Índice general

3. Transformaciones lineales . . . **2**

3.7. Suma y producto por escalar de transformaciones lineales . . . 2

3.8. Matriz asociada a la compuesta de transformaciones lineales . . . 3

3.9. Producto de matrices . . . 8

3.10. Núcleo y rango de una matriz . . . 11

3.11. Propiedades del producto de matrices . . . 14

3.12. Ecuaciones matriciales . . . 18

Capítulo 3

Transformaciones lineales

3.7. Suma y producto por escalar de transformaciones lineales

Supongamos que U y V son e.v. sobre un mismo cuerpo $\mathbb{K}$, U es un e.v. de dimensión n, la dimensión de V es m y $T : U \to V$, $L : U \to V$ son transformaciones lineales de U a V.

Ya sabemos cómo calcular $T + L : U \to V$, ella es la transformación que a cada $u \in U$ le hace corresponder el vector $T(u) + L(u)$.

Pero, si conocemos las matrices asociadas a T y L con respecto a bases B_U de U y B_V de V, ¿cuál es la matriz asociada a $T + L$?

Esta pregunta nos la hacemos porque a partir de ahora queremos operar con transformaciones lineales solo a través de matrices asociadas a ellas, lo que nos va a permitir estudiar las transformaciones sin importarnos cuáles son los espacios U y V.

La matriz asociada a $T + L$ con respecto a B_U y B_V es tal que

$$[(T + L)(u)]_{B_V} = [T + L]_{B_U}^{B_V}[u]_{B_U},$$

pero, como $(T + L)(u) = T(u) + L(u)$ se tiene que

$$\begin{aligned}
[(T + L)(u)]_{B_V} &= [T(u) + L(u)]_{B_V} = [T(u)]_{B_V} + [L(u)]_{B_V}, \\
&= [T]_{B_U}^{B_V}[u]_{B_U} + [L]_{B_U}^{B_V}[u]_{B_U}, \\
&= \left([T]_{B_U}^{B_V} + [L]_{B_U}^{B_V}\right)[u]_{B_U},
\end{aligned}$$

y entonces la matriz asociada a $T + L$ con respecto a las bases B_U de U y B_V de V es la suma de las matrices asociadas a T y L con respecto a las mismas bases.

Por otro lado, si $\alpha \in \mathbb{K}$ ya sabemos que la transformación $\alpha T : U \to V$ es la que a cada $u \in U$ le hace corresponder el vector $\alpha T(u)$ y si T es lineal, ésta es también una transformación lineal.

Pero, si conocemos la matriz asociada a T con respecto a bases B_U de U y B_V de V, ¿cuál es la matriz asociada a αT? Ella debe cumplir que

$$[(\alpha T)(u)]_{B_V} = [\alpha T]_{B_U}^{B_V}[u]_{B_U},$$

pero, como $(\alpha T)(u) = \alpha T(u)$ se tiene que

$$\begin{aligned}[(\alpha T)(u)]_{B_V} &= [\alpha T(u)]_{B_V} = \alpha[T(u)]_{B_V}, \\ &= \alpha[T]_{B_U}^{B_V}[u]_{B_U},\end{aligned}$$

y entonces la matriz asociada a αT con respecto a las bases B_U de U y B_V de V es el producto por α de la matriz asociada a T con respecto a las mismas bases.

Veamos cómo determinar la matriz asociada a la compuesta de dos transformaciones lineales a partir de las matrices asociadas a las transformaciones que componemos. Ésta es una operación más interesante que la suma de transformaciones lineales o el producto de una transformación lineal por un escalar y nos permitirá definir una nueva operación aritmética con matrices, el producto de matrices.

3.8. Matriz asociada a la compuesta de transformaciones lineales

Sean $T : U \to V$ y $S : V \to W$ transformaciones lineales. Supongamos además que U, V y W tienen dimensiones n, m y k respectivamente. Por último, sean B_U, una base de U, B_V, una base de V y B_W, una base de W.

Denotemos por $[T]_{B_U}^{B_V}$ a la matriz asociada a T con respecto a las bases B_U y B_V y por $[S]_{B_V}^{B_W}$ a la matriz asociada a S con respecto a las bases B_V y B_W de V y W respectivamente. Es importante que notemos dos cosas: el espacio de llegada de T es igual al espacio de partida de S, lo que nos permite definir la transformación lineal $S \circ T : U \to W$ y para construir las matrices asociadas a S y T hemos tomado la misma base de V.

Construyamos la matriz $[S \circ T]_{B_U}^{B_W}$ del mismo modo que hemos construido las matrices asociadas a otras transformaciones lineales. Supongamos que $B_U = \{u_1, u_2, \ldots, u_n\}$.

Contruiremos la matriz asociada a$S \circ T$ en forma general y con un ejemplo simultáneamente. Como ejemplo tomemos $T : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^3$ y $S : \mathbb{R}^3 \to \mathbb{R}^2$ tales que

$$T(p) = (p(-1), p(0), p(1))^{\mathrm{T}}, \qquad S((x, y, z)^{\mathrm{T}}) = (x + z, y - z)^{\mathrm{T}}.$$

La transformación $S \circ T : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^2$ es tal que

$$(S \circ T)(p) = S((p(-1), p(0), p(1))^{\mathrm{T}}) = (p(-1) + p(1), p(0) - p(1))^{\mathrm{T}}.$$

Si $B_1 = \{1, x, x^2\}$,

$$B_2 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \qquad B_3 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

son las bases canónicas de $\mathcal{P}_2(\mathbb{R})$, $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente, entonces

$$[T]_{B_1}^{B_3} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \qquad [S]_{B_3}^{B_2} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \qquad [S\circ T]_{B_1}^{B_2} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & -1 & -1 \end{pmatrix}.$$

Queremos determinar $[S\circ T]_{B_1}^{B_2}$ directamente a partir de las matrices asociadas a S y T. Veamos cómo construir cada columna de la matriz asoaciada a $S\circ T$, primero en general y después para nuestro ejemplo particular.

1. La primera columna de $[S\circ T]_{B_U}^{B_W}$ son las coordenadas de $(S\circ T)(u_1)$ con respecto a B_W, pero, como
$$(S\circ T)(u_1) = S(T(u_1)),$$
las coordenadas de $(S\circ T)(u_1)$ con respecto a B_W son las coordenadas de $S(T(u_1))$ con respecto a B_W y ésas podemos determinarlas con ayuda de $[S]_{B_V}^{B_W}$ como $[S]_{B_V}^{B_W}[T(u_1)]_{B_V}$. Como las coordenadas de $T(u_1)$ con respecto a B_V es la primera columna de $[T]_{B_U}^{B_V}$, se tiene que la primera columna de $[S\circ T]_{B_U}^{B_W}$ es el producto de la matriz asociada a S por la primera columna de la matriz asociada a T, éste es un producto matriz-vector que ya sabemos calcular.
En nuestro ejemplo, la primera columna de $[S\circ T]_{B_1}^{B_2}$ son las coordenadas de $(S\circ T)(1)$ con respecto a B_2, pero, como
$$(S\circ T)(1) = S(T(1)),$$
las coordenadas de $(S\circ T)(1)$ con respecto a B_2 son las coordenadas de $S(T(1))$ con respecto a B_2 y ésas son $[S]_{B_3}^{B_2}[T(1)]_{B_3}$. Como las coordenadas de $T(1)$ con respecto a B_3 es la primera columna de $[T]_{B_1}^{B_3}$, se tiene que la primera columna de $[S\circ T]_{B_1}^{B_2}$ es el producto de la matriz asociada a S por la primera columna de la matriz asociada a T, éste es un producto matriz-vector que ya sabemos calcular, es decir, la primera columna de $[S\circ T]_{B_1}^{B_2}$ tiene que ser
$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$
que es efectivamente lo que obtuvimos al calcular la matriz utilizando la expresión de $(S\circ T)(1)$, pero sin utilizar las matrices asociadas a S y T.
Continuemos con la segunda columna de $[S\circ T]_{B_U}^{B_W}$.

2. La segunda columna de $[S\circ T]_{B_U}^{B_W}$ son las coordenadas de $(S\circ T)(u_2)$ con respecto a B_W, pero, como
$$(S\circ T)(u_2) = S(T(u_2)),$$
las coordenadas de $(S\circ T)(u_2)$ con respecto a B_W son las coordenadas de $S(T(u_2))$ con respecto a B_W y ésas son $[S]_{B_V}^{B_W}[T(u_2)]_{B_V}$. Como las coordenadas de $T(u_2)$ con respecto a B_V es la segunda columna de $[T]_{B_U}^{B_V}$, se tiene que la segunda columna de $[S\circ T]_{B_U}^{B_W}$ es el producto de la matriz asociada a S por la segunda columna de la matriz asociada a T, éste es un producto matriz-vector que ya sabemos calcular.

En nuestro ejemplo, la segunda columna de $[S \circ T]_{B_1}^{B_2}$ son las coordenadas de $(S \circ T)(x)$ con respecto a B_2, pero, como

$$(S \circ T)(x) = S(T(x)),$$

las coordenadas de $(S \circ T)(x)$ con respecto a B_2 son las coordenadas de $S(T(x))$ con respecto a B_2 y ésas son $[S]_{B_3}^{B_2}[T(x)]_{B_3}$. Como las coordenadas de $T(x)$ con respecto a B_3 es la segunda columna de $[T]_{B_1}^{B_3}$, se tiene que la segunda columna de $[S \circ T]_{B_1}^{B_2}$ es el producto de la matriz asociada a S por la segunda columna de la matriz asociada a T, éste es un producto matriz-vector que ya sabemos calcular, es decir, la segunda columna de $[S \circ T]_{B_1}^{B_2}$ tiene que ser

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

que es efectivamente lo que obtuvimos al calcular la matriz utilizando la expresión de $(S \circ T)(x)$, pero sin utilizar las matrices asociadas a S y T.
Repitamos este proceso para la tercera columna.

3. La tercera columna de $[S \circ T]_{B_U}^{B_W}$ son las coordenadas de $(S \circ T)(u_3)$ con respecto a B_W, pero, como

$$(S \circ T)(u_3) = S(T(u_3)),$$

las coordenadas de $(S \circ T)(u_3)$ con respecto a B_W son las coordenadas de $S(T(u_3))$ con respecto a B_W y ésas son $[S]_{B_V}^{B_W}[T(u_3)]_{B_V}$. Como las coordenadas de $T(u_3)$ con respecto a B_V es la tercera columna de $[T]_{B_U}^{B_V}$, se tiene que la tercera columna de $[S \circ T]_{B_U}^{B_W}$ es el producto de la matriz asociada a S por la tercera columna de la matriz asociada a T, éste es un producto matriz-vector que ya sabemos calcular.
En nuestro ejemplo, la tercera y última columna de $[S \circ T]_{B_1}^{B_2}$ son las coordenadas de $(S \circ T)(x^2)$ con respecto a B_2, pero, como

$$(S \circ T)(x^2) = S(T(x^2)),$$

las coordenadas de $(S \circ T)(x^2)$ con respecto a B_2 son las coordenadas de $S(T(x^2))$ con respecto a B_2 y ésas son $[S]_{B_3}^{B_2}[T(x^2)]_{B_3}$. Como las coordenadas de $T(x^2)$ con respecto a B_3 es la tercera columna de $[T]_{B_1}^{B_3}$, se tiene que la tercera columna de $[S \circ T]_{B_1}^{B_2}$ es el producto de la matriz asociada a S por la tercera columna de la matriz asociada a T, éste es un producto matriz-vector que ya sabemos calcular, es decir, la tercera columna de $[S \circ T]_{B_1}^{B_2}$ tiene que ser

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix},$$

que es efectivamente lo que obtuvimos al calcular la matriz utilizando la expresión de $(S \circ T)(x)$, pero sin utilizar las matrices asociadas a S y T.

4. Repitiendo este proceso para cada una de las n columnas de $[S \circ T]_{B_U}^{B_W}$ obtenemos que la columna i-ésima de esta matriz es el producto de $[S]_{B_V}^{B_W}$ por la columna i-ésima de la matriz

asociada a T, es decir, de $[T]_{B_U}^{B_V}$. De este modo, si denotamos

$$[T]_{B_U}^{B_V} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1,n-1} & t_{1,n} \\ t_{21} & t_{22} & \cdots & t_{2,n-1} & t_{2,n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{m,n-1} & t_{m,n} \end{pmatrix} \in \mathcal{M}_{m\times n}(\mathbb{K}),$$

entonces

$$[S\circ T]_{B_U}^{B_W} = \left(\underbrace{[S]_{B_V}^{B_W}\begin{pmatrix} t_{11} \\ t_{21} \\ \vdots \\ t_{m1} \end{pmatrix}}_{\text{columna } 1} \quad \underbrace{[S]_{B_V}^{B_W}\begin{pmatrix} t_{12} \\ t_{22} \\ \vdots \\ t_{m2} \end{pmatrix}}_{\text{columna } 2} \quad \cdots \quad \underbrace{[S]_{B_V}^{B_W}\begin{pmatrix} t_{1,n-1} \\ t_{2,n-1} \\ \vdots \\ t_{m,n-1} \end{pmatrix}}_{\text{columna } n-1} \quad \underbrace{[S]_{B_V}^{B_W}\begin{pmatrix} t_{1n} \\ t_{2n} \\ \vdots \\ t_{m,n} \end{pmatrix}}_{\text{columna } n} \right).$$

A esta manera en que hemos construido la matriz $[S\circ T]_{B_U}^{B_V}$ (cada columna de esta matriz es el producto de $[S]_{B_V}^{B_W}$ por la columna correspondiente de $[T]_{B_U}^{B_V}$) le llamamos *producto* de las matrices $[S]_{B_V}^{B_W}$ y $[T]_{B_U}^{B_V}$ y lo denotamos mediante $[S]_{B_V}^{B_W}[T]_{B_U}^{B_V}$.
La matriz asociada a la compuesta de dos transformaciones lineales es el producto de las matrices asociadas a cada transformación.

Veamos algunos ejemplos sobre cómo calcular la matriz asociada a compuestas de transformaciones lineales.

Ejemplo 3.1. *Supongamos que U y V son e.v. sobre un mismo cuerpo $\mathbb{K}$ y $T: U\to V$ y $S: V\to U$ son transformaciones lineales. Pueden definirse $T\circ S: V\to V$ y $S\circ T: U\to U$.*

Sean $\mathcal{B}_V = \{v_1, v_2, v_3\}$ una base de V y $\mathcal{B}_U = \{u_1, u_2, u_3, u_4\}$ una base de U.

Las matrices asociadas a estas transformaciones son las siguientes

$$[T]_{\mathcal{B}_U}^{\mathcal{B}_V} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \qquad [S]_{\mathcal{B}_V}^{\mathcal{B}_U} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Estas matrices indican que

$$\begin{aligned} T(u_1) &= v_1 - v_2, & S(v_1) &= u_2 - u_3, \\ T(u_2) &= \theta_V, & S(v_2) &= u_2 + u_4, \\ T(u_3) &= v_2 + v_3, & S(v_3) &= u_2 + u_3, \\ T(u_4) &= v_1 - v_3. \end{aligned}$$

Sabemos entonces que, por ejemplo,

$$(S\circ T)(u_1) = S(T(u_1)) = S(v_1 - v_2) = S(v_1) - S(v_2) = -u_3 - u_4.$$

Para calcular $[S\circ T]_{\mathcal{B}_U}^{\mathcal{B}_U}$ *podríamos repetir el procedimiento anterior y evaluar* $S\circ T$ *en cada vector de* $\mathcal{B}_U$*, pero ya sabemos que* $[S\circ T]_{\mathcal{B}_U}^{\mathcal{B}_U} = [S]_{\mathcal{B}_V}^{\mathcal{B}_U}[T]_{\mathcal{B}_U}^{\mathcal{B}_V}$ *y esta matriz es*

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 0 & 1 & -2 \\ -1 & 0 & 1 & 0 \end{pmatrix}. \tag{3.1}$$

Por otro lado,

$$[T\circ S]_{\mathcal{B}_V}^{\mathcal{B}_V} = [T]_{\mathcal{B}_U}^{\mathcal{B}_V}[S]_{\mathcal{B}_V}^{\mathcal{B}_U} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & -1 & 1 \end{pmatrix}. \tag{3.2}$$

¿Qué transformación lineal es $S\circ T : U \to U$*? De la matriz* (3.1) *sabemos que*

$$(S\circ T)(u_1) = -u_3 - u_4, \qquad (S\circ T)(u_2) = \theta_U,$$
$$(S\circ T)(u_3) = 2u_2 + u_3 + u_4, \qquad (S\circ T)(u_4) = -2u_3,$$

y, por tanto, si $u \in U$ *es tal que* $u = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4$*, se tiene que*

$$\begin{aligned}(S\circ T)(u) &= \alpha_1(S\circ T)(u_1) + \alpha_2(S\circ T)(u_2) + \alpha_3(S\circ T)(u_3) + \alpha_4(S\circ T)(u_4),\\ &= \alpha_1(-u_3 - u_4) + \alpha_2\theta_U + \alpha_3(2u_2 + u_3 + u_4) + \alpha_4(-2u_3),\\ &= 2\alpha_3 u_2 + (-\alpha_1 + \alpha_3 - 2\alpha_4)u_3 + (-\alpha_1 + \alpha_3)u_4.\end{aligned}$$

También pudimos calcular $(S\circ T)(u)$ *utilizando directamente* (3.1) *pues si* $u = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4$*, entonces*

$$[u]_{B_U} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix}$$

y

$$[(S\circ T)(u)]_{B_U} = [S\circ T]_{B_U}^{B_U}[u]_{B_U} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 0 & 1 & -2 \\ -1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 2\alpha_3 \\ -\alpha_1 + \alpha_3 - 2\alpha_4 \\ -\alpha_1 + \alpha_3 \end{pmatrix}$$

y, por tanto, $(S\circ T)(u) = 0u_1 + 2\alpha_3 u_2 + (-\alpha_1 + \alpha_3 - 2\alpha_4)u_3 + (-\alpha_1 + \alpha_3)u_4$.

¿Qué transformación sería $T\circ S$ *si* U *hubiera sido* $\mathcal{P}_3(\mathbb{R})$, V, $\mathcal{P}_2(\mathbb{R})$, $\mathcal{B}_U = \{1, x, x^2, x^3\}$ *y* $\mathcal{B}_V = \{1, x, x^2\}$*, las bases canónicas de ambos espacios? Recuerda que el orden en que escribimos los vectores en* B_U *y* B_V *es importante. La primera columna de* $[T\circ S]_{B_V}^{B_V}$ *son las coordenadas respecto a* B_V *de* $(T\circ S)(1)$*, la segunda columna corresponde a las coordenadas con respecto a* B_V *de* $(T\circ S)(x)$ *y, por último, la tercera columna son las coordenadas con respecto a* B_V *de* $(T\circ S)(x^2)$*. Entonces, utilizando la matriz en* (3.2) *obtenemos que*

$$\begin{aligned}(T\circ S)(ax^2 + bx + c) &= a(T\circ S)(x^2) + b(T\circ S)(x) + c(T\circ S)(1),\\ &= a(x + x^2) + b(1 - x^2) + c(-x - x^2),\\ &= (a - b - c)x^2 + (a - c)x + b.\end{aligned}$$

3.9. Producto de matrices

Definamos formalmente el producto de matrices y veamos sus propiedades.

Recuerda que el producto de matrices es la operación que nos permite calcular la matriz asociada a la compuesta de dos transformaciones lineales $T : U \to V$ y $S : V \to W$. Para definir $S \circ T$ es necesario que el espacio de llegada de T sea igual al de partida de S.

Del mismo modo que no es posible componer dos transformaciones lineales cualesquiera tampoco es posible multiplicar dos matrices cualesquiera.

Sea $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Si $B \in \mathcal{M}_{n\times p}(\mathbb{K})$, podemos calcular el producto AB y es una matriz que pertenece al conjunto $\mathcal{M}_{m\times p}(\mathbb{K})$. Veamos que tiene sentido que para que podamos realizar el producto AB tiene que ocurrir que el número de columnas de A sea igual al número de filas de B y que la matriz resultante tiene tantas filas como A y tantas columnas como B. Supongamos que $T : \mathbb{K}^n \to \mathbb{K}^m$ y $L : \mathbb{K}^p \to \mathbb{K}^n$ son las transformaciones que podemos definir con A y B, es decir, T es la transformación que a cada $x \in \mathbb{K}^n$ le asigna $T(x) = Ax$ y L es tal que para cada $x \in \mathbb{K}^p$, $L(x) = Bx$. La matriz asociada a T con respecto a las bases canónicas de $\mathbb{K}^n$ y $\mathbb{K}^m$ es A. La matriz asociada a L con respecto a las bases canónicas de $\mathbb{K}^p$ y $\mathbb{K}^n$ es B. Como el espacio de llegada de L es igual al de partida de T (como la cantidad de filas de B es igual a la cantidad de columnas de A), entonces podemos definir $T \circ L$ y es una transformación de $\mathbb{K}^p$ a $\mathbb{K}^m$ cuya matriz con respecto a las bases canónicas de $\mathbb{K}^p$ y $\mathbb{K}^m$ es AB, tiene sentido entonces que el producto sea una matriz con m filas y p columnas.

Definición 3.2 (Producto de matrices)**.** Dadas dos matrices $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, $B \in \mathcal{M}_{n\times p}(\mathbb{K})$, la matriz producto de ellas $C = AB$ es tal que $C \in \mathcal{M}_{m\times p}(\mathbb{K})$ y

$$C = \begin{pmatrix} A\,B(:,1) & A\,B(:,2) & \cdots & A\,B(:,p) \end{pmatrix}$$

En esta definición $B(:,i)$ denota la columna $i-$ésima de la matriz B. La igualdad anterior indica que la columna i-ésima de AB es la matriz columna que resulta de multiplicar A por la columna i-ésima de B.

Nota que con esta definición del producto de matrices, el producto matriz-vector Ax también podemos verlo como el resultado de multiplicar la matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ por la matriz columna $x \in \mathcal{M}_{n\times 1}(\mathbb{K})$.

Ejemplo 3.3.

1. *Sean*

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & -1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

de donde se observa que $A \in \mathcal{M}_{2\times 3}(\mathbb{R})$, $B \in \mathcal{M}_{3\times 1}(\mathbb{R})$, $C \in \mathcal{M}_{3\times 3}(\mathbb{R})$.

Es posible entonces realizar los productos $A\,B \in \mathcal{M}_{2\times 1}(\mathbb{R})$ *(es un producto matriz-vector),* $A\,C \in \mathcal{M}_{2\times 3}(\mathbb{R})$, $C\,B \in \mathcal{M}_{3\times 1}(\mathbb{R})$ *(es un producto matriz-vector),* $C\,C \in \mathcal{M}_{3\times 3}(\mathbb{R})$, *obteniéndose*

$$A\,B = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \quad C\,B = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix},$$

mientras que

$$A\,C = \left(A \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad A \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad A \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 3 & 2 & 3 \\ 3 & 1 & 0 \end{pmatrix}$$

y

$$C\,C = \left(C \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad C \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad C \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}.$$

Sin embargo, los productos $B\,A$, $C\,A$, $B\,B$ *y* $A\,A$ *no son posibles.*

2. *Si tomamos* $y \in \mathcal{M}_{1\times 3}(\mathbb{R}), x \in \mathcal{M}_{3\times 1}(\mathbb{R})$,

$$y = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} \frac{1}{2} \\ -1 \\ 3 \end{pmatrix}.$$

Es posible calcular yx *(es un producto matriz-vector y es, por tanto,la combinación lineal de las columnas de* y *con los escalares en* x*)*

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

y xy *(cada columna de* xy *es el producto de* x *por la columna correspondiente de* y*),*

$$\begin{aligned} xy &= \begin{pmatrix} \frac{1}{2} \\ -1 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}, \\ &= \left(\begin{pmatrix} \frac{1}{2} \\ -1 \\ 3 \end{pmatrix} (1) \qquad \begin{pmatrix} \frac{1}{2} \\ -1 \\ 3 \end{pmatrix} (2) \qquad \begin{pmatrix} \frac{1}{2} \\ -1 \\ 3 \end{pmatrix} (0) \right), \\ &= \begin{pmatrix} \frac{1}{2} & 1 & 0 \\ -1 & -2 & 0 \\ 3 & 6 & 0 \end{pmatrix}. \end{aligned}$$

Observación 3.4. *Podemos escribir fórmulas para los productos matriz-vector y matriz-matriz que nos permitan calcularlos más rápido, pero no debemos olvidar que si* $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, $x \in \mathbb{K}^n$ *y* $B \in \mathcal{M}_{n\times p}(\mathbb{K})$, *el producto* Ax *es la combinación lineal de las columnas en* A *con los escalares en* x *y la columna i-ésima de* AB *es el producto matriz-vector de* A *por la columna i-ésima de* B.

Si $y \in \mathcal{M}_{1\times n}(\mathbb{K})$ *y* $x \in \mathcal{M}_{n\times 1}(\mathbb{K})$*, el producto* yx *(matriz fila por vector) es igual a*

$$(y_1 \; y_2 \; \cdots \; y_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = y_1x_1 + y_2x_2 + \cdots + y_nx_n.$$

Puede ser que éste te resulte fácil de memorizar. Los restantes productos (matriz-vector y matriz-matriz) podemos re-escribirlos como varios productos matriz fila - vector. Si, por ejemplo, $A \in \mathcal{M}_{2\times 3}(\mathbb{K})$ *y* $x \in \mathbb{K}^3$ *son tales que*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

entonces

$$Ax = x_1\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3\begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{pmatrix}.$$

Nota que entonces la primera componente de Ax *es el producto de la primera fila de* A *por* x

$$(a_{11} \; a_{12} \; a_{13})\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

y la segunda componente es el producto de la segunda fila de A *por* x*,*

$$(a_{21} \; a_{22} \; a_{23})\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

En general, si $A \in \mathcal{M}_{m\times n}(\mathbb{K})$*,* $x \in \mathbb{K}^n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad y \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

la i-ésima componente de Ax *es el producto de la i-ésima fila de* A *por* x*,*

$$Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}.$$

También podemos escribir de otra forma el producto matriz-matriz. Tomemos

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix},$$

y calculemos $C = AB$, que es una matriz de 2 filas y 4 columnas. La primera columna de C es

$$A\begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix}$$

y, a su vez, la primera componente de ese producto, es decir, $C(1,1)$, es el producto matriz fila - vector: primera fila de A por primera columna de B. El elemento en $C(2,1)$ es el producto matriz fila - vector: segunda fila de A por primera columna de B. Obtenemos entonces que la primera columna de C es

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \end{pmatrix}.$$

Repitiendo este proceso para cada columna de C se tiene

$$C = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} & a_{11}b_{14} + a_{12}b_{24} + a_{13}b_{34} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} & a_{21}b_{14} + a_{22}b_{24} + a_{23}b_{34} \end{pmatrix}.$$

En general, si $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, $B \in \mathcal{M}_{n\times p}(\mathbb{K})$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{i,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m\times n} \quad y \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2p} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{n,j} & \cdots & b_{np} \end{pmatrix}_{n\times p}$$

entonces $AB \in \mathcal{M}_{m\times p}(\mathbb{K})$, para cada $i \in \{1,2,\ldots,m\}$ y cada $j \in \{1,2,\ldots,p\}$ se tiene que el elemento en la posición (i,j) (fila i, columna j) de AB es el producto de la i-ésima fila de A por la j-ésima columna de B,

$$AB(i,j) = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \ldots + a_{i,n-1}b_{n-1,j} + a_{i,n}b_{n,j}.$$

3.10. Núcleo y rango de una matriz

Antes de continuar trabajando con matrices veamos un ejemplo de cómo podemos utilizar la matriz asociada a una transformación lineal para estudiar las propiedades de la transformación.

Ejemplo 3.5. *Supongamos nuevamente que U y V son e.v. sobre un mismo cuerpo $\mathbb{K}$ y $T : U \to V$ es una transformación lineal. Sean $\mathcal{B}_V = \{v_1, v_2, v_3\}$ una base de V y $\mathcal{B}_U = \{u_1, u_2, u_3, u_4\}$ una base de U. Supongamos además que la matriz asociada a T con respecto a las bases $\mathcal{B}_U$ y $\mathcal{B}_V$ es*

$$[T]_{\mathcal{B}_U}^{\mathcal{B}_V} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Veamos que con ayuda de esta matriz podemos determinar núcleo e imagen de la transformación, sin importarnos qué espacios son U y V.

Comencemos con el núcleo.

$$\ker(T) = \{u \in U \ : \ T(u) = \theta_V\} .$$

Pero,

$$T(u) = \theta_V \ \Leftrightarrow \ [T(u)]_{\mathcal{B}_V} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Entonces $u \in \ker(T)$ *si y solo si*

$$[T(u)]_{\mathcal{B}_V} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \ \Leftrightarrow \ [T]_{\mathcal{B}_U}^{\mathcal{B}_V} [u]_{\mathcal{B}_U} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ,$$

es decir, $u \in \ker(T)$ *si y solo* $[u]_{\mathcal{B}_U}$*, que es un vector de* $\mathbb{K}^4$*, satisface*

$$[T]_{\mathcal{B}_U}^{\mathcal{B}_V} [u]_{\mathcal{B}_U} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Hemos convertido el problema de determinar el núcleo de T *en el problema de determinar qué vectores* $x \in \mathbb{K}^4$*,* $x = (x_1, x_2, x_3, x_4)^T$ *satisfacen*

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

A este conjunto se le llama ***núcleo de la matriz*** $[T]_{\mathcal{B}_U}^{\mathcal{B}_V}$ *y es un s.e.v. de* $\mathbb{K}^4$ *(problema en listado 7),*

$$\begin{aligned} \ker([T]_{\mathcal{B}_U}^{\mathcal{B}_V}) &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{K}^4 \ : \ \begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} , \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{K}^4 \ : \ x_1 + x_4 = 0 \ \wedge \ -x_1 + x_3 = 0 \ \wedge \ x_3 - x_4 = 0 \right\} . \end{aligned}$$

De la primera ecuación obtenemos $x_1 = -x_4$*. Insertando este valor en la segunda ecuación se tiene que* $x_3 + x_4 = 0$*, de donde,* $x_3 = -x_4$*. Si insertamos este valor de* x_3 *en la última ecuación obtenemos que* $x_4 = 0$ *con lo que* $x_3 = 0$ *y* $x_1 = 0$*. Por tanto,*

$$\ker([T]_{\mathcal{B}_U}^{\mathcal{B}_V}) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{K}^4 \ : \ x_1 = x_3 = x_4 = 0 \right\} = \left\langle \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\} \right\rangle .$$

¿Cuál es entonces el núcleo de T? El generado por los vectores cuyas coordenadas con respecto a $\mathcal{B}_U$ son iguales a $\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$,

$$\ker(T) = \langle \{u_2\} \rangle \,.$$

Si conocemos la matriz asociada a una transformación lineal con respecto a bases de sus espacios de partida y llegada, los vectores en el núcleo de la matriz son las coordenadas con respecto a la base del espacio de partida de los vectores en el núcleo de la transformación.

Más adelante veremos un algoritmo para determinar qué vectores de $\mathbb{K}^n$ pertenecen al núcleo de una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Con esto tendremos un algoritmo para determinar el núcleo de cualquier transformación lineal entre e.v. de dimensión finita.

Ya sabemos que la imagen de T es el siguiente s.e.v. de V

$$\begin{aligned} \operatorname{im}(T) &= \{T(u) \;:\; u \in U\} \,, \\ &= \langle \{T(u_1), T(u_2), T(u_3), T(u_4)\} \rangle \,, \\ &= \langle \{v_1 - v_2, v_2 + v_3, v_1 - v_3\} \rangle \end{aligned}$$

¿Cómo saber si, por ejemplo, $2v_1 - v_3$ pertenece a la imagen de T? Esto se cumple si y solo si $2v_1 - v_3$ es combinación lineal de $T(u_1), T(u_2), T(u_3), T(u_4)$, es decir, si y solo si existen escalares $a, b, c, d \in \mathbb{K}$ de modo que

$$2v_1 - v_3 = aT(u_1) + bT(u_2) + cT(u_3) + dT(u_4)$$

lo que es equivalente a que existan $a, b, c, d \in \mathbb{K}$ de modo que

$$\begin{aligned} [2v_1 - v_3]_{B_V} &= [aT(u_1) + bT(u_2) + cT(u_3) + dT(u_4)]_{B_V} \,, \\ &= a[T(u_1)]_{B_V} + b[T(u_2)]_{B_V} + c[T(u_3)]_{B_V} + +d[T(u_4)]_{B_V} \,, \\ &= a \begin{pmatrix}1\\-1\\0\end{pmatrix} + b \begin{pmatrix}0\\0\\0\end{pmatrix} + c \begin{pmatrix}0\\1\\1\end{pmatrix} + d \begin{pmatrix}1\\0\\-1\end{pmatrix} . \end{aligned}$$

Hemos llegado a que $2v_1 - v_3 \in \operatorname{im}(T)$ si y solo si $[2v_1 - v_3]_{\mathcal{B}_V}$ es combinación lineal de las columnas de $[T]_{\mathcal{B}_U}^{\mathcal{B}_V}$. Al conjunto de los vectores que se escriben como combinación lineal de las columnas de una matriz se le llama ***imagen de la matriz***. *Más adelante veremos un algoritmo para determinar si un vector de $\mathbb{K}^m$ pertenece o no a la imagen de una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Con esto tendremos un algoritmo para determinar si un vector cualquiera del espacio de llegada de una transformación lineal pertenece o no a la imagen de la transformación lineal.*

En el ejemplo anterior hemos definido dos conjuntos importantes asociados a una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$: el núcleo y la imagen de A.

Definición 3.6. Sea $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Entonces, el *núcleo de A*, que se denota mediante $\ker(A)$, es el conjunto de los $x \in \mathbb{K}^n$ para los que Ax es el vector nulo de $\mathbb{K}^m$,

$$\ker(A) = \{x \in \mathbb{K}^n \;:\; Ax = \theta_{\mathbb{K}^m}\} \,.$$

Éste es el núcleo de la transformación lineal $T : \mathbb{K}^n \to \mathbb{K}^m$ definida mediante $T(x) = Ax$, él es un s.e.v. de $\mathbb{K}^n$ y a su dimensión se le llama *nulidad de* A y se denota $\eta(A)$.

La *imagen de* A es el conjunto de los vectores en $\mathbb{K}^m$ que se escriben como combinación lineal de las columnas de A y se denota por $\operatorname{im}(A)$,

$$\operatorname{im}(A) = \{Ax \ : \ x \in \mathbb{K}^n\} .$$

Este conjunto es la imagen de la transformación lineal $T : \mathbb{K}^n \to \mathbb{K}^m$ definida mediante $T(x) = Ax$. Nota entonces que

$$\operatorname{im}(A) = \operatorname{im}(T) = \left\langle \left\{ A \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, A \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, A \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\} \right\rangle$$

y, dado que el producto de A por el j-ésimo vector en la base canónica de $\mathbb{K}^n$ es la columna j-ésima de A, se tiene que, si $A(:, j)$ denota la columna j-ésima de A,

$$\operatorname{im}(A) = \langle \{A(:, 1), A(:, 2), \ldots, A(:, n)\} \rangle .$$

Este conjunto es un s.e.v. de $\mathbb{K}^m$, a su dimensión se le llama *rango de* A. A la imagen de A, dado que es el subespacio de $\mathbb{K}^m$ generado por las columnas de A, también se le llama *espacio columna de* A.

3.11. Propiedades del producto de matrices

A partir de este momento trabajaremos más con matrices. Antes de ver, como prometimos en la sección anterior, algoritmos para determinar el núcleo de una matriz o para averiguar si un vector pertenece a su imagen, debemos entender mejor el producto de matrices y sus propiedades.

1. Una diferencia importante entre el producto de matrices y el producto de números reales o de polinomios es que no es cierto que para cualquier par de matrices $A \in \mathcal{M}_{m\times n}(\mathbb{K}), B \in \mathcal{M}_{n\times k}(\mathbb{K})$ se cumpla que $AB = \Theta$ si y solo si $A = \Theta$ o $B = \Theta$, es decir,

$$A\,B = \Theta \quad \not\Rightarrow \quad A = \Theta \quad \vee \quad B = \Theta$$

   Si, por ejemplo,

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

   entonces $AB = \Theta$, pero $A \neq \Theta$ y $B \neq \Theta$. Esto significa, en particular, que si $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ y $B, C \in \mathcal{M}_{n\times k}(\mathbb{K})$ son tales que $A\,B = A\,C$, que puede escribirse de manera equivalente como $A\,(B - C) = \Theta$, entonces no tiene que ocurrir que $A = \Theta$ o $B = C$.

2. El producto de matrices no es una operación conmutativa. Por ejemplo, si

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

entonces

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

A pesar de las propiedades anteriores, que diferencian al producto de matrices de, por ejemplo, el producto de números reales, sí hay semejanzas entre ambos productos, por ejemplo, el producto de matrices es una operación asociativa y distribuye con respecto a la suma.

3. El producto de matrices es una operación asociativa, es decir, para matrices cualesquiera $A \in \mathcal{M}_{m\times n}(\mathbb{K}), B \in \mathcal{M}_{n\times k}(\mathbb{K})$ y $C \in \mathcal{M}_{k\times r}(\mathbb{K})$ con $m, n, k, r \in \mathbb{N}$ se tiene que

$$A\,(B\,C) = (A\,B)\,C.$$

4. El producto de matrices distribuye con respecto a la suma, es decir, si $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ y $B, C \in \mathcal{M}_{n\times k}(\mathbb{K})$ con $m, n, k \in \mathbb{N}$, entonces

$$A\,(B + C) = A\,B + A\,C.$$

De igual modo, si $A \in \mathcal{M}_{n\times k}(\mathbb{K})$ y $B, C \in \mathcal{M}_{m\times n}(\mathbb{K})$ con $m, n, k \in \mathbb{N}$, entonces

$$(B + C)\,A = B\,A + C\,A.$$

Recuerde que, en general, no se cumple que $A\,(B + C) = (B + C)\,A$, ni siquiera si A, B y C son tales que los productos $A\,B$, $B\,A$, $A\,C$ y $C\,A$ puedan realizarse.

5. Ya hemos visto que igualdades de la forma $AB = \Theta$ no implican que $A = \Theta$ o $B = \Theta$. El hecho de que el producto de matrices además no sea una operación conmutativa significa que debemos ser muy cuidadosos al realizar operaciones aritméticas entre matrices.
Por ejemplo, si $x, y \in \mathbb{R}$, entonces

$$(x + y)^2 = (x + y)\,(x + y) = x^2 + 2\,x\,y + y^2.$$

En cambio, si $A, B \in \mathcal{M}_{n\times n}(\mathbb{K})$, entonces

$$(A + B)\,(A + B) = A\,(A + B) + B\,(A + B) = A\,A + A\,B + B\,A + B\,B$$

y esta última expresión, en general, no es igual a $A\,A + 2\,A\,B + B\,B$ pues, no necesariamente $A\,B$ y $B\,A$ son iguales.

Recordemos que si $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, podemos definir una transformación lineal $T : \mathbb{K}^n \to \mathbb{K}^m$ mediante $T(x) = Ax$ y A es la matriz asociada a T con respecto a las bases canónicas de $\mathbb{K}^n$ y $\mathbb{K}^m$.

Definición 3.7. La matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ es invertible si la transformación lineal $T : \mathbb{K}^n \to \mathbb{K}^m$ definida mediante $T(x) = Ax$ es invertible.

Ya sabemos que T es invertible si y solo si T es biyectiva. Si $n \neq m$, T no puede ser biyectiva, por tanto, si $n \neq m$, A no es invertible.

¿Cuándo es entonces $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ una matriz invertible?

Consideremos la transformación $T : \mathbb{K}^n \to \mathbb{K}^n$ tal que $T(x) = Ax$. Hemos visto tres formas de decidir si T es invertible:

- Existe $L : \mathbb{K}^n \to \mathbb{K}^n$ de modo que $T \circ L = L \circ T = \mathrm{id}_{\mathbb{K}^n}$.
- T es inyectiva, es decir, si $\ker(T) = \{\theta_{\mathbb{K}^n}\}$ o $\eta(T) = 0$. Como $\ker(T) = \ker(A)$, A es invertible si y solo si $\ker(A) = \{\theta_{\mathbb{K}^n}\}$ o $\eta(A) = 0$.
- T es sobreyectiva, es decir, si $\mathrm{im}(T) = \mathbb{K}^n$ o $\mathrm{r}(T) = n$. Como $\mathrm{im}(T) = \mathrm{im}(A)$ (espacio generado por columnas de A), A es invertible si y solo si $\mathrm{im}(A) = \mathbb{K}^n$ o $\mathrm{r}(A) = n$.

Estos mismos criterios podemos utilizarlos para decidir si una matriz A es invertible. El primero de ellos nos lleva al siguiente lema:

Lema 3.8. La matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ es invertible si y solo si existe $B \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que

$$AB = BA = \underbrace{\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}}_{=I_n}.$$

La matriz anterior, que denotamos I_n, recibe el nombre de matriz identidad de tamaño n. Por ejemplo, si $n = 2$,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Recuerda que I_n es la matriz asociada a la transformación $\mathrm{id}_{\mathbb{K}^n}$ con respecto a las bases canónicas de $\mathbb{K}^n$.

Demostración. Sea $T : \mathbb{K}^n \to \mathbb{K}^n$ tal que $T(x) = Ax$.

Si existe $B \in \mathcal{M}_{n\times n}(\mathbb{K})$ de modo que $AB = BA = I_n$, entonces T es invertible porque la transformación $L : \mathbb{K}^n \to \mathbb{K}^n$ definida mediante $L(x) = Bx$ es tal que $T \circ L = L \circ T = \mathrm{id}_{\mathbb{K}^n}$. Esto es cierto porque

$$(T \circ L)(x) = T(L(x)) = T(Bx) = A(Bx) = (AB)x = I_n x = x$$

y

$$(L \circ T)(x) = L(T(x)) = L(Ax) = B(Ax) = (BA)x = I_n x = x.$$

Las igualdades $A(Bx) = (AB)x$ y $B(Ax) = (BA)x$ se cumplen porque el producto de matrices es una operación asociativa.

Supongamos ahora que T es invertible y demostremos que existe $B \in \mathcal{M}_n(\mathbb{K})$ de modo que $AB = BA = I_n$.

Si T es invertible, existe $L : \mathbb{K}^n \to \mathbb{K}^n$ tal que $T \circ L = \mathrm{id}_{\mathbb{K}^n}$ y $L \circ T = \mathrm{id}_{\mathbb{K}^n}$. Llamemos B a la matriz asociada a L con respecto a la base canónica, $\mathcal{B}_c$, de $\mathbb{K}^n$. Entonces,

$$\mathrm{id}_{\mathbb{K}^n} = T \circ L \;\Leftrightarrow\; [\mathrm{id}_{\mathbb{K}^n}]_{\mathcal{B}_c} = [(T \circ L)]_{\mathcal{B}_c} = [T]_{\mathcal{B}_c}[L]_{\mathcal{B}_c} = AB$$

y

$$\mathrm{id}_{\mathbb{K}^n} = L \circ T \;\Leftrightarrow\; [\mathrm{id}_{\mathbb{K}^n}]_{\mathcal{B}_c} = [(L \circ T)]_{\mathcal{B}_c} = [L]_{\mathcal{B}_c}[T]_{\mathcal{B}_c} = BA.$$

Como $[\mathrm{id}_{\mathbb{K}^n}]_{\mathcal{B}_c} = I_n$, la matriz B es tal que $AB = BA$. □

La matriz nula en $\mathcal{M}_n(\mathbb{K})$ es tal que $\forall\, A \in \mathcal{M}_n(\mathbb{K}) \;:\; A\,\Theta = \Theta\,A = \Theta$, es decir, no es posible encontrar $B \in \mathcal{M}_n(\mathbb{K})$ tal que $B\,\Theta = \Theta\,B = I_n$. La matriz nula no es invertible.

Pero, ¿es cualquier $A \in \mathcal{M}_n(\mathbb{K}), A \neq \Theta$ una matriz invertible? Desafortunadamente, la respuesta es NO. Ya sabemos que A es invertible si y solo si la transformación $T : \mathbb{K}^n \to \mathbb{K}^n$ definida mediante $T(x) = Ax$, es invertible y esto, a su vez, es cierto si y solo si el espacio nulo de T, que es igual al espacio nulo de A, solo contiene al vector nulo de $\mathbb{K}^n$. Podemos construir muchos ejemplos de matrices cuadradas que no sean invertibles. Sea, por ejemplo,

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

El espacio nulo de esta matriz es

$$\ker(A) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{K}^2 \;:\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{K}^2 \;:\; x_2 = 0 \right\} = \left\langle \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \right\rangle.$$

Ella no es, por tanto, una matriz invertible.

Lema 3.9. Si $A \in \mathcal{M}_n(\mathbb{K})$ es invertible, es decir, si existe $B \in \mathcal{M}_n(\mathbb{K})$ tal que $AB = BA = I_n$, B es única.

Demostración. Supongamos que A es invertible, es decir, existe cierta matriz B tal que $A\,B = B\,A = I_n$. Supongamos además que existe otra matriz $\hat{A}$ con esta misma propiedad. Entonces, multiplicando $B\,A = I_n$ por la derecha por $\hat{A}$ obtenemos,

$$B\,A\,\hat{A} = \hat{A} \quad \underset{A\,\hat{A} = I_n}{\Rightarrow} \quad BI_n = \hat{A} \quad \Rightarrow \quad B = \hat{A}.$$

Es decir, si existe B tal que $BA = AB = I_n$, B es la única matriz que cumple estas condiciones. □

A partir de ahora también utilizaremos solamente la letra I para denotar a la matriz identidad, es decir, no especificaremos su tamaño, éste será claro en cada caso.

Definición 3.10. Si $A \in \mathcal{M}_n(\mathbb{K})$ es invertible, es decir, si existe una matriz, denotada A^{-1}, $A^{-1} \in \mathcal{M}_n(\mathbb{K})$, tal que

$$A\,A^{-1} = A^{-1}\,A = I,$$

la matriz A^{-1} se denomina *inversa de* A. Cuando A sea invertible también se dice que A es *no singular* (o *regular).*

Nota que en la definición anterior, en lugar de I_n, escribimos I, pero está claro que si $A \in \mathcal{M}_n(\mathbb{K})$, la matriz I es I_n.

Observación 3.11. *Sean $A, B \in \mathcal{M}_n(\mathbb{K})$ invertibles, entonces*

1. *A^{-1} es única,*
2. *A^{-1} es invertible y $(A^{-1})^{-1} = A$.*
3. *$A\,B$ es invertible y $(A\,B)^{-1} = B^{-1}\,A^{-1}$.*
   Probemos que efectivamente $B^{-1}A^{-1}$ es la inversa de $A\,B$, para lo cual basta demostrar que $(A\,B)(B^{-1}\,A^{-1}) = I$ y $(B^{-1}\,A^{-1})(A\,B) = I$. Esto se cumple, pues
   $$A\,\underbrace{B\,B^{-1}}_{=I}\,A^{-1} = A\,I\,A^{-1} = A\,A^{-1} = I \quad \wedge \quad B^{-1}\,\underbrace{A^{-1}\,A}_{=I}\,B = B^{-1}\,I\,B = B^{-1}\,B = I.$$
4. *para todo $\alpha \in \mathbb{K}$, $\alpha \neq 0$, se cumple que $\alpha\,A$ es invertible y $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.*
5. *aunque podemos asegurar que el producto de matrices invertibles es invertible y el producto de una matriz invertible por un escalar distinto de cero es invertible, no es cierto que la suma de matrices invertibles sea invertible. Si, por ejemplo, $A = I_n$ y $B = -I_n$, A y B son invertibles, pero $A + B = \Theta$, que no es invertible.*

3.12. Ecuaciones matriciales

Ecuaciones matriciales son ecuaciones que involucran matrices. Veamos algunos ejemplos.

Ejemplo 3.12. *Consideremos*

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} -1 & 3 \\ 4 & -2 \end{pmatrix}.$$

Determinemos X de modo que $2X + C = -2B$.

La matriz X debe ser una matriz de tamaño 2×2, de lo contrario, $2X + C$ no podría calcularse.

Utilizando las propiedades de la suma de matrices y del producto de matriz por escalar sabemos que

$$\begin{aligned}
2X + C = -2B \;&\Leftrightarrow\; (2X + C) - C = -2B - C, && \\
&\Leftrightarrow\; 2X + (C - C) = -2B - C, && \textit{suma de matrices es asociativa} \\
&\Leftrightarrow\; 2X + \Theta = -2B - C, && -C \textit{ es inverso aditivo de } C \textit{ y } \Theta \textit{ es neutro para la suma} \\
&\Leftrightarrow\; 2X = -2B - C, && \Theta \textit{ es neutro para la suma} \\
&\Leftrightarrow\; \frac{1}{2}2X = \frac{1}{2}(-2B - C), && \\
&\Leftrightarrow\; X = -B - \frac{1}{2}C. &&
\end{aligned}$$

Con esto tenemos que

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

En general, cuando resolvemos ecuaciones matriciales no necesitamos realizar tantos pasos intermedios, pero es aconsejable los realices al principio, hasta que te acostumbres al trabajo con matrices.

Veamos otras ecuaciones matriciales, pero antes debemos escribir un par de definiciones.

Definición 3.13 (Transpuesta de una matriz)**.** Dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$,

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

(m y n no necesariamente distintos entre sí), la *transpuesta de* A, que se denota mediante A^{T}, es la matriz en $\mathcal{M}_{n\times m}(\mathbb{K})$, $A^{\mathrm{T}} = (b_{ij})_{1\leq i\leq n,\, 1\leq j\leq m}$ con coeficientes

$$b_{ij} = a_{ji}$$

para todo $i, j,\ 1 \leq i \leq n,\ 1 \leq j \leq m$, es decir, la matriz A^{T} es tal que el elemento en la posición (i, j) de A^{T} es igual al elemento en la posición (j, i) de A.

Ejemplo 3.14.

1. *Si* $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$, A^T *es la siguiente matriz de 2 filas y 1 columna*

$$A^T = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

2. *Si*

$$B = \begin{pmatrix} 1 & 5 \\ \sqrt{2} & -i \\ 0 & 2+i \end{pmatrix},$$

B^T es la siguiente matriz de 2 filas y 3 columnas

$$B^T = \begin{pmatrix} 1 & \sqrt{2} & 0 \\ 5 & -i & 2+i \end{pmatrix}$$

3. *Por último, si*

$$C = \begin{pmatrix} 0 & 1 & -1 \\ 2 & -2 & 1 \\ 10 & 100 & 0 \end{pmatrix},$$

C^T es la siguiente matriz de 3 filas y 3 columnas

$$C^T = \begin{pmatrix} 0 & 2 & 10 \\ 1 & -2 & 100 \\ -1 & 1 & 0 \end{pmatrix}.$$

Observación 3.15. *La matriz transpuesta tiene las siguientes propiedades:*

1. $\left(A^T\right)^T = A$.
2. $(\lambda\, A)^T = \lambda\, A^T$,
3. $(A + B)^T = A^T + B^T$,

4. $(A\,B)^T = B^T A^T$, *(*$A \in \mathcal{M}_{m\times n}(\mathbb{K}) \Rightarrow B \in \mathcal{M}_{n\times k}(\mathbb{K})$*, por tanto,* $AB \in \mathcal{M}_{m\times k}(\mathbb{K})$*,* $(AB)^T \in \mathcal{M}_{k\times m}(\mathbb{K})$*).*
5. *Si* A *es invertible,* A^T *es invertible y* $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$.
   Para demostrar esta propiedad debemos comprobar que la matriz $\left(A^{-1}\right)^T$ *satisface*
   $$\left(A^{-1}\right)^T A^T = I \quad y \quad A^T \left(A^{-1}\right)^T = I$$
   sabiendo que
   $$AA^{-1} = I \quad y \quad A^{-1}A = I,$$
   para lo cual basta observar que
   $$\left(A^{-1}\right)^T A^T = \left(A\,A^{-1}\right)^T = I^T = I$$
   y
   $$A^T \left(A^{-1}\right)^T = \left(A^{-1}A\right)^T = I^T = I.$$

Si bien se definió una matriz como un arreglo de escalares en un cuerpo $\mathbb{K}$, en algunas ocasiones dichos escalares hacen que la matriz tenga estructuras específicas que la convierten en un objeto de estudio (como por ejemplo, en Cálculo Numérico cuando se estudie desde un *punto de vista computacional* los llamados *sistemas de ecuaciones lineales*. Es por esto que en este curso veremos algunas matrices con estructuras típicas, que se suelen dar en la práctica con cierta frecuencia.

Definición 3.16 (Matrices simétricas y antisimétricas)**.** Una matriz $A \in \mathcal{M}_n(\mathbb{K})$ se dice

- *simétrica* si y solo si se cumple que ella y su transpuesta $A^{\mathrm{T}} \in \mathcal{M}_n(\mathbb{K})$ son iguales, es decir, si y solo si $A = A^{\mathrm{T}}$, o bien, de forma equivalente,
  $$\forall\, i, j \in \{1, 2, \cdots, n\} \;:\; a_{ij} = a_{ji}.$$
- *antisimétrica* si y solo si se cumple que ella y el inverso aditivo de su transpuesta $A^{\mathrm{T}} \in \mathcal{M}_n(\mathbb{K})$ son iguales, es decir, si y solo si $A = -\,A^{\mathrm{T}}$, o bien, de forma equivalente,
  $$\forall\, i, j \in \{1, 2, \cdots, n\} \;:\; a_{ij} = -\,a_{ji}.$$
  Observe que en este caso, los elementos de las posiciones (i, i), $i \in \{1, 2, \cdots, n\}$ deben ser cero, pues
  $$i = j \quad \Rightarrow \quad a_{ii} = -\,a_{ii} \quad \Rightarrow \quad a_{ii} = 0.$$

Ejemplo 3.17. *La matriz* $A \in \mathcal{M}_3(\mathbb{R})$ *definida como*
$$A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 0 & 3 \\ 5 & 3 & 10 \end{pmatrix},$$
es simétrica, pues no es difícil darse cuenta que $A^T = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 0 & 3 \\ 5 & 3 & 10 \end{pmatrix} = A.$

La matriz

$$B = \begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & -i \\ -3 & i & 0 \end{pmatrix}$$

es antisimétrica, pues

$$B^T = \begin{pmatrix} 0 & 2 & -3 \\ -2 & 0 & i \\ 3 & -i & 0 \end{pmatrix} = - \begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & -i \\ -3 & i & 0 \end{pmatrix} = -B.$$

Ejemplo 3.18. *Mostremos que toda matriz cuadrada (matriz con misma cantidad de filas que de columnas) se puede escribir como suma de una matriz simétrica y una antisimétrica.*

Para ello, debemos buscar, dada una matriz A cuadrada, matrices B y C, siendo B simétrica y C antisimétrica, tales que

$$A = B + C$$

y, como B es simétrica y C, antisimétrica, ellas deben cumplir,

$$B = B^T, \qquad C = -\,C^T \qquad \Rightarrow \qquad A^T = B - C.$$

De $A = B + C$, se obtiene que $B = A - C$ y, por tanto,

$$A^T = B - C = A - C - C = A - 2\,C \qquad \Rightarrow \qquad C = \frac{1}{2}\left(A - A^T\right)$$

y con ello $B = A - \frac{1}{2}\left(A - A^T\right) = \frac{1}{2}\left(A + A^T\right)$.

Aplicando este resultado podemos escribir

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & -1 & -\frac{1}{2} \\ 1 & 0 & 1 \end{pmatrix}$$

como suma de una matriz simétrica B y una antisimétrica C,

$$B = \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & -\frac{1}{4} \\ 3 & -\frac{1}{4} & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -\frac{1}{4} \\ -2 & \frac{1}{4} & 0 \end{pmatrix}.$$

Veamos otros ejemplos de ecuaciones matriciales. Pero, antes de hacerlo, definamos, para un número natural $m \in \mathbb{N} \cup \{0\}$ y una matriz $A \in \mathcal{M}_n(\mathbb{K})$ la potencia m-ésima de A.

Definición 3.19. Sea $m \in \mathbb{N} \cup \{0\}$ y $A \in \mathcal{M}_n(\mathbb{K})$, la potencia m-ésima de A se define como

$$A^0 = I, \qquad A^1 = A$$

y si $m > 1$,

$$A^m := A^{m-1}\,A = A\,A^{m-1}.$$

Ejemplo 3.20. *Consideremos*

$$A = \begin{pmatrix} 3 & 0 \\ -1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} -1 & 3 \\ 4 & -2 \end{pmatrix}.$$

Determinemos X de modo que

1. $2C^T + XA = B^2$,
2. $\left(A + \frac{1}{3}X\right)^T = C - A$.

Comencemos con la primera ecuación

$$2C^T + XA = B^2 \Leftrightarrow XA = B^2 - 2C^T = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} -2 & 8 \\ 6 & -4 \end{pmatrix} = \begin{pmatrix} 3 & -8 \\ -5 & 4 \end{pmatrix}.$$

Supongamos que $X \in \mathcal{M}_{m\times n}(\mathbb{K})$. Notemos que para que el producto XA esté definido debe cumplirse $n = 2$. Pero además, como $XA \in \mathcal{M}_{m\times 2}(\mathbb{K})$, la suma $2C^T + XA$ puede realizarse si también se cumple que $m = 2$.

Llamemos x_1, x_2, x_3, x_4 a los elementos de X,

$$X = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}.$$

Entonces

$$XA = \begin{pmatrix} 3x_1 - x_2 & 2x_2 \\ 3x_3 - x_4 & 2x_4 \end{pmatrix}.$$

Se tiene que

$$XA = \begin{pmatrix} 3 & -8 \\ -5 & 4 \end{pmatrix} \Leftrightarrow 3x_1 - x_2 = 3,\ 2x_2 = -8,\ 3x_3 - x_4 = -5,\ 2x_4 = 4$$

y, por tanto,

$$X = \begin{pmatrix} -\frac{1}{3} & -4 \\ -1 & 2 \end{pmatrix}.$$

Esta primera ecuación matricial tiene solución única y es la matriz X anterior.

Nota que la ecuación $2C^T + XB = B^2$, por el contrario, no tiene solución.

Resolvamos la segunda ecuación matricial.

$$\left(A + \frac{1}{3}X\right)^T = C - A \Leftrightarrow A + \frac{1}{3}X = (C - A)^T \Leftrightarrow A + \frac{1}{3}X = C^T - A^T.$$

Se tiene entonces que X debe ser tal que

$$A + \frac{1}{3}X = \begin{pmatrix} -4 & 5 \\ 3 & -4 \end{pmatrix} \Leftrightarrow \frac{1}{3}X = \begin{pmatrix} -7 & 5 \\ 4 & -6 \end{pmatrix} \Leftrightarrow X = \begin{pmatrix} -21 & 15 \\ 12 & -18 \end{pmatrix}.$$

La segunda ecuación matricial también tiene solución única.

Antes de escribir los nombres de otras matrices con estructuras especiales necesitamos saber cuál es la diagonal principal de una matriz cuadrada.

Definición 3.21 (Diagonal principal)**.** En una matriz cuadrada $A \in \mathcal{M}_n(\mathbb{K})$ los elementos de la forma a_{ii}, $1 \leq i \leq n$ forman la *diagonal principal de A*.

A partir de lo anterior, se consideran las siguientes matrices con estructuras (propiedades) especiales sobre el cuerpo $\mathbb{K}$,

1. *matriz diagonal:* $A \in \mathcal{M}_n(\mathbb{K}), A = (a_{ij})_{1\leq i,j \leq n}$ es diagonal si y solo si los elementos de A que no están en la diagonal principal de A son iguales a cero, es decir, si y solo si

   $$\forall\, i,j \in \{1,2,\cdots,n\} \;:\; i \neq j \quad \Rightarrow \quad a_{ij} = 0.$$

   Una matriz diagonal $A \in \mathcal{M}_n(\mathbb{K})$ tiene la forma general

   $$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

   con $a_{11}, a_{22}, \dots, a_{nn} \in \mathbb{K}$.
   Esta matriz también se representa mediante

   $$A = \operatorname{diag}\left(a_{11}\; a_{22}\; \dots\; a_{nn}\right).$$

   De esta forma, por ejemplo, $D = \operatorname{diag}(1\;5\;3\;0)$ representa la matriz diagonal

   $$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

2. *matriz escalar:* $A \in \mathcal{M}_n(\mathbb{K}), A = (a_{ij})_{1\leq i,j \leq n}$ es escalar si y solo si es diagonal y todos sus elementos sobre la diagonal principal son iguales, es decir, si y solo si

   $$\forall\, i,j \in \{1,2,\dots,n\} \;:\; i \neq j \quad \Rightarrow \quad a_{ij} = 0$$

   y existe $\lambda \in \mathbb{K}$ tal que

   $$\forall\, i \in \{1,2,\dots,n\} \;:\; a_{ii} = \lambda.$$

   De manera similar, $A \in \mathcal{M}_n(\mathbb{K})$ es escalar si y solo si existe $\lambda \in \mathbb{K}$ tal que $A = \lambda \operatorname{diag}\left(1\;1\;\dots\;1\right)$.

3. *triangular superior:* $A \in \mathcal{M}_n(\mathbb{K}), A = (a_{ij})_{1\leq i,j \leq n}$ es triangular superior si y solo si todos los elementos debajo de la diagonal principal de A son ceros, es decir, si y solo si

   $$\forall\, i,j \in \{1,2,\dots,n\} \;:\; i > j \quad \Rightarrow \quad a_{ij} = 0.$$

4. *triangular inferior:* $A \in \mathcal{M}_n(\mathbb{K}), A = (a_{ij})_{1\leq i,j\leq n}$ es triangular inferior si y solo si todos los elementos encima de la diagonal principal de A son ceros, es decir, si y solo si

$$\forall\, i, j \in \{1, 2, \ldots, n\} \;:\; i < j \quad \Rightarrow \quad a_{ij} = 0.$$

Ejemplo 3.22. *Sea*

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

1. *Muestre que A es invertible y*

$$A^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. *Halle, si es posible, una matriz invertible triangular superior $B \in \mathcal{M}_3(\mathbb{R})$ tal que $(AB)^{-1} = (AB)^T$.*

Solución: *Para responder la primera pregunta basta comprobar que la matriz A^{-1} satisface $A\,A^{-1} = I$ y $A^{-1}\,A = I$. Si estas igualdades no se cumplieran, pudiera ocurrir que A es invertible, pero su inversa no es la matriz dada o que A no es invertible.*

En la segunda parte del problema debemos determinar B de modo que $(AB)^{-1} = (AB)^T$. La igualdad dada es equivalente a las siguientes

$$\begin{aligned} (AB)^{-1} = (AB)^T &\Leftrightarrow B^{-1}A^{-1} = B^T A^T, \\ &\Leftrightarrow (B^T)^{-1}B^{-1}A^{-1} = A^T \\ &\Leftrightarrow (B\,B^T)^{-1} = A^T A, \\ &\Leftrightarrow B\,B^T = A^{-1}(A^{-1})^T. \end{aligned}$$

Dado que B es triangular superior, B tiene la forma general

$$B = \begin{pmatrix} b_1 & b_2 & b_3 \\ 0 & b_4 & b_5 \\ 0 & 0 & b_6 \end{pmatrix},$$

y así

$$B\,B^T = \begin{pmatrix} b_1^2 + b_2^2 + b_3^2 & b_2b_4 + b_3b_5 & b_3b_6 \\ b_2b_4 + b_3b_5 & b_4^2 + b_5^2 & b_5b_6 \\ b_3b_6 & b_5b_6 & b_6^2 \end{pmatrix}.$$

Por otro lado,

$$A^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A^{-1}(A^{-1})^T = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

y, por tanto, B debe cumplir las siguientes condiciones

$$\begin{gathered} b_6^2 = 1, \quad b_5 = 0, \quad b_4^2 = 1, \\ b_3\, b_6 = -1, \quad b_2 = 0, \quad b_1^2 + b_3^2 = 2. \end{gathered}$$

Los valores $b_6 = 1$, $b_5 = 0$, $b_4 = 1$, $b_3 = -1$, $b_2 = 0$, $b_1 = 1$ *satisfacen estas condiciones y con ello tenemos la matriz*

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

Se observa que ésta no es la única matriz que cumple las condiciones dadas (si calculas el núcleo de B puedes convencerte de que B es invertible). ¿Podrías construir una matriz diferente que cumpla lo pedido?