Repaso de Transformaciones lineales

Juan Piccini

Contents

1 Espacio Vectorial

El concepto de espacio vectorial $(V, K, +, .)$ (EV de ahora en más), esencialmente consiste de un conjunto no vacío V de objetos (que llamaremos vectores), otro conjunto no vacío K de objetos llamados escalares (que además es un cuerpo numérico, o sea que tiene definidas dos operaciones que cumplen con una serie de propiedades), y dos operaciones $+$ y $.$

1. Una de esas operaciones $(+)$ se llama suma, es una operación que toma dos vectores y entrega un vector.

2. La otra operación $(.)$ se llama producto de un vector con un escalar, y toma un vector, un escalar y entrega un vector.

3. Ambas operaciones cumplen con una serie de propiedades (asociativas, conmutativas, existencias de neutros, etc.) que hacen posible trabajar con ellas.

4. Además hay una propiedad que nos dice como es la interacción entre ambas operaciones (distributiva).

Estas dos operaciones, el sumar vectores o multiplicarlos por escalares (que están en la definición misma de espacio vectorial), pueden condensarse en un concepto fundamental: el de **combinación lineal**, CL de ahora en más.

En efecto, si lo pensamos un poco vemos que lo único que hacemos en un espacio vectorial son combinaciones lineales de vectores.

Luego vienen conceptos tales como el de **conjunto generador**:
$A \underset{g}{\rightarrow} S$ sii todo elemento de S puede fabricarse a partir de los elementos de A, donde "fabricarse" significa poderse obtener como CL de elementos de A.

También el concepto de **conjunto linealmente independiente** (LI de ahora en más) es consecuencia del concepto de CL: un conjunto A es LI cuando ninguno de sus elementos puede obtenerse como CL de los restantes.

Caso contrario el conjunto se llamará **linealmente dependiente** (LD).

Intuitivamente, que un conjunto sea LI nos dice que c/u de sus elementos (vectores) no puede ser "explicado" u obtenido (mediante CL) a partir de los restantes, cada elemento tiene algo que los demás no, perderlo implica perder una cierta información que los demás no podrán suplir.

Luego vienen los conceptos de **base** de un EV (un generador sin elementos redundantes, o sea un generador LI), el darnos cuenta de que todas las bases de un mismo EV tienen algo en común (la cantidad de elementos), y de ahí el concepto de dimensión de un EV.

El concepto de **subespacio vectorial** (SEV de ahora en más) también es hijo directo del concepto de CL: es un subconjunto no vacío que además es cerrado o estable bajo CL (esto es, dado un EV V, $S \subset V$ es un SEV de V sii las CL de elementos de S vuelven a dar elementos de S.

En suma, un EV es una estructura en la que podemos hacer CL de vectores, lo que a su vez permite hacerse una serie de preguntas y definir una serie de conceptos, pero en suma todos ellos vienen como consecuencia de que existe algo llamado CL (de hecho la palabra Lineal en " Algebra Lineal" ya debería ponernos sobre aviso, y hay bibliografía donde a los EV se les llama Espacios Lineales).

2 Funciones entre espacios vectoriales

Los EV son conjuntos en los cuales podemos hacer CL, y prácticamente todos los conceptos y definiciones que se estudian se construyen sobre las CL.

Por otro lado, cuando uno tiene dos conjuntos V, W es natural estudiar funciones $f : V \rightarrow W$ (las que ya vienen con sus propios conceptos tales como inyectividad, sobreyectividad, biyectividad, etc.).

Algunos de esos conceptos pueden carecer de sentido cuando V, W no tienen estructuras compatibles con los mismos.

Por ejemplo, carece de sentido preguntarnos sobre si $f : V \rightarrow W$ es continua si no tenemos el concepto de distancia en V y en W, o si $f : V \rightarrow W$ es creciente o decreciente si no tenemos una estructura de orden en V y W.

Otros conceptos de la teoría de funciones siguen siendo aplicables al caso particular en que el dominio y el codominio de f no son conjuntos cualesquiera sino que son EV.

A nadie sorprenderá que ocupen un lugar destacado aquellas $f : V \rightarrow W$ que "preserven" o "respeten" las CL, que son la columna vertebral o el cimiento sobre el que se estudian los EV.

Estas funciones son precisamente las que llamaremos **Transformaciones Lineales**, TL de ahora en más.

Así como en los cursos de cálculo es habitual denotar a las funciones mediante f, la notación habitual para una TL es T.

3 Transformaciones Lineales

Dados dos EV V, W sobre un mismo cuerpo K, una función $T : V \to W$ se llama TL cuando cumple dos condiciones:

1. T preserva la suma: Cualesquiera sean $u, v \in V$, se cumple que $T(u) + T(v) = T(u + v)$.

2. T preserva el producto: Cualesquiera sean $\alpha \in K, v \in V$, se cumple que $T(\alpha v) = \alpha T(v)$.

En otras palabras, si T es una TL, da lo mismo operar primero en V y luego aplicarle T al resultado que aplicar primero T y operar luego en W.

Aunque la notación no lo explicita, la suma que aplicamos al hacer $T(u+v)$ es la suma de vectores en V, mientras que la suma que aplicamos al hacer $T(u) + T(v)$ es la suma en W, otro tanto para el producto de un escalar y un vector.

Por ejemplo, V podría ser un EV donde los vectores son polinomios, mientras que W podría ser un EV donde los vectores son matrices.

El que T preserve o conserve la suma y el producto puede condensarse en una sola condición: que T preserve las CL, o sea que $T : V \to W$ es una TL sii $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$ cualesquiera sean los escalares $\alpha, \beta \in K$ y los vectores $u, v \in V$, o en otros términos, si "T de una CL es igual a la CL de los T".

3.1 Particularidades de las Transformaciones Lineales

En lo que sigue tendremos en mente una TL $T : V \to W$, donde V, W son dos EV sobre el mismo cuerpo K.

1. Toda TL lleva el vector nulo del dominio en el vector nulo del codominio.

   Esto es, $T(\mathbf{o}_V) = \mathbf{o}_W$ para cualquier $T : V \to W$ que sea una TL.

2. Si tenemos dos TL $S, T : V \to W$, entonces $S + T : V \to W$ también es una TL, lo mismo sucede con αT.

   Dicho en otros términos, la CL de TL es una nueva TL.

3. Si tenemos dos TL $S : V \to W$ y $T : W \to X$ (V, W, X son EV sobre el mismo cuerpo K), la función compuesta $(T \circ S) : V \to X$ definida como $(T \circ S)(v) = T(S(v))$, también es una TL.

4. La inversa de una TL también es una TL.

   Esto es, si $T : V \to W$ es biyectiva, entonces $T^{-1} : W \to V$ también es una TL.

   Esto es similar a algo que ya hemos visto varias veces en cursos de cálculo, por ejemplo que la suma de funciones continuas es también una función continua, que la composición de funciones continuas es también una función continua y que si f es continua e invertible entonces su inversa también es continua.

   En este caso la propiedad que se conserva cuando sumamos, multiplicamos por escalares, componemos o invertimos, es la de ser una TL.

5. Los ceros de una TL (esto es, los $v \in V$ tales que $T(v)$ es el vector nulo de W) son un SEV de V.

   A este conjunto de ceros o raíces le llamaremos Núcleo o Kernel de T.

   La notación habitual es $N(T)$ o $Ker(T)$.

6. El conjunto de imágenes de T (esto es, los elementos de W que son T de algún elemento de V) es un SEV de W.

   La notación habitual es $Im(T)$.

7. Si T es inyectiva, llevará conjuntos LI en V a conjuntos LI en W.

   Esto es, si A es LI en V y T es inyectiva, entonces $T(A)$ será LI en W.

8. Si T es sobreyectiva, llevará conjuntos generadores en V a conjuntos generadores en W.

   Esto es, si $A \underset{g}{\to} V$ y T es sobreyectiva, entonces $T(A) \underset{g}{\to} W$.

9. Como consecuencia de los dos puntos anteriores, si T es biyectiva, llevará bases en bases.

   Esto es, si $A \overset{b}{\to} V$ y T es biyectiva, entonces $T(A) \overset{b}{\to} W$.

10. Una TL queda definida si conocemos lo que le hace a una base del dominio.

    En efecto, sea $A = \{v_1, \dots, v_n\} \overset{b}{\to} V$, tomemos $v \in V$ cualquiera.

    Como $v = \displaystyle\sum_{i=1}^{n} \alpha_i v_i$, entonces $T(v) = \displaystyle\sum_{i=1}^{n} \alpha_i T(v_i)$ (usamos que T es una TL), de modo que para conocer a $T(v)$ lo que realmente necesitamos es conocer $T(A) = \{T(v_1), \dots, T(v_n)\}$.

    Tener $T(A) = \{T(v_1), \dots, T(v_n)\}$ equivale a tener "la fórmula "$T(v) = w$.

    Tomemos por ejemplo $T : \mathbb{R}^3 \to \mathbb{R}^2$, $A = \{(1,0,0), (1,1,0), (1,1,1) \overset{b}{\to} \mathbb{R}^3$.

    Es lo mismo que nos digan $T(1,0,0) = (1,2), T(1,1,0) = (2,2), T(1,1,1) = (2,1)$ a que nos digan $T(x,y,z) = x+y, 2x-z$.

Recordemos que en los cursos de cálculo cuando nos dan una función f siempre nos la dan diciéndonos lo que f le hace a todo el dominio, siempre nos la dan mediante la fórmula $f(x) =$ algo. En el caso de TL tenemos la opción de la fórmula y también la de decir lo que la función le hace no a todos los vectores, sino solamente a algunos (los que forman una base).

11. **Teorema de las dimensiones:**

    Si $T : V \to W$ es una TL, entonces $dim(N(T)) + dim(Im(T)) = dim(V)$

3.2 Coordenadas de un vector en una base dada

Fijemos una base $A = \{v_1, \dots, v_n\}$ del EV V. Entonces, para cada vector $v \in V$, existen (porque A es generador de V) escalares $\alpha_1, \dots, \alpha_n$ tales que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$.

Además estos coeficientes $\alpha_1, \dots, \alpha_n$ son únicos (porque A es LI).

Por tanto, fijada una base A, cada vector $v \in V$ tiene asociado un vector $(\alpha_1, \dots, \alpha_n) \in K^n$, los coeficientes que permiten fabricar a v como CL de A.

Este vector de coeficientes se llama **coordenadas de v en la base** A, es una suerte de "huella digital "del vector.

Notemos que lo que tenemos aquí es una función biyectiva $f : V \to K^n$ que a cada $v \in V$ le asocia su vector de coeficientes.

Usaremos la notación $f(v) = coord_A(v)$, de modo que $coord_A(v) = (\alpha_1, \dots, \alpha_n)$

Esta correspondencia es la que permite luego representar gráficamente vectores en un sistema de ejes, nos permite "ver"lo que estamos haciendo.

Puede probarse que esta función $coord_A$ es una TL biyectiva entre V y K^n.

Veremos a continuación que si tenemos una TL $T : V \to W$ y fijamos A, B bases de V y W respectivamente, podremos asociar a T una matriz (su huella digital, que llamaremos **matriz asociada a T en las bases** A, B) que es el reflejo de T en el mundo de las coordenadas.

Esto nos permitirá operar con T ya sea en su ámbito natural de dominio y codominio como en el ámbito de las coordenadas, como veremos a continuación.

4 Transformaciones Lineales y Matrices

4.1 Matriz asociada a una TL

Sean V, W dos EV sobre el mismo cuerpo K, tomemos $A = \{v_1, \dots, v_n\} \xrightarrow{b} V$ y $B = \{w_1, \dots, w_m\} \xrightarrow{b} W$.

La idea es estudiar qué es lo que realmente necesitamos para operar con T, qué es lo que queda al final del día.

Así pues, digamos que nos interesa hallar $T(v)$ para un $v \in V$ dado.

Como A es una base de V, sabemos que existe un vector $(\alpha_1, \dots, \alpha_n) \in K^n$ tal que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ (además es el único vector de K^n que permite reconstruír a ese vector v).

Por tanto $T(v) = T(\alpha_1 v_1 + \cdots + \alpha_n v_n)$, pero como T es una TL, tenemos que

$$T(v) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) \tag{1}$$

Notemos algo importante: para conocer $T(v)$ necesitamos conocer dos cosas: los alfas (esto es, $coord_A(v)$) y las imágenes por T de la base A (esto es, $T(v_1), \ldots, T(v_n)$).

Si tenemos A y v, podemos hallar los alfas resolviendo un sistema de ecuaciones lineales.

Veamos qué podemos decir sobre $T(v_1), \ldots, T(v_n)$.

Como los $T(v_i)$ son vectores de W y tenemos una base B de W, entonces cada $T(v_i)$ tendrá su vector $coord_B(T(v_i)) \in K^m$.

Llamemos $(a_{1i}, a_{2i}, \ldots, a_{mi})$ al vector con las coordenadas de $T(v_i)$ en la base B.

Entonces $T(v_i) = a_{1i}w_1 + a_{2i}w_2 + \cdots + a_{mi}w_m$.

Si sustituímos en (1), obtendremos

$$T(v) = \alpha_1(\underbrace{a_{11}w_1 + \cdots + a_{m1}w_m}_{T(v_1)}) + \alpha_2(\underbrace{a_{12}w_1 + \cdots + a_{m2}w_m}_{T(v_2)}) + \cdots + \alpha_n(\underbrace{a_{1n}w_1 + \cdots + a_{mn}w_m}_{T(v_n)}) \tag{2}$$

Como $T(v)$ es un vector de W, si queremos saber cómo fabricarlo a partir de la base B deberemos hallar sus coordenadas $coord_B(T(v)) = (\beta_1, \ldots, \beta_m)$.

Para ello aplicamos distributiva y agrupamos en $w_1, w_2, \ldots, w_m$, obteniendo

$$T(v) = \underbrace{(\alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n})}_{\beta_1} w_1 + \cdots + \underbrace{(\alpha_1 a_{m1} + \alpha_2 a_{m2} + \cdots + \alpha_n a_{mn})}_{\beta_m} w_m \tag{3}$$

Estos coeficientes, estos betas que necesitamos para (usando la base B) poder reconstruír el vector $T(v)$, no son otra cosa que el producto

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{a_{m1}}_{coord_B(T(v_1))} & \underbrace{a_{m2}}_{coord_B(T(v_2))} & \cdots & \underbrace{a_{mn}}_{coord_B(T(v_n))} \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}}_{coord_A(v)} = \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}}_{coord_B(T(v))} \tag{4}$$

Notemos que la primera columna de la matriz (llamémosla M por ahora) son las $coord_B(T(v_1))$, la segunda columna son las $coord_B(T(v_2))$, la n-sima columna son las $coord_B(T(v_n))$, mientras que el vector de alfas (matriz columna) son las $coord_A(v)$.

Entonces, si conocemos las coordenadas de v en la base A y las coordenadas en la base B de $T(v_1), T(v_2), \ldots, T(v_n)$, tendremos todo lo necesario para calcular $T(v)$.

Esto es:

1. Primero construímos la matriz cuya columna i-sima son las $coord_B(T(v_i)),\ i = 1,\dots,n$ (que llamaremos **matriz asociada a T en las bases A, B**).
2. Luego multiplicamos esa matriz por la columna $coord_A(v)$.
3. El resultado son las $coord_B(T(v))$, esto es, los coeficientes $\beta_1, \beta_2, \dots, \beta_m$ necesarios para reconstruír al vector $T(v)$ utilizando la base B.

   Esta reconstrucción es posible dado que la TL $coord_B : W \to K^m$ es biyectiva y por tanto existe la función inversa $coord_B^{-1} : K^m \to W$, a la que alimentamos con $(\beta_1, \dots, \beta_m) \in K^m$ y nos devuelve el vector $T(v) = \beta_1 w_1 + \cdots + \beta_m w_m \in W$.

La figura (1) ilustra la relación entre ambos mundos, el de “arriba”y el de “abajo”.

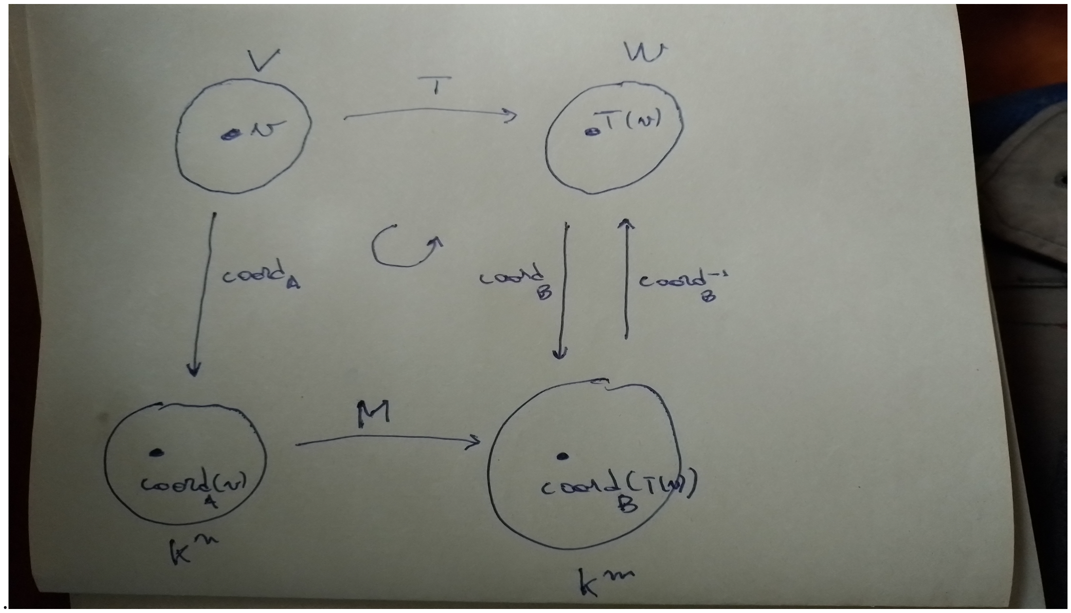


Fig. 1: Relación entre T y su matriz asociada

La figura (2) muestra otra forma de verlo.

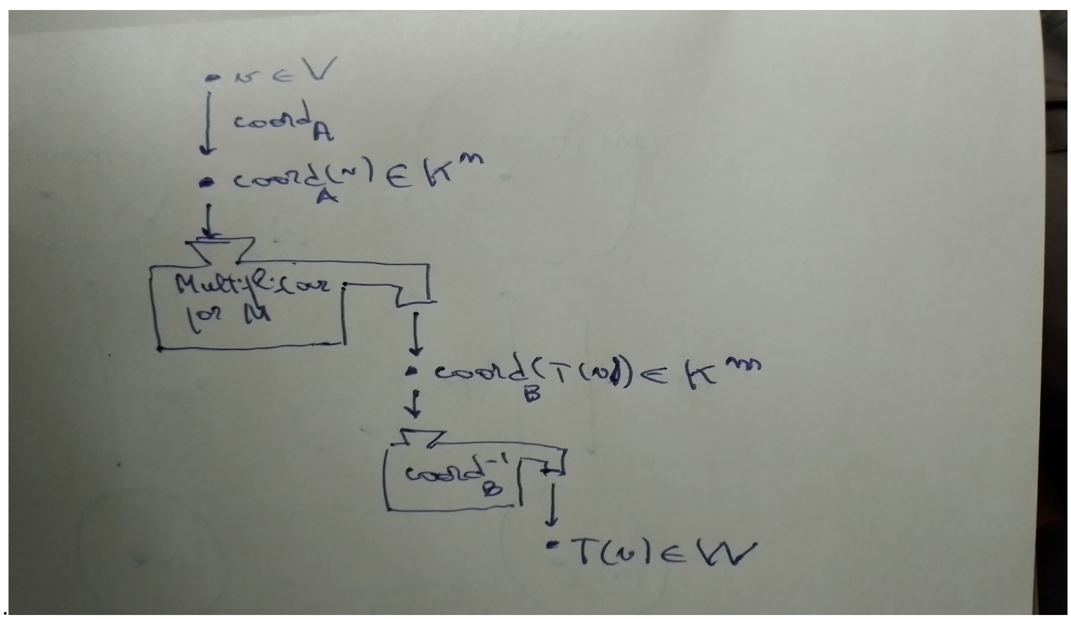


Fig. 2: Cálculo de $T(v)$ a través de su matriz asociada

Observaciones

- Notemos que la operación "multiplicar por M" define una función $T_M : K^n \to K^m$ que toma el vector $coord_A(v) \in K^n$ y devuelve el vector $coord_B(T(v)) \in K^m$.

  Esta función T_M es en sí misma una TL que queda unívocamente determinada por la matriz M.

- La cantidad de columnas de M es la dimensión de V, ya que tenemos una columna para $T(v_1)$, otra para $T(v_2)$,..., otra para $T(v_n)$ y $A = \{v_1, v_2, \ldots, v_n\}$ es una base de V.

- La cantidad de filas de M es la dimensión de W, ya que cada columna $i = 1, \ldots, n$ está formada por las coordenadas en base B de $T(v_i)$.

  Esto es, cada columna son los coeficientes que permiten expresar al vector asociado con dicha columna como CL de $\{w_1, w_2, \ldots, w_m\}$.

  Por tanto cada columna tiene m elementos y eso nos da m filas.

- Todo encaja, tenemos $\underbrace{M_{m\times n}}_{\text{matriz asociada a T}}$, $\underbrace{\alpha_{n\times 1}}_{coord_A(v)}$ y $\underbrace{\beta_{m\times 1}}_{coord_B(T(v))}$, cuyo producto es $M\alpha = \beta$.

  En síntesis, **la cantidad de filas es la dimensión del codominio, la cantidad de columnas es la dimensión del dominio**.

- Notemos que la matriz que hasta ahora hemos llamado M depende de las bases A, B de V y W que hemos elegido.

- Si cambiamos una o ambas bases, la misma T puede tener otra matriz asociada, por lo que necesitamos mejorar la notación para tener presente que la matriz depende de T pero también de A y de B.

- Por tanto usaremos la notación $\underset{B\leftarrow A}{(T)}$ en vez de M.

  En palabras, "es la matriz asociada a la TL T en las bases A del dominio y B del codominio".

Con esta notación lo que vimos puede expresarse como

$$coord_B(T(v)) = \underset{B\leftarrow A}{(T)} .coord_A(v) \tag{5}$$

Recordemos que las coordenadas de un vector en una base son columnas, por abuso de notación y comodidad escribiremos p.ej. $coord_A$ en lugar de escribir $\begin{matrix} c \\ o \\ o \\ r \\ d \end{matrix}_A$

La figura (3) ilustra la idea: tenemos una matriz $m \times n$ (la matriz $\underset{B\leftarrow A}{(T)}$) que alimentamos con un vector (matriz columna) de K^n (las coordenadas en base A del vector v).

Luego multiplicamos la matriz asociada y el vector de coordenadas, obteniendo un vector de K^m, que son las coordenadas en base B del vector $T(v)$.

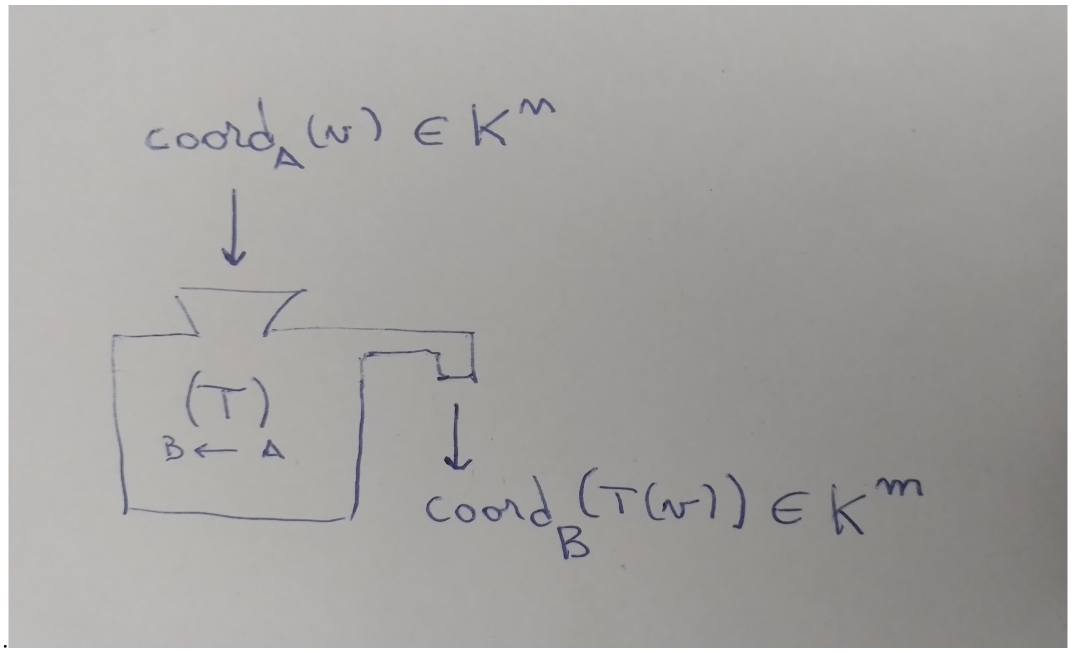
..

Fig. 3: Cómo opera la función "multiplicar por la matriz $\underset{B\leftarrow A}{(T)}$ "

4.2 Operaciones con TL, consecuencias sobre las matrices asociadas

Suma: Sabemos que si tenemos dos TL $S, T : V \to W$, entonces $(S+T) : V \to W$ es también una TL.

Supongamos que fijamos una base A de V y una base B de W, tendremos entonces dos matrices $m \times n$: $\underset{B\leftarrow A}{(S)}$ y $\underset{B\leftarrow A}{(T)}$.

Cabe preguntarse: Existirá alguna relación entre la matriz asociada a $(S+T)$ y las matrices asociadas a S y T ?

Esto es, podremos fabricar la matriz asociada a $(S+T)$ a partir de las matrices asociadas a S y a T?

Tal relación existe y además es bien sencilla:

$$\underset{B\leftarrow A}{(S+T)} = \underset{B\leftarrow A}{(S)} + \underset{B\leftarrow A}{(T)} \tag{6}$$

En palabras, la matriz asociada a la suma es la suma de las matrices.

Producto: Para cualquier escalar $a \in K$, tenemos que $aT : V \to W$ es también una TL.

Cabe preguntarse si existirá alguna relación entre la matriz $\underset{B\leftarrow A}{(aT)}$ asociada a aT y la matriz $\underset{B\leftarrow A}{(T)}$ asociada a T.

Una vez más, tal relación existe y es bien sencilla:

$$\underset{B\leftarrow A}{(aT)} = a \underset{B\leftarrow A}{(T)} \tag{7}$$

En palabras, la matriz asociada a aT es igual a a por la matriz asociada a T.

Composición: Consideremos ahora dos TL $S : V \to W$ y $T : W \to X$ con bases A, B, C respectivamente, donde $dim(V) = n$, $dim(W) = m$, $dim(X) = p$

Sabemos que la función compuesta $(T \circ S) : V \to X$ es también una TL, por lo que nuevamente cabe preguntarse acerca de si existirá alguna relación entre las matrices asociadas a $(T \circ S)$, T y S.

Notemos que ahora (a diferencia de los dos casos anteriores) las matrices tienen distintos tamaños: $\underset{C\leftarrow A}{(T \circ S)}$ es $p \times n$, $\underset{C\leftarrow B}{(T)}$ es $p \times m$ y $\underset{B\leftarrow A}{(S)}$ es $m \times n$.

En caso de existir algún tipo de relación que vincule o permita fabricar $\underset{C\leftarrow A}{(T \circ S)}$ a partir de $\underset{C\leftarrow B}{(T)}$ y $\underset{B\leftarrow A}{(S)}$, tendrá que ser una operación que tome en cuenta estos tamaños.

Al igual que antes, tal relación existe, en este caso la matriz asociada a la composición es el producto de las matrices asociadas, o más precisamente:

$$\underset{C\leftarrow A}{(T \circ S)} = \underset{C\leftarrow B}{(T)} \underset{B\leftarrow A}{(S)} \tag{8}$$

Una forma de probarlo es viendo que la matriz $\underset{C\leftarrow A}{(T\circ S)}$ hace el mismo trabajo que la matriz producto $\underset{C\leftarrow B}{(T)}\underset{B\leftarrow A}{(S)}$:

En efecto, tomemos $v \in V$ cualquiera, si hallamos sus coordenadas $coord_A(v) \in K^n$, el producto $\underset{C\leftarrow A}{(T\circ S)}.coord_A(v)$ nos dará $coord_C(T\circ S)(v)$.

Por otro lado, si hacemos primero el producto $\underset{B\leftarrow A}{(S)}.coord_A(v)$, obtendremos el vector $coord_B(S(v)) \in K^m$, y si a continuación hacemos el producto $\underset{C\leftarrow B}{(T)}.coord_B(S(v))$, obtendremos el vector de $coord_C(T(S(v))) \in K^p$.

Pero $T(S(v))$ es lo mismo que $(T\circ S)(v)$, y la TL coordenadas en una base es biyectiva, por tanto ambas matrices hacen el mismo trabajo y por tanto las TL "multiplicar por $\underset{C\leftarrow A}{(T\circ S)}$"y "multiplicar por $\underset{C\leftarrow B}{(T)}\underset{B\leftarrow A}{(S)}$ "son la misma.

La identidad La función identidad $I_V : V \to V$ definida por $I_V(v) = v\ \forall v \in V$ es una TL, por lo que si fijamos dos bases A, B de V, tendremos su matriz asociada $\underset{B\leftarrow A}{(I_V)}$.

Veamos qué podemos decir sobre la matriz asociada.

Si $A = B$: Supongamos que elegimos la misma base para el dominio (también llamada base de partida) y para el codominio (base de llegada).

Sea $A = \{v_1, v_2, \dots, v_n\}$ esta base común.

Entonces la columna i-sima de la matriz $\underset{A\leftarrow A}{(I_V)}$ estará formada por las $coord_A(I_V(v_i))$, o sea por las $coord_A(v_i)$.

Es claro que si queremos escribir a v_i como CL de un conjunto donde él mismo aparece en el lugar i-simo, alcanzará con escribir $v_i = 0v_1 + 0v_2 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_n$.

Como A es una base, dicha forma de escritura es única, por lo que $coord_A(v_i) = (0, 0, \dots, 0, 1, 0, \dots, 0)$, es la n-upla de K^n que tiene un 1 en el lugar i-simo y ceros en todo el resto.

Como esto se cumple para cualquier columna i=1...n, tenemos que

$$\underset{A\leftarrow A}{(I_V)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{9}$$

Cuando las bases de partida y de llegada son la misma, la matriz asociada a la TL identidad es la matriz identidad $\mathbb{I}$.

Si $A \neq B$: Ahora tenemos $A = \{v_1, v_2, \dots, v_n\}$, $B = \{v'_1, v'_2, \dots, v'_n\}$.

Como siempre, la columna i-sima de $\underset{B\leftarrow A}{(I_V)}$ serán las $coord_B(I_V(v_i))$, o sea las $coord_B(v_i)$.

Esto es, los coeficientes que necesitamos para fabricar al i-simo vector de la base A usando los vectores de la base B.

En este caso lo único que podemos decir es que será una matriz nxn, aunque estas matrices jugarán un rol importante, como veremos a continuación.

Inversión: Si $T : V \to W$ es una TL biyectiva (implica que V y W tienen la misma dimensión y que por tanto la matriz asociada a T es cuadrada), sabemos que $T^{-1} : W \to V$ también es una TL.

Puede probarse que la matriz asociada a T es invertible y que la matriz asociada a T^{-1} es precisamente la inversa de la matriz asociada a T.

En términos más precisos:

$$\underset{A\leftarrow B}{(T^{-1})} = (\underset{B\leftarrow A}{(T)})^{-1} \tag{10}$$

Una forma de ver esto es recordar que si $T : V \to W$ es biyectiva, entonces $T^{-1} : W \to V$.

La TL compuesta $T^{-1} \circ T : V \to V$ es la identidad $I_V : V \to V$ mientras que $T \circ T^{-1} : W \to W$ es la identidad $I_W : W \to W$.

Llamemos $M = \underset{B\leftarrow A}{(T)}$ y $N = \underset{A\leftarrow B}{(T^{-1})}$, veremos que $N = M^{-1}$ (esto es, veremos que $MN = \mathbb{I}$ y que $NM = \mathbb{I}$).

Aplicando lo que ya vimos de la matriz asociada a la TL identidad cuando tomamos la misma base en dominio y codominio, tenemos que

$$\underset{A\leftarrow A}{(T^{-1} \circ T)} = \underset{A\leftarrow A}{(I)} = \mathbb{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Por otro lado, aplicando que la matriz asociada a una composición de TL es el producto de sus respectivas matrices asociadas, tenemos que

$$\underset{A\leftarrow A}{(T^{-1} \circ T)} = \underbrace{\underset{A\leftarrow B}{(T^{-1})}}_{N} . \underbrace{\underset{B\leftarrow A}{(T)}}_{M} = \underset{A\leftarrow A}{(I_V)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Esto es, $NM = \mathbb{I}$. De forma similar puede probarse que $MN = \mathbb{I}$ (queda como ejercicio), de donde se deduce que $N = M^{-1}$.

Importante: Es un error pensar que cada vez que aparece la matriz identidad automáticamente corresponde a la TL identidad.

En efecto, supongamos dos EV V, W de la misma dimensión, sean $A = \{v_1, \ldots, v_n\} \xrightarrow{b} V$ y $B = \{w_1, \ldots, w_n\} \xrightarrow{b} W$ bases de V, W respectivamente.

Consideremos la TL $T : V \to W$ dada por $T(v_i) = w_i \ \forall i = 1, 2, \ldots, n$.

Entonces $T(v_1) = w_1$ y por tanto $coord_B(T(v_1)) = coord_B(w_1) = (1, 0, \ldots, 0)$, $T(v_2) = w_2$ y por tanto $coord_B(T(v_2)) = coord_B(w_2) = (0, 1, \ldots, 0)$ y en general $T(v_i) = w_i$ por lo que $coord_B(T(v_i)) = coord_B(w_i) = (0, \ldots, 0, 1, 0, \ldots, 0)$, donde el 1 se encuentra en el lugar i-simo.

Por tanto la matriz

$$\underset{A \leftarrow B}{(T)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{11}$$

es la matriz identidad pese a que T no es la TL identidad.

La matriz sola no dice mucho si no está acompañada del par de bases y la TL.

4.3 Matrices de cambio de base

Supongamos ahora un EV V con bases A y B y consideremos la TL $I_V : V \to V$.

A la matriz $\underset{B \leftarrow A}{(I_V)}$ la llamaremos **matriz de cambio de base**, de la base A a la base B.

Veremos el porqué de tal nombre:

Supongamos que tenemos un vector v expresado en estas dos bases distintas, A y B.

En la base A, el vector v queda identificado por sus coordenadas $coord_A(v) = (\alpha_1, \ldots, \alpha_n)$, mientras que en la base B el mismo vector v queda identificado por sus coordenadas $coord_B(v) = (\alpha'_1, \ldots, \alpha'_n)$.

Si hacemos el producto $\underset{B \leftarrow A}{(I_V)}(coord_A(v))$, obtendremos como resultado $coord_B(I_V(v)) = coord_B(v)$, (porque I es la identidad).

Esto es,

$$\underset{B \leftarrow A}{(I_V)}. \underbrace{\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}}_{coord_A(v)} = \underbrace{\begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix}}_{coord_B(v)} \tag{12}$$

Tenemos pues una matriz $\underset{B \leftarrow A}{(I_V)}$ a la que "alimentamos"con las $coord_A(v)$ y nos "entrega"las $coord_B(v)$, o sea que nos da las coordenadas del mismo vector v pero en otra base.

De ahí que a esta matriz se le llame **matriz de cambio de base**.

Podemos pensar a esta matriz como una especie de traductor, que toma una palabra (vector) v (que "en el idioma A"se representa, se escribe, por medio de $coord_A(v)$, que serían las letras utilizadas para escribir dicha palabra) y devuelve la misma palabra v pero ahora escrita en el idioma B, esto es, devuelve $coord_B(v)$.

Es como escribir la misma palabra en dos idiomas: tendremos distintas letras, distintos caracteres que representan la misma palabra.