathematica

G. Strang: *Algebra linel y sus aplicaciones*, Addison–Wesley.

Este libro tiene un enfoque algo diferente de los otros libros recomendados. Su objeto de estudio son los sistemas lineales y las matrices, los espacios vectoriales y las transformaciones lineales solo aparecen secundariamente. No obstante tiene un punto de vista interesante y claro que puede ayudar a ver desde otra óptica los problemas analizados en el curso, es especialmente recomendado para los primeros capítulos del curso. Cuenta con un gran número de aplicaciones interesantes incluyendo códigos de computadora en Fortran.

Marzo del 2000
Marcelo Cerminara
Roberto Markarian
Responsables de esta edición

CAPíTULO 0

RELACIONES Y FUNCIONES

0.1. Relaciones

En este capítulo introduciremos algunas definiciones y notaciones básicas para el desarrollo de estas notas. Supondremos que el lector tiene cierta familiaridad con las nociones de relación y de función, de modo que no nos detendremos en la explicación del significado de estos conceptos, sino que partiremos de sus definiciones estudiando luego sólo aquellas propiedades que serán aplicadas en los capítulos que siguen.

Comenzaremos introduciendo la noción de relación mediante una definición que esta motivada por la idea usual de una relación entre pares de objetos. Observamos que dar una relación entre pares de elementos de un conjunto, en la acepción corriente del término, equivale a dar una lista formada por los elementos que verifican dicha relación, esto es, a especificar un conjunto de pares.

Dados dos conjuntos, A y B, indicaremos con A $\times$ B el conjunto constituido por los pares ordenados (a,b) tales que **a** pertenece al conjunto A y b pertenece al conjunto B. Con las notaciones usuales de la teoría de conjuntos esto se traduce en la igualdad:

Presuponiendo solo los conceptos de conjunto, subconjunto, elemento y par ordenado de elementos de un conjunto, podemos definir una relación como sigue;

DEFINICIÓN 0.1. Una **relación** en un conjunto A, es un subconjunto R de A $\times$ A.

Si (a,b) $\in$ R diremos que **a** esta en relación R con **b** y escribiremos aRb.

EJEMPLO 0.1. En todo conjunto A la relación de igualdad esta dada por el conjunto $R = \{(a, a) \,/ a \in A\}$, según esta relación, cada elemento de A solo esta en relación consigo mismo; luego en este caso aRb significa a = b. ■

EJEMPLO 0.2. Sea Z el conjunto de los números enteros, consideremos:

$$R = \{(a,b) \ / \ a \ , \ b \in Z, \ a - b \text{ es múltiplo entero de } 2\}.$$

En este caso, el entero **a** esta en relación R con el entero **b**, cuando son ambos pares o bien son ambos impares. ■

0.2. Relaciones de equivalencia

DEFINICIÓN 0.2. Una relación R en A se dice **reflexiva** si aRa para todo a $\in$ A; **simétrica** si aRb implica bRa y **transitiva** si aRb y bRc implica aRc. Una **relación de equivalencia** en un conjunto A es una relación reflexiva, simétrica y transitiva.

Cuando R es una relación de equivalencia, si aRb diremos que a es **equivalente** a b y usaremos la notación a $\sim$ b.

Las propiedades que **definen** una relación de equivalencia en un conjunto A, son entonces:

a) a $\sim$ a $\forall$ a $\in$ A.
b) a $\sim$ b $\Rightarrow$ b $\sim$ a.
c) a $\sim$ b y b $\sim$ c $\Rightarrow$ a $\sim$ c.

Mencionaremos algunos ejemplos de relaciones de equivalencia que son útiles, pues permiten definir ciertos conceptos importantes. En cada caso el lector verificará que valen a), b) y c).

EJEMPLO 0.3. En todo conjunto A se tiene una relación de equivalencia trivial: R = $\{$(a,a) / a $\in$ A$\}$, es decir a $\sim$ b $\Leftrightarrow$ a = b. ■

EJEMPLO 0.4. Sea **N** el conjunto de los números naturales. A fin de obtener los números enteros Z a partir de los naturales, se definen en **N** $\times$ **N** la relación (a,b) $\sim$ (a',b') $\Leftrightarrow$ a + b' = a' + b. Observar que cada par representará al entero que surge de restar a la primera componente la segunda. ■

EJEMPLO 0.5. Sea **Z** el conjunto de los enteros. A fin de obtener los racionales, se define en el conjunto **Z** $\times$ **Z** $-\{0,0\}$, la relación: (a,b) $\sim$ (a',b') $\Leftrightarrow$ a.b' = a'.b . ■

EJEMPLO 0.6. Dado un entero fijo m, se define una relación de equivalencia en **Z**, poniendo a $\sim$ a' cuando **a - a'** es múltiplo de m. En este caso se escribe a = a' (mod. m) para indicar que a $\sim$ a' y se dice que **a** y **a'** son **congruentes módulo m**. Observamos que **m** y **-m** definen la misma relación, de modo que se puede suponer m $\geqslant$ 0. ■

EJEMPLO 0.7. En el conjunto de los reales **R**, se tiene una relación de equivalencia tomando r $\sim$ r' cuando r - r' entero. ■

EJEMPLO 0.8. Se puede definir un ángulo como una figura plana que es la intersección de dos semiplanos del mismo plano. Esta noción de ángulo tiene la limitación de que para estos ángulos no es posible definir una suma con todas las propiedades deseables para las aplicaciones de este concepto, las que se pueden resumir diciendo que los ángulos con la suma deben formar un grupo abeliano. Para esto es necesario dar otra definición de ángulo. Uno de los caminos posibles para llegar a esa definición, consiste en introducir en el conjunto de los pares ordenados de semirrectas del plano con origen en un punto O, la siguiente relación de equivalencia. Suponemos sabido que: para cada par ordenado de semirrectas de origen O, (a,b), existe una única simetría axial del plano que transforma **a** en **b**, la indicaremos con la notación s(a,b). Dados dos pares ordenados de semirrectas de origen O, (a,b) y (a',b'), definimos (a,b) $\sim$ (a',b') si la simetría axial que transforma **a** en **b'**, también transforma **b** en **a'**, es decir si s(a,b) = s(b,a). Las propiedades reflexiva y simétrica se verifican trivialmente. Para demostrar que la relación es transitiva, debemos probar que si s(a,b') = s(b,a') y s(a',b") = s(b',a") entonces s(a,b") = s(b,a"). .*El resultado de aplicar sucesivamente tres simetrías cuyos ejes pasan por un punto fijo O, es una simetría".* De lo anterior surge que al aplicar sucesivamente s(a,b'), s(b',a") y s(a",b"), se lleva a **a** en **b**", por lo que esta composición es, efectivamente, s(a,b"). Análogamente, aplicando s(b,a'), s(a',b") y s(a",b"), se obtiene s(b,a"). De lo anterior y observando que: s(a,b') = s(b',a)
s(a",b') = s(b',a")
s(a",b") = s(b",a")
tenemos que: s(a,b") = s(b",a). ■

FIGURA 2.1

Anotamos para usarlo mas adelante, que, de la definición de esta relación (a,b) $\sim$ (a',b') si y solo si (a,a') $\sim$ (b,b').

EJEMPLO 0.9. Llamaremos **figura** del plano a todo conjunto de puntos del plano. La **relación de congruencia** de figuras en el plano, se puede definir como sigue; Dadas dos figuras planas F1 y F2 definimos F1 $\sim$ F2 cuando hay alguna sucesión finita de simetrías axiales del plano que transforma F1 en F2. Para mostrar que esta es una relación de equivalencia en el conjunto de las figuras planas, es suficiente verificar lo que sigue: *a)* Para toda figura plana F se tiene F $\sim$ F, porque la sucesión de dos simetrías iguales transforma todo punto del plano en si mismo.- *b)* Si la sucesión de simetrías $s1, \ldots, sn$, transforma F1 en F2 y $s'1, \ldots, s'm$ transforma F2 en F3, entonces la sucesión $s1, \ldots, sn, s'1, \ldots, s'm$ transforma F1 en F3.- ∎

DEFINICIÓN 0.3. Dada una relación de equivalencia en un conjunto A y un elemento a $\in$ A. Llamaremos "**clase de equivalencia de a**", al conjunto de todos los elementos x $\in$ A que son equivalentes con **a**, y que notaremos:

$$\mathrm{cl(a)}, \, a.$$

De las condiciones b) y c) de la definición de relación de equivalencia se deduce que si a $\sim$ b , entonces x $\sim$ a si y solo si x $\sim$ b . Luego a $\sim$ b implica cl(b) = cl(a) . Es claro que vale el recíproco, esto es: Si cl(b) = cl(a) entonces a $\sim$ b .

Todo elemento b $\in$ cl(a) se dice que es un **representante de esa clase**. Es decir que b es un representante de cl(a) si y solo si b $\sim$ a.

El siguiente resultado muestra que el conjunto de las clases de equivalencia determina la relación.

PROPOSICIÓN 0.1. *Dadas dos clases de equivalencia cl(a) y cl(b) en un conjunto A se tiene que, o bien cl(a) $\equiv$ cl(b), o bien son disjuntas (sin elementos comunes). Además, una relación de equivalencia clasifica los elementos de un conjunto A en clases de equivalencia, disjuntas dos a dos. Todo elemento de A pertenece a alguna clase.*

PROPOSICIÓN 0.2. *Recíprocamente, dada una clase cualquiera C de subconjuntos de A, no vacíos y sin elementos comunes dos a dos, tales que todo elemento de A pertenezca a alguno de ellos, existe una única relación de equivalencia en A cuyas clases son dichos subconjuntos.*

Demostración de 0.1. En efecto, supongamos que $c \in cl(a)$ y $c \in cl(b)$, entonces $c \sim a$ y $c \sim b$, de donde $a \sim b$, por lo tanto $cl(a) = cl(b)$. Por otra parte, para todo $a \in A$ se tiene que $a \in cl(a)$.

□

Demostración de 0.2. Se puede demostrar, que definiendo $a \sim b$ cuando a y b pertenecen a un mismo subconjunto de la clase C se obtiene una relación de equivalencia. □

0.3. Conjunto Cociente

DEFINICIÓN 0.4. Llamaremos "**conjunto cociente de A por R**", a aquel cuyos elementos son las clases de equivalencia de A definidas por la relación de equivalencia R en A, que notaremos:

$$A/R, \; A.$$

Según esta definición, cada elemento de A junto con todos sus equivalentes, constituyen un elemento de A/R. Es por esto que a veces se dice que A/R se obtiene **identificando** cada elemento de A con todos sus equivalentes. La utilidad de las relaciones de equivalencia proviene precisamente de que permiten en esta forma construir un conjunto de objetos nuevos, las clases de equivalencia, a partir de ciertos objetos dados, los elementos de A. Para individualizar un elemento de A/R, es decir una clase de A, se puede tomar un representante cualquiera de esa clase. Usualmente se trata de elegir un representante que facilite la notación o el cálculo.-

- Veamos cuales son los conjuntos cociente que se obtienen de algunas de las relaciones de la sección 2.

EJEMPLO 0.10. En este caso cada a $\in$ A solo es equivalente a sí mismo por lo tanto la clase de equivalencia de a es el conjunto {a}. Quiere decir que cada elemento de A/R es un conjunto formado únicamente por algún elemento de A; luego en este ejemplo al formar el conjunto cociente no tenemos nada escencialmente nuevo. ∎

EJEMPLO 0.11. En este caso la cl(a,b) = {(a',b') $\in$ **N** $\times$ **N** / a + b' = a' + b}, por lo tanto, en cada clase de equivalencia hay infinitos pares. Observando que (a,b) $\sim$ (a + 1, b + 1), se puede demostrar que si a = b, en la cl(a,b) hay: o bien un único representante de la forma (c,0) o bien un único representante de la forma (0,c) con c $\neq$ 0. Las clases correspondientes denotan c y -c. Si a = b, (a,b) $\sim$ (0,0). La clase de (0,0) se denota 0. Se definen los **números enteros** como los elementos del conjunto cociente obtenido. ∎

EJEMPLO 0.12. En este ejemplo las clases de equivalencia son de la forma:

$$cl\,(a,b) = \left\{ \left(a',b'\right)/a',b' \in Z, b' \neq 0, a.b' = a'.b \right\}.$$

Usaremos la notación a / b para la clase de (a,b). La igualdad de clase a / b = a' / b' significa la equivalencia de los representantes, (a,b) $\sim$ (a',b'), o sea, según la definición de la relación, a.b' = a'.b. Se demuestra fácilmente que para todo natural k $\neq$ 0 es (a.k,b.k) $\sim$ (a,b) , de aquí se deduce que se puede tomar para cada clase un representante (a,b) con a y b relativamente primos, es decir, que toda clase se puede escribir en la forma a / b con a y b primos entre sí. Se definen los **números racionales** como los elementos del conjunto cociente. De acuerdo con esta definición, cada racional, es una clase de equivalencia a / b, es decir un conjunto de pares enteros.- Designaremos al conjunto de los números racionales con la letra Q. ∎

EJEMPLO 0.13. En este caso cl(a) = {a , a ± m , a ± 2.m , . . .}. Sabemos que todo elemento a se puede escribir en la forma a = r ± q.m con 0 ⩽ r ⩽ m , por lo tanto a ∼ r y cl(a) ∼ cl(r). Quiere decir que para cada clase de equivalencia se puede tomar como representante uno de los números 0,1, ... , m-1. Por otra parte, si 0 ⩽ r < r' < m , es facil ver que r no es equivalente a r', en consecuencia hay exactamente m clases distintas. Usaremos la notación a para indicar la clase de a e indicaremos con Z_m el conjunto cociente, luego $Z_m = \left\{\overline{0}, \overline{1}, ..., \overline{m-1}\right\}$. Los elementos de Z_m, se llaman **enteros módulo m**. Vamos a introducir la siguiente notación: Si X es un conjunto finito cualquiera, indicaremos el número de elementos de X con — X —. En este ejemplo tenemos $|Z_m|$ = m. ■

EJEMPLO 0.14. En este ejemplo cl(r) = { r + n / n ∈ Z }. Como todo número real se puede escribir en la forma r + n con $0 \leqslant r < 1$ y n entero, para cada clase de equivalencia hay un representante tal que $0 \leqslant r < 1$. Además es claro que dos reales r y r', con $0 \leqslant r < r' < 1$ no pueden tener por diferencia un entero, luego pertenecen a clases distintas. Por lo tanto el conjunto de los reales r, tales que $0 \leqslant r < 1$, contiene un solo representante de cada clase de equivalencia. El conjunto cociente que resulta se llama el de los **números reales** de **modulo 1**. ■

EJEMPLO 0.15. Llamaremos ángulo de vértice O, a toda la clase de equivalencia de los pares de semirrectas de origen O según esta relación. Cada par de semirrectas de origen O es entonces un representante de un ángulo. Usaremos la notación aob para designar el ángulo que tiene como representante el par (a,b). Dada una semirrecta cualquiera a de origen O, para cada ángulo xoy existe una única semirrecta de origen O tal que $aob = xoy$. Para demostrarlo basta observar que para que b tenga esa propiedad tiene que ser (a,b) ∼ (x,y), o sea s(x,b) = s(y,a), luego la semirrecta b buscada es la imagen de x según s(y,a). En consecuencia si tomamos una semirrecta fija a, el conjunto de todos los pares de la forma (a,b) contiene uno y solo uno de los representantes de cada ángulo. ■

0.4. Funciones

Una función se determina dando dos conjuntos y asociando a cada elemento del primero un único elemento del segundo. Este concepto de función incluye mucho más que la idea de ser un procedimiento en el cual dados ciertos números permite calcular otros. Dándole al término el significado más amplio, una función puede no estar definida por ninguna expresión numérica o analítica. Además, puede asociar a elementos de un conjunto cualquiera, no necesariamente de números, elementos de otro conjunto cualquiera. La consideración de funciones en este sentido, sin restricciones sobre los conjuntos ni sobre la forma de asociar elementos del primer conjunto a elementos del segundo, se justifica por su utilidad en todas las ramas de la matemática. Veamos como se puede dar una definición formal del concepto de función que hemos esbozado, presuponiendo solo las nociones de conjuntos, subconjunto, elemento, par ordenado y terna ordenada. Esta definición esta motivada por la idea de "gráfico de una función". Observamos que una función que a cada número real x le asocia un real y, está determinada dando todos los pares de números (x,y) que constituyen una función de esa forma siempre que para cada número $x \in \mathrm{R}$, haya uno y un solo $y \in \mathrm{R}$ tal que $(x,y) \in \mathrm{G}$. Teniendo esto en cuenta, hacemos la siguiente definición:

DEFINICIÓN 0.5. Una **función** es una terna ordenada $F = (A,B,G)$, donde A y B son conjuntos cualquiera y G es un subconjunto de $A \times B$, tal que para cada $x \in A$ hay un $y \in B$ y uno solo para el cual $(x,y) \in G$.

A se llama el **dominio** o conjunto de partida, B el **codominio** o conjunto de llegada y G el **gráfico** de f. Para abreviar diremos que que f es una función de A en B y escribiremos $f : A \to B$.

Usaremos las palabras **transformación** y **aplicación** como sinónimo de función.

Una función $f =$ *(A,B,G)* puede tener dominio o codominio vacio, aunque estos casos tienen poco interés. Si $A = \emptyset$, entonces para todo B, $A \times B = \emptyset$, de donde $G = \emptyset$ y f es de la forma $f = (\emptyset$,B,$\emptyset)$.

Luego para cada conjunto B hay una única función con dominio vacío y codominio B. Si una función $f = (A, B, G)$ tiene codominio $B = \emptyset$ es $G = A \times B = \emptyset$, y como para cada $x \in A$ existe un $y \in B$, debe ser $A = \emptyset$. En consecuencia $f = (\emptyset,\emptyset,\emptyset)$ es la única función con codominio vacío.

DEFINICIÓN 0.6. Si $(x,y) \in G$ diremos que y es la imágen de x por f o el valor de f en x.

Indicaremos con *f(x)* la imagen de x por f, luego $G = \{\ (x\ ,\ f(x));\ x \in A\ \}$. Dar el gráfico de f equivale entonces a dar , para cada x del dominio, su imagen *f(x).*

Nótese que dos elementos diferentes del dominio pueden tener la misma imagen. Por ejemplo, sea

$$A{=}B{=}\boldsymbol{R} \text{ y } G = \left\{\left(x, x^2\right);\ x \in R\right\}$$

entonces para todo $x \neq 0$ es$x \neq -x$pero $f(x) = f(-x) = x^2$.

DEFINICIÓN 0.7. Dado un subconjunto del dominio, $X \subset A$, llamaremos **imagen de *x*** por f al subconjunto de B formado por las imágenes de todos los elementos de X.

La imagen de X se denotará $f(X)$, luego $f(X) = \{f(x); x \in X\}$. A veces se llama **recorrido** de f a la imagen del dominio, $f(A)$.

Nótese que $f(A) \subset B$ pero $f(A)$ puede no coincidir con B. Así en el último ejemplo el recorrido está formado por los reales positivos y no coincide con el codominio R.

DEFINICIÓN 0.8. Llamaremos imagen inversa o preimagen por f de un conjunto $Y \subset B$ al subconjunto, de A formado por todos los x que tienen imagen en Y, es decir:

$\{x \in A; f(x) \in Y\}$. Se llama imagen inversa de un elemento $y \in B$ a la imagen inversa del conjunto $\{y\}$.

Por ejemplo, si $f: R \to R$ es la función tal que $f(x) = x^2$ para todo $x \in R$, entonces la imagen inversa de 0 por f es $\{0\}$, la de 1 es $\{-1, 1\}$ y la de -1 es el conjunto vacío.

DEFINICIÓN 0.9. Una función $f : A \to B$ se dice **sobreyectiva** si $f(A) = B$ es decir, si para todo $y \in B$ existe por lo menos un $x \in A$ tal que $f(x) = y$.

EJEMPLO 0.16. La función $p: R \times R \to R$ definida por $p(x, x') = x$ es sobreyectiva. ■

EJEMPLO 0.17. La función $f : N \to N$ dado por $f(x) = 2x$ no es sobreyectiva, pues la imagen $f(N)$ está formada sólo por los números pares. En cambio, es sobreyectiva la función g de Nen el conjunto de los naturales pares, dada por $g(x) = 2x$. ■

El ejemplo 0.17 es un caso particular del siguiente. Dada una función cualquiera $f : A \to B$ la función $g : A \to f(A)$ definida por $g(x) = f(x)$ para todo $x \in A$, es sobreyectiva.

DEFINICIÓN 0.10. Una función $f : A \to B$ se dice **inyectiva** si para todo $y \in B$ hay al lo sumo un $x \in A$ tal que $f(x) = y$. Dicho de otra manera, f es inyectiva cuando $x \neq x'$ implica $f(x) \neq f(x')$.

EJEMPLO 0.18. La función $f : N \rightarrow N$ dada por $f(n) = n + 1$ es inyctiva, pues si $n \neq n'$ entonces $n + 1 \neq n' + 1$. No es sobreyectiva, porque no existe ningún natural n que verifique $f(n) = 0$. ■

EJEMPLO 0.19. Indicaremos con la letra C al conjunto de los números complejos, y con C^olos complejos diferentes de 0. Para todo real x está definido e^x; suponemos sabido que para cada real p¿0 existe un x tal que $e^x = p$. Se puede definir la función exponencial, $exp : C \rightarrow C^0$, tomando como valor en $x + iy$ el complejo:

$$exp(x + iy) = e^x (cosy + iseny)$$

Denotaremos este número con e^{x+iy}. Esta función no es inyectiva porque para todo real x y todo entero k vale:

$$e^{x+2k\pi i} = e^x (cos2k\pi + isen2k\pi) = e^x$$

Mostraremos que es sobreyectiva. Para esto observamos que todo complejo $u + iv \neq 0$ se puede expresar en la forma

$$u + iv = \rho (cos\phi + isen\phi)$$

donde

$$\rho = \sqrt{u^2 + v^2} > 0, \ \ cos\phi = \frac{u}{\sqrt{u^2 + v^2}} \ , \ \ sen\phi = \frac{v}{\sqrt{u^2 + v^2}}$$

Por otra parte, como $\rho > 0$, existe un real x tal que $\rho = e^x$, luego:

$$u + iv = \rho (cos\phi + isen\phi) = e^{x+i\phi}.$$

■

EJEMPLO 0.20. Sean $A \subset B$ conjuntos cualquiera. La función inclusión $i : A \rightarrow B$, definida por $i(x) = x$ para todo $x \in A$, es inyectiva. Si $A \neq B$, i no es sobreyectiva. ■

Dada una función $f : A \to B$, consideremos en A la relación $x\ x'$ si $f(x) = f(x')$. Es claro, que esta es una relación de equivalencia; sea A el conjunto cociente de A por esa relación. Todos los $x \in A$ que están en una clase de equivalencia $\alpha \in A$ tienen la misma imagen según f, luego se puede definir una función $f : A \to B$ eligiendo para cada $\alpha \in A$ un representante cualquiera $x \in A$ y tomando $f(\alpha) = f(x)$. La función f es inyectiva porque si $\beta \in A$ verifica $f(\beta) = f(\alpha)$, entonces para cada representante $y \in B$ se tiene:

$$f(y) = f(\beta) = f(\alpha) = f(x)$$

luego $y\ x$, por lo tanto $\beta = \alpha$.

DEFINICIÓN 0.11. Una función se dice **biyectiva** si es sobreyectiva e inyectiva.

Se usa **correspondencia biunívoca** como sinónimo de función biyectiva.

EJEMPLO 0.21. La función $f : Z \to Z$ definida por $f(n) = n + 1$ es biyectiva. ■

EJEMPLO 0.22. Es biyectiva la función $f : R \to R$ dada por $f(x) = x^3$? ■

EJEMPLO 0.23. La función logaritmo definida de los reales positivos en R es biyectiva. ■

EJEMPLO 0.24. Dado un conjunto cualquiera A definimos la función **identidad** en A, Id_A, como la función de A en A tal que para todo $x \in A$, $Id_A(x) = x$. Esta función es biyectiva. A veces escribiremos simplemente Id en vez de Id_A. ■

Sea $f : A \to B$ una función cualquiera y sea A el conjunto cociente de A por la relación $x\ x'$ si $f\,(x) = f\,(x')$; de consideraciones anteriores resulta que la función $\overline{f} : \overline{A} \to f\,(A)$ definida por $\overline{f}\,(\alpha) = f\,(x)$ donde x es un representante de α, es biyectiva.

DEFINICIÓN 0.12. Dada $f : A \to B$ y un conjunto $C \subset A$, se llama **restricción** de f a C a la función $g\ : C \to B$ definida por $g\ (x) = f\,(x)$ para todo $x \in C$. Se dice que f es una **prolongación de *g* en *A***. La restricción g se notará: q/c.

DEFINICIÓN 0.13. Una función $f\ :\ N\ \ \to\ \ A$ se llama **sucesión** de elementos de A. Un **elemento** de la sucesión es una imagen $f\,(i)$ de un $i \in N$.

Lo usual es denotar la imagen i como a_i, e indicar la sucesión en alguna de la siguientes formas:

$$(a_1, a_2, ...)\,; (a_i)_{i \in N}\,; (a_i)$$

Una n-upla de elementos de A es una función $f : \{1, ..., n\} \to A$. Se escribe $\{a_1, ..., a_n\}$, donde $a_1 = f\,(1)\,, ..., a_n = f\,(n)$. En general, dado un conjunto I, una función $f : I \to A$ se llama también una **familia** de elementos de A **indizada** por I.

De acuerdo a la definición una función es una terna, por lo tanto dos funciones, $f = (A, B, G)$ y $g = (A', B', G')$, son iguales sólo cuando *A=A'*, *B=B'* y *G=G'*, es decir cuando tienen igual dominio, igual codominio y el mismo valor, $f\,(x) = g\,(x)$, para cada x del dominio. De modo que, por ejemplo, la función $f\ :\ Z\ \ \to\ \ Z$ tal que $f\,(x) = -x$ para todo $x \in\ R$, es diferente a la función $f\ :\ Z\ \ \to\ \ R$ definida por $f\,(x) = -x$ para todo $x \in\ R$, pues los codominios no coinciden. (Observar que la primera es sobreyectiva y la segunda no lo es).

0.5. Composición de funciones

DEFINICIÓN 0.14. Dadas dos funciones $f : A \to B$ y $g : B \to C$, llamaremos **composición** de f y g a la función $h : A \to C$ tal que para todo $x \in A$ $h(x) = g(f(x))$. La función compuesta se notará $g\,f$ o $g \circ f$.

Tres funciones de la forma $f : A \to B$, $g : B \to C$, $h : A \to C$, se pueden indicar en un mismo diagrama del modo siguiente:

Figura 5.1

Cuando h=gf se dice que el diagrama es **conmutativo**. Obsérvese que esto significa que para todo $a \in A$ se tiene el mismo resultado si "se sigue la flecha"de A a C y se halla $h(a)$, o si se "siguen las flechas"de A a B y de B a C para llegar a $gf(a)$.

Para que esté definida gf es necesario que el codominio de f esté incluido en el dominio de g. En particular, puede estar definida gf y no estarlo fg. Si se consideran funciones $f : A \to A$ y $g : A \to A$ entonces fg y gf están definidas y ambas son funciones de A en A, pero no tienen por qué ser iguales. Por ejemplo, si $f : R \to R$ está dada por $f(x) = x^2$ y $g : R \to R$ por $g(x) = -x$ entonces $gf(x) = -x^2$ y $fg(x) = x^2$.

Dadas tres funciones, $f : A \to B$, $g : B \to C$ y $h : C \to D$, de la definición de composición de funciones resulta que

$h(gf)$ $y\,(hg)\,f$ están definidas y son iguales pues ambas son funciones de A en D y para todo $x \in A$:

$$(hg)(f(x)) = h(g(f(x))) = h((gf)(x)) = (h(gf))(x)$$

Observemos finalmente que para toda función $f : A \to B$ se verifica $id_B\,f = f\,id_A = f$.

En los teoremas que siguen veremos algunas propiedades de las funciones ineyectivas y sobreyectivas que podrían utilizarse para definir inyectividad y sobreyectividad.

TEOREMA 0.3. *Una función $f : A \to B$ con dominio A no vacío es inyectiva si y sólo si existe una función $g : B \to A$ tal que $gf = id_A$.*

DEMOSTRACIÓN. Supongamos que f es inyectiva. Definimos una función $g : B \to A$ como sigue.
a)Si $y \in f(A)$, por ser f inyectiva la imagen inversa de y se reduce a un elemento $x \in A$, entonces tomamos $g(y) = x$.
b) Si $y \in B$ pero $y \notin f(A)$, como A es no vacío, podemos tomar un $x \in A$ cualquiera y definir $g(y) = x$. Entonces $gf = id_A$ pues para todo $x \in A$, según a):

$$(gf)(x) = (g(f(x))) = g(y) = x.$$

Recíprocamente, supongamos que $gf = id_A$. De $f(x) = f(x')$ sigue $x = gf(x) = gf(x') = x'$, luego f es inyectiva. $\square$

Si $gf = id_A$ diremos que g es la **inversa a la izquerda de *f***. Según este teorema, cuando A es no vacío, f es inyectiva si y sólo si tiene inversa a la izquerda. Obsérvese que para todo B la función $f : \phi \to B$ es inyectiva pero si $B \neq \phi$, f no tiene inverso izquerdo pues no existe funciones con dominio $B \neq \phi$ y con codominio ϕ. Por esta razón el teorema 0.3 no vale sin la hipótesis de que el dominio A no sea vacío.

De la demostración del teorema 0.3 resulta que si f es inyectiva pero no sobreyectiva y si A tiene más de un elemento entonces el inverso izquierdo de f no es único. En efecto, en este caso existe **ciertamente** algún y perteneciente a B, $y \notin f(A)$, al cual puede asignarse por imagen según g diferentes elementos de A, resultando en cada caso un inverso izquierdo g diferente.

Dadas dos funciones *h, h'* de C en A, si $f : A \to B$ es inyectiva *fh=fh'* implica *h=h'*, porque si A no es vacío se tiene:

$$g\,(fh) = id_A\, h = h = g\left(fh'\right) = id_A\, h' = h'$$

Si $A = \phi$, de $h : C \to A$ sigue $C = \phi$ luego, también en este caso $h = h' = (\phi, \phi, \phi)$.

Mostraremos que esta propiedad caracteriza las funciones inyectivas, es decir que toda f para la cual *fh=fh'* implica *h=h'* es inyectiva. Supongamos que $f : A \to B$ no sea inyectiva y sean $x, x' \in A$, $x \neq x'$, tales que $f\,(x) = f\,(x')$. Sea C un conjunto formado por un único elemento c, definimos $h : C \to A$ con $h\,(c) = x$ y $h' : C \to A$ con $h'\,(c) = x'$. Entonces $(fh)\,(c) = (fh')\,(c) = f\,(c)$, es decir *fh=fh'*, pero $h \neq h'$, contrario a la hipótesis. Luego f es inyectiva.

TEOREMA 0.4. *Una función $f : A \to B$ es sobreyectiva si y sólo si existe $g : B \to A$ tal que $f \circ g = id_B$.*

Demostración. Supongamos que f es sobreyectiva; entonces para todo $y \in B$ la imagen inversa por f es un conjunto no vacío del cual podemos elegirun elemento x. (Esto requiere usar el axioma de elección). Definimos $g : B \to A$ tomando $g\,(y) = x$. Con esto tenemos para todo $y \in B$: $(fg)\,(y) = f\,(x) = y$ pues x fue elegido en la imagen inversa de y.

Recíprocamente, sea $g : B \to A$ tal que $fg = id_B$. De aquí sigue para todo $y \in B$, $y = id_B\,(y) = f\,(g\,(y))$, quere decir que la imagen por f del ekemento $g\,(y) \in A$ es y. Kuego f es sobreyectiva. □

Una función $g : B \to A$ tal que $fg = id_B$ se llama **inversa a la derecha de *f***. Según el teorema 0.4, f es sobreyectiva si y sólo si tiene inverso derecho. De la demostración de este teorema puede deducirse que si f es sobreyectiva pero no inyectiva entonces tiene más de un inverso derecho.

También se caracterizan las funciones sobreyectivas $f : A \to B$, por la propiedad de que para todo par de funciones *h, h'* de B en C, *hf=h'f* implica *h=h'*. La demostración es similar a la que hicimos para la propiedad análoga de las funciones inyectivas. Si $gf = id_A$ y $fg' = id_B$, entonces:

$$g = g\ id_B = f\left(fg'\right) = (gf)\,g' = id_A\,g' = g'$$

Esto demuestra que si f tiene inversa a la izquierda e inversa a la derecha entonces cualquier inversa a la izquierda es igual a cualquier inversa a la derecha, por lo tanto hay un único inversa a la izquierda, un único inversa a la derecha y son iguales. En tal caso decimos que este es la **inversa de *f*** y lo denotaremos f^{-1}.

La definición de inversa de $f : A \to B$ muestra que f^{-1} asocia a cada $y \in B$ el único $x \in A$ tal que $f(x) = y$. Por ejemplo si *A=R, * $B = \{r \in R; r > 0\}$ y $f : A \to B$ es la función exponencial de base *a¿*0, entonces $f^{-1} : B \to A$ es la función que cada real positivo le asocia su logaritmo de base *a*.

Si f tiene inversa, según los teoremas 0.3 y 0.4, f es biyectiva, por los mismos teoremas tiene inversa a la izquierda e inversa a la derecha que, como acabamos de observar, son iguales entre sí. Si f es biyectiva con dominio vacío, debe ser de codominio vacío, entonces $f = (\phi, \phi, \phi)$. Esta función tiene inverso $f^{-1} = f$, por lo tanto también en este caso vale el enunciado recíproco. Así queda demostrado el corolario siguiente.

COROLARIO 0.5. *Una función $f : A \to B$ es biyectiva si y sólo si existe $g : B \to A$ tal que $gf = id_A$ y $fg = id_B$.*

Es claro que si f^{-1} es el inverso de f entonces f es el inverso de f^{-1}. De aquí y del corolario, resulta que si f es biyectiva, también lo es f^{-1}. Esto también se puede deducir directamente de la definición de función inversa.

CAPíTULO 1

SISTEMAS LINEALES DE ECUACIONES Y MATRICES

1.1. Introducción

Gran parte de lo que estudiaremos en este curso y en su continuación en el segundo semestre se vincula con los sistemas de ecuaciones lineales por lo tanto vale la pena dedicar algún tiempo a presentarlos, discutir sus propiedades y ciertas estrategias para resolverlos. Quienes ya conozcan estos temas tendrán en todo caso una oportunidad para repasarlos.

Un **sistema lineal de** m **ecuaciones con** n **incógnitas** $x_1, x_2, \ldots, x_n$ es un problema del tipo:

$$(S) = \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{cases}$$

donde a_{ij} con $i = 1, \ldots, m$ y $j = 1, \ldots, n$ (los *coeficientes* del sistema) y b_j con $j = 1, \ldots, m$ (los *términos independientes*) son números[1] Diremos que el sistema es $m \times n$ indicando siempre el primer número la cantidad de ecuaciones y el segundo la de incógnitas.

Una solución del sistema (S) es una sucesión 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.

[1]En nuestro cursos los números que consideraremos pertenecen al cuerpo de los reales o de los complejos

Llamaremos **conjunto solución** del sistema (S) y lo notaremos $Sol(S)$ al conjunto de todas las soluciones de (S). Resolver un sistema es determinar su conjunto solución.[2]

Clasificaremos los sistemas de acuerdo a número de soluciones que tenga.

Si A es un conjunto escribiremos $card(A)$ para representa su cardinal. Si el conjunto es finito $card(A)$ indica el número de elementos e infinito en otro caso. Diremos entonces que un sistema es **compatible determinado** si $card(Sol(S)) = 1$ (tiene solución única), **compatible indeterminado** si $card(Sol(S)) > 1$ (tiene solución pero no es única)[3] e **incompatible** si $card(Sol(S)) = 0$ (no tiene ninguna solución).

EJEMPLO 1.1. El ejemplo más sencillo de un sistema lineal es el 1×1

$$ax = b$$

en pocos renglones podemos discutir completamente como es el conjunto solución. Si $a \neq 0$ entonces el sistema es compatible determinado y su única solución es $x = \frac{b}{a} = a^{-1}b$. Si en cambio $a = 0$ hay dos posibilidades: si $b \neq 0$ el sistema es incompatible y si $b = 0$ trivialmente el sistema es compatible indeterminado y cualquier número es solución del mismo.

Naturalmente este problema es extremadamente sencillo pero como acontece muchas veces las ideas que aquí se utilizan convenientemente adaptadas resultaran útiles en casos más generales y complicados. ■

EJEMPLO 1.2.

$$(S)\begin{cases} x = 1 \\ y = 3 \\ z = 2 \end{cases}$$

es un sistema 3×3. Trivialmente se observa que $Sol(S) = \{(1, 2, 3)\}$.

■

[2] En general, salvo que se indique lo contrario, las soluciones se consideraran en el mismo conjunto de números que los coeficientes

[3] Veremos más adelante que si un sistema lineal admite más de una solución entonces necesariamente debe admitir infinitas

EJEMPLO 1.3.

$$(S)\begin{cases} x+y+2z=8 \\ \quad\ 2y+z=8 \\ \qquad\quad z=2 \end{cases}$$

es también un sistema 3×3. Aquí con algo más de trabajo también se puede deducir fácilmente el conjunto solución. Para esto basta observar que de abajo hacia arriba cada ecuación tiene exactamente una incógnita más que la siguiente, por este motivo se dice que el sistema esta "escalerizado". De la última ecuación se despeja

$$\boxed{\text{z=2}}$$

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

$$2y+2=8 \Longrightarrow 2y=6 \Longrightarrow \boxed{\text{y=3}}$$

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

$$x+3+2(2)=8 \Longrightarrow \boxed{\text{x=1}}$$

En consecuencia $Sol(S)=\{(1,3,2)\}$.
Este procedimiento de despejar una incógnita en una ecuación y sustituirla en otra de arriba se denomina **sustitución hacia atrás** ■

Observemos que en los ejemplos 1.2 y 1.3 los conjuntos solución son iguales, cuando esto ocurra diremos que los respectivos sistemas son **equivalentes.**

EJEMPLO 1.4.

$$(S)\begin{cases} x+y+2z=8 \\ x-y+z=0 \\ x+y+z=6 \end{cases}$$

Las cosas aquí no son tan claras. Una estrategia para resolverlo, inspirada en los ejemplos anteriores podría ser, la de sustituir el sistema (S) por otro, (S') equivalente y "escalerizado" de modo de poder resolver (S') (y por ende (S) ya que al ser equivalentes los conjuntos solución de ambos sistemas coinciden) con el procedimiento utilizado en el ejemplo 1.3. Analicemos un poco estas ideas antes de retomar la resolución concreta del ejemplo. ■

La pregunta que naturalmente surge es ¿cómo hacer para sustituir un sistema por otro que resulte equivalente?. Para responderla introducimos la siguiente

DEFINICIÓN 1.1. Llamaremos **transformaciones elementales** a cualquiera de las siguientes operaciones efectuadas a las ecuaciones de un sistema lineal:

1. Intercambiar de lugar dos ecuaciones. ($F_i \leftrightarrow F_j$)
2. Multiplicar una ecuación por un número $\alpha \neq 0$. (αF_i)
3. Sumar a una ecuación un múltiplo de otra. ($F_i + \alpha F_j$)

Entre paréntesis se indica la notación que utilizaremos para especificar cada operación en los ejemplos. La ecuación i-esima será notada con F_i.

La siguiente proposición responde la pregunta formulada arriba.

PROPOSICIÓN 1.1. *Si en un sistema lineal de ecuaciones se efectúan una* ***transformación elemental*** *el sistema resultante es equivalente al original.*

DEMOSTRACIÓN. La demostración es muy sencilla y la haremos sólo en el caso que la transformación elemental sea del tipo 3. El resto queda a cargo del lector.
Supongamos que el sistema original sea

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

Supongamos que sumamos a la ecuación i un múltiplo de la ecuación j. El sistema resultante es

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

Para probar que ambos son equivalentes deberemos ver que $Sol(S) = Sol(S')$. Como se trata de mostrar la igualdad de dos conjuntos tendremos que demostrar que están mutuamente incluidos.
Veamos primero que $Sol(S) \subset Sol(S')$. Sea $(\alpha_1, \alpha_2, \cdots, \alpha_n) \in Sol(S)$ es claro que $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ satisface todas las ecuaciones de (S') (pues son las mismas que las de (S)) salvo tal vez la i-esima. Como $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ debe verificar la i-esima y j-esima ecuación de (S) se tiene que

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + \alpha_{in}x_n = b_i \text{ y } a_{j1}\alpha_1 + a_{j2}\alpha_2 + \cdots + a_{jn}\alpha_n = b_j$$

por lo tanto

$$(a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n) + \beta(a_{j1}\alpha_1 + a_{j2}\alpha_2 + \cdots + a_{jn}\alpha_n) = b_i + \beta b_j$$

de donde $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in Sol(S')$.
Finalmente probemos que $Sol(S') \subset Sol(S)$. Consideremos ahora

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \in Sol(S').$$

Igual que antes es claro $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ debe verificar todas las ecuaciones de (S) salvo tal vez la i-esima pero como

$$(a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n) + \beta(a_{j1}\alpha_1 + a_{j2}\alpha_2 + \cdots + a_{jn}\alpha_n) = b_i + \beta b_j$$

y

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

se deduce que

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

con lo cual $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in Sol(S)$ y la prueba esta concluida. $\square$

Naturalmente las cuestiones que se imponen ahora son las de saber si cualquier sistema puede ser "escalerizado" mediante transformaciones elementales y en caso afirmativo determinar un método o algoritmo que permita hacerlo. Volvamos a analizar el ejemplo 1.4.

EJEMPLO 1.4 (continuación). Para que el sistema quede "escalerizado en la segunda y tercera ecuación no debe aparecer la incógnita x entonces dejemos la primera ecuación igual y realicemos transformaciones elementales en la segunda y tercera. La idea es combinar ambas con la primera de modo que los coeficientes de la incógnita x se anulen. Como el coeficiente de x en las tres ecuaciones es 1 basta en ambos casos restar la primera.

$$\left\{\begin{array}{r} x+y+2z=8 \\ x-y+z=0 \\ x+y+z=6 \end{array}\right. \sim \left\{\begin{array}{rl} x+y+2z=8 & \\ -2y-z=-8 & \longleftarrow F_2-F_1 \\ -z=-2 & \longleftarrow F_3-F_1 \end{array}\right.$$

El sistema esta ya "escalerizado", si multiplicamos las dos últimas ecuaciones por -1 se obtiene el sistema del ejemplo 1.3 por lo cual $Sol(S) = \{(1,3,2)\}$ $\blacksquare$

1.2. Matrices

Al trabajar con un sistema de ecuaciones la representación de las incógnitas (x, y, z o $x_1, \ldots, x_n$, etc.) no juega ningún papel y puede usarse cualquiera; de hecho para determinar 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 mantener el orden que ocupan (a qué ecuación pertenecen y a qué incógnita multiplican). Por esto resulta conveniente dar la siguiente

DEFINICIÓN 1.2. Llamaremos **matriz** A de m filas por n columnas $(m \times n)$ de entradas a_{ij} a un ordenamiento rectangular de números

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

Notación: En general la matriz A con entradas a_{ij} la indicaremos por $A = ((a_{ij}))_{j=1,\dots,n}^{i=1,\dots,m}$ o más brevemente $A = ((a_{ij}))$ cuando las dimensiones estén claras. El primer índice i indica la fila y el segundo j la columna a la que pertenece la entrada.
El conjunto de todas las matrices $m \times n$ lo indicaremos como $\mathcal{M}_{m\times n}$.
Llamaremos **matriz columna (fila)** a una matriz $X \in \mathcal{M}_{m\times 1}(\mathcal{M}_{1\times n})$.

Diremos que dos matrices son iguales si tienen el mismo tamaño y los mismas entradas en las mismas posiciones. Dicho de otro mondo si A y B son dos matrices $m \times n$ entonces $A = B \Leftrightarrow a_{ij} = b_{ij} \;\; \forall i = 1, \dots, m, \; j = 1, \dots, n$

EJEMPLO 1.5. Sea

$$A = \begin{pmatrix} 1 & 2 \\ \sqrt{2} & 1 \end{pmatrix} \text{ y } B = \begin{pmatrix} 2 & 1 \\ \sqrt{2} & 1 \end{pmatrix}$$

Entonces $A \neq B$ pues $a_{11} = 1 \neq b_{11} = 2$ ■

OBSERVACIÓN 1.2. El lector atento notará que nuestra definición de matriz si bien es clara no es del todo formal. La frase "ordenamiento rectangular.$^{\text{es}}$ intuitiva pero no está definida con precisión, una definición formal sería:

Una **matriz** A *de* m *filas por* n *columnas es una función* $A : \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{R}$.
El lector verificará que esta definición es coherente, chequeando por ejemplo, el criterio de igualdad de matrices.

DEFINICIÓN 1.3. Dado un sistema lineal de ecuaciones

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

llamaremos **matriz del sistema** a la matriz $m \times n$ A de coeficientes del mismo, esto es:

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

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

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

Diremos que dos matrices son equivalente si los respectivos sistemas lo son

La definición anterior introduce una notación más compacta (matricial) para un sistema de ecuaciones. Veamos algunos ejemplos.

EJEMPLO 1.6. Consideremos el sistema

$$(S)\left\{\begin{array}{r} x + y + 2z = 9 \\ 3x + 6y - 5z = 0 \\ 2x + 4y - 3z = 1 \end{array}\right.$$

entonces la matriz del sistema y la matriz ampliada son:

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 6 & -5 \\ 2 & 4 & -3 \end{pmatrix} \text{ y } A|b = \left(\begin{array}{rrr|r} 1 & 1 & 2 & 9 \\ 3 & 6 & -5 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right)$$

■

1.3. El método de escalerización

Resulta natural ahora analizar , cómo se aplica el procedimiento de "escalerización" del ejemplo 1.4 a la matriz ampliada de un sistema. Las transformaciones elementales deben aplicarse sobre las filas de la matriz.

EJEMPLO 1.7. Intentemos "escalerizar.[el] sistema del ejemplo 1.6

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 3 & 6 & -5 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 3 & -11 & -27 \\ 0 & 2 & -7 & -17 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 - 3F_1 \\ \longleftarrow F_3 - 2F_1 \end{array} \sim$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 3 & -11 & -27 \\ 0 & 0 & \frac{1}{3} & 1 \end{array}\right) \begin{array}{l} \\ \\ \longleftarrow F_3 - \frac{2}{3}F_2 \end{array} \sim$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 3 & -11 & -27 \\ 0 & 0 & 1 & 3 \end{array}\right) \begin{array}{l} \\ \\ \longleftarrow 3F_3 \end{array}$$

Por lo tanto el sistema (S) es equivalente al sistema

$$\begin{cases} x + y + 2z = 9 \\ 3y - 11z = -27 \\ z = 3 \end{cases}$$

Ahora podemos proceder a realizar sustitución hacia atrás. De la última ecuación se tiene

$$\boxed{\text{z=3}}$$

Con este valor sustituimos en la segunda y tenemos

$$3y - 33 = -27 \Longrightarrow 3y = 6 \Longrightarrow \boxed{\text{y=2}}$$

y finalmente sustituyendo en la primera

$$x + 2 + 6 = 9 \Longrightarrow \boxed{\text{x=1}}$$

En consecuencia el sistema es compatible determinado y su única solución es la 3-upla $(1, 2, 3)$. ■

Hasta ahora no hemos sido precisos para definir a que le llamamos un sistema o matriz escalerizada

DEFINICIÓN 1.4. 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.

Llamaremos **matriz escalerizada reducida** a una matriz escalerizada que además cumple las siguientes condiciones:

1. La primera entrada no nula de una fila es 1.
2. La columna correspondiente a la primera entrada no nula de una fila tiene el resto de las entradas todas nulas.

Diremos que un **sistema está escalerizado** si su matriz ampliada lo está. **Escalerizar un sistema (o matriz)** es encontrar un sistema (matriz) escalerizado equivalente. Si A es una matriz y E es una matriz escalerizada equivalente a A diremos que E es una **forma escalerizada** de A

EJEMPLO 1.8.

$$(1.8.a)\quad \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \qquad (1.8.b)\quad \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(1.8.c)\quad \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (1.8.d)\quad \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Los ejemplos a y b son matrices escalerizadas y los ejemplos c y d son escalerizadas reducidas. ■

En el ejemplo 1.6 se aprecia que una vez obtenida la forma escalerizada el método de sustitución hacia atrás permite resolver el sistema. Cabe entonces preguntarse el motivo para introducir la definición de matriz escalerizada reducida. En el siguiente ejemplo se aclara en parte esto.

EJEMPLO 1.9. Resolver el sistema:

$$(S)\begin{cases} 2x+4y+2z=-2 \\ x+y+2z=0 \\ -x+y-2z=2 \end{cases}$$

La matriz ampliada del sistema es:

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

Procedamos a escalerizarla, para esto en primer lugar obtengamos una matriz equivalente pero con ceros en la segunda y tercera posición de la primera columna. Dejamos la primera fila fija y la utilizamos como "pivot" para combinar con las otras dos.

$$\left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 1 & 1 & 2 & 0 \\ -1 & 1 & -2 & 2 \end{array}\right) \sim \left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 0 & -1 & 1 & 1 \\ 0 & 3 & -1 & 1 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2-\frac{1}{2}F_1 \\ \longleftarrow F_3+\frac{1}{2}F_1 \end{array}$$

Ahora la primera y segunda fila tiene el aspecto adecuado. El segundo paso consiste en sustituir la tercera fila por otra con dos ceros en los primeros lugares. Para esto combinamos convenientemente tercera con la segunda fila.

$$\left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 0 & -1 & 1 & 1 \\ 0 & 3 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 2 & 4 \end{array}\right) \begin{array}{l} \\ \\ \longleftarrow F_3+3F_2 \end{array} \tag{1.1}$$

Nótese que al dar el segundo paso la primera fila ya no interviene y se procede de manera análoga al primer paso pero utilizado la segunda fila como fila "pivot".

Ahora que obtuvimos una forma escalerizada de la matriz original podríamos proceder con la sustitución hacia atrás pero en lugar de esto intentemos obtener la forma reducida para esto fijamos la tercera fila como "pivot" y la combinamos con la primera y segunda de modo que las entradas en las posiciones $1,3$ y $2,3$ se anulen

$$\left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 2 & 4 \end{array}\right) \sim \left(\begin{array}{rrr|r} 2 & 4 & 0 & -6 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 4 \end{array}\right) \begin{array}{l} \longleftarrow F_1 - F_3 \\ \longleftarrow F_2 - \frac{1}{2}F_3 \\ \end{array}$$

Combinando la segunda fila con la primera obtenemos un cero en la posición $1,2$

$$\left(\begin{array}{rrr|r} 2 & 4 & 0 & -6 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 4 \end{array}\right) \sim \left(\begin{array}{rrr|r} 2 & 0 & 0 & -10 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 4 \end{array}\right) \begin{array}{l} \longleftarrow F_1 + 4F_2 \\ \\ \end{array}$$

Para obtener la forma reducida solo resta obtener unos en las entradas no nulas de cada fila

$$\left(\begin{array}{rrr|r} 2 & 0 & 0 & -10 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 4 \end{array}\right) \sim \left(\begin{array}{rrr|r} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right) \begin{array}{l} \longleftarrow \frac{1}{2}F_1 \\ \longleftarrow -F_2 \\ \longleftarrow \frac{1}{2}F_3 \end{array}$$

El sistema resultante, equivalente a (S) es

$$(S')\begin{cases} x = -5 \\ y = 1 \\ z = 2 \end{cases}$$

y por lo tanto el sistema es compatible determinado con $Sol(S) = \{(-5, 1, 2)\}$.

Obsérvese que en 1.1 la escalera que se obtuvo tiene todos sus "escalones" de ancho una columna. Esto implica que en el correspondiente sistema cada ecuación tenga exactamente una incógnita más que la inmediata inferior. ■

El proceso de obtener la forma reducida sustituye entonces al algoritmo de sustitución hacia atrás, además el lector puede observar que la forma escalerizada de una matriz no es única (para obtener otra basta sustituir una fila por una combinación de ella con una fila inferior) pero la reducida si lo es.

EJEMPLO 1.10. Resolver el sistema:

$$(S)\begin{cases} 2x + 4y + 2z = -2 \\ x + y + 2z = 0 \\ -x + y - 4z = 2 \end{cases}$$

La matriz ampliada del sistema es:

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

Procedamos como en el ejemplo anterior

$$\left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 1 & 1 & 2 & 0 \\ -1 & 1 & -4 & 2 \end{array}\right) \sim \left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 0 & -1 & 1 & 1 \\ 0 & 3 & -3 & 1 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 - \frac{1}{2}F_1 \\ \longleftarrow F_3 + \frac{1}{2}F_1 \end{array} \sim$$

$$\left(\begin{array}{rrr|r} 2 & 4 & 2 & -2 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 4 \end{array}\right) \begin{array}{l} \\ \\ \longleftarrow F_3 + 3F_2 \end{array}$$

A diferencia del ejemplo precedente la forma escalerizada de la matriz ampliada tiene un escalón más que la de la matriz del sistema, por lo tanto la última ecuación del sistema escalerizado es

$$0 = 4$$

y consecuentemente el sistema resulta incompatible pues ningún valor de x, y y z la satisface. ■

EJEMPLO 1.11. Resolver el sistema:

$$(S)\begin{cases} x_1 + 2x_2 + x_3 + x_5 + x_6 = 1 \\ -x_1 - 2x_2 + x_3 + x_4 - x_6 = 0 \\ x_1 + 2x_2 + x_3 + 2x_4 + 5x_5 + 3x_6 = 1 \\ x_1 + 2x_2 + x_3 + x_5 + 3x_6 = 3 \\ x_1 + 2x_2 + x_3 + 4x_4 + 9x_5 + 3x_6 = 1 \end{cases}$$

La matriz del sistema ampliada es:

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

Como antes el proceso comienza fijando la primera fila como pivot y utilizando su primera entrada para lograr anular el resto de la primera columna.

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 & 0 \\ 1 & 2 & 1 & 2 & 5 & 3 & 1 \\ 1 & 2 & 1 & 0 & 1 & 3 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 & -1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & -1 & -3 & -2 & 1 \\ 0 & 0 & 2 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & 4 & 8 & 2 & 0 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 + F_1 \\ \longleftarrow F_3 + 2F_1 \\ \longleftarrow F_4 - F_1 \\ \longleftarrow F_5 - F_1 \end{array}$$

En los ejemplos anteriores en el segundo paso, se utilizaba la entrada $2,2$ para conseguir ceros en el resto de la 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. Utilizamos entonces la entrada de la posición $2,3$ para generar ceros en el resto de la columna tres.

$$
(1.2)\qquad \left(\begin{array}{cccccc|c}
1 & 2 & 1 & 0 & 1 & 1 & 1\\
0 & 0 & 2 & 1 & 1 & 0 & 1\\
0 & 0 & 2 & -1 & -3 & -2 & 1\\
0 & 0 & 2 & 1 & 1 & -2 & -1\\
0 & 0 & 0 & 4 & 8 & 2 & 0
\end{array}\right) \sim
$$

$$
\sim \left(\begin{array}{cccccc|c}
1 & 2 & 1 & 0 & 1 & 1 & 1\\
0 & 0 & 2 & 1 & 1 & 0 & 1\\
0 & 0 & 0 & -2 & -4 & -2 & 0\\
0 & 0 & 0 & 0 & 0 & -2 & -2\\
0 & 0 & 0 & 4 & 8 & 2 & 0
\end{array}\right) \begin{array}{l} \\ \\ \longleftarrow F_3 - F_2 \\ \longleftarrow F_4 - F_2 \\ \longleftarrow F_5 \end{array}
$$

En el tercer paso escalerizaremos la columna siguiente (la cuarta)

$$
\left(\begin{array}{cccccc|c}
1 & 2 & 1 & 0 & 1 & 1 & 1\\
0 & 0 & 2 & 1 & 1 & 0 & 1\\
0 & 0 & 0 & -2 & -4 & -2 & 0\\
0 & 0 & 0 & 0 & 0 & -2 & -2\\
0 & 0 & 0 & 4 & 8 & 2 & 0
\end{array}\right) \sim
$$

$$
\sim \left(\begin{array}{cccccc|c}
1 & 2 & 1 & 0 & 1 & 1 & 1\\
0 & 0 & 2 & 1 & 1 & 0 & 1\\
0 & 0 & 0 & -2 & -4 & -2 & 0\\
0 & 0 & 0 & 0 & 0 & -2 & -2\\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5 + 2F_3 \end{array}
$$

Ya tenemos la forma escalerizada, pero como en el ejemplo 1.9 continuemos hasta obtener la forma reducida

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

$$\sim \left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & -2 & -4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \longleftarrow F_1 - \frac{1}{2}F_4 \\ \\ \longleftarrow F_3 - 2F_4 \\ \\ \\ \end{array} \sim$$

$$\sim \left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & -2 & -4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 + \frac{1}{2}F_3 \\ \\ \\ \\ \end{array} \sim$$

$$\sim \left(\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & \frac{3}{2} & 0 & -1 \\ 0 & 0 & 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & -2 & -4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \longleftarrow F_1 - \frac{1}{2}F_2 \\ \\ \\ \\ \\ \end{array} \sim \tag{1.3}$$

$$\sim \left(\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & \frac{3}{2} & 0 & -1 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \longleftarrow \frac{1}{2}F_2 \\ \longleftarrow -\frac{1}{2}F_3 \\ \longleftarrow -\frac{1}{2}F_4 \\ \\ \end{array}$$

El sistema de ecuaciones asociado queda entonces

$$(S')\left\{\begin{array}{r} x_1 + 2x_2 + \frac{3}{2}x_5 = -1 \\ x_3 - \frac{1}{2}x_5 = 1 \\ x_4 + 2x_5 = -1 \\ x_6 = 1 \end{array}\right.$$

Despejando se tiene

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

Las incógnitas x_2 y x_5 no pueden determinarse, las otras se obtienen como función de ellas. Cada valor real elegido libremente de x_2 y x_5 determina una solución del sistema (S) Por ejemplo si $x_2 = 1$ 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 $\left(-\frac{5}{2}, 0, \frac{1}{2}, -3, 1, 1\right)$. El sistema es entonces compatible indeterminado y en virtud de que dos de sus incógnitas pueden elegirse libremente diremos que tiene **dos grados de libertad**. El conjunto solución es

$$\left\{\left(-2x_2 - \frac{3}{2}x_5 - 1, x_2, -\frac{1}{2}x_5 + 1, -2x_5 - 1, x_5, 1\right), x_2 \in \mathbb{R}, x_5 \in \mathbb{R}\right\}$$

Vale la pena observar que al igual que en el ejemplo 1.9 el número de escalones de la matriz ampliada y el de la matriz del sistema coinciden. Por esta razón el sistema resulta compatible. Sin embargo, aquí la "escalera" presenta "descansos" o escalones largos (el primero y el tercero). Esto implica que en el sistema escalerizado algunas ecuaciones tienen por lo menos dos incógnitas más que la ecuación inmediata inferior, por esto resulta imposible determinar todas las incógnitas. ■

Los ejemplos anteriores parecen indicar que cualquier matriz y por ende cualquier sistema puede ser escalerizado. Establezcamos este hecho en el siguiente resultado

PROPOSICIÓN 1.3. *Toda matriz puede ser transformada en una matriz escalerizada (o escalerizada reducida) mediante una cantidad finita de transformaciones elementales*

Demostración. La prueba es algorítmica[4] . Indicaremos esquemáticamente el procedimiento, los detalles quedan a cargo del lector.
Sea $A = ((a_{ij}))$ una matriz $m \times n$ cualquiera. Indiquemos por $A_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ su fila i-esima.
El objetivo del primer paso es obtener una matriz con una primera columna que tenga nulas todas sus entradas excepto la primera. Si toda la primera columna es nula se pasa a la segunda y así sucesivamente hasta tener una primera columna con alguna entrada distinta de cero. La idea para anular el resto de las entradas de la columna es ubicar la entrada no nula (a la que llamremos **pivot**) en la primera fila y utilizar esta para combinar convenientemente con cada fila a los efectos de lograr ceros en todas las posiciones.
Para simplificar la notación aquí supondremos que ya la primera columna tiene alguna entrada no nula. Con la suposición indicada comenzamos reconociendo si $a_{11} \neq 0$. Si $a_{11} = 0$ cambiamos la primera fila por otra cuya primera entrada sea no nula.[5] Una vez que el pivot esta ubicado realizamos

[4]Un algoritmo es, groseramente hablando, un conjunto de acciones expresadas en un lenguaje formal y sin ambigüedades que pueden ser comprendidos y ejecutados por una persona o una máquina.
Este algoritmo es conocido con el nombre de Método de Gauss.

[5]Por razones numéricas y para evitar errores de redondeo al resolver sistemas con ayuda de la 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**.

el primer paso de la escalerización:

$$\begin{pmatrix} A_1 & \text{------} \\ A_2 & \text{------} \\ \vdots & \vdots \\ A_i & \text{------} \\ \vdots & \vdots \\ A_m & \text{------} \end{pmatrix} \sim \begin{pmatrix} A_1 & \text{------} \\ A'_2 & \text{------} \\ \vdots & \vdots \\ A'_i & \text{------} \\ \vdots & \vdots \\ A'_m & \text{------} \end{pmatrix} \begin{array}{l} \\ \longleftarrow A_2 - \frac{a_{21}}{a_{11}} A_1 \\ \vdots \\ \longleftarrow A_i - \frac{a_{i1}}{a_{11}} A_1 \\ \vdots \\ \longleftarrow A_m - \frac{a_{m1}}{a_{11}} A_1 \end{array}$$

La matriz que obtuvimos tiene el siguiente aspecto:

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

En el segundo paso el objetivo es obtener una nueva matriz que tenga las entradas de la segunda columna todas nulas salvo tal vez las dos primeras. Para eso repetimos el procedimiento del paso uno a la sub-matriz que aparece recuadrada.

De nuevo el primer paso es determinar el segundo pivot para esto chequeamos que $a'_{22} \neq 0$ de lo contrario cambiamos la segunda fila por otra (de abajo) con segunda entrada no nula. Podría ocurrir que todas las entradas de la segunda fila fuesen nulas en ese caso la segunda columna ya tiene el aspecto adecuado por lo cual el procedimiento se continúa ubicando un pivot en al posición $2, 3$ (ver (1.2)en el ejemplo 1.11). Supongamos, para fijar ideas, que $a'_{22} \neq 0$ entonces el segundo paso de la escalerización es:

$$\begin{pmatrix} A_1 & \text{———} \\ A'_2 & \text{———} \\ A'_3 & \text{———} \\ \vdots & \vdots \\ A'_i & \text{———} \\ \vdots & \vdots \\ A'_m & \text{———} \end{pmatrix} \sim \begin{pmatrix} A_1 & \text{———} \\ A'_2 & \text{———} \\ A''_3 & \text{———} \\ \vdots & \vdots \\ A''_i & \text{———} \\ \vdots & \vdots \\ A''_m & \text{———} \end{pmatrix} \begin{array}{c} \\ \\ \longleftarrow A'_3 - \frac{a_{32}}{a_{22}} A'_2 \\ \vdots \\ \longleftarrow A'_i - \frac{a_{i2}}{a_{22}} A'_2 \\ \vdots \\ \longleftarrow A'_m - \frac{a_{m2}}{a_{22}} A'_2 \end{array}$$

La matriz obtenida en el paso dos tiene entonces el aspecto:

$$\left(\begin{array}{cc|ccc|} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a'_{22} & a'_{23} & \dots & a'_{2n} \\ \cline{3-5} 0 & 0 & a''_{33} & \dots & a''_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a''_{m3} & \dots & a''_{mn} \\ \cline{3-5} \end{array}\right)$$

Repetimos nuevamente los pasos previos aplicados a la sub-matriz del recuadro. El procedimiento culmina cuando se llega a la última fila o columna.

Para obtener la matriz escalerizada reducida a partir de la escalerizada, basta con repetir el procedimiento pero tomando como pivots la entrada que ocupa el lugar inferior derecho de la matriz y yendo hacia arriba. $\square$

COROLARIO 1.4. *Todo sistema lineal es equivalente a uno escalerizado.*

1.4. Teorema de Rouche-Frobenius

En las discusiones que realizamos de los ejemplos 1.9, 1.10 y 1.11 se apreciaba cierta relación entre el número de escalones de la forma escalerizada y la cantidad de soluciones del sistema, en esta sección establecemos con precisión esta relación describiendo completamente el conjunto solución de un sistema lineal de ecuaciones en termino de la cantidad de "escalones" del mismo.

DEFINICIÓN 1.5. Si E es una matriz escalerizada el **número de escalones** de E es la cantidad de filas no nulas.

OBSERVACIÓN 1.5. Si E es una matriz escalerizada y E' es una matriz reducida equivalente el número de escalones de ambas es el mismo.

OBSERVACIÓN 1.6. Ya hemos observado que la forma escalonada de una matriz no es única sin embargo (aunque no lo hemos probado todavía) la forma reducida si lo es. De este hecho y de la observación precedente se deduce que la cantidad de escalones de todas las formas escalerizadas de una misma matriz coinciden.

TEOREMA 1.7 (Rouche-Frobenius). *Sea (S) un sistema lineal de m ecuaciones con n incógnitas con matriz ampliada $A|b$. Sea p el número de escalones de A y p' el número de escalones de $A|b$ entonces*

1. *(S) es compatible si y solo si $p = p'$*
2. *Si (S) es compatible entonces*
   - *a) (S) es determinado si y solo si $p = n$.*
   - *b) (S) es indeterminado si y solo si $p < n$.*

Demostración. Comencemos probando los recíprocos de las tres afirmaciones. Observemos en primer lugar que como A tiene n columnas y cada escalón ocupa por lo menos una columna deducimos que para cualquier sistemas se cumple que

$$p \leq n \tag{1.4}$$

Supongamos que $p = p'$. De (1.4) se deduce que hay dos posibilidades $p' = p = n$ o $p' = p < n$. Analicemos el primer caso. La forma reducida debe tener tantos escalones como columnas por lo tanto debe ser:

$$\left(\begin{array}{cccccc|c}
1 & 0 & 0 & \dots & \dots & 0 & c_1 \\
0 & 1 & 0 & \dots & \dots & 0 & c_2 \\
0 & 0 & 1 & \ddots & & \vdots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
\vdots & \vdots & & \ddots & 1 & 0 & \vdots \\
0 & 0 & \dots & \dots & 0 & 1 & c_n \\
\hline
0 & 0 & \dots & \dots & 0 & 0 & 0
\end{array}\right)$$

Donde la columna de entradas c_i representa el termino independiente luego de efectuado el proceso de escalerización. El sistema correspondiente es

$$\begin{cases} x_1 = c_1 \\ x_2 = c_2 \\ \vdots \qquad \vdots \\ x_n = c_n \end{cases}$$

el cual es compatible determinado.

Veamos ahora el segundo caso, $p' = p < n$. Como la forma escalerizada tiene más columnas que escalones necesariamente algún escalón debe tener “descanso” y por lo tanto la forma reducida debe ser

$$
\left(\begin{array}{cccccccccccccccc|c}
1 & 0 & \star & \dots & \star & 0 & \dots & 0 & 0 & \star & \dots & \star & 0 & \dots & 0 & c_1 \\
0 & 1 & \star & \dots & \star & 0 & \dots & 0 & 0 & \star & \dots & \star & 0 & \dots & 0 & c_2 \\
\vdots & 0 & \dots & \dots & 0 & 1 & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots \\
\vdots & \vdots & \dots & \dots & \vdots & \ddots & \ddots & 0 & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots \\
\vdots & \vdots & \dots & \dots & \vdots & \dots & \ddots & 1 & 0 & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots \\
\vdots & \vdots & \dots & \dots & \vdots & \dots & \dots & 0 & 1 & \star & \dots & \star & 0 & \dots & \vdots & c_s \\
0 & 0 & \dots & \dots & \vdots & \dots & \dots & \vdots & 0 & 0 & \dots & 0 & 1 & \dots & \vdots & c_{s+1} \\
\vdots & \vdots & \dots & \dots & \vdots & \dots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \ddots & 0 & \vdots \\
0 & 0 & \dots & \dots & 0 & \dots & \dots & \dots & \vdots & 0 & \dots & 0 & 0 & \dots & 1 & c_p \\
0 & 0 & \dots & \dots & 0 & \dots & \dots & \dots & \vdots & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \dots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots \\
0 & 0 & \dots & \dots & 0 & \dots & \dots & \dots & \vdots & 0 & \dots & 0 & 0 & \dots & 0 & 0
\end{array}\right)
$$

donde las zonas de estrellas representan submatrices con entradas arbitrarias.

Para analizar la solución de este sistema intentemos aplicar el algoritmo de sustitución hacia atrás. Supongamos que s es la primera fila (comenzando desde abajo) con descanso (escalón de ancho mayor que una columna). Las ecuaciones $s+1, s+2, \dots, p$ son entonces

$$
\begin{array}{rcl}
x_{s+1} & = & c_{s+1} \\
x_{s+2} & = & c_{s+2} \\
 & \vdots & \\
x_p & = & c_p
\end{array}
$$

la ecuación s es

$$x_h + a_{s\,h+1}x_{h+1} + \dots + a_{ss}x_s = c_s$$

despejando se tiene

$$x_h = c_s - (a_{s\,h+1}x_{h+1} + \dots + a_{ss}x_s)$$

La incógnita x_h no puede entonces determinarse sino en función de los valores de $x_{h+1}, \ldots, x_s$. Fijemos valores arbitrarios para estas últimas $s-h$ incógnitas, x_h quedará determinado y con estos valores podemos ahora sustituir en las ecuación $s-1, \ldots, 1$ y continuar el proceso de sustitución hacia atrás. Esto demuestra que el sistema es compatible, pues hay soluciones pero si cambiamos la elección hecha para los valores de $x_{h+1}, \ldots, x_s$ se obtiene otro valor para x_h y consecuentemente otra solución con lo cual el sistema resulta compatible indeterminado. De hecho como las variables $x_{h+1}, \ldots, x_s$ se pueden elegir arbitrariamente se obtienen infinitas soluciones. Obsérvese que estas elecciones arbitrarias se pueden hacer sólo en los descansos y que el número de variables que pueden elegirse libremente es $n-p$ (la suma de los anchos de los descansos es igual al número de columnas menos el de escalones).

Demostremos ahora que si (S) compatible entonces $p' = p$. Para probarlo demostraremos su contra recíproco[6] y para esto basta con ver que $p < p'$ implica (S) incompatible pues en cualquier sistema $A|b$ tiene una columna más que A y por lo tanto se cumple que $p \leq p'$
Si $p < p'$, la última columna de la forma reducida de la matriz ampliada debe tener un escalón por lo tanto la forma es:

$$\left(\begin{array}{ccccccccccccc}
1 & 0 & \star & \ldots & \star & 0 & \ldots & 0 & \star & \star & 0 \\
0 & 1 & \star & \ldots & \star & 0 & \ldots & 0 & \star & \star & 0 \\
0 & 0 & 0 & \ldots & 0 & 1 & \ldots & 0 & \star & \star & 0 \\
\vdots & \vdots & \vdots & \ldots & \vdots & & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 1 & \star & \star & 0 \\
0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & 0 & 0
\end{array}\right)$$

En consecuencia la ultima ecuación del sistema es

$$0 = 1$$

[6]Recordemos que un resultado de lógica elemental establece que $A \Rightarrow B$ si y solo si negación de $B \Rightarrow$ negación de A

de donde resulta que (S) debe ser incompatible.
Si S es compatible determinado debe ser $p = n$ pues si $p < n$ ya hemos visto que (S) resultaría indeterminado.
De manera análoga se razona para probar que S compatible indeterminado implica $p < n$ $\square$

En el transcurso de la demostración se observo que si $p = p' < n$ el sistema es indeterminado y se pueden elegir $n - p$ variables libremente quedando las otras determinadas por estas. Por este motivo diremos que el sistema es indeterminado con $n - p$ **grados de libertad**.

1.5. Sistemas homogéneos

Un sistema con todas las entradas del término independiente nulas se denomina **sistema homogéneo**. En virtud del importante rol que estos sistemas desempeñan vale la pena discutir alguna de sus propiedades.

PROPOSICIÓN 1.8. *Un sistema lineal homogéneo es siempre compatible.*

DEMOSTRACIÓN. En efecto como el término independiente de un sistema homogéneo tiene todas sus entradas nulas debe ser $p = p'$ y por lo tanto el teorema de Rouche-Frobenius asegura la compatibilidad. $\square$

Otra manera más directa de probar la afirmación anterior es observar que la solución $x_1 = 0, \dots, x_n = 0$ a la que llamaremos solución trivial verifica cualquier sistema homogéneo. Naturalmente los sistemas homogéneos determinados únicamente admiten solución trivial. Veamos con un ejemplo que ocurre en el caso de un sistema homogéneo indeterminado

EJEMPLO 1.12. Resolver el sistema

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

La matriz ampliada es:

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

Al escalerizarla se obtiene la matriz

$$(S)\ \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

en consecuencia el sistema asociado es

$$\begin{cases} x + y - z = 0 \\ \quad\;\; y - z = 0 \end{cases}$$

La solución es entonces

$$Sol(S) = \{(0, z, z) \text{ con } z \in \mathbb{R}\} = \{z(0, 1, 1) \text{ con } z \in \mathbb{R}\}.$$

Se puede observar que todas las soluciones se pueden obtener como combinación lineal de ciertas soluciones fijas. ■

OBSERVACIÓN 1.9. Sea A una matriz $m \times n$ con $n > m$, el sistema homogéneo asociado tiene más incógnitas que ecuaciones y resulta indeterminado. En efecto, si E es una forma escalerizada asociada a A, es evidente que a lo sumo puede tener m escalones y por lo tanto el teorema de Rouche-Frobenius implica la afirmación.

1.6. Una interpretación geométrica

El lector ha estudiado en el curso de Geometría Analítica la ecuación

$$ax + by = c.$$

Si a y b no se anulan simultáneamente los puntos del plano de coordenadas x, y que la verifican pertenecen a una recta. Por lo tanto en el sistema

$$(S)\begin{cases} ax + by = c \\ a'x + b'y = c \end{cases}$$

cada ecuación representa una recta y su solución no es otra cosa que las coordenadas de los puntos que pertenecen a la intersección de ambas. De ese modo un sistema 2×2 compatible determinado (una única solución) se corresponde con dos rectas que se cortan en un único punto (ver figura 1.13), un sistema indeterminado (infinitas soluciones) se corresponde con dos rectas coincidentes y un sistema incompatible con dos rectas paralelas distintas (ver figuras 1.14 y 1.15 respectivamente).

EJEMPLO 1.13. El sistema

$$\begin{cases} 2x + y = 2 \\ x + y = 2 \end{cases}$$

es compatible determinado con única solución $(0, 2)$ ■

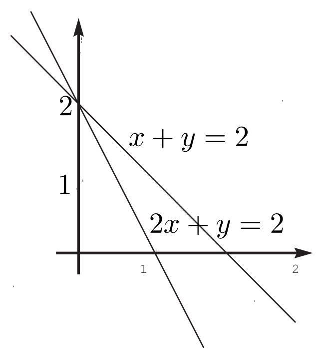


FIGURA 1

EJEMPLO 1.14. El sistema

$$\begin{cases} 2x + 2y = 4 \\ x + y = 2 \end{cases}$$

es compatible indeterminado con solución $\{(x, 2 - x)\, x \in \mathbb{R}\}$ ■

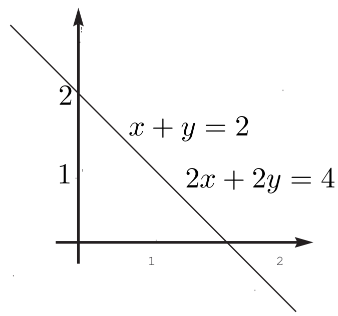


Figura 2

EJEMPLO 1.15. El sistema

$$\begin{cases} x + y = 1 \\ x + y = 2 \end{cases}$$

es incompatible ■

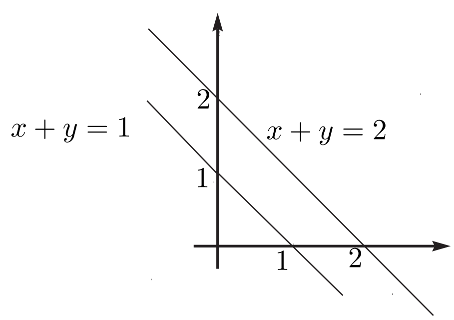


Figura 3

Esta interpretación geométrica de los sistemas 2×2 puede también extenderse a los sistemas de tres incógnitas salvo que, como veremos en el próximo capítulo, las ecuaciones representan planos en el espacio y la solución sus intersecciones.

En cierta medida la teoría que desarrollaremos en capítulos posteriores nos permitirá extender también la interpretación geométrica para sistemas de mayor tamaño.

CAPíTULO 2

ÁLGEBRA DE MATRICES

2.1. Operaciones con matrices

Ya hemos visto que en un sistema lineal de ecuaciones los coeficientes del mismo forman un "ordenamiento rectangular de números" al que denominamos matriz. Hasta el momento las matrices no son otra cosa que una nueva manera de representar al sistema de ecuaciones, simplemente se omite escribir las incógnitas, lo cual resulta más cómodo y compacto. Sin embargo las matrices son objetos matemáticos con "vida propia". Como veremos más adelante, las matrices comparten con los números muchas de sus propiedades aunque claro esta, son más "complicados" que estos. No olvidemos que guardan más información que una simple lista de números: son números ordenados en un damero.

Comencemos definiendo algunas operaciones con matrices.

DEFINICIÓN 2.1 (Suma de matrices). Sea $A = ((a_{ij}))$ y $B = ((b_{ij}))$ dos matrices $m \times n$, definimos la **suma** $+ : \mathcal{M}_{m\times n} \times \mathcal{M}_{m\times n} \to \mathcal{M}_{m\times n}$ de modo que $A + B = C$ siendo $C = ((c_{ij}))$, donde $c_{ij} = a_{ij} + b_{ij}$,

EJEMPLO 2.1. Sean $A = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 1 & 2 & -1 \\ 0 & -1 & \sqrt{2} \end{pmatrix}$ y $B = \begin{pmatrix} 0 & \frac{1}{2} & -1 \\ 2 & 2 & 0 \\ 1 & -2 & 2 \end{pmatrix}$, entonces

$$\begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 1 & 2 & -1 \\ 0 & -1 & \sqrt{2} \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{2} & -1 \\ 2 & 2 & 0 \\ 1 & -2 & 2 \end{pmatrix} = \begin{pmatrix} 1+0 & \frac{1}{2}+\frac{1}{2} & 0-1 \\ 1+2 & 2+2 & -1+0 \\ 0+1 & -1-2 & \sqrt{2}+2 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 1 & -1 \\ 3 & 4 & -1 \\ 1 & -3 & \sqrt{2}+2 \end{pmatrix}$$

■

DEFINICIÓN 2.2 (Producto de números por matrices). Sea $\lambda \in \mathbb{R}$ un número y $A \in \mathcal{M}_{m\times n}$ una matriz definimos el **producto de λ por A** como
$\lambda \cdot A = B \quad \text{donde } B = ((b_{ij})) \quad \text{con} \quad b_{ij} = \lambda a_{ij}$

EJEMPLO 2.2. Sea $\lambda = \sqrt{2}$ y $A = \begin{pmatrix} \sqrt{2} & 1 \\ 0 & -\sqrt{2} \end{pmatrix}$ entonces

$$\sqrt{2}\cdot \begin{pmatrix} \sqrt{2} & 1 \\ 0 & -\sqrt{2} \end{pmatrix} = \begin{pmatrix} \sqrt{2}(\sqrt{2}) & \sqrt{2}(1) \\ \sqrt{2}(0) & \sqrt{2}(-\sqrt{2}) \end{pmatrix} = \begin{pmatrix} 2 & \sqrt{2} \\ 0 & -2 \end{pmatrix}$$

■

Resulta útil destacar alguna de las propiedades que tienen estas operaciones.

Propiedades:

[S1] [Conmutativa] $A + B = B + A$, $\forall\, A, B \in \mathcal{M}_{m\times n}$.

[S2] [Asociativa] $\big(A + B\big) + Z = A + \big(B + Z\big)$, $\forall\, A, B, Z \in \mathcal{M}_{m\times n}$.

[S3] [Neutro de la suma] Existe $O \in \mathcal{M}_{m\times n}$ tal que $A + O = O + A = A$, $\forall\, A \in \mathcal{M}_{m\times n}$.

[S4] [Existencia de opuesto] Para cada $A \in \mathcal{M}_{m\times n}$ existe $B \in \mathcal{M}_{m\times n}$ tal que $A + B = O$ (Notación: $B = -A$).

[P1] [Asociativa del producto] $\big(\alpha\beta\big)A = \alpha(\beta A)$, $\forall\, \alpha, \beta \in \mathbb{R}$ y $A \in \mathcal{M}_{m\times n}$.

[P2] [Neutro del producto] $1.A = A \ \forall\, A \in \mathcal{M}_{m\times n}$
[P3] [Distributiva] $(\alpha+\beta)A = \alpha A + \beta A$ y $\forall\, \alpha, \beta \in \mathbb{R}$ y $A, B \in \mathcal{M}_{m\times n}$.
[P4] [Distributiva] $\alpha(A+B) = \alpha A + \alpha B$, $\forall\, \alpha \in \mathbb{R}$ y $A, B \in \mathcal{M}_{m\times n}$.

Las pruebas quedarán a cargo del lector. Observemos no obstante que la matriz O de la propiedad [S3] se denomina **matriz nula** y tiene todas sus entradas cero.

2.2. Producto de matrices

A las operaciones de suma y multiplicación por escalares, en el caso de las matrices, vamos a agregar un producto entre ellas. En una primera mirada y por analogía con las operaciones ya definidas parece razonable definir el producto de matrices como producto entrada a entrada. No será sin embargo, ese el camino que seguiremos. Nuestro producto es algo más complicado pero como veremos más adelante extremadamente útil.

DEFINICIÓN 2.3. Sea $A \in \mathcal{M}_{m\times p}$ y $B \in \mathcal{M}_{p\times n}$ definimos el **producto entre matrices** como, $\cdot : \mathcal{M}_{m\times p} \times \mathcal{M}_{p\times n} \to \mathcal{M}_{m\times n}$ de modo que $A{\cdot}B = C$ donde $C = ((c_{ij}))$, con

$$c_{ij} = \sum_{h=1}^{p} a_{ih} b_{hj}$$

El siguiente esquema ayuda a recordar la forma de realizar la operación:

$$\begin{array}{cc}
 & \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ b_{21} & \dots & b_{2j} & \dots & b_{2n} \\ \vdots & \dots & \vdots & \dots & \dots \\ b_{p1} & \dots & b_{pj} & \dots & b_{pn} \end{pmatrix} \\
\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ip} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{pmatrix} & \begin{pmatrix} & & \downarrow & & \\ \vdots & \vdots & & \vdots & \vdots \\ \longrightarrow & & c_{ij} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}
\end{array}$$

EJEMPLO 2.3. Sea $A = \begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$ calculemos su producto

$$\begin{aligned}
\begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} &= \begin{pmatrix} 2{\cdot}1 + (-1)(-1) & 2(2) + (-1)3 \\ 0{\cdot}1 + 2(-1) & 0{\cdot}2 + 2{\cdot}3 \end{pmatrix} = \\
&= \begin{pmatrix} 3 & 1 \\ -2 & 6 \end{pmatrix}
\end{aligned}$$

■

EJEMPLO 2.4. Sean $A = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 2 & 1 & 2 & -1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$. El producto de A por B es:

$$\begin{pmatrix} 2 & 1 & -1 & 0 \\ 2 & 1 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 7 \end{pmatrix}$$

pues

$$\begin{aligned}
3 &= 2{\cdot}1 + 1{\cdot}2 + (-1)1 + 0{\cdot}1 \\
4 &= 2{\cdot}2 + 1{\cdot}1 + (-1)1 + 0{\cdot}0 \\
5 &= 2{\cdot}1 + 1{\cdot}2 + 2{\cdot}1 + (-1)1 \\
7 &= 2{\cdot}2 + 1{\cdot}1 + 2{\cdot}1 + (-1)0
\end{aligned}$$

Vale la pena observar que las matrices no deben ser necesariamente del mismo tamaño para poder multiplicarlas. De hecho si se mira con cuidado la definición se observa que para poder hacerlo la cantidad de filas de la primera debe coincidir con la cantidad de columnas de la segunda. Cuando esto ocurra se dice que las matrices son **conformables**. ■

EJEMPLO 2.5. Sea $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ y $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ entonces

$$A.X = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ 2x + y + z \\ x + y + 2z \end{pmatrix}$$

■

EJEMPLO 2.6. Sea $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ y $X' = \begin{pmatrix} x & y & z \end{pmatrix}$ entonces

$$X'A = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} x + 2y + z & 2x + y + z & x + y + 2z \end{pmatrix}$$

■

OBSERVACIÓN 2.1. Sean $A = ((a_{ij}))$ una matriz $m \times n$ y $B = ((b_{jk}))$ una matriz $n \times p$ notaremos por $A_i = (a_{i1}, \ldots, a_{in})$ a la fila i-esima de A y con $B^j = \begin{pmatrix} b_{1k} \\ \vdots \\ a_{nk} \end{pmatrix}$ la columna k-esima de B. Si $A{\cdot}B = C$ designemos como C_i la fila i-esima y por C^k la columna k- esima de C. Entonces $C^k = A{\cdot}B^k$ y $C_i = A_iB$, es decir que

$$A{\cdot}B = \begin{pmatrix} A{\cdot}B^1 & A{\cdot}B^2 & \dots & A{\cdot}B^n \\ \Bigg| & \Bigg| & & \Bigg| \\ \Bigg| & \Bigg| & \dots & \Bigg| \end{pmatrix}$$

y

$$A{\cdot}B = \begin{pmatrix} A_1{\cdot}B & \rule{3cm}{0.4pt} & \\ A_2{\cdot}B & \rule{3cm}{0.4pt} & \\ \vdots & \vdots & \vdots \\ A_m{\cdot}B & \rule{3cm}{0.4pt} & \end{pmatrix}$$

La prueba queda como ejercicio a cargo del lector pero el ejemplo siguiente ilustra las afirmaciones.

EJEMPLO 2.7. Sea $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 1 \end{pmatrix}$ entonces

$$A{\cdot}B = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 6 & 4 \end{pmatrix}$$

Las columnas de la matriz producto son

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} \text{ y } \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

y las filas son

$$\begin{pmatrix} 2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 3 \end{pmatrix} \text{ y } \begin{pmatrix} 1 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 4 \end{pmatrix}$$

■

Propiedades del producto de matrices

Analicemos ahora las propiedades del producto de matrices. Comencemos observando que **no es conmutativo**. En el ejemplo 2.5 se observa que dos matrices pueden ser conformables en un orden y no en otro. De hecho en ese ejemplo AX tiene sentido pero XA no. Pero el producto de matrices no es conmutativo ni siquiera cuando las matrices a multiplicar son cuadradas. En efecto sea $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ entonces

$$A.B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}, \text{ pero}$$

$$B.A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}$$

Naturalmente el contraejemplo solo indica que en general $A{\cdot}B \neq B{\cdot}A$ lo cual no significa que no existan matrices que conmuten.

- [**Asociativa**] Si $A = ((a_{ij})), B = ((b_{ij}))$ y $C = ((c_{ij}))$ son tres matrices conformables entonces $\big(A{\cdot}B\big)C = A\big(B{\cdot}C\big)$.
  En efecto pongamos $\big(A{\cdot}B\big)C = D = ((d_{ij}))$ y $A\big(B{\cdot}C\big) = D' = ((d'_{ij}))$ debemos probar que $d_{ij} = d'_{ij} \;\; \forall\, i, j$.

  $$d_{ij} = \sum_h \left(a_{ih} \left(\sum_s b_{hs} c_{sj} \right) \right) = \sum_s \left(\left(\sum_h a_{ih} b_{hs} \right) c_{sj} \right) = d'_{ij}$$

■

- [**Distributiva**] Si $A = ((a_{ij})), B = ((b_{ij}))$ son matrices de igual dimensión y $C = ((c_{ij}))$ es una matriz conformable con A y B entonces

$$C(A+B) = C{\cdot}A + C{\cdot}B \text{ y } (A+B)C = A{\cdot}C + B{\cdot}C^1$$

Probemos sólo la primera, la otra queda como ejercicio. Pongamos

$$\begin{aligned} &C(A+B) = D = ((d_{ij})),\\ &C{\cdot}A = E = ((e_{ij})) \text{ y } C{\cdot}B = F = ((f_{ij})). \end{aligned}$$

Debemos probar entonces que $d_{ij} = e_{ij} + f_{ij}\ \forall\, i, j$

$$\begin{aligned} d_{ij} &= \sum_h c_{ih}\big(a_{hj} + b_{hj}\big) = \sum_h c_{ih}a_{hj} + c_{ih}b_{hj} = \\ &= \sum_h c_{ih}a_{hj} + \sum_h c_{ih}b_{hj} = e_{ij} + f_{ij} \end{aligned}$$

■

DEFINICIÓN 2.4. Sea $A = ((a_{ij}))$ una matriz $n \times m$, definimos la **matriz traspuesta** de A como la matriz $m \times n$ $A^t = ((a^t_{ij}))$, donde $a^t_{ij} = a_{ji}$.

EJEMPLO 2.8. Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \end{pmatrix}$ entonces $A^t = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}$

Sea $B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ entonces $B^t = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, entonces $B = B^t$, cuando esto ocurra diremos que la matriz es **simétrica**

Sea $C = \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$ entonces $C^t = \begin{pmatrix} 0 & -2 & 1 \\ 2 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$ por lo tanto $C^t = -C$, cuando esto ocurra diremos que la matriz es **antisimétrica**.

---

[1] Observe el lector que la no conmutatividad del producto de matrices obliga a probar ambas identidades independientemente, pues en general $C(A+B) \neq (A+B)C$.

Sea $D = \begin{pmatrix} 1 \\ -2 \\ 2 \\ 0 \end{pmatrix}$ entonces $D^t = \begin{pmatrix} 1 & -2 & 2 & 0 \end{pmatrix}$. Obsérvese que la traspuesta de una matriz columna resulta ser un matriz fila. ■

PROPOSICIÓN 2.2 (Propiedades de la matriz traspuesta). *Sean A y B dos matrices y λ un número se cumplen las siguientes propiedades*

1. $\big(A+B\big)^t = A^t + B^t$.
2. $(\lambda A)^t = \lambda A^t$.
3. $(A{\cdot}B)^t = B^t{\cdot}A^t$.
4. $(A^t)^t = A$

DEMOSTRACIÓN. Ejercicio. □

2.3. Ecuaciones matriciales. Matriz inversa.

Si A es una matriz $m \times n$ y b es una matriz $m \times 1$ entonces $AX = b$ es una ecuación matricial. Resolverla es encontrar todas las matrices X $n \times 1$ que la verifiquen.

El ejemplo 2.5 nos motiva a pensar que, identificando de manera natural matrices columna con sucesiones finitas de números, resolver la ecuación matricial $AX = b$ es equivalente a resolver el sistema $A|b$.

PROPOSICIÓN 2.3. *Sea*

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

un sistema lineal de m ecuaciones con n incógnitas, A la matriz del sistema y $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ *entonces* $(\widehat{x_1}, \widehat{x_2}, \dots, \widehat{x_n})$ *es solución de* (S) *si y solo si* $\widehat{X} = \begin{pmatrix} \widehat{x_1} \\ \widehat{x_2} \\ \vdots \\ \widehat{x_n} \end{pmatrix}$ *cumple que* $A{\cdot}\widehat{X} = b$

Demostración. Observemos en primer lugar que

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \widehat{x_1} \\ \widehat{x_2} \\ \vdots \\ \widehat{x_n} \end{pmatrix} = \begin{pmatrix} a_{11}\widehat{x_1} + a_{12}\widehat{x_2} + \cdots + a_{1n}\widehat{x_n} \\ a_{21}\widehat{x_1} + a_{22}\widehat{x_2} + \cdots + a_{2n}\widehat{x_n} \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}\widehat{x_1} + a_{m2}\widehat{x_2} + \cdots + a_{mn}\widehat{x_n} \end{pmatrix}$$

Entonces $\quad (\widehat{x_1}, \widehat{x_2}, \dots, \widehat{x_n}) \in Sol(S) \Longleftrightarrow$

$$a_{i1}\widehat{x_1} + a_{i2}\widehat{x_2} + \cdots + a_{in}\widehat{x_n} = b_i, \quad \forall\, i = 1, 2, \dots, m \Longleftrightarrow$$

$$\begin{pmatrix} a_{11}\widehat{x_1} + a_{12}\widehat{x_2} + \cdots + a_{1n}\widehat{x_n} \\ a_{21}\widehat{x_1} + a_{22}\widehat{x_2} + \cdots + a_{2n}\widehat{x_n} \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}\widehat{x_1} + a_{m2}\widehat{x_2} + \cdots + a_{mn}\widehat{x_n} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \Longleftrightarrow A\widehat{X} = b$$

□

El resultado anterior permite pensar el sistema de ecuaciones lineal con matriz ampliada $A|b$ como la ecuación matricial

$$AX = b. \tag{2.5}$$

Este hecho, en primer lugar, puede considerarse como la primera aplicación importante del producto de matrices y en cierta forma justifica su definición, a priori no muy natural. Más adelante veremos otras aplicaciones que permitirán comprender más profundamente las motivaciones de la definición.

Por otra parte la ecuación 2.5 recuerda fuertemente la situación que analizamos en nuestro primer ejemplo 1.1. Al ver el sistema escrito en forma matricial nos vemos tentados a "pasar dividiendo" A para "despejar" X. Naturalmente, por el momento, esto carece de sentido pero, si analizamos con más cuidado el procedimiento utilizado para resolver la ecuación $ax = b$ observamos que la clave reside en que todo número no nulo tiene inverso, esto es dado $a \in \mathbb{R}$ existe $a^{-1} \in \mathbb{R}$ tal que $a.a^{-1} = 1$. Podríamos analizar la posibilidad de que una matriz tuviese una propiedad análoga pero para esto necesitaríamos saber qué puede representar el "1" en nuestro contexto. De nuevo la clave está en analizar que propiedad abstracta caracteriza a la unidad real. Rápidamente se observa que el 1 no es otra cosa que el neutro multiplicativo de $\mathbb{R}$ es decir 1 es el único número real que verifica $a{,}1 = a$ $\forall a \in \mathbb{R}$. Veamos que el producto de matrices tiene también un neutro multiplicativo.

DEFINICIÓN 2.5. Llamaremos **matriz identidad** de tamaño n a la matriz $n \times n$

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Esto es $I_n = ((e_{ij}))$ con $e_{ij} = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases}$.

PROPOSICIÓN 2.4. $I_n{\cdot}A = A{\cdot}I_n = A$, $\forall\, A \in \mathcal{M}_{n\times n}$.

Demostración. La prueba es una verificación inmediata y queda a cargo del lector. $\square$

OBSERVACIÓN 2.5. El lector puede verificar directamente que si $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ es una matriz columna cualquiera entonces $I_n{\cdot}X = X$.

Ahora que tenemos nuestro neutro podemos dar la siguiente

DEFINICIÓN 2.6 (Matriz invertible). Sea A una matriz $n \times n$. Diremos que la matriz A es **invertible** si y solo si existe B matriz $n \times n$ tal que

$$A.B = I_n \qquad \text{y} \qquad B.A = I_n \tag{2.6}$$

donde I_n es la matriz identidad $n \times n$

OBSERVACIÓN 2.6. Nuevamente como consecuencia de la no conmutatividad del producto de matrices debemos, en la definición anterior, exigir que $A.B$ y $B.A$ sean ambos iguales a la identidad. A priori podría ocurrir que $A.B = I_n \neq B.A$. Veremos más adelante que esto realmente no es posible, de hecho si existe B tal que $AB = I_n$ entonces $BA = I_n$ (ver lema 2.9). Observemos además, que las matrices invertibles son por definición matrices cuadradas.

A continuación veremos que la matriz B que cumple (2.6), cuando existe, es única. Por tal motivo se le llama **la inversa de** A y se utiliza la notación $B = A^{-1}$

PROPOSICIÓN 2.7. *Sean B_1 y B_2 matrices $n \times n$ que cumplen (2.6) entonces $B_1 = B_2$*

Demostración. En efecto

$$B_1 = B_1 I = B_1\left(A.B_2\right) = B_1 A.B_2 = \left(B_1 A\right).B_2 = I B_2 = B_2.$$

$\square$

EJEMPLO 2.9. La matriz $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ es invertible, pues existe $B = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ tal que $A.B = I$ y $B.A = I$ (verificarlo) ■

EJEMPLO 2.10. La matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ no es invertible, pues para cualquier matriz $B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ se cumple que

$$A.B = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a+d & b+e & c+f \\ a-g & b-h & c-i \\ 0 & 0 & 0 \end{pmatrix} \neq$$

$$\neq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

■

PROPOSICIÓN 2.8. *Sean A y B dos matrices $n \times n$ invertibles entonces $A{\cdot}B$ es invertible y $(A{\cdot}B)^{-1} = B^{-1}{\cdot}A^{-1}$*

DEMOSTRACIÓN. Basta verificar directamente que

$$(A{\cdot}B)(B^{-1}{\cdot}A^{-1}) = A\overbrace{(B{\cdot}B^{-1})}^{=I_n}A^{-1} = A{\cdot}A^{-1} = I_n$$

y que

$$(B^{-1}{\cdot}A^{-1})(A{\cdot}B) = B^{-1}\overbrace{(A^{-1}{\cdot}A)}^{=I_n}B = B^{-1}{\cdot}B = I_n$$

□

Sea $A \in \mathcal{M}_{n\times n}$ una matriz invertible y $b \in \mathcal{M}_{n\times 1}$ una matriz columna consideremos el sistema $A{\cdot}X = b\,;$ como A es invertible podemos multiplicar en la ecuación anterior a ambos miembros por A^{-1} se tiene entonces [2]

$$\begin{aligned} A^{-1}\big(A{\cdot}X\big) = A^{-1}{\cdot}b \quad &\Longleftrightarrow \quad \big(A^{-1}{\cdot}A\big)X = A^{-1}{\cdot}b \Longleftrightarrow I_n{\cdot}X = A^{-1}{\cdot}b \\ &\Longleftrightarrow \quad X = A^{-1}{\cdot}b \end{aligned}$$

Por lo tanto si A es invertible $(x_1, x_2, \ldots, x_n)$ es solución del sistema con matriz ampliada $A|b$ si y solo si

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A^{-1}b \tag{2.7}$$

EJEMPLO 2.11. Resolvamos el sistema

$$(S)\begin{cases} 2x + y = 3 \\ x + y = -1 \end{cases}$$

la matriz del sistema es

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{y} \quad b = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Es fácil ver que A es invertible (ejemplo 2.9) y que

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}; \quad \text{entonces la solución es}$$

$$A^{-1}b = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -5 \end{pmatrix}.$$

El lector como ejercicio verificará que efectivamente $(4, -5)$ satisface (S) ■

[2] Obsérvese que se debe multiplicar en ambos miembros del mismo lado, pues el producto no es conmutativo

La sencilla idea, utilizada al principio del curso, para resolver un sistema lineal 1×1, funcionó también para resolver un sistema lineal $n \times n$ con $n > 1$ La fórmula (2.7) no sólo es análoga a la obtenida en el ejemplo 1.1 sino que parece indicar un nuevo procedimiento para resolver sistemas lineales. Sin embargo hay varios asuntos a aclarar; en primer lugar todo lo hecho sólo se aplica a sistemas cuadrados; en segundo lugar la fórmula (2.7) es válida sólo bajo la hipótesis de que A sea invertible y no tenemos por el momento ningún criterio para decidir si una matriz lo es o no (trataremos esto en la sección 2.7 y de también en el Capítulo sobre Determinantes). Finalmente cuando A es invertible para poder aplicar la fórmula necesitamos saber cómo obtener la matriz A^{-1}, inversa de A. A continuación veremos un algoritmo para calcular la matriz inversa.

Para ello, admitamos momentáneamente el siguiente Lema, cuya prueba realizaremos en la sección 2.7.

LEMA 2.9. *Sea A una matriz $n \times n$ entonces A es invertible si y solo si existe $X \in \mathcal{M}_{n\times n}$ tal que $A \cdot X = I_n$*

Nos bastara entonces, hallar X tal que $AX = I_n$. Lo cual es equivalente a resolver n sistemas de ecuaciones con igual matriz asociada y diferente termino independiente, pues si indicamos por X^j la columna j-esima de X

y por e^j la columna j-esima de I_n tenemos que

$$AX = I \Longleftrightarrow$$

$$A\begin{pmatrix} X^1 & X^2 & \dots & X^n \\ | & | & & | \\ | & | & & | \\ | & | & \dots & | \end{pmatrix} = \begin{pmatrix} A{\cdot}X^1 & A{\cdot}X^2 & \dots & A{\cdot}X^n \\ | & | & & | \\ | & | & & | \\ | & | & \dots & | \end{pmatrix} =$$

$$\begin{pmatrix} e^1 & e^2 & \dots & e^n \\ | & | & & | \\ | & | & & | \\ | & | & \dots & | \end{pmatrix} \Longleftrightarrow$$

$$\Longleftrightarrow (S_j) \quad A\begin{pmatrix} x_{j1} \\ \vdots \\ x_{jj} \\ \vdots \\ x_{jn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \vdots \\ \vdots \\ \longleftarrow \text{lugar } j \\ \vdots \\ \vdots \end{matrix} \qquad (j = 1, 2, ..., n)$$

Como los n sistemas $S_1, S_2, \dots, S_n$ tienen la misma matriz A de coeficientes (con diferentes términos independientes), se pueden resolver simultáneamente escalerizando la matriz $(A|\, I)$.. Naturalmente a priori nada asegura que estos sistemas sean compatibles, podría ocurrir que alguno no lo fuera y en ese caso no se puede determinar la matriz X y consecuentemente A no resulta invertible. De hecho A es invertible si y solo si los n sistemas S_i, $i = 1, \dots, n$ son compatibles.

EJEMPLO 2.12. Consideremos la matriz $A = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$. Queremos hallar (si existe) la matriz $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$ tal que $AX = I$: $AX =$

$$= \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 2x_{11} + x_{21} & 2x_{12} + x_{22} \\ 4x_{11} + 3x_{21} & 4x_{12} + 3x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Igualando (por columna) se tienen los siguientes sistemas de ecuaciones :

$$(S_1) \quad \left\{ \begin{array}{l} 2x_{11} + x_{21} = 1 \\ 4x_{11} + 3x_{21} = 0 \end{array} \right. \iff \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\iff A \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$(S_2) \quad \left\{ \begin{array}{l} 2x_{12} + x_{22} = 0 \\ 4x_{12} + 3x_{22} = 1 \end{array} \right. \iff \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\iff A \begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Resolvemos el sistema (S_1) por escalerización

$$\left(\begin{array}{cc|c} 2 & 1 & 1 \\ 4 & 3 & 0 \end{array} \right) \sim \left(\begin{array}{cc|r} 2 & 1 & 1 \\ 0 & 1 & -2 \end{array} \right) \longleftarrow F_2 + (-2)F_1 \quad \sim$$

$$\sim \left(\begin{array}{cc|r} 2 & 0 & 3 \\ 0 & 1 & -2 \end{array} \right) \longleftarrow F_1 - F_2 \quad \sim$$

$$\sim \left(\begin{array}{cc|r} 1 & 0 & \frac{3}{2} \\ 0 & 1 & -2 \end{array} \right) \longleftarrow \tfrac{1}{2}F_1$$

Por lo tanto $\begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ -2 \end{pmatrix}$. Resolvemos el sistema (S_2) por escalerización

$$\left(\begin{array}{cc|c} 2 & 1 & 0 \\ 4 & 3 & 1 \end{array} \right) \sim \left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 1 & 1 \end{array} \right) \longleftarrow F_2 + (-2)F_1 \quad \sim$$

$$\sim \left(\begin{array}{cc|r} 2 & 0 & -1 \\ 0 & 1 & 1 \end{array} \right) \longleftarrow F_1 - F_2 \quad \sim$$

$$\sim \left(\begin{array}{cc|r} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -2 \end{array} \right) \longleftarrow \tfrac{1}{2}F_1$$

Por lo tanto $\begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -2 \end{pmatrix}$. Es decir que la inversa de A es $X = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & -2 \end{pmatrix}$. Como para ambos sistemas la matriz A es la matriz de coeficientes, los podemos resolver simultáneamente

$$\left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 4 & 3 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array}\right) \longleftarrow F_2 + (-2)F_1 \sim$$

$$\sim \left(\begin{array}{cc|cc} 2 & 0 & 3 & -1 \\ 0 & 1 & -2 & 1 \end{array}\right) \longleftarrow F_1 - F_2 \sim$$

$$\sim \left(\begin{array}{cc|cc} 1 & 0 & \frac{3}{2} & -\frac{1}{2} \\ 0 & 1 & -2 & -2 \end{array}\right) \longleftarrow \tfrac{1}{2}F_1$$

■

En general, para calcular la inversa de la matriz A de tamaño $n \times n$, se reduce la matriz $(A|\, I)$ por escalerización a la matriz $(I|\, B)$ y se tiene que $A^{-1} = B$

2.4. El espacio de n-uplas

Las filas de una matriz (al igual que las columnas y las soluciones del sistema) son "colecciones ordenadas de números". Hagamos entonces, un pequeño paréntesis para introducir la noción de "colección ordenada" o n-upla y ciertas operaciones naturales que se corresponden con las transformaciones elementales. Estas ideas resultarán de gran importancia en este capítulo y en el resto del curso.

Recordamos que si A y B son dos conjuntos el producto cartesiano de ambos se define como el conjunto de **pares ordenados** de elementos de A y B

$$A \times B = \{(a, b) \,:\, a \in A \text{ y } b \in B\}.$$

Dos elementos (a, b) y (c, d) son iguales si $a = c$ y $b = d$. Si $A = B = \mathbb{R}$ entonces a

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \,:\, x \in \mathbb{R} \text{ y } y \in \mathbb{R}\}$$

lo llamaremos espacio de 2-uplas.
De manera análoga definimos

$$\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R} = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) \,:\, x \in \mathbb{R}, y \in \mathbb{R} \text{ y } z \in \mathbb{R}\}$$

el espacio de las 3-uplas.
En general

$$\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R} = \overbrace{\mathbb{R} \times \cdots \times \mathbb{R}}^{\text{n veces}} = \{(x_1, \dots, x_n) \,:\, x_i \in \mathbb{R},\ \forall i = 1, \dots, n\}$$

es el espacio de n-uplas.
Si $(x_1, \dots, x_n)$ y $(y_1, \dots, y_n)$ son dos n-uplas y α un número definimos la suma entre n-uplas como

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

y el producto de números por n-uplas como

$$\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$$

EJEMPLO 2.13. Supongamos $n = 3$ entonces

$$(1, 2, -1) + (0, 1, 2) = (1 + 0, 2 + 1, -1 + 2) = (1, 3, 1)$$

y

$$\sqrt{2}(1, 0, \sqrt{2}) = (\sqrt{2}.1, \sqrt{2}.0, \sqrt{2}.\sqrt{2}) = (\sqrt{2}, 0, 2)$$

Las operaciones tienen una lista de propiedades que nuevamente, al igual que en el caso de las operaciones con matrices, resulta útil destacar. Nótese que en ambos casos las propiedades coinciden.
Para compactar la notación designemos cada n-upla con una letra mayúscula, $X = (x_1, \dots, x_n)$.

[S1] [Conmutativa] $X + Y = Y + X$, $\forall\, X, Y \in \mathbb{R}^n$.
[S2] [Asociativa] $\big(X + Y\big) + Z = X + \big(Y + Z\big)$, $\forall\, X, Y, Z \in \mathbb{R}^n$.
[S3] [Neutro de la suma] Existe $O \in \mathbb{R}$ tal que $X + O = O + X = X$, $\forall\, X \in \mathbb{R}^n$.
[S4] [Existencia de opuesto] Para cada $X \in \mathbb{R}^n$ existe $Y \in \mathbb{R}^n$ tal que $X + Y = O$ (Notación: $Y = -X$).

[P1] [Asociativa del producto] $\big(\alpha\beta\big)X = \alpha(\beta X)\ \forall\, \alpha, \beta \in \mathbb{R}$ y $X \in \mathbb{R}^n$.
[P2] [Neutro del producto] $1.X = X,\ \forall\, X \in \mathbb{R}^n$
[P3] [Distributiva] $(\alpha + \beta)X = \alpha X + \beta X,\ \forall\, \alpha, \beta \in \mathbb{R}$ y $X, Y \in \mathbb{R}^n$.
[P4] [Distributiva] $\alpha(X + Y) = \alpha X + \alpha Y,\ \forall\, \alpha \in \mathbb{R}$ y $X, Y \in \mathbb{R}^n$.

Probemos sólo algunas, las restantes quedan a cargo del lector

[S1] Si $X = (x_1, \ldots, x_n)$ y $Y = (y_1, \ldots, y_n)$ entonces

$$\begin{aligned} X + Y &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n) = \\ &= (y_1 + x_1, \ldots, y_n + x_n) = (y_1, \ldots, y_n) + (x_1, \ldots, x_n) = Y + X \end{aligned}$$

[S3] Basta observar que $O = (0, \ldots, 0)$ actúa como neutro. De hecho es único (verificarlo). La n-upla O se denomina n-upla nula.

[S4] Si $X = (x_1, \ldots, x_n)$ basta elegir $Y = (-x_1, \ldots, -x_n)$ entonces

$$\begin{aligned} &X + Y = (x_1, \ldots, x_n) + (-x_1, \ldots, -x_n) = \\ &= (x_1 - x_1, \ldots, x_n - x_n) = (0, \ldots, 0) = O \end{aligned}$$

Obsérvese que $-1.X = -X$

DEFINICIÓN 2.7. Decimos que $X \in \mathbb{R}^n$ es una **combinación lineal** de las n-uplas $X_1, X_2, \ldots, X_n$ si existen números reales $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que

$$X = \lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n = \sum_{i=1}^{n} \lambda_i X_i$$

Los números λ_i, $i = 1, 2, \ldots, n$ se denominan **coeficientes** de la combinación lineal.
Si $A = \{X_1, X_2, \ldots, X_n\} \subset \mathbb{R}^n$ **combinar linealmente** los vectores de A con coeficientes λ_i, $i = 1, 2, \ldots, n$ es calcular

$$X = \sum_{i=1}^{n} \lambda_i X_i$$

OBSERVACIÓN 2.10. La n-upla nula $O = (0, \ldots, 0)$ es combinación lineal de cualquier colección de n-uplas, basta tomar todos los coeficientes nulos.

Si X es combinación lineal $A = \{X_1, \dots, X_p\}$ entonces es combinación lineal de B para cualquier conjunto B finito que contenga a A. En efecto supongamos que

$$X = \lambda_1 X_1 + \cdots + \lambda_p X_p$$

y que

$$B = \{X_1, \dots, X_p, Y_1, \dots, Y_h\}$$

entonces

$$X = \lambda_1 X_1 + \cdots + \lambda_p X_p + 0Y_1 + \cdots + 0Y_h$$

EJEMPLO 2.14. 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

$$(0,6,0) = 2(1,2,1) + (-1)(2,-2,2).$$

■

EJEMPLO 2.15. Sea $A = \{(1,2,1), (2,-2,2), (1,1,1)\}$ combinemos linealmente los vectores de A con coeficientes 3,-1, 0

$$3(1,2,1) + (-1)(2,-2,2) + 0(1,1,1) = (1,8,1)$$

■

EJEMPLO 2.16. La 2-upla $(2,1)$ es combinación lineal de

$$A = \{(1,1), (1,0), (-1,1)\}$$

pues

$$(2,1) = (1,1) + (1,0) + 0(-1,1)$$

pero también

$$(2,1) = 2(1,1) + (-1)(1,0) + (-1)(-1,1)$$

■

OBSERVACIÓN 2.11. En general si X es combinación lineal de A los coeficientes no tienen por que ser únicos, tal como muestra el ejemplo anterior.

EJEMPLO 2.17. Sea $A = \{(2,1),(1,0)\}$ y $X = (1,1)$. ¿Es X combinación lineal de A? Apelando a la definición observamos que la pregunta puede formularse de manera equivalente como ¿existen números λ_1, λ_2 tales que $\lambda_1(2,1) + \lambda_2(1,0) = (1,1)$?. Ahora bien,

$$\lambda_1(2,1) + \lambda_2(1,0) = (1,1) \Longleftrightarrow (S)\begin{cases} 2\lambda_1 + \lambda_2 = 1 \\ \lambda_1 = 1 \end{cases}$$

El sistema (S) tiene matriz ampliada $A|b = \left(\begin{array}{cc|c} 2 & 1 & 1 \\ 1 & 0 & 1 \end{array}\right)$ de la última ecuación se deduce que $\lambda_1 = 1$ y sustituyendo en la primera se tiene que $\lambda_2 = -1$ Nótese que la matriz A tiene como columnas las n-uplas del conjunto B. ■

OBSERVACIÓN 2.12. El razonamiento realizado en el último ejemplo puede aplicarse en general (el lector deberá verificar esto) y muestra que el problema de determinar si una n-upla $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ es combinación lineal de un conjunto $B = \{X_1, \ldots, X_n\} \subset \mathbb{R}^n$ es equivalente a discutir la compatibilidad de un sistema lineal de ecuaciones con matriz ampliada $A|b$ donde A es la matriz de columnas $X_1, X_2, \ldots, X_n$ y $b = \begin{pmatrix} x_1, \\ \vdots \\ x_n \end{pmatrix}$

Culminemos esta sección probando un resultado que será de utilidad en futuro que básicamente establece la "transitividad" de la combinación lineal.

PROPOSICIÓN 2.13. *Sea $A = \{X_1, X_2, \ldots, X_r\}$ y $B = \{Y_1, Y_2, \ldots, Y_m\}$ dos conjuntos de n-uplas tales que $\forall i = 1, 2, \ldots, m$ Y_i es combinación lineal de A. Si $Z \in \mathbb{R}^n$ es combinación lineal de B entonces Z es combinación lineal de A*

DEMOSTRACIÓN. Como Z es combinación de B existen entonces números $\beta_1, \beta_2, \ldots, \beta_m$ tales que

$$Z = \beta_1 Y_1 + \beta_2 Y_2 + \cdots + \beta_m Y_m \tag{2.8}$$

Pero como cada Y_i con $i = 1, 2, \ldots, m$ es combinación de A, existen números α_{ij} con $i = 1, 2, \ldots, m$ y $j = 1, 2, \ldots, r$ tales que

$$Y_i = \alpha_{i1}X_1 + \alpha_{i2}X_2 + \cdots + \alpha_{ir}X_r, \;\; \forall\, i = 1, 2, \ldots, m \tag{2.9}$$

Sustituyendo (2.9) en (2.8) se tiene

$$\begin{aligned} Z &= \beta_1(\alpha_{11}X_1 + \alpha_{12}X_2 + \cdots + \alpha_{1r}X_r) + \\ &\quad \beta_2(\alpha_{21}X_1 + \alpha_{22}X_2 + \cdots + \alpha_{2r}X_r) + \cdots + \\ &\quad \beta_m(\alpha_{m1}X_1 + \alpha_{m2}X_2 + \cdots + \alpha_{mr}X_r) \end{aligned}$$

Reordenando,

$$\begin{aligned} Z &= \underbrace{(\beta_1\alpha_{11} + \beta_2\alpha_{21} + \cdots + \beta_m\alpha_{m1})}_{=\lambda_1} X_1 + \\ &\quad \underbrace{(\beta_1\alpha_{12} + \beta_2\alpha_{22} + \cdots + \beta_m\alpha_{m2})}_{=\lambda_2} X_2 + \cdots + \\ &\quad \underbrace{(\beta_1\alpha_{1r} + \beta_2\alpha_{2r} + \cdots + \beta_m\alpha_{mr})}_{=\lambda_r} X_r \end{aligned}$$

Por lo tanto

$$Z = \lambda_1X_1 + \lambda_2X_2 + \cdots + \lambda_rX_r$$

probando que Z es combinación de A. $\square$

2.5. Dependencia lineal

Introducimos ahora el concepto de dependencia lineal para n-uplas. el cual resulta ahora básico para la caracterización de las matrices invertibles y que en el futuro jugara un papel central en muchos otros problemas. Los conceptos básicas de esta sección serán vistas nuevamente en el Capítulo sobre Espacios Vectoriales, en un contexto más abstracto. Los ideas nuevas contenidos en las definiciones y demostraciones de esta Sección se basan únicamente en las **propiedades de las operaciones** (S1 -P4) entre n-uplas y no en el hecho de que éstas sean una sucesión finita de números. Por tanto si entre elementos de otro conjunto se tuvieran definidas operaciones con esas propiedades, los resultados serían también válidos. De hecho estas observaciones son las que conducen a introducir el concepto de Espacio Vectorial.

Sea (S) un sistema lineal $n \times n$ compatible con matriz ampliada $A|b$ El carácter determinado o indeterminado de (S) depende, de acuerdo al teorema de Rouché-Frobenius, de la cantidad de escalones (i.e. de filas no nulas) de la forma reducida de A. De hecho (S) es determinado si ninguna fila de la forma reducida de A es nula y es indeterminado si alguna lo es. La forma reducida se obtiene realizando combinaciones lineales con las filas de A, por lo tanto para que en el proceso de escalerización una fila se anule debe ser combinación lineal de las restantes. Para ilustrar esto puede observarse en el ejemplo 1.11 que la quinta fila de la matriz $(1 \;\; 2 \;\; 1 \;\; 4 \;\; 9 \;\; 3 \;\; -1)$ se anula, analizando las transformaciones elementales que se efectuaron al escalerizar se deduce que $F_5 = 2F_3 - F_1$, lo cual se comprueba directamente

$$(1 \;\; 2 \;\; 1 \;\; 4 \;\; 9 \;\; 3 \;\; -1) = 2(1 \;\; 2 \;\; 1 \;\; 2 \;\; 5 \;\; 3 \;\; 1) - (1 \;\; 2 \;\; 1 \;\; 0 \;\; 1 \;\; 1 \;\; 1)$$

En otras palabras la determinación o indeterminación de un sistema compatible depende de que alguna fila en la matriz del sistema sea combinación de las restantes. Como las filas de una matriz pueden pensarse como n-uplas podemos plantear el problema de una forma un poco más abstracta.
Sea $B = \{X_1, \ldots, X_n\} \subset \mathbb{R}^n$ queremos determinar si alguno de los elementos de B es combinación lineal de los restantes, veamos algunos ejemplos para ilustrar el problema.

EJEMPLO 2.18. Sea $B = \{(2,1,0),(1,1,1),(3,2,1)\}$ ¿alguna n-upla de B es combinación lineal de las restantes? Intentemos ver si $(2,1,0)$ es combinación de los otros dos.
Para esto debemos determinar si existen $\lambda_1, \lambda_2 \in \mathbb{R}$ tales que

$$\lambda_1(1,1,1) + \lambda_2(3,2,1) = (2,1,0).$$

Esto es equivalente a que el sistema

$$(S)\begin{cases} \lambda_1 + 3\lambda_2 = 2 \\ \lambda_1 + 2\lambda_2 = 1 \\ \lambda_1 + \lambda_2 = 0 \end{cases}$$

sea compatible. Para resolver el sistema (S) escribamos la matriz ampliada

$$A = \left(\begin{array}{cc|c} 1 & 3 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{array}\right),$$

al escalerizarla se obtiene

$$A = \left(\begin{array}{rr|r} 1 & 3 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{array}\right)$$

Por lo tanto el sistema (S) es compatible determinado y $Sol(S) = \{(-1,1)\}$. En consecuencia,

$$(2,1,0) = (-1)(1,1,1) + (3,2,1).$$

Además de lo anterior, se deduce fácilmente que

$$(3,2,1) = (2,1,0) + (1,1,1) \quad \text{y} \quad (1,1,1) = (3,2,1) + (-1)(2,1,0)$$

Es decir que cualquiera de las tres n-uplas es combinación lineal de las restantes. Cabe preguntarse si esto ocurre siempre. La respuesta es negativa como muestra el siguiente ejemplo

Sea $B' = \{(1,1,1),(2,2,2),(0,0,1)\}$ entonces es claro que

$$(1,1,1) = \frac{1}{2}(2,2,2) + 0(0,0,1) \quad \text{y} \quad (2,2,2) = 2(1,1,1) + 0(0,0,1) \tag{2.10}$$

pero sin embargo

$$(0,0,1) \neq \lambda_1(2,2,2) + \lambda_2(1,1,1), \quad \forall\, \lambda_1, \lambda_2 \in \mathbb{R}$$

pues el sistema

$$\begin{cases} 2\lambda_1 + \lambda_2 = 0 \\ 2\lambda_1 + \lambda_2 = 0 \\ 2\lambda_1 + \lambda_2 = 1 \end{cases}$$

es incompatible. Vale la pena observar también, que la n-upla $(0,0,1)$ no puede despejarse de (2.10) pues en cualquiera de las dos combinaciones lineales tiene coeficiente nulo. ■

El ejemplo anterior muestra que en general determinar si en un conjunto de n-uplas alguna es combinación lineal de las restantes puede resultar trabajoso si se utiliza el procedimiento del ejemplo, pues como se vio puede ocurrir que sea necesario chequear una a una las n-uplas del conjunto ya que algunas pueden no ser combinación de las restantes y otras si. Vale la pena entonces tratar de obtener un método más eficiente para resolver este problema.

DEFINICIÓN 2.8. Un conjunto $A = \{X_1, X_2, \ldots, X_m\} \subset \mathbb{R}^n$ se dice **linealmente dependiente** (L.D.) si existen $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$ no todos nulos tales que

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_m X_m = \sum_{i=0}^{m} \lambda_i X_i = O$$

DEFINICIÓN 2.9. Un conjunto $A = \{X_1, X_2, \ldots, X_m\} \subset \mathbb{R}^n$ se dice **linealmente independiente** (L.I.) si no es linealmente dependiente o equivalentemente $A = \{X_1, X_2, \ldots, X_m\} \subset \mathbb{R}^n$ es L.I. si $\lambda_1, \lambda_2, \ldots, \lambda_m$ son números reales tales que

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_m X_m = \sum_{i=0}^{m} \lambda_i X_i = O$$

entonces $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$

En otras palabras un conjunto es L.I. si la única forma de escribir la n-upla nula como combinación lineal de sus elementos es la trivial y es L.D. si hay otras formas.

Los siguientes resultados dan una caracterización de los conjuntos L.D. y L.I.

PROPOSICIÓN 2.14. *Sea* $A = \{X_1, X_2, \ldots, X_m\} \subset \mathbb{R}^n$ *entonces*

1. *A es linealmente dependiente, si y solo si existe* $X_{i_0} \in A$*, combinación lineal de los restantes elementos de* A
2. *A es linealmente independiente si ningún elemento de A es combinación lineal de los restantes.*

DEMOSTRACIÓN. Sólo es necesario probar la primera afirmación pues la segunda es su consecuencia inmediata.
($\Longrightarrow$) Supongamos que A es L.D. entonces existen números $\lambda_1, \lambda_2, \ldots, \lambda_m$ no todos nulos tales que

$$\sum_{i=0}^{m} \lambda_i X_i = 0 \tag{2.11}$$

Sea λ_{i_0} un coeficiente no nulo en la combinación lineal precedente, las propiedades de las operaciones de suma y producto de n-uplas nos habilitan a "despejar" el termino i_0-esimo en (2.11)

$$\begin{aligned}
&\lambda_1 X_1 + \cdots + \lambda_{i_0-1} X_{i_0-1} + \lambda_{i_0} X_{i_0} + \lambda_{i_0+1} X_{i_0+1} + \cdots + \lambda_m X_m = O \Longrightarrow \\
&\overbrace{-\lambda_{i_0}}^{\neq 0} X_{i_0} = \lambda_1 X_1 + \cdots + \lambda_{i_0-1} X_{i_0-1} + \lambda_{i_0+1} X_{i_0+1} + \cdots + \lambda_m X_m \Longrightarrow \\
&X_{i_0} = -\frac{\lambda_1}{\lambda_{i_0}} X_1 - \cdots - \frac{\lambda_{i_0-1}}{\lambda_{i_0}} X_{i_0-1} - \frac{\lambda_{i_0+1}}{\lambda_{i_0}} X_{i_0+1} - \cdots - \frac{\lambda_m}{\lambda_{i_0}} X_m
\end{aligned}$$

Por lo tanto X_{i_0} es combinación lineal de los restantes elementos de A como queríamos.
($\Longleftarrow$) Supongamos ahora que existe $X_{i_0} \in A$ combinación lineal de los restantes elementos de A entonces existen números $\lambda_1, \ldots, \lambda_{i_0-1}, \lambda_{i_0+1}, \ldots, \lambda_m$ tales que

$$X_{i_0} = \lambda_1 X_1 + \cdots + \lambda_{i_0-1} X_{i_0-1} + \lambda_{i_0+1} X_{i_0+1} + \cdots + \lambda_m X_m$$

entonces sumando el opuesto de X_{i_0} a ambos miembros se tiene que

$$\lambda_1 X_1 + \cdots + \lambda_{i_0-1} X_{i_0-1} + \overbrace{(-1)}^{\neq 0} X_{i_0} + \lambda_{i_0+1} X_{i_0+1} + \cdots + \lambda_m X_m = O$$

de donde hemos obtenido la n-upla nula como combinación lineal no trivial de los elementos de A y consecuentemente éste es un conjunto L.D. como queríamos. □

La proposición anterior y su prueba dan un método más eficiente para determinar si en un conjunto de n-uplas un elemento es combinación de los restantes: basta investigar si es L.D o no. Además si se analiza con cuidado en la demostración del directo se observa cuales son las n-uplas que son combinación de las otras, son exactamente aquellas que en (2.11) tienen coeficiente no nulo, esto es lo que permite "despejarlas".

Veamos ahora algunos ejemplos de conjuntos L.I. y L.D.

EJEMPLO 2.19. El ejemplo más sencillo de un conjunto L.D. es el que tiene como único elemento a la n-upla nula. Si $A = \{(0, 0, \ldots, 0)\} \subset \mathbb{R}^m$ entonces $\lambda(0, 0, \ldots, 0) = O \ \forall \lambda \in \mathbb{R}$, en particular si $\lambda \neq 0$.
Es fácil ver que este es el único conjunto de un elemento L.D.
En efecto si $A = \{X\}$ con $X \neq O$ entonces $\lambda X = O$ implica $\lambda = 0$ (probarlo), con lo cual A resulta L.I.
Por otra parte cualquier conjunto que contenga a O debe ser L.D.. Sea $A = \{O, X_1, \ldots, X_m\} \subset \mathbb{R}^n$ entonces

$$100.O + 0X_1 + \ldots + 0X_m = O$$

es una combinación lineal de los elementos de A que da el nulo con no todos los coeficientes nulos. ■

EJEMPLO 2.20. Determinar en cada caso la dependencia lineal del conjunto A y en caso de ser L.D. indicar cuáles vectores son combinación lineal de los restantes.

1. Sea $A = \{(1,1,-1),(1,0,1),(1,1,1)\}$ supongamos que $\lambda_1, \lambda_2, \lambda_3$ son tres números reales tales que

$$\begin{aligned}
&\lambda_1(1,1,-1) + \lambda_2(1,0,1) + \lambda_3(1,1,1) = (0,0,0) \iff \\
&(\lambda_1 + \lambda_2 + \lambda_3,\ \lambda_1 + \lambda_3,\ -\lambda_1 + \lambda_2 + \lambda_3) = (0,0,0) \iff \\
&(H)\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ \lambda_1 + \lambda_3 = 0 \\ -\lambda_1 + \lambda_2 + \lambda_3 = 0 \end{cases}
\end{aligned}$$

En consecuencia el conjunto es L.I si el sistema homogéneo (H) es determinado (sólo admite solución trivial) y es L.D. si (H) es indeterminado (admite soluciones no triviales). La matriz del sistema es

$$\left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 \end{array}\right)$$

escalerizándola se obtiene

$$\left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \end{array}\right)$$

por lo tanto (H) es determinado y la única solución es la trivial $\lambda_1 = \lambda_2 = \lambda_3 = 0$ de donde A es L.I..

2. Sea $A = \{(1,1,2,0),(1,0,1,1),(0,1,1,-1),(2,1,3,1),(1,0,1,0)\}$ y sean $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ números reales tales que

$$\begin{aligned}
\lambda_1(1,1,2,0) + \lambda_2(1,0,1,1) + \lambda_3(0,1,1,-1) + \\
\lambda_4(2,1,3,1) + \lambda_5(1,0,1,0) = (0,0,0,0) \iff
\end{aligned} \tag{2.12}$$

$$\begin{aligned}
(\lambda_1 + \lambda_2 + 2\lambda_4 + \lambda_5, \lambda_1 + \lambda_3 + \lambda_4, 2\lambda_1 + \lambda_2 + \lambda_3 + 3\lambda_4 + \lambda_5, \lambda_2 - \lambda_3 + \lambda_4) \\
= (0,0,0,0) \iff
\end{aligned}$$

$$\begin{cases} \lambda_1 + \lambda_2 + 2\lambda_4 + \lambda_5 = 0 \\ \lambda_1 + \lambda_3 + \lambda_4 = 0 \\ 2\lambda_1 + \lambda_2 + \lambda_3 + 3\lambda_4 + \lambda_5 = 0 \\ \lambda_2 - \lambda_3 + \lambda_4 = 0 \end{cases}$$

La matriz del sistema es:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 2 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 3 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 \end{array}\right)$$

Escalerizándola se obtiene:

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

y en consecuencia la solución es $\lambda_1 = -\lambda_3 - \lambda_4$, $\lambda_2 = \lambda_3 - \lambda_4$, $\lambda_5 = 0$, $\lambda_3, \lambda_4 \in \mathbb{R}$ El sistema es entonces compatible indeterminado y por lo tanto A es L.D. En efecto sustituyendo en (2.12) se tiene que para cualquier valor real de λ_3 y λ_4 se cumple que

$$(-\lambda_3 - \lambda_4)(1,1,2,0) + (\lambda_3 - \lambda_4)(1,0,1,1) + \lambda_3(0,1,1,-1) + \\ \lambda_4(2,1,3,1) + 0(1,0,1,0) = (0,0,0,0)$$

en particular poniendo por ejemplo $\lambda_3 = 1$ y $\lambda_4 = 2$ tenemos que

$$(-3)(1,1,2,0) + (-1)(1,0,1,1) + (0,1,1,-1) + \\ 2(2,1,3,1) + 0(1,0,1,0) = (0,0,0,0)$$

de donde podemos "despejar" cualquier n-upla salvo la última (tiene coeficiente cero) como combinación lineal de las restantes. Por ejemplo,

$$(1,1,2,0) = \frac{1}{3}(1,0,1,1) - \frac{1}{3}(0,1,1,-1) \\ -\frac{2}{3}(2,1,3,1) + 0(1,0,1,0)$$

■

Supongamos que (S) es un sistema $n \times n$ indeterminado, cada una de sus ecuaciones pueden pensarse en cierto sentido como "datos" o "información" que permite resolver el problema. Si (S) es indeterminado ya hemos visto que algunas ecuaciones (o filas de su matriz asociada) son combinación de las restantes, esto puede considerarse como "redundancia en la información" que estas aportan. Nos interesa saber cuanta es la "información" no redundante con que contamos y esto es cuantas ecuaciones independientes tienen el sistema. De nuevo planteamos el problema en el contexto de n-uplas.

DEFINICIÓN 2.10. Sea $A \subset \mathbb{R}^n$ un conjunto finito, llamaremos **rango del conjunto** A a la mayor cantidad de elementos de A que forman un conjunto L.I.

Naturalmente si A es L.I. entonces $rango(A) = card(A)$. El siguiente resultado indica un camino para calcular el rango de un conjunto.

PROPOSICIÓN 2.15. *Sea* $A = \{X_1, X_2, \ldots, X_m\} \subset \mathbb{R}^n$ *un conjunto cualquiera y sea* $B = \{X_{h_1}, X_{h_2}, \ldots, X_{h_r}\} \subset A$ *un subconjunto de* A *que cumple las siguiente dos condiciones:*

1. B *es L.I.*
2. X_s *es combinación lineal de* B $\forall\ s \neq h_1, h_2, \ldots, h_r$

entonces $rango(A) = r$

DEMOSTRACIÓN. Como B es L.I. y $card(B) = r$ entonces $rango(A) \geq r$. El resultado estará probado si verificamos que cualquier subconjunto de A con más de r elementos es L.D.
Sea $C = \{X_{k_1}, X_{k_2}, \ldots, X_{k_q}\} \subset A$ un subconjunto cualquiera de A con q elementos, $q > r$.
Observemos primeramente que la segunda condición de la hipótesis implica que todos los elementos de A son combinación lineal de B pues cualquier n-upla es combinación de si misma. Por lo tanto existen números a_{ij}, con $i = 1, \ldots, r$, $j = 1, \ldots, q$ tales que

$$X_{k_j} = a_{1j}X_{h_1} + a_{2j}X_{h_2} + \ldots + a_{rj}X_{h_r},\ \ \forall\, j = 1, 2, \ldots, q \tag{2.13}$$

Analicemos la dependencia lineal del conjunto C. Sean $\lambda_1, \lambda_2, \ldots, \lambda_q$ números tales que

$$\lambda_1 X_{k_1} + \lambda_2 X_{k_2} + \ldots + \lambda_q X_{k_q} = O \tag{2.14}$$

Sustituyendo (2.13) en (2.14) se tiene que $\lambda_1, \lambda_2, \ldots, \lambda_q$ verifica (2.14) si y solo si

$$\lambda_1 \big(a_{11} X_{h_1} + a_{21} X_{h_2} + \ldots + a_{r1} X_{h_r}\big) + \lambda_2 \big(a_{12} X_{h_1} + a_{22} X_{h_2} + \ldots + a_{r2} X_{h_r}\big) + \\ + \cdots + \lambda_q \big(a_{1q} X_{h_1} + a_{2q} X_{h_2} + \ldots + a_{rq} X_{h_r}\big) = O$$

Reordenando y sacando X_{h_i} de factor común se tiene que $\lambda_1, \lambda_2, \ldots, \lambda_q$ verifican (2.14) si y solo si

$$\overbrace{\big(a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1q}\lambda_q\big)}^{=0} X_{h_1} + \overbrace{\big(a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2q}\lambda_q\big)}^{=0} X_{h_2} + \\ + \cdots + \underbrace{\big(a_{r1}\lambda_1 + a_{r2}\lambda_2 + \cdots + a_{rq}\lambda_q\big)}_{=0} X_{h_r} = O$$

Los paréntesis en la ecuación anterior resultan todos nulos pues, el conjunto $B = \{X_{h_1}, X_{h_2}, \ldots, X_{h_r}\}$ es L.I.. Tenemos entonces que $\lambda_1, \lambda_2, \ldots, \lambda_q$ verifican (2.14) si y solo si verifican el sistema lineal

$$(S) \begin{cases} a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1q}\lambda_q = 0 \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2q}\lambda_q = 0 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{r1}\lambda_1 + a_{r2}\lambda_2 + \cdots + a_{rq}\lambda_q = 0 \end{cases}$$

Pero (S) es un sistema homogéneo y $q > r$ (tienen más incógnitas que ecuaciones) por lo cual es indeterminado y consecuentemente tiene soluciones no triviales. De donde resulta que existen $\lambda_1, \lambda_2, \ldots, \lambda_q$ no todos nulos que verifican (2.14) y por lo tanto C es L.D. como se deseaba. $\square$

EJEMPLO 2.21. Analicemos el rango del conjunto

$$A = \{(1,1,2,0), (1,0,1,1), (0,1,1,-1), (2,1,3,1), (1,0,1,0)\}$$

del ejemplo 2.20. Ya sabemos que es L.D. por lo cual $rango(A) < 5$. Nuestra intención es construir con los elementos de A un conjunto B en las condiciones de la proposición anterior. Para construir un conjunto L.I. con los elementos de A debemos quitar las n-uplas que sean combinación lineal de las restantes.
Como ya vimos, en nuestro caso, $(1,1,2,0)$ es combinación de las otras así que consideramos el conjunto

$$A' = \{(1,0,1,1),(0,1,1,-1),(2,1,3,1),(1,0,1,0)\}$$

y estudiamos su dependencia lineal. Sean $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ números tales que

$$\lambda_1(1,0,1,1) + \lambda_2(0,1,1,-1) + \lambda_3(2,1,3,1) + \lambda_4(1,0,1,0) = (0,0,0,0) \Longleftrightarrow$$

$$(S)\begin{cases} \lambda_1 + 2\lambda_3 + \lambda_4 = 0 \\ \lambda_2 + \lambda_3 = 0 \\ \lambda_1 + \lambda_2 + 3\lambda_3 + \lambda_4 = 0 \\ \lambda_1 - \lambda_2 + \lambda_3 = 0 \end{cases}$$

La matriz ampliada de este sistema es

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 3 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 \end{array}\right)$$

Escalerizándola se obtiene

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

Por lo tanto (S) tiene solución $\lambda_1 = -2\lambda_3$, $\lambda_2 = -\lambda_3$, $\lambda_4 = 0$ y $\lambda_3 \in \mathbb{R}$. A' es L.D. y además $\forall\, \lambda_3 \in \mathbb{R}$ se tiene que

$$-2\lambda_3(1,0,1,1) - \lambda_3(0,1,1,-1) + \lambda_3(2,1,3,1) + 0(1,0,1,0) = (0,0,0,0)$$

En particular si $\lambda_3 \neq 0$, despejando resulta

$$(2,1,3,1) = -\frac{2\lambda_3}{\lambda_3}(1,0,1,1) - \frac{\lambda_3}{\lambda_3}(0,1,1,-1) + \frac{0}{\lambda_3}(1,0,1,0) \Longrightarrow$$
$$(2,1,3,1) = -2(1,0,1,1) - (0,1,1,-1) + 0(1,0,1,0)$$

Con lo cual tenemos que en A' $(2,1,3,1)$ es combinación lineal de los restantes elementos. Consideramos ahora el conjunto

$$A'' = \{(1,0,1,1),(0,1,1,-1),(1,0,1,0)\} \subset A' \subset A$$

el cual es L.I. (verificarlo). Entonces $B = A''$ cumple las dos condiciones de la proposición anterior, esto es:

1. B es L.I.
2. $(2,1,2,1)$ y $(1,1,2,0)$ son combinación lineal de B

Por lo tanto $rango(A) = card(B) = 3$.
Obsérvese que habíamos probado, en el ejemplo 2.20, que $(1,1,2,0)$ es combinación de A' pero como los elementos de A' son a su vez combinación de los de A'' se deduce aplicando la proposición 2.13 que $(1,1,2,0)$ es combinación de A''. ■

2.6. El rango de una matriz

Si $A = ((a_{ij}))$ es una matriz $m \times n$ notaremos por $A_i = (a_{i1}, \ldots, a_{in}) \in \mathbb{R}^n$ a la fila i-esima de A y con $A^j = (a_{1j}, \ldots, a_{mj}) \in R^m$ la columna j-esima de A.

DEFINICIÓN 2.11. Sea A una matriz $m \times n$

1. Definimos el **rango por filas** de A $(rango_f(A))$ como el rango del conjunto $\{(A_1, \ldots, A_n\}$ de filas de A.
2. Definimos el **rango por columnas** de A $(rango_c(A))$ como el rango del conjunto $\{A^1, \ldots, A^m\} \subset \mathbb{R}^m$ de columnas de A.

LEMA 2.16. *Sea $E \in \mathcal{M}_{m\times n}$ una matriz escalerizada reducida con p escalones ubicados en las columnas $j_1, j_2, \ldots, j_p$ entonces los conjuntos $B = \{E^{j_1}, E^{j_2}, \ldots, E^{j_p}\}$ de columnas y $B' = \{E_1, E_2, \ldots, E_p\}$ de filas cumplen*

ambos con las hipótesis del lema 2.15.
En particular, entonces $rango_c(E) = rango_f(E) = p$

Demostración. La prueba es sencilla y queda a cargo del lector. □

LEMA 2.17. *Sea* A *una matriz* $m \times n$ *y* p *el número de escalones de la forma escalerizada reducida de* A *entonces* $rango_c(A) = p$

Demostración. Sea A una matriz y E la forma escalerizada reducida de A con p escalones (esto es con p filas no nulas). Como siempre indicaremos por A^j y E^j las j-esimas columnas de A y E respectivamente.
Supongamos que los escalones de E están ubicados en las columnas $j_1, j_2, \ldots, j_p$ entonces para probar que $rango_c(A) = p$ alcanza con demostrar que el conjunto $B = \{A^{j_1}, A^{j_2}, \ldots, A^{j_p}\}$ cumple las condiciones de la proposición 2.15 esto es B es L.I y A^j es combinación lineal de B $\forall\, j \neq j_1, j_2, \ldots, j_p$.
Veamos primero que B es L.I. Sean $\lambda_{j_1}, \lambda_{j_2}, \ldots, \lambda_{j_p}$ números tales que

$$\lambda_{j_1} A^{j_1} + \lambda_{j_2} A^{j_2} + \cdots + \lambda_{j_p} A^{j_p} = O$$

entonces la n-upla $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ tal que

$$\alpha_h = \begin{cases} \lambda_{j_i}, \text{ si } h = j_i \text{ para algún } i. \\ 0 \text{ si } h \neq j_i \text{ para todo } i. \end{cases}$$

es una solución del sistema homogéneo con matriz $A|O$ y por lo tanto es solución del sistema homogéneo equivalente con matriz $E|O$. Esto último implica que

$$O = \alpha_1 E^1 + \alpha_2 E^2 + \cdots + \alpha_n E^n = \lambda_{j_1} E^{j_1} + \lambda_{j_2} E^{j_2} + \cdots + \lambda_{j_p} E^{j_p},$$

pero el conjunto $\{E^{j_1}, E^{j_2}, \ldots, E^{j_p}\}$ es L.I. en virtud de lo visto en el lema 2.16, entonces $\lambda_{j_1} = \lambda_{j_2} = \cdots = \lambda_{j_p} = 0$ probando que B es L.I.
Veamos ahora que si $j \neq j_1, \ldots, j_p$ entonces A^j es combinación lineal de B. Sea entonces $j \neq j_1, \ldots, j_p$, del lema 2.16 se sabe que E^j es combinación lineal del conjunto $\{E^{j_1}, E^{j_2}, \ldots, E^{j_p}\}$ por lo tanto existen $\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_p}$ números tales que

$$E^j = \beta_{j_1} E^{j_1} + \beta_{j_2} E^{j_2} + \cdots + \beta_{j_p} E^{j_p}$$

entonces

$$(-1)E^j + \beta_1 E^{j_1} + \beta_2 E^{j_2} + \cdots + \beta_p E^{j_p} = O.$$

Por lo tanto la n-upla $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ con

$$\alpha_h = \begin{cases} -1, \text{ si } h = j. \\ \beta_{j_i}, \text{ si } h = j_i \text{ para algún } i. \\ 0 \text{ si } h \neq j_i \text{ para todo } i. \end{cases}$$

es solución del sistema homogénea con matriz $E|O$ y consecuentemente también del sistema equivalente $A|O$ por lo tanto

$$O = \alpha_1 A^1 + \alpha_2 A^2 + \cdots + \alpha_n A^n = (-1)A^j + \beta_{j_1} A^{j_1} + \beta_{j_2} A^{j_2} + \cdots + \beta_{j_p} A^{j_p}$$

de donde se deduce que

$$A^j = \beta_{j_1} A^{j_1} + \beta_{j_2} A^{j_2} + \cdots + \beta_{j_p} A^{j_p}$$

culminando la prueba del lema. □

TEOREMA 2.18. *Sea A una matriz $m \times n$ entonces $rango_f(A) = rango_c(A)$*

DEMOSTRACIÓN. La prueba se reduce a mostrar que $rango_f(A) = p$ donde p es el número de escalones de E,la forma reducida de A.
Es claro que en el proceso de escalerización una fila se anula si y solo si es combinación lineal de las restantes en consecuencia si la forma escalerizada tiene p filas no nulas la matriz A tiene p filas que forman un conjunto L.I. (ninguna de ellas puede ser combinación de las restantes) y además las otras $m-p$ son combinación de ellas (por eso al escalerizar A se anularon). En consecuencia el lema 2.15 asegura que $rango_f(A) = p$. □

El resultado anterior nos habilita entonces a definir el **rango de una matriz** como el rango por filas o el rango por columnas.

OBSERVACIÓN 2.19. El rango de una matriz es igual al rango de su transpuesta, pues las filas de una son las columnas de la otra.

TEOREMA 2.20 (Rouche-Frobenius (2^{da} versión)). *Sea (S) un sistema lineal de m ecuaciones con n incógnitas con matriz ampliada $A|b$. entonces*

1. (S) *es compatible si y solo si* $rango(A) = rango(A|b)$
2. *Si* (S) *es compatible entonces*
   a) (S) *es determinado si y solo si* $rango(A) = n$.
   b) (S) *es indeterminado si y solo si* $rango(A) < n$.

Demostración. La prueba es consecuencia inmediata del Teorema de Rouche-Frobenius (1era. versión) y del lema 2.17 □

2.7. Matrices invertibles y rango

Todo número real no nulo es invertible, sin embargo tal como se vio en el ejemplo 2.10 existen matrices no nulas que no poseen inversa, se impone por lo tanto obtener algún resultado que caracterice a las matrices invertibles, observemos que de (2.7) se deduce que si $A \in \mathcal{M}_{n\times n}$ es una matriz invertible el sistema $AX = b$ es compatible determinado por lo tanto el teorema de Rouche-Frobenius asegura que $rango(A) = n$, probaremos en esta sección como principal resultado que el recíproco es también cierto.

Comencemos mostrando que las matrices de rango n son exactamente aquellas que admiten una “inversa a derecha”.

LEMA 2.21. *Sea* $A \in \mathcal{M}_{n\times n}$ *entonces* $rango(A) = n$ *si y solo si existe* $B \in \mathcal{M}_{n\times n}$ *tal que* $A{\cdot}B = I_n$.

Demostración. ($\Rightarrow$) Recordemos en primer lugar que de acuerdo a la observación 2.1 si A y B son dos matrices $n \times n$ cualquiera y B^j indica la columna j-esima de B entonces

$$A{\cdot}B = \begin{pmatrix} A{\cdot}B^1 & A{\cdot}B^2 & \dots & A{\cdot}B^n \\ \Big| & \Big| & & \Big| \\ \Big| & \Big| & \dots & \Big| \end{pmatrix}$$

Por lo tanto

$$A{\cdot}B = I_n \Longleftrightarrow A{\cdot}B^j = e^j, \ \ \forall j = 1, 2, \dots, n$$

Donde

$$e^j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \longleftarrow \text{fila } j\text{-esima}$$

Ahora bien, como $rango(A) = n$ el teorema de Rouché-Frobenius asegura que n sistemas de ecuaciones $AX = e^j$ con $j = 1, 2, \ldots, n$ son compatibles y consecuentemente se tiene que $\forall\, j = 1, 2, \ldots, n$ existe B^j tal que $A{\cdot}B^j = e_j$. Por lo tanto eligiendo B como la matriz cuya columna j-esima es B^j se tiene probado el directo.

($\Leftarrow$) Veamos ahora el recíproco. Supongamos que existe B tal que $A{\cdot}B = I_n$. Entonces, llamando B^j a la columna j-esima de la matriz B se tiene

$$\left(\begin{array}{cccc} A{\cdot}B^1 & A{\cdot}B^2 & \ldots & A{\cdot}B^n \\ \Big| & \Big| & & \Big| \\ \Big| & \Big| & \ldots & \Big| \end{array}\right) = A{\cdot}B = I_n = \left(\begin{array}{cccc} e^1 & e^2 & \ldots & e^n \\ \Big| & \Big| & & \Big| \\ \Big| & \Big| & \ldots & \Big| \end{array}\right)$$

Mirado los extremos de la igualdad anterior se deduce que los sistemas

$$AX = e^j \quad \forall j = 1, 2, \ldots, n$$

son todos compatibles.

El $rango(A) \leq n$ pues A es $n \times n$, supongamos por absurdo que $rango(A) = r < n$. Entonces, en virtud del lema 2.17, una forma escalerizada de la matriz A debe tener $r < n$ escalones (filas no nulas). Por lo tanto, al aplicar el algoritmo de Gauss a la matriz A alguna fila, supongamos que la h, debió anularse. Pero entonces el sistema $AX = e^h$ resulta incompatible pues al escalerizarlo, la h- esima ecuación se transforma en $0 = 1$ y esto naturalmente es absurdo pues todos los sistemas $AX = e^j$ eran compatibles $\forall\, j = 1, 2, \ldots, n$, así que debe ser $rango(A) = n$ como queríamos. $\square$

Estamos ahora en condiciones de probar que que para que una matriz sea invertible es suficiente que tenga un “inverso a derecha”

Demostración del lema 2.9. El directo es consecuencia inmediata de la definición de matriz invertible por lo cual sólo probaremos el recíproco.

($\Leftarrow$) Supongamos que existe $X \in \mathcal{M}_{n\times n}$ tal que $A{\cdot}X = I_n$ entonces por el lema 2.21 $rango(A) = n$. De la observación 2.19 se deduce que

$$rango(A^t) = rango(A) = n,$$

por lo que, utilizando nuevamente el lema 2.21, existe $C \in \mathcal{M}_{n\times n}$ tal que $A^t{\cdot}C = I_n$. Transponiendo se tiene que

$$C^t{\cdot}A = \left(A^t{\cdot}C\right)^t = (I_n)^t = I_n.$$

Por lo tanto $D = C^t$ es una "inversa a izquierda" de A.
Veamos que $D = X$, en efecto

$$D = D{\cdot}I_n = D\big(A{\cdot}X\big) = \big(D{\cdot}A\big)X = I_n{\cdot}X = X$$

entonces

$$A{\cdot}X = X{\cdot}A = I_n$$

y consecuentemente A es invertible. $\square$

TEOREMA 2.22. *Sea $A \in \mathcal{M}_{n\times n}$, A es invertible si y solo si $rango(A) = n$.*

Demostración. ($\Rightarrow$) Si A es invertible en particular existe B tal que $A{\cdot}B = I_n$ entonces lema 2.9 $rango(A) = n$.

($\Leftarrow$) Si $rango(A) = n$ el lema 2.21 asegura que existe $B \in \mathcal{M}_{n\times n}$ tal que $A{\cdot}B = I_n$ pero entonces por el lema 2.9 A es invertible. $\square$

2.8. Matrices elementales

Culminemos el capítulo mostrando que el proceso de escalerización puede representarse mediante el producto por ciertas matrices especiales llamadas matrices elementales.

Sea e una de las transformaciones elemental de la definición 1.1, esto es

1- multiplicar una fila por una constante no nula

2- intercambio de filas

3- sumarle a una fila un múltiplo de otra fila

DEFINICIÓN 2.12. La **matriz elemental** E **asociada con la transformación** e es la que se obtiene al aplicar la transformación e a la matriz identidad $n \times n$

EJEMPLO 2.22. La matriz

$$E_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 & \alpha & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{array}{l} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \longleftarrow \text{Fila i} \\ \vdots \\ \vdots \\ \vdots \end{array} \tag{2.15}$$

está asociada a la transformación elemental e_1 **"multiplicar la fila** i **por la constante** $\alpha \neq 0$**"**, pues E_1 se obtiene de la matriz identidad I aplicando dicha transformación

$$I \quad \xrightarrow{\text{operación } e_1} \quad E_1$$

EJEMPLO 2.23. La matriz

(2.16)
$$E_2 = \begin{pmatrix}
1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
\cdots & \cdots & 0 & \cdots & \cdots & \cdots & 1 & \cdots & \vdots \\
\vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \cdots & 1 & \cdots & \cdots & \vdots \\
\vdots & \vdots & 1 & \cdots & \cdots & \cdots & 0 & \cdots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1
\end{pmatrix}
\begin{array}{l}
\quad\vdots \\
\quad\vdots \\
\longleftarrow \text{Fila j} \\
\quad\vdots \\
\quad\vdots \\
\quad\vdots \\
\longleftarrow \text{Fila i} \\
\quad\vdots \\
\quad\vdots
\end{array}$$

está asociada a la transformación elemental e_2 **"intercambiar la fila** i **por la fila** j**"**, pues E_2 se obtiene de la matriz identidad I aplicando dicha transformación.

$$I \quad \xrightarrow{\text{operación } e_2} \quad E_2$$

La matriz

(2.17)
$$E_3 = \begin{pmatrix}
1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
\cdots & \cdots & 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \cdots & 1 & \cdots & \cdots & \vdots \\
\vdots & \vdots & \alpha & \cdots & \cdots & \cdots & 1 & \cdots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1
\end{pmatrix}
\begin{array}{l}
\quad\vdots \\
\quad\vdots \\
\longleftarrow \text{Fila j} \\
\quad\vdots \\
\quad\vdots \\
\quad\vdots \\
\longleftarrow \text{Fila i} \\
\quad\vdots \\
\quad\vdots
\end{array}$$

está asociada a la transformación elemental e_3 **"sumarle a fila i un múltiplo de la fila j"**, pues E_3 se obtiene de la matriz identidad I aplicando dicha transformación

$$I \quad \xrightarrow{\text{operación } e_3} \quad E_3$$

PROPOSICIÓN 2.23. *Sean: B la matriz (de tamaño $n \times m$) que se obtiene a partir de la matriz A (de tamaño $n \times m$) realizando la transformación elemental e con las filas de A*

$$A \quad \xrightarrow{\text{operación } e} \quad B;$$

E la matriz elemental (de tamaño $n \times n$) asociada con la transformación e

$$I \quad \xrightarrow{\text{operación } e} \quad E$$

Se cumple que $\quad B = EA.$

Demostración. Sea B la matriz que se obtiene de A multiplicando por α la fila i de A

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Si E_1 es la matriz (2.15) asociada a la transformación elemental "multiplicar la fila i por la constante $\alpha \neq 0$", es inmediato verificar que

$$\begin{aligned} E_1 A &= \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & \cdots & \alpha & \cdots & 0 \\ \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \\ &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = B \end{aligned}$$

Análogamente se prueba la propiedad para las dos restantes operaciones elementales. $\square$

EJEMPLO 2.24. Consideremos la matriz

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix}, \text{ y la matriz elemental } E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

asociada a la transformación elemental e de sumarle a la tercera fila 3 veces la primer fila. Sea B la matriz que se obtiene de A si se le aplica la transformación elemental considerada

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} \xrightarrow{\text{operación } e} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{pmatrix} = B$$

$$\text{y operando se verifica que } B = EA = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{pmatrix}$$

■

Sea A una matriz cualquiera, en virtud de la proposición 1.3 (Método de Gauss) puede escalerizarse mediante una cantidad finita, $e_1, e_2 \dots, e_k$ de transformaciones elementales. Sea B una forma escalerizada de A, esquematizamos como sigue el proceso de escalerización:

$$A \quad \xrightarrow{\text{operación } e_1} \quad \cdots\cdots \quad \xrightarrow{\text{operación } e_k} \quad B.$$

Si representamos estas operaciones elementales por sus respectivas matrices elementales asociadas, se puede enunciar que

PROPOSICIÓN 2.24. *Toda matriz A puede reducirse mediante una sucesión de matrices elementales $E_1, E_2, \dots, E_k$ a su forma escalerizada B, esto es:* $E_k \dots E_2 E_1 A = B$.

PROPOSICIÓN 2.25. *Sea A una matriz $n \times n$. Si $rango(A) = n \Rightarrow$ existe una sucesión de matrices elementales $E_1, E_2 \dots, E_h$ tales que*

$$E_h \dots E_2 E_1 A = I.$$

(es decir que al aplicarle el proceso de escalerización se obtiene la matriz identidad)

Demostración. Recordemos primeramente que el rango de una matriz es el número de escalones de su forma escalerizada (filas no nulas). Sea A' la forma reducida de A. Como $rango(A) = n$ y A (y por lo tanto A') es una matriz $n \times n$ se tiene que

$$A' = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{array}{l} \qquad\qquad\quad \vdots \\ \qquad\qquad\quad \vdots \\ \qquad\qquad\quad \vdots \\ \leftarrow n \text{ escalones no nulos y } n \text{ filas} \end{array}$$

Así por la proposición 2.24 existen matrices elementales $E_1, E_2 \dots, E_h$ tales que $E_1 E_2 \dots E_h A = I$. □

Cada transformación elemental puede ser deshecha por su transformación elemental inversa, esto es

1- multiplicar una fila por una constante $\alpha \neq 0$ se deshace multiplicando dicha fila por $\frac{1}{\alpha}$

2- intercambio de filas se deshace intercanbiándolas nuevamente

3- sumarle a una fila un múltiplo de otra fila se deshace restándole a la fila el múltiplo de la otra

PROPOSICIÓN 2.26. *Toda matriz elemental es invertible y su inversa es una matriz elemental*

DEMOSTRACIÓN. Sea E la matriz elemental asociada a la transformación elemental e

$$I \xrightarrow{\text{operación } e} E$$

Indiquemos por e^{-1} a la transformación elemental inversa de e, es decir que

$$I \xrightarrow{\text{operación } e} E \xrightarrow{\text{operación } e^{-1}} I$$

Si E' es la matriz elemental asociada a la transformación elemental e^{-1}, esto es

$$I \xrightarrow{\text{operación } e^{-1}} E'$$

aplicando la proposición 2.23 a la matriz E con la transformación e^{-1} se tiene que $E'E = I$ Análogamente se prueba que $EE' = I$. Esto es E es invertible y su inversa es $E^{-1} = E'$ (es una matriz elemental correspondiente a la transformación inversa). □

Obtengamos una nueva caracterización de las matrices invertibles

COROLARIO 2.27. *Sea A una matriz $n \times n$ entonces A es invertible si y solo si existen $E_1, E_2, \ldots, E_k$ matrices elementales tales que $A = E_1 \cdot E_2 \cdot \ldots \cdot E_k$*

DEMOSTRACIÓN. ($\Rightarrow$) Si A es invertible entonces en virtud de la proposición anterior existen $E'_1, E'_2, \ldots, E'_k$ una sucesión de matrices elementales tales que

$$E'_1 \cdot E'_2 \cdot \ldots \cdot E'_k \cdot A = I_n$$

pero como las matrices elementales son invertibles y sus inversas son también matrices elementales se tiene

$$\left({E'_k}^{-1} \cdot \ldots \cdot {E'_2}^{-1} \cdot {E'_1}^{-1}\right) E'_1 \cdot E'_2 \cdot \ldots \cdot E'_k \cdot A = \left({E'_k}^{-1} \cdot \ldots \cdot {E'_2}^{-1} \cdot {E'_1}^{-1}\right) I_n$$

de donde se tiene que

$$A = {E'_k}^{-1} \cdot \ldots \cdot {E'_2}^{-1} \cdot {E'_1}^{-1}$$

Por lo tanto poniendo $E_1 = {E'_k}^{-1}$, $E_2 = {E'_{k-1}}^{-1}, \ldots, E_k = {E'_1}^{-1}$ se tiene probado el directo.

($\Leftarrow$) El recíproco es consecuencia de que el producto de matrices invertibles es invertible. $\square$

CAPíTULO 3

DETERMINANTES

El determinante de una matriz cuadrada es un único número que se asocia a dicha matriz; es por tanto una función del conjunto de las matrices cuadradas en el conjunto numérico al que pertenecen los elementos de las matrices. En nuestro caso estos números serán los reales o los complejos, pero se puede dar sobre conjuntos “numéricos” más generales.

El uso del determinante surgió de las fórmulas que dan las soluciones de sistemas de n ecuaciones con n incógnitas, luego fue identificado (en el caso 3 por 3) como área de paralelogramo o volumen de un paralelepípedo, hasta extenderse a definiciones más generales que la que nosotros daremos en este curso (función multilineal alternada).

Más adelante se verá que podemos hablar del determinante de una transformación lineal entre espacios vectoriales, a la que se puede asociar matrices de una manera sencilla. En particular las matrices invertibles son las únicas que tienen determinantes distinto de cero. Así, una propiedad tan definitoria de una matriz (su invertibilidad) estará caracterizada por la no nulidad de su determinante (o sea por el valor de un **único** número).

3.1. Definición

La definición de determinante de una matriz cuadrada será dada de manera inductiva en el número de filas (o de columnas). O sea, daremos la definición de determinante de una matriz $n \times n$ a partir del conocimiento de los determinantes de matrices $(n-1) \times (n-1)$. El determinante de una matriz A se representa con el símbolo $|A|$; tratándose de una matriz dada por sus coeficientes, en general no escribiremos los paréntesis con los que en general encerramos el çuadrado" de los números.

DEFINICIÓN 3.1. Sea $A = ((a_{ij}))$ una matriz $n \times n$, se define la **matriz adjunta del elemento** a_{ij} como la submatriz A_{ij} de la matriz A que se obtiene eliminado la fila i y la columna j de A

EJEMPLO 3.1. Si $A = \begin{pmatrix} 3 & -4 & -3 \\ -1 & 2 & 7 \\ -2 & 4 & 0 \end{pmatrix}$, la matriz adjunta del elemento $a_{21} = -1$, se obtiene eliminando la fila 2 y la columna 1 de A:
$\begin{pmatrix} \times & -4 & -3 \\ \times & \times & \times \\ \times & 4 & 0 \end{pmatrix}$, obteniéndose la matriz adjunta $A_{21} = \begin{pmatrix} -4 & -3 \\ 4 & 0 \end{pmatrix}$ ■

OBSERVACIÓN 3.1. Si la matriz cuadrada A es de tamaño n, las matrices adjuntas A_{ij} son de tamaño $(n-1) \times (n-1)$.

DEFINICIÓN 3.2. **(Inductiva en el tamaño de la matriz)** El determinante de una matriz 1×1 es el propio número. El determinante de la matriz 2×2, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es el número $|A| \overset{def}{=} ad - bc$
El determinante de una matriz A $n \times n$ se define como el número

$$|A| \overset{def}{=} (-1)^{1+1} a_{11} |A_{11}| + \ldots + (-1)^{i+1} a_{i1} |A_{i1}| + \ldots + (-1)^{n+1} a_{n1} |A_{n1}|$$

EJEMPLO 3.2. (a) Calculemos el determinante de la matriz $A = \begin{pmatrix} 3 & 1 \\ 7 & 2 \end{pmatrix}$

$$|A| = (3)(2) - (7)(2) = -8$$

(b) Calculemos el determinante de la matriz $A = \begin{pmatrix} 3 & -4 & -3 \\ -1 & 2 & 7 \\ -2 & 4 & 0 \end{pmatrix}$

$$\begin{aligned}
|A| &= (-1)^{1+1} a_{11} |A_{11}| + (-1)^{2+1} a_{21} |A_{21}| + (-1)^{3+1} a_{31} |A_{31}| = \\
&= 3 \begin{vmatrix} 2 & 7 \\ 4 & 0 \end{vmatrix} - (-1) \begin{vmatrix} -4 & -3 \\ 4 & 0 \end{vmatrix} + (-2) \begin{vmatrix} -4 & -3 \\ 2 & 7 \end{vmatrix} = \\
&= 3(-28) + (-12) + (-2)(-22) = -52
\end{aligned}$$

(c) Calculemos el determinante de la matriz $A = \begin{pmatrix} 2 & 0 & 1 & 1 \\ 2 & 2 & -4 & -1 \\ -3 & 1 & 1 & 1 \\ 2 & -2 & 0 & 0 \end{pmatrix}$

$$\begin{aligned}
|A| = (-1)^{1+1} a_{11} |A_{11}| + (-1)^{2+1} a_{21} |A_{21}| + \\
+ (-1)^{3+1} a_{31} |A_{31}| + (-1)^{4+1} a_{41} |A_{41}| = \\
2 \begin{vmatrix} 2 & -4 & -1 \\ 1 & 1 & 1 \\ -2 & 0 & 0 \end{vmatrix} - 2 \begin{vmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ -2 & 0 & 0 \end{vmatrix} + \\
+ (-3) \begin{vmatrix} 0 & 1 & 1 \\ 2 & -4 & -1 \\ -2 & 0 & 0 \end{vmatrix} - 2 \begin{vmatrix} 0 & 1 & 1 \\ 2 & -4 & -1 \\ 1 & 1 & 1 \end{vmatrix}
\end{aligned}$$

Ahora bien, los determinantes de tamaño 3×3 se calculan del mismo modo que en (b). Verificar que respectivamente, los determinantes tres por tres valen $6, 0, -6, 3$ por lo cual

$$|A| = 2(6) - 3(-6) - 2(3) = 24$$

■

3.2. Propiedades de los determinantes

1. **Linealidad (respecto a una fila)**

a) **Aditividad (respecto a una fila)**

$$|C| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = |A| + |B|$$

Demostración. Por inducción completa

Para $n = 2$. Por un lado

$$|C| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix} = a_{11}a_{22} + a_{11}b_{22} - a_{12}a_{21} - a_{12}b_{21}$$

y por otro

$$|A|+|B| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{vmatrix} = (a_{11}a_{22} - a_{12}a_{21}) + (a_{11}b_{22} - a_{12}b_{21})$$

Con lo cual $|C| = |A| + |B|$.

Hipótesis inductiva: La propiedad es válida para matrices de orden menor que n

Tesis inductiva: La propiedad es válida para matrices de orden n. En efecto

$$(3.18) \qquad \begin{gathered} |C| \overset{def}{=} (-1)^{1+1} a_{11} |C_{11}| + \ldots + \\ + (-1)^{i+1} (a_{i1} + b_{i1}) |C_{i1}| + \ldots + (-1)^{n+1} a_{n1} |C_{n1}| \end{gathered}$$

Observemos que si $k \neq i$, como la matriz C_{k1} es de orden $n-1$, por la hipótesis inductiva se tiene que

$$|C_{k1}| = \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{k-12} & \cdots & a_{k-1n} \\ a_{k+12} & \cdots & a_{k+1n} \\ \vdots & \ddots & \vdots \\ a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{k-12} & \cdots & a_{k-1n} \\ a_{k+12} & \cdots & a_{k+1n} \\ \vdots & \ddots & \vdots \\ a_{i2} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{k-12} & \cdots & a_{k-1n} \\ a_{k+12} & \cdots & a_{k+1n} \\ \vdots & \ddots & \vdots \\ b_{i2} & \cdots & b_{in} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = |A_{k1}| + |B_{k1}|\,;$$

luego si sustituimos en (3.18) se tiene que

$$\begin{aligned} |C| &= (-1)^{1+1}\, a_{11}\, [|A_{11}| + |B_{11}|] + \ldots + (-1)^{i+1}\, (a_{i1} + b_{i1})\, |C_{i1}| + \ldots \\ &\qquad \ldots + (-1)^{n+1}\, a_{n1}\, [|A_{n1}| + |B_{n1}|] \\ &= (-1)^{1+1}\, a_{11}\, |A_{11}| + \ldots + (-1)^{i+1}\, a_{i1}\, |C_{i1}| + \ldots + (-1)^{n+1}\, a_{n1}\, |A_{n1}| + \\ &\quad + (-1)^{1+1}\, a_{11}\, |B_{11}| + \ldots + (-1)^{i+1}\, b_{i1}\, |C_{i1}| + \ldots + (-1)^{n+1}\, a_{n1}\, |B_{n1}| \end{aligned}$$

$$\text{Pero como } |C_{i1}| = \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{i-12} & \cdots & a_{i-1n} \\ a_{i+12} & \cdots & a_{i+1n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = |A_{i1}| = |B_{i1}| \text{ resulta que}$$

$$\begin{aligned} |C| = (-1)^{1+1} a_{11} |A_{11}| + \ldots + (-1)^{i+1} a_{i1} |A_{i1}| + \ldots + (-1)^{n+1} a_{n1} |A_{n1}| + \\ + (-1)^{1+1} a_{11} |B_{11}| + \ldots + (-1)^{i+1} b_{i1} |B_{i1}| + \ldots + (-1)^{n+1} a_{n1} |B_{n1}| \\ = |A| + |B| \end{aligned}$$

□

OBSERVACIÓN 3.2. *No es cierto que* $|A+B| = |A| + |B|$

$$|A| = \begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix} = -10, \quad |B| = \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = 2 \text{ pero}$$

$$|A+B| = \begin{vmatrix} 3 & 3 \\ 4 & 4 \end{vmatrix} = 0 \neq -7 = |A| + |B|.$$

b) **Homogeneidad (respecto a una fila)**
Si B *es una matriz que se obtiene de* A *multiplicando una de sus filas por un número* α *entonces*

$$|B| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \alpha |A|$$

Demostración. Por inducción completa
Para $n = 2$. Por un lado

$$|B| = \begin{vmatrix} a_{11} & a_{12} \\ \alpha a_{21} & \alpha a_{22} \end{vmatrix} = a_{11}\alpha a_{22} - a_{12}\alpha a_{21} =$$
$$= \alpha\left(a_{11}a_{22} - a_{12}a_{21}\right) = \alpha \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \alpha\,|A|$$

Hipótesis inductiva: La propiedad es válida para matrices de orden menor que n
Tesis inductiva: La propiedad es válida para matrices de orden n. En efecto

$$|B| \overset{def}{=} (-1)^{1+1}\,a_{11}\,|B_{11}| + \ldots + \\ + (-1)^{i+1}\,\alpha a_{i1}\,|B_{i1}| + \ldots + (-1)^{n+1}\,a_{n1}\,|B_{n1}| \tag{3.19}$$

Observemos que si $k \neq i$, como la matriz C_{k1} es de orden $n-1$, por la hipótesis inductiva se tiene que

$$|B_{k1}| = \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{k-12} & \cdots & a_{k-1n} \\ a_{k+12} & \cdots & a_{k+1n} \\ \vdots & \ddots & \vdots \\ \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{k-12} & \cdots & a_{k-1n} \\ a_{k+12} & \cdots & a_{k+1n} \\ \vdots & \ddots & \vdots \\ a_{i2} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = \alpha\,|A_{k1}|$$

$$\text{y que } |B_{i1}| = \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{i-12} & \cdots & a_{i-1n} \\ a_{i+12} & \cdots & a_{i+1n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = |A_{i1}|$$

luego si sustituimos en (3.19) se tiene que

$$\begin{aligned} |B| = (-1)^{1+1} a_{11}\alpha |A_{11}| + \ldots + \\ + (-1)^{i+1} \alpha a_{i1} |A_{i1}| + \ldots + (-1)^{n+1} a_{n1}\alpha |A_{n1}| \\ = \alpha \Big[(-1)^{1+1} a_{11} |A_{11}| + \ldots + (-1)^{i+1} a_{i1} |A_{i1}| + \ldots + \\ + (-1)^{n+1} a_{n1} |A_{n1}| \Big] = \alpha |A| \end{aligned}$$

□

COROLARIO 3.3. *Se deduce de la propiedad anterior que* $|\alpha A| = \alpha^n |A|$.

Demostración. En efecto

$$|\alpha A| = \begin{vmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{n1} & \alpha a_{n2} & \cdots & \alpha a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{n1} & \alpha a_{n2} & \cdots & \alpha a_{nn} \end{vmatrix} =$$

$$= \alpha\alpha \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{n1} & \alpha a_{n2} & \cdots & \alpha a_{nn} \end{vmatrix} = \alpha\alpha\ldots\alpha \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \\ = \alpha^n |A|$$

□

2. **Intercambio de filas**

   Si B es la matriz que se obtiene de A intercambiando dos de sus filas entonces

$$|B| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = -|A|$$

   DEMOSTRACIÓN. Por inducción completa

   Para $n = 2$. Por un lado $|B| = \begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix} = a_{12}a_{21} - a_{11}a_{22}$,

   y por otro $|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}a_{22} - a_{12}a_{21})$.

   Así $\quad |B| = -|A|$.

   Hipótesis inductiva: La propiedad es válida para matrices de orden menor que n

   Tesis inductiva: La propiedad es válida para matrices de orden n

   Primer caso: La matriz B se ha obtenido de A intercambiando las filas consecutivas i e $i+1$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Por la definición, se tiene que

$$|B| = (-1)^{1+1}\, a_{11}\, |B_{11}| + \ldots + (-1)^{i+1}\, a_{i+11}\, |B_{i1}| \\ + (-1)^{i+2}\, a_{i1}\, |B_{i+11}| \ldots + (-1)^{n+1}\, a_{n1}\, |B_{n1}| \tag{3.20}$$

Pero siendo $B = \begin{array}{cl} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \quad\vdots \\ \quad\vdots \\ \longleftarrow \text{Fila i} \\ \longleftarrow \text{Fila i}+1 \\ \quad\vdots \\ \quad\vdots \end{array} \end{array}$

se tiene que $B_{i1} = A_{i+11}$ y $B_{i+11} = A_{i1}$, por lo cual

$$|B_{i1}| = |A_{i+11}|\,, \quad |B_{i+11}| = |A_{i1}|$$

y si $k \neq i$ y $k \neq i+1$, se tiene que B_{k1} es una matriz de orden $n-1$ con dos filas intercambiadas respecto de A, por lo cual, por la hipótesis inductiva $|B_{k1}| = -\,|A_{k1}|$. Sustituimos en (3.20)

$$|B| = (-1)^{1+1}\, a_{11}\, [-\,|A_{11}|] + \ldots + (-1)^{i+1}\, a_{i+11}\, |A_{i+11}| \\ + (-1)^{i+2}\, a_{i1}\, |A_{i1}| + \ldots + (-1)^{n+1}\, a_{n1}\, [-\,|A_{n1}|]$$

Pero

$$(-1)^{i+1}\, a_{i+11}\, |A_{i+11}| = (-1)^{i+1}\, (-1)^2\, a_{i+11}\, |A_{i+11}| = \\ = (-1)^{i+1}\, (-1)\, a_{i+11}\, (-1)\, |A_{i+11}| = (-1)^{i+2}\, a_{i+11}\, [-\,|A_{i+11}|] = \\ = (-1)^{i+1}\, (-1)\, a_{i1}\, |A_{i1}| = (-1)^{i+1}\, a_{i1}\, [-\,|A_{i1}|]$$

con lo cual

$$\begin{aligned}
|B| = (-1)^{1+1}\, a_{11}\, [-\,|A_{11}|] + \ldots + (-1)^{i+2}\, a_{i+11}\, [-\,|A_{i+11}|] + \\
+ (-1)^{i+1}\, a_{i1}\, [-\,|A_{i1}|] + \ldots + (-1)^{n+1}\, a_{n1}\, [-\,|A_{n1}|] \\
= -[(-1)^{1+1}\, a_{11}\, |A_{11}| + \ldots + (-1)^{i+2}\, a_{i+11}\, |A_{i+11}| + \\
+ (-1)^{i+1}\, a_{i1}\, |A_{i1}| + \ldots + (-1)^{n+1}\, a_{n1} - |A_{n1}|] = -\,|A|
\end{aligned}$$

Segundo caso La matriz B se ha obtenido de A intercambiando las filas i y $j = k{+}1$. La matriz B se puede obtener de A realizando varios cambios de filas consecutivas. En primer lugar realizamos k cambios consecutivos de la fila i de A hasta colocarla en la fila $j = i + k$.

$$A = \left(\begin{array}{cccc}
a_{11} & a_{12} & \cdots & a_{1n} \\
\vdots & \vdots & \cdots & \vdots \\
a_{i1} & a_{i2} & \cdots & a_{in} \\
a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\
a_{i+21} & a_{i+22} & & a_{i+2n} \\
\vdots & \vdots & \cdots & \vdots \\
a_{k+11} & a_{k+12} & \cdots & a_{k+1n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{array}\right)
\begin{array}{l}
\vdots \\
\vdots \\
\longleftarrow \text{ fila i de } A \\
\longleftarrow \text{ fila i} + 1 \text{ de } A \\
\longleftarrow \text{ fila i} + 2 \text{ de } A \\
\vdots \\
\longleftarrow \text{ fila i} + \text{k de } A \\
\vdots \\
\vdots
\end{array}$$

$$\underset{\longrightarrow}{\text{cambio de la fila i por la i} + 1 \text{ de } A}$$

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i+21} & a_{i+22} & \cdots & a_{i+2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{array}{l} \vdots \\ \vdots \\ \longleftarrow \text{ fila i de } A_1 \\ \longleftarrow \text{ fila i} + 1 \text{ de } A_1 \\ \longleftarrow \text{ fila i} + 2 \text{ de } A_1 \\ \vdots \\ \longleftarrow \text{ fila i} + \text{k de } A_1 \\ \vdots \\ \vdots \end{array}$$

$$\xrightarrow{\text{cambio de la fila i} + 1 \text{ por la i} + 2 \text{ de } A_1}$$

$$A_2 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\ a_{i+21} & a_{i+22} & \cdots & a_{i+2n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \dots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{array}{l} \vdots \\ \vdots \\ \longleftarrow \text{ fila i de } A_2 \\ \longleftarrow \text{ fila i} + 1 \text{ de } A_2 \\ \longleftarrow \text{ fila i} + 2 \text{ de } A_2 \\ \vdots \\ \longleftarrow \text{ fila i} + \text{k de } A_2 \\ \vdots \\ \vdots \end{array}$$

$$\longrightarrow \cdots \longrightarrow \cdots \longrightarrow$$

$$A_{k-1} = \begin{array}{cl} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\ a_{i+21} & a_{i+22} & \cdots & a_{i+2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i+k1} & a_{i+k2} & \cdots & a_{i+kn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \vdots \\ \vdots \\ \longleftarrow \text{ fila i de } A_{k-1} \\ \longleftarrow \text{ fila i} + 1 \text{ de } A_{k-1} \\ \vdots \\ \longleftarrow \text{ fila i} + \text{k} - 1 \text{ de } A_{k-1} \\ \longleftarrow \text{ fila i} + \text{k de } A_{k-1} \\ \vdots \\ \vdots \end{array} \end{array}$$

$$\xrightarrow{\text{cambio de la fila i} + \text{k} - 1 \text{ por la i} + \text{k de } A_{k-1}}$$

$$A_{k} = \begin{array}{cl} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1n} \\ a_{i+21} & a_{i+22} & \cdots & a_{i+2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i+k1} & a_{i+k2} & \cdots & a_{i+kn} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \vdots \\ \vdots \\ \longleftarrow \text{ fila i de } A_{k} \\ \longleftarrow \text{ fila i} + 1 \text{ de } A_{k} \\ \vdots \\ \longleftarrow \text{ fila i} + \text{k} - 1 \text{ de } A_{k} \\ \longleftarrow \text{ fila i} + \text{k de } A_{k} \\ \vdots \\ \vdots \end{array} \end{array}$$

Como los k cambios de filas son consecutivos, por la primer parte, se tiene que

$$\begin{aligned} |A_1| &= -|A| \\ |A_2| &= -|A_1| \\ &\vdots \\ |A_k| &= -|A_{k-1}| \end{aligned}$$

Con lo cual $|A_k| = (-1)^k |A|$.

En segundo lugar, (repetimos el procedimiento "hacia arriba") realizamos $k-1$ cambios consecutivos de la fila $i+k-1$ de A_K hasta colocarla en la fila i, obteniéndose la matriz B y se cumple, aplicando la primer parte $|B| = (-1)^{k-1} |A_k|$ De las dos últimas igualdades se deduce $|B| = (-1)^{2k-1} |A| = -|A|$. $\square$

3. **Determinantes nulos**

   a) *Si A tiene una fila de ceros entonces $|A| = 0$*

   Demostración. En efecto

$$
\begin{aligned}
|A| &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ 1-1 & 1-1 & \cdots & 1-1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\
&= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ -1 & -1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \\
&= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0
\end{aligned}
$$

$\square$

   b) *Si A tiene dos filas iguales entonces $|A| = 0$*

DEMOSTRACIÓN. En efecto, si se intercambian dos filas, sabemos que el determinante cambia de signo, con lo cual

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = -|A|$$

Así $2|A| = 0 \Rightarrow |A| = 0$ □

c) *Si A tiene dos filas proporcionales entonces* $|A| = 0$

DEMOSTRACIÓN. En efecto

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$$

□

d) *Si A tiene una fila combinación lineal de las otras filas entonces* $|A| = 0$

DEMOSTRACIÓN. Sin pérdida de generalidad, para simplificar la notación, supongamos que la última fila es combinación lineal de las anteriores. Aplicando la propiedad de linealidadad

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{i=1}^{i=n-1} \alpha_i a_{i1} & \sum\limits_{i=1}^{i=n-1} \alpha_i a_{i2} & \cdots & \sum\limits_{i=1}^{i=n-1} \alpha_i a_{in} \end{vmatrix} =$$

$$= \sum_{i=1}^{i=n-1} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_i a_{i1} & \alpha_i a_{i2} & \cdots & \alpha_i a_{in} \end{vmatrix} =$$

$$= \sum_{i=1}^{i=n-1} \alpha_i \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \end{vmatrix} = 0$$

(pues los determinantes tienen dos filas iguales) □

4. **Combinación lineal de filas**

   a) *Si B es una matriz que se obtiene de A sumado a una fila de A un múltiplo de otra fila de A entonces $|B| = |A|$*

Demostración. En efecto

$$|B| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1}+\alpha a_{i1} & a_{j2}+\alpha a_{i2} & \cdots & a_{jn}+\alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = |A|$$

□

b) *Si B es una matriz que se obtiene de A cambiando la fila A_j por la combinación lineal $\beta A_j + \alpha A_i$ con $\beta \neq 0$ ($i \neq j$).entonces* $|B| = \beta\,|A|$

Demostración. En efecto

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ \beta a_{j1} + \alpha a_{i1} & \beta a_{j2} + \alpha a_{i2} & \cdots & \beta a_{jn} + \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ \beta a_{j1} & \beta a_{j2} & \cdots & \beta a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \beta \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \beta\,|A|$$

$\square$

Ejercicio Probar que si $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$ (ceros por debajo de la diagonal principal), entonces $|A| = a_{11}a_{22}\ldots a_{nn}$

EJEMPLO 3.3. $\begin{vmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{vmatrix} \overset{(A)}{=} - \begin{vmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix} \overset{(B)}{=} -3 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix} \overset{(C)}{=}$
$-3 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 10 & -5 \end{vmatrix} \overset{(D)}{=}$

$$-(3)(5) \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 2 & -1 \end{vmatrix} \overset{(E)}{=} -(3)(5) \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -11 \end{vmatrix} = -(3)(5) \begin{vmatrix} 1 & 5 \\ 0 & -11 \end{vmatrix} =$$
$-(3)(5)(-11) = 165$

(A) Se intercambian las filas 1 y 2
(B) La fila 1 es múltiplo de 3
(C) A la fila 3 se le restó 2 veces la fila 1
(D) La fila 3 es múltiplo de 5
(E) A la fila 3 se le restó 2 veces la fila 2 ■

NOTAS:

1) Desarrollo por cualquier fila o columna

El determinante se ha definido usando los elementos de la primer columna de la matriz A con sus respectivas matrices adjuntas

$$|A| \overset{def}{=} (-1)^{1+1} a_{11} |A_{11}| + \ldots + (-1)^{i+1} a_{i1} |A_{i1}| + \ldots + (-1)^{n+1} a_{n1} |A_{n1}|$$
(desarrollo por la primera columna)

Pero se puede probar [1] que dicha definición coincide con cualquier desarrollo por cualquier columna o fila de A:

$$|A| \overset{def}{=} (-1)^{i+1} a_{i1} |A_{i1}| + (-1)^{i+2} a_{i2} |A_{i2}| + \ldots + (-1)^{i+n} a_{in} |A_{in}|$$

(desarrollo por la i – ésima fila)

$$|A| \overset{def}{=} (-1)^{i+1} a_{1i} |A_{1i}| + (-1)^{i+2} a_{2i} |A_{2i}| + \ldots + (-1)^{i+n} a_{ni} |A_{ni}|$$

(desarrollo por la i – ésima columna)

2) Propiedades "por columnas"

También se puede probar que

$$|A| = \left|A^t\right|$$

con lo cual en todas las propiedades anteriores se puede cambiar la palabra "fila" por "columna", es decir que las propiedades valen para las columnas

3.3. Matrices elementales y determinantes

PROPOSICIÓN 3.4. *Toda matriz elemental tiene determinante no nulo*

Demostración. 1- La matriz E_1 (ver 2.15) que está asociada a la operación elemental e_1 "multiplicar la fila i por la constante $\alpha \neq 0$" tiene determiante

$$|E_1| = \alpha |I| = \alpha \neq 0$$

pues E_1 se obtiene de la matriz identidad I aplicando dicha propiedad

2- La matriz E_2 (ver 2.16) que está asociada a la operación elemental e_2 "intercambiar la fila i por la fila j" tiene determinante

$$|E_2| = -|I| = -1$$

pues E_2 se obtiene de la matriz identidad I aplicando dicha propiedad

3- La matriz E_3 (ver 2.17) que está asociada a la operación elemental e_3 "sumarle a fila i un múltiplo de la fila j" tiene determinante

$$|E_3| = |I| = 1$$

[1] Es engorroso y no lo haremos aquí pero el lector puede consultar por ejemplo (*Teorema 3 y 5 pag 87 del texto Algebra y geometría de E. Hernandez*)

pues E_3 se obtiene de la matriz identidad I aplicando dicha propiedad. $\square$

PROPOSICIÓN 3.5. *Si A es una matriz $n \times n$ y E una matriz elemental $n \times n$ se cumple que*

$$|EA| = |E|\,|A|$$

DEMOSTRACIÓN. 1- Sea E_1 la matriz que está asociada a la operación elemental "multiplicar la fila i por la constante $\alpha \neq 0$" (ver 2.15)

Por la Proposición 3.4

$$|E_1| = \alpha\,|I| = \alpha \tag{3.21}$$

Si B es la matriz que se obtiene al aplicar la operación e_1, por las propiedades de determinantes

$$|B| = \alpha\,|A| \tag{3.22}$$

y por la Proposición 2.23 se cumple que $E_1 A = B$, con lo cual

$$|E_1 A| = |B| \tag{3.23}$$

al sustituir (3.21) y (3.22) en (3.23)

$$|E_1 A| = |E_1|\,|A|$$

2- Sea E_2 la matriz que está asociada a la operación elemental e_2 "intercambiar la fila i por la fila j" (ver 2.16)

Por la Proposición 3.4

$$|E_2| = -\,|I| = -1 \tag{3.24}$$

Si B es la matriz que se obtiene al aplicar la operación e_2, por las propiedades de determinantes

$$|B| = -\,|A| \tag{3.25}$$

y por la Proposición 2.23 se cumple que $E_2 A = B$, con lo cual

$$|E_2 A| = |B| \tag{3.26}$$

al sustituir (3.24) y (3.25) en (3.26)

$$|E_1 A| = |E_1|\,|A|$$

3- Sea E_3 la matriz que está asociada a la operación elemental e_3 "sumarle a fila i un múltiplo de la fila j" (ver 2.17)

Por la Proposición 3.4

$$|E_3| = |I| = 1 \tag{3.27}$$

Si B es la matriz que se obtiene al aplicar la operación e_3, por las propiedades de determinantes

$$|B| = |A| \tag{3.28}$$

y por la Proposición 2.23 se cumple que $E_3 A = B$, con lo cual

$$|E_3 A| = |B| \tag{3.29}$$

al sustituir (3.27) y (3.28) en (3.29)

$$|E_1 A| = |E_1|\,|A|$$

pues E_3 se obtiene de la matriz identidad I aplicando dicha propiedad. $\square$

3.4. Matrices invertibles y determinantes

Ya vimos que una matriz A $n \times n$ es invertible si y solo si $rango(A) = n$

Veamos otra condición equivalente de invertibilidad

LEMA 3.6. *Si A no es invertible la forma escalerizada de A tiene por lo menos una fila de ceros*

Demostración. Como A no es invertible, se tiene que $rango(A) < n$
Al aplicar el proceso de escalerización de Gauss, (es decir que aplicamos una sucesión de operaciones elementales) obtenemos la matriz B (forma escalerizada de A)

$$A \quad \xrightarrow{\text{operación } e_1} \quad \cdots\cdots \quad \xrightarrow{\text{operación } e_k} \quad B$$

Es decir que

$$E_k, \ldots E_2 E_1 A = B$$

donde E_i es la matriz asociada a la operación elemental e_i y por lo tanto

$$A = E_1^{-1} E_2^{-1} \dots E_k^{-1} B$$

Como

$$rango(A) = \text{“cantidad de escalones no nulos de } B\text{”}$$

Concluimos que “cantidad de escalones no nulos de B”$< n$, es decir que B tiene por lo menos una fila de ceros. □

TEOREMA 3.7. *Sea A una matriz cuadrada de orden $n \times n$.*

$$A \text{ es invertible} \quad \Leftrightarrow \quad |A| \neq 0$$

DEMOSTRACIÓN. ($\Rightarrow$) Si A es invertible, por el corolario 2.27, A puede puede factorizarse como un producto de matrices elementales

$$A = E_1 E_2 \dots E_k$$

Luego, por la proposición 3.5

$$|A| = |E_1 E_2 \dots E_k| = |E_1|\,|E_2| \dots |E_k|$$

por la proposición 3.4 $|E_i| \neq 0$, con lo cual $|A| \neq 0$

($\Leftarrow$) Supongamos, por absurdo, que A no es invertible, Si B es la forma escalerizada de A, existe una sucesión de matrices elementales tales que

$$E_k, \dots E_2 E_1 A = B$$

Pero como $|E_i| \neq 0$ y $|B| = 0$ (pues por el Lema 3.6 B tiene por lo menos una fila de ceros), concluimos que $|A| = 0$, lo cual contradice nuestra hipótesis □

3.5. Determinante de un producto de matrices

TEOREMA 3.8 (Binet-Cauchy). *Si A y B son matrices $n \times n$ se cumple que*

$$|AB| = |A|\,|B|$$

Demostración. *Primer caso* A no es invertible

Si A no es invertible $\Rightarrow |A| = 0 \Rightarrow |A|\,|B| = 0$

Por otro lado, si A' es la forma escalerizada de A, existe una sucesión de matrices elementales tales

que

$$E_k, \dots E_2 E_1 A = A'$$

es decir que

$$A = E_1^{-1} E_2^{-1} \dots E_k^{-1} A'$$

Luego

$$|AB| = \left|E_1^{-1} E_2^{-1} \dots E_k^{-1} A'B\right|$$

y aplicando la proposición 3.5

$$|AB| = \left|E_1^{-1}\right| \left|E_2^{-1}\right| \dots \left|E_k^{-1}\right| \left|A'B\right|$$

Pero por el Lema 3.6, A' tiene una fila de ceros, y por lo tanto $A'B$ también (por la definición del producto de matrices), con lo cual $|A'B| = 0 \Rightarrow |AB| = 0$

Hemos probado que

$$|AB| = |A|\,|B| = 0$$

Segundo caso A es invertible

Si A es invertible, por el corolario 2.27

$$A = E_1 E_2 \dots E_k$$

con E_i matriz elemental

$$|AB| = |(E_1 E_2 \dots E_k)\, B|$$

y aplicando la proposición 3.5

$$|AB| = |E_1|\,|E_2| \dots |E_k|\,|B| = |E_1 E_2 \dots E_k|\,|B| = |A|\,|B|\,.$$

□

3.6. Cálculo de la matriz inversa por determinantes

El desarrollo del determinante de la matriz $A = ((a_{ij}))$ por su i-esima fila es

$$|A| = (-1)^{i+1} a_{i1} |A_{i1}| + (-1)^{i+2} a_{i2} |A_{i2}| + \ldots + \\ + (-1)^{i+j} a_{ij} |A_{ij}| + \ldots + (-1)^{i+n} a_{in} |A_{in}| \tag{3.30}$$

El número $c_{ij} = (-1)^{i+j} |A_{ij}|$ se denomina **cofactor** del elemento a_{ij} de la matriz A. Por lo tanto podemos reescribir el desarrollo (3.30) de la siguiente manera

$$|A| = a_{i1}c_{i1} + a_{i2}c_{i2} + \ldots + a_{ij}c_{ij} + \ldots + a_{in}c_{in} \tag{3.31}$$

$$\text{A la matriz} \quad cof\,(A) \overset{def}{=} \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix}$$

se le llama **matriz de cofactores de** A

LEMA 3.9. *Si* $[cof\,(A)]^t$ *es la matriz transpuesta de la matriz de cofactores de* A*, esto es*

$$[cof\,(A)]^t = \begin{pmatrix} c_{11} & c_{21} & \ldots & c_{n1} \\ c_{12} & c_{22} & \ldots & c_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \ldots & c_{nn} \end{pmatrix}$$

$$\text{se cumple que} \quad A\,[cof\,(A)]^t = \begin{pmatrix} |A| & 0 & \ldots & 0 \\ 0 & |A| & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & |A| \end{pmatrix}.$$

Demostración. Si $A\,[cof\,(A)]^t = ((d_{ii}))$ queremos probar que:

1. Los elementos de la diagonal de la matriz $A\,[cof\,(A)]^t$ son todos iguales a $|A|$, esto es $d_{ii} = |A|$.

2. Los elementos que no son de la diagonal de la matriz $A\,[cof\,(A)]^t$ son todos nulos, esto es $\quad d_{ij} = 0 \quad (i \neq j)$

Probemos 1. Para ello calculamos el elemento d_{ii} de la matriz $A\,[cof\,(A)]^t$

$$
\begin{array}{cc}
 & \begin{pmatrix} \cdots & \cdots & c_{i1} & \cdots & \cdots \\ \cdots & \cdots & c_{i2} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & c_{in} & \cdots & \cdots \end{pmatrix} \\
\begin{pmatrix} \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \vdots \end{pmatrix} & \begin{pmatrix} \cdots & \cdots & \downarrow & \cdots & \cdots \\ \cdots & \cdots & \downarrow & \cdots & \cdots \\ \rightarrow & \rightarrow & d_{ii} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}
\end{array}
$$

Por lo tanto $\quad d_{ii} = a_{i1}c_{i1} + a_{i2}c_{i2} + \ldots + a_{ij}c_{ij} + \ldots + a_{in}c_{in}$, pero por (3.31) $|A| = a_{i1}c_{i1} + a_{i2}c_{i2} + \ldots + a_{ij}c_{ij} + \ldots + a_{in}c_{in}$, con lo cual $\quad d_{ii} = |A|$.

Probemos 2. Calculamos el elemento d_{ij} de la matriz $A\,[cof\,(A)]^t$

$$
\begin{array}{cc}
 & \begin{pmatrix} \cdots & \cdots & \cdots & c_{j1} & \cdots \\ \cdots & \cdots & \cdots & c_{j2} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & c_{jn} & \cdots \end{pmatrix} \\
\begin{pmatrix} \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \vdots \end{pmatrix} & \begin{pmatrix} \cdots & \cdots & \cdots & \downarrow & \cdots \\ \cdots & \cdots & \cdots & \downarrow & \cdots \\ \rightarrow & \rightarrow & \rightarrow & d_{ij} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}
\end{array}
$$

Por lo tanto $\quad d_{ij} = a_{i1}c_{j1} + a_{i2}c_{j2} + \ldots + a_{ik}c_{jk} + \ldots + a_{in}c_{jn}$.

Por otro lado consideremos la matriz B que se obtiene de A cambiando la fila j por la fila i:

$$B = \begin{array}{cl} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \cdots \\ \cdots \\ \longleftarrow \text{fila } i \\ \cdots \\ \longleftarrow \text{fila } j \\ \ddots \\ \cdots \end{array} \end{array}$$

Si calculamos el determinante de B desarrollando por la j-esima fila se tiene que

$$\begin{aligned} |B| = (-1)^{j+1} b_{j1} |B_{j1}| + (-1)^{j+2} b_{j2} |B_{j2}| + \ldots + \\ + (-1)^{j+k} b_{jk} |B_{jk}| + \ldots + (-1)^{j+n} b_{jn} |B_{jn}| \end{aligned}$$

pero como $b_{jk} = a_{ik}$ y $B_{jk} = A_{jk}$, resulta que

$$\begin{aligned} |B| = (-1)^{j+1} a_{i1} |A_{j1}| + (-1)^{j+2} a_{i2} |A_{j2}| + \ldots + \\ + (-1)^{j+k} a_{ik} |A_{jk}| + \ldots + (-1)^{j+n} a_{in} |A_{jn}| \end{aligned}$$

pero $c_{jk} = (-1)^{j+k} |A_{jk}|$, por lo cual

$$|B| = a_{i1}c_{j1} + a_{i2}c_{j2} + \ldots + a_{ik}c_{jk} + \ldots + a_{in}c_{jn}.$$

Esto prueba que $|B| = d_{ij}$. Como $|B| = 0$ (pues tiene dos filas iguales), se prueba lo deseado. □

PROPOSICIÓN 3.10. *Si A es invertible, se tiene que*

$$A^{-1} = \frac{1}{|A|} [cof\,(A)]^t$$

DEMOSTRACIÓN. Por el lema anterior

$$A\left[cof\,(A)\right]^t = \begin{pmatrix} |A| & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |A| \end{pmatrix} = |A| \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = |A|\, I$$

Siendo A invertible se tiene que $|A| \neq 0$, con lo cual $\frac{1}{|A|}A\left[cof\,(A)\right]^t = I$; esto es

$$A\left(\frac{1}{|A|}\left[cof\,(A)\right]^t\right) I,$$

lo cual prueba que $A^{-1} = \frac{1}{|A|}\left[cof\,(A)\right]^t$. □

La matriz $A = \begin{pmatrix} 1 & -1 & 2 \\ 4 & 2 & 4 \\ 5 & 0 & 1 \end{pmatrix}$ es invertible pues $|A| = -34$; calculemos los cofactores de A

$$\begin{pmatrix} \times & \times & \times \\ \times & 2 & 4 \\ \times & 0 & 1 \end{pmatrix} \Rightarrow A_{11} = \begin{pmatrix} 2 & 4 \\ 0 & 1 \end{pmatrix} \Rightarrow |A_{11}| = \begin{vmatrix} 2 & 4 \\ 0 & 1 \end{vmatrix} = 2 \Rightarrow c_{11} = 2$$

y trabajando de la misma manera se obtienen

$$|A_{21}| = \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -1 \Rightarrow c_{21} = 1; \quad |A_{31}| = \begin{vmatrix} -1 & 2 \\ 2 & 4 \end{vmatrix} = -8 \Rightarrow c_{31} = -8;$$

$$|A_{12}| = \begin{vmatrix} 4 & 4 \\ 5 & 1 \end{vmatrix} = -16 \Rightarrow c_{12} = 16; \quad |A_{22}| = \begin{vmatrix} 1 & 2 \\ 5 & 1 \end{vmatrix} = -9 \Rightarrow c_{22} = -9;$$

$$|A_{32}| = \begin{vmatrix} 1 & 2 \\ 4 & 4 \end{vmatrix} = -4 \Rightarrow c_{32} = 4; \quad |A_{13}| = \begin{vmatrix} 4 & 2 \\ 5 & 0 \end{vmatrix} = -10 \Rightarrow c_{13} = -10;$$

$$|A_{23}| = \begin{vmatrix} 1 & -1 \\ 5 & 0 \end{vmatrix} = \Rightarrow c_{23} = -5; \quad |A_{33}| = \begin{vmatrix} 1 & -1 \\ 4 & 2 \end{vmatrix} = 6 \Rightarrow c_{33} = 6$$

Por lo tanto $cof\,(A) = \begin{pmatrix} 2 & 16 & -10 \\ 1 & -9 & -5 \\ -8 & 4 & 6 \end{pmatrix}$ con lo cual

$$A^{-1} = \frac{1}{|A|}\left[cof\,(A)\right]^t = \frac{1}{-34}\begin{pmatrix} 2 & 1 & -8 \\ 16 & -9 & 4 \\ -10 & -5 & 6 \end{pmatrix} = \begin{pmatrix} -\frac{1}{17} & -\frac{1}{34} & \frac{4}{17} \\ -\frac{8}{17} & \frac{9}{34} & -\frac{2}{17} \\ \frac{5}{17} & \frac{5}{34} & -\frac{3}{17} \end{pmatrix}.$$

3.7. Regla de Cramer

TEOREMA 3.11 (Cramer 1). *Si $|A| \neq 0$ entonces el sistema $Ax = b$ es compatible y determinado y su solución (única) es*

$$x_j = \frac{|A_j|}{|A|} \qquad j = 1, 2 \ldots, n$$

donde A_j es la matriz que se obtiene de A sustituyendo la columna j por b

Demostración. Como $|A| \neq 0$, existe A^{-1}, por lo tanto $Ax = b \Leftrightarrow x = A^{-1}b$ lo que prueba la existencia y la unicidad de la solución

$$x = A^{-1}b \Leftrightarrow x = \left(\frac{1}{|A|}\left[cof\,(A)\right]^t\right) b \Leftrightarrow$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} c_{11} & c_{21} & \ldots & c_{n1} \\ \vdots & \vdots & \vdots & \vdots \\ c_{1j} & c_{2j} & \ldots & c_{nj} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \ldots & c_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} c_{11}b_1 + c_{21}b_2 + \ldots + c_n b_n \\ \vdots \\ c_j b_1 + c_{2j}b_2 + \ldots + c_{nj}b_n \\ \vdots \\ c_n b_1 + c_{2n}b_2 + \ldots + c_{n^2}b_n \end{pmatrix}$$

Por lo tanto

$$x_j = \frac{1}{|A|}\left(c_{1j}b_1 + c_{2j}b_2 + \ldots + c_{nj}b_n\right) \qquad j = 1, 2 \ldots, n \tag{3.32}$$

Por otro lado si consideramos la matriz

$$A_j = \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{i1} & \cdots & b_2 & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix}$$

que se obtiene de A sustituyendo la columna j por b, y desarrollamos el determinante de A_j por la columna j se tiene que

$$\begin{aligned} |A_j| &= (-1)^{j+1} b_1 |A_{1j}| + (-1)^{j+2} b_2 |A_{2j}| + \ldots + (-1)^{j+n} b_j |A_{nj}| \\ &= c_{1j}b_1 + c_{2j}b_2 + \ldots + c_{nj}b_n \end{aligned}$$

Al sustituir en (3.32) se obtiene $\quad x_j = \frac{|A_j|}{|A|} \qquad j = 1, 2 \ldots, n.$ □

EJEMPLO 3.4. Vamos a resolver el sistema $\begin{cases} x + 2y - z = 7 \\ 2x + y + z = 6 \\ x - y + 3z = -1 \end{cases}$ usando la regla de Cramer. En este caso $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ 1 & -1 & 3 \end{pmatrix}$ y $b = \begin{pmatrix} 7 \\ 6 \\ -1 \end{pmatrix}$

Como $|A| = -3 \neq 0$ el sistema es compatible y determinado y sus soluciones son

$$x = \frac{|A_1|}{|A|} = \frac{\begin{vmatrix} 7 & 2 & -1 \\ 6 & 1 & 1 \\ -1 & -1 & 3 \end{vmatrix}}{-3} = \frac{5}{3},$$

$$y = \frac{|A_2|}{|A|} = \frac{\begin{vmatrix} 1 & 7 & -1 \\ 2 & 6 & 1 \\ 1 & -1 & 3 \end{vmatrix}}{-3} = \frac{8}{3}, \quad z = \frac{|A_3|}{|A|} = \frac{\begin{vmatrix} 1 & 2 & 7 \\ 2 & 1 & 6 \\ 1 & -1 & -1 \end{vmatrix}}{-3} = 0$$

■

TEOREMA 3.12 (Cramer 2). *Si $|A| = 0$ y existe j_0 tal que $|A_{j_0}| \neq 0$ entonces el sistema $Ax = b$ es incompatible*

EJEMPLO 3.5. Como $|A| = 0$ se tiene que $rango\,(A) < n$. Por otro lado si $\widetilde{A} = (\,A|\,b)$ es la matriz ampliada del sistema, como las columnas de A_{j_0} son también columnas de $\widetilde{A}$ se tiene que

$$rango\left(\widetilde{A}\right) \geq rango\,(A_{j_0})$$

Además $rango\,(A_{j_0}) = n$, por ser $|A_{j_0}| \neq 0$, con lo cual $n \leq rango\left(\widetilde{A}\right)$.

Por lo tanto $rango\,(A) < rango\left(\widetilde{A}\right)$ y por el teorema de Rouché - Frobenius el sistema $Ax = b$ es incompatible.

OBSERVACIÓN 3.13. Un **error frecuente** es decir que: *Si* $|A| = 0$ *y* $|A_j| = 0$ *entonces el sistema* $Ax = b$ *es indeterminado*. Veamos ejemplos.

EJEMPLO 3.6. Para el sistema de ecuaciones $\begin{cases} x+y+z=1 \\ x+y+z=1 \\ x+y+z=0 \end{cases}$ se tiene que

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Por lo tanto $|A| = |A_1| = |A_2| = |A_3| = 0$, sin embargo es evidente que el sistema es incompatible. Mientras que para el sistema de ecuaciones $\begin{cases} x+y+z=1 \\ x+y+z=1 \\ x+y+z=1 \end{cases}$ se tiene que $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ y también se cumple que $|A| = |A_1| = |A_2| = |A_3| = 0$, pero el sistema es compatible e indeterminado. ■

CAPíTULO 4

RECTAS Y PLANOS EN EL ESPACIO

4.1. Introducción

El propósito del presente capítulo -y del capítulo denominado *Producto Escalar y Vectorial*- es dar una presentación sucinta de la Geometría Analítica del Espacio, que permita trabajar matemáticamente con los objetos del **espacio ambiente**. En un sentido muy vago y deliberadamente informal entendemos por tal al modelo espacial en que transcurren la mayoría de los fenómenos analizados por la Mecánica y otras teorías científicas Clásicas. Estas teorías se refieren a los acontecimientos más o menos directamente percibidos por los seres humanos.

Esta presentación incluirá la formalización del concepto de vector que el lector conoce de sus cursos de Nivel Secundario. Por un lado, en los cursos de Geometría ha trabajado con las traslaciones que, como recordará, están determinadas por vectores. En general el vector se define en esos cursos como una clase de equivalencia de segmentos orientados con la misma dirección, sentido y congruentes (dos figuras son congruentes si se corresponden por una isometría).

Por otro lado, en Física ha trabajado con ciertas "magnitudes" (fuerza, velocidad, aceleración, etc.) llamadas "vectoriales". Estas se distinguen claramente de otras llamadas "escalares" (masa, temperatura, etc.) porque las últimas quedan determinadas por un único dato numérico, en cambio las primeras requieren más información para caracterizarlas: "módulo", "dirección" y "sentido". Las magnitudes vectoriales se representan además gráficamente mediante una "flecha" y a pesar de las diferencias que ya marcamos con los "escalares" tienen algo en común: con ambos se realizan operaciones.

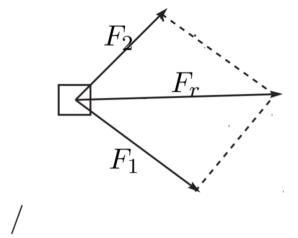


Figura 1

Por ejemplo cuando se aplican dos fuerzas $\vec{F_1}$ y $\vec{F_2}$ sobre un mismo cuerpo el resultado es el mismo que si se aplicara una única fuerza $\vec{F_r}$ con módulo, dirección y sentido igual a la diagonal del paralelogramo de lados $\vec{F_1}$ y $\vec{F_2}$ (ver figura 4.1). Se dice entonces que

$$\vec{F_r} = \vec{F_1} + \vec{F_2}$$

o que $\vec{F_r}$ es la resultante de $\vec{F_1}$ y $\vec{F_2}$, de este modo la "suma de vectores" cobra sentido. Si $\vec{F_1} = \vec{F_2}$ entonces $\vec{F_r} = 2\vec{F_1}$ con lo cual también cobra sentido la multiplicación de escalares por vectores.

Nuestra intención es introducir el concepto de "vector" de una manera más rigurosa y utilizarlo para desarrollar algunas aplicaciones a la Geometría Analítica del Espacio. Al mismo tiempo queremos destacar que este concepto de vector servirá de fuente de inspiración para uno de los objetos fundamentales de este curso. Más adelante daremos una definición más general y abstracta de vector perteneciente a un **espacio vectorial** con múltiples aplicaciones dentro y fuera de la matemática. Este tipo de procesos de generalización desde ejemplos hacia conceptos abstractos permite dar un tratamiento unificado a múltiples objetos a menudo de apariencia disímil pero con propiedades esenciales comunes. Para poder hacerlos se debe estudiar con cuidado los ejemplos sencillos tratando de destacar estas propiedades.

4.2. Vectores

El espacio ambiente, que denotaremos con la letra E, está formado por puntos designados en general con letras mayúsculas $A, B, P, Q, ...$ Dos puntos distintos definen una única recta y tres puntos no contenidos en una recta definen un plano. En este capítulo se dará una descripción analítica de recta, plano y sus relaciones de intersección.

La longitud del segmento entre dos puntos de E define la **distancia entre dos puntos** $d : E \times E \longrightarrow \mathbb{R}^+$, que satisface
$d(A, B) = d(B, A) \quad \forall A \in E$
$d(A, B) = 0 \quad$ si y sólo si $A = B$
$d(A, C) \leq d(A, B) + d(B, C) \quad \forall A, B, C \in E.$

A este concepto elemental de distancia agregamos ahora el de **vector**. Pretendemos formalizar lo que antes hemos denominado una "flecha" como un objeto que tenga "modulo", "dirección" y "sentido". ¿Qué datos son necesarios para dibujarla? En principio basta indicar un par de puntos ordenados (A, B) el punto A será el origen y el punto B el extremo del segmento orientado ("flecha"). Consideremos dos segmentos orientados como en la figura

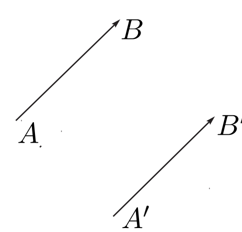
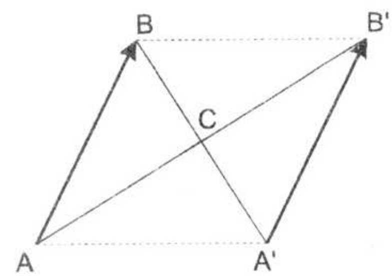


FIGURA 2

4.2 de modo que $AA'B'B$ sea un paralelogramo, el par (A, B) es distinto de (A', B'). Sin embargo por ser $AA'B'B$ un paralelogramo ambas segmentos orientados tienen igual dirección (las rectas AB y $A'B'$ son paralelas), módulo $(d(A, B) = d(A', B'))$ y sentido. Solo se diferencian en el punto de base o de "aplicación". No estamos interesados en que nuestro concepto de vector distinga objetos por su punto de aplicación así que deberemos ser más cuidadosos en nuestra definición y adoptar una que **no** distinga entre dos flechas como (A, B) y (A', B'). Una manera de lograr esto es identificándolas mediante una relación de equivalencia.

Para ello definimos una relación de equivalencia entre pares ordenados de puntos de E: $AB \sim /A'B'$ si y sólo si $ABB'A'$ es un paralelogramo. Ésta es una relación de equivalencia de acuerdo con la definición dada en el Capítulo 0.



Obsérvese que:

1) Es fácil probar que $AB \sim A'B'$ si y sólo si el punto medio C de AB' / coincide con el de $A'B$; o sea que una misma simetría central de centro C lleva A en B' y A' en B.

2) También se puede dar una definición de la relación de equivalencia en término de traslaciones.

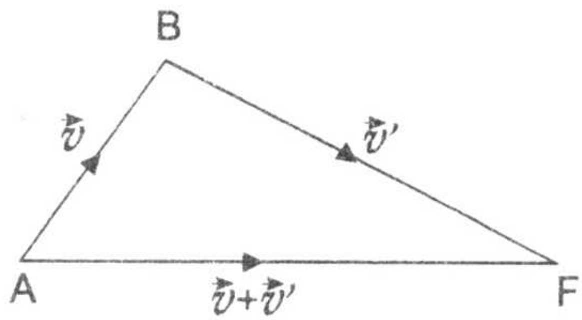
3) Todas estas definiciones sólo utilizan elementos de la geometría plana.

Llamaremos **vector del espacio** E a una clase de equivalencia determinada por esta relación. Cualquier pareja de puntos de una clase de equivalencia representa al correspondiente vector. Si la pareja es AB, el vector por ella representado se escribirá de cualquiera de las siguientes maneras $v = \overrightarrow{AB} = B - A$. Dada una representación AB de un vector v llamaremos **origen (o punto base)** al punto A, y **extremo** al punto B.

Denominaremos con la letra V al conjunto de estos vectores. Obsérvese que el espacio V incluye al vector nulo $\overrightarrow{o} = \overrightarrow{AA}$, para cualquier $A \in E$. Dado un punto $P \in E$ y un vector $v \in V$, es claro que existe un único punto $Q \in E$ tal que $\overrightarrow{PQ} = v$, o sea que hay un único representante de v que tiene origen en P: ese representante es PQ. Dado P, y si v está dado por una pareja AB (o sea $v = \overrightarrow{AB}$) para hallar Q alcanza con construir el paralelogramo $PABQ$. De esta manera hemos definido una función que a pareja de punto de E y vector de V, le hace corresponder otro punto de E , de la siguiente manera: $(P, v) \longrightarrow Q \Longleftrightarrow \overrightarrow{PQ} = v$. Escribiremos / $Q = P + v$ y diremos que Q es la **suma de P más** v.

4.3. Operaciones con vectores.

La **suma de dos vectores** es otro vector, o sea que se trata de definir una función que a parejas de vectores le haga corresponder otro vector: $(v, w) \longrightarrow v + w$. Dados dos vectores $v, w \in V$, se tomarán representantes de cada uno de ellos de modo que el extremo del representante de v sea el origen del de w. O sea, que si $v = \overrightarrow{AB}$, tomamos $w = \overrightarrow{BC}$. Decimos que $\overrightarrow{AC}$ representa al vector suma $v + w$; o sea que $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.



Es fácil ver que el vector suma no depende del representante elegido para v ; o sea que si $\overrightarrow{A'B'} = v$ (y $\overrightarrow{B'C'} = w$) entonces $\overrightarrow{A'B'} + \overrightarrow{B'C'} = \overrightarrow{AC}$. Para verificar esta independencia de los representantes, se recomienda -como para casi todas las pruebas de este capítulo- hacer los dibujos correspondientes.

Conviene observar desde ya que para cualquier $v = \overrightarrow{AB}$, resulta $\overrightarrow{o} + v = v + \overrightarrow{o} = v$ (alcanza con tomar el representante AA o BB, según el caso, de $\overrightarrow{o}$). Además el vector $\overrightarrow{BA}$ sumado a v da $\overrightarrow{o}$: $\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \overrightarrow{o}$.

Llamaremos **opuesto de v** al vector representado por BA, y lo escribiremos $-v$. Además (háganse los dibujos), para todo $A, B, C, D \in E$ vale $A - B = C - D \Longrightarrow A - C = B - D$.

Para definir el **producto de un número real por un vector** se debe andar con más cuidado, distinguiendo diferentes casos y utilizando el concepto de distancia. Se trata de definir una función $(a, v) \longrightarrow av$, para todo $a \in \mathbb{R},\ v \in V$. Obsérvese que no hemos indicado el producto del número real a por el vector v con ningún símbolo; sencillamente hemos colocado av juntos. Esto podría dar lugar a confusiones, pero siendo a y v elementos de conjuntos en general distintos, hemos preferido la simplificación de no usar símbolo explícito para este producto. Sea $v = \overrightarrow{AB}$,
- si $a \geq 0$, en la semirrecta AB se elige el único C tal que $d(A, C) = ad(A, B)$. O sea que C está del mismo lado que B respecto a A, en la recta determinada por A y B
- Si $a < 0$, se elige C en el lado opuesto al que está B respecto de A, de modo que $d(A, C) = |a|d(A, B)$.

En ambos casos se define $\overrightarrow{AC} = a\overrightarrow{AB} = av$. Se observa fácilmente que esta definición no depende del representante de v que se haya elegido. Obsérvese también que la multiplicación del número cero por cualquier vector da el vector nulo: $0v = \overrightarrow{o}$, para todo $v \in V$. Igualmente, el vector opuesto $-v$ se obtiene multiplicando v por el número -1; o sea que $(-1)v + v = \overrightarrow{o}$. Diremos que dos vectores u, v son **colineales** (o **paralelos** o que tienen la **misma dirección**) si $u = av$ para algún número real a. También diremos que u es **múltiplo** de v. Si $a > 0$ diremos que además tienen el **mismo sentido**. Si $a = \pm 1$ tienen el mismo **módulo**. Esta definición lleva a que el vector nulo sea colineal con todo vector de V, porque el número real puede ser el cero.

Estas dos operaciones de suma de vectores y producto por un real, verifican las siguientes ocho propiedades (u, v, w son cualesquiera vectores de V y a, b son cualesquiera números reales):

[S1] [Conmutativa] $u + v = v + u$,
[S2] [Asociativa] $u + (v + w) = (u + v) + w$, $\forall\, X, Y, Z \in \mathbb{R}^n$.

[S3] [Neutro de la suma] Existe un vector $\overrightarrow{o}$ tal que $v + \overrightarrow{o} = v$,
[S4] [Existencia de opuesto] Existe $-v$ tal que $v + (-v) = \overrightarrow{o}$;
[P1] [Asociativa del producto] $a(bv) = (ab)v$,
[P2] [Neutro del producto] $1v = v$
[P3] [Distributiva] $(a + b)v = av + bv$,
[P4] [Distributiva] $a(u + v) = au + av$.

Obsérvese los distintos significados que tienen los productos (sus símbolos están omitidos) en P1; y las sumas, en P3. Para probar estas propiedades para los vectores de V recomendamos una vez más hacer dibujos claros, aplicar bien las definiciones de cada una de las operaciones y luego aplicar propiedades elementales de geometría plana y de los números reales. Como ejemplo daremos la prueba de P1. Si $v = \overrightarrow{AB}$, sean $(ab)v = \overrightarrow{AC}$ y $a(bv) = \overrightarrow{AD}$ con C, D en la recta AB. Si a ó b es cero, el resultado es $A = C = D$ o sea $\overrightarrow{AC} = \overrightarrow{AD} = \overrightarrow{o}$. Si a, b tienen el mismo signo, C y D están en la semirrecta de origen A por B y $d(A, C) = d(A, D) = ab$, por lo que $C = D$. Si $a > 0,\ b < 0,\ \ C$ está en el lado opuesto a B y D también porque bv tiene sentido opuesto a v y $a(bv)$ tiene el sentido de bv; además $d(A, C) = d(A, D) = |ab|$; por tanto C y D coinciden. Más adelante se considerarán entes matemáticos entre cuyos elementos se pueden definir operaciones que verifican este conjunto de propiedades. Se verá que con sólo ellas se puede desarrollar la teoría más simple, abstracta y abarcativa de **espacios vectoriales**. También es importante tener presente las siguientes propiedades que relacionan la suma de punto y vector con la suma de vectores ($\forall\ \ A, B, C, D \in E,\ v, w \in V$):
1. $A + (v + w) = (A + v) + w$.
2. $A + (B - A) = B$; $\quad (A + w) - A = w$.
3. $A + \overrightarrow{o} = A$; $A - A = \overrightarrow{o}$.
4. $(B - A) + u = (B + u) - A = B - [A + (-u)]$. Como ejemplo, veremos la prueba de 4. Si $w = (B - A) + u$, resulta $A + w = [A + (B - A)] + u = B + u$, por 2. Luego $w = (B + u) - A$ por la segunda parte de 2. Esto prueba la primera igualdad. Además $[A + (-u)] + w = A + [-u + (B + u) - A] = A + [-u + (B - A) + u] = A + (B - A) = B$; se han usado sucesivamente

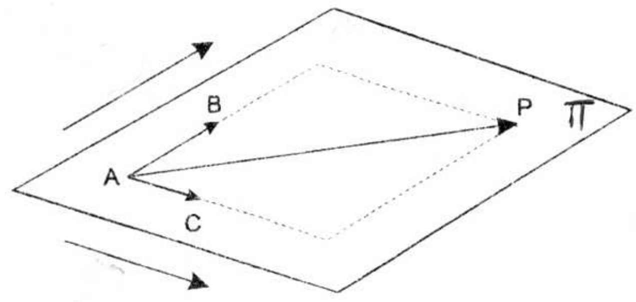
1., la definición de w, la primera igualdad de 4. y la primera de 2. Entonces $w = B - (A + (-u))$ y queda demostrada la segunda igualdad de 4.

4.4. Ecuaciones vectoriales de rectas y planos. Paralelismo

Recta. Sea r la recta de E definida por los puntos A y B. De acuerdo con lo visto en la definición de producto de un vector por un número real, todos los puntos de esa recta quedan descriptos por $A + \lambda\overrightarrow{AB}$, al variar λ entre los números reales. O sea que todo punto $P \in r$ es de la forma $P = A + \lambda\overrightarrow{AB}$ para algún $\lambda \in \mathbb{R}$. También se puede dar una recta s por un punto $A \in E$ y un vector no nulo $v \in V$, llamado **vector director** de la recta s. En este caso resulta $s = \{P \in E : P = A + \lambda v, \lambda \in \mathbb{R}\}$. Dos rectas con vectores directores $v,\ w$ son **paralelas** si y sólo si v es colineal con w. Si además, ambas rectas paralelas tienen un punto en común, ellas son la misma recta. En efecto si las rectas son de la forma $A + \lambda v,\quad B + \mu w$ con $w = a_1 v$ y $B = A + \lambda_1 v$, resulta que todo punto de la segunda recta es de la forma $B + \mu w = A + (\lambda_1 + a_1\mu)v$, o sea que es de la primera recta.

Plano. Dados tres puntos no pertenecientes a la misma recta (no colineales) $A, B, C \in E$, el plano π por ellos definido esta formado por el conjunto de puntos de la forma $A + \lambda\overrightarrow{AB} + \mu\overrightarrow{AC}$ con $\lambda, \mu \in \mathbb{R}$. De igual manera que en caso de las rectas, un plano también queda bien definido por un punto A y dos vectores u, v no colineales. En este caso $\pi = \{P \in E : P = A + \lambda u + \mu v, \lambda, \mu \in \mathbb{R}\}$. Cualquier pareja de vectores u, v no colineales que definan un plano de esa manera se denominarán **vectores directores** del plano.

Dos planos $\pi,\ \pi'$ son **paralelos** si todo vector con representantes definidos por dos puntos de π, tiene también representantes con dos puntos de π'. Diremos, un tanto imprecisamente, que los vectores **son** de π y π'. Hallaremos una condición necesaria y suficiente para que dos planos π, π' de las formas $P = A + \lambda u + \mu v,\quad P' = A' + au' + bv',\quad u, v$ no colineales, u', v' no colineales, sean paralelos. Obsérvese primero que todos los vectores determinados por dos puntos del plano π son de la forma $\overrightarrow{AP} = \lambda u + \mu v$. Para que los dos planos sean paralelos no es necesario que u, u' y v, v' sean colineales,



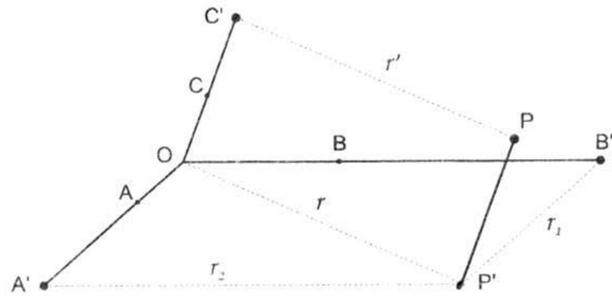
sino que los vectores u', v' se puedan escribir como suma de múltiplos de u, v. En efecto si $u' = \lambda_1 u + \mu_1 v$, $v' = \lambda_2 u + \mu_2 v$ resulta que todo vector de π' es de la forma $au' + bv' = (a\lambda_1 + b\lambda_2)u + (a\mu_1 + b\mu_2)v$; o sea que también lo es de π. Si u' y v' se escriben de aquella manera como combinaciones lineales de u, v, se pueden despejar u y v de la siguiente manera

$$v = \frac{\lambda_2 u' - \lambda_1 v'}{\lambda_2 \mu_1 - \lambda_1 \mu_2}, \quad u = \frac{-\mu_2 u' + \mu_1 v'}{\lambda_2 \mu_1 - \lambda_1 \mu_2}$$

(que tiene sentido porque si fuera $\lambda_2\mu_1 - \lambda_1\mu_2 = 0$ resultaría $\lambda_2 u' - \lambda_1 v' = 0$ y serían u', v' colineales, lo cual va contra una suposición inicial); por un razonamiento igual que el anterior resulta que todo vector de π también es vector de π'.

4.5. Sistemas de coordenadas

Comencemos con dos definiciones de gran trascendencia. v es una **combinación lineal de los vectores** $v_1, v_2, \cdots, v_n \in V$ si es de la forma $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ con $a_1, a_2, \cdots, a_n \in \mathbb{R}$. Una **base** de V es un conjunto de vectores $\{v_1, v_2, \cdots, v_n\} \subset V$ tal que todo vector $v \in V$ se escribe *de manera única* como combinación lineal de ese conjunto. O sea, si para cada $v \in V$ existen únicos $a_1, a_2, \cdots, a_n \in \mathbb{R}$ tales que $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Destacamos la *unicidad* de la combinación lineal de vectores de la base, para generar cada uno de los vectores de V. En un sentido que se aclarará en capítulos posteriores, una/ base es un conjunto mínimo que permite escribir todos los vectores de V como combinación lineal de ellos. A continuación mostraremos cómo son las bases de V.



De acuerdo con las definiciones que venimos dando, es claro que se pueden elegir cuatro puntos $A, B, C, D \in E$ **no coplanares** (no pertenecientes a un mismo plano). Diremos también que los vectores por ellos determinados son no coplanares. O sea, los vectores $u, v, w \in V$ son no coplanares si $w \neq au + bv$ para cualesquiera dos números reales a, b. Esto equivale, si se toman representantes de los vectores con el mismo origen $O: \quad u = \overrightarrow{OA},\ v = \overrightarrow{OB}, w = \overrightarrow{OC}$, a que el punto C no esté en el plano determinado por O, A, B, o sea que $C \neq O + a\overrightarrow{OA} + b\overrightarrow{OB}$ para cualesquiera números reales a, b. Sean cuatro puntos no coplanares $O, A, B, C \in E$. Dado cualquier punto $P \in E$, consideremos la recta s paralela a OC por P. Ella corta al plano OAB en P' (que coincide con O si P esta en la recta OC). Por P' trazamos las paralelas r_1, r_2 a las rectas OA, OB, respectivamente, que cortan a OB y OA en B' y A', respectivamente. Por otro lado, el plano paralelo a OAB por P corta a la recta OC en el punto C' (que coincide con O si P esta en el plano OAB). No es difícil ver, utilizando propiedades elementales de paralelismo e intersecciones, que se habría llegado a los mismos puntos A', B', C' si se hubieran trazado recta y plano paralelos a por ejemplo la recta OA y el plano OBC, y seguido el procedimiento indicado anteriormente. Por la definición de suma de vectores aplicada dos veces, se deduce que $\overrightarrow{OP} = \overrightarrow{OA'} + \overrightarrow{OB'} + \overrightarrow{OC'}$, y por la definición de producto de número real por vector, se ve el vector $\overrightarrow{OA'}$ es colineal con $\overrightarrow{OA}$, por lo que

$\overrightarrow{OA'} = a\overrightarrow{OA}$ para algún $a \in \mathbb{R}$. De igual manera $\overrightarrow{OB'} = b\overrightarrow{OB}$, $\overrightarrow{OC'} = c\overrightarrow{OC}$ para $b, c \in \mathbb{R}$. Por tanto $v = \overrightarrow{OP} = a\overrightarrow{OA} + b\overrightarrow{OB} + c\overrightarrow{OC}$. Para cada $P \in E$ (o lo que es lo mismo, cada $v \in V$) los números reales a, b, c son únicos pues si fuera $v = a'\overrightarrow{OA} + b'\overrightarrow{OB} + c'\overrightarrow{OC}$ resultaría, suponiendo que, por ejemplo $a \neq a'$: $(a - a')\overrightarrow{OA} = (b' - b)\overrightarrow{OB} + (c' - c)\overrightarrow{OC}$ por lo que

$$\overrightarrow{OA} = \frac{b' - b}{a - a'}\overrightarrow{OB} + \frac{c' - c}{a' - a}\overrightarrow{OC}.$$

Entonces el punto A estaría en el plano OBC lo cual contradice la hipótesis de que los puntos O, A, B, C no son coplanares. Esta demostración vale para cualesquiera tres vectores $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ no coplanares. Por tanto hemos demostrado que *tres vectores no coplanares cualesquiera forman una base del espacio* V. Recíprocamente, *si tres vectores forman una base de* V *ellos deben ser no coplanares.* En efecto, si fueran coplanares, tomando representantes con un mismo origen O, sus extremos determinarían un plano π, y ningún vector $\overrightarrow{OP}$ con $P \in \pi$ se podría escribir como combinación lineal de ellos.

A continuación generalizaremos el concepto de sistema de coordenadas –que el estudiante ha estudiado en el plano– al espacio E, que es el objeto principal de esta parte del curso. El estudiante recordará que los puntos de una recta se ponen en correspondencia biunívoca con los números reales, eligiendo un origen y una unidad; y que los puntos de un plano se ponen en correspondencia biunívoca con las parejas ordenadas de números reales ($\mathbb{R}^2$) eligiendo un origen y dos ejes transversales que se cortan en ese origen.

Un **sistema de coordenadas o referencial de E** está formado por un punto O, el **origen de coordenadas** y una base $\{v_1, v_2, v_3\}$ de V. Las rectas determinadas por O y v_1, v_2, v_3 se denominan **ejes coordenados**. Si los ejes coordenados son perpendiculares se dice que el sistema de coordenadas es **ortogonal**. Se suele representar por Ox, Oy, Oz a los ejes coordenados definidos por O y v_1, v_2, v_3, respectivamente. Se llamará **plano coordenado xOy** al plano determinado por Ox, Oy. De igual manera se definen los planos coordenados yOz, zOx.

Por lo que antecede resulta que dado un sistema de coordenadas $\{O, v_1, v_2, v_3\}$, para cada $P \in E$ existen números únicos x_1, x_2, x_3 tales que

$P-O = x_1v_1+x_2v_2+x_3v_3$; luego, $P = O+x_1v_1+x_2v_2+x_3v_3$. Decimos que x_1, x_2, x_3 son las **coordenadas de P en el sistema** $\{O, v_1, v_2, v_3\}$. De igual manera, si $v = a_1v_1 + a_2v_2 + a_3v_3$, llamaremos **coordenadas de v en la base** $\{v_1, v_2, v_3\}$ a la terna (a_1, a_2, a_3) por lo que esta terna, dependiendo del contexto al que se refiera, indicará a las coordenadas del punto $A = O+v \in E$, o del vector $v \in V$. La determinación de un sistema de coordenadas establece una correspondencia biunívoca entre el conjunto de los puntos del espacio E y $\mathbb{R}^3$, el conjunto de las ternas ordenadas de números reales. Como ejemplos destacamos que el punto elegido como origen de coordenadas tiene coordenadas $(0, 0, 0) \in \mathbb{R}^3$ y el punto $O + v_1$ tiene coordenadas $(1, 0, 0)$. Se suele escribir $P = (x_1, x_2, x_3)$, para indicar que x_1, x_2, x_3 son las coordenadas de P en un sistema de coordenadas en el que se está trabajando, o que se da por sobreentendido.

Observaciones

1. Si $A = O + a_1v_1 + a_2v_2 + a_3v_3$ y $v = x_1v_1 + x_2v_2 + x_3v_3$, entonces $A-O = a_1v_1+a_2v_2+a_3v_3$ y $A-O+v = (a_1+x_1)v_1+(a_2+x_2)v_2+(a_3+x_3)v_3$, por lo que $A + v = O + (a_1 + x_1)v_1 + (a_2 + x_2)v_2 + (a_3 + x_3)v_3$. Por tanto las coordenadas del punto $A + v$ son: $(a_1 + x_1), (a_2 + x_2), (a_3 + x_3)$.

2. Si $A = O + a_1v_1 + a_2v_2 + a_3v_3//$ y $B = O + b_1v_1 + b_2v_2 + b_3v_3$, entonces $B - A = (b_1 - a_1)v_1 + (b_2 - a_2)v_2 + (b_3 - a_3)v_3$ y las coordenadas del vector $B - A$ son $(b_1 - a_1,\ b_2 - a_2,\ b_3 - a_3)$.

4.6. Ecuaciones paramétricas de rectas y planos

Veamos que forma toman las ecuaciones de rectas y planos en un referencial $\{O, v_1, v_2, v_3\}$. Sea la recta $r: \;\; P = A + \lambda(B - A)$. Supongamos $A = (a_1, a_2, a_3),\ B = (b_1, b_2, b_3)$; entonces por la última observación de la sección anterior resulta $B - A = (b_1 - a_1)v_1 + (b_2 - a_2)v_2 + (b_3 - a_3)v_3$, y por la observación 1., si P tiene coordenadas (x, y, z) los puntos de la recta r tienen coordenadas

$$x = a_1+\lambda(b_1-a_1), \qquad y = a_2+\lambda(b_2-a_2), \qquad z = a_3+\lambda(b_3-a_3). \tag{4.33}$$

Si r es de la forma $P = A + \lambda v$ con $v = x_1 v_1 + x_2 v_2 + x_3 v_3$ se obtiene:

$$x = a_1 + \lambda x_1, \qquad y = a_2 + \lambda x_2, \qquad z = a_3 + \lambda x_3. \tag{4.34}$$

Estas ecuaciones se suelen llamar **ecuaciones paramétricas de la recta**, y se suele pedir, por ejemplo "hallar las ecuaciones paramétricas de la recta"que pasa por los puntos $P = (1, 0, 1)$ y $Q = (-2, 1, 2)$. En realidad debiera decirse, "hallar unas ecuaciones paramétricas de la recta"porque dependiendo del punto que se tome como base, y del vector director elegido (uno o cualquier múltiplo -no nulo- de él) se obtendrán diferentes ecuaciones paramétricas. Sin embargo, dado que la recta es la misma, independientemente de la presentación analítica elegida, seguiremos refiriéndonos a *las* ecuaciones, en todos los casos. La respuesta al ejemplo que antecede es

$$(x, y, z) = (1, 0, 1) + \lambda(-3, 1, 1)$$

Sean tres puntos no alineados $A = (a_1, a_2, a_3)$, $B = (b_1, b_2, b_3)$, $C = (c_1, c_2, c_3)$. Consideramos el plano de ecuación $P = A + \lambda(B - A) + \mu(C - A)$. Llamando (x, y, z) a las coordenadas de P, por un razonamiento análogo al utilizado para obtener (4.33), se obtiene

$$\begin{aligned} x &= a_1 + \lambda(b_1 - a_1) + \mu(c_1 - a_1), \\ y &= a_2 + \lambda(b_2 - a_2) + \mu(c_2 - a_2), \\ z &= a_3 + \lambda(b_3 - a_3) + \mu(c_3 - a_3). \end{aligned} \tag{4.35}$$

Si el plano es de la forma $P = A + \lambda v + \mu w$ con $v = x_1 v_1 + x_2 v_2 + x_3 v_3$, $w = y_1 v_1 + y_2 v_2 + y_3 v_3$ se obtiene:

$$x = a_1 + \lambda x_1 + \mu y_1, \qquad y = a_2 + \lambda x_2 + \mu y_2, \qquad z = a_3 + \lambda x_3 + \mu y_3. \tag{4.36}$$

4.7. Ecuaciones implícitas de rectas y planos

Rectas Las ecuaciones (4.33), si $a_1 \neq b_1$, $a_2 \neq b_2$, $a_3 \neq b_3$, se pueden escribir de la siguiente manera

$$\frac{x - a_1}{b_1 - a_1} = \frac{y - a_2}{b_2 - a_2} = \frac{z - a_3}{b_3 - a_3} \tag{4.37}$$

Análogamente, de (4.34) se deduce, si $x_1, x_2, x_3 \neq 0$:

$$\frac{x-a_1}{x_1} = \frac{y-a_2}{x_2} = \frac{z-a_3}{x_3} \tag{4.38}$$

En ambos casos, si alguno de los denominadores fuera cero, el correspondiente numerador también es cero. También por eliminación de λ, μ en (4.35), se obtiene la forma general de la ecuación de un plano (usando determinantes de matrices dos por dos):

$$(x-a_1)\begin{vmatrix} b_2-a_2 & b_3-a_3 \\ c_2-a_2 & c_3-a_3 \end{vmatrix} - (y-a_2)\begin{vmatrix} b_1-a_1 & b_3-a_3 \\ c_1-a_1 & c_3-a_3 \end{vmatrix} +$$

$$+(z-a_3)\begin{vmatrix} b_1-a_1 & b_2-a_2 \\ c_1-a_1 & c_2-a_2 \end{vmatrix} = 0.$$

La expresión que antecede al signo de igual puede tomarse como la definición del determinante de una matriz 3 por 3 -que será vista en detalle más adelante en el curso-, pudiéndose escribir más compactamente de la siguiente manera:

$$\begin{vmatrix} x-a_1 & y-a_2 & z-a_3 \\ b_1-a_1 & b_2-a_2 & b_3-a_3 \\ c_1-a_1 & c_2-a_2 & c_3-a_3 \end{vmatrix} = 0 \tag{4.39}$$

Partiendo de (4.36) se obtiene otra forma de la ecuación implícita de un plano. Así, si el plano está dado por el punto $A = (a_1, a_2, a_3)$ y los vectores $v = x_1v_1 + x_2v_2 + x_3v_3,\ w = y_1v_1 + y_2v_2 + y_3v_3$, la ecuación implícita toma la forma

$$\begin{vmatrix} x-a_1 & y-a_2 & z-a_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0 \tag{4.40}$$

Desarrollando estos determinantes tenemos, por ejemplo para (4.39)

$$a(x-a_1) + b(y-a_2) + c(z-a_3) = 0, \quad \text{o también } \ ax + by + cz + d = 0,$$

$$\text{donde } a = k\begin{vmatrix} b_2-a_2 & b_3-a_3 \\ c_2-a_2 & c_3-a_3 \end{vmatrix}, b = k\begin{vmatrix} b_1-a_1 & b_3-a_3 \\ c_1-a_1 & c_3-a_3 \end{vmatrix},$$

$$c = k \begin{vmatrix} b_1 - a_1 & b_2 - a_2 \\ c_1 - a_1 & c_2 - a_2 \end{vmatrix}.$$

Recíprocamente toda ecuación de la última forma, en la que por lo menos uno de los coeficientes a, b, c es distinto de cero, representa un plano porque, suponiendo, por ejemplo $c \neq 0$, ella equivale a las tres ecuaciones

$$x = \lambda, \quad y = \mu, \quad z = -\frac{d}{c} - \lambda\frac{a}{c} - \mu\frac{b}{c}$$

que son de la forma(4.36). Obsérvese que los coeficientes a, b, c sólo dependen de los vectores directores del plano. Dado que planos paralelos se pueden determinar con los mismos vectores directores, resulta que dos planos

$$ax + by + cz + d = 0, \qquad a'x + b'y + cz + d' = 0.$$

son paralelos si y sólo si $a = pa'$; $b = pb'$; $c = pc'$ para algún real p. Si, además, $d = pd'$ ambas ecuaciones representan el *mismo* plano. Observamos también que, por ejemplo de (4.38) se puede deducir

$$\frac{x - a_1}{x_1} = \frac{y - a_2}{x_2}, \quad \frac{x - a_1}{x_1} = \frac{z - a_3}{x_3}$$

que se reduce a

$$ax + by + cz + d = 0, \quad a'x + b'y + c'y + d' = 0. \tag{4.41}$$

Estas dos ecuaciones indican la propiedad bien conocida de toda recta puede determinarse por dos planos que se cortan en ella.

Ej. 1. En el ejemplo de la sección anterior, las ecuaciones implícitas de la recta por P y Q (recordar las disquisiciones sobre *las* y *unas* ecuaciones) son

$$\frac{x - 1}{-3} = y = z - 1, \text{ o bien } x + 3y = 1, \;\; y - z = -1$$

Ej. 2. Las ecuaciones implícitas de la recta por $(2, 1, 0)$, $(3, 1, 5)$ son $x - 2 = z - 5/5$, $y = 1$

Ej. 3. Las ecuaciones pàramétricas del plano por $(1, 1, 0)$ que tiene por vectores directores $u = (-1, 0, 2)$ y $v = (1, 3, 3)$ son

$$x = 1 - \lambda + \mu, \;\; y = 1 + 3\mu \;\; z = 2\lambda + 3\mu.$$

Su ecuación implícita es $-6x + 5y - 3z + 1 = 0$.
Ej. 4. $x - 2y = 0$ es la ecuación de un plano paralelo al eje coordenado Oz, que corta al plano xOy según la recta $z = 0,\ \ x - 2y = 0$.
Ej. 5. Ecuación del plano paralelo al de ecuación $x+y+z = 0$ por el punto (1,1,1). Será de la forma $x + y + z + d = 0$, por el punto $(1, 1, 1)$; por tanto $1 + 1 + 1 + d = 0$ y resulta $d = -3$.

El lector atento ya habrá percibido que si (x, y, z) designa las coordenadas genéricas de cualquier punto del espacio, dado un cierto sistema de coordenadas,la ecuación implícita de un plano viene dado por *una* ecuación lineal en (x, y, z); en general la ecuación de cualquier superficie vendrá dada por una ecuación. Y una recta, vendrá dada por *dos* ecuaciones. Los puntos que verifican las dos ecuaciones serán los de la recta.

4.8. Posiciones relativas de rectas y planos.

En esta sección se verá que el estudio de las posiciones relativas de rectas y planos llevan naturalmente a plantearse la resolución de sistemas de ecuaciones.

Dadas dos rectas en el espacio, podemos simplificar sus posiciones relativas en tres categorías. O bien se cruzan sin cortarse (esta es la posición genérica, la más común), o bien se cortan en un solo punto, o bien son paralelas (este caso incluye el de las rectas coincidentes. El caso de que sean paralelas ya fue analizado: son rectas con vectores directores colineales (uno múltiplo del otro) y coinciden o no según tengan algún punto común o no. Si los vectores directores no son colineales se trata de encontrar puntos comunes a las dos rectas. Para ello se estudia el sistema de ecuaciones formado al igualar las ecuaciones de sus coordenadas. Si las rectas se dan en forma paramétrica se tendrá un sistema con tres ecuaciones y dos incógnitas (los parámetros de cada una de las rectas). Este sistema, si los vectores directores no son colineales, puede ser compatible determinado (una solución formada por una pareja de números correspondientes a cada parámetro) o incompatible (sin solución). El primer caso corresponde al caso de corte (y la solución en cada parámetro dará el mismo punto en el espacio, pero al moverse por cada una de las rectas); el segundo caso corresponde a las rectas que se cruzan.

Ej. 1. Determinar las posiciones relativas de las rectas

$$(x, y, z) = (1, 0, 1) + \lambda(-1, 1, 0)$$
$$(x, y, z) = (0, 1, 2) + \mu(2, 0, 1).$$

Al igualar las coordenadas resulta

$$1 - \lambda = 2\mu; \quad \lambda = 1; \quad 0 = 2 + \mu$$

que no posee ninguna solución: las rectas se cruzan.
Ej. 2. Determinar las posiciones relativas de las rectas

$$(x, y, z) = (1, 0, 1) + \lambda(-1, 1, 0); \quad (x, y, z) = (0, 1, 2) + \mu(-4, 4, 0).$$

Los vectores directores son colineales y al igualar las coordenadas se obtiene

$$(1, 0, 1) + \lambda(-1, 1, 0) = (0, 1, 2) + \mu(-4, 4, 0)$$

que no tiene solución pues en la tercera coordenada resulta la igualdad de los números distintos 1 y 2: las rectas son paralelas.

Dados dos planos en el espacio podemos simplificar sus posiciones relativas en dos categorías. O bien se cortan en una recta, o bien son paralelos (este caso incluye el de los planos coincidentes). La posición relativa se manifestará analíticamente en que el sistema dado por las dos ecuaciones implícitas sea o bien indeterminado con las soluciones dependientes de un parámetro (un grado de libertad) en el caso de corte en una recta, o bien incompatible (no tiene soluciones) en e caso de planos paralelos no coincidentes, o bien indeterminado con las soluciones dependientes de dos parámetros (dos grados de libertad) en el caso de coincidencia. Si los planos estuvieran dados en forma paramética tendríamos, al igualar las coordenadas de los dos planos, un sistema de tres ecuaciones con cuatro incógnitas (dos por cada plano).

Ej. 3. Posición relativa de los planos

$$2x+3y-z=1,\ \ -x-2y+z=0.$$

Para resolver el sistema dado por las dos ecuaciones, se escaleriza, obteniéndose el sistema $2x+3y-z=1,\ -y+z=1$. Este sistema tiene infinitas soluciones que dependen de un solo parámetro (un grado de libertad) y por tanto los planos se cortan en una recta. Para determinar las ecuaciones paramétricas de esa recta, se puede tomar $y=\lambda$, resultando $z=1+\lambda,\ x=1-\lambda$.

Ej. 4. Posiciones relativas de los planos de ecuaciones paramétricas

$$\begin{aligned}(x,y,z)&=(1,0,1)+\lambda(1,1,0)+\mu(0,1,-1)\\(x,y,z)&=(0,0,3)+t(1,2,-1)+s(2,1,1).\end{aligned}$$

Igualando las correspondientes coordenadas resulta

$$\begin{array}{lcl}1+\lambda=t+2s & & \lambda+0\mu-t-2s=-1\\ \lambda+\mu=2t+s & \Longrightarrow & \lambda+\mu-2t-s=0\\ 1-\mu=3-t+s & & 0\lambda-\mu+t-s=2.\end{array}$$

Si ahora omitimos de escribir las variables, pero mantenemos su orden λ,μ,t,s, y agregamos una columna con los términos independientes, tendremos las siguientes matrices de coeficientes (primero la original, y luego la escalerizada)

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

El sistema es por tanto incompatible dado que sumando cero de cada una de las variables deberíamos obtener 3. Los planos no se cortan, ellos son paralelos. Observar que de acuerdo con lo observado en la sección tres los vectores directores del segundo plano son combinación lineal de los vectores directores del primer plano:

$$\begin{aligned}(1,2,-1)&=(1,1,0)+(0,1,-1),\\(2,1,1)&=2(1,1,0)-(0,1,-1).\end{aligned}$$

CAPíTULO 5

PRODUCTO ESCALAR Y VECTORIAL

5.1. Producto escalar

DEFINICIÓN 5.1. Dado un vector $v \in V$ llamamos **norma** de v al número $d(A, B)$ con $v = \overrightarrow{AB}$,que notaremos: $\|v\|$ (Observar que $d(A, B)$ no depende del par de puntos (A, B) que se elija como representante de v).

Propiedades:
1) $\|v\| \geqslant 0$; $\|v\| = 0$ si y solo si $v = 0$.
2) $\|av\| = |a|\,\|v\|$ para todo $a \in \mathbb{R}$ y para todo $v \in V$
3) $\|u + v\| \leqslant \|u\| + \|v\|$, para todo $u,\ v \in V$
DEMOSTRACIÓN
1) Es inmediato. ☐

2) Si $a = 0$ es trivial. Si $a \neq 0$ y $v = \overrightarrow{AB}$, por definición del producto de un número por un vector , $av = \overrightarrow{AC}$, donde $d\,(A, C) = |a|\, d\,(A, B)$. Luego $\|av\| = d\,(A, C) = |a|\, d\,(A, B) = |a|\,\|v\|$. ☐

3) Un lado de un triángulo es menor o igual que la suma de los otros dos (es la propiedad triangular de geometría métrica). ☐

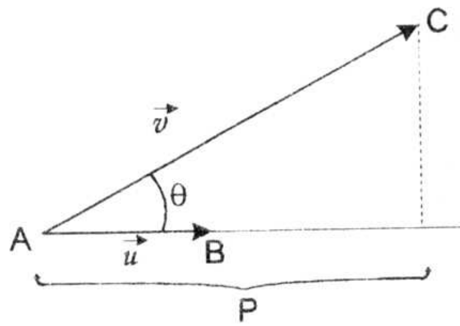
Diremos que v es un **versor** si $\|v\| = 1$. Para todo $v \neq 0$ se tiene que $\frac{v}{\|v\|}$ es un versor pues $\left\|\frac{v}{\|v\|}\right\| = \frac{1}{\|v\|}\,\|v\| = 1$.

DEFINICIÓN 5.2. Dados los vectores u y v de V llamamos **producto interno o producto escalar** de u por v al número $\|u\|\,.\,\|v\|\,cos\,\theta$ donde θ es el ángulo BAC de tres puntos tales que $\overrightarrow{AB} = u$ y $\overrightarrow{AC} = v$. Observar que el producto escalar no depende de los puntos A, B, C que se elijan para representar u y v.

Notación: u.v ; $< u, v >$, (u, v) ; etc.

En particular $v.v = \|v\|^2$; luego $\|v\| = \sqrt{v.v}$.

Si u es un versor, es decir, si $\|u\| = 1$, entonces $u.v = \|v\| \, cos\, \theta = p$ es la proyección del segmento AC sobre la recta AB de la figura.



Propiedades:

4) $u.v = v.u$

5) $(a.u)\,.v = u.\,(a.v) = a.\,(u.v)$

6) $(u_1 + u_2)\,.v = u_1.v + u_2.v \quad ; \quad u.\,(v_1 + v_2) = u.v_1 + u.v_2$.

7) $v.v \geqslant 0$ y $v.v = 0$ si y sólo si $v = 0$

Demostración:

4) Es inmediato. □

5) Si $a = 0$ la demostración es trivial . Si $a \geqslant 0$, $\text{áng}\,(au, v) = \text{áng}\,(u, v) = \theta$, luego

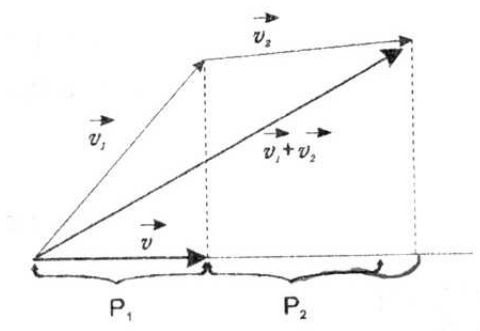
$(\,a.u\,)\,.v = \|au\|\,.\,\|v\|\,cos\theta = |a|\,.\,\|u\|\,.\,\|v\|\,cos\,\theta = a\,(u.v)$.

Si $a < 0$, $\text{áng}\,(au, v) = \pi + \text{áng}\,(u, v) = \pi + \theta$, luego $(\,a.u\,)\,v = \|a.u\|\,\|v\|\,cos\,(\pi + \theta) = -\,|a|\,\|u\|\,\|v\| \cos\,\theta = a\,\|u\|\,\|v\|\,cos\,\theta = a(u.v)$. □

6) Consideremos primero el caso $\|u\| = 1$. Entonces $v_1.u = p_1$, $v_2.u = p_2$. $(v_1 + v_2)\,u = p$ (proyección de $u_1 + u_2$). Es claro que $p = p_1 + p_2$, luego $(v_1 + v_2)\,.u = v_1.u + v_2.u$. (ver figura)

Consideremos ahora el caso general. Por el razonamiento anterior, como $\left\|\frac{u'}{\|u\|}\right\| = 1$, tenemos $(v_1 + v_2) \cdot \frac{u}{\|u\|} = v_1 \cdot \frac{u}{\|u\|} + v_2 \cdot \frac{u}{\|u\|}$, luego $(v_1 + v_2) \cdot u = v_1 \cdot u + v_2 \cdot u$ $\square$

7) Es inmediata pues $v.v = \|v\|^2$. $\square$



DEFINICIÓN 5.3. Dos vectores $u,\ v$ se dicen **ortogonales** si $u\,.v\ = 0$.

Una base $\left\{\vec{i}, \vec{j}, \vec{k}\right\}$ (a partir de aquí no escribiremos la flecha $\vec{\cdot}$) se dice **ortogonal** si $i \cdot j = i \cdot k = j \cdot k = 0$. La base se dice **ortonormal** si además de ser ortogonal verifica que $\|i\| = \|j\| = \|k\| = 1$.

Diremos que $\{O, i, j, k\}$ es un **sistema ortogonal de coordenadas** si $\{i, j, k\}$ es ortonormal.

En el resto de esta sección supondremos que todos los sistemas de coordenadas son ortonormales. En tal caso dados:

$$v = a_1 i + a_2 j + a_3 k \text{ y } w = b_1 i + b_2 j + b_3 k \text{ se tiene} v.w = \sum_{i=1,2,3} a_i b_i.$$

En particular: $\|v\| = \sqrt{v.v} = \sqrt{a_1^2 + a_2^2 + a_3^2}$.

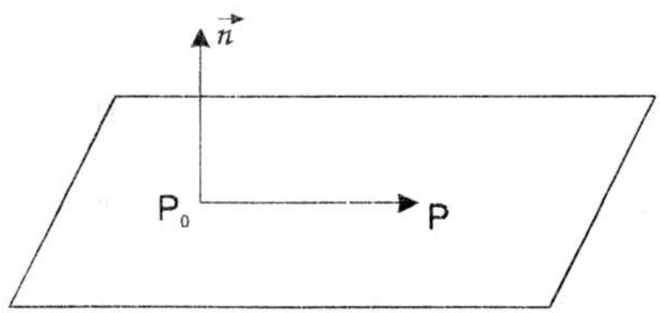
Si $P = O + a_1 i + a_2 j + a_3 k$ y $Q = O + b_1 i + b_2 j + b_3 k$ entonces $\overrightarrow{PQ} = Q - P = (b_1 - a_1)\, v_1 + (b_2 - a_2)\, v_2 + (b_3 - a_3)\, v_3$, luego
$d\,(P, Q) = \|PQ\| = \sqrt{(b_1 - a_1)^2 + (b_1 - a_1)^2 + (b_3 - a_3)^2}$.

5.2. Aplicaciones a la geometría: ángulos y distancias

Ya vimos que toda ecuación de la forma $ax + by + cz + d = 0$ representa un plano y recíprocamente. Suponiendo que el sistema de coordenadas sea ortogonal volveremos a obtener esa ecuación y mostraremos que en ese caso se puede obtener más información de esta ecuación.

Sea $\{O, i, j, k\}$ un sistema ortogonal de coordenadas, $P_0 = (x_0, y_0, z_0)$ un punto de un plano π y $n = ai + bj + ck \neq 0$ un vector de la dirección de una perpendicular a π (para abreviar diremos que n es un vector normal a π). Entonces, para todo $P \in \pi$, $P = (x, y, z)$, vale $n \cdot (P - P_0) = 0$, de aquí resulta:

$$a\,(x - x_0) + b\,(y - y_0) + c\,(z - z_0) = 0$$



Esta es entonces la ecuación del plano dado por un punto y la dirección de la normal. Es claro que esa ecuación también se escribe en la forma: $ax + by + cz + d = 0$.

Recíprocamente, dada una ecuación $ax+by+cz+d = 0$, con $a^2+b^2+c^2 \neq 0$ (por ejemplo, c$\neq$ 0), resulta : $ax + by + c\left(z + \frac{d}{c}\right) = 0$. Si el sistema de coordenadas es ortogonal esto equivale a:

$$(av_1 + bv_2 + cv_3)\,.\left(xv_1 + yv_2 + \left(z + \frac{d}{c}\right) v_3\right) = 0.$$

Poniendo $P = (x, y, z)$, $P_0 = (0, 0, -d/c)$ y $n = av_1+bv_2+cv_3$, esta ecuación se escribe: $n.(P - P_0) = 0$. Por tanto, es la ecuación del plano perpendicular a π por P_0.

OBSERVACIÓN 5.1. Si el sistema de coordenadas no es ortogonal, la ecuación $ax + by + cz + d = 0$ también representa un plano como vimos antes, pero en ese caso el vector $av_1 + bv_2 + cv_3$ no tiene por que ser normal al plano.

Veremos ahora algunas consecuencias de esto, siempre con la hipótesis de que el sistema de coordenadas es ortonormal.

a) condición para que dos planos sean paralelos: $ax+by+cz+d = 0$ y $a'x + b'y + c'z + d' = 0$ son paralelos si y sólo si $a' = \lambda a$, $b' = \lambda b$, $c' = \lambda c$ para algún $\lambda \in \mathbb{R}, (\lambda \neq 0)$.

b) condición para que una recta y un plano sean paralelos: Dados el plano y la recta de ecuaciones $ax + by + cz + d = 0$ y $\frac{x-x_0}{p} = \frac{y-y_0}{q} = \frac{z-z_0}{r}$, son paralelos si y sólo si $pa + qb + rc = 0$, pues esto significa que $n.v = 0$, donde n es normal al plano y v es un vector de la dirección de la recta.

c) ángulo entre dos planos: Para dos vectores cualesquiera, $u \neq 0$, $v \neq 0$, tenemos $\frac{u.v}{\|u\|.\|v\|} = cos\,\theta$. Como el ángulo θ entre dos planos es igual al que forman sus vectores normales respectivos, tendremos: $cos\,\theta = \frac{n_1.n_2}{\|n_1\|\ \|n_2\|}$ donde n_1 y n_2 son vectores normales a esos planos: luego el ángulo θ entre los planos de ecuaciones *ax+by+cz+d*=0 y *a'x+b'y+c'z+d'*=0 verifica:

$$cos\,\theta = \frac{aa' + bb' + cc'}{\sqrt{a^2 + b^2 + c^2}.\sqrt{a'^2 + b'^2 + c'^2}}$$

d) Ángulo de una recta con un plano: Sean n un vector normal al plano y v un vector de la dirección de la recta . El ángulo del plano con la recta es entonces $\theta = \pi/2 - \varphi$ donde φ es el ángulo determinado por n

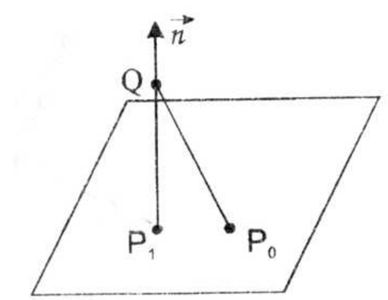
y v. Luego el ángulo de $ax{+}by{+}cz{+}d\ {=}0$ con $\frac{x-x_0}{p} = \frac{y-y_0}{q} = \frac{z-z_0}{r}$ se puede calcular mediante:

$$sen\,\theta = cos\,\varphi = \frac{ap + bq + cr}{\sqrt{a^2 + b^2 + c^2}\sqrt{p^2 + q^2 + r^2}}$$

e) Ángulo de dos rectas: El ángulo θ de $\frac{x-x_0}{p} = \frac{y-y_0}{q} = \frac{z-z_0}{r}$ con $\frac{x-x_1}{p'} = \frac{y-y_1}{q'} = \frac{z-z_1}{r'}$ se calcula mediante la fórmula

$$cos\,\theta = \frac{pp' + qq' + rr'}{\sqrt{p^2 + q^2 + r^2}\sqrt{p'^2 + q'^2 + r'^2}}.$$

f) Distancia de un punto a un plano: Definimos la distancia $d\,(Q, \pi)$ de Q al plano π como el mínimo de $d(Q,P)$ donde $P \in \pi$. Es claro que el mínimo de $d(Q,P)$ se obtiene cuando $P{=}P_1{=}\pi \cap r$, donde r es la perpendicular a π por Q. Luego : $d\,(Q, \pi) = d\,(Q, P_1) = \left|(Q - P_0)\,.\frac{n}{\|n\|}\right|$, siendo P_0 un punto cualquiera del plano π.



Como la ecuación del plano es $(P - P_0)\,.n = 0$, esto significa que d se obtiene reemplazando P por $O{=}(0{,}0{,}0)$ en el primer miembro de la ecuación del plano.

Consideremos ahora un sistema ortonormal de coordenadas. Si $P_0 = (x_0, y_0, z_0)$, $Q = (\overline{x}, \overline{y}, \overline{z})$, $P = (x, y, z)$, $n = ai + bj + ck$ entonces:

$$(Q - P_o)\ \cdot\ n\ =\ a\,(\overline{x} - x_o) + b\,(\overline{y} - y_o) + c\,(\overline{z} - z_o)\ =\ a\overline{x} + b\overline{y} + c\overline{z} + d.$$

Luego : $d(Q,\pi) = \left|(Q-P_0)\cdot\frac{n}{\|n\|}\right| = \left|\frac{\overline{a}x+\overline{b}y+\overline{c}z+d}{\sqrt{a^2+b^2+c^2}}\right|$

5.3. Aplicaciones: ecuaciones de algunas superficies

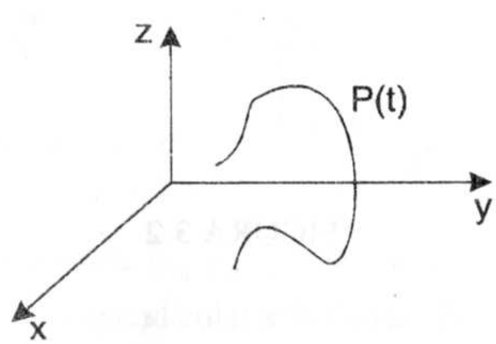
a) **Esferas de centro** (x_0, y_0, z_0) **y radio *r:*** En un sistema ortogonal de coordenadas la ecuación es: $(x-x_0)^2+(y-y_0)^2+(z-z_0)^2=r^2$

b) **Superficies de revolución:** Son las engendradas por una curva (generatriz), que gira alrededor de una recta (eje).

OBSERVACIÓN 5.2. Una **curva** en el espacio puede darse:

- O bien en forma **paramétrica**, por medio de tres ecuaciones de la forma $\begin{cases} x = x(t) \\ y = y(t) \quad t\in\mathbb{R} \\ z = z(t) \end{cases}$

con tres funciones contínuas de un parámetro real (para cada valor del parámetro t se tendrá un punto *(x(t),y(t),z(t))=P(t)* que está en la curva).



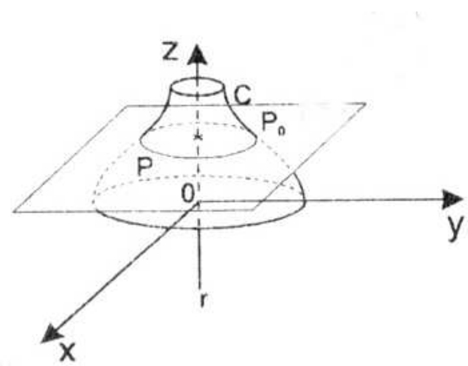
- O bien en la forma reducida $\begin{cases} f(x,y,z)=0 \\ g(x,y,z)=0 \end{cases}$

Cuando se tienen las ecuaciones paramétricas, **a veces** puede eliminarse el parámetro t y obtenerse una forma reducida. Generalmente hay muchos sistemas de ecuaciones posibles que representan a la misma curva.

OBSERVACIÓN 5.3. Una **superficie** en el espacio puede darse:
- O bien en forma parametrizada, por medio de tres ecuaciones con dos parámetros (como la ecuación paramétrica del plano).
- O bien en forma reducida por medio de una ecuación del tipo: $f(x, y, z) = 0$.

OBSERVACIÓN 5.4. La curva *C:* $\begin{cases} f(x, y, z) = 0 \\ g(x, y, z) = 0 \end{cases}$ es la **intersección** de dos superficies S y S' donde *S: f(x,y,z) = 0* y *S': g(x,y,z) =0.*

Ecuación de la superficie de revolución con eje $\vec{Oz}$ Sea r el eje $r : \begin{cases} x = 0 \\ y = 0 \end{cases}$, y C la generatriz *C:*$\begin{cases} x = 0 \\ z = f(y) \end{cases}$



La superficie de revolución S está constituida por **todos** los puntos P del espacio que cumplen, a la vez, dos condiciones:

$$\begin{cases} P \in \pi \perp r \ \ por \text{ algún } \quad P_0 \in C \\ d(P, r) = d(P_0, r) \end{cases}$$

Las condiciones se traducen en:

$$\begin{cases} z - z_0 = 0 & P \in \pi\perp\ , r,\ por\ P_0 \\ x_0 = 0 & P_0 \in C \\ z_0 = f\ (y_0) & P_0 \in C \\ \sqrt{x^2+y^2} = \sqrt{x_0^2+y_0^2}, & d\,(P,r) = d\,(P_0,r) \end{cases}$$

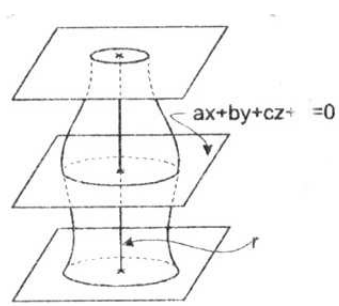
de donde, eliminando x_0, y_0, z_0 se obtiene: $z = f\left(\pm\sqrt{x^2+y^2}\right)$.

OBSERVACIÓN 5.5. Dada una ecuación *F(x,y,z)* $= 0$ **no** intentaremos, en general, reconocer si es o no una superficie de revolución. Observaremos sólo que si se halla una familia de planos paralelos (es decir $ax+by+cz+\alpha = 0$ con $(a,\ b,\ c)$ fijo y α variable) que cumplan las condiciones de más abajo, entonces *F(x,y,z)* $= 0$ es una superficie de revolución que verifica (ver figura 3.3):

i) La intersección $\begin{cases} F\,(x,y,z) = 0 \\ ax+by+cz+\alpha = 0 \end{cases}$ es una **circunferencia** para cada valor de α.

ii) El centro de la circunferencia $(x_\alpha, y_\alpha, z_\alpha)$ pertenece, para todo α, a una recta fija r colineal al vector *(a,b,c)* (o sea perpendicular a la familia de planos paralelos).

EJEMPLO 5.1. Demostrar que $xy+yz+zx = 7$ es una superficie de revolución.



Se puede escribir como: $(x+y+z)^2-\left(x^2+y^2+z^2\right)=14$ Cortándola en el plano:$x+y+z+\alpha=0$ se obtiene:

$$\begin{cases} x+y+z+\alpha=0 \\ (x+y+z)^2-\left(x^2+y^2+z^2\right)=14 \end{cases}$$

i) Es una **circunferencia** porque es la intersección de la esfera $x^2+y^2+z^2=14+\alpha^2$ con el plano $x+y+z+\alpha=0$. **ii)** El centro C_α de la circunferencia se halla trazando la perpendicular al plano $x+y+z+\alpha=0$ que pasa por el centro de la esfera $x^2+y^2+z^2=14+\alpha^2$, o sea:

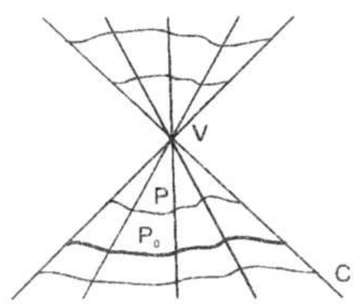
$$C_\alpha=\begin{cases} x+y+z+\alpha=0 \\ x=y=z \ \ recta\ \ por\,(0,0,0)\perp \\ \ \ al\ \ plano,\ \ x+y+z+\alpha=0 \end{cases}$$

$$C_\alpha=\begin{cases} x=-\alpha/3 \\ y=-\alpha/3 \\ z=-\alpha/3 \end{cases}$$

es una recta colineal al vector $(-1/3,-1/3,-1/3)$ o sea, colineal con *(*1,1,1*).*

■

c) Conos: Se llama "cono" o "superficie cónica" de vértice V y directriz C (donde V es un punto y C es una curva dada , tales que V$\notin$C), a la superficie engendrada por todas las rectas que pasan por V y cortan a C.



Si $V=(a,b,c)$ y $C=\begin{cases} x=f(t) \\ y=g(t) \\ z=h(t) \end{cases}$

La ecuación paramétrica del cono se puede obtener con los parámetros t y λ como:

$$S = \begin{cases} x = a + \lambda(f(t) - a) \\ y = b + \lambda(g(t) - b) \\ z = c + \lambda(h(t) - c) \end{cases}$$

Para escribir la ecuación reducida habrá que eliminar los parámetros t, λ, para obtener una ecuación del tipo

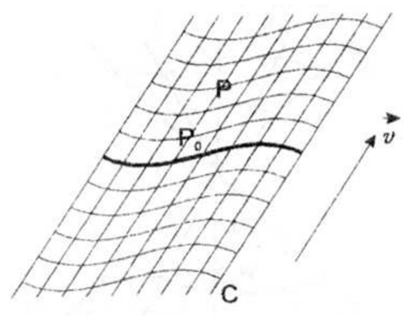
$$\mathrm{F}(x,y,z) = 0.$$

EJEMPLO 5.2. Hallar las ecuaciones reducidas y paramétrica del cono de vértice $V = (0,0,0)$ y generatriz $\begin{cases} x = sen\,t \\ y = cos\,t \\ z = 1 \end{cases}$. La ecuación parametrizada es: $\begin{cases} x = \lambda\, sen\,t \\ y = \lambda\, cos\,t \\ z = 1 + \lambda \end{cases}$ Despejando t y λ se obtiene la ecuación reducida $x^2+y^2=(z-1)^2$. ■

d) Cilindros: Se llama "cilindro" o "superficie cilíndrica" de directriz 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 C.

Si $v=(a,b,c)$ y $C=\begin{cases} x=f(t) \\ y=g(t) \\ z=h(t) \end{cases}$ Un punto $P=(x,y,z)$ pertenece al cilindro si y sólo si existe algún punto $P_0=(x_0,y_0,z_0)$ tal que:

$$\begin{cases} P_0 \in C \\ P \in,\ recta\ por\ P_0\ colineal\ a\ v \end{cases}$$



. Entonces, las ecuaciones paramétricas del cilindro son:

$$S = \begin{cases} x = f(t) + \lambda a \\ y = g(t) + \lambda b \\ z = h(t) + \lambda c \end{cases}$$

de las que habrá que eliminar los parámetros $t,\ \lambda,$ para obtener la ecuación reducida del tipo $F(x, y, z) = 0$.

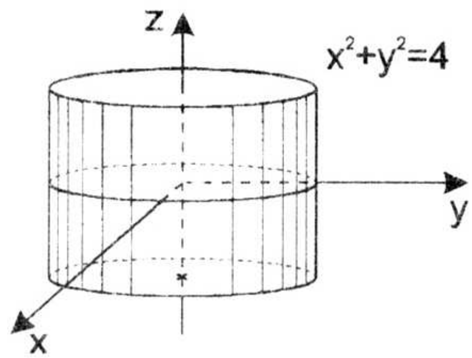
OBSERVACIÓN 5.6. La curva C también puede estar dada como

$$C : \begin{cases} f(x, y, z) = 0 \\ g(x, y, z) = 0 \end{cases}$$

en cuyo caso se trabajará análogamente , pero con un parámetro menos.

OBSERVACIÓN 5.7. Una ecuación $F(x, y, z) = 0$ en la que no figura la variable z, por ejemplo, representa un cilindro de generatrices colineales con el eje Oz. Si $P_0 = (x_0, y_0, z_0)$ verifica la ecuación $F(x_0, y_0, z_0) = 0$, entonces tomando $x = x_0,\ y = y_0,\ z = z_0 + \lambda$ (eso es la recta por P_o colineal con *(0,0,1)*), se tiene $F(x_0, y_0, z_0 + \lambda) = F(x_0, y_0, z_0) = 0$ porque F es, en realidad, independiente de z. O sea: cuando P_o está en la superficie, toda la recta por la colineal al eje Oz también la está.

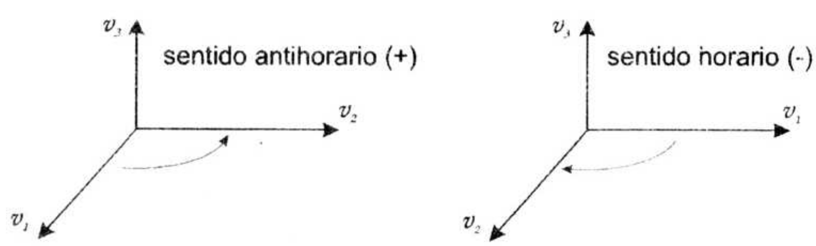
EJEMPLO 5.3. $x^2 + y^2 = 4$ es un **cilindro** con generatrices paralelas al eje Oz y directriz, por ejemplo, en la circunferencia $\begin{cases} x^2 + y^2 = 4 \\ z = 0 \end{cases}$



■

5.4. Producto vectorial.

Consideremos la terna ordenada de vectores de V, $\{\vec{i}, \vec{j}, \vec{k}\}$, que constituye una base de V. Estas las vamos a clasificar en dos grupos: uno formado por las ternas $\{i, j, k\}$ tales que para un observador parado en v_3, la rotación de ángulo convexo (o sea, de medida $< \pi$) que lleva i en j es de sentido antihorario; el otro formado por las ternas en las que esa rotación tiene sentido horario.



Para definir el producto vectorial necesitamos elegir uno de los dos tipos como terna preferida. Nos quedaremos para esto con las ternas del primer tipo que llamaremos **positivas**.

Esta convención sirve para definir el producto vectorial como una función de $V \times V$ *en* V (así como el producto escalar era una función de $V \times V$ *en* $\mathbb{R}$) que a cada par de vectores v, w le asocia un vector, que denotaremos $v \wedge w$, que cumple:
a) $\|v \wedge w\| = \|v\| . \|w\| .sen\, \theta$ (θ ángulo de v con w)
b) $(v \wedge w) .v = 0 \ y \ (v \wedge w) .w = 0$
c) Si $v \neq 0 \ \ y \ \ w \neq 0$ la terna $\{v, w, v \wedge w\}$ es positiva.

Propiedades:

1) $\|v \wedge w\|$ es el doble del área del triángulo determinado por esos vectores.

2) $v \wedge w = 0$ si y sólo si v y w son colineales (en particular, si alguno de ellos es el vector 0).

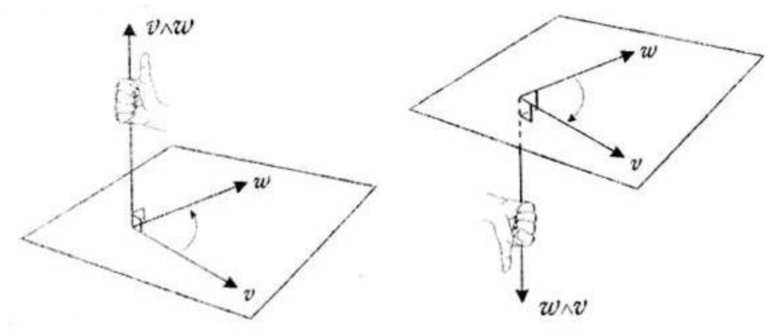
3) Respecto de la condición c) de la definición, corresponde observar que si $v \wedge w \neq 0$ la terna $\{v, w, v \wedge w\}$ es una base, pues por b) esos vectores no son coplanares salvo que v y w sean colineales, en cuyo caso $v \wedge w = 0$

4) $v \wedge w = -(w \wedge v)$ (en particular, esta operación no es conmutativa) Para verificar esto basta notar que si para cierto observador la rotación del ángulo $< \pi$ que lleva v en w es de sentido antihorario, para el mismo observador la que lleva w en v es de sentido horario. De modo que el observador debe ubicarse del lado opuesto del plano u,w para que esa rotación aparezca de sentido trigonométrico . Luego esa debe ser la ubicación de $w \wedge v$ para que la terna $[w, v, w \wedge v]$ sea positiva.

5) Este producto no es asociativo. Se puede deducir que:

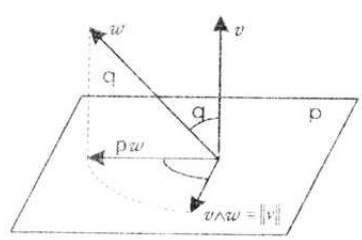
$$(u \wedge v) \wedge w + (v \wedge w) \wedge u + (w \wedge u) \wedge v = 0$$

de donde, usando 4), $(u \wedge v) \wedge w = u \wedge (v \wedge w) + v \wedge (w \wedge u)$, y como en general $v \wedge (w \wedge u) \neq 0$, la propiedad asociativa no vale.



6) Para todo $\lambda \in \mathbb{R}$: $\lambda\,(v \wedge w) = (\lambda v) \wedge w = v \wedge (\lambda w)$. Esta propiedad se deduce directamente de la definición en el caso $\lambda \geqslant 0$. Para el caso $\lambda < 0$, hay que observar que del hecho que la terna $\{v, w, v \wedge w\}$ es positiva se deduce que $\{-v, w, -\,(v \wedge w)\}$ y
$\{v, -w, -\,(v \wedge w)\}$ son también positivas.

7) El producto vectorial es distributivo respecto de la suma de vectores. Esto es:
a) $(v_1 + v_2) \wedge w = v_1 \wedge w + v_2 \wedge w$.
b) $v \wedge (w_1 + w_2) = v \wedge w_1 + v \wedge w_2$.
Para demostrar esta propiedad hacemos algunas observaciones previas. En primer lugar, observamos que llamando π al plano con vector normal v e indicando por $p\ w$ la proyección de w sobre ese plano, y con $r(pw)$ el vector que se obtiene rotando esa proyección un ángulo de medida $\pi/2$ en sentido antihorario observado desde v, se verifica que: $v \wedge w = \|v\|\,.r\,(pw)$ pues $\|r\,(pw)\| = \|pw\| = \|w\|\,.sen\,\theta$ (ver figura)



Como $p\,(w_1 + w_2) = pw_1 + pw_2$ y $r\,(pw_1 + pw_2) = r\,(pw_1) + r\,(pw_2)$ tendremos que: $v \wedge (w_1 \wedge w_2) = \|v\|\,.r\,(p\,(w_1 + w_2)) = \|v\|\,.r\,(pw_1) + \|v\|\,.r\,(pw_2) = v \wedge w_1 + v \wedge w_2$.

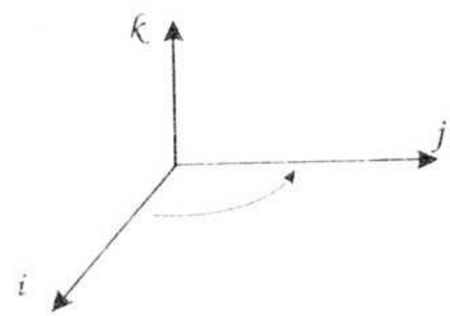
Así se prueba b) y análogamente se obtiene a).

8) Si $\{i, j, k\}$ es una base ortonormal positiva de V , entonces:

$$\begin{aligned} &i \wedge i = j \wedge j = k \wedge k = 0; \\ &i \wedge j = k; \quad j \wedge k = i; \quad k \wedge i = j \end{aligned}$$

(es decir, en la sucesión $(i, j, k, i, j, \cdots)$ el producto de dos vectores sucesivos da el que le sigue):

$$j \wedge i = -k,\; k \wedge j = -i;\; i \wedge k = -j$$



9) De 4.6), 4.7) y 4.8) resulta que si $v = a_1 i + a_2 j + a_3 k$ y $w = b_1 i + b_2 j + b_3 k$. entonces:

$$v \wedge w = (a_2 b_3 - a_3 b_2)\, i - (a_1 b_3 - a_3 b_1)\, j + (a_1 b_2 - a_2 b_1)\, k.$$

Obsérvese que este resultado puede recordarse pensando en el desarrollo por la primer fila de un determinante:

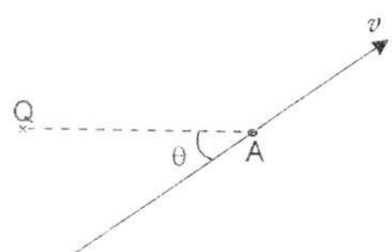
$$\begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2)\, i - (a_1 b_3 - a_3 b_1)\, j + (a_1 b_2 - a_2 b_1)\, k$$

Esta expresión del producto vectorial puede también tomarse como su definición. Más precisamente, dada una terna ortonormal cualquiera $\{i,j,k\}$ (positiva o negativa), puede definirse el producto de dos vectores por la expresión dada en 4.9). Se comprueba entonces sin dificultad que el vector: $u = (a_2b_3 - a_3b_2)\, i - (a_1b_3 - a_3b_1)\, j + (a_1b_2 - a_2b_1)\, k$ verifica $\|u\| = \|v\| \,.\, \|w\| \,.sen\,\theta$ y $u.v = u.w = 0$. Si además $\{i,j,k\}$ es una terna positiva puede verificarse, que $\{v,w,u\}$ es también positiva. Luego $u = v \wedge w$, por lo tanto esta definición de $v \wedge w$ coincide con la inicial.

5.5. Aplicaciones geométricas.

a) Distancia de un punto a una recta: sea r una recta dada por un punto A y un vector v. La distancia de Q a r es:

$$d\,(Q,r) = |d\,(Q,A)\,.sen\,\theta| = \|AQ\|\,.sen\,\theta = \frac{\|AQ \wedge v\|}{\|v\|}$$



Si $A = (x_o, y_o, z_o);\ Q = (\overline{x}, \overline{y}, \overline{z});\ v = ai + bj + ck$ entonces la distancia es $d(Q,r) =$

$$\frac{\sqrt{[(\overline{y}-y_0)c-(\overline{z}-z_0)b]^2+[(\overline{x}-x_0)c-(\overline{z}-z_0)a]^2+[(\overline{x}-x_0)b-(\overline{y}-y_0)a]^2}}{\sqrt{a^2+b^2+c^2}}$$

EJEMPLO 5.4. En el punto 3 de este capítulo se vieron algunas superficies de revolución particulares. Sea $r \begin{cases} x = y \\ z = 0 \end{cases}$ $C : \begin{cases} x = 0 \\ zy = 1 \end{cases}$. Hallar la superficie de revolución de eje r y generatriz C.

Las ecuaciones son:

$$\begin{cases} x_0 = 0, \;\; P_0 \in C \\ z_0 y_0 = 1, \;\; P_0 \in C \\ x - x_0 + y - y_0 = 0, \;\; plano, \;\; \pi, \;\; por, \;\; P_0 \perp r \\ 2z^2 + (x-y)^2 = 2z_0^2 + (x_0 - y_0)^2, \;\; dist\,(P, r) = dist\,(P_0, r) \end{cases}$$

eliminando $x_0, y_0,\, z_0$ resulta $(x+y)^2 \left(z^2 - 2xy\right) = 1$ ■

b)**Intersección de dos planos no paralelos:**

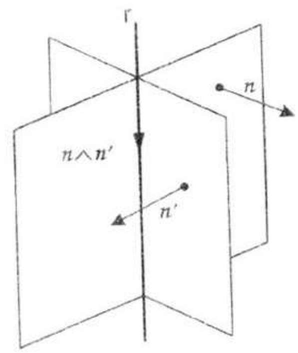
Sea $r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$ una recta (dada como intersección de dos planos no paralelos). Supongamos que queremos las ecuaciones paramétricas de esa recta, es decir, las de la forma: $x = x_0 + \lambda p,\; y = y_0 + \lambda q,\; z = z_0 + \lambda r$ (o también $P = P_0 + \lambda v$). Para esto se necesita hallar un punto $P_0 = (x_0, y_0, z_0)$ de la recta y un vector $v = p\,i + q\,j + r\,k$ de la dirección de r. Esto se puede hacer sin usar el producto vectorial, resolviendo el

sistema de ecuaciones

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

Usando el producto vectorial: tomar $P_0 = (x_0, y_0, z_0)$ una solución particular del sistema de ecuaciones. Si el sistema de coordenadas es ortogonal, entonces: $n = ai + bj + ck$ y $n' = a'i + b'j + c'k$ son vectores normales a los planos, y entonces $n \wedge n'$ es un vector de la intersección de ambos planos (luego, está en la dirección de la recta de intersección de ambos).

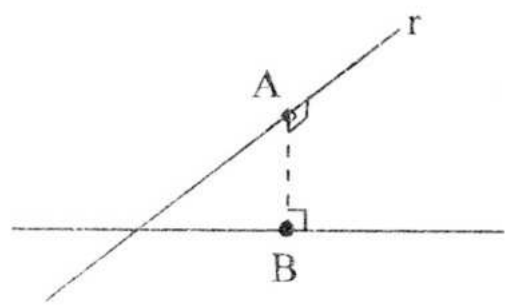
Luego: $n \wedge n' = (bc' - b'c)\, i - (ac' - ca')\, j + (ab' - ba')\, k$ es de la dirección de r.



c) Distancia entre dos rectas: Se define la distancia $d\,(r, r')$ entre dos rectas r, r' como el mínimo de $d\,(P, P')$ donde $P \in r$ y $P' \in r'$.

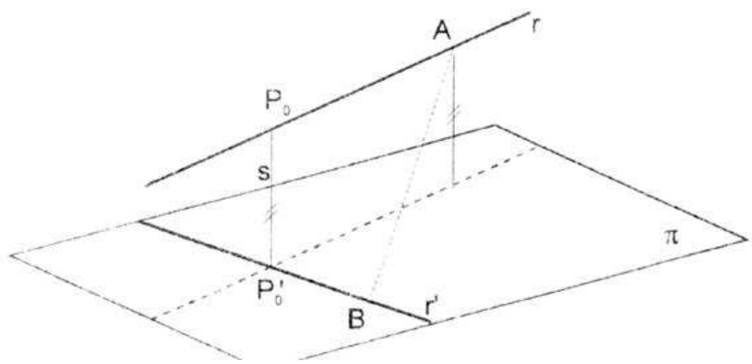
Es claro que este mínimo se obtiene cuando la recta PP' es perpendicular al mismo tiempo a r y a r'. Sea r dada por un punto $A \in r$ y un vector v de su dirección y r' por $B \in r$ y un vector w.

Para tener $d\,(r, r')$: si r y r' son paralelas: basta tomar el punto A y hallar $dist\,(A, r')$; si r y r' se cortan: $d\,(r, r')$= 0.



Supongamos ahora que r y r' no son coplanares. Si s es la perpendicular común , $P_o \;=\; s \cap r$ y $P_0' = s{\cap}r'$, entonces: $d\,(r, r') = d\left(P_0, P_0'\right) = d\,(A, \pi)$ donde π es el plano paralelo a r que contiene a r'. (ver figura) . Como versor de la dirección de s puede tomarse $n = \frac{v \wedge w}{\|v \wedge w\|}$. Luego :

$$d\left(r, r'\right) = d\left(A, \pi\right) = \frac{1}{\|v \wedge w\|}\left|\left(v \wedge w\right).\overrightarrow{AB}\right|$$

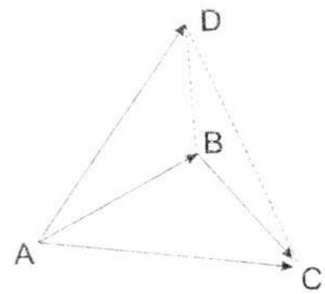


d) Perpendicular común a dos rectas: Si las dos rectas son paralelas, una perpendicular común es la recta perpendicular a una de ellas trazada por un punto de la otra. Si se interceptan , el problema es también fácil de resolver. Supongamos ahora que las dos rectas no son coplanares. La perpendicular común s está en el plano π de s y r, y en el π' determinado por s y r'. Luego $s = \pi \cap \pi'$. Dada la dirección v de r y la w de r', y $A \in r,\ B \in r'$, el plano π queda determinado por A,v y $v \wedge w$. El π' por B,w,$v \wedge w$, pues $v \wedge w$ es un vector de la dirección de s. Las ecuaciones de π y π' así obtenidas constituyen un par de ecuaciones que representan la recta s.

e) Volumen de un tetraedro: Consideremos un tetraedro dado por sus vértices A, B, C, D. Su volumen es $V = \frac{1}{3}$ área de la base $\times$ altura. Tendremos: área base $= 1/2\,\|(B-A) \wedge (C-A)\|$.

Si n es el versor normal a la base, la altura es $|n.(D-A)|$ con $n = \frac{(B-A)\wedge(C-A)}{\|(B-A)\wedge(C-A)\|}$. Luego $V = 1/6.\,\|(B-A) \wedge (C-A)\|\,.\frac{|[(B-A)\wedge(C-A)].(D-A)|}{\|(B-A)\wedge(C-A)\|}$

$$\text{por lo que } V = 1/6.\,\left|\left[(B-A) \wedge (C-A)\right].(D-A)\right|$$



5.6. Producto mixto

Es una operación definida de $V \times V \times V$ en $\mathbb{R}$. Dados tres vectores u, v, y w, llamamos producto mixto al número $(v \wedge w).u$ que indicamos $v \wedge w.u$, o también $u.(v \wedge w)$. Consideremos un sistema ortonormal $\{i, j, k\}$ y

sean:

$$v = a_1 i + a_2 j + a_3 k, w = b_1 i + b_2 j + b_3 k, u = c_1 i + c_2 j + c_3 k.$$

Entonces:

$$v \wedge w = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} i - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} j + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} k.$$

$$v \wedge w.u = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} c_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} c_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} c_3 = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

(Por desarrollo por la primer fila de ese determinante)

Usando el hecho de que si se permutan dos filas de una matriz entre sí el determinante sólo cambia de signo, resulta que dos de esas permutaciones no cambian el determinante; luego:

$$\begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

En consecuencia: $(v \wedge w).u = (u \wedge v).w = (w \wedge u).v$. Es decir, que en la sucesión (u,v,w,u,v,w) el producto vectorial de dos vectores sucesivos multiplicado escalarmente por el siguiente, da lo mismo cualesquiera sean los tres vectores sucesivos. Se puede entonces hablar del producto mixto de u, v, w sin indicar si se trata de $(u \wedge v).w$ o de $u.(v \wedge w)$, pues el resultado es el mismo. Es por esto que se escribe (u,v,w) como notación para $(u \wedge v).w = u.(v \wedge w)$.

Observamos que $(v,w,u) = 0$ si y sólo si áng$(v \wedge w, u) = \pi/2$ o alguno de los vectores es el nulo. Como áng $(v \wedge w, v) =$ áng $(v \wedge w, w) = \pi/2$, para que $(v,w,u) = 0$ es necesario y suficiente que v, w y u sean coplanares. (Obsérvese que esto podría sacarse como conclusión del cálculo del volumen del tetraedro. $(v, w, u) = 0$ si y sólo si el volumen del tetraedro determinado por esos vectores es 0, esto es equivalente a decir que los vectores son coplanares).

Esta condición permite escribir la ecuación vectorial del plano dado por tres puntos A, B, C en otra forma. Decir que P pertenece a ese plano equivale a decir que los vectores $P - A, B - A$ y $C - A$ son coplanares, o sea $(P - A, B - A, C - A) = 0$.

Pasando a coordenadas, si P=(x,y,z),$A=(a_1, a_2, a_3)$, $B = (b_1, b_2, b_3)$ y $C = (c_1, c_2, c_3)$ esta ecuación se escribe:

$$\begin{vmatrix} x - a_1 & y - a_2 & z - a_3 \\ b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \\ c_1 - a_1 & c_2 - a_2 & c_3 - a_3 \end{vmatrix} = 0, \quad \text{ó también:} \begin{vmatrix} x & y & z & 1 \\ a_1 & a_2 & a_3 & 1 \\ b_1 & b_2 & b_3 & 1 \\ c_1 & c_2 & c_3 & 1 \end{vmatrix} = 0.$$

Para ver esto obsérvese que si se le resta la 2da. fila a la 1ra., la 3da. y la 4ta. filas, el determinante no cambia. Así se tiene:

$$\begin{vmatrix} x - a_1 & y - a_2 & z - a_3 & 0 \\ a_1 & a_2 & a_3 & 1 \\ b_1 - a_1 & b_2 - a_2 & b_3 - a_3 & 0 \\ c_1 - a_1 & c_2 - a_2 & c_3 - a_3 & 0 \end{vmatrix} = \begin{vmatrix} x - a_1 & y - a_2 & z - a_3 \\ b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \\ c_1 - a_1 & c_2 - a_2 & c_3 - a_3 \end{vmatrix} = 0.$$

CAPíTULO 6

ESPACIOS VECTORIALES

En este Capítulo se introducirá el concepto básico de este curso, que es el del título . Trataremos de hacer hincapié en las nociones geométricas simples que son comunes a otras ramas de la matemática y motivar en todo lo posible en base al concepto de vectores en el espacio ambiente E estudiado en el capítulo sobre **Rectas y planos en el espacio**.

En muchas ramas de la matemática aparecen conjuntos entre cuyos elementos se realizan *combinaciones lineales*. Ya hemos visto tres ejemplos de esto en nuestro curso: las matrices, las n-uplas y los vectores del espacio ambiente. Pero el lector puede recordar otros ejemplos como las funciones continuas y los polinomios con los cual también se opera. En los ejemplos desarrollados hasta el momento se hizo notar que muchos de los resultados obtenidos solo dependían de las propiedades que tenían las operaciones y no de otras particularidades del ejemplo en cuestión. También hemos observado que estas propiedades son las mismas en los tres casos abordados hasta ahora. Eso nos motiva a intentar un desarrollo abstracto de la teoría. Es decir a introducir un conjunto arbitrario con cuyos elementos se supone se puede operar con propiedades análogas a las que por ejemplo tienen las operaciones con n-uplas. De este modo se desarrolla de una sola vez una teoría que puede aplicarse a múltiples y a priori disímiles ejemplos. Naturalmente los ejemplos que nos motivaron y serán una constante fuente de inspiración para encontrar las definiciones, enunciados y demostraciones de nuestra teoría.

En particular el lector podrá recurrir frecuentemente a motivaciones basadas en los vectores de E, en las ternas de números o en las matrices, pero el objeto en estudio es un concepto abstracto, mucho más general, cuya comprensión y manejo fluido es una de las finalidades principales de este curso.

Un concepto fundamental que re-aparecerá en este capítulo es el de *independencia lineal*, cuya definición ya hemos establecido antes y que parece particularmente simple, pero cuya comprensión en toda su riqueza -nuestra experiencia lo indica- merece una atención especial. Realizar muchos ejercicios con este nuevo concepto abstracto, comprender bien como se le utiliza en diversas demostraciones, es una recomendación que hacemos amistosamente al lector.

6.1. Espacios vectoriales

DEFINICIÓN 6.1. Un **espacio vectorial** sobre un cuerpo $\mathbb{K}$ (que en este curso será el de los números reales R, o el de los números complejos C), es una cuádrupla ordenada $\{V, \mathbb{K}, +, \cdot\}$

1) Un conjunto $V \neq \varnothing$ cuyos elementos se llamarán vectores.

2) El cuerpo $\mathbb{K}$, cuyos elementos se llamarán escalares.

3) Una operación llamada suma de vectores, y denotada con el símbolo $+$ cuyo dominio es el producto cartesiano $V \times V$ y cuyo codominio es V ($+ : V \times V \to V$ quiere decir que u+v está definida para todo par (u,v) $\in V \times V$ y da como resultado un elemento V) .

4) Una operación llamada producto de un escalar por un vector, y denotada con el símbolo $\cdot$; cuyo dominio es $\mathbb{K} \times V$ y cuyo codominio es V $(\cdot : \mathbb{K} \times V \to V)$, que verifica las siguientes propiedades:

Para la suma:

S1) Asociativa: $(u + v) + w = u + (v + w)\ \forall\ u, v, w \in V$

S2) Conmutativa: $u + v = v + u\ \forall\ u, v \in V$

S3) Neutro: Existe un elemento $\vec{0} \in V$, llamado “vector nuloz denotado con $\vec{0}$, tal que $\vec{0} + u = u + \vec{0} = u, \forall u \in V$.

S4) Opuesto: Para cada vector $u \in V$, existe un vector llamado "opuesto de u", denotado como $-u$, tal que $u + (-u) = (-u) + u = \vec{0}$.

Para el producto:

P1) $\alpha.(\beta.u) = (\alpha.\beta).u$

P2) Multiplicación por la unidad del cuerpo: $1.u = u \,, \forall\, u \in V$

P3) Distributiva respecto de suma de vectores:$\alpha.(u + v) = \alpha.u{+}\alpha.v \,, \forall \in \mathbb{K}, \forall\, u, v \in V$

P4) Distributiva respecto de suma de escalares $(\alpha + \beta).u = \alpha.u + \beta.u \,, \ \forall \alpha, \beta \in \mathbb{K}, \forall\, u \in V$

OBSERVACIÓN 6.1. En P4) la suma de escalares (suma dentro de $\mathbb{K}$) se escribió con el símbolo $+$ pero no hay que confundirla con la suma dentro de V, que se representa también con el símbolo $+$.

OBSERVACIÓN 6.2. El lector comparará estas propiedades de la definición de espacio vectorial con las propiedades de los vectores de E, de las n-uplas de números reales, o de las matrices con sus operaciones de suma y producto por un escalar.

PROPOSICIÓN 6.3. *En todo espacio vectorial* $\{V, \mathbb{K}, +, \cdot\}$*se cumple:*

1) el neutro es único.

2) Para cada vector $u \in V$ *su opuesto* $-u$ *es único..*

DEMOSTRACIÓN. 1) Sean $\vec{0_1}$ y $\vec{0_2}$ dos neutros: $\vec{0_1} = \vec{0_1} + \vec{0_2} = \vec{0_2}$, luego son iguales.

2) Sean $(-u)_1$ y $(-u)_2$ dos opuestos de u:

$(-u)_1 = \vec{0} + (-u)_1 = ((-u)_2 + u) + (-u)_1 = (-u)_2 + (u + (-u)_1) = (-u)_2 + \vec{0} = (-u)_2$, luego coinciden. □

OBSERVACIÓN 6.4. A los espacios vectoriales sobre $\mathbb{R}$ se los llamará espacios vectoriales reales, o $\mathbb{R}$–espacios vectoriales. A los espacios vectoriales sobre $\mathbb{C}$, espacios vectoriales complejos, o $\mathbb{C}$–espacios vectoriales. A veces, por abuso de lenguaje los espacios vectoriales se indican por el conjunto de vectores V, sobreentendiendo el cuerpo $\mathbb{K}$ y las operaciones $+$ y $\cdot$.

6.2. Ejemplos de espacios vectoriales

EJEMPLO 6.1. Si V es el conjunto de vectores del espacio, $\mathbb{K} = V$ y las operaciones $+$ y $\cdot$ son las definidas en el capítulo 2, obtenemos un espacio vectorial real.

EJEMPLO 6.2. V es el conjunto de ternas ordenadas de números reales $V = \mathbb{R}^3$ y $\mathbb{K} =\mathbb{R}$ y definimos las operaciones $+$ y $\cdot$ como sigue:
suma: $(a_1, a_2, a_3) + (b_1, b_2, b_3) \stackrel{def}{=} (a_1 + b_1, a_2 + b_2, a_3 + b_3)$
producto: $\lambda \cdot (a_1, a_2, a_3) \stackrel{def}{=} (\lambda.a_1, \lambda.a_2, \lambda.a_3)$ para todo $\lambda \in \mathbb{R}$ y todas las ternas de $\mathbb{R}^3$.

Notación: El símbolo $\stackrel{def}{=}$ quiere decir que el miembro de la izquierda esta siendo definido mediante la expresión de la derecha.

EJEMPLO 6.3. En general si V es el conjunto de las n-uplas ordenadas de números reales, $V = \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R} = \mathbb{R}^n$,$\mathbb{K} = \mathbb{R}$ y las operaciones son: $(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) \stackrel{def}{=} (a_1 + b_1, a_2 + b_1, \ldots, a_n + b_1)$ y además $\lambda \cdot (a_1, a_2, \ldots, a_n) \stackrel{def}{=} (\lambda.a_1, \lambda.a_2, \ldots, \lambda.a_n)$. Es fácil ver que $\{\mathbb{R}^n, \mathbb{R}, +, \cdot\}$ es un espacio vectorial sobre $\mathbb{R}$, con $\vec{0} = (0, 0, \cdots, 0)$ y $-(a_1, a_2, \cdots, a_n) = (-a_1, -a_2, \cdots, -a_n)$. Las operaciones definidas en este ejemplo se llaman **usuales** de $\mathbb{R}^n$.

EJEMPLO 6.4. Sea V el conjunto de las sucesiones de números reales, $\mathbb{K} = \mathbb{R}$, y las operaciones $+$ y $\cdot$ definidas como sigue:

Dadas $\{a_n\}$ y $\{b_n\} \in V$, y dado $\lambda \in \mathbb{R}$:

$$\{a_n\} + \{b_n\} \stackrel{def}{=} (a_1 + b_1, a_2 + b_2, ..., a_n + b_n, ...) = \{a_n + b_n\}$$

(la suma de dos sucesiones es definida como la sucesión que se obtiene sumando los términos n-ésimos de las sucesiones dadas),

$$\lambda \cdot \{a_n\} = \lambda \cdot (a_1, a_2, \ldots, a_n) \stackrel{def}{=} (\lambda \cdot a_1, \lambda \cdot a_2, \ldots, \lambda \cdot a_n) = \{\lambda \cdot a_n\}$$

Es fácil comprobar que es un espacio vectorial real . El vector nulo es la sucesión que tiene todos sus términos iguales a cero. Se verifica además $-\{a_n\} = \{-a_n\}$

EJEMPLO 6.5. Sea $\mathcal{P}$ el conjunto de todos los polinomios en una variable con coeficientes reales, sea $\mathbb{K} = \mathbb{R}$, y sean la suma usual de polinomios y el producto usual de un número por un polinomio. $\{\mathcal{P}, \mathbb{R}, +, \cdot\}$ es un $\mathbb{R}$–espacio vectorial.

EJEMPLO 6.6. El conjunto de todas las funciones $f : [a, b] \to \mathbb{R}$, forma un $\mathbb{R}$-espacio vectorial, con las operaciones usuales de suma de funciones y producto de un número real por una función. Es decir:

$f{+}g \stackrel{def}{=} h$ *donde* $h(x) \stackrel{def}{=} f(x){+}g(x)$ $\forall x \in [a, b]$.

$\lambda \cdot f \stackrel{def}{=} k$ *donde* $k\,(x) \stackrel{def}{=} \lambda f\,(x) \quad \forall\, x \in [a, b]\,.$

EJEMPLO 6.7. El conjunto de las n-uplas ordenadas de números complejos $V = \mathbb{C}^n$ con las operaciones definidas como en el ejemplo 6.3 tomando $\mathbb{K} = \mathbb{C}$ forma un $\mathbb{C}$-espacio vectorial $\{\mathbb{C}^n, \mathbb{C}, +, \cdot)\}$.

EJEMPLO 6.8. Si $\mathbb{K} = \mathbb{R}$, y las operaciones en $\mathbb{C}^n$ se definen en forma análoga al ejemplo anterior (y en forma análoga al ejemplo 6.3), resulta $\mathbb{C}^n, \mathbb{R}, +, \cdot\}$ un espacio vectorial real (distinto del ejemplo 6.7 , puesto que difieren en el cuerpo $\mathbb{K}$).

EJEMPLO 6.9. Las sucesiones de números complejos con la suma de sucesiones y el producto de una sucesión por un número complejo definidos en forma similar al ejemplo 6.4 constituyen un $\mathbb{C}$-espacio vectorial.

OBSERVACIÓN 6.5. En todos los ejemplos anteriores, para verificar que efectivamente son espacios vectoriales, hay que verificar que se cumplen las propiedades incluidas en la definición 1.1, para la suma y el producto por escalares.

EJEMPLO 6.10. Sea V el conjunto de sucesiones de números complejos. Si $\mathbb{K} = \mathbb{R}, + : V \times V \to V$ y $\cdot : \mathbb{R} \times V \to V$, operaciones éstas definidas análogamente al ejemplo 6.4 Entonces $\{V, \mathbb{R}, +, \cdot\}$ es un $\mathbb{R}$-espacio vectorial distinto al del ejemplo 6.4 y la del ejemplo 6.9

EJEMPLO 6.11. El conjunto $\{\vec{0}\}$ es un $\mathbb{K}$-espacio vectorial, donde se define $\vec{0}+\vec{0}=\vec{0}$ y $\alpha \cdot \vec{0} = \vec{0}$, $\forall\, \alpha \in \mathbb{K}$.

EJEMPLO 6.12. El conjunto de los polinomios de grado menor o igual a n, en una variable x (este conjunto se indica con $\mathcal{P}_n$), $\mathbb{K} = \mathbb{R}$, y la suma y producto por un real definido como en el ejemplo 6.5, forman un espacio vectorial real.

EJEMPLO 6.13. El conjunto de los polinomios en una variable x, de grado igual a n, con coeficientes reales, $\mathbb{K} = \mathbb{R}$ y la suma y el producto definidos como en el ejemplo 6.5, no forman un espacio vectorial (Estudiar por qué no).

EJEMPLO 6.14. Con las operaciones usuales $\{\mathbb{R}, \mathbb{C}, +, \cdot\}$ no es un espacio vectorial.

OBSERVACIÓN 6.6. Se denota $u - v$ a la suma $u + (-v)$ con $u, v \in V$.

PROPOSICIÓN 6.7. $0 \cdot u = \vec{0}\ ,\ \ \forall u \in V$

DEMOSTRACIÓN. $0 \cdot u = 0 \cdot u + \vec{0} = 0 \cdot u + (u - u) = (0 \cdot u + 1 \cdot u) - u = (0 + 1) \cdot u - u = 1 \cdot u - u = u - u = \vec{0}$ □

PROPOSICIÓN 6.8. $\alpha \cdot \vec{0} = \vec{0}\ ,\ \forall \alpha \in \mathbb{K}$

DEMOSTRACIÓN. Sea u un vector cualquiera de V, $\vec{0} = \alpha \cdot u - \alpha \cdot u =$
$= \alpha \cdot \left(\vec{0} + u\right) - \alpha \cdot u = \left(\alpha \cdot \vec{0} + \alpha \cdot u\right) - \alpha \cdot u = \alpha \cdot \vec{0} + (\alpha \cdot u - \alpha \cdot u) =$
$= \alpha \cdot \vec{0} + \vec{0} = \alpha \cdot \vec{0}$. □

PROPOSICIÓN 6.9. *Sean $\alpha \in \mathbb{K}$, $u \in V$. Entonces , $\alpha \cdot u = \vec{0}$ si y sólo si : $\alpha = 0$ y/o $u = \vec{0}$.*

DEMOSTRACIÓN. Directo: Supongamos que $\alpha \neq 0$. Como $\alpha \cdot u = \vec{0}$, $\frac{1}{\alpha} \cdot (\alpha \cdot u) = \frac{1}{\alpha} \cdot \vec{0}$ de donde $\left(\frac{1}{\alpha} \cdot \alpha\right) \cdot u = \vec{0}$, o sea $1 \cdot u = \vec{0}$ entonces, $u = \vec{0}$.

Recíproco: son las Proposiciones 6.7 y 6.8. □

PROPOSICIÓN 6.10. $(-1)\,.v = -v\ \forall v \in V$

DEMOSTRACIÓN. Basta ver que $(-1)\ v$ es el opuesto de v, es decir

$$v + (-1)\,.v = 1.v + (-1)\,v = (1 + (-1))\,v = 0.v = \vec{0}$$

□

6.3. Subespacios

DEFINICIÓN 6.2. Sea $\{V, \mathbb{K}, +, \cdot\}$ un espacio vectorial y $W \neq \varnothing$ un subconjunto no vacio de V. Diremos que W es un **subespacio** de V, si se cumple:
a) $w_1 + w_2 \in W$ *para todo* $w_1 \in W$ *y* $w_2 \in W$
b) $\lambda w \in W$ *para todo* $w \in W$ *y* $\lambda \in \mathbb{K}$

Es decir, un subespacio de un espacio vectorial es un conjunto no vacío, "cerrado" frente a la suma y al producto de escalares por vectores.

OBSERVACIÓN 6.11. Si W es un subespacio, entonces $\vec{0} \in W$. En efecto como $W \neq \varnothing$ existe $w \in W$ y como W es cerrado para la multiplicación por escalares $\lambda w \in W$, $\forall \lambda \in \mathbb{K}$. En particular $0.w = \vec{0} \in W$. La última igualdad es consecuencia de la Proposición 6.7.

EJEMPLOS DE SUBESPACIOS VECTORIALES

EJEMPLO 6.15. Los primeros ejemplos y los más sencillos, de subespacio vectoriales, son los llamados **subespacios triviales**. Uno de ellos es $W = \left\{\vec{0}\right\}$ (conjunto formado por un único elemento, el vector nulo). El otro subespacio trivial es $W = V$. En ambos casos, W es subespacio de V. ■

EJEMPLO 6.16. Sea π un plano del espacio por el origen.

$$\pi = \{(x, y, z) \in \mathbb{R}^3 \,:\, ax + by + cz = 0\}.$$

Entonces π es un subespacio de $\mathbb{R}^3$ En efecto $\pi \neq \varnothing$, pues $\vec{0} = (0, 0, 0) \in \pi$ y es cerrado para las operaciones. Sean $p_1 = (x_1, y_1, z_1)$ y $p_2 = (x_2, y_2, z_2)$ dos puntos de π y λ un escalar cualquiera. Verifiquemos primero que la suma es cerrada

$$p_1 + p_2 = (x_1, x_2, x_3) + (x_2, y_2, z_2) = (x_1 + x_2,\, y_1 + y_2,\, z_1 + z_2) = q$$

y $q \in \pi$. Para ver que el elemento q pertenece al conjunto π es necesario, dado que éste esta definido por comprensión, verificar que q cumple las condición lógica de pertenencia, esto es sus coordenadas deben verificar la ecuación $ax + by + cz = 0$

$$a(x_1+x_2)+b(y_1+y_2)+c(z_1+z_2) = \overbrace{\left(ax_1 + by_1 + cz_1\right)}^{=0, \text{ pues } p_1 \in \pi} + \overbrace{\left(ax_2 + by_2 + cz_2\right)}^{=0, \text{ pues } p_2 \in \pi} = 0$$

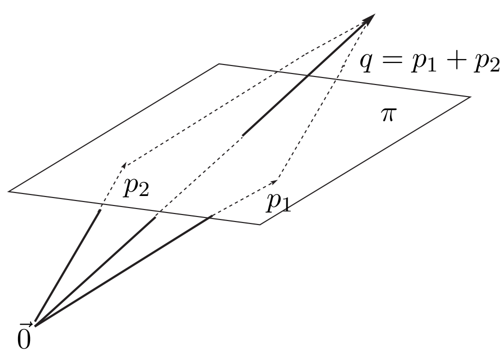
Veamos ahora que π es cerrado para la multiplicación por escalares.

$$\lambda p_1 = \lambda(x_1, y_1, z_1) = (\lambda x_1, \lambda y_1, \lambda z_1) = q'$$

y $q' \in \pi$ pues

$$a(\lambda x_1) + b(\lambda y_2) + c(\lambda z_1) = \lambda \overbrace{\left(ax_1 + by_1 + cz_1\right)}^{=0, \text{ pues } p_1 \in \pi} = 0$$

Si en cambio se considera un plano π que no pase por el origen entonces su ecuación será $ax + by + cz = d$ con $d \neq 0$. Obsérvese que $\pi \neq \varnothing$ pero $\vec{0} \notin \pi$ en consecuencia no es un subespacio. Es además fácil ver que no es cerrado para las operaciones si p_1 y p_2 son dos puntos de π su suma $q = p_1 + p_2$ no pertenece a π tal como se ve en el dibujo. Compruebe el lector esto analíticamente. ■



EJEMPLO 6.17. Sea r un recta por el origen entonces r es un subespacio vectorial de $\mathbb{R}^3$.

$$r = \{(x, y, z) \in \mathbb{R}^3 \, : \, ax + by + cz = 0; \, a'x + b'y + c'z = 0\}$$

Obsérvese que $(x, y, z) \in r$ si y solo si

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{6.42}$$

Sean $p_1 = (x_1, y_1, z_1)$ y $p_2 = (x_2, y_2, z_2)$ dos puntos de r y $\lambda \in \mathbb{R}$ entonces $q = p_1 + p_2$ pertenece a r, para verificar esto veremos que las coordenadas de $q = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ verifican la condición 6.42

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix} = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \left(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \right) =$$

$$= \underbrace{\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}}_{=\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \text{ pues } p_1 \in r} + \underbrace{\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}}_{=\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \text{ pues } p_2 \in r} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

El lector verificara que $\lambda p_1 = q' \in r$.
Es claro que para verificar que $p_1 + p_2$ pertenece a r sólo tuvimos que usar la propiedad distributiva del producto de matrices por lo tanto en general si $A \in \mathcal{M}_{m \times n}$ es una matriz cualquiera y $S = \{X \in \mathbb{R}^n \, : \, A \cdot X = \vec{0}\} \subset \mathbb{R}^n$ es el conjunto solución del sistema homogéneo de matriz A entonces S es un subespacio de $\mathbb{R}^n$.
De hecho veremos más adelante que el recíproco es verdadero. Es decir que si S es un subespacio cualquiera de $\mathbb{R}^n$ existe $A \in \mathcal{M}_{m \times n}$ tal que S es el conjunto solución del sistema homogéneo $AX = \vec{0}$.
Es correcto, entonces afirmar que los subespacios de $\mathbb{R}^n$ son exactamente los conjuntos solución de los sistemas homogéneos de ecuaciones.
Por otra parte, desde un punto de vista más geométrico es natural entonces

pensar esquemáticamente los subespacios de $\mathbb{R}^n$ (y de hecho, como veremos más adelante, los de un espacio vectorial de dimensión finita cualquiera [1] como la generalización de los planos y rectas por el origen de E. ■

EJEMPLO 6.18. En $\{\mathbb{R}^3, \mathbb{R}, +, \cdot\}$ $W{=}\big\{(a_1, a_2, a_3) \in \mathbb{R}^3$ tales que $a_3 = 0\big\}$ entonces W es un subespacio de V.

EJEMPLO 6.19. En $\{\mathbb{R}^n, \mathbb{R}, +, \cdot\}$ sea $W = \{(a_1, ..., a_n) \in \mathbb{R}^n : a_i = 0\}$ con i fijo $0 < i \leqslant n$, es decir W es el conjunto de las n-uplas con la i-ésima componente nula. Entonces W es un subespacio de $\mathbb{R}^n$.

EJEMPLO 6.20. En las sucesiones de números reales (ejemplo 6.4) sea W el conjunto formado por las sucesiones que a partir de algún término, tienen todos los términos restantes iguales a cero.

Verificar que W es un subespacio del espacio vectorial V.

EJEMPLO 6.21. Sea $\mathcal{P}$ el espacio vectorial de los polinomios definido en el ejemplo 6.5. Sea $\mathcal{P}_n$ el subconjunto de los polinomios de grado menor o igual que n. Entonces $\mathcal{P}_n$ es un subespacio de $\mathcal{P}$. Nótese que si consideramos sólo el conjunto de polinomios de grado exactamente n con $n \geq 1$ no constituyen subespacio de $\mathcal{P}$ pues el polinomio nulo tiene grado 0.

EJEMPLO 6.22. Sea V el espacio vectorial del ejemplo 6.6, y sea $C[a, b]$ el conjunto de las funciones f que son continuas. $f : [a, b] \to \mathbb{R}$. Entonces $C[a, b]$ es un subespacio de V. Recordar que la suma de funciones continuas es una función continua, así como también lo es el producto de una función continua por un escalar.

---

[1] Para entender que es un espacio de dimensión finita véase más adelante la definición 6.8)

EJEMPLO 6.23. En el espacio de las n-$u$plas de complejos (ejemplo 6.7) sea W un subconjunto de las n-uplas cuya suma de términos da cero. Se puede verificar que es un subespacio de $\mathbb{C}^n$.

EJEMPLO 6.24. En el espacio del ejemplo 6.7 $\{\mathbb{C}^n, \mathbb{C}, +, \cdot\}$ sea W el subconjunto de las n-$u$plas cuyo primer término es 5+2i . Este subconjunto W no es un subespacio, porque al sumar dos elementos de W, el resultado no pertenece a W.

EJEMPLO 6.25. Sea $\{\mathbb{C}^n, \mathbb{R}, +, \cdot\}$ el espacio vectorial del ejemplo 6.8 Sea W**'** el subconjunto formado por las n-$u$plas de números complejos cuyos términos tienen parte real igual a cero. Es un subespacio. En este ejemplo es importante que el cuerpo $\mathbb{K}$ sea $\mathbb{R}$ y no $\mathbb{C}$. Observar que el mismo subconjunto W**'**, considerado como subconjunto del espacio vectorial $\{\mathbb{C}^n, \mathbb{C}, +, \cdot\}$ no es un subespacio.

EJEMPLO 6.26. En el ejemplo 6.9 de las sucesiones de números complejos sobre el cuerpo $\mathbb{C}$, con las operaciones usuales, sea W el subconjunto de las sucesiones que tienen igual a cero los términos que están en lugar par: $W = \{\{a_n\} : a_n \in \mathbb{C}, a_{2k} = 0\}$. Entonces W es un subespacio.

TEOREMA 6.12. *Si W es un subespacio de $\{V, \mathbb{K}, +, \cdot\}$ entonces $\{W, \mathbb{K}, +, \cdot\}$ con las operaciones + y · "heredadas"de V, es un espacio vectorial sobre el mismo cuerpo $\mathbb{K}$.*

Demostración. Hay que probar que si $\{V, \mathbb{K}, +, \cdot\}$ es espacio vectorial y $W \subset V$ es subespacio de V entonces $\{W, \mathbb{K}, +, \cdot\}$ también es un espacio vectorial es decir se cumplen todas las condiciones de la definición 6.1. Las operaciones + y · en W son las mismas que en V sólo que restringidas a W (esto es definidas sólo para los elementos de W).
Como W es cerrado para las operaciones estas quedan bien definidas es decir

$+ : W \times W \rightarrow W$ y $\cdot : \mathbb{K} \times W \rightarrow W$. Por otra parte las propiedades asociativa y conmutativa de la suma y todas las del producto se verifican en W pues se verificaban para cualesquiera vectores de V y $W \subset V$.
Además en la observación 6.11 vimos que $\vec{0} \in W$ y como $\vec{0} + v = v + \vec{0} = v$, $\forall v \in V$ entonces también será esto cierto $\forall v \in W$ pues W es un subconjunto de V. Por lo tanto la suma en W tiene neutro. Resta entonces sólo verificar la existencia del opuesto. Naturalmente que si $w \in W \subset V$ es un elemento cualquiera de W entonces también lo es de V y por lo tanto existe $-w \in V$, pero a priori podría ocurrir que $-w \notin W$. Veremos que esto no es posible si W es no vacío y cerrado para las operaciones, en efecto sea $w \in W$, como W es cerrado para las operaciones se tiene que $\lambda w \in W$, $\forall \lambda \in \mathbb{K}$, en particular si $\lambda = -1$. Por lo tanto $(-1)w = -w \in W$. En ésta última igualdad se uso la proposición 6.10. Esto prueba la existencia del opuesto en W y concluye la prueba. $\square$

OBSERVACIÓN 6.13. Este teorema permite ver en forma fácil si un cierto conjunto W sobre un cuerpo $\mathbb{K}$, con ciertas operaciones $+$ y $\cdot$ es un espacio vectorial. En efecto, en lugar de verificar todas las condiciones de la definición 6.1, alcanza ver si W es un subespacio de algún V, que ya se sepa tiene estructura de espacio vectorial. Para lo cual alcanza con probar que W es cerrado con respecto a la suma y cerrado con respecto al producto de un escalar por un vector.

TEOREMA 6.14. *Sean $W_1, W_2, ..., W_n$, subespacios de un mismo espacio V. Entonces $W = \bigcap_{i=1}^{n} W_i$ es también un subespacio de V.*

DEMOSTRACIÓN. En primer lugar $W \neq \varnothing$; en efecto: el vector $\vec{0}$ pertenece a todos los subespacios W_i, y por lo tanto pertenece también a su intersección W.

Además $W \subset V$ puesto que $W \subset W_i \subset V \; \forall \, i = 1, 2, ..., n$

Falta probar que W es cerrado frente a la suma de vectores y al producto de un escalar por un vector. Sean w y $w' \in W$, entonces w y $w' \in W_i$ para cada $i = 1, 2, ..., n$. Luego $w + w' \in W_i \, \forall \, i = 1, 2, ..., n$ (puesto que W_i son subespacios). Entonces $w + w' \in W$.

Ahora sean $\alpha \in \mathbb{K}$ y $w \in W$. Se tiene $w \in W_i \, \forall \, i = 1, 2, ..., n$. De allí $\alpha \cdot w \in W_i \, \forall \, i = 1, 2, ..., n$ (puesto que W_i son subespacios). Entonces $\alpha \cdot w$ pertenece a la intersección de todos los W_i, o sea a W. $\square$

6.4. Subespacio generado e independencia lineal

En esta Sección daremos definiciones de combinación lineal, dependencia e independencia lineal de vectores en espacios vectoriales cualesquiera, que generalizan los conceptos vistos en el Capítulo 2

DEFINICIÓN 6.3. Sea V un espacio vectorial, $v_1, v_2, ..., v_n$ vectores de V. Decimos que v es **combinación lineal (c.l.)** de $v_1, v_2, ..., v_n$ cuando existen escalares $\lambda_1, \lambda_2, ..., \lambda_n$ tales que : $v = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_n v_n$

Si A es un subconjunto no vacío de V, se indica con $[A]$ al conjunto de todas las combinaciones lineales de elementos de A, es decir:

$$[A] = \left\{ v \in V \text{ tales que} : v = \lambda_1 v_1 + ... + \lambda_p v_p \text{ con } \lambda_i \in \mathbb{K}, v_i \in A \ \forall i \right\}.$$

TEOREMA 6.15. *Sean V un espacio vectorial y $A \neq \varnothing$, $A \subset V$. Entonces $[A]$ es un subespacio de V.*

Demostración. $[A]$ es distinto de vacío porque dado algún vector v de A, se tiene $\vec{0} = 0 \cdot v \Rightarrow \vec{0}$ es combinación lineal de $v \Rightarrow \vec{0} \in [A]$. Sean ahora $v,w \in [A]$. Esto quiere decir que existen vectores $v_1, v_2, ..., v_n$; $w_1, w_2, ..., w_m$ pertenecientes a A (no necesariamente distintos) y existen $\alpha_1, \alpha_2, ..., \alpha_n$; $\beta_1, \beta_2, ..., \beta_m$ tales que $v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$ y $w = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_m w_m$. Entonces $v + w = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n + \beta_1 w_1 + \beta_2 w_2 + ... + \beta_m w_m$. O sea, $v + w$ es combinación lineal de elementos de A. Hasta ahora hemos probado que $[A]$ es cerrado respecto a la suma.

Sea ahora $\lambda \in \mathbb{K}$ y sea $v \in [A]$. Existen escalares $\alpha_1, \alpha_2, ..., \alpha_n$ y vectores $v_1, v_2, ..., v_n \in A$, tales que: $v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$

Entonces: $\lambda \cdot v = \lambda \cdot (\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) = (\lambda \cdot \alpha_1) v_1 + ... + (\lambda \cdot \alpha_n) v_n$ y obtenemos que λv también es combinación lineal de vectores de A. O sea, $[A]$ es cerrado respecto al producto de un escalar por un vector. Luego, $[A]$ es un subespacio. □

DEFINICIÓN 6.4 (Subespacio generado). Sea V un espacio vectorial, A un subconjunto de V. Decimos que $[A]$ es el **subespacio generado** por A.

DEFINICIÓN 6.5 (Conjunto generador). Sea V un espacio vectorial y $S \subset V$ un subespacio decimos que $A \subset S$ es un **generador** de S si $[A] = S$. Es decir si cualquier elemento de S es combinación lineal de A. (ver definición 6.3).

Notación: Si A es un generador de S pondremos $A \xrightarrow{g} S$

EJEMPLO 6.27. Sea $V = \mathbb{R}^3$ y $A = \{(1,0,0),(0,1,0)\}$ ¿$A \xrightarrow{g} V$?.
Sea $v = (a,b,c) \in \mathbb{R}^3$ un vector cualquiera de $\mathbb{R}^3$ para que A genere a $\mathbb{R}^3$ debemos poder escribir v como combinación lineal de A cuales quiere sean a,b y c. Sea α_1 y α_2 reales tales que

$$(a,b,c) = \alpha_1(1,0,0) + \alpha_2(0,1,0) \iff \begin{cases} \alpha_1 = a \\ \alpha_2 = b \\ 0 = c \end{cases}$$

Es claro que el sistema anterior es compatible si y solo si $c = 0$ entonces A no genera $\mathbb{R}^3$. ■

EJEMPLO 6.28. Sea $V = \mathbb{R}^3$ y $S = \{(a,b,c) \in \mathbb{R}^3 : c = 0\}$ determinar un generador de S.
Como $v \in \mathbb{R}^3$ si y solo si existen a y b reales tal

$$v = (a,b,0) = (a,0,0) + (0,b,0) = a(1,0,0) + b(0,1,0)$$

entonces $A = \{(1,0,0),(0,1,0)\} \xrightarrow{g} S$. ■

EJEMPLO 6.29. Sea $V = \mathcal{M}_{2\times 2}$ y $S = \{A \in \mathcal{M}_{2\times 2} \,:\, A = A^t\} \subset \mathcal{M}_{2\times 2}$ el subespacio (verificarlo) de las **matrices simétricas** 2×2. Determinemos un generador. Para eso observemos que $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in S$ si y solo si $b = c$. Entonces $A \in S$ si y solo si existen a, b, d reales tales que

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} =$$

$$a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Entonces $G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \xrightarrow{g} S$ ■

EJEMPLO 6.30. Sea $V = \mathcal{P}_2$ y

$$S = \{p \in \mathcal{P}_2 \,:\, p \text{ tiene termino independiente nulo}\}$$

Entonces, sea p tal que $p(x) = ax^2 + bx + c$, $p \in S$ si y solo si $c = 0$ es decir $p \in S$ si y solo si existen a, b reales tales que $p(x) = ax^2 + bx$ pero entonces si ponemos $p_i \in \mathcal{P}_2$ tal que $p_i(x) = x^i$, con $i = 0, 1, 2$ tenemos que $p \in S$ si y solo si existen a y b reales tales que $p = ap_2 + bp_1$. Entonces $G = \{p_2, p_1\} \xrightarrow{g} S$ ■

PROPOSICIÓN 6.16. *Sea V un subespacio vectorial y A un subconjunto finito de V, o sea $A = \{v_1, v_2, ..., v_n\} \subset V$. Si v_k es combinación lineal de los demás vectores de A, entonces* $[A] = \left[v_1, v_2, ...v_{k-1}, v_{k+1}, ..., v_n\right]$

DEMOSTRACIÓN. Toda combinación lineal de vectores de

$$\{v_1, v_2, ...v_{k-1}, v_{k+1}, ..., v_n\}$$

es también una c.l. de vectores de A (basta agregar el vector v_k multiplicado por el escalar 0), es decir $[v_1, v_2, ...v_{k-1}, v_{k+1}, ..., v_n] \subset [A]$.

Ahora hay que probar al revés: que toda c.l. de vectores de A es c.l. de vectores de $\{v_1, v_2, ...v_{k-1}, v_{k+1}, ..., v_n\}$. Se sabe que

$$v_k = \alpha_1 v_1 + \alpha_2 v_2 + ...\alpha_{k-1} v_{k-1} + \alpha_{k+1} v_{k+1} + ... + \alpha_n v_n. \tag{6.43}$$

Sea $u \in [A]$ entonces u es c.l. de A, tendremos:

$$u = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1} + \lambda_k v_k + \lambda_{k+1} v_{k+1} + \cdots + \lambda_n v_n.$$

Sustituyendo v_k por la expresión (6.43), resulta:

$$u = (\lambda_1 + \lambda_k \alpha_1) v_1 + \cdots + (\lambda_{k-1} + \lambda_k \alpha_{k-1}) v_{k-1} + \\ + (\lambda_{k+1} + \lambda_k \alpha_{k+1}) v_{k+1} + \cdots + (\lambda_n + \lambda_k \alpha_n) v_n.$$

O sea, u es c.l. de $\{v_1, v_2, ...v_{k-1}, v_{k+1}, ..., v_n\}$ como queríamos probar. $\square$

DEFINICIÓN 6.6 (Independencia y dependencia lineal). Sea V un espacio vectorial, $A = \{v_1, v_2, ..., v_n\} \subset V$. Decimos que el conjunto es **linealmente independiente (L.I.)** si y sólo si dados escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que:

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = \vec{0}$$

implica $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Decimos que el conjunto A es **linealmente dependiente (L.D.)** cuando no es L.I..

Es decir, un conjunto es L.I. cuando la única combinación lineal de sus elementos que da el nulo es la **trivial**, esto es la que tiene todos los coeficientes nulos.

Por el contrario es L.D. si existe alguna combinación no trivial (algún coeficiente distinto de cero) de sus elementos que de el nulo.

El lector debe remitirse a la sección 2.5 del capítulo 2 por ejemplos en $\mathbb{R}^n$, no obstante esto veremos aquí algunos ejemplos nuevos.

EJEMPLO 6.31. Sea V un espacio vectorial, $v \neq \vec{0}$ un vector de V. Entonces $\{v\}$ es L.I..

EJEMPLO 6.32. Sea V un espacio vectorial, $\left\{\vec{0}\right\}$ es L.D.

EJEMPLO 6.33. Sean $v_1, v_2, \ldots, v_n \in V$ (V espacio vectorial), entonces $\left\{\vec{0}, v_1, v_2, ..., v_n\right\}$ es l.d ■

EJEMPLO 6.34. Sea $V = \mathcal{M}_{2\times 2}$ y

$$G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

entonces G es L.I. en efecto sean $\lambda_1, \lambda_2, \lambda_3$ reales tales que

$$\lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Longleftrightarrow$$

$$\Longleftrightarrow \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Longleftrightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 0 \\ \lambda_3 = 0 \end{cases}$$

entonces G es L.I. ■

EJEMPLO 6.35. Sea $A = \{q_1, q_2, q_3, q_4\} \subset \mathcal{P}_2$ donde $q_1(x) = x+1$, $q_2(x) = x^2 + x - 1$, $q_3(x) = x^2 + 2x$ y $q_4(x) = x^2 - 2$, investiguemos la dependencia lineal de A.

Sean $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ reales tales que

$$\lambda_1 q_1 + \lambda_2 q_2 + \lambda_3 q_3 + \lambda_4 q_4 = \vec{0} \Longleftrightarrow \tag{6.44}$$

$$\lambda_1 q_1(x) + \lambda_2 q_2(x) + \lambda_3 q_3(x) + \lambda_4 q_4(x) = \vec{0}, \quad \forall x \in \mathbb{R} \Longleftrightarrow$$

$$\lambda_1(x+1) + \lambda_2(x^2 + x - 1) + \lambda_3(x^2 + 2x) + \lambda_4(x^2 - 2) = 0, \quad \forall x \in \mathbb{R} \Longleftrightarrow$$

$$(\lambda_2 + \lambda_3 + \lambda_4)x^2 + (\lambda_1 + \lambda_2 + 2\lambda_3)x + (\lambda_1 - \lambda_2 - 2\lambda_4) = 0, \quad \forall x \in \mathbb{R}$$

de donde utilizando el teorema de identidad de polinomios[2] se deduce que (6.44) se cumple si y solo si λ_1,λ_2, λ_3 y λ_4 verifican el sistema homogéneo:

$$(S)\begin{cases} \lambda_2+\lambda_3+\lambda_4=0 \\ \lambda_1+\lambda_2+2\lambda_3=0 \\ \lambda_1-\lambda_2-2\lambda_4=0 \end{cases}.$$

Escalerizando y resolviendo (S) se deduce que es un sistema compatible indeterminado y que

$$Sol(S)=\left\{(\lambda_1,\lambda_2,\lambda_3,\lambda_4)\in\mathbb{R}^4 \,:\, \lambda_1=\lambda_4-\lambda_3, \lambda_2=-(\lambda_3+\lambda_4), \lambda_3,\lambda_4\in\mathbb{R}\right\}$$

Por lo tanto (S) admite soluciones no triviales y consecuentemente también (6.44) por lo cual G es L.D.
Observemos también que sustituyendo la solución obtenida en (6.44) se tiene que $\forall\,\lambda_3$ y $\lambda_4\in\mathbb{R}$

$$(\lambda_4-\lambda_3)q_1-(\lambda_3+\lambda_4)q_2+\lambda_3q_3+\lambda_4q_4=\vec{0}, \implies$$

$$\text{poniendo } \lambda_3=0 \text{ y } \lambda_4=1 \text{ tenemos, } q_4=-q_1+q_2$$

Esto es un vector de A resultó combinación lineal de los restantes. ■

OBSERVACIÓN 6.17. Las definiciones L.I. e L.D. se dieron para conjuntos finitos, pero son aplicables también a conjuntos con una cantidad infinita de vectores en un espacio vectorial. Si A es infinito., $A\subset V$. V espacio vectorial, decimos que A es L.I. cuando cualquier subconjunto finito de A es L.I. Decimos que A es L.D. cuando no es L.I. (es decir, cuando algún subconjunto finito de A es L.D.).

EJEMPLO 6.36. Sea $V=\mathcal{P}$ el espacio vectorial de todos los polinomios de una variable x con coeficientes reales y sea $A=\{p_i\in\mathcal{P} \,:\, i\in\mathbb{N}\}$, donde $p_i(x)=x^i$. Entonces A es un conjunto L.I. con infinitos elementos, pues cualquier subconjunto $B\subset A$ finito que consideremos contiene una

[2]Recordamos al lector el enunciado de este resultado: dos polinomios de coeficientes reales o complejos son iguales si y solo si los coeficientes de las potencias respectivas son iguales

cantidad finita de polinomios de diferentes grados lo cual implica que B es L.I. (verificarlo). ■

Veremos ahora un resultado que generaliza la proposición 2.14 del capítulo 2 y lo visto en el ejemplo 6.35; un conjunto de más de un vector es L.D. si y solo si uno de los vectores es combinación lineal de los demás. Más aún, probaremos que la dependencia lineal se da sólo cuando un vector del conjunto se puede despejar en función (lineal) de los anteriores.

PROPOSICIÓN 6.18. *Sea V un espacio vectorial, y A un conjunto ordenado contenido en V, que tiene más de un elemento, tal que v_1 no es el vector nulo: $A = \{v_1, v_2, ..., v_n\} \subset V$ con $n > 1,\ v_1 \neq \vec{0}$.*
Entonces: A es L.D. si y sólo si existe $v_k \in A$ $(1 \leqslant k \leqslant n)$ tal que v_k es combinación lineal de los k-1 vectores anteriores $\left\{v_1, ..., v_{k-1}\right\}$.

Demostración. $(\Leftarrow)$

Sabemos que existe $v_k \in A$, $1 \leqslant k \leqslant n$, tal que v_k es combinación lineal de los anteriores:

$$v_k = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_{k-1} v_{k-1}$$

o sea

$$\vec{0} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1} + (-1) v_k + 0 \cdot v_{k+1} + \cdots + 0 \cdot v_n.$$

Tenemos así una combinación lineal de los vectores de A, con no todos los coeficientes iguales a cero (porque el coeficiente de v_k es -1), que da el vector nulo. Por definición A es L.D.

$(\Rightarrow)$

Ahora, sabiendo que A es L.D. y que $v_1 \neq \vec{0}$, hay que probar que un v_k es combinación de los anteriores vectores de A.

Por ser L.D. existen $\lambda_1, \lambda_2, \ldots, \lambda_n$ escalares no todos nulos tales:

$$\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_n v_n = \vec{0}.$$

Sea k el mayor i tal que $\lambda_i \neq 0 \Rightarrow \lambda_i = 0\, \forall i > k$ y $\lambda_k \neq 0$. Resulta $k > 1$, porque sino sería $\lambda_1 v_1 = \vec{0},\ \lambda_1 \neq 0 \ \Rightarrow\ v_1 = \vec{0}$ (contrario a la hipótesis).

Entonces $\lambda_1 v_1 + \cdots + \lambda_k v_k = \vec{0}$ y despejando, $\lambda_k v_k = -\lambda_1 v_1 - \cdots - \lambda_{k-1} v_{k-1}$ y como $\lambda_k \neq 0$ se deduce que:

$$v_k = -\frac{\lambda_1}{\lambda_k} v_1 - ... - \frac{\lambda_{k-1}}{\lambda_k} v_{k-1}.$$

Entonces v_k es combinación lineal de $v_1, ..., v_{k-1}$ como se quería probar. □

OBSERVACIÓN 6.19. De lo anterior se deduce en particular que un conjunto es L.D. si y sólo si un vector es combinación lineal de los demás.

COROLARIO 6.20. *Sea V un espacio vectorial y $A = \{v_1, v_2, \ldots, v_n\}$ un subconjunto cualquiera entonces A es L.I. si y solo si ningún vector de A es combinación lineal de los restantes.*

6.5. Base de un espacio vectorial y dimensión

La idea central de la Geometría Analítica es que, fijado un sistema de coordenadas a cada punto del espacio le corresponde una única terna ordenada de números reales. Esto correspondencia permite tratar de manera analítica los objetos geométricos del espacio (i.e. asociar a cada uno de ellos una ecuación) tal como vimos en el capítulo 4. En la construcción que allí hicimos, no solo asociábamos coordenadas a los puntos del espacio E sino que también asociábamos coordenadas a los vectores del espacio V. Veamos como es posible generalizar esto a un espacio vectorial cualquiera.

DEFINICIÓN 6.7 (Base de un espacio vectorial). Sea V un espacio vectorial y A= $\{v_1, v_2, ..., v_n\} \subset V$ (A finito). Decimos que A es **base** de V si y sólo si

a) A es un conjunto L.I.

b) A genera a V, es decir $[A] = V$.

Notación: Si V es un espacio vectorial y B es una base de V pondremos $B \xrightarrow{b} V$.

OBSERVACIÓN 6.21. Sea $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $S \subset V$ un subespacio de V. Ya hemos visto que $(S, \mathbb{K}, +, \cdot)$ es él mismo un espacio vectorial, donde $+$ y $\cdot$ son las operaciones que hereda de V . Una **base del subespacio** S es una base del espacio vectorial $(S, \mathbb{K}, +, \cdot)$.

PROPOSICIÓN 6.22. *Sea V un espacio vectorial y $A = \{v_1, v_2, \ldots, v_n\}$ un subconjunto cualquiera, entonces A es base de V si y solo si, todo vector del espacio se escribe en forma única como c.l. de los vectores de A (es decir, los coeficientes de la c.l. son únicos).*

DEMOSTRACIÓN. $(\Rightarrow)$

En primer lugar, como A genera a V, entonces cualquier vector de V se puede escribir como c.l. de vectores de A. Veremos ahora que esa combinación lineal es única.

En efecto, supongamos que pueda escribirse un vector v como combinación lineal de A, en dos formas distintas:

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \quad \text{y} \quad v = \beta_1 v_1 + \cdots + \beta_n v_n$$

Restando ambas igualdades tenemos:

$$\vec{0} = (\alpha_1 - \beta_1)\, v_1 + \cdots + (\alpha_n - \beta_n)\, v_n$$

La anterior es entonces una forma de escribir el nulo como combinación lineal de A. Como por hipótesis A es base, entonces en particular A es L.I.. Por definición de conjunto L.I. los coeficientes de la combinación lineal anterior deben ser todos ceros y por lo tanto:

$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$$

Consecuentemente las dos formas de escribir v como combinación lineal de A, son la misma.

($\Leftarrow$) Supongamos ahora que cada vector de V se escribe de manera única como combinación lineal de A, entonces en particular $A \xrightarrow{g} V$, para verifica que A es base sólo resta ver que A es L.I.. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ escalares tales que

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = \vec{0}$$

Hemos escrito el vector nulo como combinación lineal de A Por otra parte también se tienen que

$$0v_1 + 0v_2 + \cdots + 0v_n = \vec{0}.$$

Pero como cada vector (incluyendo al nulo) sólo se puede escribir de única manera como combinación lineal de A debe ser $\lambda_i = 0$. $\forall i = 1, 2, \ldots, n$ y consecuentemente A es L.I. $\square$

Sea V un espacio vectorial de dimensión finita cualquiera sobre el cuerpo $\mathbb{K}$ y $B = \{v_1, v_2, \ldots, v_n\}$ una base de V. Dado un vector $v \in V$ cualquiera la proposición anterior establece que existen únicos escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

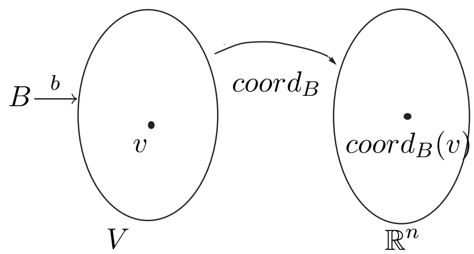
llamaremos a estos números las **coordenadas en la base** B del vector v. De hecho con un poco más de formalidad podemos definir la función $coord_B : V \longrightarrow \mathbb{K}^n$, tal que que a cada vector v le hace corresponder los coeficientes de la combinación lineal de B que da v. Esto es

$$coord_B(v) \stackrel{def}{=} (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

Obsérvese que el vector v es un elemento del espacio V pero sus coordenadas sin importar quien sea V son siempre un n-upla de escalares.

EJEMPLO 6.37. Sea $V = (\mathbb{R}^n, \mathbb{R}, +, \cdot)$ (o $(\mathbb{C}^n, \mathbb{C}, + \cdot)$) el conjunto

$$\mathcal{C} = \{(1, 0, 0, ..., 0), (0, 1, 0, ..., 0), (0, 0, 1, ..., 0), ..., (0, 0, 0, ..., 1)\}$$



es una base de V pues es L.I. y genera V (comprobarlo). El conjunto $\mathcal{C}$ se denomina **base canónica** de $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$ o de $(\mathbb{C}^n, \mathbb{C}, +, \cdot)$.
En general el i-esimo vector de la base canónica lo notaremos

$$e_i = \quad (0, \ldots, 0, \overset{\substack{\text{i-esimo} \\ \text{lugar} \\ \downarrow}}{1}, 0, \ldots, 0)$$

Sea $v = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ un vector cualquiera entonces

$$v = (x_1, x_2, \ldots, x_n) = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

entonces

$$coord_{\mathcal{C}}(v) = (x_1, x_2, \ldots, x_n).$$

Esto es el vector y sus coordenadas en la base canónica coinciden. De hecho lo que ocurre es que la función $coord_{\mathcal{C}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es la función identidad. En general, como veremos en el próximo ejemplo, un vector de $\mathbb{R}^n$ y sus coordenadas en una base arbitraria no coinciden esto sólo ocurre si la base considerada es la canónica. ■

EJEMPLO 6.38. Sea $V = \mathbb{R}^3$ y $B = (1, 1, 0), (0, 1, 0), (0, 0, 1)$ es fácil comprobar y queda a cargo del lector que B es L.I. verifiquemos que también genera $\mathbb{R}^3$. En efecto sea $(a,b,c) \in R^3$ un vector arbitrario comprobemos que existen $\alpha_1, \alpha_2, \alpha_3$ escalares tales que

$$(a,b,c) = \alpha_1(1,1,0) + \alpha_2(0,1,0) + \alpha_3(0,0,1) \Longleftrightarrow$$

$$(a,b,c) = (\alpha_1, \alpha_1 + \alpha_2, \alpha_3) \Longleftrightarrow$$

$$\begin{cases} \alpha_1 = a \\ \alpha_1 + \alpha_2 = b \\ \alpha_3 = c \end{cases}$$

Y este último sistema es compatible determinado y tiene solución $\alpha_1 = a$, $\alpha_2 = b - a$ y $\alpha_3 = c$ de donde $B \xrightarrow{g} \mathbb{R}^3$ y por lo tanto es base. Además

$$(a,b,c) = a(1,1,0) + (b-a)(0,1,0) + c(0,1,0)$$

por lo tanto

$$coord_B\big((a,b,c)\big) = (a, b-a, c)$$

por ejemplo si $v = (1,2,1)$ entonces $coord_B\big((1,2,1)\big) = (1,1,1)$.
En el otro sentido, supongamos que $w \in \mathbb{R}^3$ es un vector tal que sus coordenadas en la base B son $coord_B(w) = (1,2,1)$ entonces el vector

$$w = 1(1,1,0) + 2(0,1,0) + 1(0,0,1) = (1,3,1).$$

■

Es claro que si $V = \mathbb{R}^n$ tanto los vectores como sus coordenadas son n-uplas por lo cual cabe la posibilidad de confundirlos lo cual naturalmente constituye un grave error.
Hay que aclarar en cada caso, cuando se da una n-upla de $\mathbb{R}^n$, si se trata de los vectores en si, o si se trata de las componentes en alguna base distinta de la canónica. Sólo puede omitirse esta aclaración cuando la base es la canónica. Cuando decimos por ejemplo sea v el vector de $\mathbb{R}^3$ tal que $v = (1,2,3)$ estamos indicando la 3-upla $(1,2,3)$ por lo que no tiene sentido preguntar en que base lo estamos expresando.

EJEMPLO 6.39. Sea $\{\mathbb{C}^n, \mathbb{R}, +, \cdot\}$. Verifíquese que
$\{(1,0,...,0), (0,1,...,0), ..., (0,0,...,1)\}$ **no** es base.

Sugerencia: El vector $(i, 0, ..., 0)$ donde i es la unidad imaginaria no es c.l. del conjunto anterior. Verifíque que:

$$\mathcal{C}' = \{(1, 0, \dots, 0), (i, 0, \dots, 0), (0, 1, \dots, 0), (0, i, \dots, 0), \dots$$
$$\dots, (0, 0, \dots, 1), (0, 0, \dots, i)\}$$

es base de $\{\mathbb{C}^n, \mathbb{R}, +, \cdot\}$.

EJEMPLO 6.40. Sea $V = \mathcal{M}_{2\times2}$ y

$$A = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

entonces $A \xrightarrow{b} \mathcal{M}_{2\times2}$, pues es un conjunto L.I. (verificarlo) y además si $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es una matriz 2×2 cualquiera se tiene que

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

de donde $A \xrightarrow{g} \mathcal{M}_{2\times2}$ y por lo tanto es base de $\mathcal{M}_{2\times2}$. Además $coord_A(M) = (a, b, c, d)$. Obsérvese que aquí no hay posibilidad de confundir los vectores con sus coordenadas pues los primeros son matrices y los segundos 4-uplas de números.

EJEMPLO 6.41. Sea $V = \mathcal{P}_n$ y $A = \{p_0, p_1, \dots, p_n\}$, donde $p_i(x) = x^i$, con $i = 0, \dots, n$. Entonces A es L.I. (verificarlo) y además si $p \in \mathcal{P}_n$ es un polinomio cualquiera de grado menor o igual a n tal que $p(x) = a_n x^n + \cdots + a_1 x + a_o$ entonces

$$p = a_0 p_0 + a_1 p_1 + \cdots + a_n p_n$$

por lo tanto $A \xrightarrow{g} \mathcal{P}_n$ y en consecuencia A es base de $\mathcal{P}_n$. Además se tiene que $coord_A(p) = (a_0, a_1, \dots, a_n)$. Nuevamente aquí no hay posibilidad de confundirse coordenadas y vectores las primeras son $(n + 1)$-uplas y los segundos son polinomios de grado menor o igual a n. ■

EJEMPLO 6.42. Sea $S = \{(x, y, z) : 2x - y + z = 0\}$ deseamos determinar una base de S. Para esto comencemos hallando un generador. Observemos que $v \in S$ si y solo si existen x y y reales tales que

$$v = (x, y, y - 2x) = (x, 0, -2x) + (0, y, y) = x(1, 0, -2) + y(0, 1, 1)$$

entonces $A = \{(1, 0, -2), (0, 1, 1)\} \subset S \xrightarrow{g} S$. Además es fácil de ver (hacerlo) que A es L.I. y por lo tanto es un base de S. ■

EJEMPLO 6.43. Sea $V = \mathcal{M}_{2\times 2}$ y $S = \{A \in \mathcal{M}_{2\times 2} : A = A^t\}$ hemos visto en el ejercicio 6.29 que el conjunto

$$G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \xrightarrow{g} S$$

y de lo visto en el ejercicio 6.34 se sabe que A es L.I. por lo tanto $G \xrightarrow{b} S$.
■

TEOREMA 6.23. *Sea V un espacio vectorial. Sea $A = \{v_1, \ldots, v_n\}$ un conjunto que genera a V y $B = \{w_1, \ldots, w_m\}$ un conjunto L.I. de V. Entonces $m \leqslant n$.*

Demostración. La idea de la pruebe consiste en construir un algoritmo que permita sustituir uno a uno los vectores de A con los de B, en cada paso al sustituir un vector de A por uno de B se obtiene un nuevo conjunto generador del espacio V. Como el proceso de sustitución no culmina mientras queden elementos en el conjunto B se prueba que en B hay por lo menos tantos elementos como en A.

En el primer paso del algoritmo elegiremos un vector conveniente de A y lo sustituiremos el vector w_m. Para esto comencemos observando que, $A \xrightarrow{g} V$ y por lo tanto $w_m \in B \subset V$ es combinación lineal de A. En consecuencia el conjunto

$$A_1' = \{w_m, v_1, \ldots, v_n\}$$

es L.D.. Ahora bien, $w_m \neq \vec{0}$, pues $w_m \in B$ y B es L.I. por lo tanto la proposición 6.18 asegura que existe en A_1' un vector (que no es el primero) que resulta combinación de los anteriores. Designemos por $v_i \in A$ a ese vector, que será el primero a sustituir. Pongamos, ahora

$$A_1 = A_1' - \{v_i\} = \{w_m, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}.$$

Entonces $A_1 \xrightarrow{g} V$, pues, en virtud de la proposición 6.16 $[A_1] = [A_1'] = V$. En definitiva al cabo del primer paso se ha obtenido un nuevo conjunto generador de V substituyendo v_i por w_m.
Para el segundo paso razonamos en forma análoga eligiendo otro vector de A para ser sustituido por w_{m-1}: como $A_1 \xrightarrow{g} V$, $w_{m-1} \in B$ es combinación lineal de A_1 por lo tanto el conjunto

$$A_2' = A_1 \cup \{w_{m-1}\} = \{w_{m-1}, w_m, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$$

resulta L.D.. De nuevo observamos que $w_{m-1} \neq \vec{0}$, pues $w_{m-1} \in B$ y B es L.I., por lo tanto la proposición 6.18 asegura que existe un vector en A_2' (que no es el primero) que resulta combinación lineal de los restantes, este vector no puede ser w_m pues si lo fuera debería ser combinación de w_{m-1} y esto contradice la independencia lineal de B. Designemos por $v_k \in A$ a dicho vector (será el segundo vector a sustituir). La proposición 6.16 implica nuevamente que el conjunto

$$A_2 = A_2' - \{v_k\} = \{w_{m-1}, w_m, v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$$

es un generador de V. Al cabo de dos pasos hemos sustituido dos vectores de A por dos vectores de B y lo hemos hecho de modo que el conjunto resultante A_2 es un generador de V, podemos continuar este procedimiento de sustitución y el mismo sólo concluirá cuando se terminen los vectores de B. Cuando esto ocurra o bien los vectores de A se han acabado en simultáneo con los de B y entonces $m = n$ o bien aún restan vectores de A sin sustituir y por lo tanto $m < n$ lo cual concluye la prueba. □

COROLARIO 6.24. *Sea V un espacio vectorial. Si una base tiene n elementos, entonces todas las demás bases tienen n elementos.*

Demostración. Sean A y B bases de V, A con n elementos y B con p elementos. Como A es L.I. y B es generador, aplicando el teorema 6.23 tenemos $n \leqslant p$. Por otra parte, como B es L.I. y A es generador, aplicando nuevamente el teorema 6.23, tenemos $p \leqslant n$. O sea $n = p$. $\square$

DEFINICIÓN 6.8 (Dimensión de un espacio vectorial). Sea V un espacio vectorial. Si existe una base de V que tenga n elementos, se dice que n es la **dimensión** de V (obsérvese en virtud del corolario anterior, n no depende de la base que se halla encontrado). Notaremos $dim\ (V) = n$.

Si $V = \left\{\vec{0}\right\}$ (ejemplo 6.11), se conviene en decir que V tiene dimensión 0. Si $V \neq \left\{\vec{0}\right\}$ y no existe una base de Vcon una cantidad finita de elementos, entonces decimos que la dimensión de V es infinita. Obsérvese que si un espacio vectorial V contiene un subconjunto L L.I. con infinitos elementos entonces $dim(V) = \infty$. En efecto, V no puede tener una base finita pues si $A \subset V$ es un subconjunto finito y $A \xrightarrow{b} V$ entonces en particular $A \xrightarrow{g} V$ pero como L es L.I. y pose infinitos elementos podemos encontrar un subconjunto finito $L' \subset L$ L.I. con $card(L') > card(A)$ lo cual contradice el teorema 6.23.

OBSERVACIÓN 6.25. Naturalmente si $S \subset V$ es un subespacio del espacio $(V, \mathbb{K}, +, \cdot)$ entonces su dimensión es la dimensión de $(S, \mathbb{K}, +, \cdot)$. Esto es el número de elementos de una base de S (ver observación 6.21).

EJEMPLO 6.44. En el ejemplo 6.37 vimos que $\mathcal{C}$ es una base (la base canónica) de $\mathbb{R}^n$ por lo tanto $dim(\mathbb{R}^n) = n$. De manera análoga $\mathcal{C} \xrightarrow{b} (\mathbb{C}^n, \mathbb{C}, +\cdot)$ por lo tanto $dim\big((\mathbb{C}^n, \mathbb{C}, +, \cdot)\big) = n$. Por el contrario en el ejemplo 6.39 se vio que $\mathcal{C}$ no es base de $(\mathbb{C}^n, \mathbb{R}, +, \cdot)$ pero si lo es

$$\mathcal{C}' = \{(1, 0, \ldots, 0), (i, 0, \ldots, 0), (0, 1, \ldots, 0), (0, i, \ldots, 0), \ldots \\ \ldots, (0, 0, \ldots, 1), (0, 0, \ldots, i)\}$$

de donde $dim\big((\mathbb{C}^n, \mathbb{R}, +, \cdot)\big) = 2n$. Obsérvese que al cambiar el cuerpo, aunque no hayamos cambiado el conjunto de vectores, el espacio vectorial cambia radicálmente al punto que cambia su dimensión. En los ejercicios veremos que incluso hay ejemplos en los cuales un espacio de dimensión finita puede transformarse en uno de dimensión infinita cambiando sólo el cuerpo. ■

EJEMPLO 6.45. De lo visto en los ejemplos 6.40 y 6.41 se deduce que $dim(\mathcal{M}_{2\times 2}) = 4$ y $dim(\mathcal{P}_n) = n + 1$. ¿Cuál es la dimensión de $\mathcal{M}_{m\times n}$?. ■

EJEMPLO 6.46. Sean $S = \{(x, y, z) : 2x - y + z = 0\} \subset \mathbb{R}^3$ y S' el subespacio de las matrices 2×2 simétricas de lo visto en el ejemplos 6.42 y 6.43 se deduce respectivamente que existe una base de S con 2 elementos y una base de S' con 3 elementos por lo tanto $dim(S) = 2$ y $dim(S') = 3$. ■

El siguiente resultado da otra caracterización de una base, de hecho afirma que un base es un conjunto L.I. "maximal" en el sentido de que cualquier otro conjunto más grande debe ser L.D.

TEOREMA 6.26. *Sea $A = \{v_1, v_2, ..., v_n\} \subset V$. Entonces A es base si y sólo si A es L.I. y todo conjunto con más de n vectores es L.D.*

Demostración. ($\Leftarrow$)

Para probar que A es base , alcanza probar que A es generador de todo el espacio, puesto que A es L.I. por hipótesis. Para esto veremos que cualquier vector puede ponerse como combinación lineal de A. Es claro que los vectores de A son ellos mismos combinación de A así que consideremos un vector $v \in V$ cualquiera tal que $v \notin A$, entonces el conjunto $A' = A \cup \{v\}$ tiene $n + 1$ elementos y en consecuencia es por hipótesis un conjunto L.D. Entonces existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n, \lambda$ no todos nulos tales que

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n + \lambda v = \vec{0} \tag{6.45}$$

Veamos que debe ser $\lambda \neq 0$, ya que si $\lambda = 0$ se tendría que

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = \vec{0}$$

con algún escalar distinto de cero y esto contradice la independencia lineal de A asegurada por hipótesis.

Entonces tenemos $\lambda \neq 0$. Despejando v en 6.45 se tiene:

$$v = -\frac{\lambda_1}{\lambda}v_1 - \frac{\lambda_2}{\lambda}v_2 - \dots - \frac{\lambda_n}{\lambda}v_n$$

Hemos probado que v es c.l. de A. Como v es un vector cualquiera de V, se tiene que $A \xrightarrow{g} V$, como queríamos probar.

($\Rightarrow$)

Tenemos que probar que todo conjunto con más de n vectores es L.D. Sabemos por hipótesis que A es base y por lo tanto en particular A es generador de V. Sea $W \subset V$ un conjunto cualquiera con m vectores $m > n$. Si W fuera L.I. tendríamos un conjunto L.I. con mas elementos que un generador y esto contradice el teorema 6.23, por lo tanto W debe ser L.D. □

EJEMPLO 6.47. Sea $A = \{x^2+x-1, x+1, x^2+3x-3, x^2-1, -x^2+2x+1\} \subset \mathcal{P}_2$ entonces A es L.D. pues $dim(\mathcal{P}_2) = 3$ y $card(A) = 5$ ■

OBSERVACIÓN 6.27. De hecho si $m > n$ cualquier conjunto de $\mathbb{R}^n$ con m vectores es necesariamente L.D. Obsérvese que si $m < n$ un conjunto de $\mathbb{R}^n$ con m vectores puede ser tanto L.D. como L.I. (verifique esto construyendo ejemplos de ambas situaciones).

OBSERVACIÓN 6.28. Sea A un subconjunto de finito de vectores de $\mathbb{R}^n$ en el capítulo 2 habíamos definido el concepto de rango del conjunto como la mayor cantidad de vectores de A que fueran linealmente independientes. Ahora podemos observar que el rango de A no es otra cosa que la dimensión del subespacio generado por A. En particular el rango de una matriz es la dimensión del subespacio generado por las filas o por las columnas de la misma.

Los siguiente resultados son útiles para construir bases de subespacios.

TEOREMA 6.29. *Sea V un espacio vectorial de dimensión $n \geqslant 1$, A y B subconjuntos de V (finitos).*

Se cumplen las siguientes propiedades:

1) Si B genera a V, existe B' contenido en B, tal que B' es base de V.

2) Si A es L.I., existe A' que contiene a A, tal que A' es base de V (dicho de otra forma, todo conjunto generador, contiene a alguna base; y todo L.I. se puede agrandar hasta formar una base).

DEMOSTRACIÓN. 1) Sea $B = \{v_1, v_2, ..., v_k\}$. Si B es L.I., entonces es base (porque por hipótesis B genera a V). Si B es L.D., o bien está formado por un único vector, que tiene que ser el vector nulo (porque B es L.D.), o bien está formado por más de un vector. En el caso en que $B = \left\{\vec{0}\right\}$, como por hipótesis B genera a V, sería el caso en que $V = \left\{\vec{0}\right\}$. Pero este caso está excluido, porque por hipótesis $n \geqslant 1$. De esta forma, resta sólo estudiar el caso en que B es L.D. y está formado por más de un vector. Por la Proposición 6.18, hay un vector de B, que es combinación lineal de los demás vectores de B. Quitando ese vector v_r de B, se obtiene un conjunto B_1. Por la proposición 6.16 $[B] = [B_1]$ y entonces B genera a V.

Si B_1 es L.I: ya está probado. Si es L.D. aplicando el mismo razonamiento a B_1 en lugar de B, se obtiene un nuevo conjunto $B_2 \subset B_1 \subset B$ tal que genera a V. Si se repite el razonamiento, se obtendrá finalmente un conjunto B' contenido en B, que es L.I. y que se obtuvo de B retirando sucesivamente aquellos vectores que se expresaban como combinación lineal de los restantes. Por la proposición 6.16, este conjunto B' también genera a V. Por definición es base de V, probando 1).

2) Supongamos $A = \left\{w_1, ..., w_p\right\}$ es un conjunto L.I. Si A genera a V, entonces A es base y tomamos $A' = A$. Si A no genera a V, entonces existe algún vector de V que no es combinación lineal de A. Llamemos w_{p+1} a este vector. Mostremos que: $\left\{w_1, w_2, ..., w_p, w_{p+1}\right\}$ es L.I.

En efecto: $\lambda_1 w_1 + \lambda_2 w_2 + ... + \lambda_p w_p + \lambda_{p+1} w_{p+1} = \vec{0} \;\Rightarrow\; \lambda_{p+1} = 0$, porque sino se podría despejar w_{p+1} en función de los demás, y eso no es posible porque w_{p+1} no es c.l. de $w_1, w_2, ..., w_p$. Entonces: $\lambda_1 w_1 + \lambda_2 w_2 + ... + \lambda_p w_p = \vec{0}$. Luego, todos los coeficientes λ_i son ceros, porque A es L.I. Tenemos entonces que el conjunto $A = \{w_1, w_2, ..., w_p, w_{p+1}\}$ es L.I. Repitiendo la construcción anterior a A_1 en lugar de A, vamos agregando vectores hasta obtener un conjunto L.I. y que genere a V (sea hasta obtener una base de V). Siempre se llega, en una cantidad finita de pasos, a un conjunto que genere a V, porque por el teorema 6.26 los conjuntos L.I. no pueden tener mas de n elementos, donde n es la dimensión del espacio. □

El siguiente resultado es particularmente útil cuando se desea encontrar una base de un espacio del cual se conoce la dimensión, de hecho en este caso solo es necesario chequear la cantidad de vectores del conjunto candidato y su independencia lineal

COROLARIO 6.30. *Sea V un espacio vectorial de* ***dimensión n*** *. Se cumplen las siguientes afirmaciones:*

1) Todo conjunto linealmente independiente A de n vectores, es base.

2) Todo conjunto de n vectores A que genera a V es base.

DEMOSTRACIÓN. 1) Si $A = \{v_1, ..., v_n\}$ no genera V, existe $v \in V$ tal que $v \notin [A]$. Entonces $\{v_1, ..., v_n, v\}$ es L.I. (como en la demostración del teorema anterior) pero tiene $n + 1$ vectores, lo que contradice el teorema 6.26, absurdo. Luego $[A] = V$.

2) Sea $A = \{v_1, ..., v_n\}$, con $A \xrightarrow{g} V$. Si A es L.D. aplicando el teorema 6.29 se sabe que existe $A_1 \subsetneq A$ tal que $A_1 \xrightarrow{b} V$, y $card(A_1) = p < n$. Luego+ $dim\ (V) = p < n$. Absurdo. □

EJEMPLO 6.48. Sea $A = \{(1,0,1,0), (-1,0,1,0), (0,1,0,1), (0,-1,0,1)\} \subset \mathbb{R}^4$ como es fácil de verifica A es L.I. y como $card(A) = 4 = dim(\mathbb{R}^4)$ entonces A es base de $\mathbb{R}^4$. ■

EJEMPLO 6.49. Sea $S = [(1,1,1),(2,1,0),(-1,0,1)]$, hallar una base de S y determinar su dimensión. Para esto comenzamos por encontrar un generador, lo cual obviamente es inmediato, el conjunto

$$A = \{(1,1,1),(2,1,0),(-1,0,1)\}$$

es por definición un generador de S como desconocemos la dimensión del subespacio no podemos aplicar la proposición 6.30 y debemos verifica si A es L.I. o no. Es fácil de ver que A es L.D. pues $(1,1,1)-(2,1,0)=(-1,0,1)$. Entonces A no es base pero de la proposición 6.29 existe $A' \subset A$ que si lo es. Como el vector $(-1,0,1)$ es c.l. de los restantes elementos de A la proposición 6.16 $A' = \{(1,1,1),(2,1,0)\} \xrightarrow{g} S$. Es fácil verificar además que A' es L.I. en consecuencia $A' \xrightarrow{b} S$ y por lo tanto $dim(S) = 2$. ■

EJEMPLO 6.50. Sea $A = \{q_1, q_2, q_3, q_4\}$ donde $q_1(x) = x+1$, $q_2(x) = x^2+x-1$, $q_3(x) = x^2+2x$ y $q_4(x) = x^2-2$. Analicemos ahora el subespacio $S = [A] \subset \mathcal{P}_2$. Queremos determinar la dimensión de S. Para eso encontraremos una base del subespacio. $A \xrightarrow{g} S$ pero por lo visto en el ejemplo 6.35 A es L.D., por lo cual debe contener propiamente un subconjunto que sea base. Para determinar este subconjunto procedemos como en la prueba del teorema 6.29 y determinemos un conjunto $A' \subsetneq A$ tal que $[A'] = [A]$ eliminando de A un vector que sea combinación de los restantes. Sabemos del ejemplo 6.35 que q_4 es combinación de los restantes elementos de A en consecuencia consideramos $A = \{q_1, q_2, q_3\}$ en este caso es fácil ver que $q_3 = q_1 + q_2$ por lo tanto A' es también L.D. Eliminando q_3 tenemos que $A'' = \{q_1, q_2\} \xrightarrow{g} S$ y como el lector comprobará A'' es L.I. y por lo tanto una base de S. Entonces $dim(S) = 2$. ■

PROPOSICIÓN 6.31. *Sea V un espacio vectorial de dimensión finita n. Sea W un subespacio de V. Entonces W tiene también dimensión finita, que es menor o igual que n.*

DEMOSTRACIÓN. Si $W = \left\{\vec{0}\right\}$, entonces dim $W = 0$ y ya está probado, pues $0 \leqslant n$.

Si $W \neq \left\{\vec{0}\right\}$, alcanza probar que existe una base de W con una cantidad finita de elementos p. Una vez demostrado eso, por el teorema 6.23 resulta $p \leqslant n$, pues una base de W, es un conjunto L.I., con p elementos, y una base de V es un conjunto que genera a todo el espacio, con n elementos.

Como $W \neq \left\{\vec{0}\right\}$ entonces existe algún vector $w_1 \in W$, $w_1 \neq \vec{0}$. Si $\{w_1\}$ genera a W, es base de W, y ya está probado. Si $\{w_1\}$ no genera a W, existe un vector $w_2 \in W$ que no es combinación lineal de w_1.

Afirmamos que $\{w_1, w_2\}$ es L.I. . En efecto, si $\lambda_1 w_1 + \lambda_2 w_2 = \vec{0}$, $\lambda_2 = 0$ pues de lo contrario se podría despejar w_2 en función de w_1, lo cual no es posible porque w_2 no es combinación lineal de w_1. Entonces $\lambda_2 = 0$, y entonces $\lambda_1 w_1 = \vec{0}$. Como $w_1 \neq \vec{0}$ se deduce $\lambda_1 = 0$.Tenemos ahora que $\{w_1, w_2\}$ es L.I. . Si genera a W, ya tenemos una base. Si no genera a W, existe un tercer vector $w_3 \in W$ que no es combinación lineal de los dos primeros. Igual que antes se prueba que el conjunto $\{w_1, w_2, w_3\}$ es L.I.. Se puede repetir el razonamiento en una cantidad menor que $n + 1$ de pasos hasta construir una base de W, porque de lo contrario tendríamos un conjunto L.I. de $n + 1$ vectores, contradiciendo el teorema 6.26 que dice que los conjuntos L.I. no pueden tener más de n elementos, donde n es la dimensión del espacio. $\square$

PROPOSICIÓN 6.32. *Sea W un subespacio de V, con $dim\,(V) = n$. Si $dim\,(W) = dim\,(V)$, entonces $W = V$.*

DEMOSTRACIÓN. Sea $\{w_1, ..., w_n\}$ base de W. Entonces es L.I. y está formado por n vectores de $W \subset V$. Si $v \in V$; $\{w_1, ..., w_n, v\}$ es L.D. (por Teorema 6.26) y como antes, vemos que v se escribe como combinación lineal de $\{w_1, ..., w_n\}$, luego $\{w_1, ..., w_n\} = V$. $\square$

6.6. Suma de subespacios y suma directa

DEFINICIÓN 6.9. Suma de subespacios: Sea V un espacio vectorial. se llama **suma** de los subespacios $V_1, V_2, ..., V_n$ de V, (y se denota como $V_1 +$

$V_2 + ... + V_n$) al subconjunto de vectores W definido por:

$$W = \{v \in V : \exists,\ v_1 \in V_1, v_2 \in V_2, ..., v_n \in V_n\ ,\ v = v_1 + v_2 + ... + v_n\}$$

TEOREMA 6.33. *La suma de subespacios de V es unr subespacio de V.*

DEMOSTRACIÓN. Es fácil ver que $\vec{0} \in W$, porque $\vec{0} = \vec{0} + \cdots + \vec{0}$ y $\vec{0} \in V_i \forall i = 1, 2, ..., n$. También es sencillo verificar que el vector u se puede descomponer como suma de un vector de V_1, uno de $V_2, \cdots$, uno de V_n; entonces λu también se descompone de esa forma. Es claro que W es cerrado con respecto al producto por escalares. Para verificar que W es cerrado respecto a la suma sean u y u' tales que $u = v_1 + v_2 + ... + v_n$ y $u' = v_1' + v_2' + ... + v_n'$ con $v_1\, y\, v_1' \in V_1$; $v_2\, y\, v_2' \in V_2$; ... ; $v_n\, y\, v_n' \in V_n$. Entonces $u + u' = (v_1 + v_1') + (v_2 + v_2') + ... + (v_n + v_n')$, donde $v_1 + v_1' \in V_1$; $v_2 + v_2' \in V_2$; ... ; $v_n + v_n' \in V_n$, porque $V_1, V_2, ..., V_n$ son subespacios. □

DEFINICIÓN 6.10. Suma directa: Dados los subespacios $V_1, V_2, ..., V_n$, el subespacio suma de ellos se llama **suma directa** cuando todo vector de él puede expresarse de forma única como suma de vectores de los subespacios $V_1, V_2, ..., V_n$. Cuando la suma es directa se representa $V_1 \oplus V_2 \oplus ... \oplus V_n$.

EJEMPLO 6.51. Sean en $\mathbb{R}^3$ los subespacios $V_1 = \{(a_1, a_2, a_3) : a_3 = 0\}$ y $V_2 = \{(a_1, a_2, a_3) : a_1 = 0\}$. Todo vector de $\mathbb{R}^3$ se puede escribir como suma de un vector de V_1 y uno de V_2, pero no de forma única:

$$(1, 1, 1) = (1, 1, 0) + (0, 0, 1)$$

$$(1, 1, 1) = (1, 1/2, 0) + (0, 1/2, 1)$$

Entonces la suma de V_1 y V_2 es todo $\mathbb{R}^3$. Pero esta suma no es directa.

TEOREMA 6.34. *Si $V = V_1 \oplus V_2 \oplus ... \oplus V_n$ y si $B_1, B_2, ..., B_n$ son las bases de $V_1, V_2, ..., V_n$, respectivamente entonces: $B = B_1 \cup B_2 \cup ... \cup B_n$ es una base de V.*

DEMOSTRACIÓN. Probaremos que B genera a V, y que B es L.I.

Todo vector de V es suma de vectores de $V_1, V_2, ..., V_n$, y a su vez, estos son combinaciones lineales de sus respectivas bases. Entonces v es c.l.de la unión de esas bases, que es B. Hemos probado que todo vector de V es combinación lineal de B, o sea B genera a V. Veremos que B es L.I.. Sea $B_1 = \left\{v_{11}, v_{12}, ..., v_{1p_1}\right\}$; $B_2 = \{v_{21}, v_{22}, ..., v_{1p_2}\}$;$\cdots$ $B_n = \left\{v_{n1}, v_{n2}, ..., v_{np_n}\right\}$. Tomemos una combinación lineal de B igualada a $\vec{0}$:

$$\begin{aligned} a_{11}v_{11} + a_{12}v_{12} + ... &+ a_{1p_1}v_{1p_1} + a_{21}v_{21} + a_{22}v_{22} + ... \\ +a_{2p_2}v_{2p_2} + ... + a_{n1}v_{n1} &+ a_{n2}v_{n2} + ... + a_{np_n}v_{np_n} = \vec{0}; \end{aligned}$$

como $\vec{0} \in V$ y $V = V_1 \oplus V_2 \oplus ... \oplus V_n$, existe una única manera de expresar el vector nulo como suma de vectores de $V_1, V_2, ..., V_n$. Esa única manera tiene que ser entonces $\vec{0} = \vec{0} + \vec{0} + ... + \vec{0}$.

La combinación lineal de más arriba implica entonces que:

$$a_{11}v_{11} + a_{12}v_{12} + ... + a_{1p_1}v_{1p_1} = \vec{0}$$

$$a_{21}v_{21} + a_{22}v_{22} + ... + a_{2p_2}v_{2p_2} = \vec{0}$$

$$\cdots$$

$$a_{n1}v_{n1} + a_{n2}v_{n2} + ... + a_{np_n}v_{np_n} = \vec{0}.$$

Siendo B_1 una base de V_1, se tiene que B_1 es L.I., y de la primera igualdad de mas arriba se deduce $a_{11} = a_{12} = ... = a_{1n} = \vec{0}$. Análogamente para los demás coeficientes a_{ij}. Entonces tenemos probado que B es L.I. $\square$

COROLARIO 6.35. *Si V es suma directa de $V_1, V_2, ..., V_n$ entonces $dimV = dimV_1 + dimV_2 + ... + dimV_n$*

DEMOSTRACIÓN. Es inmediata a partir de la definición de dimensión y del teorema anterior. $\square$

PROPOSICIÓN 6.36. *Sean V_1 y V_2 dos subespacios de V tales que $V = V_1 + V_2$. Entonces, $V_1 \cap V_2 = \left\{\vec{0}\right\}$ si y sólo si $V = V_1 \oplus V_2$.*

DEMOSTRACIÓN. $(\Rightarrow)$

Supongamos $v = v_1 + v_2$ y $v' = v_1' + v_2'$ con $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$. Luego, $v_1 - v_1' = v_2 - v_2' \in V_1$ y V_2, por lo que , por hipótesis $v_1 - v_1' = v_2 - v_2' = \vec{0}$, lo que prueba que la suma es directa.

$(\Leftarrow)$

Sea $v \in V_1 \cap V_2$ *luego* $-v \in V_1 \cap V_2$. Por lo tanto:

$\vec{0} = v + (-v) = \vec{0} + \vec{0}$ y por definición de suma directa, $\vec{0} = v = -v$. Entonces $V_1 \cap V_2 = \left\{\vec{0}\right\}$. $\square$

CAPíTULO 7

TRANSFORMACIONES LINEALES.

Las transformaciones lineales son funciones entre espacios vectoriales. Ellas preservan (mantienen) las estructuras lineales que caracterizan a esos espacios: los transformados de combinaciones lineales de vectores son las correspondientes combinaciones de los transformados de cada uno de los vectores.

Las funciones que mantienen las estructuras características de los espacios entre los que actúan son de especial importancia porque permiten "moverse" entre ellos comparando las estructuras que los definen; en particular, permiten analizar cuán diferentes o semejantes son desde el punto de vista de esas propiedades (lineales, en el presente caso). En este sentido las transformaciones lineales son tan importantes en el estudio de los espacios vectoriales, como las isometrías o movimientos lo son en la geometría métrica.

7.1. Transformaciones lineales

DEFINICIÓN 7.1. Sean V y W dos espacios vectoriales sobre el mismo cuerpo K y $T: V \to W$ una función. Diremos que T es una **transformación lineal** si satisface :

$$\begin{array}{ll} i)\ T(u+v) = T(u) + T(v) & \forall\, u, v \in V. \\ ii)\ T(a.v) = a.T(v) & \forall\, a \in K,\ \forall\, v \in V. \end{array}$$

EJEMPLO 7.1. Sean

$$C^1 = \{f: R \to R \ / \ f \text{ es derivable}\} \subset F = \{f: \mathbb{R} \to \mathbb{R}\}.$$

La transformación $T: C^1 \to F$ tal que $T(f) = f'$ es una transformación lineal, pues se cumple que
i) $T(f_1 + f_2) = (f_1 + f_2)' = f_1' + f_2' = T(f_1) + T(f_2)$
ii) $T(\alpha f) = (\alpha f)' = \alpha f' = \alpha T(f)$

EJEMPLO 7.2. Sea V un espacio vectorial y sea $\lambda \in K$. La función $T: V \to V$ tal que $T(v) = \lambda v$ es una transformación lineal. Efectivamente:
i) $T(v_1 + v_2) = \lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 = T(v_1) + T(v_2)$
ii) $T(a.v) = \lambda a v = a \lambda v = a T(v)$

EJEMPLO 7.3. Sea V un espacio vectorial y $v_0 \neq \vec{0}$ un vector fijo de V. La función $T: V \to V$ tal que: $T(v) = v + v_0$ no es una transformación lineal, pues $T(v_1 + v_2) = v_1 + v_2 + v_0 \neq v_1 + v_0 + v_2 + v_0 = T(v_1) + T(v_2)$.

Las propiedades i) y ii) de la definición anterior se pueden resumir de la siguiente forma:

PROPOSICIÓN 7.1. *$T: V \to W$ es una transformación lineal si y sólo si $T(a_1 v_1 + a_2 v_2) = a_1 T(v_1) + a_2 T(v_2)$ $\forall a_1, a_2 \in K$ $\forall v_1, v_2 \in V$.*

Demostración. Sean a_1 y a_2 escalares y v_1 y v_2 vectores.
$(\Rightarrow)$ $T(a_1 v_1 + a_2 v_2) = T(a_1 v_1) + T(a_2 v_2) = a_1 T(v_1) + a_2 T(v_2)$.
$(\Leftarrow)$ En el caso particular en que $a_1 = a_2 = 1$ se obtiene $T(v_1 + v_2) = T(v_1) + T(v_2)$, y en el caso particular en que $a_2 = 0$ se obtiene $T(a_1 v_1) = a_1 T(v_1)$. $\square$

PROPOSICIÓN 7.2. *Sea $T: V \to W$ una transformación lineal. Entonces, $T\left(\vec{0}\right) = \vec{0}$.*

Demostración. $T\left(\vec{0}\right) = T(0.v) = 0.T(v) = \vec{0}$. $\square$

PROPOSICIÓN 7.3. *Sean V y W espacios vectoriales sobre un mismo cuerpo con* $dim\,(V) = n$, $T : V \to W$ *y* $S : V \to W$ *transformaciones lineales. Consideremos una base* $B = \{v_1, \ldots, v_n\}$ *de* V. *Entonces:* $T\,(v) = S\,(v)\ \ \forall\, v \in V$ *si y solo si* $T\,(v_i) = S\,(v_i) \quad \forall i = 1, 2, \ldots, n$.

DEMOSTRACIÓN. ($\Rightarrow$) Si T y S coinciden en todo el espacio V, en particular coinciden en los vectores de la base B.

($\Leftarrow$) Sea $v \in V$. Como B es base de V, existen $\alpha_1, \ldots, \alpha_n \in K$ tales que $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Luego, por ser T y S lineales

$$\begin{aligned} T\,(v) = T\,(\alpha_1 v_1 + \cdots + \alpha_n v_n) &= \alpha_1 T\,(v_1) + \cdots + \alpha_n T\,(v_n) \\ = \alpha_1 S\,(v_1) + \cdots + \alpha_n S\,(v_n) &= S\,(\alpha_1 v_1 + \cdots + \alpha_n v_n) = S\,(v)\,. \end{aligned}$$

□

TEOREMA 7.4. *Sean V, W espacios vectoriales sobre un mismo cuerpo con* $dim\,(V) = n$. *Sea* $\{v_1, \ldots, v_n\}$ *una base de V y* $w_1, \ldots, w_n \in W$ *arbitrarios. Entonces, existe una única transformación lineal* $T : V \to W$ *tal que* $T\,(v_i) = w_i \quad \forall\, i = 1, \ldots, n$.

DEMOSTRACIÓN. <u>Existencia</u>: Si $v \in V$ existen únicos $a_1, \ldots, a_n$ tales que $v = \sum\limits_{i=1}^{n} a_i\, v_i$. Definimos entonces $T : V \to W$ mediante $T\,(v) = \sum\limits_{i=1}^{n} a_i\, w_i$. Hay que verificar que T es lineal. Sean v y $\bar{v}$ vectores y λ, μ escalares. Tenemos $v = \sum\limits_{i=1}^{n} a_i\, v_i$ y $\bar{v} = \sum\limits_{i=1}^{n} \bar{a}_i\, v_i$. Entonces $T\,(\lambda\, v + \mu\, \bar{v}) =$

$= T\left(\sum\limits_{i=1}^{n} (\lambda\, a_i + \mu\, \bar{a}_i)\, v_i\right) = \sum\limits_{i=1}^{n} (\lambda\, a_i + \mu\, \bar{a}_i)\, w_i = \lambda \sum\limits_{i=1}^{n} a_i\, v_i + \mu \sum\limits_{i=1}^{n} \bar{a_i} v_i =$
$= \lambda\, T\,(v) + \mu\, T\,(\bar{v})$. Luego T es lineal por la proposición 7.1

<u>Unicidad:</u> Supongamos que existan $T : V \to W$ y $S : V \to W$ lineales tales que $T\,(v_i) = w_i$ y $S\,(v_i) = w_i$. Entonces T y S coinciden en una base de V; y por la proposición 7.3, resulta que T=S. □

Hemos probado, pues, que una transformación lineal queda determinada por los valores que toma sobre una base.

EJEMPLO 7.4. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ lineal tal que $T(1,2) = (3,-1,5)$ y $T(0,1) = (2,1,-1)$ Como $\{(1,2),(0,1)\}$ es una base de $\mathbb{R}^2$, T está bien determinada.

Hallemos $T(x,y)$ con $(x,y) \in \mathbb{R}^2$.

$$\begin{aligned} T(x,y) = T(x\ (1,2) + (-2x+y)(0,1)) &= xT(1,2) + (-2x+y)T(0,1) = \\ = x(3,-1,5) + (-2x+y)(2,1,-1) &= (-x+2y,-3x+y,7x-y). \end{aligned}$$

Así $T(x,y) = (-x+2y,-3x+y,7x-y)$.

EJEMPLO 7.5. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ lineal tal que $T(1,0) = T(0,1) = (2,3)$.

Como $\{(1,0),(0,1)\}$ es la base canónica de $\mathbb{R}^2$, T esta determinada.

Luego

$$\begin{aligned} T(x,y) = T(x(1,0) + y(0,1)) &= xT(1,0) + yT(0,1) = \\ = x(2,3) + y(2,3) &= (x+y)(2,3). \end{aligned}$$

7.2. Operaciones con transformaciones lineales.

DEFINICIÓN 7.2 (Suma de Transformaciones Lineales). Sean V, W espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$ y $S : V \to W$ y $T : V \to W$ dos transformaciones lineales. Se define la **suma de S y T** como la transformación $S+T : V \to W$ tal que

$$(S+T)(v) = S(v) + T(v) \quad \forall v \in V.$$

.

EJEMPLO 7.6. Sean $S:\mathbb{R}^3\to\mathbb{R}^2$ tal que $S(x,y,z)=(x+y+z,x-z)$ y $T:\mathbb{R}^3\to\mathbb{R}^2$ tal que $T(x,y,z)=(x-y,y+z)$. Entonces la suma de S y T es la transformación $S+T:\mathbb{R}^3\to\mathbb{R}^2$ tal que

$$\begin{aligned}(S+T)(x,y,z) &= S(x,y,z)+T(x,y,z)\\ &= (x+y+z,x-z)+(x-y,y+z)\\ &= (2x+z,x+y).\end{aligned}$$

PROPOSICIÓN 7.5. *Si S y T son transformaciones lineales, entonces $S+T$ es una transformación lineal.*

Demostración. Sean v_1 y v_2 vectores, λ y μ escalares. Probemos

$$(S+T)(\lambda v_1+\mu v_2)=\lambda\ (S+T)(v_1)+\mu\ (S+T)(v_2).$$

En efecto

$$(S+T)(\lambda v_1+\mu v_2)=S(\lambda v_1+\mu v_2)+T(\lambda v_1+\mu v_2)=$$
$$=\lambda S(v_1)+\mu S(v_2)+\lambda T(v_1)+\mu T(v_2)=$$
$$=\lambda\ [S(v_1)+T(v_1)]+\mu\ [S(v_2)+T(v_2)]=$$
$$=\lambda\ [(S+T)(v_1)]+\mu\ [(S+T)(v_2)].$$

□

DEFINICIÓN 7.3 (Producto por un escalar). Se define **el producto del escalar** $\lambda\in K$ **por la transformación** $\boldsymbol{T}$, como la transformación $(\lambda.T):V\to W$ tal que $(\lambda.T)(v)=\lambda.T(v)\quad\forall v\in V$.

EJEMPLO 7.7. Sea $T:\mathbb{R}^2\to\mathbb{R}^3$ tal que $T(x,y)=(x-2y,x+y,x)$

Entonces el producto del escalar $\lambda=5$ por la transformación T es la transformación $5.T:\mathbb{R}^2\to\mathbb{R}^3$ tal que

$$(5.T)(x,y)=5.T(x,y)=5.(x-2y,x+y,x)=$$
$$=(5-10y,5x+5y,5x).$$

PROPOSICIÓN 7.6. *Si T es una transformación lineal y λ un escalar entonces λT es una transformación lineal.*

Demostración. Probemos que

$(\lambda T)(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 (\lambda T)(v_1) + \alpha_2 (\lambda T)(v_2)\ \ \forall v_1, v_2 \ \in V\ \ \forall \alpha_1, \alpha_2 \ \in K$

En efecto $(\lambda T)(\alpha_1 v_1 + \alpha_2 v_2) = \lambda T (\alpha_1 v_1 + \alpha_2 v_2) =$

$$= \lambda [\alpha_1 T(v_1) + \alpha_2 T(v_2)] = \lambda \alpha_1 T(v_1) + \lambda \alpha_2 T(v_2) =$$

$$= \alpha_1 \ \lambda T(v_1) + \alpha_2 \ \lambda T(v_2) = \alpha_1 (\lambda T)(v_1) + \alpha_2 (\lambda T)(v_2).$$

□

DEFINICIÓN 7.4 (Composición de Transformaciones Lineales.). Sean U, V y W espacios vectoriales sobre un mismo cuerpo K, y $S : U \to V$ y $T : V \to W$ dos transformaciones lineales. Se define la composición S y T como la transformación $T \circ S : U \to W$ tal que $(T \circ S)(u) = T(S(u))$.

EJEMPLO 7.8. Sean $S : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $S(x, y, z) = (2x, y + z)$ y $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(x, y) = (y, x)$.

Luego $T \circ S : \mathbb{R}^3 \to \mathbb{R}^2$ es tal que $(T \circ S)(x, y, z) = (y + z, 2x)$.

PROPOSICIÓN 7.7. *Si S y T son lineales, entonces $T \circ S$ es lineal*

Demostración. Probemos que $\forall \alpha_1, \alpha_2 \in K \ \ \forall u_1, u_2 \in V$.

$$(T \circ S)(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 \ (T \circ S)(u_1) + \alpha_2 \ (T \circ S)(u_2)$$

En efecto $(T \circ S)(\alpha_1 u_1 + \alpha_2 u_2) = T(S(\alpha_1 u_1 + \alpha_2 u_2)) =$

$$= \alpha_1 T(\ S(u_1)) + \alpha_2 T(\ S(u_2)) = \alpha_1 \ (T \circ S)(u_1) + \alpha_2 \ (T \circ S)(u_2).$$

□

7.3. Imagen y núcleo de una transformación lineal.

DEFINICIÓN 7.5. Sean V y W espacios vectoriales sobre el mismo cuerpo K y $T : V \to W$ una transformación lineal.
i) Llamaremos núcleo de T al conjunto de V cuya imagen por T es el vector nulo, es decir $N(T) = \left\{ v \in V : T(v) = \vec{0} \right\}$ (también se usará la notación $Ker(T)$ para $N(T)$).
ii) Llamaremos imagen de T al conjunto:
$Im(T) = \{ w \in W : \exists v \in V \text{ con } w = T(v) \}$ (también se utilizará la notación $T(V)$ para $Im(T)$.)

EJEMPLO 7.9. Se considera la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $T(x, y, z) = (x + y + z, 2x + 2y + 2z)$.

1) Hallemos el núcleo de T.

$$(x, y, z) \in N(T) \Leftrightarrow T(x, y, z) = (0, 0) \Leftrightarrow (x + y + z, 2x + 2y + 2z) = (0, 0)$$

$$\Leftrightarrow \begin{cases} x + y + z = 0 \\ 2x + 2y + 2z = 0 \end{cases} \Leftrightarrow x + y + z = 0$$

Así el $N(T) = \left\{ (x, y, z) \in \mathbb{R}^3 : x + y + z = 0 \right\}$.

2) Hallemos la imagen de T.

$$(a, b) \in Im(T) \Leftrightarrow \text{ existe } (x_0, y_0, z_0) \in \mathbb{R}^3 / \, T(x_0, y_0, z_0) = (a, b)$$

$$\Leftrightarrow \text{ existe } (x_0, y_0, z_0) \in \mathbb{R}^3 \text{ tal que } (x_0 + y_0 + z_0, 2x_0 + 2y_0 + 2z_0) = (a, b)$$

$$\Leftrightarrow \text{ existe } (x_0, y_0, z_0) \in \mathbb{R}^3 \text{ tal que } \begin{cases} x_0 + y_0 + z_0 = a \\ 2x_0 + 2y_0 + 2z_0 = b \end{cases}$$

$$\Leftrightarrow \text{ el sistema } \begin{cases} x + y + z = a \\ 2x + 2y + 2z = b \end{cases} \text{ es compatible}$$

$$\Leftrightarrow \text{ el sistema } \begin{cases} x + y + z = a \\ 0 = b - 2a \end{cases} \text{ es compatible}$$

$$\Leftrightarrow b = 2a$$

Luego: $Im\,(T) = \left\{(a,b) \in \mathbb{R}^2 \ : b = 2a \ \right\}$.

EJEMPLO 7.10. Se considera la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $T\,(x,y,z) = (x+y, y+z)$.

1) Hallemos el núcleo de T.

$$(x,y,z) \in \ N\,(T) \Leftrightarrow T\,(x,y,z) = \vec{0} \Leftrightarrow (x+y, y+z) = 0$$
$$\Leftrightarrow \begin{cases} x+y=0 \\ y+z=0 \end{cases}$$

Así el $N\,(T) = \left\{(x,y,z) \in \mathbb{R}^3 \ : x = -y \,, z = -y \text{ con } y \in \mathbb{R}\right\}$.

2) Hallemos la imagen de T.

$$(a,b) \ \in \ Im\,(T) \Leftrightarrow \ \text{existe} \ (x_0,y_0,z_0) \in \mathbb{R}^3 / \ T\,(x_0,y_0,z_0) = (a,b)$$
$$\Leftrightarrow \ \text{existe} \ (x_0,y_0,z_0) \in \mathbb{R}^3 \text{ tal que } \begin{cases} x_0+y_0=a \\ y_0+z_0=b \end{cases}$$
$$\Leftrightarrow \ \text{el sistema} \begin{cases} x+y=a \\ y+z=b \end{cases} \ \text{es compatible.}$$

Como el sistema es siempre compatible, resulta que $Im\,(T) = \mathbb{R}^2$.

EJEMPLO 7.11. Se considera la transformación lineal $T : P_2 \to \mathbb{R}^2$ tal que $T\,(p) = (p\,(0)\,, p\,(1))$.

1) Hallamos el núcleo $N\,(T)$.

Sea $p : p\,(t) = xt^2 + yt + z$ un polinomio de P_2

$$p \ \in N\,(T) \Leftrightarrow \ T\,(p) = (0,0) \Leftrightarrow (p\,(0)\,, p\,(1)) = (0,0)$$
$$\Leftrightarrow \begin{cases} p\,(0) = 0 \\ p\,(1) = 0 \end{cases} \Leftrightarrow \begin{cases} z = 0 \\ x+y+z=0 \end{cases} \Leftrightarrow \begin{cases} x \text{ cualquiera} \\ y = -x \\ z = 0 \end{cases}$$

Luego $N\,(T) = \left\{p \in P_2 / \ p : p\,(t) = xt^2 - xt \text{ con } x \in \mathbb{R}\right\}$

2) Hallemos la imagen de T, $Im\,(T)$.

$$(a,b) \in Im(T) \Leftrightarrow \text{ existe } p_0 \in P_2 \text{ tal que } T(p_0) = (a,b) \Leftrightarrow$$

$$\Leftrightarrow \text{ existe } p_0 \in P_2 \text{ tal que } (p_0(0), p_0(1)) = (a,b)$$

$$\Leftrightarrow \text{ existe } p_0 : p_0(t) = x_0t^2 + y_0t + z_0 \in P_2 \text{ tal que } (x_0, x_0 + y_0 + z_0) = (a,b)$$

$$\Leftrightarrow \text{ existen } x_0, y_0, z_0 \in \mathbb{R} \text{ tales que } \begin{cases} x_0 = a \\ x_0 + y_0 + z_0 = b \end{cases}$$

$$\Leftrightarrow \text{ el sistema } \begin{cases} x = a \\ x + y + z = b \end{cases} \quad \text{es compatible.}$$

Como el sistema anterior es siempre compatible, resulta que $Im(T) = \mathbb{R}^2$

EJEMPLO 7.12. Se considera la transformación lineal

$T : M_{2x2}(\mathbb{R}) \to M_{2x2}(\mathbb{R})$ tal que $T(M) = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} \cdot M.$

1) Hallemos el $N(T)$ Sea $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ un matriz en $M_{2x2}(\mathbb{R})$

$$M \in N(T) \Leftrightarrow T(M) = \vec{0} \Leftrightarrow \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} M = \vec{0}_{M_{2x2}(\mathbb{R})}$$

$$\Leftrightarrow \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} a - c & b - d \\ -2a + 2c & -2b + 2d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} a - c = 0 \\ b - d = 0 \\ -2a + 2c = 0 \\ -2b + 2d = 0 \end{cases} \Leftrightarrow \begin{cases} a = c \\ b = d \\ c \quad \text{cualquiera} \\ d \quad \text{cualquiera} \end{cases}$$

Así $N(T) = \left\{ M \in M_{2x2}(\mathbb{R}) : \ M = \begin{pmatrix} c & d \\ c & d \end{pmatrix} \quad \text{con } c, d \in \mathbb{R} \right\}.$

2) Hallar la $Im\,(T)$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Im\,(T) \Leftrightarrow \text{ existe } \begin{pmatrix} x_0 & y_0 \\ z_0 & t_0 \end{pmatrix} \in M_{2x2}\,(\mathbb{R}) \text{ tal que}$$

$$T\begin{pmatrix} x_0 & y_0 \\ z_0 & t_0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\Leftrightarrow \text{ existe } \begin{pmatrix} x_0 & y_0 \\ z_0 & t_0 \end{pmatrix} \in M_{2x2}\,(\mathbb{R}) \text{ tal que}$$

$$\begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x_0 & y_0 \\ z_0 & t_0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\Leftrightarrow \text{ existe } \begin{pmatrix} x_0 & y_0 \\ z_0 & t_0 \end{pmatrix} \in M_{2x2}\,(\mathbb{R}) \text{ tal que}$$

$$\begin{pmatrix} x_0 - z_0 & y_0 - t_0 \\ -2x_0 + 2z_0 & -2y_0 + 2t_0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$\Leftrightarrow$ existen $x_0, y_0, z_0, t_0 \in \mathbb{R}$ tales que

$$\begin{cases} x_0 - z_0 &= a \\ y_0 - t_0 &= b \\ -2x_0 + 2z_0 &= c \\ -2x_0 + 2z_0 &= d \end{cases}$$

$$\Leftrightarrow \text{ el sistema } \begin{cases} x - z &= a \\ y - t &= b \\ -2x + 2z &= c \\ -2x + 2z &= d \end{cases} \text{ tiene solución.}$$

$$\Leftrightarrow \text{ el sistema } \begin{cases} x - z &= a \\ y - t &= b \\ 0 &= 2a + c \\ 0 &= 2b + d \end{cases} \text{ tiene solución}$$

$$\Leftrightarrow 2a + c = 0 \text{ y } 2b + d = 0.$$

Así la $Im\,(T) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2x2}\,(\mathbb{R}) : 2a + c = 0 \text{ y } 2b + d = 0 \right\}.$

PROPOSICIÓN 7.8. *$N\;(T)$ es un subespacio vectorial de V.*

Demostración. Veamos que $N\,(T) \neq \phi$ y que es cerrado respecto de las operaciones del espacio vectorial.

1) $T\,(0) \;=\; \vec{0} \;\Rightarrow\; \vec{0} \;\in\; N\;(T)$.

2) Si $\alpha\,,\,\beta \;\in\; K$ y $v_1\,,\,v_2 \;\in\; N\;(T)$ entonces:

$$T\,(\alpha\,v_1 \;+\; \beta\,v_2) \;=\; \alpha\,T\,(v_1) \;+\; \beta\,T\,(v_2) \;=\; \alpha\,.\,\vec{0} \;+\; \beta\,.\,\vec{0} \;=\; \vec{0} \;\Rightarrow$$

$$\Rightarrow \;\; \alpha\,v_1 \;+\; \beta\,v_2 \;\in\; N\;(T)$$

Luego $N\;(T)$ es un subespacio de V. □

PROPOSICIÓN 7.9. *$Im\;(T)$ es un subespacio vectorial de W.*

Demostración. Análogamente,

1) $T\left(\vec{0}\right) \;=\; \vec{0} \;\Rightarrow\; \vec{0} \;\in\; Im\;(T)$

2) Si $\alpha\,,\,\beta \;\in\; K$ y $w_1\,,\,w_2 \;\in\; Im\;(T)$ entonces existen $v_1\,,\,v_2$ tales que $T\,(\,v_1) \;=\; w_1$ y $T\,(\,v_2) \;=\; w_2$. Entonces si llamamos $v \;=\; \alpha\,v_1 \;+\; \beta\,v_2$, se tiene $T\,(v) \;=\; \alpha\,w_1 + \beta\,w_2$ y por lo tanto $\alpha\,w_1 + \beta\,w_2 \;\in\; Im\;(T)$. Luego $Im\;(T)$ es un subespacio de W. □

EJEMPLO 7.13. Hallemos una base del núcleo y una base de la imagen de la transformación lineal $T : \mathbb{R}^4 \;\rightarrow \mathbb{R}^3$ definida por

$$T\,(x, y, z, t) = (x - y + z + t, x + 2z - t, x + y + 3z - 3t)$$

1) $(x, y, z, t) \in N\,(T) \Leftrightarrow T\,(x, y, z, t) = (0, 0, 0) \Leftrightarrow$

$$\Leftrightarrow (x - y + z + t, x + 2z - t, x + y + 3z - 3t) = (0, 0, 0)$$

$$\Leftrightarrow \begin{cases} x-y+z+t &= 0 \\ x+2z-t &= 0 \\ x+y+3z-3t &= 0 \end{cases} \Leftrightarrow \begin{cases} x-y+z+t &= 0 \\ y+z-2t &= 0 \\ 2y+2z-4t &= 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x-y+z+t &= 0 \\ y+z-2t &= 0 \end{cases} \Leftrightarrow \begin{cases} x &= -2z+t \\ y &= -z+2t \end{cases}$$

Luego los vectores del núcleo de T son de la forma

$$(-2z+t,-z+2t,z,t) \text{ con } z,t \in \mathbb{R}. \text{ Como}$$

$$(-2z+t,-z+2t,z,t) = z\,(-2,-1,1,0)+t\,(1,2,0,1) \text{ con } z,t \in \mathbb{R},$$

todo vector del núcleo de T es combinación lineal de los vectores $(-2,-1,1,0)$ y $(1,2,0,1)$. Por ser $\{(-2,-1,1,0),\ (1,2,0,1)$ un conjunto L.I., es una base de *N(T)*.

2) $w \in Im\,(T) \Leftrightarrow \exists v \in \mathbb{R}^4$ tal que $w = T\,(v)$, es decir

$$\exists\,(x,y,z,t) \in \mathbb{R}^4 \text{ tal que } w = T\,(x,y,z,t)\,,$$

$$w = (x-y+z+t, x+2z-t, x+y+3z-3t)$$

$\Leftrightarrow \exists x,y,z,t \in \mathbb{R}$ tal que $v = (x,x,x)+(-y,0,y)+z\,(1,2,3)+t\,(1,-1,-3)\,.$

Luego todo vector de la imagen de T es combinación lineal de

$$\{(1,1,1)\,,(-1,0,1)\,,(1,2,3)\,,(1,-1,-3)\}$$

(observar que dichos vectores están en la imagen pues son imagen respectivamente de los vectores de la base canónica).

Pero el conjunto anterior es L.D. pues tiene 4 vectores y $\dim\left(\mathbb{R}^3\right)$=3. Observemos que

$$\begin{aligned} &(1,2,3) = 2\,(1,1,1)+(-1,0,1) \\ &(1,-1,-3) = -1\,(1,1,1)+(-2)\,(-1,0,1) \end{aligned}$$

Luego reducimos el generador L.D. hasta obtener el generador

$$\{(1,1,1)\,,(-1,0,1)\}$$

que sí es L.I Luego $\{(1,1,1)\,,(-1,0,1)\}$ es base de $Im\,(T)$.

PROPOSICIÓN 7.10. *Sea V, W espacios vectoriales de dimensión finita sobre un mismo cuerpo y $T : V \to W$ lineal.*
i) Si $\{v_1, \dots, v_n\}$es un generador de V, entonces $\{T(v_1), \dots, T(v_n)\}$ es un generador de la $Im(T)$ (es decir que la imagen de un conjunto generador de V es un generador de la $Im(T)$).
ii) Si $\{T(v_1), \dots, T(v_n)\}$ es un conjunto L.I. en $Im(T)$ entonces $\{v_1, \dots, v_n\}$ es conjunto L.I. en V (es decir que la pre-imagen de un conjunto L.I. en la $Im(T)$ es un conjunto L.I. en V)

Demostración. i) Sea $w \in Im(T)$. Entonces existe $v \in V$ tal que $w = T(v)$.

Como $v \in V$ y siendo $\{v_1, \dots, v_n\}$ un generador de V, existen $\alpha_1, \dots, \alpha_n$ escalares tales que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Luego

$$w = T(v) = T(\alpha_1 v_1 + \cdots + \alpha_n v_n) == \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n)$$

Así todo vector $w \in Im(T)$ es combinación lineal de $\{T(v_1), \dots, T(v_n)\} \subseteq Im(T)$ es decir que $\{T(v_1), \dots, T(v_n)\}$ es un generador de la $Im(T)$.

ii) Sean $\alpha_1, \dots, \alpha_n \in K$ tales que

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \vec{0}.$$

Aplicando T y usando la linealidad $\alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) = \vec{0}$ y siendo $\{T(v_1), \dots, T(v_n)\}$ un conjunto L.I. resulta que $\alpha_1 = \ldots = \alpha_n = 0$. Esto prueba que $\{v_1, \dots, v_n\}$ es L.I. □

PROPOSICIÓN 7.11. *Sean V y W espacios vectoriales sobre un mismo cuerpo con $dim(V) = n$ y $T : V \to W$ una transformación lineal.*

Si $\{e_1, \dots, e_d\}$ es una base del $N(T)$ y $\{e_1, \dots, e_d, e_{d+1}, \dots, e_n\}$ es una base de V entonces $\{T(e_{d+1}), \dots, T(e_n)\}$ es una base de la $Im(T)$.

Demostración. Probaremos que $\{T(e_{d+1}), \dots, T(e_n)\}$ es base de $Im(T)$.

1) Es L.I. pues $\alpha_{d+1} T(e_{d+1}) + \cdots + \alpha_n T(e_n) = \vec{0}$

$$\Rightarrow T(\alpha_{d+1} e_{d+1} + \cdots + \alpha_n e_n) = \vec{0} \Rightarrow \alpha_{d+1} e_{d+1} + \cdots + \alpha_n e_n \in N(T)$$

$\alpha_{d+1}e_{d+1} + \cdots + \alpha_n e_n$ es C.L de la base de $N(T)$.

$$\alpha_{d+1}e_{d+1} + \cdots + \alpha_n e_n = \alpha_1 e_1 + \cdots + \alpha_d e_d$$

$$-\alpha_1 e_1 - \cdots - \alpha_d e_d + \alpha_{d+1}e_{d+1} + \cdots + \alpha_n e_n = \vec{0}.$$

Pero $\{e_1, \ldots, e_n\}$ es L.I. porque es base de V; luego todos los coeficientes α_i son ceros. Hemos probado que $\{T(e_{d+1}), \ldots, T(e_n)\}$ es L.I.

2) $\{T(e_{d+1}), \ldots, T(e_n)\}$ generan la $Im(T)$ pues:

Sea $w \in Im(T) \Rightarrow \exists v \in V$ tal que $T(v) = w$

$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, porque $\{e_1, \ldots, e_n\}$ es base de V,

$$T(v) = w = T(\alpha_1 e_1 + \cdots + \alpha_n e_n) =$$

$$= \alpha_1 T(e_1) + \cdots + \alpha_d T(e_d) + \alpha_{d+1} T(e_{d+1}) + \cdots + \alpha_n T(e_n) =$$

$$= \alpha_{d+1} T(e_{d+1}) + \cdots + \alpha_n T(e_n).$$

Luego todo vector $w \in Im(T)$ es C.L. de $\{T(e_{d+1}), \ldots, T(e_n)\}$, es decir, $\{T(e_{d+1}), \ldots, T(e_n)\}$ genera a $Im(T)$. □

TEOREMA 7.12 (De las dimensiones). *Sean V y W espacios vectoriales sobre un mismo cuerpo K, con $dim(V) = n$ y $T : V \to W$ una transformación lineal. Entonces $dim\ (V) = dim\ (N(T)) + dim\ (Im(T))$.*

Demostración. Como N(T) es subespacio de V, si $d = dim(N(T))$, tenemos $0 \leqslant d \leqslant n$.

Discutiremos tres casos.

<u>1er. Caso</u>. d=n. Como $dim(N(T)) = V$ obtenemos $N(T) = V$.

Entonces $T(v) = \vec{0}\ \ \forall v \in V$, es decir $Im(T) = \vec{0}$ y $dim(Im(T)) = 0$, cumpliéndose la tesis en este caso.

<u>2do. Caso</u>. $0 < d < n$. El resultado es una consecuencia directa de la proposición anterior.

<u>3er. Caso</u>. $d = dim(N(T)) = 0$. Luego $N(T) = \left\{\vec{0}\right\}$.

Sea $\{v_1, \ldots, v_n\}$ una base de V. Por ser $\{v_1, \ldots, v_n\}$ un generador de V, por la Proposición 7.10(i) $\{T(v_1), \ldots, T(v_n)\}$ un generador de la $Im(T)$.

Veamos que también es L.I. En efecto

$$\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n) = \vec{0} \Leftrightarrow T(\alpha_1 v_1 + \ldots + \alpha_n v_n) = \vec{0}$$
$$\Leftrightarrow \alpha_1 v_1 + \ldots + \alpha_n v_n \quad \in N(T) \Leftrightarrow \alpha_1 v_1 + \ldots + \alpha_n v_n = \vec{0}$$
$$\Leftrightarrow \alpha_1 = \ldots. = \alpha_n = 0$$

por ser $\{v_1, \ldots, v_n\}$ base de V. Por lo tanto $\{T(v_1), \ldots, T(v_n)\}$ es base de la $Im(T)$.

Así $dim(Im(T)) = n$ y se cumple que $dim\ V = dim\ N(T) + dim\ Im(T)$.

□

7.4. Transformaciones lineales inyectivas y sobreyectivas.

Sean V y W espacios vectoriales sobre un mismo cuerpo K y $T : V \to W$ una transformación lineal.

Recordemos que T es inyectiva cuando $T(v_1) = T(v_2)$ *entonces* $v_1 = v_2$; o equivalentemente $v_1 \neq v_2$ entonces $T(v_1) \neq T(v_2)$.

PROPOSICIÓN 7.13. *Son equivalentes las siguientes afirmaciones:*
i) T es inyectiva.
ii) Para todo conjunto A linealmente independiente de V se cumple que $T(A)$ es linealmente independiente en W.
iii) $N(T) = \left\{\vec{0}\right\}$.

DEMOSTRACIÓN. $i) \Rightarrow ii)$ Sabemos que T es inyectiva. Sea $A = \{v_1, \ldots, v_n\}$ un conjunto L.I. de V. Probemos que $T(A) = \{T(v_1), \ldots, T(v_n)\}$ es L.I. en W.

Sean $\alpha_1, \ldots., \alpha_n \in K$ tales que $\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n) = \vec{0}_w$. Debemos probar que $\alpha_1 = \ldots. = \alpha_n = 0$. Siendo T lineal obtenemos
$T(\alpha_1 v_1 + \ldots + \alpha_n v_n) = \vec{0}_w$.

Así $T(\alpha_1 v_1 + \ldots + \alpha_n v_n) = T\left(\vec{0}_v\right)$ y como T es inyectiva se obtiene que

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = \vec{0}_v$$

y siendo $A = \{v_1, \ldots, v_n\}$ L.I., se cumple que $\alpha_1 = \ldots. = \alpha_n = 0$.

$ii) \Rightarrow iii)$ Supongamos, por absurdo, que $N(T) \neq \left\{\vec{0}_v\right\}$.Entonces existe $v \in N(T)$ con $v \neq \vec{0}_v$.

Luego $A = \{v\}$ es L.I. en V $\underset{porhiptesis}{\Rightarrow} T(A) = T(v)$ es L.I. en W y por ii) $T(v) = \vec{0}_w$ es L.I. Absurdo.

$iii) \Rightarrow i) Si T(v_1) = T(v_2) \ \Rightarrow T(v_1) - T(v_2) = \vec{0}_w$ luego $T(v_1 - v_2) = \vec{0}_w \Rightarrow v_1 - v_2 \ \in \ N(T)$ por iii) $v_1 - v_2 = \vec{0}_v$ es decir $v_1 = v_2$. □

Recordemos ahora que T es sobreyectiva cuando $\forall w \ \in W$ existe $v \in V$ con $T(v) = w$, o equivalentemente $Im(T) = W$.

PROPOSICIÓN 7.14. *Son equivalentes las siguientes afirmaciones: i) T es sobreyectiva*
ii) Para todo A *generador de* V *se cumple que* $T(A)$ *es un generador de* W
iii) Existe un generador A_0 *de* V *tal que* $T(A_0)$ *es generador de* W.

Demostración. $i) \Rightarrow ii)$ Sea A un generador de V. Por la Proposición 7.10 se cumple que $T(A)$ es un generador de la $Im(T)$.Pero $Im(T) = W$, por ser T sobreyectiva. Así $T(A)$ es un generador de W.

$ii) \Rightarrow iii)$ Inmediato. $iii) \Rightarrow i)$ Sabemos que existe $A_0 = \{v_1, \ldots, v_n\}$ generador de V tal que $T(A_0) = \{T(v_1), \ldots, T(v_n)\}$ es generador de W.Sea $w \ \in W$, como $T(A_0)$ es generador de W, existen $\alpha_1, \ldots, \alpha_n \in K$ tales que

$$w = \alpha_1 T(v_1) + \ldots + \alpha_n T(v_n)$$

Siendo T lineal se cumple que

$$w = T(\alpha_1 v_1 + \ldots . + \alpha_n v_n)$$

Hemos probado que dado $w \ \in W$ existe $v_0 = \alpha_1 v_1 + \ldots . + \alpha_n v_n \ \in V$ tal que

$w = T(v_0)$ luego T es sobreyectiva. □

Recordemos por último que T es biyectiva cuando es inyectiva y sobreyectiva.

PROPOSICIÓN 7.15. *Son equivalentes las siguientes afirmaciones:*
i) T es biyectiva
ii) Para toda base B del espacio V se cumple que $T(B)$es una base de W
iii) Existe una base B_0 del espacio V tal que $T(B_0)$ es una base de W.

DEMOSTRACIÓN.

$$i) \Rightarrow ii) T \text{ es lineal y biyectiva } \Leftrightarrow \begin{cases} \text{T es lineal e inyectiva } (1) \\ \text{T es lineal y sobreyectiva } (2) \end{cases}$$

Por otro lado, B es una base de $V \Leftrightarrow \begin{cases} B \text{ es L.I. } (1') \\ B \text{ es generador de V } (2') \end{cases}$

De (1) y (1') aplicando la Proposición 7.13 se obtiene que $T(B)$ es L.I.(1")

De (2) y (2') aplicando la Proposición 7.14 se obtiene que $T(B)$ es generador de W (2") De (1") y (2") se tiene que $T(B)$ es una base de W.

$ii) \Rightarrow iii)$ Inmediato. $iii) \Rightarrow i)$ Sabemos que existe B_0 es base de V tal que $T(B_0)$ es base de W, en particular existe B_0 es generador de V tal que $T(B_0)$ es generador de W, luego por la proposición 7.14 T es sobreyectiva.Probemos que T es inyectiva, lo cual es equivalente a probar que $N(T) = \left\{\vec{0}\right\}$, si $v \in N(T) \subseteq V$por ser $B_0 = \{v_1, \ldots, v_n\}$ base de V se cumple que existen $\alpha_1, \ldots, \alpha_n \in K$ tales que $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$.

Luego usando la linealidad de T se obtiene que $T(v) = \alpha_1 T(v_1) + \ldots + \alpha_n T(v_n)$. Pero como $v \in N(T)$, se tiene que $T(v) = \vec{0}$

$$\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n) = \vec{0}$$

Pero $T(B_0) = \{T(v_1), \ldots, T(v_n)\}$ es L.I. (por ser base) entonces $\alpha_1 = \ldots = \alpha_n = 0$ Así $v = \vec{0}$. $\square$

Recordemos que T es biyectiva si y solo si T es invertible, es decir, que existe la función inversa T^{-1}.

PROPOSICIÓN 7.16. *Sean V, W espacios vectoriales sobre un mismo cuerpo y $T : V \to W$ una transformación lineal y biyectiva.*

Entonces la función inversa $T^{-1} : W \to V$ es una transformación lineal.

DEMOSTRACIÓN. Sean w_1 y $w_2 \in W$ y λ y $\mu \in K$.

Observando que $TT^{-1}(w) = w$ y que T es lineal,

$$\lambda\, w_1 + \mu\, w_2 = \lambda TT^{-1}(w_1) + \mu TT^{-1}(w_2) =$$

$$= T\left(\lambda T^{-1}(w_1) + \mu T^{-1}(w_2)\right)$$

Aplicando T^{-1} a ambos miembros

$$T^{-1}(\lambda w_1 + \mu\, w_2) = \lambda T^{-1}(w_1) + \mu T^{-1}(w_2).$$

□

7.5. Isomorfismos entre espacios vectoriales

DEFINICIÓN 7.6. Sean V, W espacios vectoriales sobre un mismo cuerpo K.
a) Diremos que una transformación lineal $T : V \to W$ biyectiva es un isomorfismo.
b) Diremos que V y W son isomorfos cuando existe un isomorfismo que lleva V en W.

OBSERVACIÓN 7.17. De acuerdo a la proposición 7.16 si $T : V \to W$ es un isomorfismo entonces $T^{-1} : W \to V$ también lo es.

PROPOSICIÓN 7.18. *Sean V y W dos espacios vectoriales de dimensión finita sobre un mismo cuerpo K. Entonces las proposiciones a) y b) son equivalentes:*

a) dim (V) = dim (W)
b) V es isomorfo a W.

DEMOSTRACIÓN. $a) \Rightarrow b)$ Sea $dim\,(V) = dim\,(W) = n$. Sean $A = \{v_1\,, \ldots\,, v_n\}$ base de V y $B = \{w_1\,, \ldots\,, w_n\}$ base de W. Luego definimos $T : V \to W$ lineal tal que $T\,(v_i) = w_i$ (T está "bien definida"por el teorema 7.4). Como existe una base A de V tal que $T\,(A) = B$ es base de W, T es un isomorfismo por la proposición 7.15.

$b) \Rightarrow a)$ Como V es isomorfo a W, existe un isomorfismo: $T : V \to W$. Por el teorema de las dimensiones se tiene que $dim\,(V) = dim\,(N\,(T)) + dim\,(Im\,(T))$. Por ser T inyectiva $dim\,(N\,(T)) = 0$.Por ser T sobreyectiva $Im\,(T) = W$. Luego $dim\,(V) = dim\,(W)$. □

COROLARIO 7.19. *Si V es un espacio vectorial, con $dim\,(V) = n$, sobre el cuerpo **K**, entonces V es isomorfo a K^n.*

DEMOSTRACIÓN. En efecto el espacio vectorial K^n sobre el cuerpo K tiene dimensión n. Así $dim\,(V) = dim\,(K^n) = n \Rightarrow V y K^n$ son isomorfos. □

Las coordenadas como transformación lineal

Sea V un espacio vectorial de dimensión n sobre el cuerpo K. Fijemos una base $B = \{v_1\,, \ldots\,, v_n\}$ del espacio V, entonces $\forall v \in V$, existen y son únicos $\alpha_1, \ldots, \alpha_n \in K$ tales que $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$.

Luego definimos $coord_B : V \to K^n$ tal que

$$coord_B\,(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

(es decir que a cada vector $v \in V$ le asociamos sus coordenadas en la base B).

PROPOSICIÓN 7.20. *$coord_B$ es un isomorfismo entre V y $\mathbb{K}^n$.*

DEMOSTRACIÓN. 1) $coord_B$ es lineal. Sean $v, w \in V$, luego existen y son únicos $\alpha_1, \ldots, \alpha_n \in K$ y $\beta_1, \ldots, \beta_n \in K$ tales que

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n \Rightarrow coord_B(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

$$w = \beta_1 v_1 + \ldots + \beta_n v_n \Rightarrow coord_B(w) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

Luego $v + w = (\alpha_1 + \beta_1) v_1 + \ldots. + (\alpha_n + \beta_n) v_n$

$$\Rightarrow \ coord_B(v + w) = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix}$$

Así $coord_B(v) + \ coord_B(w) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} =$ $coord_B(v + w)$

Por otro lado si $\lambda \in K$, se tiene que $\lambda v = \lambda (\alpha_1 v_1 + \ldots + \alpha_n v_n) =$

$$= \lambda \alpha_1 v_1 + \ldots + \lambda \alpha_n v_n \Rightarrow coord_B(\lambda v) = \begin{pmatrix} \lambda \alpha_1 \\ \vdots \\ \lambda \alpha_n \end{pmatrix}$$

Así $\lambda \ coord_B(v) = \lambda \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \lambda \alpha_1 \\ \vdots \\ \lambda \alpha_n \end{pmatrix} = coord_B(\lambda v)$

2) Como $dim(V) = dim(K^n) = n$ para probar que $coord_B$ es biyectiva, alcanza con probar que $coord_B$ es inyectiva.

Como $coord_B(v) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow v = 0v_1 + \ldots + 0v_n \Leftrightarrow v = \vec{0}$

$N(coord_B) = \left\{\vec{0}\right\} \Rightarrow coord_B$ es inyectiva. □

7.6. Matriz asociada a una transformación lineal.

Sean V, W espacios vectoriales de dimensión finita sobre un mismo cuerpo K y $T : V \to W$ una transformación lineal. Supongamos que $A = \{v_1, v_2, \ldots, v_n\}$ es una base de V y $B = \{w_1, w_2, \ldots, w_m\}$ es una base de W. Ahora bien, $T(v_1), T(v_2), \ldots, T(v_n)$ son vectores de W y por lo tanto cada uno de ellos es una combinación lineal de los vectores de la base B:

$$\begin{array}{l} T(v_1) = a_{11}w_1 + a_{12}w_2 + \ldots + a_{1m}w_m \\ T(v_2) = a_{21}w_1 + a_{22}w_2 + \ldots + a_{2m}w_m \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ T(v_n) = a_{n1}w_1 + a_{n2}w_2 + \ldots + a_{nm}w_m \end{array}$$

En otras palabras

$$coord_B(T(v_1)) = \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1m} \end{pmatrix},$$

$$coord_B(T(v_2)) = \begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2m} \end{pmatrix}, \ \ldots, \ coord_B(T(v_n)) = \begin{pmatrix} a_{n1} \\ a_{n2} \\ \vdots \\ a_{nm} \end{pmatrix}$$

Luego definimos:

DEFINICIÓN 7.7. Se llama representación matricial de T en las bases A y B o matriz asociada a T en las bases A y B, a la matriz que representaremos por ${}_B\left((T)\right)_A$ y cuya i-ésima columna son las coordenadas del vector $T\left(v_i\right)$ en la base B.

Esto es

$${}_B\left((T)\right)_A = \begin{pmatrix} [coord_B T\left(v_1\right)], & [coord_B T\left(v_2\right)], & \ldots, & [coord_B T\left(v_n\right)] \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

Nota. La matriz asociada se obtiene çolgando" las coordenadas respecto a la base de W de los vectores de la base de V. Obsérvese que los subíndices de esta matriz están colocados de manera distinta a la habitual; alcanzaría con cambiar las denominaciones de los subíndices de las coordenadas de $T(v_j)$ para obtener la forma habitual.

EJEMPLO 7.14. Consideremos la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que

$T\left(x, y\right) = \left(4x - 2y, 2x + y, x + y\right) \quad \forall x \in \mathbb{R}^2$, y las bases canónicas $A = \{(1,0), (0,1)\}$ de $\mathbb{R}^2$ y $B = \{(1,0,0), (0,1,0), (0,0,1)\}$ de $\mathbb{R}^3$. Para hallar la matriz asociada a T en dichas bases

1) Hallamos las imágenes de los vectores de la base A

$$T\left(1,0\right) = \left(4,2,1\right)$$

$$T\left(0,1\right) = \left(-2,1,1\right)$$

2) Calculamos las coordenadas de estos en la base B

$$T\left(1,0\right) = \left(4,2,1\right) = 4\left(1,0,0\right) + 2\left(0,1,0\right) + 1\left(0,0,1\right)$$

$$\Rightarrow coord_B\left(T\left(1,0\right)\right)=\begin{pmatrix}4\\2\\1\end{pmatrix}$$

$$T\left(0,1\right)=\left(-2,1,1\right)=-2\left(1,0,0\right)+1\left(0,1,0\right)+1\left(0,0,1\right)$$

$$\Rightarrow coord_B\left(T\left(0,1\right)\right)=\begin{pmatrix}-2\\1\\1\end{pmatrix}.\text{ Luego }{}_B\left(\left(T\right)\right)_A=\begin{pmatrix}4&-2\\2&1\\1&1\end{pmatrix}.$$

EJEMPLO 7.15. Consideremos la transformación lineal $T:P_2\to\mathbb{R}^2$ tal que $T\left(p\right)=\left(2a+b,a+b+4c\right)\quad\forall p\in P_2$ tal que $p\left(t\right)=at^2+bt+c\ \ \forall t\in\mathbb{R}$ y las bases $A=\{p_1,p_2,p_3\}$ de P_2 donde $p_1:p_1\left(t\right)=t^2,\ p_2:p_2\left(t\right)=t\ \ ,p_3:p_3\left(t\right)=1\ \ \forall t\in\mathbb{R}$; y $B=\{\left(1,1\right),\left(1,0\right)\}$ de $\mathbb{R}^2$.

Para hallar la matriz asociada a T en dichas bases,

1) Hallamos las imágenes de los vectores de la base A

$$T\left(p_1\right)=\left(2,1\right)$$

$$T\left(p_2\right)=\left(1,1\right)$$

$$T\left(p_3\right)=\left(0,4\right)$$

2) Calculamos las coordenadas de estos en la base B

$$T\left(p_1\right)=\left(2,1\right)=1\left(1,1\right)+1\left(1,0\right)\Rightarrow coord_B\left(T\left(p_1\right)\right)=\begin{pmatrix}1\\1\end{pmatrix}$$

$$T\left(p_2\right)=\left(1,1\right)=1\left(1,1\right)+0\left(1,0\right)\Rightarrow coord_B\left(T\left(p_2\right)\right)=\begin{pmatrix}1\\0\end{pmatrix}$$

$$T\left(p_3\right)=\left(0,4\right)=4\left(1,1\right)-4\left(1,0\right)\Rightarrow coord_B\left(T\left(p_3\right)\right)=\begin{pmatrix}4\\-4\end{pmatrix}$$

Luego ${}_B\left(\left(T\right)\right)_A=\begin{pmatrix}1&1&4\\1&0&-4\end{pmatrix}$.

EJEMPLO 7.16. Consideremos la transformación lineal $T : P_2 \to P_2$ tal que $T(p) = p'$ y la base canónica de P_2 $A = \{p_0, p_1, p_2\}$ donde $p_i(t) = t^i, \quad i = 0, 1, 2$.
1) Hallemos las imágenes de los vectores de la base A:

$$T(p_0) = 0,\ T(p_1) = p_0,\ T(p_2) = 2p_1$$

2)Calculemos las coordenadas de estos en la base A

$$T(p_0) = 0 = 0p_0 + 0p_1 + 0p_2 \Rightarrow coord_A(T(p_0)) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$T(p_1) = p_0 = 1p_0 + 0p_1 + 0p_2 \Rightarrow coord_A(T(p_1)) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$T(p_2) = 2p_1 = 0p_0 + 2p_1 + 0p_2 \Rightarrow coord_A(T(p_2)) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

Luego $_A((T))_A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$

OBSERVACIÓN 7.21. Si *dim(V)=n* y *dim(W)=m* la matriz asociada tiene dimensión $m \times n$.

OBSERVACIÓN 7.22. La matriz $_B((T))_A$ como recién vimos, queda completamente determinada conocidas la transformación lineal T y las bases A y B del dominio y codominio respectivamente.

Recíprocamente, dada una matriz M de tamaño $m \times n$ y dos bases A y B de los espacios V y W respectivamente, queda completamente determinada una transformación lineal T tal que $_B((T))_A$=*M*.

En efecto, al conocer la matriz M, sus columnas son las coordenadas en la base B de las imágenes de dos vectores de la base A.

Esto nos permite conocer a las imágenes de los vectores de la base A, y por el Teorema 7.4, esto es suficiente para determinar T.

EJEMPLO 7.17. Hallar la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$ sabiendo que ${}_B\left((T)\right)_A = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & 2 \end{pmatrix}$ donde $A = \{(1,0,1), (2,0,0), (0,1,0)\}$ es base de $\mathbb{R}^3$ y $B = \{(2,-1), (0,2)\}$ es base de $\mathbb{R}^2$

De acuerdo a la definición de matriz asociada

$$coord_B\left(T(1,0,1)\right) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$coord_B\left(T(2,0,0)\right) = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

$$coord_B\left(T(0,1,0)\right) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Luego

$$T(1,0,1) = 2(2,-1) + 1(0,2) = (4,0)$$

$$T(2,0,0) = 3(2,-1) + 0(0,2) = (6,-3)$$

$$T(0,1,0) = -1(2,-1) + 2(0,2) = (-2,5)$$

Ahora bien siendo $A = \{(1,0,1), (2,0,0), (0,1,0)\}$ una base de $\mathbb{R}^3$ se cumple que $\forall (x,y,z) \in \mathbb{R}^3$ que

$$(x,y,z) = z(1,0,1) + \left(\frac{x-z}{2}\right)(2,0,0) + y(0,1,0)$$

Luego por la linealidad de T

$$T(x,y,z) = z\ T(1,0,1) + \left(\frac{x-z}{2}\right)\ T(2,0,0) + y\ T(0,1,0) =$$

$$z(4,0) + \ \frac{x-z}{2} \cdot (6,-3) + y(-2,5) = \left(3x - 2z - 2y, -\frac{3}{2}x + 5y + \frac{3}{2}z\right)$$

Así $T:\mathbb{R}^3\to\mathbb{R}^2$ es tal que

$$T\left(x,y,z\right)=\left(3x-2z-2y,-\frac{3}{2}x+5y+\frac{3}{2}z\right)$$

PROPOSICIÓN 7.23. *Considere los K-espacios vectoriales U, V y W, con dim(U)=s, dim(V)=n y dim(W)=t, y las transformaciones lineales $S:U\to V$y$T:V\to W$. Sean $A=\{u_1,\ldots,u_s\}$, $B=\{v_1,\ldots,v_n\}$ y $C=\{w_1,\ldots,w_t\}$ bases de U, V y W respectivamente. Entonces la matriz asociada a la composición $T\circ S$ es el producto de las matrices asociadas. Es decir*

$${}_C\left((T\circ S)\right)_A={}_C\left((T)\right)_B\,{}_B\left((S)\right)_A.$$

DEMOSTRACIÓN. Sea ${}_C\left((T)\right)_B=(a_{i\,j})$, ${}_B\left((S)\right)_A=(b_{j\,k})$ y ${}_C\left((T)\right)_B\,{}_B\left((S)\right)_A=(c_{i\,j})$ con $1\leqslant i\leqslant t,\ \ 1\leqslant j\leqslant n,\ \ 1\leqslant k\leqslant s.$

Por definición de producto de matrices $c_{i\,k}=\sum\limits_{j=1}^{n}a_{i\,j}b_{j\,k}$.

Calculemos ${}_C\left((T\circ S)\right)_A$. Sea $u_k\in A$.

$$\left(T\circ S\right)\,(u_k)=T\left(S\left(u_k\right)\right)=T\left(\sum_{j=1}^{n}b_{j\,k}v_j\right)=$$

$$\sum_{j=1}^{n}b_{j\,k}T\left(v_j\right)=\sum_{j=1}^{n}b_{j\,k}\sum_{i=1}^{t}a_{i\,j}w_i=\sum_{i=1}^{t}\left(\sum_{j=1}^{n}a_{i\,j}b_{j\,k}\right)w_i=\sum_{i=1}^{t}c_{i\,k}w_i.$$

Resulta, por definición de matriz asociada ${}_C\left((T\circ S)\right)_A=(c_{i\,j})$. □

OBSERVACIÓN 7.24. Sea $T\ :\ V\to W$ un isomorfismo, B y B' bases de V y W respectivamente, y sea $T^{-1}\ :\ W\to V$ la inversa de T.

Como $T\circ T^{-1}\ =\ id_W$

$$\Rightarrow\quad {}_{B'}\left((T)\right)_B\cdot{}_B\left(\left(T^{-1}\right)\right)_{B'}\ =\ B'\left((id_W)\right)_{B'}\ =I$$

$$\text{y de } T^{-1}\circ T\ =\ id_V$$

$$\Rightarrow\quad {}_B\left(\left(T^{-1}\right)\right)_{B'}\cdot{}_{B'}\left((T)\right)_B\ =\ {}_B\left((id_V)\right)_B\ =I$$

y por lo tanto la matriz asociada a la transformación inversa es la inversa de la matriz asociada a la transformación. Es decir si ${}_{B'}((T))_B = A \Rightarrow {}_B\left((T^{-1})\right)_{B'} = A^{-1}$

PROPOSICIÓN 7.25. *Sean dos transformaciones lineales de espacios vectoriales sobre un cuerpo K, $T : V \to W$ y $S : V \to W$ y α un escalar de K. Sea $B = \{v_1, \ldots, v_n\}$ base de V y $E = \{w_1, \ldots, w_m\}$ base de W. Entonces:*

$${}_E((T + S))_B = {}_E((T))_B + {}_E((S))_B, \quad {}_E((\lambda T))_B = \lambda\, {}_E((T))_B,$$

Demostración. Sean $A = (a_{ij}) = {}_E((T))_B \quad B = (b_{ij}) = {}_E((S))_B$. Entonces: $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad S(v_j) = \sum_{i=1}^{m} b_{ij} w_i$. De donde $(T + S)(v_j) = T(v_j) + S(v_j) = \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i = \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i \Rightarrow$ ${}_E((T + S))_B = (a_{ij} + b_{ij}) = A + B$. Además $(\lambda T) v_j = \lambda T(v_j) = \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \lambda a_{ij} w_i \Rightarrow {}_E((\lambda T))_B = (\lambda a_{ij}) = \lambda A$. □

OBSERVACIÓN 7.26. Demostración de la propiedad asociativa del producto de matrices usando transformaciones lineales.
Se trata de ver que si A, B, C son matrices, entonces:

$$(AB)\,C = A\,(BC)$$

si las dimensiones relativas de las matrices hacen que la expresión tenga sentido.

Ya vimos (observación 7.22) que fijando bases hay una correspondencia biunívoca entre matrices y transformaciones lineales, por lo tanto, el producto anterior se puede interpretar como un producto (o composición) de transformaciones lineales.

Consideremos los espacios vectoriales K^m, K^n, K^p, K^r y sus respectivas bases $E_m, E_n, E_p\ E_r$. Sean $T : K^n \to K^m$ tal que ${}_{E_m}((T))_{E_n} =$

$A_{m\times n}$, $S : K^p \to K^n$ tal que ${}_{E_n}((S))_{E_p} = B_{n\times p}$ y $L : K^r \to K^p$ tal que ${}_{E_p}((L))_{E_r} = C_{p\times r}$.

Luego, en virtud de que la composición de funciones es asociativa tenemos que:

$$(A\,.\,B)\,.\,C = [((T))\,.\,((S))]\,.\,((L)) = ((T\circ S))\,.\,((L)) =$$

$$= (((T\circ S)\;L)) = ((T\circ(S\circ L))) =$$

$$= ((T))\,.\,((S\,L)) = ((T))\,[((S))\,.\,((L))] = A.\,(B.C)$$

(Hemos omitido para simplificar la notación, las bases en cuestión.)

TEOREMA 7.27. *Sean V, W espacios vectoriales sobre un mismo cuerpo K, $A = \{v_1, v_2, \dots, v_n\}$ y $B = \{w_1, w_2, \dots, w_m\}$ bases ordenadas de V y W respectivamente; y $T : V \to W$ una transformación lineal. Entonces se cumple que ${}_B((T))_A\ coord_A(v) = coord_B(T(v))$*

Demostración. Usaremos las siguientes notaciones

$${}_B((T))_A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \text{ y } coord_A(v) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ siendo}$$

$${}_B((T))_A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{y} \quad B = \{w_1, w_2, \dots, w_m\},$$

por definición de matriz asociada

$$\text{(I)}\quad \begin{array}{l} T(v_1) = a_{11}w_1 + a_{21}w_2 + \dots + a_{m1}w_m \\ T(v_2) = a_{12}w_1 + a_{22}w_2 + \dots + a_{m2}w_m \\ \vdots \\ T(v_n) = a_{1n}w_1 + a_{2n}w_2 + \dots + a_{mn}w_m \end{array}$$

y siendo $coord_A(v) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ y $A = \{v_1, v_2, \ldots, v_n\}$ tenemos que $v = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n$. Luego aplicando T y usando la linealidad

$$\text{(II) } T(v) = x_1\ T(v_1) + x_2\ T(v_2) + \ldots + x_n\ T(v_n)$$

Sustituimos (I) en (II) obteniendo

$$\begin{aligned} & T(v) = \\ & x_1(a_{11}w_1 + a_{21}w_2 + \ldots + a_{m1}w_m) + x_2(a_{12}w_1 + a_{22}w_2 + \ldots + a_{m2}w_m) \\ & \qquad + \ldots + x_n(a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{mn}w_m) = \\ & = (x_1a_{11} + x_2a_{12} + \ldots + x_na_{1n})\,w_1 + (x_1a_{21} + x_2a_{22} + \ldots + x_na_{2n})\,w_2 \\ & \qquad + \ldots + (x_1a_{m1} + x_2a_{m2} + \ldots + x_na_{mn})\,w_m. \end{aligned}$$

$$\text{Luego } coord_B(T(v)) = \begin{pmatrix} x_1a_{11} + x_2a_{12} + \ldots. + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \ldots. + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \ldots. + x_na_{mn} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = {}_B((T))_A\ coord_A(v)$$

□

EJEMPLO 7.18. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que ${}_B((T))_A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 0 \\ 3 & 4 & 1 \end{pmatrix}$ siendo $A = B = \{(1,0,0), (1,1,0), (1,1,1)\}$. Hallar $T(2,0,-1)$.

Como $(2,0,-1) = 2(1,0,0) + (-1)(1,1,0) + 1(1,1,1)$

resulta que $coord_A(2,0,-1) = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$

Luego de acuerdo al Teorema 7.27 se cumple que

$$coord_B(T((2,0,-1))) =_B ((T))_A \quad coord_A(2,0,-1) =$$
$$= \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 0 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -3 \end{pmatrix}$$

Así $T(2,0,-1) = 3(1,0,0) + 4(1,1,0) + (-3)(1,1,1) = (4,1,-3)$

7.7. Núcleo e imagen de una matriz.

Dada una matriz $A \in M_{nxm}(K)$, definimos la transformación $T_A : K^m \to K^n$ tal que

$$T_A(x_1,\ldots,x_m) = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \qquad \forall (x_1,\ldots,x_m) \in K^m$$

PROPOSICIÓN 7.28. *(a) T_A es lineal*

(b) Si E_m y E_n son las bases canónicas de K^m y K^n respectivamente entonces ${}_{E_n}((T_A))_{E_m} = A$

Demostración. (Ejercicio). □

DEFINICIÓN 7.8. Sea $A \in M_{nxm}(K)$. Definimos el **núcleo de la matriz** A como el núcleo de la transformación lineal T_A.

$$N(A) \stackrel{def}{=} N(T_A) = \left\{x \in \mathbb{R}^m : T_A\ (x) = \vec{0}\right\} = \left\{x \in \mathbb{R}^m : A\ x = \vec{0}\right\}$$

En otras palabras el $N(A)$ son las soluciones del sistema de ecuaciones homogéneo $A\ x = \vec{0}$.

DEFINICIÓN 7.9. Sea $A \in M_{nxm}(K)$. Definimos la **imagen de la matriz** A como la imagen de la transformación lineal T_A.

$$Im(A) \overset{def}{=} Im(T_A) = \{y \in \mathbb{R}^n : \text{ existe } x \in \mathbb{R}^m \text{ tal que } T_A\ (x) = y\} =$$
$$\{y \in \mathbb{R}^n : \text{ existe } x \in \mathbb{R}^m \text{ tal que } A\ x = y\}$$

En otras palabras la $Im(A)$ son los "términos independientes"$y \in \mathbb{R}^n$ para los cuales el sistema $A\ x = y$ es compatible.

OBSERVACIÓN 7.29. Resulta obvio de las definiciones 7.2. y 7.3. que $N(A)$ es un subespacio de $\mathbb{R}^m$ y $Im(A)$ es un subespacio de $\mathbb{R}^n$ (por serlo $N(T_A)$ e $Im(T_A)$)

PROPOSICIÓN 7.30. *Las columnas de A generan a la* $Im(A)$.

DEMOSTRACIÓN. Sea $T_A : K^m \to K^n$ tal que $T_A(x_1, \ldots, x_m) = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$.

Si $\{e_1, e_2, \ldots, e_m\}$ es la base canónica de $\mathbb{R}^m$. Entonces de acuerdo a la proposición 7.10(i) $\{T_A(e_1), T_A(e_2), \ldots, T_A(e_m)\}$ es un generador de la $Im(T_A) = Im(A)$. Pero $T_A(e_i) = A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = A^{(i)}$ (i-ésima columna de A) Así las columnas de A $\left\{A^{(1)}, \ldots, A^{(m)}\right\}$ constituyen un generador de la $Im(A)$. $\square$

COROLARIO 7.31. *El rango de una matriz A es la dimensión de su imagen, es decir* $r(A) = dim(Im(A))$.

EJEMPLO 7.19. Hallemos bases del núcleo y de la imagen de la matriz $A=\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & -2 \end{pmatrix}$. Sea $(x,y,z)\in N(A)$. Entonces

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\Leftrightarrow\begin{pmatrix} 1 & 01 & \\ 2 & 1 & 1 \\ -1 & 1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\Leftrightarrow$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\Leftrightarrow\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow\begin{cases} x+z=0 \\ y-z=0 \end{cases}\Leftrightarrow\begin{cases} x=-z \\ y=z \\ z \text{ cualquiera} \end{cases}$$

Luego los vectores del núcleo de A son de la forma $(-z,z,z)=z(-1,1,1)$. Así $\{(-1,1,1)\}$ es base del $N(A)$ (¿por qué?)

Por otro lado de acuerdo a la Proposición 7.30 resulta que $\{(1,2,-1),(0,1,1),(1,1,-2)\}$ genera a la $Im(A)$. Como $(1,1,-2)=(1,2,-1)-(0,1,1)$, el generador es L.D. y lo reducimos hallando $\{(1,2,-1),(0,1,1)\}$ que es un generador L.I. –y por tanto una base– de $Im(A)$ (verifique).

EJEMPLO 7.20. Hallar una base del núcleo y una base de la imagen de la matriz

$$A=\begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 \\ 1 & 1 & 3 & -3 \end{pmatrix}.\ \text{Sea}\,(x,y,z,t)\in N(T).\ \text{Entonces}$$

$$A\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\Leftrightarrow\begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 \\ 1 & 1 & 3 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\Leftrightarrow$$

$$\begin{pmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 2 & 2 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow$$

$$\begin{pmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} x - y + z + t = 0 \\ y + z - 2t = 0 \end{cases} \Leftrightarrow \begin{cases} x = -2z + t \\ y = -z + 2t \\ z = \text{ cualquiera} \\ t = \text{ cualquiera} \end{cases}$$

Así los vectores del núcleo de A son de la forma

$$(-2z + t, -z + 2t, z, t) = z\,(-2, 1, 0, 1) + t\,(1, 2, 0, 1)$$

Luego $\{(-2, 1, 0, 1), (1, 2, 0, 1)\}$ es un generador L.I. del $N(A)$ (verifique) y por lo tanto es base.

Por la proposición 7.30 $\{(1, 1, 1), (-1, 0, 1), (1, 2, 3), (1, -1, -3)\}$ es un generador de la $Im(A)$ que claramente es L.D.

Observando que $(1, 2, 3) = 2\,(1, 1, 1) + (-1, 0, 1)$

$$(1, -1, -3) = -1\,(1, 1, 1) + (-2)\,(-1, 0, 1)$$

reducimos el generador anterior hasta obtener $\{(1, 1, 1), (-1, 0, 1)\}$ que es un generador L.I. de la $Im(A)$ y por lo tanto es base.

7.8. Relación entre núcleo e imagen de una transformación lineal y de una matriz.

PROPOSICIÓN 7.32. *Sean V, W espacios vectoriales sobre un mismo cuerpo $A = \{v_1, v_2, \ldots, v_n\}$ una base de V, $B = \{w_1, w_2, \ldots, w_n\}$ una base de W, $T : V \to W$ lineal. Entonces:*

1) Si $v \in N(T) \Rightarrow coord_A(v) \in N\left({}_B((T))_A\right)$
2) Si $x \in N\left({}_B((T))_A\right) \Rightarrow coord_A^{-1}(x) \in N(T)$

Demostración. 1) Si $v \in N(T)$ entonces $T(v) = \vec{0}$. Luego como $T(v)$ y $\vec{0}$ coinciden sus coordenadas en la base B también: $coord_B(T(v)) = coord_B\left(\vec{0}\right)$. Siendo $coord_B : W \to \mathbb{R}^m$ una transformación lineal se cumple: $coord_B\left(\vec{0}\right) = \vec{0}$. Así $coord_B(T(v)) = \vec{0}$. Pero, por el Teorema 7.27: $coord_B(T(v)) = {}_B((T))_A\, coord_A(v)$

Así ${}_B((T))_A\, coord_A(v) = \vec{0}$ esto es $coord_A(v) \in N\left({}_B((T))_A\right)$

2) Si $x = (x_1, \ldots, x_n) \in N\left({}_B((T))_A\right) \Rightarrow {}_B((T))_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \vec{0}$. Entonces, teniendo en cuenta que las columnas de la matriz asociada son las coordenadas de los transformados de la base del dominio se tiene que

$$\left([coord_B(T(v_1))] \ldots [coord_B(T(v_1))]\right) . \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \vec{0}$$

$$\Rightarrow coord_B(T(v_1)).x_1 + \ldots + coord_B(T(v_n)).x_n = \vec{0} \Rightarrow$$

$$\overset{(\text{por ser } coord_B \text{ lineal})}{\Rightarrow} \quad coord_B(x_1T(v_1) + \ldots\ldots + x_nT(v_n)) = \vec{0}$$

$$\overset{(\text{por ser } coord_B \text{ inyectiva})}{\Rightarrow} \quad x_1T(v_1) + \ldots\ldots + x_nT(v_n) = \vec{0}$$

$$\overset{(\text{linealidad de } T)}{\Rightarrow} \quad T(x_1v_1 + \ldots + x_nv_n) = o_V$$

$$\Rightarrow \underbrace{x_1v_1 + \ldots + x_nv_n}_{coord_A^{-1}(x)} \in N(T).$$

$\square$

OBSERVACIÓN 7.33. Con las notaciones de la proposición anterior recordemos que $coord_A : V \to \mathbb{R}^n$ es un isomorfismo entre V y $\mathbb{R}^n$.

Pero la proposición anterior nos asegura que todo vector del $N(T)$, al aplicarle la transformación lineal $coord_A$ se obtiene un vector del $N\left({}_B((T))_A\right)$ y recíprocamente todo vector del $N\left({}_B((T))_A\right)$ al aplicarle la transformación lineal inversa $coord_A^{-1}(x)$ se obtiene un vector del $N(T)$.

Luego $coord_A$ es un isomorfismo entre el $N(T)$ y $N\left({}_B((T))_A\right)$. Esto nos permite enunciar la siguiente

PROPOSICIÓN 7.34. *Sean V, W espacios vectoriales sobre un mismo cuerpo $A = \{v_1, v_2, \ldots, v_n\}$ una base de V, $B = \{w_1, w_2, \ldots, w_n\}$ una base de W, $T : V \to W$ lineal. Entonces:*

1) $N(T)$ y $N\left({}_B((T))_A\right)$ son isomorfos

2) $dim(N(T)) = n - rango({}_B((T))_A)$.

3) Si $\{e_1, e_2, \ldots, e_k\}$ es base del $N(T) \Rightarrow$
$\{coord_A(e_1), coord_A(e_2), \ldots, coord_A(e_k)\}$ es base del $N\left({}_B((T))_A\right)$

4) Si $\{x_1, x_2, \ldots, x_k\}$ es base del $N\left({}_B((T))_A\right) \Rightarrow$
$\left\{coord_A^{-1}(x_1), coord_A^{-1}(x_2), \ldots, coord_A^{-1}(x_k)\right\}$ es base del $N(T)$.

Demostración. (Ejercicio) Sugerencias:

1) Vimos en la observación anterior que $coord_A$ es un isomorfismo entre $N(T)$ y $N\left({}_B((T))_A\right)$.

2) Basta recordar que dos espacios isomorfos tienen la misma dimensión y que si M es una matriz, $dim\,(Im\,(M)) = rango\,(M)$.

3) Los isomorfismos "llevan bases en bases".

4) La inversa de un isomorfismo es un isomorfismo. $\square$

Recomendamos prestar particular atención a los ejemplos que siguen.

EJEMPLO 7.21. Se considera una transformación lineal $T : M_{2x2}(\mathbb{R}) \rightarrow \mathbb{R}^3$ tal que ${}_B((T))_A = \begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 \\ 1 & 1 & 3 & -3 \end{pmatrix}$ siendo

$$A = \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

$$\text{y } B = \{(1,1,1), (0,1,1), (0,0,1)\}.$$

Queremos hallar una base del núcleo de T. Primero calculamos una base del núcleo de la matriz ${}_B((T))_A$. Sea $(x,y,z,t) \in N({}_B((T))_A)$. Entonces

$${}_B((T))_A \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 \\ 1 & 1 & 3 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow$$

$$\begin{pmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 2 & 2 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow$$

$$\begin{pmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} x - y + z + t = 0 \\ y + z + t = 0 \end{cases} \Leftrightarrow \begin{cases} x = -2z - 2t \\ y = -z - t \end{cases}$$

Así los vectores del núcleo de la matriz ${}_B((T))_A$ 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 $N({}_B((T))_A)$.

Luego, de acuerdo con la Proposición 7.34

$$\left\{ coord_A^{-1}(-2,1,1,0), coord_A^{-1}(-2,1,0,1) \right\} =$$

$$= \left\{ -2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} + 1 \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \right.$$

$$\left. -2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\} =$$

$$= \left\{ \begin{pmatrix} -1 & -3 \\ -3 & -2 \end{pmatrix}, \begin{pmatrix} -1 & -2 \\ -3 & -2 \end{pmatrix} \right\}$$

es una base del núcleo de T.

EJEMPLO 7.22. Se considera una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que ${}_B((T))_A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ siendo $A = B = \{(1,0), (1,1)\}$. Queremos hallar una base del núcleo de T.

Primero calculemos una base del núcleo de la matriz ${}_B((T))_A$:

$$(x,y) \in N\left({}_B((T))_A\right) \Leftrightarrow {}_B((T))_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \ x + 2y = 0 \ \Leftrightarrow x = -2y$$

Así los vectores del núcleo de la matriz ${}_B((T))_A$ son de la forma: $(-2y, y) = y(-2,1)$. Entonces $\{(-2,1)\}$ es una base de $N\left({}_B((T))_A\right)$. Luego, de acuerdo a la Proposición 7.34, $\left\{coord_A^{-1}(-2,1)\right\} = \{-2(1,0) + 1(1,1)\} = \{(-1,1)\}$ es una base del $N(T)$.

EJEMPLO 7.23. Se considera una transformación lineal $T : P_2 \to \mathbb{R}^3$ tal que ${}_B((T))_A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -2 \end{pmatrix}$ siendo $A = \{p_0, p_1, p_2\}$ con $p_i : p_i(t) = t^i \ \ (i = 0,1,2)$ y $B = \{(2,0,-1), (0,1,4), (0,0,3)\}$.

Queremos hallar una base del núcleo de T.

Primero calculamos una base del núcleo de la matriz ${}_B((T))_A$

$$(x,y,z) \in N\left({}_B((T))_A\right) \Leftrightarrow {}_B((T))_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow$$

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 3 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x + 2y - z = 0 \\ y + z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 3z \\ y = -z \end{cases}$$

Así los vectores del núcleo de ${}_B((T))_A$ son de la forma:
$(3z, -z, z) = z(3, -1, 1)$. Entonces $\{(3, -1, 1)\}$ es una base del $N\left({}_B((T))_A\right)$. Luego, de acuerdo a la proposición 7.34, una base del núcleo de $N(T)$ es

$$\left\{coord_A^{-1}(3,-1,1)\right\} = \{3p_0 + (-1)p_1 + 1p_2\} = \{3p_0 - p_1 + p_2\}.$$

PROPOSICIÓN 7.35. *Sean V, W espacios vectoriales sobre un mismo cuerpo K, $A = \{v_1, v_2, \ldots, v_n\}$ una base de V. $B = \{w_1, w_2, \ldots, w_m\}$ una base de W, $T : V \to W$ lineal. Entonces*

1) Si $w \in Im(T) \Rightarrow coord_B(w) \in Im\left({}_B((T))_A\right)$
2) Si $y \in Im\left({}_B((T))_A\right) \Rightarrow coord_B^{-1}(y) \in Im(T)$

DEMOSTRACIÓN. 1) Si $w \in Im(T) \Rightarrow$ existe $v \in V$ tal que $T(v) = w$- Luego $coord_B(T(v)) = coord_B(w)$. Pero por el Teorema 7.27 $coord_B(T(v)) = {}_B((T))_A\ coord_A(v)$. Así $coord_B(w) = {}_B((T))_A\ coord_A(v)$

Hemos probado que existe $coord_A(v) \in \mathbb{R}^n$ tal que

$${}_B((T))_A\ coord_A(v) = coord_B(w)$$

$$\Rightarrow coord_B(w) \in Im\left({}_B((T))_A\right)$$

2) Si $y = (y_1, \ldots, y_m) \in Im\left({}_B\left((T)\right)_A\right) \Rightarrow$ existe $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tal que ${}_B\left((T)\right)_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \Rightarrow$

$$\Rightarrow x_1 coord_B\left(T\left(v_1\right)\right) + \ldots + x_n coord_B\left(T\left(v_n\right)\right) = y$$

$$\Rightarrow coord_B\left(x_1 T\left(v_1\right) + \ldots + x_n T\left(v_n\right)\right) = y$$

$$x_1 T\left(v_1\right) + \ldots + x_n T\left(v_n\right) = coord_B^{-1}\left(y\right) \Rightarrow$$
$$T(\underbrace{x_1 v_1 + \ldots + x_n v_n}_{= v_0}) = coord_B^{-1}\left(y\right).$$

Así existe $v_0 \in V$ tal que $T\left(v_0\right) = coord_B^{-1}\left(y\right) \Rightarrow coord_B^{-1}\left(y\right) \in Im\left(T\right)$. □

OBSERVACIÓN 7.36. Al igual que la observación 7.33 concluimos que $coord_B$ <u>es un isomorfismo entre</u> $Im\left(T\right)$ <u>y</u> $Im\left({}_B\left((T)\right)_A\right)$.

PROPOSICIÓN 7.37. *Sean V, W espacios vectoriales sobre un mismo cuerpo $A = \{v_1, v_2, \ldots, v_n\}$ una base de V, $B = \{w_1, w_2, \ldots, w_n\}$ una base de W, $T : V \to W$ lineal. Entonces:*

1) $Im\left(T\right)$ e $Im\left({}_B\left((T)\right)_A\right)$ son isomorfos.

2) $dim\left(Im\left(T\right)\right) = rango\left(A\right)$

3) Si $\{h_1, \ldots, h_k\}$ es una base de la $Im\left(T\right) \Rightarrow$ $\{coord_B\left(h_1\right), coord_B\left(h_2\right), \ldots\ldots, coord_B\left(h_k\right)\}$ es base de la $Im\left({}_B\left((T)\right)_A\right)$

4) si $\{y_1, \ldots, y_k\}$ es una base de la $Im\left({}_B\left((T)\right)_A\right)$ entonces $\left\{coord_B^{-1}\left(y_1\right), coord_B^{-1}\left(y_2\right), \ldots, coord_B^{-1}\left(y_k\right)\right\}$ es base de la $Im\left(T\right)$.

DEMOSTRACIÓN. (Ejercicio). □

EJEMPLO 7.24. Consideremos la transformación lineal del ejemplo 7.21 Hallemos una base de la $Im\,(T)$.

Primero hallemos una base de la imagen de la matriz ${}_B\left((T)\right)_A$.

Sabemos de la proposición 7.30 que las columnas de ${}_B\left((T)\right)_A$:

$$\{(1,1,1)\,,(-1,0,1)\,,(1,2,3)\,,(1,-1,-3)\}$$

son un generador de la $Im\left({}_B\left((T)\right)_A\right)$. Claramente es L.D. Reducimos dicho generador hasta obtener el generador L.I.: $\{(1,1,1)\,,(-1,0,1)\}$. Así $\{(1,1,1)\,,(-1,0,1)\}$ es una base de la $Im\left({}_B\left((T)\right)_A\right)$.

Luego de acuerdo a la proposición 7.37

$$\left\{coord_B^{-1}\,(1,1,1)\,,coord_B^{-1}\,(-1,0,1)\right\}=$$

$$=\{1\,(1,1,1)+1\,(0,1,1)+1\,(0,0,1)\quad,\quad -1\,(1,1,1)+0\,(0,1,1)+1\,(0,0,1)\}$$

$=\{(1,2,3)\,,(-1,-1,0)\}$ es base de la $Im\,(T)$.

EJEMPLO 7.25. Consideremos la transformación lineal del ejemplo 7.22

Hallemos una base de la $Im\,(T)$. Primero hallemos una base de la imagen de la matriz ${}_B\left((T)\right)_A$. Por la proposición 7.30 las columnas de ${}_B\left((T)\right)_A$, $\{(1,2)\,,(2,4)\}$ son un generador de la $Im\left({}_B\left((T)\right)_A\right)$. Claramente es L.D. Lo reducimos y hallamos $\{(1,2)\}$ base de la $Im\left({}_B\left((T)\right)_A\right)$. Luego por la proposición 7.37 $\left\{coord_B^{-1}\,(1,2)\right\}=\{1\,(1,0)+2\,(1,1)\,\}=\{(3,2)\}$ es una base de la $Im\,(T)$.

EJEMPLO 7.26. Consideremos la transformación lineal del ejemplo 7.23

Hallemos una base de la $Im\,(T)$. De acuerdo a la proposición 7.30 $\{(1,0,-1)\,,(2,1,1)\,,(-1,1,-2)\}$ es un generador de la $Im\left({}_B\left((T)\right)_A\right)$.

Como es L.D., lo reducimos hasta obtener $\{(1,0,-1)\,,(2,1,1)\}$ que es una base de la $Im\left({}_B\left((T)\right)_A\right)$. Luego por la proposición 7.37

$$\left\{coord_B^{-1}\,(1,0,-1)\ ,\quad coord_B^{-1}\,(2,1,1,)\right\}=$$

$$\{1\,(2,0,-1)+1\,(0,1,4)+(-1)\,(0,0,3)\ ,\ 2\,(2,0,-1)+1\,(0,1,4)+1\,(0,0,3)\}$$

$\{(2,0,-4)\,,(5,1,5)\}$ es base de la $Im\,(T)$.

7.9. Cambio de base.

Sea V un espacio vectorial de dimensión finita sobre un cuerpo K. Supongamos que $dim\,(V) = n$. Consideremos una base A del espacio V. Hemos visto que todo vector $v \in V$, se puede representar por la n-úpla (vector columna) $coord_A\,(v) \in K^n$.

En esta sección responderemos la siguiente pregunta natural ¿cómo cambian nuestras representaciones si seleccionamos otras bases? Es decir si consideramos otra base A' del espacio V, ¿cómo se relaciona $coord_{A'}\,(v)$ con $coord_A\,(v)$?

Para responder estas preguntas necesitamos la siguiente

DEFINICIÓN 7.10. Sean $A = \{v_1, v_2, \ldots, v_n\}$ y $A' = \{v_1', v_2', \ldots, v_n'\}$ bases del espacio V e $I : V \to V$ la transformación identidad (esto es $I\,(v) = v \;\; \forall v \in V$). Llamaremos matriz de cambio de la base ("vieja") A a la base ("nueva") A' a la matriz: $\;_{A'}\,((I))_A$.

La relación entre las coordenadas está dada por la siguiente

PROPOSICIÓN 7.38. $coord_{A'}\,(v) =_{A'} ((I))_A \;\; coord_A\,(v)$

Demostración. Por Teorema 7.27

$$coord_{A'}\,(I\,(v)) =_{A'} ((I))_A \;\; coord_A\,(v)$$

Pero siendo $I\,(v) = v$, se obtiene la hipótesis. $\square$

PROPOSICIÓN 7.39. *Sea V un espacio vectorial sobre un cuerpo K. A y A' bases de V. $I : V \to V$ la transformación lineal identidad. Entonces:*

1) $_A\,((I))_A = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$ *(matriz identidad)*

2) $_{A'}\,((I))_A$ *es invertible y* $\left[_{A'}\,((I))_A\right]^{-1} =_A ((I))_{A'}$

DEMOSTRACIÓN. (Ejercicio). □

PROPOSICIÓN 7.40. *Sean V y W espacios vectoriales sobre un mismo cuerpo K, A, A' bases de V, B, B' bases de W, $T : V \to W$ una transformación lineal. Entonces*

$$_{B'}((T))_{A'} =_{B'} ((I_W))_B \;\; _B((T))_A \;\; _A((I_V))_{A'}$$

donde $I_V : V \to V$ y $I_W : W \to W$ son las transformaciones lineales identidad en V y W respectivamente.

DEMOSTRACIÓN. Como $I_W \circ T \circ I_V \equiv T$ se tiene, aplicando la proposición 7.23 reiteradamente

$$_{B'}((I_W \circ T \circ I_V))_{A'} =_{B'} ((T))_{A'} \Rightarrow \quad _{B'}((I_W))_B \, _B((T \circ I_V))_{A'} =_{B'} ((T))_{A'}$$

$$\Rightarrow \; _{B'}((I_W))_B \, _B((T))_A \;\; _A((I_V))_{A'} =_{B'} ((T))_{A'} .$$

□

7.10. Operadores y matrices semejantes.

Sea V un espacio vectorial de dimensión finita sobre un cuerpo K. Supondremos que $dim_K(V) = n$.

DEFINICIÓN 7.11. Llamaremos **operador en** V a toda transformación lineal $T : V \to V$.

DEFINICIÓN 7.12. Sean $A \; yB \in M_{nxn}(K)$. Diremos que A y B son semejantes cuando existe $P \in M_{n\times n}(K)$ invertible tal que $B = P^{-1}A\ P$.

EJEMPLO 7.27. La matriz $B = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}$ y la matriz $A = \begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix}$ son semejantes, pues existe $P = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ cuya inversa es $P^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$ tal que $\begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ (verifique!).

PROPOSICIÓN 7.41. *Sean $A,\ B \in M_{nxn}(K)$. A y B son semejantes $\Leftrightarrow$ A y B son representaciones matriciales de un mismo operador T en V, para algún espacio vectorial V.*

Demostración. ($\Rightarrow$) Si consideramos la transformación lineal $T : K^n \to K^n$ tal que $T(x_1, x_2, \ldots, x_n) = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, entonces de acuerdo a la Proposición 7.28(b) tenemos $A =_E ((T))_E$ siendo $E = \{e_1, e_2, \ldots, e_n\}$ la base canónica de K^n.

Por otro lado, siendo A y B semejantes, existe $P \in M_{n\times n}(\mathbb{R})$ invertible tal que:

$$B = P^{-1}AP$$

Pero se cumple que: $P =_E ((I))_H$ donde $H = \{Pe_1, Pe_2, \ldots, Pe_n\}$ y por la Proposición 7.39

$$P^{-1} =_H ((I))_E$$

Así

$$B =_H ((I))_E\, {}_E((T))_E\, {}_E((I))_H$$

Pero por la Proposición 7.40

$${}_H((I))_E\, {}_E((T))_E\, {}_E((I))_H =_H ((T))_H$$

Por lo tanto

$$B =_H ((T))_H$$

($\Leftarrow$) Supongamos que $B =_H ((T))_H$, $A =_E ((T))_E$. Por la Proposición 7.40 sabemos que

$$_H((T))_H =_H ((I))_E \ \ _E((T))_E \ \ _E((I))_H$$

Sea $P =_E ((I))_H$, por la Proposición 7.39 tenemos que $P^{-1} =_H ((I))_E$. Hemos probado entonces que $B = P^{-1}A\ P$, es decir A y B son semejantes.

□

PROPOSICIÓN 7.42. *Sean A y B matrices semejantes en $M_{n\times n}(K)$. Entonces*

$$\begin{array}{l} 1)\, rango\,(A) = rango\,(B) \\ 2)\, traza\,(A) = traza\,(B) \\ 3)\ det\,(A) = det\,(B) \end{array}$$

DEMOSTRACIÓN. 1) Por la proposición anterior, existe un operador lineal en V y bases A y B en dicho espacio tales que $A =_A ((T))_A\, y B =_B ((T))_B$.

Luego $rango\,(A) = rango\,(_A((T))_A) = dim\,(Im\,(T))$

$$rango\,(B) = rango\,(_B((T))_B) = dim\,(Im\,(T))$$

Así $rango\,(A) = rango\,(B)$

2) Existe $P \in M_{nxn}(\mathbb{R})$ invertible tal que $B = P^{-1}A\ P$. Luego

$$traza\,(B) = traza\left(P^{-1}A\ P\right) = traza\left(A\ P\ P^{-1}\right) = traza\,(A\ I) = traza\,(A)$$

(Recordar que $traza(MN)$=$traza(NM)$)

3)$det\,(B) = det\left(P^{-1}A\ P\right) = det\left(P^{-1}\right)\ det\,(A)\ det\,(P) =$
$= det\,(A)\ \ det\left(P^{-1}\right)\ \ det\,(P) = det\,(A)$

(Recordar que $det(MN)$=$det(M).det(N)$ y $det\left(M^{-1}\right) = (det\,(M))^{-1}$).

□

OBSERVACIÓN 7.43. No vale el recíproco de la proposición anterior pues $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ cumplen

$$\begin{array}{l} rango\,(A) = rango\,(B) = 2 \\ traza\,(A) = traza\,(B) = 2 \\ det\,(A) = det\,(B) = 1 \end{array}$$

Sin embargo, no puede existir P invertible tal que $B = P^{-1}AP$ pues $B \neq A$.

7.11. El espacio vectorial $\mathcal{L}(V,W)$

Sean V, W espacios sobre un mismo cuerpo K. Notaremos $\mathcal{L}(V,W)$ al conjunto de todas las transformaciones lineales $T : V \to W$.

PROPOSICIÓN 7.44. *El conjunto $\mathcal{L}(V,W)$ con la suma de transformaciones lineales y el producto por un escalar por una transformación lineal definidas en la sección 2, forman un espacio vectorial.*

DEMOSTRACIÓN. (Ejercicio). $\square$

PROPOSICIÓN 7.45. *Sean V,W espacios vectoriales de dimensión finita sobre un mismo cuerpo K. Supongamos que $dim\,(V) = n$ y $dim\,(W) = m$, entonces $\mathcal{L}(V,W)$ es isomorfo a $M_{mxn}(K)$.*

DEMOSTRACIÓN. Probaremos que existe una transformación lineal

$$F : \mathcal{L}(V,W) \to M_{mxn}(K)$$

biyectiva.

Definimos F de la siguiente manera: Fijemos A base de V y B base de W. Luego, a cada transformación $T \in \mathcal{L}(V,W)$ le asignamos por imagen mediante F a la matriz ${}_B((T))_A \in M_{mxn}(K)$, esto es $F(T) ={}_B ((T))_A$ Probemos que F es lineal. Sean $T_1, T_2 \in \mathcal{L}(V,W)$ y λ, $\mu \in K$.

$$F(\lambda T_1 + \mu T_2) ={}_B ((\lambda T_1 + \mu T_2))_A = \lambda_B ((T_1))_A + \mu_B ((T_2))_A =$$
$$\lambda F(T_1) + \mu F(T_2).$$

Probemos que F es biyectiva.

1) F es inyectiva. Sean $T_1, T_2 \in \mathcal{L}(V,W)$ tales que $T_1 \neq T_2$. Debemos probar que $F(T_1) \neq F(T_2)$.

En efecto siendo A una base de V por la proposición 7.3 existe $v_i \in A$ tal que $T_1(v_i) \neq T_2(v_i)$, entonces $coord_B(T_1(v_i)) \neq coord_B(T_2(v_i))$. Pero de acuerdo a la definición de matriz asociada $coord_B(T_1(v_i))$ y $coord_B(T_2(v_i))$ son las i-ésimas columnas de las matrices ${}_B((T_1))_A$ y ${}_B((T_2))_A$ respectivamente.

Así ${}_B\ ((T_1))\ {}_A \neq_B\ ((T_2))\ {}_A$; esto es $F(T_1) \neq F(T_2)$.

2) F es sobreyectiva. Dada $M \in M_{nxn}(\mathbb{R})$, debemos probar que existe $T \in \mathcal{L}(V,W)$ tal que $F(T) = M$.

Vimos en la observación 7.22 que dada una matriz $M \in M_{nxn}(K)$ existe una transformación lineal $T : V \to W$ tal que ${}_B\ ((T))\ {}_A = M$, es decir que existe $T \in \mathcal{L}(V,W)$, $F(T) = M$. $\square$

COROLARIO 7.46. *Sean V, W espacios vectoriales de dimensión finita sobre un mismo cuerpo K.*

Se cumple que $dim(\mathcal{L}(V,W)) = dim(V).dim(W)$.