

CAPÍTULO 0

Números complejos

0.1. Cuerpos

En el cuatrimestre pasado se ha visto el concepto de cuerpo, del cual haremos un repaso (ver también [https://es.wikipedia.org/wiki/Cuerpo_\(matemáticas\)](https://es.wikipedia.org/wiki/Cuerpo_(matemáticas))).

DEFINICIÓN 0.1.1. Un conjunto \mathbb{K} es un *cuerpo* si es un anillo de división conmutativo, es decir, un anillo conmutativo con unidad en el que todo elemento distinto de cero es invertible respecto del producto. Por tanto, un cuerpo es un conjunto \mathbb{K} en el que se han definido dos operaciones, '+' y '·', llamadas *adición* y *multiplicación* respectivamente, que cumplen las siguientes propiedades. En la siguiente lista de axiomas a, b, c denotan elementos arbitrarios de \mathbb{K} , y 0 y 1 denotan elementos especiales de \mathbb{K} que cumplen las propiedades especificadas más abajo.

- I1.** $a + b$ y $a \cdot b$ pertenecen a \mathbb{K} .
- I2.** *Conmutatividad.* $a + b = b + a$; $ab = ba$.
- I3.** *Asociatividad.* $(a + b) + c = a + (b + c)$; $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- I4.** *Existencia de elemento neutro.* Existen números $0, 1 \in \mathbb{K}$ con $0 \neq 1$ tal que $a + 0 = a$; $a \cdot 1 = a$.
- I5.** *Distributividad.* $a \cdot (b + c) = a \cdot b + a \cdot c$.
- I6.** *Existencia del inverso aditivo.* Por cada a en \mathbb{K} existe un único $-a$ en \mathbb{K} tal que $a + (-a) = 0$.
- I7.** *Existencia de inverso multiplicativo.* Si a es distinto de 0, existe un único elemento $a^{-1} \in \mathbb{K}$ tal que $a \cdot a^{-1} = 1$.

Muchas veces denotaremos el producto yuxtaponiendo los elementos, es decir $ab := a \cdot b$, para $a, b \in \mathbb{K}$. Debido a la ley de asociatividad para la suma (axioma **I3**) $(a + b) + c$ es igual a $a + (b + c)$ y por lo tanto podemos eliminar los paréntesis sin ambigüedad. Es decir, denotamos

$$a + b + c := (a + b) + c = a + (b + c).$$

De forma análoga, usaremos la notación

$$abc = (ab)c = a(bc).$$

Debido a la ley de conmutatividad (axioma **I2**), es claro que del axioma **I4** se deduce que $0 + a = a + 0 = a$ y $1a = a1 = a$. Análogamente, por **I2** e **I6** obtenemos que $-a + a = a + (-a) = 0$, y por **I6** que $aa^{-1} = a^{-1}a = 1$.

Todos los axiomas corresponden a propiedades familiares de los cuerpos que ya conocemos, como ser el cuerpo de los números reales, denotado \mathbb{R} y el cuerpo de los números racionales (fracciones), denotado \mathbb{Q} . De ellas pueden deducirse la mayoría de las reglas comunes a los cuerpos. Por ejemplo, podemos *definir* la operación de sustracción diciendo que $a - b$ es lo mismo que $a + (-b)$; y deducir las reglas elementales por ejemplo,

$$a - (-b) = a + b, \quad -(-a) = a.$$

También podemos deducir

$$(ab)^{-1} = a^{-1}b^{-1}$$

con tal que a y b sean diferentes de cero. Otras reglas útiles incluyen

$$-a = (-1)a$$

y más generalmente

$$-(ab) = (-a)b = a(-b),$$

y también

$$ab = (-a)(-b),$$

así como

$$a \cdot 0 = 0,$$

todas reglas familiares de la aritmética elemental.

0.1.1. Un cuerpo finito. A modo de ejemplo, y para entrenar la intuición de que un cuerpo no necesariamente tiene un número infinito de elementos, consideremos el conjunto con dos elementos $\mathbb{F}_2 = \{0, 1\}$. Definimos la suma $+: \mathbb{F}_2 \times \mathbb{F}_2 \rightarrow \mathbb{F}_2$ mediante la regla

$$0 + 0 = 0, \quad 0 + 1 = 1, \quad 1 + 0 = 1, \quad 1 + 1 = 0$$

y el producto $\cdot: \mathbb{F}_2 \times \mathbb{F}_2 \rightarrow \mathbb{F}_2$ como

$$0 \cdot 0 = 0, \quad 0 \cdot 1 = 0, \quad 1 \cdot 0 = 0, \quad 1 \cdot 1 = 1.$$

Dejamos como ejercicio para el lector comprobar que estas operaciones así definidas satisfacen los axiomas **I1.** a **I7.** y por lo tanto \mathbb{F}_2 es un cuerpo, con dos elementos.

OBSERVACIÓN 0.1.1. El lector suspicaz reconocerá en estas operaciones a la suma y el producto definidos en el conjunto $\mathbb{Z}_2 = \{\bar{0}, \bar{1}\}$ de congruencias módulo 2 definido en Álgebra 1/Matemática Discreta. En efecto, resultados desarrollados en ese curso permiten demostrar que los conjuntos \mathbb{Z}_p , con p primo, son ejemplos de cuerpos, en este caso con p elementos.

0.2. Números complejos

La ecuación polinómica $x^2 + 1 = 0$ (¿cuál es el número que elevado al cuadrado y adicionado 1 da 0?) no tiene solución dentro del cuerpo de los números reales, pues todos sabemos que $x^2 \geq 0$ para todo $x \in \mathbb{R}$ y por lo tanto $x^2 + 1 > 0 \forall x \in \mathbb{R}$. Podemos extender \mathbb{R} a otro cuerpo, de tal forma que *toda* ecuación polinómica con coeficientes en \mathbb{R} tenga solución.

DEFINICIÓN 0.2.1. Los *números complejos* es el conjunto \mathbb{C} de los pares ordenados (a, b) , denotados $a + ib$, con a, b en \mathbb{R} , con las operaciones '+' y '·', definidas

$$(1) \quad (a + ib) + (c + id) := (a + c) + i(c + d),$$

$$(2) \quad (a + ib) \cdot (c + id) := (ac - bd) + i(ad + bc).$$

Al número complejo $i = 0 + i \cdot 1$ lo llamamos el *imaginario puro*. Si $z = a + ib$ es un número complejo, diremos que a es la *parte real* de z y la denotamos $a = \operatorname{Re} z$. Por otro lado, b es la *parte imaginaria* de z que es denotada $b = \operatorname{Im} z$.

Es claro que $z = a + ib$ es igual a $w = c + id$ si coinciden su parte real e imaginaria, es decir

$$a + bi = c + di \quad \Leftrightarrow \quad a = c \wedge b = d.$$

Podemos ver a \mathbb{R} contenido en \mathbb{C} , con la correspondencia $a \rightarrow a + i \cdot 0$ y observamos que si nos restringimos a \mathbb{R} , tenemos las reglas de adición y multiplicación usuales.

La definición de la suma de dos números complejos no debería sorprendernos, pues es la suma “coordenada a coordenada”. La definición del producto se basa en que deseamos que $i^2 = -1$ y que el producto sea distributivo.

Primero, comprobemos que $i^2 = -1$. Esto es debido a que

$$i^2 = (0 + i \cdot 1)(0 + i \cdot 1) = (0 \cdot 0 - 1 \cdot 1) + i(0 \cdot 1 + 1 \cdot 0) = -1,$$

y por lo tanto $i^2 + 1 = -1 + 1 = 0$, es decir i es solución de la ecuación polinómica $x^2 + 1 = 0$.

Sean $0 = 0 + i \cdot 0, 1 = 1 + i \cdot 0 \in \mathbb{C}$, es fácil comprobar que son los elementos neutros de la suma y el producto, respectivamente. Por otro lado, si $z = a + ib$, entonces $-z = -a - ib$ es el opuesto aditivo de z . El inverso multiplicativo es un poco más complicado. Primero observemos que dado $a + ib \in \mathbb{C}$,

$$(a + ib)(a - ib) = aa - b(-b) = a^2 + b^2 \in \mathbb{R}.$$

Supongamos que $a + ib \neq 0$, encontremos a partir de las reglas de adición y multiplicación la inversa de z . Sea $c + id$ tal que $(a + ib)(c + id) = 1$, luego

$$c + id = \frac{1}{a + ib} = \frac{1}{a + ib} \frac{a - ib}{a - ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}$$

(observar que como $a + ib \neq 0$, entonces $a^2 + b^2 > 0$.)

Usando lo anterior, y un poco más de trabajo, obtenemos

PROPOSICIÓN 0.2.1. Sean $0 = 0 + i \cdot 0, 1 = 1 + i \cdot 0 \in \mathbb{C}$. Entonces, \mathbb{C} con las operaciones '+' y '·', definidas en (1) y (2), respectivamente, es un cuerpo con elementos neutros 0 y 1, y

$$-(a + ib) = -a - ib$$

$$(a + ib)^{-1} = \frac{a - ib}{a^2 + b^2}, \quad \text{para } a + ib \neq 0.$$

DEMOSTRACIÓN. Ejercicio. □

Hemos definido los números complejos como pares ordenados y como tales es posible representarlos en el plano $\mathbb{R} \times \mathbb{R}$:

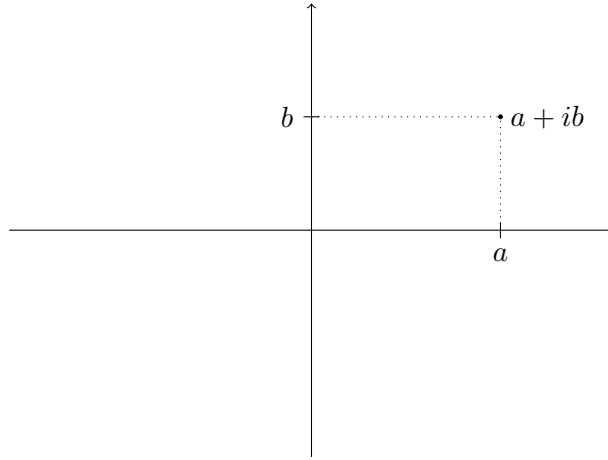


FIGURA 1. Representación gráfica de los números complejos.

Por el teorema de Pitágoras, la distancia del número complejo $a + ib$ al 0 es $\sqrt{a^2 + b^2}$.

DEFINICIÓN 0.2.2. Sea $z = a + ib \in \mathbb{C}$. El *módulo* de z es

$$|z| = \sqrt{a^2 + b^2}.$$

El *conjugado* de z es

$$\bar{z} = a - ib.$$

EJEMPLO 0.2.1. $|4 + 3i| = \sqrt{4^2 + 3^2} = \sqrt{25} = 5, \overline{4 + 3i} = 4 - 3i$.

PROPOSICIÓN 0.2.2. Sean z y w números complejos.

(1) $z\bar{z} = |z|^2$.

(2) Si $z \neq 0$, $z^{-1} = \frac{\bar{z}}{|z|^2}$.

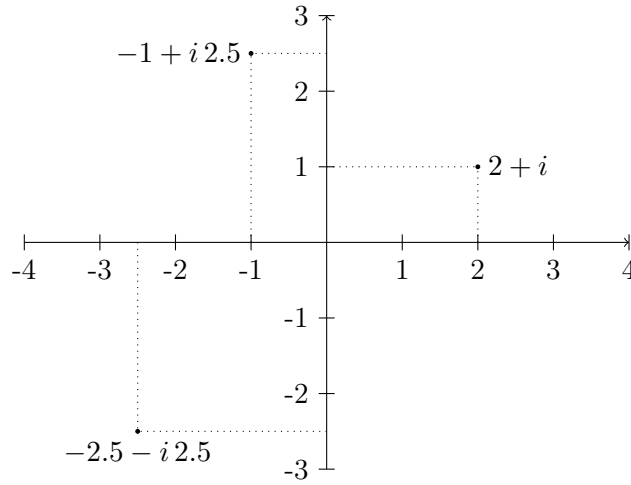


FIGURA 2. Ejemplos de la representación gráfica de los números complejos.

$$(3) \overline{z + w} = \overline{z} + \overline{w}.$$

$$(4) \overline{z\overline{w}} = \overline{z} \overline{\overline{w}}.$$

DEMOSTRACIÓN. Son comprobaciones rutinarias. Para ejemplificar, hagamos la demostración de (4).

Si $z = a + bi$ y $w = c + di$, entonces $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$. Por lo tanto,

$$\overline{z\overline{w}} = (ac - bd) - (ad + bc)i.$$

Como $\overline{z} = a - bi$ y $\overline{\overline{w}} = c - di$,

$$\overline{z} \overline{\overline{w}} = (ac - (-b)(-d)) + (a(-d) + b(-c))i = (ac - bd) - (ad + bc)i.$$

Por lo tanto $\overline{z\overline{w}} = \overline{z} \overline{\overline{w}}$. □

EJERCICIO 0.2.1. Determinar el número complejo $2 - 3i + \frac{i}{1 - i}$.

SOLUCIÓN. El ejercicio nos pide que escribamos el número en el formato $a + bi$. En general, para eliminar un cociente donde el divisor tiene parte imaginaria no nula, multiplicamos arriba y abajo por el conjugado del divisor, como $z\overline{z} \in \mathbb{R}$, obtenemos un divisor real. En el ejemplo:

$$\begin{aligned} 2 + 3i + \frac{i}{1 - i} &= 2 + 3i + \frac{i}{1 - i} \times \frac{1 + i}{1 + i} = 2 + 3i + \frac{i(1 + i)}{(1 - i)(1 + i)} \\ &= 2 + 3i + \frac{i - 1}{2} = 2 + 3i + \frac{i}{2} - \frac{1}{2} = \frac{3}{2} + i\frac{7}{2} \end{aligned}$$

□

Un poco de trigonometría. Recordemos que dado un punto $p = (x, y)$ en el plano, la recta que une el origen con p determina un ángulo θ con el eje x y entonces

$$x = r \operatorname{sen}(\theta), \quad y = r \cos(\theta)$$

donde r es la longitud del segmento determinado por $(0, 0)$ y (x, y) . En el lenguaje de los números complejos, si $z = a + bi$ y θ el ángulo determinado por z y el eje horizontal, entonces

$$a = |z| \operatorname{sen}(\theta), \quad b = |z| \cos(\theta),$$

es decir

$$(3) \quad z = |z|(\cos(\theta) + i \operatorname{sen}(\theta)).$$

Si $z \in \mathbb{C}$, la fórmula (3) es llamada la *forma polar* de z y θ es llamado el *argumento* de z .

Notación exponencial. Otra notación para representar a los números complejos es la *notación exponencial*, en la cual se denota

$$e^{i\theta} := \cos(\theta) + i \operatorname{sen}(\theta).$$

Por lo tanto si $z \in \mathbb{C}$ y θ es el argumento de z ,

$$z = r e^{i\theta}$$

donde $r = |z|$. No perder de vista, que la notación exponencial no es más que una notación (por ahora).

PROPOSICIÓN 0.2.3. Sean $z_1 = r_1 e^{i\theta_1}$, $z_2 = r_2 e^{i\theta_2}$, entonces

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}.$$

DEMOSTRACIÓN. $z_1 = r_1(\cos(\theta_1) + i \operatorname{sen}(\theta_1))$, $z_2 = r_2(\cos(\theta_2) + i \operatorname{sen}(\theta_2))$, luego

$$\begin{aligned} z_1 z_2 &= r_1 r_2 (\cos(\theta_1) + i \operatorname{sen}(\theta_1)) (\cos(\theta_2) + i \operatorname{sen}(\theta_2)) \\ &= r_1 r_2 (\cos(\theta_1) \cos(\theta_2) + i \cos(\theta_1) \operatorname{sen}(\theta_2) + i \operatorname{sen}(\theta_1) \cos(\theta_2) + i^2 \operatorname{sen}(\theta_1) \operatorname{sen}(\theta_2)) \\ &= r_1 r_2 ((\cos(\theta_1) \cos(\theta_2) - \operatorname{sen}(\theta_1) \operatorname{sen}(\theta_2)) + i (\operatorname{sen}(\theta_1) \cos(\theta_2) + \cos(\theta_1) \operatorname{sen}(\theta_2))) \\ &\stackrel{(*)}{=} r_1 r_2 (\cos(\theta_1 + \theta_2) + i \operatorname{sen}(\theta_1 + \theta_2)) = r_1 r_2 e^{i(\theta_1 + \theta_2)}. \end{aligned}$$

La igualdad (*) se debe a las tradicionales fórmulas trigonométrica del coseno y seno de la suma de ángulos. □

CAPÍTULO 1

Sistemas lineales

1.1. Sistemas de ecuaciones lineales

El problema a resolver será el siguiente: busquemos números x_1, \dots, x_n en el cuerpo \mathbb{K} ($= \mathbb{R}$ o \mathbb{C}) que satisfagan las siguientes condiciones

$$(4) \quad \begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array}$$

donde y_1, \dots, y_m y $a_{i,j}$ ($1 \leq i \leq m$, $1 \leq j \leq n$) son números en \mathbb{K} .

Llamaremos a (4) un *sistema de m ecuaciones lineales con n incógnitas*. A una n -upla (x_1, \dots, x_n) de elementos de \mathbb{K}^n que satisface cada una de las ecuaciones de (4) la llamaremos una *solución del sistema*. Si $y_1 = \dots = y_m = 0$, el sistema se llamará *homogéneo*. En caso contrario el sistema se denominará *no-homogéneo*.

EJEMPLO 1.1.1. Los siguientes son sistemas de 2 ecuaciones lineales con 2 incógnitas:

$$(1) \quad \begin{array}{rcl} 2x_1 + 8x_2 & = & 0 \\ 2x_1 + x_2 & = & 1 \end{array} \quad (2) \quad \begin{array}{rcl} 2x_1 + x_2 & = & 0 \\ 2x_1 - x_2 & = & 1 \end{array} \quad (3) \quad \begin{array}{rcl} 2x_1 + x_2 & = & 1 \\ 4x_1 + 2x_2 & = & 2 \end{array}$$

Resolvamos ahora un sistema de ecuaciones homogéneo sencillo:

$$\begin{array}{rcl} \textcircled{1} & 2x_1 - x_2 + x_3 & = 0 \\ \textcircled{2} & x_1 + 3x_2 + 4x_3 & = 0. \end{array}$$

Observar que $(0, 0, 0)$ es solución. Busquemos otras soluciones manipulando las ecuaciones.

Si hacemos $-2\textcircled{2} + \textcircled{1}$ obtenemos: $\textcircled{1'} \quad -7x_2 - 7x_3 = 0 \Rightarrow x_2 = -x_3$

Si hacemos $3\textcircled{1} + \textcircled{2}$ obtenemos: $\textcircled{2'} \quad 7x_1 + 7x_3 = 0 \Rightarrow x_1 = -x_3,$

y esto nos dice que las soluciones son de la forma $\{(-x_3, -x_3, x_3) : x_3 \in \mathbb{R}\}$, por ejemplo $(-1, -1, 1)$ es solución y $(1, 2, 3)$ no es solución.

Hemos encontrado soluciones por *eliminación de incógnitas*, es decir multiplicando por constantes adecuadas las ecuaciones y sumándolas hemos eliminado en $\textcircled{1'}$ a x_1 y en $\textcircled{2'}$ a x_2 , con lo cual el resultado se deduce inmediatamente por pasaje de término.

Veremos ahora que este método se puede extender a sistemas de ecuaciones más generales, como en (4): si en este sistema multiplicamos cada ecuación por c_i ($1 \leq i \leq m$) y sumamos miembro a miembro obtenemos

$$\sum_{i=1}^m c_i(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) = \sum_i c_i y_i.$$

Expandiendo la ecuación y tomando como factor común los x_j ($1 \leq j \leq n$) obtenemos la ecuación

$$\begin{aligned} (c_1 a_{11} + c_2 a_{21} + \cdots + c_m a_{m1})x_1 + \cdots + (c_1 a_{1n} + c_2 a_{2n} + \cdots + c_m a_{mn})x_n = \\ = c_1 y_1 + c_2 y_2 + \cdots + c_m y_m, \end{aligned}$$

o, escrito de otra forma,

$$(5) \quad \left(\sum_{i=1}^m c_i a_{i1} \right) x_1 + \cdots + \left(\sum_{i=1}^m c_i a_{in} \right) x_n = \sum_{i=1}^m c_i y_i,$$

la cual es una *combinación lineal* de las ecuaciones dadas en (4). Observar que la ecuación (5), es una ecuación lineal con n incógnitas, es decir es del mismo tipo que cada una de las ecuaciones que componen el sistema de ecuaciones original.

PROPOSICIÓN 1.1.1. Si $(x_1, \dots, x_n) \in \mathbb{K}^n$ es solución del sistema de ecuaciones (4), entonces (x_1, \dots, x_n) también es solución de la ecuación (5).

DEMOSTRACIÓN. Por hipótesis

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = y_i, \quad \text{para } 1 \leq i \leq m.$$

Luego,

$$\sum_{i=1}^m c_i(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) = \sum_i c_i y_i$$

y esta es otra escritura de la ecuación (5). □

La idea de hacer combinaciones lineales de ecuaciones es fundamental en el proceso de eliminación de incógnitas. En principio no es cierto que si obtenemos un sistema de ecuaciones por combinaciones lineales de otro sistema, ambos tengan las mismas soluciones (por ejemplo hacer combinaciones lineales triviales con todos los coeficientes iguales a 0).

Decimos que dos sistemas son *equivalentes* si cada ecuación de un sistema es combinación lineal del otro.

TEOREMA 1.1.2. *Sistemas de ecuaciones lineales equivalentes tienen las mismas soluciones.*

DEMOSTRACIÓN. Sea

$$\begin{array}{rclcl}
 (*) & a_{11}x_1 & + & a_{12}x_2 & + \cdots + a_{1n}x_n & = & y_1 \\
 & \vdots & & \vdots & & & \vdots \\
 & a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + a_{mn}x_n & = & y_m
 \end{array}$$

equivalente a

$$\begin{array}{rclcl}
 (**) & b_{11}x_1 & + & b_{12}x_2 & + \cdots + b_{1n}x_n & = & z_1 \\
 & \vdots & & \vdots & & & \vdots \\
 & b_{k1}x_1 & + & b_{k2}x_2 & + \cdots + b_{kn}x_n & = & z_k,
 \end{array}$$

En particular, las ecuaciones de $(**)$ se obtienen a partir de combinaciones lineales de las ecuaciones del sistema $(*)$. Luego, por proposición 1.1.1, si (x_1, \dots, x_n) es solución de $(*)$, también será solución de cada una de las ecuaciones de $(**)$ y por lo tanto solución del sistema.

Recíprocamente, como también las ecuaciones de $(*)$ se obtienen a partir de combinaciones lineales de las ecuaciones del sistema $(**)$, toda solución de $(**)$ es solución de $(*)$. \square

Observar que la equivalencia de sistemas lineales es una relación de equivalencia, en particular vale la propiedad transitiva: si el sistema (A) es equivalente al sistema (B) y el sistema (B) es equivalente al sistema (C), entonces (A) es equivalente a (C). Esto nos permite, ir paso a paso para eliminar las incógnitas.

EJEMPLO 1.1.2. Encontrar las soluciones del siguiente sistema de ecuaciones

$$\begin{array}{rclcl}
 (S0) & \textcircled{1} & 2x_1 & + & 4x_2 & - & 6x_3 & = & 0 \\
 & \textcircled{2} & 3x_1 & - & x_2 & + & 5x_3 & = & 0.
 \end{array}$$

SOLUCIÓN. Si reemplazamos la ecuación $\textcircled{1}$ por $\textcircled{1}/2$, obtenemos el sistema

$$\begin{array}{rclcl}
 (S1) & \textcircled{1'} & x_1 & + & 2x_2 & - & 3x_3 & = & 0 \\
 & \textcircled{2'} & 3x_1 & - & x_2 & + & 5x_3 & = & 0.
 \end{array}$$

Reemplazando $\textcircled{2'}$ por $\textcircled{2'} - 3\textcircled{1'}$, obtenemos

$$\begin{array}{rclcl}
 (S2) & \textcircled{1''} & x_1 & + & 2x_2 & - & 3x_3 & = & 0 \\
 & \textcircled{2''} & & & -7x_2 & + & 14x_3 & = & 0.
 \end{array}$$

Reemplazando $\textcircled{2''}$ por $\textcircled{2''}/(-7)$, obtenemos

$$\begin{array}{rclcl}
 (S3) & \textcircled{1'''} & x_1 & + & 2x_2 & - & 3x_3 & = & 0 \\
 & \textcircled{2'''} & & & x_2 & - & 2x_3 & = & 0.
 \end{array}$$

Luego, reemplazando x_3 por $\frac{1}{4}$ y despejando en $(1'')$ y $(2'')$:

$$\begin{array}{rcl} (1''') & -7x_2 & = -3 + 5\frac{1}{4} \\ (2''') & x_3 & = \frac{1}{4} \\ (3''') & x_1 & = 1 - 2\frac{1}{4}. \end{array}$$

Por lo tanto, $x_1 = \frac{1}{2}$, $x_2 = \frac{1}{4}$, $x_3 = \frac{1}{4}$. □

1.2. Matrices, operaciones elementales por fila y método de Gauss

Estudiaremos primeramente la solución de un sistema de ecuaciones lineales homogéneo

$$(6) \quad \begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0. \end{array}$$

Observemos que podemos escribir el sistema omitiendo las incógnitas, las sumas y los iguales (que estarán implícitos), como un arreglo rectangular de m filas y n columnas:

$$(7) \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

DEFINICIÓN 1.2.1. Sea \mathbb{K} cuerpo. Una *matriz* $m \times n$ o de *orden* $m \times n$ es un arreglo rectangular de elementos de \mathbb{K} con m filas y n columnas. A cada elemento de la matriz la llamamos *entrada* o *coeficiente*. Si A es una matriz $m \times n$, denotamos $[A]_{ij}$ la entrada que se ubica en la fila i y la columna j . Al conjunto de matrices de orden $m \times n$ con entradas en \mathbb{K} lo denotamos $M_{m \times n}(\mathbb{K})$, o simplemente $M_{m \times n}$ si \mathbb{K} está sobreentendido.

OBSERVACIÓN 1.2.1. Más formalmente, podemos ver una matriz como un elemento del producto cartesiano $(\mathbb{K}^n)^m$, es decir como m -uplas donde en cada coordenada hay una n -upla. Esta es la forma usual de describir una matriz en los lenguajes de programación modernos.

EJEMPLO 1.2.1. El siguiente es un ejemplo de una matriz 2×3 :

$$A = \begin{bmatrix} 2 & -1 & 4 \\ -3 & 0 & 1 \end{bmatrix}$$

Usualmente escribiremos a una matriz $m \times n$ con entradas $[A]_{ij} = a_{ij}$ como en (7). A esta matriz también la podemos denotar como $A = [a_{ij}]$. Dos matrices $A = [a_{ij}]$ y $B = [b_{ij}]$, de orden $m \times n$, son iguales si $a_{ij} = b_{ij}$ para todo $i = 1, \dots, m$ y $j = 1, \dots, n$. Es decir, dos matrices son iguales si los elementos que ocupan la misma posición en ambas matrices coinciden.

Como hicimos al comienzo de la sección, a un sistema de m ecuaciones con n incógnitas le asignaremos una matriz $m \times n$ lo cual nos permitirá trabajar en forma más cómoda y, como veremos en la próxima sección, podremos resolver los sistemas en forma algebraica (usando “sumas” y “productos” entre matrices). En esta sección haremos operaciones en la matriz asociada al sistema de ecuaciones para pasar a un sistema equivalente más “fácil” de resolver.

Sean

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{y} \quad 0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

podemos escribir el sistema de ecuaciones (6) en forma resumida como

$$(8) \quad AX = 0.$$

Más adelante, veremos que esta notación tiene un sentido algebraico (el término de la izquierda es un “producto de matrices”).

1.2.1. Operaciones elementales por fila. Sea $A = [a_{ij}]$ una matriz $m \times n$, entonces la fila i es

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix},$$

y la denotamos $F_i(A)$ o simplemente F_i si A está sobreentendido. Si $c \in \mathbb{K}$, entonces

$$cF_i = \begin{bmatrix} ca_{i1} & ca_{i2} & \cdots & ca_{in} \end{bmatrix}$$

y

$$F_r + F_s = \begin{bmatrix} a_{r1} + a_{s1} & a_{r2} + a_{s2} & \cdots & a_{rn} + a_{sn} \end{bmatrix}.$$

Diremos que la fila i es nula si

$$F_i = \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix},$$

DEFINICIÓN 1.2.2. Sea $A = [a_{ij}]$ una matriz $m \times n$, diremos que e es una *operación elemental por fila* si aplicada a la matriz A se obtiene $e(A)$ de la siguiente manera:

- (1) multiplicando una fila por una constante $c \neq 0$, o
- (2) cambiando la fila F_r por $F_r + tF_s$ con $r \neq s$, para algún $t \in \mathbb{K}$, o
- (3) permutando la fila r por la fila s .

Más precisamente sea

$$A = \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix},$$

entonces

- (1) si multiplicamos la fila r por $c \neq 0$,

$$e(A) = \begin{bmatrix} F_1 \\ \vdots \\ cF_r \\ \vdots \\ F_m \end{bmatrix}$$

con $c \neq 0$, o

- (2) si $r \neq s$, multiplicamos la fila s por $t \in \mathbb{K}$ y la sumamos a la fila r ,

$$e(A) = \begin{bmatrix} F_1 \\ \vdots \\ F_r + tF_s \\ \vdots \\ F_m \end{bmatrix}.$$

- (3) La última operación elemental es permutar la fila r por la fila s :

$$A = \begin{bmatrix} F_1 \\ \vdots \\ F_r \\ \vdots \\ F_s \\ \vdots \\ F_m \end{bmatrix} \Rightarrow e(A) = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_r \\ \vdots \\ F_m \end{bmatrix}.$$

Podemos describir en forma más compacta una operación elemental por fila de la matriz $A = [a_{ij}]$:

- (1) multiplicar la fila r por $c \neq 0$

$$e(A)_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r \\ ca_{ij} & \text{si } i = r \end{cases}$$

o

- (2) si $r \neq s$, multiplicar la fila s por $t \in \mathbb{K}$ y sumarla a la fila r

$$e(A)_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r \\ a_{rj} + ta_{sj} & \text{si } i = r \end{cases}$$

con $t \in \mathbb{K}$, o

(3) permutar la fila r por la fila s

$$e(A)_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r, s \\ a_{sj} & \text{si } i = r \\ a_{rj} & \text{si } i = s \end{cases}$$

EJEMPLO 1.2.2. Sea

$$A = \begin{bmatrix} 2 & 1 \\ -1 & 0 \\ 4 & -5 \end{bmatrix}.$$

Ejemplificaremos las operaciones elementales

(1) Multipliquemos la fila 2 por -2, obtenemos

$$e(A) = \begin{bmatrix} 2 & 1 \\ 2 & 0 \\ 4 & -5 \end{bmatrix}.$$

(2) Sumemos a la fila 3 dos veces la fila 1,

$$e(A) = \begin{bmatrix} 2 & 1 \\ -1 & 0 \\ 8 & -3 \end{bmatrix}.$$

(3) Permutemos la fila 2 con la fila 3.

$$e(A) = \begin{bmatrix} 2 & 1 \\ 4 & -5 \\ -1 & 0 \end{bmatrix}.$$

Una característica importante de las operaciones elementales es que cada una tiene como “inversa” otra operación elemental.

TEOREMA 1.2.1. *A cada operación elemental por fila e le corresponde otra operación elemental e' (del mismo tipo que e) tal que $e'(e(A)) = A$ y $e(e'(A)) = A$. En otras palabras, la operación inversa de una operación elemental es otra operación elemental del mismo tipo.*

DEMOSTRACIÓN.

- (1) La operación inversa de multiplicar la fila r por $c \neq 0$ es multiplicar la misma fila por $1/c$.
- (2) La operación inversa de multiplicar la fila s por $t \in \mathbb{K}$ y sumarla a la fila r es multiplicar la fila s por $-t \in \mathbb{K}$ y sumarla a la fila r .
- (3) La operación inversa de permutar la fila r por la fila s es la misma operación.

□

DEFINICIÓN 1.2.3. Sean A y B dos matrices $m \times n$. Diremos que B es *equivalente por filas* a A , si B se puede obtener de A por un número finito de operaciones elementales por fila.

OBSERVACIÓN 1.2.1. Denotamos $A \sim B$, si B es equivalente a A por filas. Entonces esta relación es una *relación de equivalencia*, es decir es reflexiva, simétrica y transitiva. En nuestro caso, sean A , B y C matrices $m \times n$, entonces “ \sim ” cumple:

- (1) $A \sim A$ (reflexiva),
- (2) $A \sim B$, entonces $B \sim A$ (simétrica), y
- (3) si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

Claramente “ \sim ” es reflexiva (admitamos que no hacer nada es una equivalencia por filas).

Si podemos obtener B de A por operaciones elementales por fila, entonces,

$$B = e_k(e_{k-1}(\cdots(e_1(A))\cdots),$$

con e_1, \dots, e_k operaciones elementales por fila. Por el teorema 1.2.1, tenemos $e'_1, \dots, e'_{k-1}, e'_k$ operaciones elementales inversas de e_1, \dots, e_{k-1}, e_k , respectivamente. Luego,

$$A = e'_1(e'_2(\cdots(e'_k(B))\cdots).$$

Es decir, podemos obtener A de B por operaciones elementales por fila, luego “ \sim ” es simétrica. Observar que para obtener A a partir de B tenemos que hacer las operaciones inversas en orden inverso.

Finalmente, si podemos obtener B de A por operaciones elementales por fila y podemos obtener C de B por operaciones elementales por fila, entonces podemos obtener C de A por operaciones elementales por fila (haciendo las primeras operaciones y luego las otras).

EJEMPLO 1.2.3. Veamos que la matriz

$$A = \begin{bmatrix} 3 & 9 & 6 \\ 4 & 8 & 4 \\ 0 & 2 & 2 \end{bmatrix}$$

es equivalente a la matriz

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

Hasta ahora, no hemos aprendido ningún algoritmo o método que nos lleve una matriz a otra por operaciones elementales por fila, pero no es difícil, en este caso, encontrar una forma de llevar la

matriz A a la matriz B :

$$\begin{aligned} \begin{bmatrix} 3 & 9 & 6 \\ 4 & 8 & 4 \\ 0 & 2 & 2 \end{bmatrix} &\xrightarrow{F_1/3} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 8 & 4 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{F_2-4F_1} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -4 & -4 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{F_2/4} \\ &\xrightarrow{F_2/4} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{F_1+3F_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{F_3+2F_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{F_3 \leftrightarrow F_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & -1 \end{bmatrix}. \end{aligned}$$

Comprobamos fácilmente la propiedad reflexiva, pues podemos llegar de la matriz B a la matriz A haciendo, sucesivamente, la operaciones inversas en orden inverso:

$$\begin{aligned} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & -1 \end{bmatrix} &\xrightarrow{F_3 \leftrightarrow F_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{F_3-2F_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{F_1-3F_2} \\ &\xrightarrow{F_1-3F_2} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{4F_3} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -4 & -4 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{F_2+4F_1} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 8 & 4 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{3F_1} \begin{bmatrix} 3 & 9 & 6 \\ 4 & 8 & 4 \\ 0 & 2 & 2 \end{bmatrix}. \end{aligned}$$

TEOREMA 1.2.2. Sean A, B matrices $m \times n$ tal que A y B son equivalentes por filas. Entonces, los sistemas homogéneos $AX = 0$ y $BX = 0$ tiene las mismas soluciones.

DEMOSTRACIÓN. Recordemos que dos sistemas de ecuaciones lineales son equivalentes si cada ecuación de un sistema es combinación lineal de las ecuaciones del otro. El teorema 1.1.2 nos dice que dos sistemas equivalentes tienen las mismas soluciones. Probaremos, entonces, que los sistemas $AX = 0$ y $BX = 0$ son equivalentes, es decir que cada ecuación de AX es combinación lineal de las ecuaciones de BX y viceversa.

Ahora bien, como A y B son equivalentes por filas existe una sucesión de operaciones elementales por fila que lleva la matriz A a la matriz B :

$$A = A_0 \longrightarrow A_1 \longrightarrow \cdots \longrightarrow A_k = B,$$

donde $A_i = e_i(A_{i-1})$ con e_i es una operación elemental ($1 \leq i \leq k$). Como las operaciones elementales permutan filas o realizan combinaciones lineales de filas, es claro que las ecuaciones de $A_i X = 0$ son combinación lineal de las ecuaciones de $A_{i-1} X = 0$. Por teorema 1.2.1, existe una operación elemental e'_i inversa de e_i tal que $A_{i-1} = e'_i(A_i)$, luego las ecuaciones de $A_{i-1} X = 0$ son combinación lineal de las ecuaciones de $A_i X = 0$ y concluimos que A_i y A_{i-1} determinan sistemas de ecuaciones equivalentes para $1 \leq i \leq k$.

Luego A_0X es equivalente a A_1X , que es equivalente a A_2X y así sucesivamente, concluyendo que $AX = A_0X$ es equivalente a $BX = A_kX$. Luego ambos sistemas tienen las mismas soluciones. \square

EJEMPLO 1.2.4. Resolvamos el siguiente sistema:

$$(9) \quad \begin{aligned} 2x_1 - x_2 + x_3 + 2x_4 &= 0 \\ x_1 - 4x_2 - x_4 &= 0 \\ 2x_1 + 6x_2 - x_3 + 3x_4 &= 0, \end{aligned}$$

para $x_i \in \mathbb{R}$ ($1 \leq i \leq 4$). En todo sistema homogéneo $x_i = 0$ ($1 \leq i \leq 4$) es solución. Veamos si hay otras soluciones.

La matriz correspondiente a este sistema de ecuaciones es

$$\begin{bmatrix} 2 & -1 & 1 & 2 \\ 1 & -4 & 0 & -1 \\ 2 & 6 & -1 & 3 \end{bmatrix}.$$

Encontraremos una matriz que nos dará un sistema de ecuaciones equivalente, pero con soluciones mucho más evidentes:

$$\begin{aligned} \begin{bmatrix} 2 & -1 & 1 & 2 \\ 1 & -4 & 0 & -1 \\ 2 & 6 & -1 & 3 \end{bmatrix} &\xrightarrow{F_1 \leftrightarrow F_2} \begin{bmatrix} 1 & -4 & 0 & -1 \\ 2 & -1 & 1 & 2 \\ 2 & 6 & -1 & 3 \end{bmatrix} \xrightarrow{F_2 - 2F_1} \begin{bmatrix} 1 & -4 & 0 & -1 \\ 0 & 7 & 1 & 4 \\ 2 & 6 & -1 & 3 \end{bmatrix} \\ &\xrightarrow{F_3 - 2F_1} \begin{bmatrix} 1 & -4 & 0 & -1 \\ 0 & 7 & 1 & 4 \\ 0 & 14 & -1 & 5 \end{bmatrix} \xrightarrow{F_3 - 2F_2} \begin{bmatrix} 1 & -4 & 0 & -1 \\ 0 & 7 & 1 & 4 \\ 0 & 0 & -3 & -3 \end{bmatrix} \xrightarrow{F_3 / (-3)} \begin{bmatrix} 1 & -4 & 0 & -1 \\ 0 & 7 & 1 & 4 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ &\xrightarrow{F_2 - F_3} \begin{bmatrix} 1 & -4 & 0 & -1 \\ 0 & 7 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{F_2 / 7} \begin{bmatrix} 1 & -4 & 0 & -1 \\ 0 & 1 & 0 & \frac{3}{7} \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{F_1 + 4F_2} \begin{bmatrix} 1 & 0 & 0 & \frac{5}{7} \\ 0 & 1 & 0 & \frac{3}{7} \\ 0 & 0 & 1 & 1 \end{bmatrix}. \end{aligned}$$

Volvamos a las ecuaciones: el nuevo sistema de ecuaciones, equivalente al original, es

$$\begin{aligned} x_1 + \frac{5}{7}x_4 &= 0 \\ x_2 + \frac{3}{7}x_4 &= 0 \\ x_3 + x_4 &= 0, \end{aligned}$$

luego

$$\begin{aligned}x_1 &= -\frac{5}{7}x_4 \\x_2 &= -\frac{3}{7}x_4 \\x_3 &= -x_4.\end{aligned}$$

Por lo tanto, el conjunto de soluciones del sistema de ecuaciones (9) es

$$\left\{ \left(-\frac{5}{7}t, -\frac{3}{7}t, -t, t \right) : t \in \mathbb{R} \right\}.$$

Luego, el sistema tiene infinitas soluciones parametrizadas por una variable $t \in \mathbb{R}$.

EJEMPLO 1.2.5. Consideremos ahora el siguiente sistema sobre los números complejos:

$$\begin{aligned}(10) \quad & 2x_1 + ix_2 = 0 \\ & -ix_1 + 3x_2 = 0 \\ & x_1 + 2x_2 = 0.\end{aligned}$$

Al ser un sistema homogéneo $x_1 = x_2 = 0$ es solución. Veamos si hay otras soluciones:

$$\begin{aligned}\begin{bmatrix} 2 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix} &\xrightarrow{F_1 \leftrightarrow F_3} \begin{bmatrix} 1 & 2 \\ -i & 3 \\ 2 & i \end{bmatrix} \xrightarrow{F_2 + iF_1} \begin{bmatrix} 1 & 2 \\ 0 & 3 + 2i \\ 2 & i \end{bmatrix} \xrightarrow{F_3 - 2F_1} \begin{bmatrix} 1 & 2 \\ 0 & 3 + 2i \\ 0 & -4 + i \end{bmatrix} \\ &\xrightarrow{F_2/(3+2i)} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & -4 + i \end{bmatrix} \xrightarrow{F_3 - (-4+i)F_2} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \xrightarrow{F_1 - 2F_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.\end{aligned}$$

Luego el sistema (10) es equivalente al sistema $x_1 = x_2 = 0$, que resulta ser la única solución.

1.2.2. Matrices reducidas por filas. Ahora avanzaremos en una forma sistemática para hallar todas las soluciones de un sistema de ecuaciones.

DEFINICIÓN 1.2.4. Una matriz A de $m \times n$ se llama *reducida por filas* o *MRF* si

- (a) la primera entrada no nula de una fila de A es 1. Este 1 es llamado *1 principal*.
- (b) Cada columna de A que contiene un 1 principal tiene todos los otros elementos iguales a 0.

EJEMPLO 1.2.6. Las siguientes matrices son MRF:

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix};$$

y las siguientes matrices, no son MRF:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \text{ no cumple (a),} \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \text{ no cumple (b).}$$

DEFINICIÓN 1.2.5. Sea I_n la matriz $n \times n$ definida

$$[I_n]_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases} \quad \text{o bien} \quad I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

(la matriz cuadrada con 1's en la diagonal y 0's en las otras entradas). Llamaremos a I_n la *matriz identidad* $n \times n$.

Observar que I_n es una matriz reducida por fila.

TEOREMA 1.2.3. *Toda matriz $m \times n$ sobre \mathbb{K} es equivalente por fila a una matriz reducida por fila.*

DEMOSTRACIÓN. Sea $A = [a_{ij}]$ una matriz $m \times n$. Trabajaremos fila por fila, de la primera a la última, de tal forma de ir encontrando matrices equivalentes por fila en cada paso, con ciertas características que ya detallaremos. Cuando terminemos llegaremos a una MRF.

Si la primera fila es nula pasamos a la segunda fila. Si la primera fila no es nula sea a_{1k} la primera entrada no nula, es decir

$$A = \begin{bmatrix} 0 & \cdots & 0 & a_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k-1} & a_{2k} & \cdots & a_{2n} \\ \vdots & & & & & \vdots \\ a_{m1} & \cdots & a_{m,k-1} & a_{mk} & \cdots & a_{mn} \end{bmatrix}.$$

Esta matriz es equivalente por fila a A_1 donde A_1 se obtiene dividiendo la fila 1 por a_{1k} . Luego

$$A_1 = \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k-1} & a_{2k} & \cdots & a_{2n} \\ \vdots & & & & & \vdots \\ a_{m1} & \cdots & a_{m,k-1} & a_{mk} & \cdots & a_{mn} \end{bmatrix}$$

(donde los nuevos a_{1j} son los originales divididos por a_{1k}).

Haciendo $m - 1$ equivalencias por fila (reemplazamos F_i por $F_i - a_{ik}F_1$) podemos hacer nulas todas las entradas debajo del 1 principal y obtener la matriz equivalente por fila

$$A_2 = \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k-1} & 0 & \cdots & a_{2n} \\ \vdots & & & & & \vdots \\ a_{m1} & \cdots & a_{m,k-1} & 0 & \cdots & a_{mn} \end{bmatrix}$$

(obviamente los nuevos a_{ij} están transformados por las equivalencias).

El mismo procedimiento que hicimos arriba lo podemos hacer en la fila 2, de tal forma que la fila 2 es cero u obtenemos otro 1 principal y todas las demás entradas de la columna donde se encuentra el 1 principal son nulas.

Repetiendo este procedimiento en todas las filas hasta la última, obtenemos que cada fila es, o bien 0, o bien la primera entrada no nula es 1 y todas las entradas en la misma columna de este 1 principal son nulas. Esto, claramente, nos dice que hemos obtenido una matriz reducida por fila. \square

El procedimiento que usamos en la demostración del teorema anterior nos da un algoritmo para obtener una MRF a partir de una matriz arbitraria.

OBSERVACIÓN 1.2.2 (MÉTODO PARA REDUCIR A MRF).

- (1) Nos ubicamos en la primera fila.
- (2) Si la fila es 0 y no es la última, pasar a la fila siguiente.
- (3) Si la fila no es 0,
 - (a) si el primera entrada no nula está en la columna k y su valor es c , dividir la fila por c (ahora la primera entrada no nula vale 1),
 - (b) con operaciones elementales del tipo $F_r + tF_s$ hacer 0 todas las entradas en la columna k (menos la de la columna actual).
- (4) Si la fila es la última, terminar. Si la fila no es la última, pasar a la fila siguiente e ir al paso (2).

Observar que en este procedimiento nunca se utiliza la operación elemental de permutar filas.

EJEMPLO 1.2.7. Ejemplifiquemos con la matriz que aparece en el ejemplo 1.2.4, es decir con la matriz

$$\begin{bmatrix} 2 & -1 & 1 & 2 \\ 1 & -4 & 0 & -1 \\ 2 & 6 & -1 & 3 \end{bmatrix}.$$

Siguiendo estrictamente el algoritmo:

$$\begin{aligned}
 \begin{bmatrix} 2 & -1 & 1 & 2 \\ 1 & -4 & 0 & -1 \\ 2 & 6 & -1 & 3 \end{bmatrix} &\xrightarrow{F_1/2} \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 1 & -4 & 0 & -1 \\ 2 & 6 & -1 & 3 \end{bmatrix} \xrightarrow[F_3-2F_1]{F_2-F_1} \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 0 & -\frac{7}{2} & -\frac{1}{2} & -2 \\ 0 & 7 & -2 & 1 \end{bmatrix} \\
 &\xrightarrow{F_2/(-\frac{7}{2})} \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 1 & \frac{1}{7} & \frac{4}{7} \\ 0 & 7 & -2 & 1 \end{bmatrix} \xrightarrow[F_3-7F_2]{F_1+\frac{1}{2}F_2} \begin{bmatrix} 1 & 0 & \frac{4}{7} & \frac{9}{7} \\ 0 & 1 & \frac{1}{7} & \frac{4}{7} \\ 0 & 0 & -3 & -3 \end{bmatrix} \\
 &\xrightarrow{F_3/(-3)} \begin{bmatrix} 1 & 0 & \frac{4}{7} & \frac{9}{7} \\ 0 & 1 & \frac{1}{7} & \frac{4}{7} \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow[F_2-\frac{1}{7}F_3]{F_1-\frac{4}{7}F_3} \begin{bmatrix} 1 & 0 & 0 & \frac{5}{7} \\ 0 & 1 & 0 & \frac{3}{7} \\ 0 & 0 & 1 & 1 \end{bmatrix}.
 \end{aligned}$$

Observemos que llegamos a la misma matriz que en el ejemplo 1.2.4, pese a que hicimos otras operaciones elementales.

DEFINICIÓN 1.2.6. Una matriz A de $m \times n$ es *escalón reducida por fila* o *MERF* si

- a) A es reducida por fila,
- b) todas las filas cuyas entradas son todas iguales a cero están al final de la matriz, y
- c) en dos filas consecutivas no nulas el 1 principal de la fila inferior está más a la derecha que el 1 principal de la fila superior.

EJEMPLO 1.2.8. Las siguientes son MERF:

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Las siguientes no son MERF:

$$\begin{bmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

En general una matriz MERF tiene la forma

$$(11) \quad \begin{bmatrix} 0 & \cdots & 1 & * & 0 & * & * & 0 & * & * \\ 0 & \cdots & 0 & \cdots & 1 & * & * & 0 & * & * \\ \vdots & & \vdots & & \vdots & & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 1 & * & * \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

TEOREMA 1.2.4. *Toda matriz $m \times n$ es equivalente por filas a una matriz escalón reducida por fila.*

DEMOSTRACIÓN. Ya vimos en el teorema 1.2.3 que toda matriz es equivalente por fila a una matriz reducida por fila. Intercambiando filas podemos entonces obtener una matriz escalón reducida por fila. \square

Llevar una matriz a una matriz escalón reducida por fila equivalente nos provee un método para encontrar todas las soluciones de sistemas de ecuaciones homogéneos.

EJEMPLO 1.2.9. Supongamos que queremos encontrar las soluciones del sistema $AX = 0$ donde A es una matriz 4×5 y que A es equivalente por filas a la matriz

$$\begin{bmatrix} 1 & 0 & 0 & 2 & -3 \\ 0 & 1 & 0 & 0 & -5 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Entonces el sistema de ecuaciones original, de 4 ecuaciones y 5 incógnitas, es equivalente al sistema

$$\begin{array}{rclcl} x_1 + 2x_4 - 3x_5 & = & 0 & & x_1 & = & -2x_4 + 3x_5 \\ x_2 - 5x_5 & = & 0 & \Rightarrow & x_2 & = & 5x_5 \\ x_3 - x_4 + 2x_5 & = & 0 & & x_3 & = & x_4 - 2x_5 \end{array}.$$

Por lo tanto el conjunto de soluciones del sistema es

$$\{(-2s + 3t, 5t, s - 2t, s, t) : s, t \in \mathbb{R}\}$$

1.2.3. Método de Gauss. Consideremos el siguiente sistema homogéneo de m ecuaciones lineales con n incógnitas:

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0 \end{array}$$

con $a_{ij} \in \mathbb{K}$. Sea $A = [a_{ij}]$ la matriz asociada a este sistema. Mediante operaciones elementales de fila obtenemos una matriz $B = [b_{ij}]$ escalón reducida por fila (teoremas 1.2.3 y 1.2.4). Sea r el número de filas no nulas y k_1, \dots, k_r las columnas donde aparecen los primeros 1's en las primeras r filas. Entonces, $k_1 < k_2 < \cdots < k_r$ y el sistema de ecuaciones asociado a B es:

$$(12) \quad \begin{array}{lcl} x_{k_1} & + & \sum_{j \neq k_1, \dots, k_r} b_{1j} x_j = 0 \\ x_{k_2} & + & \sum_{j \neq k_1, \dots, k_r} b_{2j} x_j = 0 \\ \vdots & & \vdots \\ x_{k_r} & + & \sum_{j \neq k_1, \dots, k_r} b_{rj} x_j = 0 \end{array}$$

y, por lo tanto,

$$(13) \quad \begin{array}{lcl} x_{k_1} & = & - \sum_{j \neq k_1, \dots, k_r} b_{1j} x_j \\ x_{k_2} & = & - \sum_{j \neq k_1, \dots, k_r} b_{2j} x_j \\ \vdots & & \vdots \\ x_{k_r} & = & - \sum_{j \neq k_1, \dots, k_r} b_{rj} x_j \end{array}$$

Llamaremos a $x_{k_1}, x_{k_2}, \dots, x_{k_r}$ las *variables principales* del sistema y las $n - r$ variables restantes son las *variables libres*. Es claro entonces que variando de forma arbitraria todas las variables libres obtenemos todas las soluciones del sistema.

Observar que las variables libres son $\{x_i : 1 \leq i \leq n \text{ y } i \neq k_j, j = 1, \dots, r\}$ que es igual a

$$\{x_i : 1 \leq i \leq k_r \text{ y } i \neq k_j, j = 1, \dots, r\} \cup \{x_i : i > k_r\}$$

(hacemos la partición para recalcar que algunas variables libres están intercaladas con las variables principales). Las soluciones del sistema son, entonces, los $(x_1, \dots, x_n) \in \mathbb{K}^n$ tal que x_{k_1}, \dots, x_{k_r} satisfacen las ecuaciones (13) y $x_i \in \mathbb{K}$ para x_i libre.

El método anterior es denominado *método de Gauss*. En particular, hemos probado

TEOREMA 1.2.5. *Sea A matriz $n \times n$ con coeficientes en \mathbb{K} . Entonces, si A es equivalente por filas a R una MERF y R tiene filas nulas, el sistema $AX = 0$ tiene más de una solución.*

DEMOSTRACIÓN. Sea r el número de filas no nulas de R , como R tiene n filas y tiene filas nulas, $r < n$. Por ser A y R equivalentes por fila, AX y RX tienen las mismas soluciones. Como $r < n$, entonces $n - r \geq 1$, por lo tanto existe al menos una variable libre y por lo tanto alguna solución no trivial. \square

TEOREMA 1.2.6. *Sea A una matriz $m \times n$ con $m < n$. Entonces, el sistema de ecuaciones lineales asociado a esta matriz tiene soluciones no triviales.*

DEMOSTRACIÓN. La matriz A es equivalente por fila a una matriz R escalón reducida por fila y R tiene r filas no nulas con $r \leq m < n$. Por lo tanto hay $n - r \geq 1$ variables libres, y en consecuencia RX (y AX) tiene otras soluciones además de la trivial $(0, \dots, 0)$. \square

LEMA 1.2.7. *Sea R una matriz $n \times n$ escalón reducida por fila tal que no tiene filas nulas. Entonces $R = I_n$.*

DEMOSTRACIÓN. Como R es reducida por fila y no tiene filas nulas, cada fila tiene un 1 en alguna entrada y en la columna donde está el 1 todas las otras entradas son nulas, por lo tanto hay n 1's principales distribuidos en n columnas. Concluyendo: hay un 1 por columna y en esa columna todas las demás entradas son nulas.

Ahora bien como R es una MERF, la primera fila contiene el 1 que está más a la izquierda, que no puede estar en otra ubicación que no sea la primera (pues si no la primera columna sería nula). Con el mismo razonamiento vemos que en la segunda fila hay un 1 en la columna 2 y en general en la fila k -ésima hay un 1 en la columna k . Luego $R = I_n$. \square

TEOREMA 1.2.8. *Sea A una matriz $n \times n$. Entonces, A es equivalente por filas a la matriz I_n si y sólo si el sistema $AX = 0$ tiene una única solución.*

DEMOSTRACIÓN. (\Rightarrow) Como A es equivalente por filas a la matriz I_n , las soluciones de $AX = 0$ son las mismas que las de $I_n X = 0$. Ahora bien, en la fila i de la matriz I_n tenemos $[I_n]_{ii} = 1$ y las otras entradas son cero, luego la ecuación correspondiente a esa fila es $x_i = 0$, y esto ocurre en todas las filas, luego el sistema de ecuaciones es

$$x_1 = 0$$

$$x_2 = 0$$

$$\vdots$$

$$x_n = 0$$

cuya única solución es la solución trivial.

(\Leftarrow) Sea R la matriz escalón reducida por filas asociada a A . Por hipótesis, $AX = 0$ tiene una sola solución y por lo tanto $RX = 0$ tiene una sola solución. Como la solución trivial es la única que existe, no hay variables libres, es decir hay n filas no nulas en R , como R tiene n filas, lo anterior implica que R no tiene filas nulas. Entonces, por el lema anterior, $R = I_n$. \square

1.2.4. Sistemas no homogéneos. Estudiaremos ahora sistemas lineales de ecuaciones donde las ecuaciones no están necesariamente igualadas a 0. En general, un sistema de este

tipo se escribe

$$(14) \quad \begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array}$$

donde y_1, \dots, y_m y $a_{i,j}$ ($1 \leq i \leq m$, $1 \leq j \leq n$) son números en \mathbb{K} . Si denotamos

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix},$$

entonces el sistema, para simplificar, se puede escribir como

$$AX = Y.$$

A diferencia de los sistemas homogéneos, los sistemas no homogéneos a veces no tiene solución.

EJEMPLO 1.2.10. Comprobemos que el sistema

$$2x_1 + x_2 = 0$$

$$2x_1 + x_2 = 1$$

no tiene solución. Supongamos que (x_1, x_2) es solución del sistema.

Como $2x_1 + x_2 = 0$, entonces $2x_1 = -x_2$ y finalmente $x_1 = -\frac{1}{2}x_2$.

Como $2x_1 + x_2 = 1$, entonces $2x_1 = -x_2 + 1$ y finalmente $x_1 = -\frac{1}{2}x_2 + \frac{1}{2}$.

Luego $-\frac{1}{2}x_2 = -\frac{1}{2}x_2 + \frac{1}{2}$ y entonces $0 = \frac{1}{2}$, lo cual es un absurdo que vino de suponer que existía una solución.

DEFINICIÓN 1.2.7. Consideremos un sistema como en (14) y sea A la matriz correspondiente al sistema. La *matriz ampliada* del sistema es

$$(15) \quad A' = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & y_1 \\ a_{21} & \cdots & a_{2n} & y_2 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & y_m \end{array} \right]$$

que también podemos denotar

$$A' = [A|Y].$$

El método de resolución de los sistemas no homogéneos es similar al de los sistemas homogéneos. Utilizamos el método de Gauss para la matriz A pero realizando las operaciones en toda la fila, inclusive en los elementos a la derecha de la línea: el método de Gauss se basa en intercambiar y operar con ecuaciones, así que para no cambiar las soluciones debemos trabajar con

ambos miembros de las ecuaciones. En el caso homogéneo, esto no era necesario porque siempre los segundos miembros eran idénticamente iguales a cero.

TEOREMA 1.2.9. *Sea $[A|Y]$ la matriz ampliada de un sistema no homogéneo y sea $[B|Z]$ una matriz que se obtiene a partir de $[A|Y]$ por medio de una operación elemental. Entonces, los sistemas correspondientes a $[A|Y]$ y $[B|Z]$ tiene las mismas soluciones.*

DEMOSTRACIÓN. Como $[B|Z]$ se obtiene por una operación elemental por fila a partir de $[A|Y]$, entonces las ecuaciones de $[B|Z]$ son combinaciones lineales de las ecuaciones de $[A|Y]$. Como toda operación elemental por fila tiene inversa, podemos obtener $[A|Y]$ a partir de $[B|Z]$ y por lo tanto las ecuaciones de $[A|Y]$ son combinaciones lineales de las ecuaciones de $[B|Z]$. Es decir $[A|Y]$ y $[B|Z]$ determinan sistemas de ecuaciones lineales equivalentes y por lo tanto tiene las mismas soluciones (teorema 1.1.2). \square

COROLARIO 1.2.10. *Sean $[A|Y]$ y $[B|Z]$ matrices ampliadas equivalentes por fila. Entonces, los sistemas correspondientes tiene las mismas soluciones.*

Estos resultados nos permiten usar el método de Gauss para encontrar las soluciones de un sistema no homogéneo: hacemos operaciones elementales por fila en la matriz ampliada $[A|Y]$ hasta obtener $[B|Z]$, donde $B = [b_{ij}]$ es la MERF equivalente por filas a A . Traducido a ecuaciones, obtenemos

$$\begin{aligned}
 (16) \quad & \begin{array}{rcl}
 x_{k_1} + \sum_{j \neq k_1, \dots, k_r} b_{1j} x_j & = & z_1 \\
 x_{k_2} + \sum_{j \neq k_1, \dots, k_r} b_{2j} x_j & = & z_2 \\
 \vdots & & \vdots \\
 x_{k_r} + \sum_{j \neq k_1, \dots, k_r} b_{rj} x_j & = & z_r \\
 0 & = & z_{r+1} \\
 \vdots & & \vdots \\
 0 & = & z_m
 \end{array}
 \end{aligned}$$

En el caso que para algún $i > r$ se cumpla que $z_r \neq 0$, entonces el sistema no tiene ninguna solución. Si $z_{r+1} = \dots = z_m = 0$, las soluciones son

$$\begin{aligned}
 (17) \quad & \begin{array}{rcl}
 x_{k_1} & = & -\sum_{j \neq k_1, \dots, k_r} b_{1j} x_j + z_1 \\
 x_{k_2} & = & -\sum_{j \neq k_1, \dots, k_r} b_{2j} x_j + z_2 \\
 \vdots & & \vdots \\
 x_{k_r} & = & -\sum_{j \neq k_1, \dots, k_r} b_{rj} x_j + z_r
 \end{array}
 \end{aligned}$$

Como en el caso homogéneo, x_{k_1}, \dots, x_{k_r} serán llamadas las variables principales y $\{x_j\}$ con $j \neq k_1, \dots, k_r$ son las variables libres.

EJEMPLO 1.2.11. Resolvamos el sistema

$$x_1 - 2x_2 + x_3 = 1$$

$$2x_1 + x_2 + x_3 = 2$$

$$5x_2 - x_3 = 0.$$

La matriz aumentada correspondiente a este sistema es

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 5 & -1 & 0 \end{array} \right]$$

apliquemos el método de Gauss:

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 5 & -1 & 0 \end{array} \right] &\xrightarrow{F_2 - 2F_1} \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 5 & -1 & 0 \\ 0 & 5 & -1 & 0 \end{array} \right] &\xrightarrow{F_3 - F_1} \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 5 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \\ &\xrightarrow{F_2/5} \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 1 & -1/5 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] &\xrightarrow{F_1 + 2F_2} \left[\begin{array}{ccc|c} 1 & 0 & 3/5 & 1 \\ 0 & 1 & -1/5 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]. \end{aligned}$$

Luego, el sistema se reduce a

$$x_1 + 3/5x_3 = 1$$

$$x_2 - 1/5x_3 = 0.$$

Es decir,

$$x_1 = -3/5x_3 + 1$$

$$x_2 = 1/5x_3.$$

En consecuencia, las soluciones de esta ecuación son

$$\left\{ \left(-\frac{3}{5}s + 1, \frac{1}{5}s, s \right) : s \in \mathbb{K} \right\}.$$

1.3. Álgebra de matrices

Ahora estudiaremos propiedades algebraicas de las matrices, en particular veremos que dado $n \in \mathbb{N}$, entonces podemos definir una suma y un producto en el conjunto de matrices $n \times n$ con la propiedad de que estas operaciones satisfacen muchos de los axiomas que definen a \mathbb{Z} .

1.3.1. Algunos tipos de matrices.

Matriz cuadrada. Es aquella que tiene igual número de filas que de columnas, es decir si es una matriz $n \times n$ para algún $n \in \mathbb{N}$. En ese caso, se dice que la matriz es de orden n . Por ejemplo, la matriz

$$A = \begin{bmatrix} 1 & 3 & 0 \\ -1 & 4 & 7 \\ -2 & 0 & 1 \end{bmatrix}$$

es cuadrada de orden 3.

Denotaremos el conjunto de todas las matrices cuadradas de orden n con entradas en \mathbb{K} por $M_n(\mathbb{K})$ o simplemente M_n si \mathbb{K} está sobreentendido. Así, en el ejemplo anterior $A \in M_3$.

Los elementos de la *diagonal principal* de una matriz cuadrada son aquellos que están situados en la diagonal que va desde la esquina superior izquierda hasta la inferior derecha. En otras palabras, la diagonal principal de una matriz $A = [a_{ij}]$ está formada por los elementos $a_{11}, a_{22}, \dots, a_{nn}$. En el ejemplo anterior la diagonal principal está compuesta por los elementos: $a_{11} = 1, a_{22} = 4, a_{33} = 1$.

Matriz diagonal y matriz escalar. Una matriz cuadrada, $A = [a_{ij}]$ de orden n , es *diagonal* si $a_{ij} = 0$, para $i \neq j$. Es decir, si todos los elementos situados fuera de la diagonal principal son cero. Por ejemplo, la siguiente matriz es diagonal:

$$(18) \quad \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Una matriz $n \times n$ es *escalar* si es diagonal y todos los elementos de la diagonal son iguales, por ejemplo, en el caso 4×4 las matrices escalares son

$$(19) \quad \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \end{bmatrix},$$

con $c \in \mathbb{K}$.

Matriz unidad o identidad. Esta matriz ya la hemos definido anteriormente. Recordemos que es una matriz diagonal cuya diagonal principal está compuesta de 1's.

Más adelante veremos que la matriz identidad, respecto a la multiplicación de matrices, juega un papel similar al número 1 respecto a la multiplicación de números reales o enteros (elemento neutro del producto).

Matriz nula. La *matriz nula* de orden $m \times n$, denotada $0_{m \times n}$ o simplemente 0 si m y n están sobreentendidos, es la matriz $m \times n$ cuyas entradas son todas nulas ($= 0$).

Por ejemplo, la matriz nula 2×3 es

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Veremos luego que la matriz nula juega un papel similar al número 0 en el álgebra de matrices (elemento neutro de la suma).

Matriz triangular. Una matriz cuadrada es *triangular superior* o *escalón* si todos los elementos situados por debajo de la diagonal principal son cero. Por ejemplo, la siguiente matriz es triangular superior:

$$(20) \quad \begin{bmatrix} 2 & -1 & 3 & 1 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Análogamente, una matriz cuadrada es *triangular inferior* si todos los elementos situados por encima de la diagonal principal son cero. Una matriz triangular (superior o inferior) se dice *estricta* si la diagonal principal es 0 .

En forma más precisa, sea $A = [a_{ij}] \in M_n(\mathbb{K})$, entonces

- A es *triangular superior* (*triangular superior estricta*) si $a_{ij} = 0$ para $i < j$ (resp. $i \leq j$),
- A es *triangular inferior* (*triangular inferior estricta*) si $a_{ij} = 0$ para $i > j$ (resp. $i \geq j$).

Por ejemplo, cualquier matriz diagonal es triangular superior y también triangular inferior.

No es difícil comprobar que si R es una matriz cuadrada $n \times n$ que es una MERF, entonces R es triangular superior.

1.3.2. Suma de matrices. Sean $A = [a_{ij}]$, $B = [b_{ij}]$ matrices $m \times n$. La matriz $C = [a_{ij} + b_{ij}]$ de orden $m \times n$, es decir la matriz cuyo valor en la posición ij es $a_{ij} + b_{ij}$, es llamada *la suma de las matrices A y B* y se denota $A + B$. En otras palabras, la suma de dos matrices es la matriz que resulta de sumar “coordenada a coordenada” ambas matrices.

Veamos un ejemplo, consideremos las siguientes matrices:

$$A = \begin{bmatrix} -1 & 3 & 5 \\ 2 & 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 2 & -3 \\ 0 & -1 & 1 \end{bmatrix}, \quad M = \begin{bmatrix} 2 & 1 & -3 \\ 3 & 0 & -1 \\ 0 & 8 & 5 \end{bmatrix}.$$

Las matrices A y B son de orden 2×3 , mientras la matriz M es cuadrada de orden 3 . Por tanto, no podemos calcular la suma de A y M y tampoco la suma de B y M , en cambio, sí podemos

sumar A y B ya que tienen el mismo orden. Esto es,

$$A + B = \begin{bmatrix} -1 & 3 & 5 \\ 2 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 5 & 2 & -3 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1+5 & 3+2 & 5-3 \\ 2+0 & 0-1 & -1+1 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 2 \\ 2 & -1 & 0 \end{bmatrix}$$

Dadas A , B y C matrices $m \times n$, podemos deducir fácilmente las siguientes propiedades de la suma de matrices de matrices:

- Conmutativa: $A + B = B + A$,
- Asociativa: $A + (B + C) = (A + B) + C$,
- Elemento neutro (la matriz nula): $A + 0 = 0 + A = A$,
- Elemento opuesto: existe una matriz $-A$ de orden $m \times n$ tal que $A + (-A) = (-A) + A = 0$.

Debemos explicitar la matriz opuesta: si $A = [a_{ij}]$, entonces $-A = [-a_{ij}]$. Usualmente denotaremos $A + (-B)$ como $A - B$ y $(-A) + B$ como $-A + B$.

La demostración de las propiedades anteriores se deduce de que las mismas propiedades valen coordenada a coordenada y se dejan a cargo del lector.

1.3.3. Multiplicación de matrices. Sean $A = [a_{ij}]$ matriz $m \times n$ y $B = [b_{ij}]$ matriz $n \times p$, entonces $C = [c_{ij}]$ matriz $m \times p$ es el *producto* de A y B , si

$$(21) \quad c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj}.$$

Es decir, los elementos que ocupan la posición ij en la matriz producto, se obtienen sumando los productos que resultan de multiplicar los elementos de la fila i en la primera matriz por los elementos de la columna j de la segunda matriz. Al producto de A por B lo denotamos AB . Observar que una matriz $m \times n$ por una $n \times p$ resulta en una $m \times p$.

OBSERVACIÓN 1.3.1. Sean $A = [a_{ij}]$ matriz $m \times n$ y $B = [b_{ij}]$ matriz $n \times p$, entonces si multiplicamos la matriz que se forma con la fila i de A por la matriz que determina la columna j de B , obtenemos el coeficiente ij de AB . Esquemáticamente

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = \sum_{k=1}^n a_{ik}b_{kj} = c_{ij}.$$

Por lo tanto diremos a veces, que el coeficiente ij de la matriz AB es la fila i de A por la columna j de B .

Veamos un ejemplo: si

$$A = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & -1 & 2 \\ 15 & 4 & 8 \end{bmatrix},$$

como A es 2×2 y B es 2×3 , la matriz AB será 2×3 y aplicando la regla (21), obtenemos:

$$AB = \begin{bmatrix} 1 \times 5 + 0 \times 15 & 1 \times (-1) + 0 \times 4 & 1 \times 2 + 0 \times 8 \\ -3 \times 5 + 1 \times 15 & -3 \times (-1) + 1 \times 4 & -3 \times 2 + 1 \times 8 \end{bmatrix} = \begin{bmatrix} 5 & -1 & 2 \\ 0 & 7 & 2 \end{bmatrix}.$$

Observemos que, debido a nuestra definición, no es posible multiplicar B por A , pues no está definido multiplicar una matriz 2×3 por una 2×2 . Por la definición, se puede multiplicar una matriz $m \times n$ por una matriz $r \times s$, sólo si $n = r$ y en ese caso, la multiplicación resulta ser una matriz $m \times s$.

Hay casos, como veremos en el siguiente ejemplo, en los que se pueden calcular ambos productos aunque se obtienen resultados diferentes. Consideremos las siguientes matrices:

$$A = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix}$$

Entonces, por un lado,

$$AB = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 7 \\ -4 & -8 \end{bmatrix},$$

y por otro lado,

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

Según se pudo comprobar a través del ejemplo anterior, la multiplicación de matrices no cumple la propiedad conmutativa. Veamos algunas propiedades que sí cumple esta operación:

- Asociativa:

$$A(BC) = (AB)C, \quad \forall A \in M_{m \times n}, B \in M_{n \times p}, C \in M_{p \times q},$$

- Elemento neutro: si A es matriz $m \times n$, entonces

$$AI_n = A = I_m A,$$

- Distributiva:

$$A(B + C) = AB + AC, \quad \forall A \in M_{m \times n}, B, C \in M_{n \times p},$$

y

$$(A + B)C = AC + BC, \quad \forall A, B \in M_{m \times n}, C \in M_{n \times p}.$$

Como en el caso de la suma, la demostración las propiedades anteriores se deja a cargo del lector.

En virtud de estas propiedades y de las anteriores de la suma de matrices, resulta que el conjunto $(M_n, +, \cdot)$ de las matrices cuadradas de orden n , respecto a las dos leyes de composición interna, “+” y “·”, tiene estructura de anillo unitario no conmutativo (ver [https://es.wikipedia.org/wiki/Anillo_\(matemática\)](https://es.wikipedia.org/wiki/Anillo_(matemática))).

Cuando las matrices son cuadradas podemos multiplicarlas por si mismas y definimos, de forma análoga a lo que ocurre en los productos de números, la potencia de una matriz: sea A matriz $n \times n$, y sea $m \in \mathbb{N}$ entonces

$$A^0 = I_n, \quad A^m = A^{m-1}A,$$

es decir A^m es multiplicar A consigo mismo m -veces.

OBSERVACIÓN 1.3.1. Un caso especial de multiplicación es la multiplicación por matrices diagonales. Sea

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

matriz $n \times n$ diagonal con valor d_i en la posición ii , entonces si A es matriz $n \times p$, tenemos que DA es la matriz que en la fila i tiene a la fila i de A multiplicada por d_i . Es decir,

$$\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} = \begin{bmatrix} d_1 a_{11} & d_1 a_{12} & \cdots & d_1 a_{1p} \\ d_2 a_{21} & d_2 a_{22} & \cdots & d_2 a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ d_n a_{n1} & d_n a_{n2} & \cdots & d_n a_{np} \end{bmatrix}.$$

Esto es claro, pues el coeficiente ij de DA es la fila i de D por la columna j de A , es decir

$$[DA]_{ij} = 0.a_{1j} + \cdots + 0.a_{i-1,j} + d_i.a_{ij} + 0.a_{i+1,j} + \cdots + 0.a_{nj} = d_i a_{ij}.$$

Observar que en el caso de que D sea una matriz escalar (es decir $d_1 = d_2 = \cdots = d_n$), DA es multiplicar por el mismo número todos los coeficientes de A . En particular, en este caso, si A es $n \times n$, $DA = AD$.

Si B es $m \times n$, ¿qué se obtiene haciendo BD ?

Otras observaciones importantes:

- multiplicar cualquier matriz por la matriz nula resulta la matriz nula,
- existen divisores de cero: en general, $AB = 0$ no implica que $A = 0$ o $B = 0$ o, lo que es lo mismo, el producto de matrices no nulas puede resultar en una matriz nula. Por ejemplo,

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 8 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

- En general no se cumple la propiedad cancelativa: si $A \neq 0$ y $AB = AC$ no necesariamente se cumple que $B = C$. Por ejemplo,

$$\begin{bmatrix} 2 & 0 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 8 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 5 & 3 \end{bmatrix}$$

- No se cumple la fórmula del binomio: sean A, B matrices $n \times n$, entonces

$$\begin{aligned} (A + B)^2 &= (A + B)(A + B) \\ &= A(A + B) + B(A + B) \\ &= AA + AB + BA + BB \\ &= A^2 + AB + BA + B^2, \end{aligned}$$

y esta última expresión puede no ser igual a $A^2 + 2AB + B^2$ ya que el producto de matrices no es conmutativo (en general).

1.3.4. Multiplicación de una matriz por un escalar. Otra operación importante es la multiplicación de una matriz por un elemento de \mathbb{K} : sea $A = [a_{ij}]$ matriz $m \times n$ y $c \in \mathbb{K}$, entonces el producto de c por A es la matriz

$$cA = [ca_{ij}].$$

Por ejemplo,

$$2 \begin{bmatrix} -1 & 0 & 3 & 4 \\ 5 & 1 & -2 & 1 \\ 3 & 2 & 1 & -3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 6 & 8 \\ 10 & 2 & -4 & 2 \\ 6 & 4 & 2 & -6 \end{bmatrix}.$$

Observar que multiplicar por c una matriz $m \times n$, es lo mismo que multiplicar por la matriz escalar $m \times m$ con los coeficientes de la diagonal iguales a c , es decir

$$(22) \quad cA = \begin{bmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

Debido a esta observación y a las propiedades del producto de matrices, se cumple lo siguiente:

$$\begin{aligned} c(AB) &= (cA)B, & \forall c \in \mathbb{K}, A \in M_{m \times n}, B \in M_{n \times p}, \\ 1.A &= A, & \forall c \in \mathbb{K}, A \in M_{m \times n} \\ c(A + B) &= cA + cB, & \forall c \in \mathbb{K}, A, B \in M_{m \times n}, \\ (c + d)A &= cA + dA, & \forall c, d \in \mathbb{K}, A \in M_{m \times n}. \end{aligned}$$

Si A es $n \times n$, entonces $DA = AD$ cuando D es una matriz escalar. Por lo tanto

$$c(AB) = (cA)B = A(cB), \quad \forall c \in \mathbb{K}, A \in M_{n \times n}, B \in M_{n \times n}.$$

1.4. Matrices elementales

Veremos ahora la relación entre el álgebra de matrices y la solución de sistemas de ecuaciones lineales.

Primero recordemos que dado un sistema de m ecuaciones lineales con n incógnitas

$$(23) \quad \begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array}$$

donde y_1, \dots, y_m y $a_{i,j}$ ($1 \leq i \leq m$, $1 \leq j \leq n$) son números en \mathbb{K} . Si denotamos

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix},$$

entonces

$$AX = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = Y$$

(producto de matrices). Es decir, la notación antes utilizada es consistente con el, ahora definido, producto de matrices.

DEFINICIÓN 1.4.1. Una matriz $m \times m$ se dice *elemental* si fue obtenida por medio de una única operación elemental a partir de la matriz identidad I_m .

EJEMPLO 1.4.1. Veamos cuales son las matrices elementales 2×2 :

- (1) Si $c \neq 0$, multiplicar por c la primera fila y multiplicar c por la segunda fila son, respectivamente,

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix},$$

- (2) si $c \in \mathbb{K}$, sumar a la fila 2 la fila 1 multiplicada por c o sumar a la fila 1 la fila 2 multiplicada por c son, respectivamente,

$$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}.$$

(3) Finalmente, intercambiando la fila 1 por la fila 2 obtenemos la matriz

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

En el caso de matrices $m \times m$ tampoco es difícil encontrar las matrices elementales:

- (1) Si $c \neq 0$, multiplicar por c la fila k de la matriz identidad, resulta en la matriz elemental que tiene todos 1's en la diagonal, excepto en la posición k, k donde vale c , es decir si $e(I_m) = [a_{ij}]$, entonces

$$(24) \quad a_{ij} = \begin{cases} 1 & \text{si } i = j \text{ e } i \neq k, \\ c & \text{si } i = j = k, \\ 0 & \text{si } i \neq j. \end{cases}$$

Gráficamente,

$$\begin{array}{c} k \\ \downarrow \\ \xrightarrow{k} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & c & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{bmatrix} \end{array}$$

- (2) si $c \in \mathbb{K}$, sumar a la fila r la fila s multiplicada por c , resulta en la matriz elemental que tiene todos 1's en la diagonal, y todos los demás coeficientes son 0, excepto en la fila r y columna s donde vale c , es decir si $e(I_m) = [a_{ij}]$, entonces

$$(25) \quad a_{ij} = \begin{cases} 1 & \text{si } i = j \\ c & \text{si } i = r, j = s, \\ 0 & \text{otro caso.} \end{cases}$$

Gráficamente,

$$\begin{array}{c} \begin{array}{cc} r & s \\ \downarrow & \downarrow \end{array} \\ \xrightarrow{r} \begin{bmatrix} 1 & 0 & \cdots & & & 0 \\ \vdots & \ddots & & & & \vdots \\ 0 & \cdots & 1 & \cdots & c & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & & & 1 \end{bmatrix} \end{array}$$

(3) Finalmente, intercambiar la fila r por la fila s resulta ser

$$(26) \quad a_{ij} = \begin{cases} 1 & \text{si } (i = j, i \neq r, i \neq s) \text{ o } (i = r, j = s) \text{ o } (i = s, j = r) \\ 0 & \text{otro caso.} \end{cases}$$

Gráficamente,

$$\begin{array}{c} \begin{array}{cc} r & s \\ \downarrow & \downarrow \end{array} \\ \begin{array}{c} \xrightarrow{r} \\ \xrightarrow{s} \end{array} \begin{bmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} \end{array}$$

Veamos ahora que, dada una matriz A , hacer una operación elemental en A es igual a multiplicar A a izquierda por una matriz elemental. Más precisamente:

TEOREMA 1.4.1. *Sea e una operación elemental por fila y sea E la matriz elemental $E = e(I)$. Entonces $e(A) = EA$.*

DEMOSTRACIÓN. Hagamos la prueba para matrices 2×2 . La prueba en general es similar, pero requiere de un complicado manejo de índices.

(1) Si $c \neq 0$, multiplicar I_2 por c la primera fila resulta en la matriz elemental

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}.$$

Luego

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} c \cdot a_{11} + 0 \cdot a_{21} & c \cdot a_{12} + 0 \cdot a_{22} \\ 0 \cdot a_{11} + 1 \cdot a_{21} & 0 \cdot a_{12} + 1 \cdot a_{22} \end{bmatrix} = \begin{bmatrix} c \cdot a_{11} & c \cdot a_{12} \\ a_{21} & a_{22} \end{bmatrix} = e(A).$$

De forma análoga se demuestra en el caso que la operación elemental sea multiplicar la segunda fila por c .

- (2) Si $c \in \mathbb{K}$, cambiar en I_2 la fila 2 por la fila 1 multiplicada por c resulta en la matriz elemental

$$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}.$$

Luego

$$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ c \cdot a_{11} + a_{21} & c \cdot a_{12} + a_{22} \end{bmatrix} = e(A).$$

La demostración es análoga si la operación elemental es cambiar la fila 1 por la fila 2 multiplicada por c .

- (3) Finalmente, si intercambiamos en I_2 la fila 1 por la fila 2 obtenemos la matriz elemental

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Luego

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix} = e(A).$$

□

COROLARIO 1.4.2. Sean A y B matrices $m \times n$. Entonces B equivalente por filas a A si y sólo si $B = PA$ donde P es producto de matrices elementales. Más aún, si $B = e_k(e_{k-1}(\cdots(e_1(A))\cdots))$ con e_1, e_2, \dots, e_k operaciones elementales de fila y $E_i = e_i(I)$ para $i = 1, \dots, k$, entonces $B = E_k E_{k-1} \cdots E_1 A$.

DEMOSTRACIÓN.

(\Rightarrow) Si B equivalente por filas a A existen operaciones elementales e_1, \dots, e_k tal que $B = e_k(e_{k-1}(\cdots(e_1(A))\cdots))$, más formalmente

si $A_1 = e_1(A)$ y $A_i = e_i(A_{i-1})$ para $i = 2, \dots, k$, entonces $e_k(A_{k-1}) = B$.

Sea $E_i = e_i(I_m)$, entonces, por el teorema anterior $A_1 = E_1 A$ y $A_i = E_i A_{i-1}$ ($i = 2, \dots, k$). Por lo tanto $B = E_k A_{k-1}$, en otras palabras $B = E_k E_{k-1} \cdots E_1 A$, luego $P = E_k E_{k-1} \cdots E_1$.

(\Leftarrow) Si $B = PA$, con $P = E_k E_{k-1} \cdots E_1$ donde $E_i = e_i(I_m)$ es una matriz elemental, entonces

$$B = PA = E_k E_{k-1} \cdots E_1 A \stackrel{\text{Teor. 1.4.1}}{=} e_k(e_{k-1}(\cdots(e_1(A))\cdots)).$$

Por lo tanto, B es equivalente por filas a A .

□

1.5. Matrices inversibles

DEFINICIÓN 1.5.1. Sea A una matriz $n \times n$ con coeficientes en \mathbb{K} . Una matriz $B \in M_{n \times n}(\mathbb{K})$ es *inversa de A* si $BA = AB = I_n$. En ese caso, diremos que A es *inversible*.

EJEMPLO 1.5.1. La matriz $\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$ tiene inversa $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$ pues es fácil comprobar que

$$\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

PROPOSICIÓN 1.5.1. Sea $A \in M_{n \times n}(\mathbb{K})$,

- (1) sean $B, C \in M_{n \times n}(\mathbb{K})$ tales que $BA = I_n$ y $AC = I_n$, entonces $B = C$;
- (2) si A inversible la inversa es única.

DEMOSTRACIÓN. (1)

$$B = BI_n = B(AC) = (BA)C = I_n C = C.$$

(2) Sean B y C inversas de A , es decir $BA = AB = I_n$ y $CA = AC = I_n$. En particular, $BA = I_n$ y $AC = I_n$, luego, por (1), $B = C$. \square

DEFINICIÓN 1.5.2. Sea $A \in M_{n \times n}(\mathbb{K})$ inversible. A la única matriz inversa de A la llamamos *la matriz inversa de A* y la denotamos A^{-1} .

Veremos más adelante que si una matriz $n \times n$ admite una inversa a izquierda, es decir si existe B tal que $BA = I_n$, entonces la matriz es inversible. Lo mismo vale si A admite inversa a derecha.

EJEMPLO 1.5.2. Sea A la matriz

$$\begin{bmatrix} 2 & 1 & -2 \\ 1 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}.$$

Entonces, A es inversible y su inversa es

$$A^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \\ 1 & -1 & 1 \end{bmatrix}.$$

Esto se resuelve comprobando que $AA^{-1} = I_3$ (por lo dicho más arriba es innecesario comprobar que $A^{-1}A = I_3$).

OBSERVACIÓN 1.5.1. No toda matriz tiene inversa, por ejemplo la matriz nula (cuyos coeficientes son todos iguales a 0) no tiene inversa pues $0 \cdot A = 0 \neq I$. También existen matrices no nulas no inversibles, por ejemplo la matriz

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

no tiene inversa. Si multiplicamos a A por una cualquier matriz $B = [b_{ij}]$ obtenemos

$$AB = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 2b_{11} + b_{21} & 2b_{12} + b_{22} \\ 0 & 0 \end{bmatrix}.$$

Luego AB , al tener una fila idénticamente nula, no puede ser nunca la identidad.

TEOREMA 1.5.2. Sean A y B matrices $n \times n$ con coeficientes en \mathbb{K} . Entonces

- (1) si A inversible, entonces A^{-1} es inversible y su inversa es A , es decir $(A^{-1})^{-1} = A$;
- (2) si A y B son inversibles, entonces AB es inversible y $(AB)^{-1} = B^{-1}A^{-1}$.

DEMOSTRACIÓN. (1) La inversa a izquierda de A^{-1} es A , pues $AA^{-1} = I_n$. Análogamente, la inversa a derecha de A^{-1} es A , pues $A^{-1}A = I_n$. Concluyendo: A es la inversa de A^{-1} .

(2) Simplemente debemos comprobar que $B^{-1}A^{-1}$ es inversa a izquierda y derecha de AB :

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n,$$

y, análogamente,

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

□

OBSERVACIÓN 1.5.2. Si A_1, \dots, A_k son inversibles, entonces $A_1 \dots A_k$ es inversible y su inversa es

$$(A_1 \dots A_k)^{-1} = A_k^{-1} \dots A_1^{-1}.$$

El resultado es una generalización del punto (2) del teorema anterior y su demostración se hace por inducción en k (usando (2) del teorema anterior). Se deja como ejercicio al lector.

OBSERVACIÓN 1.5.3. La suma de matrices inversibles no necesariamente es inversible, por ejemplo $A + (-A) = 0$ que no es inversible.

TEOREMA 1.5.3. Una matriz elemental es inversible.

DEMOSTRACIÓN. Sea E la matriz elemental que se obtiene a partir de I_n por la operación elemental e . Se e_1 la operación elemental inversa (teorema 1.2.1) y $E_1 = e_1(I_n)$. Entonces

$$EE_1 = e(e_1(I_n)) = I_n$$

$$E_1E = e_1(e(I_n)) = I_n.$$

Luego $E_1 = E^{-1}$.

□

EJEMPLO 1.5.3. Es fácil encontrar explícitamente la matriz inversa de una matriz elemental, por ejemplo, en el caso 2×2 tenemos:

(1) Si $c \neq 0$,

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1/c & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1/c \end{bmatrix},$$

(2) si $c \in \mathbb{K}$,

$$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -c \\ 0 & 1 \end{bmatrix}.$$

(3) Finalmente,

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

En el caso general tenemos:

(1)

$$\begin{array}{c} \downarrow k \\ \text{La inversa de} \end{array} \xrightarrow{k} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & c & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{bmatrix} \quad \text{es} \quad \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 1/c & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{bmatrix}$$

(2)

$$\begin{array}{c} \begin{array}{cc} \downarrow r & \downarrow s \end{array} \\ \text{La inversa de} \end{array} \xrightarrow{r} \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & 1 & \cdots & c & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & & & & 1 \end{bmatrix} \quad \text{es} \quad \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & 1 & \cdots & -c & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & & & & 1 \end{bmatrix}$$

(3)

La inversa de

$$\begin{array}{c} \xrightarrow{r} \\ \xrightarrow{s} \end{array} \begin{bmatrix} 1 & \cdots & & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & & \cdots & \cdots & & 1 \end{bmatrix}$$

es la misma matriz.

- i) A es inversible,
- ii) A es equivalente por filas a I_n ,
- iii) A es producto de matrices elementales.

$i) \Rightarrow ii)$ Sea R la matriz escalón reducida por fila equivalente por filas a A . Entonces, existen E_1, \dots, E_k matrices elementales tal que $E_1, \dots, E_k A = R$. Como las matrices elementales son inversibles, el producto de matrices elementales es inversible, luego E_1, \dots, E_k es inversible y por lo tanto $R = E_1, \dots, E_k A$ es inversible.

ii) \Rightarrow iii) Como A es equivalente por filas a I_n , al ser la equivalencia por filas una relación de equivalencia, tenemos que I_n es equivalente por filas a A , es decir existen E_1, \dots, E_k matrices elementales, tales que $E_1 E_2, \dots, E_k I_n = A$. Por lo tanto, $A = E_1 E_2, \dots, E_k$ producto de matrices elementales.

iii) \Rightarrow i) Sea $A = E_1 E_2, \dots, E_k$ donde E_i es una matriz elemental ($i = 1, \dots, k$). Como cada E_i es inversible, el producto de ellos es inversible, por lo tanto A es inversible. \square

DEMOSTRACIÓN.

(\Rightarrow) B es equivalente por filas a A , luego existe P matriz producto de matrices elementales tal que $B = PA$. Como cada matriz elemental es inversible (teorema 1.5.3) y el producto de matrices inversibles es inversible (teorema 1.5.2(2)), se deduce que P es inversible.

(\Leftarrow) Sea P matriz inversible tal que $B = PA$. Como P es inversible, por el teorema anterior, P es producto de matrices elementales, luego $B = PA$ es equivalente por filas a A . \square

COROLARIO 1.5.6. Sea A matriz $n \times n$. Sean e_1, \dots, e_k las operaciones elementales por filas que reducen a A a una MERF y esta MERF es la identidad, es decir $e_1(e_2(\dots(e_k(A))\dots)) = I_n$. Entonces, A inversible y las mismas operaciones elementales aplicadas a I_n nos llevan a A^{-1} , es decir $e_1(e_2(\dots(e_k(I_n))\dots)) = A^{-1}$.

DEMOSTRACIÓN. Por el teorema anterior, al ser A equivalente por filas a la identidad, A es inversible. Sean las matrices elementales $E_i = e_i(I_n)$ para $i = 1, \dots, k$, entonces (ver corolario 1.4.2) $E_1 E_2 \dots E_k A = I_n$, por lo tanto, multiplicando por A^{-1} a derecha en ambos miembros,

$$\begin{aligned} E_1 E_2 \dots E_k A A^{-1} &= I_n A^{-1} \quad \Leftrightarrow \\ E_1 E_2 \dots E_k I_n &= A^{-1} \quad \Leftrightarrow \\ e_1(e_2(\dots(e_k(I_n))\dots)) &= A^{-1}. \end{aligned}$$

\square

Este último corolario nos provee un método sencillo para calcular la inversa de una matriz A (inversible). Primero, encontramos $R = I_n$ la MERF equivalente por filas a A , luego, aplicando la mismas operaciones elementales a I_n , obtenemos la inversa de A . Para facilitar el cálculo es conveniente comenzar con A e I_n e ir aplicando paralelamente las operaciones elementales por fila. Veamos un ejemplo.

EJEMPLO 1.5.4. Calculemos la inversa (si tiene) de

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}.$$

SOLUCIÓN. Por lo que ya hemos demostrado 1) si A tiene inversa es reducible por filas a la identidad, 2) las operaciones que llevan a A a la identidad, llevan también la identidad a A^{-1} . Luego trataremos de reducir por filas a A y todas las operaciones elementales las haremos en paralelo partiendo de la matriz identidad:

$$\begin{aligned} [A|I] &= \left[\begin{array}{cc|cc} 2 & -1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array} \right] \xrightarrow{F_1 \leftrightarrow F_2} \left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 2 & -1 & 1 & 0 \end{array} \right] \xrightarrow{F_2 - 2F_1} \left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & -7 & 1 & -2 \end{array} \right] \xrightarrow{F_2 / (-7)} \\ &\rightarrow \left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} \end{array} \right] \xrightarrow{F_1 - 3F_2} \left[\begin{array}{cc|cc} 1 & 0 & \frac{3}{7} & \frac{1}{7} \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} \end{array} \right]. \end{aligned}$$

Luego, como A se reduce por filas a la identidad, A es inversible y su inversa es

$$A^{-1} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}.$$

El lector desconfiado podrá comprobar, haciendo el producto de matrices, que $AA^{-1} = A^{-1}A = I_2$. \square

TEOREMA 1.5.7. *Sea A matriz $n \times n$ con coeficientes en \mathbb{K} . Entonces, las siguientes afirmaciones son equivalentes.*

- i) A es inversible.*
- ii) El sistema $AX = Y$ tiene una única solución para toda matriz Y de orden $n \times 1$.*
- iii) El sistema homogéneo $AX = 0$ tiene una única solución trivial.*

DEMOSTRACIÓN.

$i) \Rightarrow ii)$ Sea X_0 solución del sistema $AX = Y$, luego

$$AX_0 = Y \quad \Rightarrow \quad A^{-1}AX_0 = A^{-1}Y \quad \Rightarrow \quad X_0 = A^{-1}Y.$$

Es decir, X_0 es único (siempre igual a $A^{-1}Y$).

$ii) \Rightarrow i)$ Es trivial, tomando $Y = 0$.

$iii) \Rightarrow i)$ Sea R la matriz escalón reducida por filas equivalente a A , es decir $R = PA$ con P inversible y R es MERF. Si R tiene una fila nula, entonces por teorema 1.2.5, el sistema $AX = 0$ tiene más de una solución, lo cual es absurdo. Por lo tanto, R no tiene filas nulas. Como es una matriz cuadrada y es MERF, tenemos que $R = I_n$. Luego A es equivalente por filas a I_n y por teorema 1.5.4 se deduce que A es inversible. \square

COROLARIO 1.5.8. *Sea A una matriz $n \times n$ con coeficientes en \mathbb{K} . Si A tiene inversa a izquierda, es decir si existe B matriz $n \times n$ tal que $BA = I_n$, entonces A es inversible. Lo mismo vale si A tiene inversa a derecha.*

DEMOSTRACIÓN. Supongamos que A tiene inversa a izquierda y que B sea la inversa a izquierda, es decir $BA = I_n$. El sistema $AX = 0$ tiene una única solución, pues $AX_0 = 0 \Rightarrow BAX_0 = B0 \Rightarrow X_0 = 0$. Luego, A es inversible (y su inversa es B).

Supongamos que A tiene inversa a derecha y que C sea la inversa a derecha, es decir $AC = I$. Por lo anterior, C es inversible y su inversa es A , es decir $AC = I$ y $CA = I$, luego A es inversible. \square

Terminaremos la sección calculando algunas matrices inversas usando el corolario 1.5.6.

EJEMPLO 1.5.5. Calcular la inversa (si tiene) de la matriz $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 2 & 4 \\ 0 & 1 & -2 \end{bmatrix}$.

SOLUCIÓN.

$$\begin{aligned}
& \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 3 & 2 & 4 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array} \right] \xrightarrow{F_2-3F_1} \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 5 & -2 & -3 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array} \right] \xrightarrow{F_2 \leftrightarrow F_3} \\
& \rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 5 & -2 & -3 & 1 & 0 \end{array} \right] \xrightarrow{F_1+F_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 5 & -2 & -3 & 1 & 0 \end{array} \right] \xrightarrow{F_3-5F_2} \\
& \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 8 & -3 & 1 & -5 \end{array} \right] \xrightarrow{F_3/8} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{array} \right] \xrightarrow{F_2+2F_3} \\
& \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{4} & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{array} \right].
\end{aligned}$$

Por lo tanto

$$A^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -\frac{3}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{bmatrix}$$

□

EJEMPLO 1.5.6. Dados $a, b, c, d \in \mathbb{R}$, determinar cuando la matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es inversible y en ese caso, cual es su inversa.

SOLUCIÓN. Para poder aplicar el método de Gauss, debemos ir haciendo casos.

1) Supongamos que $a \neq 0$, entonces

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{F_1/a} \begin{bmatrix} 1 & \frac{b}{a} \\ c & d \end{bmatrix} \xrightarrow{F_2-cF_1} \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & d - c\frac{b}{a} \end{bmatrix} = \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & \frac{ad-bc}{a} \end{bmatrix}$$

Si $ad - bc = 0$, entonces la matriz se encuentra reducida por filas y la última fila es 0, luego en ese caso no es inversible. Si $ad - bc \neq 0$, entonces

$$\begin{bmatrix} 1 & \frac{b}{a} \\ 0 & \frac{ad-bc}{a} \end{bmatrix} \xrightarrow{a/(ad-bc) F_2} \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{bmatrix} \xrightarrow{F_1-b/a F_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Luego, en el caso $a \neq 0$, $ad - bc \neq 0$ hemos reducido por filas la matriz A a la identidad y por lo tanto A es inversible. Además, podemos encontrar A^{-1} aplicando a I las mismas operaciones

elementales que reducían A a la identidad:

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} &\xrightarrow{F_1/a} \begin{bmatrix} \frac{1}{a} & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{F_2 - cF_1} \begin{bmatrix} \frac{1}{a} & 0 \\ -\frac{c}{a} & 1 \end{bmatrix} \xrightarrow{a/(ad-bc) F_2} \begin{bmatrix} \frac{1}{a} & 0 \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \xrightarrow{F_1 - b/a F_2} \\ &\rightarrow \begin{bmatrix} \frac{1}{a} + \frac{bc}{a(ad-bc)} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}. \end{aligned}$$

Concluyendo, en el caso $a \neq 0$, $ad - bc \neq 0$, A es invertible y

$$(27) \quad A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

2) Estudiemos el caso $a = 0$. Primero observemos que si $c = 0$ o $b = 0$, entonces la matriz no es invertible, pues en ambos casos nos quedan matrices que no pueden ser reducidas por fila a la identidad. Luego la matriz puede ser invertible si $bc \neq 0$ y en este caso la reducción por filas es:

$$\begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \xrightarrow{F_1 \leftrightarrow F_2} \begin{bmatrix} c & d \\ 0 & b \end{bmatrix} \xrightarrow{F_1/c} \begin{bmatrix} 1 & \frac{d}{c} \\ 0 & b \end{bmatrix} \xrightarrow{F_2/b} \begin{bmatrix} 1 & \frac{d}{c} \\ 0 & 1 \end{bmatrix} \xrightarrow{F_1 - d/c F_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Luego A es invertible y aplicando estas mismas operaciones elementales a la identidad obtenemos la inversa:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{F_1 \leftrightarrow F_2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \xrightarrow{F_1/c} \begin{bmatrix} 0 & \frac{1}{c} \\ 1 & 0 \end{bmatrix} \xrightarrow{F_2/b} \begin{bmatrix} 0 & \frac{1}{c} \\ \frac{1}{b} & 0 \end{bmatrix} \xrightarrow{F_1 - d/c F_2} \begin{bmatrix} -\frac{d}{bc} & \frac{1}{c} \\ \frac{1}{b} & 0 \end{bmatrix}.$$

Luego, en el caso que $a = 0$, entonces A invertible si $bc \neq 0$ y su inversa es

$$A^{-1} = \begin{bmatrix} -\frac{d}{bc} & \frac{1}{c} \\ \frac{1}{b} & 0 \end{bmatrix} = \frac{1}{-bc} \begin{bmatrix} d & -b \\ -c & 0 \end{bmatrix}.$$

Es decir, la expresión de la inversa es igual a (27) (considerando que $a = 0$).

Reuniendo los dos casos: A es invertible si $a \neq 0$ y $ad - bc \neq 0$ o si $a = 0$ y $bc \neq 0$, pero esto es lógicamente equivalente a pedir solamente $ad - bc \neq 0$, es decir

$$(a \neq 0 \wedge ad - bc \neq 0) \vee (a = 0 \wedge bc \neq 0) \Leftrightarrow ad - bc \neq 0$$

(ejercicio).

Resumiendo, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es inversible $\Leftrightarrow ad - bc \neq 0$ y en ese caso, su inversa viene dada por

$$(28) \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Veremos en la próxima sección que el uso de determinantes permitirá establecer la generalización de este resultado para matrices $n \times n$ con $n \geq 1$.

□

1.6. Determinante

El determinante puede ser pensado como una función que a cada matriz cuadrada $n \times n$ con coeficientes en \mathbb{K} , le asocia un elemento de \mathbb{K} . En esta sección veremos como se define esta función y algunas propiedades de la misma. Algunas demostraciones se omitirán, pues se pondrá énfasis en los usos del determinante y no tanto en sus propiedades teóricas. Las demostraciones faltantes se pueden ver en el Apéndice A.

Una forma de definir determinante es mediante una definición recursiva, es decir para calcular el determinante de una matriz $n \times n$, usaremos el cálculo del determinante para matrices $n - 1 \times n - 1$, que a su vez se calcula usando el determinante de matrices $n - 2 \times n - 2$ y así sucesivamente hasta llegar al caso base, que es el caso de matrices 1×1 .

DEFINICIÓN 1.6.1. Sea $A \in M_n(\mathbb{K})$. Sean i, j tal que $1 \leq i, j \leq n$. Entonces $A(i|j)$ es la matriz $n - 1 \times n - 1$ que se obtiene eliminando la fila i y la columna j de A .

EJEMPLO 1.6.1. Sea $A = \begin{bmatrix} 1 & -1 & 3 \\ 4 & 2 & -5 \\ 0 & 7 & 3 \end{bmatrix}$, entonces

$$A(1|1) = \begin{bmatrix} 2 & -5 \\ 7 & 3 \end{bmatrix}, \quad A(2|3) = \begin{bmatrix} 1 & -1 \\ 0 & 7 \end{bmatrix}, \quad A(3|1) = \begin{bmatrix} -1 & 3 \\ 2 & -5 \end{bmatrix}.$$

DEFINICIÓN 1.6.2. Sea $n \in \mathbb{N}$ y $A = [a_{ij}] \in M_n(\mathbb{K})$, entonces el *determinante de A*, denotado $\det(A)$ se define como:

- (1) si $n = 1$, $\det([a]) = a$;
- (n) si $n > 1$,

$$\begin{aligned} \det(A) &= a_{11} \det A(1|1) - a_{21} \det A(2|1) + \cdots + (-1)^{1+n} a_{n1} \det A(n|1) \\ &= \sum_{i=1}^n (-1)^{1+i} a_{i1} \det A(i|1). \end{aligned}$$

Si $1 \leq i, j \leq n$, al número $\det A(i|j)$ se lo llama el *menor i, j de A* y a $C_{ij}^A := (-1)^{i+j} \det A(i|j)$ se lo denomina el *cofactor i, j de A* . Si la matriz A está sobreentendida se denota, a veces, $C_{ij} := C_{ij}^A$.

Observemos, que con las definiciones introducidas tenemos

$$(29) \quad \det(A) = \sum_{i=1}^n a_{i1} C_{i1}^A.$$

A este cálculo se lo denomina *calculo del determinante por desarrollo por la primera columna*, debido a que usamos los coeficientes de la primera columna, multiplicados por los cofactores correspondientes. A veces, para simplificar, denotaremos

$$|A| := \det A.$$

OBSERVACIÓN 1.6.1. Determinantes 2×2 . Calculemos el determinante de las matrices 2×2 . Sea

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

entonces

$$\det A = a \det[d] - c \det[b] = ad - bc.$$

Hemos visto en el ejemplo 1.5.6 que A es inversible si y solo si $ad - bc \neq 0$, es decir

$$(30) \quad A \text{ es inversible si y solo si } \det A \neq 0.$$

Este resultado se generaliza para matrices $n \times n$. Más aún, la fórmula (27) se generaliza también para matrices cuadradas de cualquier dimensión (ver el corolario A.2.3).

OBSERVACIÓN 1.6.1. Determinantes 3×3 . Calculemos el determinante de las matrices 3×3 . Sea

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

entonces

$$\begin{aligned} \det A &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{13}a_{32}) + a_{13}(a_{12}a_{23} - a_{13}a_{22}) \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}. \end{aligned}$$

Luego

$$\begin{aligned}\det A &= 1 \times (-1) \times 4 + 3 \times 5 \times 2 + 2 \times (-2) \times 1 \\ &\quad - 2 \times (-1) \times 2 - 1 \times 5 \times 1 - 3 \times (-2) \times 4 \\ &= -4 + 30 - 4 + 4 - 5 + 24 \\ &= 35.\end{aligned}$$

OBSERVACIÓN 1.6.2. La regla para calcular el determinante de matrices 3×3 **no** se aplica a matrices $n \times n$ con $n \neq 3$.

PROPOSICIÓN 1.6.1. Sea $A \in M_n(\mathbb{K})$ matriz triangular superior cuyos elementos en la diagonal son d_1, \dots, d_n . Entonces $\det A = d_1 \cdot d_2 \cdot \dots \cdot d_n$.

DEMOSTRACIÓN. Podemos demostrar el resultado por inducción sobre n : es claro que si $n = 1$, es decir si $A = [d_1]$, el determinante vale d_1 . Por otro lado, si $n > 1$, observemos que $A(1|1)$ es también triangular superior con valores d_2, \dots, d_n en la diagonal principal. Entonces, usamos la definición de la fórmula (29) y observamos que el desarrollo por la primera columna solo tiene un término, pues esta columna solo tiene un coeficiente no nulo, el d_1 en la primera posición. Por lo tanto,

$$\det(A) = d_1 \det(A(1|1)) \stackrel{(\text{HI})}{=} d_1 \cdot (d_2 \cdot \dots \cdot d_n).$$

□

COROLARIO 1.6.2. $\det I_n = 1$.

DEMOSTRACIÓN. Se deduce del hecho que I_n es triangular superior y todo coeficiente de la diagonal principal vale 1. □

COROLARIO 1.6.3. Si R es una MERF, entonces

$$\det R = \begin{cases} 1 & \text{si } R \text{ no tiene filas nulas,} \\ 0 & \text{si } R \text{ tiene filas nulas.} \end{cases}$$

DEMOSTRACIÓN. Si R no tiene filas nulas es igual a I_n (lema 1.2.7), luego $\det R = 1$. En general, R es una matriz triangular superior y si tiene alguna fila nula r , entonces el coeficiente r_r es igual a 0 y por lo tanto $\det R = 0$. □

EJEMPLO 1.6.3. Veamos, en el caso de una matriz $A = [a_{ij}]$ de orden 2×2 que ocurre con el determinante cuando hacemos una operación elemental.

- (1) Si $c \neq 0$, multiplicar por c la primera fila y multiplicar c por la segunda fila obtenemos, respectivamente,

$$e(A) = \begin{bmatrix} ca_{11} & ca_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{y} \quad e(A) = \begin{bmatrix} a_{11} & a_{12} \\ ca_{21} & ca_{22} \end{bmatrix},$$

luego

$$\det \begin{bmatrix} ca_{11} & ca_{12} \\ a_{21} & a_{22} \end{bmatrix} = ca_{11}a_{22} - ca_{12}a_{21} \quad \text{y} \quad \det \begin{bmatrix} ca_{11} & ca_{12} \\ a_{21} & a_{22} \end{bmatrix} = ca_{11}a_{22} - ca_{12}a_{21}.$$

Por lo tanto, en ambos casos, $\det e(A) = c \det A$.

- (2) Sea $c \in \mathbb{K}$, si sumamos a la fila 2 la fila 1 multiplicada por c o sumamos a la fila 1 la fila 2 multiplicada por c obtenemos, respectivamente,

$$e(A) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} + ca_{11} & a_{22} + ca_{12} \end{bmatrix} \quad \text{y} \quad e(A) = \begin{bmatrix} a_{11} + ca_{21} & a_{12} + ca_{22} \\ a_{21} & a_{22} \end{bmatrix}.$$

Por lo tanto,

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} + ca_{11} & a_{22} + ca_{12} \end{bmatrix} &= a_{11}(a_{22} + ca_{12}) - a_{12}(a_{21} + ca_{11}) \\ &= a_{11}a_{22} + ca_{11}a_{12} - a_{12}a_{21} - ca_{12}a_{11} \\ &= a_{11}a_{22} - a_{12}a_{21} \\ &= \det A. \end{aligned}$$

En el otro caso también se comprueba que $\det e(A) = \det A$.

- (3) Finalmente, intercambiando la fila 1 por la fila 2 obtenemos la matriz

$$e(A) = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix},$$

por lo tanto

$$\det e(A) = \det \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix} = a_{21}a_{12} - a_{22}a_{11} = -\det A.$$

Todos los resultado del ejemplo anterior se pueden generalizar.

TEOREMA 1.6.4. Sea $A \in M_n(\mathbb{K})$ y sean $1 \leq r, s \leq n$.

- (1) Sea $c \in \mathbb{K}$ y B la matriz que se obtiene de A multiplicando la fila r por c , es decir $A \xrightarrow{cF_r} B$, entonces $\det B = c \det A$.
- (2) Sea $c \in \mathbb{K}$, $r \neq s$ y B la matriz que se obtiene de A sumando a la fila r la fila s multiplicada por c , es decir $A \xrightarrow{F_r + cF_s} B$, entonces $\det B = \det A$.
- (3) Sea $r \neq s$ y sea B la matriz que se obtiene de A permutando la fila r con la fila s , es decir $A \xrightarrow{F_r \leftrightarrow F_s} B$, entonces $\det B = -\det A$.

Este resultado nos permite calcular el determinante de matrices elementales.

COROLARIO 1.6.5. Sea $n \in \mathbb{N}$ y $c \in \mathbb{K}$. Sean $1 \leq r, s \leq n$, con $r \neq s$.

- (1) Si $c \neq 0$, la matriz elemental que se obtiene de multiplicar por c la fila r de I_n , tiene determinante igual a c .
- (2) Sea $r \neq s$. La matriz elemental que se obtiene de sumar a la fila r de I_n la fila s multiplicada por c , tiene determinante 1.
- (3) Finalmente, si $r \neq s$, la matriz elemental que se obtiene de intercambiar la fila r por la fila s de I_n tiene determinante -1 .

DEMOSTRACIÓN. Se deduce fácilmente del teorema anterior y del hecho de que $\det I_n = 1$. \square

COROLARIO 1.6.6. Sea $A \in M_n(\mathbb{K})$.

- (1) Si A tiene dos filas iguales, entonces $\det A = 0$.
- (2) Si A tiene una fila nula, entonces $\det A = 0$.

DEMOSTRACIÓN. (1) Sea A matriz donde $F_r = F_s$ con $r \neq s$. Luego, intercambiando la fila r por la fila s obtenemos la misma matriz. Es decir $A \xrightarrow{F_r \leftrightarrow F_s} A$. Por el teorema 1.6.4 (3), tenemos entonces que $\det A = -\det A$, por lo tanto $\det A = 0$.

(2) Sea F_r una fila nula de A , por lo tanto multiplicar por 2 esa fila no cambia la matriz. Es decir $A \xrightarrow{2F_r} A$. Por el teorema 1.6.4 (1), tenemos entonces que $\det A = 2\det A$, por lo tanto $\det A = 0$. \square

TEOREMA 1.6.7. Sean $A, B \in M_n(\mathbb{K})$, entonces

- (1) $\det(AB) = \det(A)\det(B)$.
- (2) A inversible si y solo si $\det(A) \neq 0$.

COROLARIO 1.6.8. $\det(AB) = \det(BA)$.

DEMOSTRACIÓN. $\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA)$. \square

DEFINICIÓN 1.6.3. Sea A una matriz $m \times n$ con coeficientes en \mathbb{K} . La *transpuesta* de A , denotada A^t , es la matriz $n \times m$ que en la fila i y columna j tiene el coeficiente $[A]_{ji}$. Es decir

$$[A^t]_{ij} = [A]_{ji}.$$

EJEMPLO 1.6.4. Si

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

entonces

$$A^t = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}.$$

EJEMPLO 1.6.5. Si

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix},$$

entonces

$$A^t = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

En general A^t es la matriz cuyas filas son las columnas de A .

PROPOSICIÓN 1.6.9. Sea A matriz $m \times n$.

(1) $(A^t)^t = A$.

(2) Si B matriz $n \times k$, entonces

$$(AB)^t = B^t A^t.$$

(3) Sea A matriz $n \times n$, entonces, A inversible si y sólo si A^t es inversible y en ese caso $(A^t)^{-1} = (A^{-1})^t$.

DEMOSTRACIÓN. (1) $[(A^t)^t]_{ij} = [A^t]_{ji} = [A]_{ij}$.

(2) Por definición de transpuesta $(AB)^t$ es una matriz $k \times m$. Ahora observemos que B^t es una matriz $k \times n$ y A^t es $n \times m$, luego tiene sentido multiplicar B^t por A^t y se obtiene también una matriz $k \times m$. La demostración de la proposición se hace comprobando que el coeficiente ij de $(AB)^t$ es igual al coeficiente ij de $B^t A^t$ y se deja como ejercicio para el lector.

(3)

$$A \text{ inversible} \Leftrightarrow \text{existe } B \text{ matriz } n \times n \text{ tal que } AB = I_n = BA$$

$$\Leftrightarrow (AB)^t = I_n^t = (BA)^t$$

$$\Leftrightarrow B^t A^t = I_n = A^t B^t$$

$$\Leftrightarrow B^t \text{ es la inversa de } A^t.$$

Es decir, A inversible si y sólo si A^t es inversible y si $B = A^{-1}$, entonces $(A^t)^{-1} = B^t$. \square

Observar que por inducción no es complicado probar que si A_1, \dots, A_k son matrices, entonces

$$(A_1 \dots A_k)^t = A_k^t \dots A_1^t.$$

EJEMPLO 1.6.6. Veamos las transpuestas de las matrices elementales 2×2 .

(1) Si $c \neq 0$, multiplicar por c la primera fila y multiplicar c por la segunda fila son, respectivamente,

$$E = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad E = \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix},$$

por lo tanto E^t es la misma matriz en ambos casos.

- (2) si $c \in \mathbb{K}$, sumar a la fila 2 la fila 1 multiplicada por c o sumar a la fila 1 la fila 2 multiplicada por c son, respectivamente,

$$E_1 = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \quad \text{y} \quad E_2 = \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix},$$

por lo tanto

$$E_1^t = \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} = E_2 \quad \text{y} \quad E_2^t = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} = E_1,$$

- (3) Finalmente, intercambiando la fila 1 por la fila 2 obtenemos la matriz

$$E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

por lo tanto $E^t = E$

OBSERVACIÓN 1.6.3. En el caso de matrices 2×2 podemos comprobar fácilmente que $\det A^t = \det A$:

$$\det A^t = \det \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{21}a_{12} = \det A.$$

También vale este resultado para matrices $n \times n$.

TEOREMA 1.6.10. *Sea $A \in M_n(\mathbb{K})$, entonces $\det(A) = \det(A^t)$*

El resultado anterior permite obtener resultados nuevos del cálculo de determinante a partir de resultados vistos anteriormente.

PROPOSICIÓN 1.6.11. *Sea $A \in M_n(\mathbb{K})$ matriz triangular inferior cuyos elementos en la diagonal son d_1, \dots, d_n . Entonces $\det A = d_1 \cdot d_2 \cdot \dots \cdot d_n$.*

DEMOSTRACIÓN. Si A es triangular inferior con elementos en la diagonal d_1, \dots, d_n , entonces A^t es triangular superior con elementos en la diagonal d_1, \dots, d_n . Por la proposición 1.6.1, $\det A^t = d_1 \cdot \dots \cdot d_n$. Por el teorema 1.6.10 obtenemos el resultado. \square

TEOREMA 1.6.12. *Sea $A \in M_n(\mathbb{K})$ y sean $1 \leq r, s \leq n$.*

- (1) *Sea $c \in \mathbb{K}$ y B la matriz que se obtiene de A multiplicando la columna r por c , entonces $\det B = c \det A$.*
- (2) *Sea $c \in \mathbb{K}$ y B la matriz que se obtiene de A sumando a la columna r la columna s multiplicada por c , entonces $\det B = \det A$.*
- (3) *Sea B la matriz que se obtiene de A permutando la columna r con la fila s , entonces $\det B = -\det A$.*

DEMOSTRACIÓN. Las operaciones por columna del enunciado se traducen a operaciones por fila de la matriz A^t . Luego, aplicando los resultados del teorema 1.6.4 y usando el hecho de que $\det(A) = \det(A^t)$ y $\det(B) = \det(B^t)$ en cada caso, se deduce el corolario. \square

COROLARIO 1.6.13. Sea $A \in M_n(\mathbb{K})$.

- (1) Si A tiene dos columnas iguales, entonces $\det A = 0$.
- (2) Si A tiene una columna nula, entonces $\det A = 0$.

DEMOSTRACIÓN. (1) Si A tiene dos columnas iguales, entonces A^t tiene dos filas iguales, luego, por corolario 1.6.6 (1), $\det A^t = 0$ y por lo tanto $\det A = 0$.

(2) Si A tiene una columna nula, entonces A^t tiene una fila nula, luego, 1.6.6 (2), $\det A^t = 0$ y por lo tanto $\det A = 0$. \square

El siguiente teorema nos dice que es posible calcular el determinante desarrollándolo por cualquier fila o cualquier columna.

TEOREMA 1.6.14. El determinante de una matriz A de orden $n \times n$ puede ser calculado por la expansión de los cofactores en cualquier columna o cualquier fila. Más específicamente,

- (1) si usamos la expansión por la j -ésima columna, $1 \leq j \leq n$, tenemos

$$\begin{aligned} \det A &= \sum_{i=1}^n a_{ij} C_{ij} \\ &= a_{1j} C_{1j} + a_{2j} C_{2j} + \cdots + a_{nj} C_{nj}. \end{aligned}$$

- (2) si usamos la expansión por la i -ésima fila, $1 \leq i \leq n$, tenemos

$$\begin{aligned} \det A &= \sum_{j=1}^n a_{ij} C_{ij} \\ &= a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in}; \end{aligned}$$

aa

Espacios vectoriales

En este capítulo estudiaremos en forma general las combinaciones lineales sobre conjuntos abstractos. En el primer capítulo desarrollamos el método de Gauss para la resolución de sistemas de ecuaciones lineales. En el método de Gauss se usan sistemáticamente las combinaciones lineales de las filas de una matriz. Podemos ver estas filas como elementos de \mathbb{K}^n y nuestro primer impulso para el estudio de las combinaciones lineales sería trabajar en este contexto, es decir en \mathbb{K}^n . Eso es lo que haremos, en forma introductoria, en la primera sección, donde trabajaremos restringidos a \mathbb{R}^2 y \mathbb{R}^3 . Sin embargo, muchos de los resultados sobre combinaciones lineales en \mathbb{K}^n son aplicables también a conjuntos más generales y de gran utilidad en la matemática. Por lo tanto, apartir de la segunda sección, nuestros “espacios vectoriales” (espacios donde pueden hacerse combinaciones lineales de vectores) serán espacios abstractos.

2.1. Definición y ejemplos de espacios vectoriales

DEFINICIÓN 2.1.1. Sea \mathbb{K} cuerpo. Un *espacio vectorial* sobre \mathbb{K} o un \mathbb{K} -*espacio vectorial*, consiste de un conjunto V no vacío, cuyos elementos son llamados *vectores*, junto a ‘+’ y ‘.’ tal que

- (a) ‘+’: $V \times V \rightarrow V$ es una operación, llamada *adición* o *suma de vectores*, tal que a dos vectores $v, w \in V$ les asigna otro vector $v + w \in V$,
- (b) ‘.’: $\mathbb{K} \times V \rightarrow V$ es una operación tal que a $\lambda \in \mathbb{K}$ y $v \in V$ le asigna el vector $\lambda.v$ (o simplemente λv). ‘.’ es llamada el *producto por escalares*.

Además, estas operaciones deben satisfacer

- S1.** $v + w = w + v$, para $v, w \in V$ (*conmutatividad de la suma*),
- S2.** $(v + w) + u = v + (w + u)$, para $v, w, u \in V$ (*asociatividad de la suma*),
- S3.** existe un único vector 0 , llamado *vector cero*, tal que $0 + v = v + 0 = v$, para todo $v \in V$ (*existencia de elemento neutro de la suma*).
- S4.** Para cada $v \in V$, existe un único vector $-v$ tal que $v + (-v) = (-v) + v = 0$ (*existencia de opuesto o inverso aditivo*).
- P1.** $1.v = v$ para todo $v \in V$.
- P2.** $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v$, para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $v \in V$.
- D1.** $\lambda(v + w) = \lambda v + \lambda w$, para todo $\lambda \in \mathbb{K}$ y todo $v, w \in V$ (*propiedad distributiva*).
- D2.** $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $v \in V$ (*propiedad distributiva*).

Debido a la ley de asociatividad para la suma $(v + w) + u$ es igual a $v + (w + u)$ y por lo tanto podemos eliminar los paréntesis sin ambigüedad. Es decir, $\forall v, w, u \in V$ denotamos

$$v + w + u := (v + w) + u = v + (w + u).$$

De forma análoga, $\forall \lambda_1, \lambda_2 \in V, \forall v \in V$ usaremos la notación

$$\lambda_1 \lambda_2 v = (\lambda_1 \lambda_2) v = \lambda_1 (\lambda_2 v).$$

Otra notación importante, e intuitiva, es la siguiente $\forall v, w \in V$

$$v - w := v + (-w),$$

y a menudo diremos que $v - w$ es la *resta* de v menos w .

EJEMPLO 2.1.1. \mathbb{K}^n . Sea \mathbb{K} cuerpo, y sea

$$V = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{K}, 1 \leq i \leq n\} = \mathbb{K}^n.$$

Entonces V es espacio vectorial con las siguientes operaciones: si $(x_1, x_2, \dots, x_n) \in \mathbb{K}^n$, $(y_1, y_2, \dots, y_n) \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$

- (a) $(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$,
- (b) $\lambda(x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$.

Observar que las sumas y productos son coordenada a coordenada y, por lo tanto, en cada coordenada son sumas y productos en \mathbb{K} .

Comprobemos las propiedades necesarias para que V sea un espacio vectorial. Como la suma de vectores y el producto por escalares es coordenada a coordenada, las propiedades se deducirán fácilmente de los axiomas para la suma y el producto en los cuerpos. Sean $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, $z = (z_1, \dots, z_n)$ en V y $\lambda, \lambda_1, \lambda_2 \in \mathbb{K}$:

- S1. $x + y = y + x$, pues $x_i + y_i = y_i + x_i$, $1 \leq i \leq n$.
- S2. $(x + y) + z = x + (y + z)$, pues $(x_i + y_i) + z_i = x_i + (y_i + z_i)$, $1 \leq i \leq n$.
- S3. Sea $0 = (0, \dots, 0)$, entonces $0 + x = (0 + x_1, \dots, 0 + x_n) = (x_1, \dots, x_n) = x$.
- S4. Sea $-x = (-x_1, \dots, -x_n)$, entonces $x + (-x) = (x_1 - x_1, \dots, x_n - x_n) = (0, \dots, 0)$.
- P1. $1.x = (1.x_1, \dots, 1.x_n) = (x_1, \dots, x_n) = x$.
- P2. $\lambda_1(\lambda_2 x) = (\lambda_1 \lambda_2)x$ pues $\lambda_1(\lambda_2 x_i) = (\lambda_1 \lambda_2)x_i$, $1 \leq i \leq n$.
- D1. $\lambda(x + y) = \lambda x + \lambda y$, pues $\lambda(x_i + y_i) = \lambda x_i + \lambda y_i$, $1 \leq i \leq n$.
- D2. $(\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$, pues $(\lambda_1 + \lambda_2)x_i = \lambda_1 x_i + \lambda_2 x_i$, $1 \leq i \leq n$.

EJEMPLO 2.1.2. MATRICES $m \times n$. Sea \mathbb{K} cuerpo, definimos en $M_{m \times n}(\mathbb{K})$ la suma y el producto por escalares de la siguiente forma. Sean $A = [a_{ij}]$, $B = [b_{ij}]$ matrices $m \times n$ y $\lambda \in \mathbb{K}$,

entonces $A + B$, λA son matrices en $M_{m \times n}(\mathbb{K})$ con coeficientes:

$$\begin{aligned}[A + B]_{ij} &= [a_{ij} + b_{ij}], \\ [\lambda A]_{ij} &= [\lambda a_{ij}].\end{aligned}$$

Es decir, la suma es coordenada a coordenada y el producto es multiplicar el escalar en cada coordenada. Este caso no es más que \mathbb{K}^{mn} presentado de otra manera.

Ejemplifiquemos, con casos sencillos, la suma de matrices y el producto por escalares

$$\begin{bmatrix} -2 & 1 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 1 \\ 2 & -5 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & -1 \end{bmatrix}, \quad 3 \begin{bmatrix} -2 & 1 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} -6 & 3 \\ 0 & 12 \end{bmatrix}.$$

EJEMPLO 2.1.3. POLINOMIOS. Sea

$$\mathbb{K}[x] = \{a_n x^n + \cdots + a_1 x + a_0 : n \in \mathbb{N}_0, a_i \in \mathbb{K}, \text{ para } 0 \leq i \leq n\}$$

el conjunto de polinomios sobre \mathbb{K} . Entonces si $P(x), Q(x) \in \mathbb{K}[x]$, definimos la suma de polinomios de la siguiente manera: sea $P(x) = a_n x^n + \cdots + a_1 x + a_0$ y $Q(x) = b_n x^n + \cdots + b_1 x + a_0$ (completamos coeficientes con 0 hasta que ambos tengan el mismo n), entonces

$$(P + Q)(x) = (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0).$$

Si $\lambda \in \mathbb{K}$,

$$(\lambda P)(x) = \lambda a_n x^n + \cdots + \lambda a_1 x + \lambda a_0.$$

Por ejemplo,

$$\begin{aligned}(3x^2 + 1) + (x^4 + 2x^3 + 5x^2 - x) &= x^4 + 2x^3 + 8x^2 - x + 1, & \text{y} \\ 3(x^4 + 2x^3 + 5x^2 - x) &= 3x^4 + 6x^3 + 15x^2 - 3x.\end{aligned}$$

EJEMPLO 2.1.4. ESPACIOS DE FUNCIONES. Sean

$$F(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} : \text{tal que } f \text{ es una función}\},$$

$$C(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} : \text{tal que } f \text{ es una función continua}\}.$$

Recordemos que si f, g son funciones, entonces la función suma de f y g está definida por

$$(f + g)(x) = f(x) + g(x).$$

Por otro lado, si $\lambda \in \mathbb{R}$, la función multiplicar f por λ está definida por

$$(\lambda f)(x) = \lambda f(x).$$

Es sencillo ver que con estas dos operaciones, $F(\mathbb{R})$ es un \mathbb{R} -espacio vectorial.

Con respecto a $C(\mathbb{R})$, hemos visto en el primer curso de análisis matemático que la suma de funciones continuas es una función continua y, por lo tanto, $f + g$ es continua si f y g lo son.

El producto de un escalar c por una función continua f , puede ser visto como el producto de una función que es constante y vale λ (y es continua) y la función f . Por lo tanto, λf es producto de funciones continuas y, en consecuencia, es una función continua. Resumiendo,

$$f, g \in C(\mathbb{R}) \Rightarrow f + g \in C(\mathbb{R}), \quad \lambda \in \mathbb{R}, f \in C(\mathbb{R}) \Rightarrow \lambda f \in C(\mathbb{R}).$$

No es difícil ver que con estas definiciones $C(\mathbb{R})$ es un \mathbb{R} -espacio vectorial.

EJEMPLO 2.1.5. Consideremos el conjunto de los números reales positivos:

$$R_{>0} = \{x \in \mathbb{R} : x > 0\}.$$

Entonces $V = R_{>0}$ es un \mathbb{R} -espacio vectorial con la suma $\oplus : V \times V \rightarrow V$ y el producto $\odot : \mathbb{R} \times V \rightarrow V$ dados por

$$x \oplus y = x \cdot y, \quad c \odot x = x^c,$$

para cada $c \in \mathbb{R}$, $x, y \in R_{>0}$. Es fácil ver que los axiomas **S1.** y **S2.** sobre la conmutatividad y asociatividad, respectivamente, de la suma \oplus se siguen de las propiedades de conmutatividad y asociatividad del producto \cdot en \mathbb{R} .

La existencia del vector **0** del axioma **S3.**, neutro para la suma \oplus , requiere de cierto cuidado. Notar que este vector debe ser un elemento **0** en V (un real positivo) que cumpla $x \oplus \mathbf{0} = \mathbf{x}$ para todo x . Ahora, $x \oplus \mathbf{0} = \mathbf{x} \cdot \mathbf{0}$ por definición, de donde se desprende que debemos tomar $\mathbf{0} = 1$. Es decir, *el vector cero es el número 1*. De manera similar, se sigue que el opuesto indicado en el axioma **S4.** debe estar dado por $-x = x^{-1}$.

Finalmente, las propiedades de los axiomas **P1.**, **P2.**, **D1.** y **D2.** se siguen de las propiedades conocidas de la exponenciación en \mathbb{R} y quedan a cargo del lector, como un interesante desafío para terminar de comprender este ejemplo.

PROPOSICIÓN 2.1.1. Sea V un espacio vectorial sobre el cuerpo \mathbb{K} . Entonces,

- (1) $\lambda \cdot \mathbf{0} = \mathbf{0}$, para todo $\lambda \in \mathbb{K}$;
- (2) $\mathbf{0} \cdot v = \mathbf{0}$, para todo $v \in V$;
- (3) si $\lambda \in \mathbb{K}$, $v \in V$, $v \neq \mathbf{0}$ y $\lambda \cdot v = \mathbf{0}$, entonces $\lambda = \mathbf{0}$;
- (4) $(-1) \cdot v = -v$, para todo $v \in V$.

DEMOSTRACIÓN. Tanto la prueba de (1), como la de (2) son similares a la demostración de que $\mathbf{0} \cdot a = \mathbf{0}$ en \mathbb{Z} (o en \mathbb{R}).

(1) Como 0 es el elemento neutro de la suma en V , entonces $0 = 0 + 0$, luego

$$\begin{aligned}\lambda.0 &= \lambda.(0 + 0) && (\text{propiedad distributiva} \Rightarrow) \\ \lambda.0 &= \lambda.0 + \lambda.0 && (\text{sumando a la izquierda } -\lambda.0 \Rightarrow) \\ \lambda.0 - \lambda.0 &= \lambda.0 + \lambda.0 - \lambda.0 && (\text{opuesto} \Rightarrow) \\ 0 &= \lambda.0 + 0 && (\text{elemento neutro} \Rightarrow) \\ 0 &= \lambda.0.\end{aligned}$$

(2) Análoga a (1).

(3) Supongamos que $\lambda.v = 0$ y $\lambda \neq 0$, entonces, por (1), $\lambda^{-1}(\lambda.v) = 0$, pero $\lambda^{-1}(\lambda.v) = (\lambda^{-1}\lambda).v = 1.v = v$. Luego $0 = v$, que contradice la hipótesis. El absurdo vino de suponer que $\lambda \neq 0$.

(4) $(-1).v + v = (-1).v + 1.v = (-1 + 1).v$, esto último es por la propiedad distributiva. Ahora bien $(-1 + 1).v = 0.v \stackrel{(2)}{=} 0$. Es decir $(-1).v + v = 0$ y por lo tanto $(-1).v$ es el opuesto de v (que es $-v$).

□

DEFINICIÓN 2.1.2. Sea V espacio vectorial sobre \mathbb{K} y v_1, \dots, v_n vectores en V . Dado $v \in V$, diremos que v es combinación lineal de los v_1, \dots, v_n si existen escalares $\lambda_1, \dots, \lambda_n$ en \mathbb{K} , tal que

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

EJEMPLO 2.1.6. (1) Sean $v_1 = (1, 0)$, $v_2 = (0, 1)$ en \mathbb{C}^2 ¿es $v = (i, 2)$ combinación lineal de v_1, v_2 ? La respuesta es sí, pues

$$v = iv_1 + 2v_2.$$

Observar además que es la única combinación lineal posible, pues si

$$v = \lambda_1 v_1 + \lambda_2 v_2,$$

entonces

$$(i, 2) = (\lambda_1, 0) + (0, \lambda_2) = (\lambda_1, \lambda_2),$$

luego $\lambda_1 = i$ y $\lambda_2 = 2$.

Puede ocurrir que un vector sea combinación lineal de otros vectores de varias formas diferentes. Por ejemplo, si $v = (i, 2)$ y $v_1 = (1, 0)$, $v_2 = (0, 1)$, $v_3 = (1, 1)$, tenemos que

$$v = iv_1 + 2v_2 + 0v_3, \quad \text{y también}$$

$$v = (i - 1)v_1 + v_2 + v_3.$$

(2) Sean $(0, 1, 0)$, $(0, 1, 1)$ en \mathbb{C}^3 ¿es $(1, 1, 0)$ combinación lineal de $(0, 1, 0)$, $(0, 1, 1)$? La respuesta es no, pues si

$$(1, 1, 0) = \lambda_1(0, 1, 0) + \lambda_2(0, 1, 1) = (0, \lambda_1, 0) + (0, \lambda_2, \lambda_2) = (0, \lambda_1 + \lambda_2, \lambda_2),$$

luego, la primera coordenada nos dice que $1 = 0$, lo cual es absurdo. Por lo tanto, no existe un par $\lambda_1, \lambda_2 \in \mathbb{K}$ tal que $(1, 1, 0) = \lambda_1(0, 1, 0) + (0, 1, 1)$.

OBSERVACIÓN 2.1.1. La pregunta de si un vector $v = (b_1, \dots, b_m) \in \mathbb{K}^m$ es combinación lineal de vectores $v_1, \dots, v_n \in \mathbb{K}^m$ se resuelve con un sistema de ecuaciones lineales: si

$$v_i = (a_{1i}, \dots, a_{mi}), \quad \text{para } 1 \leq i \leq n,$$

entonces $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ se traduce, en coordenadas, a

$$\begin{aligned} (b_1, \dots, b_m) &= \lambda_1(a_{11}, \dots, a_{m1}) + \dots + \lambda_n(a_{1n}, \dots, a_{mn}) \\ &= (\lambda_1 a_{11} + \dots + \lambda_n a_{1n}, \dots, \lambda_1 a_{m1} + \dots + \lambda_n a_{mn}). \end{aligned}$$

Luego, v es combinación lineal de los vectores $v_1, \dots, v_n \in \mathbb{K}^n$ si y sólo si el tiene solución el siguiente sistema de ecuaciones con incógnitas $\lambda_1, \dots, \lambda_n$:

$$\begin{array}{ccccccc} a_{11}\lambda_1 & + & a_{12}\lambda_2 & + & \dots & + & a_{1n}\lambda_n & = & b_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}\lambda_1 & + & a_{m2}\lambda_2 & + & \dots & + & a_{mn}\lambda_n & = & b_m \end{array}$$

EJEMPLO 2.1.7. Demostrar que $(5, 12, 5)$ es combinación lineal de los vectores $(1, -5, 2), (0, 1, -1), (1, 2, -1)$. Planteamos la ecuación:

$$\begin{aligned} (5, 12, 5) &= \lambda_1(1, -5, 2) + \lambda_2(0, 1, -1) + \lambda_3(1, 2, -1) \\ &= (\lambda_1 + \lambda_3, -5\lambda_1 + \lambda_2 + 2\lambda_3, 2\lambda_1 - \lambda_2 - \lambda_3). \end{aligned}$$

Por consiguiente, esta ecuación se resuelve con el siguiente sistema de ecuaciones

$$\begin{aligned} \lambda_1 + \lambda_3 &= 5 \\ -5\lambda_1 + \lambda_2 + 2\lambda_3 &= 12 \\ 2\lambda_1 - \lambda_2 - \lambda_3 &= 5. \end{aligned}$$

Ahora bien, usando el método de Gauss

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ -5 & 1 & 2 & 12 \\ 2 & -1 & -1 & 5 \end{array} \right] &\xrightarrow[F_3 - 2F_1]{F_2 + 5F_1} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & 7 & 37 \\ 0 & -1 & -3 & -5 \end{array} \right] &\xrightarrow{F_3 + F_1} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & 7 & 37 \\ 0 & 0 & 4 & 32 \end{array} \right] \\ &\xrightarrow{F_3/4} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & 7 & 37 \\ 0 & 0 & 1 & 8 \end{array} \right] &\xrightarrow[F_2 - 7F_3]{F_1 - F_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & -19 \\ 0 & 0 & 1 & 8 \end{array} \right]. \end{aligned}$$

Luego $\lambda_1 = -3$, $\lambda_2 = -19$ y $\lambda_3 = 8$, es decir

$$(5, 12, 5) = -3(1, -5, 2) - 19(0, 1, -1) + 8(1, 2, -1).$$

2.2. Subespacios vectoriales

DEFINICIÓN 2.2.1. Sea V un espacio vectorial sobre \mathbb{K} . diremos que $W \subset V$ es *subespacio* de V si $W \neq \emptyset$ y

- (a) si para cualesquiera $w_1, w_2 \in W$, se cumple que $w_1 + w_2 \in W$ y
- (b) si $\lambda \in \mathbb{K}$ y $w \in W$, entonces $\lambda w \in W$.

OBSERVACIÓN 2.2.1. Si W subespacio de V , entonces $0 \in W$: como $W \neq \emptyset$, tomo cualquier $w \in W$ y por (a) tenemos que $0.w \in W$. Ya vimos, proposición 2.1.1(2), que $0.w = 0$ y por lo tanto $0 \in W$.

OBSERVACIÓN 2.2.2. Si W subespacio de V y $w_1, \dots, w_k \in W$, entonces cualquier combinación lineal de los w_1, \dots, w_k pertenece a W . Es decir, para cualesquiera $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, se cumple que $\lambda_1 w_1 + \dots + \lambda_k w_k \in W$. Esto se debe a que $\lambda_i w_i \in W$ para $1 \leq i \leq k$. Por un argumento inductivo, no es difícil probar que la suma de k términos en W es un elemento de W , por lo tanto $\lambda_1 w_1 + \dots + \lambda_k w_k \in W$.

TEOREMA 2.2.1. Sea V un espacio vectorial sobre \mathbb{K} y W subespacio de V . Entonces W con las operaciones suma y producto por escalares de V es un espacio vectorial.

DEMOSTRACIÓN. Debemos verificar las propiedades de la definición de espacio vectorial.

- S1. $w, w' \in W$, entonces $w + w' = w' + w$, pues conmutan en V ;
- S2. si $w_1, w_2, w_3 \in W$, entonces $w_1 + (w_2 + w_3) = (w_1 + w_2) + w_3$, pues la suma es asociativa en V ;
- S3. ya vimos que $0 \in W$, y como 0 es el elemento neutro de la suma para V , también lo es para W .
- S4. Sea $w \in W$, entonces $(-1)w \in W$ (por S2). Pero, por proposición 2.1.1(4), $(-1)w = -w$, es decir $(-1)w$ es el opuesto aditivo de w en V , luego es el opuesto aditivo de w en W .
- P1. $1.v = v$ para todo $v \in V$, luego $1.w = w$ para todo $w \in W$.
- P2. $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v$, para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $v \in V$, luego $\lambda_1(\lambda_2 w) = (\lambda_1 \lambda_2)w$, para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $w \in W$.
- D1. $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$, para todo $\lambda \in \mathbb{K}$ y cualesquiera $v_1, v_2 \in V$, luego $\lambda(w_1 + w_2) = \lambda w_1 + \lambda w_2$, para todo $\lambda \in \mathbb{K}$ y cualesquiera $w_1, w_2 \in W$.
- D2. $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $v \in V$, luego $(\lambda_1 + \lambda_2)w = \lambda_1 w + \lambda_2 w$ para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $w \in W$.

□

EJEMPLO 2.2.1. Veremos ahora una serie de ejemplos de subespacios vectoriales.

- (1) Sea V un \mathbb{K} -espacio vectorial, entonces 0 y V son subespacios vectoriales de V . Suelen ser llamados los *subespacios triviales* de V .

- (2) Sea V un \mathbb{K} -espacio vectorial y sea $v \in V$, entonces

$$W = \{\lambda v : \lambda \in \mathbb{K}\}$$

es un subespacio vectorial. En efecto

- (a) si $\lambda_1 v, \lambda_2 v \in W$, con $\lambda_1, \lambda_2 \in \mathbb{K}$, entonces $\lambda_1 v + \lambda_2 v = (\lambda_1 + \lambda_2)v \in W$;
- (b) $\lambda_1 v \in W$, con $\lambda_1 \in \mathbb{K}$ y $\lambda \in \mathbb{K}$, entonces $\lambda(\lambda_1 v) = (\lambda\lambda_1)v \in W$.

El subespacio W suele ser denotado $\mathbb{K}v$.

- (3) Sean $V = \mathbb{K}^n$ y $1 \leq j \leq n$. Definimos

$$W = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{K} (1 \leq i \leq n), x_j = 0\}.$$

Es decir W es el subconjunto de V de todas las n -uplas con la coordenada j igual a 0.

Por ejemplo si $j = 1$

$$W = \{(0, x_2, \dots, x_n) : x_i \in \mathbb{K} (2 \leq i \leq n)\}.$$

Veamos que este último es un subespacio:

- (a) si $(0, x_2, \dots, x_n), (0, y_2, \dots, y_n) \in W$, entonces $(0, x_2, \dots, x_n) + (0, y_2, \dots, y_n) = (0, x_2 + y_2, \dots, x_n + y_n)$, el cual pertenece a W .
- (b) Por otro lado, si $\lambda \in \mathbb{K}$, $\lambda(0, x_2, \dots, x_n) = (0, \lambda x_2, \dots, \lambda x_n) \in W$.

La demostración para $j > 1$ es completamente análoga.

- (4) El conjunto $\mathbb{R}[x] = \{P(x) : P(x) \text{ es polinomio en } \mathbb{R}\}$, es subespacio de $F(\mathbb{R})$, pues $\mathbb{R}[x] \subset F(\mathbb{R})$ y las operaciones de suma y producto por un escalar son cerradas en $\mathbb{R}[x]$.
- (5) De forma análoga, el conjunto $\mathbb{R}[x]$ es subespacio de $C(\mathbb{R})$, el espacio de funciones continuas de \mathbb{R} .
- (6) Sea $W = \{A \in M_n(\mathbb{K}) : A^t = A\}$. Es claro que $A \in W$ si y sólo si $[A]_{ij} = [A]_{ji}$. Veamos que W es subespacio de $M_n(\mathbb{K})$:
- (a) sean $A = [a_{ij}]$, $B = [b_{ij}]$ tales que $A = A^t$ y $B = B^t$, entonces debemos verificar que $A + B \in W$, es decir que la transpuesta de $A + B$ es la misma matriz: ahora bien, $[A + B]_{ij} = a_{ij} + b_{ij}$, luego

$$[(A + B)^t]_{ij} = a_{ji} + b_{ji} = [A]_{ji} + [B]_{ji} = [A]_{ij} + [B]_{ij} = [A + B]_{ij},$$

por lo tanto $A + B \in W$.

- (b) Si $\lambda \in \mathbb{K}$, $[\lambda A]_{ij} = \lambda a_{ij}$, luego,

$$[\lambda A^t]_{ij} = \lambda a_{ji} = \lambda a_{ij} = [\lambda A]_{ij},$$

por lo tanto $\lambda A \in W$.

- (7) Sea $A \in M_{m \times n}(\mathbb{K})$. Si $x = (x_1, \dots, x_n) \in \mathbb{K}^n$, entonces Ax denotará la multiplicación de A por la matriz columna formada por x_1, \dots, x_n , es decir

$$Ax = A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Sea

$$W = \{x \in \mathbb{K}^n : Ax = 0\}.$$

Es decir, W es el subconjunto de \mathbb{K}^n de las soluciones del sistema $Ax = 0$. Entonces, W es un subespacio de \mathbb{K}^n :

- (a) si $x, y \in W$, es decir si $Ax = 0$ y $Ay = 0$, entonces $A(x+y) = Ax + Ay = 0 + 0 = 0$, luego $x + y \in W$;
- (b) si $\lambda \in \mathbb{K}$ y $x \in W$, entonces $A(\lambda x) = \lambda Ax = \lambda \cdot 0 = 0$, luego $\lambda x \in W$.

TEOREMA 2.2.2. Sea V un espacio vectorial sobre \mathbb{K} y sean $v_1, \dots, v_k \in V$. Entonces

$$W = \{\lambda_1 v_1 + \dots + \lambda_k v_k : \lambda_1, \dots, \lambda_k \in \mathbb{K}\}$$

es un subespacio vectorial. Es decir, el conjunto de las combinaciones lineales de v_1, \dots, v_k es un subespacio vectorial.

DEMOSTRACIÓN. (a) Sean $\lambda_1 v_1 + \dots + \lambda_k v_k$ y $\mu_1 v_1 + \dots + \mu_k v_k$ dos combinaciones lineales de v_1, \dots, v_k , entonces

$$\begin{aligned} (\lambda_1 v_1 + \dots + \lambda_k v_k) + (\mu_1 v_1 + \dots + \mu_k v_k) &= \lambda_1 v_1 + \mu_1 v_1 + \dots + \lambda_k v_k + \mu_k v_k \\ &= (\lambda_1 + \mu_1) v_1 + \dots + (\lambda_k + \mu_k) v_k, \end{aligned}$$

que es una combinación lineal de v_1, \dots, v_k y por lo tanto pertenece a W .

- (b) Si $\lambda \in \mathbb{K}$ y $\lambda_1 v_1 + \dots + \lambda_k v_k$ es una combinación lineal de v_1, \dots, v_k , entonces

$$\begin{aligned} \lambda(\lambda_1 v_1 + \dots + \lambda_k v_k) &= \lambda(\lambda_1 v_1) + \dots + \lambda(\lambda_k v_k) \\ &= (\lambda \lambda_1) v_1 + \dots + (\lambda \lambda_k) v_k, \end{aligned}$$

que es una combinación lineal de v_1, \dots, v_k y por lo tanto pertenece a W .

□

DEFINICIÓN 2.2.2. Sea V un espacio vectorial sobre \mathbb{K} y sean $v_1, \dots, v_k \in V$. Al subespacio vectorial $W = \{\lambda_1 v_1 + \dots + \lambda_k v_k : \lambda_1, \dots, \lambda_k \in \mathbb{K}\}$ de las combinaciones lineales de v_1, \dots, v_k se lo denomina *subespacio generado por v_1, \dots, v_k* y se lo denota

$$\begin{aligned} W &= \langle v_1, \dots, v_k \rangle, & \text{o} \\ W &= \text{gen} \{v_1, \dots, v_k\} & \text{o} \\ W &= \text{span} \{v_1, \dots, v_k\}. \end{aligned}$$

Además, en este caso, diremos que el conjunto $S = \{v_1, \dots, v_k\}$ genera al subespacio W .

TEOREMA 2.2.3. *Sea V un espacio vectorial sobre \mathbb{K} . Entonces la intersección de subespacios vectoriales es un subespacio vectorial.*

DEMOSTRACIÓN. Sea $\{W_i\}_{i \in I}$ una familia de subespacios vectoriales y sea

$$W = \bigcap_{i \in I} W_i,$$

entonces

- (a) si $w_1, w_2 \in W$, tenemos que $w_1, w_2 \in W_i$ para todo $i \in I$, luego, como W_i es subespacio vectorial, $w_1 + w_2 \in W_i$ para todo $i \in I$, por lo tanto $w_1 + w_2 \in W$;
- (b) si $\lambda \in \mathbb{K}$ y $w \in W$, $w \in W_i$ para todo $i \in I$ y, por lo tanto, $cw \in W_i$ para todo $i \in I$. En consecuencia $w \in W$.

□

OBSERVACIÓN 2.2.1. Si V es un \mathbb{K} -espacio vectorial, S y T subespacios de V , entonces $S \cup T$ no es necesariamente un subespacio de V . En efecto, consideremos en \mathbb{R}^2 los subespacios $S = \mathbb{R}(1, 0)$ y $T = \mathbb{R}(0, 1)$. Observamos que $(1, 0) \in S$ y $(0, 1) \in T$; luego, ambos pertenecen a $S \cup T$. Pero $(1, 0) + (0, 1) = (1, 1) \notin S \cup T$, puesto que $(1, 1) \notin S$ y $(1, 1) \notin T$.

TEOREMA 2.2.4. *Sea V un espacio vectorial sobre \mathbb{K} y sean $v_1, \dots, v_k \in V$. Entonces, la intersección de todos los subespacios vectoriales que contienen a v_1, \dots, v_k es igual a $\langle v_1, \dots, v_k \rangle$.*

DEMOSTRACIÓN. Denotemos $W_1 = \langle v_1, \dots, v_k \rangle$ y W_2 la intersección de todos los subespacios vectoriales que contienen a v_1, \dots, v_k . Probaremos que $W_1 = W_2$ con la doble inclusión, es decir probando que $W_1 \subseteq W_2$ y $W_2 \subseteq W_1$.

$(W_1 \subseteq W_2)$. Sea W subespacio vectorial que contiene v_1, \dots, v_k . Como W es subespacio, entonces W contiene a cualquier combinación lineal de los v_1, \dots, v_k , por lo tanto W contiene a W_1 . Es decir, cualquier subespacio que contiene a v_1, \dots, v_k , también contiene a W_1 , por lo tanto la intersección de todos los subespacios que contienen a v_1, \dots, v_k , contiene a W_1 . Luego $W_2 \supseteq W_1$.

$(W_2 \subseteq W_1)$. W_1 es un subespacio que contiene a v_1, \dots, v_k , por lo tanto la intersección de todos los subespacios que contienen a v_1, \dots, v_k está contenida en W_1 . Es decir, $W_2 \subseteq W_1$. □

DEFINICIÓN 2.2.3. Sea V un espacio vectorial sobre \mathbb{K} y sean S_1, \dots, S_k subconjuntos de V . definimos

$$S_1 + \dots + S_k := \{s_1 + \dots + s_k : s_i \in S_i, 1 \leq i \leq k\},$$

el conjunto *suma de los* S_1, \dots, S_k .

TEOREMA 2.2.5. Sea V un espacio vectorial sobre \mathbb{K} y sean W_1, \dots, W_k subespacios de V . Entonces $W = W_1 + \dots + W_k$ es un subespacio de V .

DEMOSTRACIÓN. Sean $v = v_1 + \dots + v_k$ y $w = w_1 + \dots + w_k$ en W y $\lambda \in \mathbb{K}$. Entonces

- (a) $v + w = (v_1 + w_1) + \dots + (v_k + w_k) \in W_1 + \dots + W_k$, pues como W_i es subespacio de V , tenemos que $v_i + w_i \in W_i$.
- (b) $\lambda v = \lambda(v_1 + \dots + v_k) = \lambda v_1 + \dots + \lambda v_k \in W_1 + \dots + W_k$, pues como W_i es subespacio de V , tenemos que $\lambda v_i \in W_i$.

□

PROPOSICIÓN 2.2.6. Sea V un espacio vectorial sobre \mathbb{K} y sean v_1, \dots, v_r elementos de V . Entonces

$$\langle v_1, \dots, v_r \rangle = \langle v_1 \rangle + \dots + \langle v_r \rangle.$$

DEMOSTRACIÓN. Probemos el resultado viendo que los dos conjuntos se incluyen mutuamente.

(\subseteq) Sea $w \in \langle v_1, \dots, v_r \rangle$, luego $w = \lambda_1 v_1 + \dots + \lambda_r v_r$. Como $\lambda_i v_i \in \langle v_i \rangle$, $1 \leq i \leq r$, tenemos que $w \in \langle v_1 \rangle + \dots + \langle v_r \rangle$. En consecuencia, $\langle v_1, \dots, v_r \rangle \subseteq \langle v_1 \rangle + \dots + \langle v_r \rangle$.

(\supseteq) Si $w \in \langle v_1 \rangle + \dots + \langle v_r \rangle$, entonces $w = w_1 + \dots + w_r$ con $w_i \in \langle v_i \rangle$ para todo i . Por lo tanto, $w_i = \lambda_i v_i$ para algún $\lambda_i \in \mathbb{K}$ y $w = \lambda_1 v_1 + \dots + \lambda_r v_r \in \langle v_1, \dots, v_r \rangle$. En consecuencia, $\langle v_1 \rangle + \dots + \langle v_r \rangle \subseteq \langle v_1, \dots, v_r \rangle$. □

EJEMPLO 2.2.2. Veremos una serie de ejemplos de subespacios, suma e intersección de subespacios.

(1) Sea $\mathbb{K} = \mathbb{C}$ y $V = \mathbb{C}^5$. Consideremos los vectores

$$v_1 = (1, 2, 0, 3, 0), \quad v_2 = (0, 0, 1, 4, 0), \quad v_3 = (0, 0, 0, 0, 1),$$

y sea $W = \langle v_1, v_2, v_3 \rangle$.

Ahora bien, $w \in W$, si y sólo si $w = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$, con $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$. Es decir

$$\begin{aligned} w &= \lambda_1(1, 2, 0, 3, 0) + \lambda_2(0, 0, 1, 4, 0) + \lambda_3(0, 0, 0, 0, 1) \\ &= (\lambda_1, 2\lambda_1, 0, 3\lambda_1, 0) + (0, 0, \lambda_2, 4\lambda_2, 0) + (0, 0, 0, 0, \lambda_3) \\ &= (\lambda_1, 2\lambda_1, \lambda_2, 3\lambda_1 + 4\lambda_2, \lambda_3) \end{aligned}$$

Luego, también podríamos escribir

$$W = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{C}^5 : x_2 = 2x_1, x_4 = 3x_1 + 4x_3\}.$$

(2) Sea $V = M_2(\mathbb{C})$ y sean

$$W_1 = \left\{ \begin{bmatrix} x & y \\ z & 0 \end{bmatrix} : x, y, z \in \mathbb{C} \right\}, \quad W_2 = \left\{ \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} : x, y \in \mathbb{C} \right\}.$$

Es claro que cada uno de estos conjuntos es un subespacio, pues,

$$W_1 = \mathbb{C} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \mathbb{C} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \mathbb{C} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad W_2 = \mathbb{C} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \mathbb{C} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Entonces, $W_1 + W_2 = V$. En efecto, sea $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in V$, entonces

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ z_1 & 0 \end{bmatrix} + \begin{bmatrix} x_2 & 0 \\ 0 & y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 & y_1 \\ z_1 & y_2 \end{bmatrix},$$

y esto se cumple tomando $x_1 = a, x_2 = 0, y_1 = b, z_1 = c, y_2 = d$.

Por otro lado

$$\begin{aligned} W_1 \cap W_2 &= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in W_1, \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in W_2 \right\} \\ &= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : \begin{array}{l} (a = x_1, b = y_1, c = z_1, d = 0) \wedge \\ (a = x_2, b = c = 0, d = y_2) \end{array} \right\} \\ &= \left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} : a \in \mathbb{C} \right\}. \end{aligned}$$

DEFINICIÓN 2.2.4. Sea $A = [a_{ij}] \in M_{m \times n}(\mathbb{K})$. El *vector fila* i es el vector $(a_{i1}, \dots, a_{in}) \in \mathbb{K}^n$. El *espacio fila* de A es el subespacio de \mathbb{K}^n generado por los m vectores fila de A . De forma análoga, se define el *vector columna* i al vector $(a_{1i}, \dots, a_{mi}) \in \mathbb{K}^m$ y el *espacio columna* de A es el subespacio de \mathbb{K}^m generado por los n vectores columna de A .

EJEMPLO 2.2.3. Sea

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

entonces, como vimos en el ejemplo 2.2.2(1), el espacio fila es

$$W = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{C}^5 : x_2 = 2x_1, x_4 = 3x_1 + 4x_3\}.$$

2.3. Bases y dimensión

DEFINICIÓN 2.3.1. Sea V un espacio vectorial sobre \mathbb{K} . Un subconjunto S de V se dice *linealmente dependiente* (o simplemente, *dependiente*) si existen vectores distintos $v_1, \dots, v_n \in S$ y escalares $\lambda_1, \dots, \lambda_n$ de \mathbb{K} , no todos nulos, tales que

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0.$$

Un conjunto que no es linealmente dependiente se dice *linealmente independiente* (o simplemente, *independiente*).

Si el conjunto S tiene solo un número finito de vectores v_1, \dots, v_n , se dice, a veces, que los v_1, \dots, v_n son dependientes (o independientes), en vez de decir que S es dependiente (o independiente).

Cuando un conjunto S es dependiente (independiente), diremos, para simplificar, que S es LD (resp. LI).

OBSERVACIÓN 2.3.1. Por definición, un conjunto $S = \{v_1, \dots, v_n\}$ es independiente si se cumple cualquiera de las dos afirmaciones siguientes:

- (a) dados $\lambda_1, \dots, \lambda_n$ en \mathbb{K} tal que $\lambda_i \neq 0$ para algún i , entonces $\lambda_1 v_1 + \dots + \lambda_n v_n \neq 0$, o ,
- (b) dados $\lambda_1, \dots, \lambda_n$ en \mathbb{K} tal que $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$, entonces $0 = \lambda_1 = \dots = \lambda_n$

El enunciado (a) se deduce intuitivamente negando la definición de linealmente dependiente y el resultado (b) es el contrarrecíproco de (a).

Para los interesados, esto es un ejercicio de lógica: ser linealmente dependiente se puede enunciar

$$(\exists \lambda_1, \dots, \lambda_n : (\exists i : \lambda_i \neq 0) \wedge (\lambda_1 v_1 + \dots + \lambda_n v_n = 0)),$$

luego, su negación (es decir ser LI), es

$$(\forall \lambda_1, \dots, \lambda_n : (\exists i : \lambda_i \neq 0) \Rightarrow \lambda_1 v_1 + \dots + \lambda_n v_n \neq 0),$$

(observar que $\neg(P \wedge Q) \equiv \neg P \vee \neg Q \equiv P \Rightarrow \neg Q$).

Como $(P \Rightarrow Q) \equiv (\neg Q \Rightarrow \neg P)$, la propiedad anterior es equivalente a

$$(\forall \lambda_1, \dots, \lambda_n : (\lambda_1 v_1 + \dots + \lambda_n v_n = 0) \Rightarrow (\forall i : \lambda_i = 0)),$$

Las siguientes afirmaciones son consecuencias fáciles de la definición.

- (1) Todo conjunto que contiene un conjunto linealmente dependiente es linealmente dependiente.
- (2) Todo subconjunto de un conjunto linealmente independiente es linealmente independiente.
- (3) Todo conjunto que contiene el vector 0 es linealmente dependiente; en efecto, $1 \cdot 0 = 0$.

EJEMPLO 2.3.1. En \mathbb{R}^3 los vectores $(1, -1, 1)$ y $(-1, 1, 1)$ son LI, pues si $\lambda_1(1, -1, 1) + \lambda_2(-1, 1, 1) = 0$, entonces $0 = (\lambda_1, -\lambda_1, \lambda_1) + (-\lambda_2, \lambda_2, \lambda_2) = (\lambda_1 - \lambda_2, -\lambda_1 + \lambda_2, \lambda_1 + \lambda_2)$, y esto es cierto si

$$\begin{aligned} \lambda_1 - \lambda_2 &= 0 \\ -\lambda_1 + \lambda_2 &= 0 \\ \lambda_1 + \lambda_2 &= 0 \end{aligned}$$

Luego $\lambda_1 = \lambda_2$ y $\lambda_1 = -\lambda_2$, por lo tanto $\lambda_1 = \lambda_2 = 0$. Es decir, hemos visto que

$$\lambda_1(1, -1, 1) + \lambda_2(-1, 1, 1) = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0,$$

y, por lo tanto, $(1, -1, 1)$ y $(-1, 1, 1)$ son LI.

EJEMPLO 2.3.2. Sea \mathbb{K} cuerpo. En \mathbb{K}^3 los vectores

$$v_1 = (3, 0, -3)$$

$$v_2 = (-1, 1, 2)$$

$$v_3 = (4, 2, -2)$$

$$v_4 = (2, 1, 1)$$

son linealmente dependientes, pues

$$2v_1 + 2v_2 - v_3 + 0.v_4 = 0.$$

Por otro lado, los vectores

$$e_1 = (1, 0, 0)$$

$$e_2 = (0, 1, 0)$$

$$e_3 = (0, 0, 1)$$

son linealmente independientes.

OBSERVACIÓN 2.3.2. En general, en \mathbb{K}^m , si queremos determinar si v_1, \dots, v_n es LI, planteamos la ecuación

$$\lambda_1 v_1 + \dots + \lambda_n v_n = (0, \dots, 0),$$

que, viéndola coordenada a coordenada, es equivalente a un sistema de m ecuaciones lineales con n incógnitas (que son $\lambda_1, \dots, \lambda_n$). Si la única solución es la trivial entonces v_1, \dots, v_n es LI. Si hay alguna solución no trivial, entonces v_1, \dots, v_n es LD.

DEFINICIÓN 2.3.2. Sea V un espacio vectorial. Una *base* de V es un conjunto $\mathcal{B} \subseteq V$ tal que

- (1) \mathcal{B} genera a V , y
- (2) \mathcal{B} es LI.

El espacio V es de *dimensión finita* si tiene una base finita, es decir con un número finito de elementos.

EJEMPLO 2.3.3 (BASE CANÓNICA DE \mathbb{K}^n). Sea el espacio vectorial \mathbb{K}^n y sean

$$e_1 = (1, 0, 0, \dots, 0)$$

$$e_2 = (0, 1, 0, \dots, 0)$$

$$\dots\dots\dots$$

$$e_n = (0, 0, 0, \dots, 1)$$

(e_i es el vector con todas sus coordenadas iguales a cero, excepto la coordenada i que vale 1). Entonces veamos que e_1, \dots, e_n es una base de \mathbb{K}^n .

(1) Si $(x_1, \dots, x_n) \in \mathbb{K}^n$, entonces

$$(x_1, \dots, x_n) = x_1 e_1 + \dots + x_n e_n.$$

Por lo tanto, e_1, \dots, e_n genera a \mathbb{K}^n .

(2) Si

$$x_1 e_1 + \dots + x_n e_n = 0,$$

entonces

$$\begin{aligned} (0, \dots, 0) &= x_1(1, 0, \dots, 0) + x_2(0, 1, \dots, 0) + \dots + x_n(0, 0, \dots, 1) \\ &= (x_1, 0, \dots, 0) + (0, x_2, \dots, 0) + \dots + (0, 0, \dots, x_n) \\ &= (x_1, x_2, \dots, x_n). \end{aligned}$$

Luego, $x_1 = x_2 = \dots = x_n = 0$ y por lo tanto e_1, \dots, e_n es LI.

Para $1 \leq i \leq n$, al vector e_i se lo denomina el *i-ésimo vector canónico* y a la base $\mathcal{B}_n = \{e_1, \dots, e_n\}$ se la denomina la *base canónica* de \mathbb{K}^n .

EJEMPLO 2.3.4. Sea P una matriz $n \times n$ inversible con elementos en el cuerpo \mathbb{K} . Entonces si C_1, \dots, C_n son los vectores columna de P (ver definición 2.2.4), estos forman una base de \mathbb{K}^n . Eso se verá como sigue. Si $X = (x_1, \dots, x_n) \in \mathbb{K}^n$, lo podemos ver como columna y

$$PX = x_1 C_1 + \dots + x_n C_n.$$

Como $PX = 0$ tiene solo la solución trivial $X = 0$, se sigue que $\{C_1, \dots, C_n\}$ es un conjunto linealmente independiente. ¿Por qué generan \mathbb{K}^n ? Sea $Y \in \mathbb{K}^n$, si $X = P^{-1}Y$, entonces $Y = PX$, esto es

$$Y = x_1 C_1 + \dots + x_n C_n.$$

Así, $\{C_1, \dots, C_n\}$ es una base de \mathbb{K}^n .

EJEMPLO 2.3.5. Sea $\mathbb{K}_n[x]$ el conjunto de polinomios de grado menor que n con coeficientes en \mathbb{K} :

$$\mathbb{K}_n[x] = \{a_0 + a_1 x + a_2 x^2 + \dots + a_{n-1} x^{n-1} : a_0, \dots, a_{n-1} \in \mathbb{K}\}.$$

Entonces $1, x, x^2, \dots, x^{n-1}$ es una base de $\mathbb{K}_n[x]$. Es claro que los $1, x, x^2, \dots, x^{n-1}$ generan $\mathbb{K}_n[x]$. Por otro lado, si $\lambda_0 + \lambda_1 x + \lambda_2 x^2 + \dots + \lambda_{n-1} x^{n-1} = 0$, tenemos que $\lambda_0 = \lambda_1 = \lambda_2 = \dots = \lambda_{n-1} = 0$.

EJEMPLO 2.3.6 (BASE CANÓNICA DE $M_{m \times n}(\mathbb{K})$). Sean $1 \leq i \leq m$, $1 \leq j \leq n$ y $E_{ij} \in M_{m \times n}(\mathbb{K})$ definida por

$$[E_{ij}]_{kl} = \begin{cases} 1 & \text{si } i = k \text{ y } j = l, \\ 0 & \text{otro caso.} \end{cases}$$

Es decir E_{ij} es la matriz cuyas entradas son todas iguales a 0, excepto la entrada ij que vale 1. En el caso 2×2 tenemos las matrices

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Volviendo al caso general, es claro que si $A = [a_{ij}] \in M_{m \times n}(\mathbb{K})$, entonces

$$(33) \quad A = \sum_{i=1}^m \sum_{j=1}^n a_{ij} E_{ij},$$

luego $\{E_{ij}\}_{1 \leq i \leq m, 1 \leq j \leq n}$ genera $M_{m \times n}(\mathbb{K})$. También, por la ecuación (33), es claro que si $\sum_{i=1}^m \sum_{j=1}^n a_{ij} E_{ij} = 0$, entonces $a_{ij} = 0$ para todo i y j . Luego, $\{E_{ij}\}_{1 \leq i \leq m, 1 \leq j \leq n}$ es LI.

Concluyendo, $\{E_{ij}\}_{1 \leq i \leq m, 1 \leq j \leq n}$ es una base de $M_{m \times n}(\mathbb{K})$ y se la denomina la *base canónica* de $M_{m \times n}(\mathbb{K})$.

Si S es un conjunto finito denotemos $|S|$ al *cardinal* de S es decir, la cantidad de elementos de S .

TEOREMA 2.3.1. *Sea V un espacio vectorial generado por un conjunto finito de vectores w_1, \dots, w_m . Entonces todo conjunto independiente de vectores de V es finito y contiene a lo más m elementos.*

DEMOSTRACIÓN. Sea $V = \langle w_1, \dots, w_m \rangle$ y $S \subset V$. El enunciado del teorema es equivalente a decir:

$$\text{si } S \text{ es LI} \Rightarrow |S| \leq m.$$

Para demostrar este teorema es suficiente probar el contrarrecíproco del enunciado, es decir:

$$\text{si } |S| > m \Rightarrow S \text{ es LD},$$

o, dicho de otra forma, todo subconjunto S de V que contiene más de m vectores es linealmente dependiente. Sea S un tal conjunto, entonces $S = \{v_1, \dots, v_n\}$ con $n > m$. Como w_1, \dots, w_m generan V , existen escalares a_{ij} en \mathbb{K} tales que

$$v_j = \sum_{i=1}^m a_{ij} w_i, \quad (1 \leq j \leq n).$$

Probaremos ahora que existen $x_1, \dots, x_n \in \mathbb{K}$ no todos nulos, tal que $x_1v_1 + \dots + x_nv_n = 0$. Ahora bien, para cualesquiera $x_1, \dots, x_n \in \mathbb{K}$ tenemos

$$\begin{aligned}
 x_1v_1 + \dots + x_nv_n &= \sum_{j=1}^n x_jv_j \\
 &= \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij}w_i \\
 &= \sum_{j=1}^n \sum_{i=1}^m (x_ja_{ij})w_i \\
 &= \sum_{i=1}^m \left(\sum_{j=1}^n x_ja_{ij} \right) w_i. \tag{*}
 \end{aligned}$$

El sistema de ecuaciones

$$\sum_{j=1}^n x_ja_{ij} = 0, \quad (1 \leq i \leq m)$$

tiene m ecuaciones y $n > m$ incógnitas, luego, por el teorema 1.2.6, existen escalares $x_1, \dots, x_n \in \mathbb{K}$ no todos nulos, tal que $\sum_{j=1}^n x_ja_{ij} = 0$, ($1 \leq i \leq m$) y, por (*), tenemos que $x_1v_1 + \dots + x_nv_n = 0$. Esto quiere decir que los v_1, \dots, v_n son LD. \square

COROLARIO 2.3.2. *Si V es un espacio vectorial de dimensión finita, entonces dos bases cualesquiera de V tienen el mismo número de elementos.*

DEMOSTRACIÓN. Como V es de dimensión finita, tiene una base finita \mathcal{B} de m vectores, es decir, \mathcal{B} es base de V y $|\mathcal{B}| = m$. Sea \mathcal{B}' otra base de V , como \mathcal{B} genera V y \mathcal{B}' es un conjunto LI, entonces, por el teorema anterior, $|\mathcal{B}'| \leq m$. Sea $n = |\mathcal{B}'|$, entonces $n \leq m$. Por otro lado \mathcal{B}' es base y, por lo tanto, genera V y \mathcal{B} es LI, luego, por el teorema anterior nuevamente, $m \leq n$, y en consecuencia $m = n$. \square

DEFINICIÓN 2.3.3. Sea V espacio vectorial de dimensión finita. Diremos que n es la *dimensión* de V y denotaremos $\dim V = n$, si existe una base de V de n vectores. Si $V = \{0\}$, entonces definimos $\dim V = 0$.

Como ya demostramos, si V es un espacio vectorial de dimensión finita y $\mathcal{B}, \mathcal{B}'$ dos bases de V , entonces $|\mathcal{B}| = |\mathcal{B}'|$.

EJEMPLO 2.3.7. Sean $m, n \in \mathbb{N}$.

- (1) $\dim \mathbb{K}^n = n$, pues la base canónica tiene n elementos.
- (2) $\dim M_{m \times n}(\mathbb{K}) = mn$, pues la base canónica de $M_{m \times n}(\mathbb{K})$ tiene mn elementos.
- (3) $\dim \mathbb{K}_n[x] = n$, pues $1, x, x^2, \dots, x^{n-1}$ es una base.

COROLARIO 2.3.3. *Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$. Entonces*

- (1) cualquier subconjunto de V con más de n vectores es linealmente dependiente;
 (2) ningún subconjunto de V con menos de n vectores puede generar V .

DEMOSTRACIÓN. (1) Sea $\{v_1, \dots, v_n\}$ una base de V , entonces v_1, \dots, v_n generan V , luego, por el teorema 2.3.1, cualquier subconjunto de V que contenga más de n vectores es LD.

- (2) Sea S subconjunto de V con $m < n$ vectores. Si S genera V , entonces todo subconjunto de más de m vectores es LD (teorema 2.3.1), por lo tanto, un subconjunto de n vectores es LD. En consecuencia, no puede haber una base de n elementos, lo cual contradice la hipótesis.

□

LEMA 2.3.4. Sea S un subconjunto linealmente independiente de un espacio vectorial V . Supóngase que w es un vector de V que no pertenece al subespacio generado por S . Entonces el conjunto que se obtiene agregando w a S , es linealmente independiente.

DEMOSTRACIÓN. Supóngase que v_1, \dots, v_n son vectores distintos de S y sean $\lambda_i, \lambda \in \mathbb{K}$ tales que

$$(34) \quad \lambda_1 v_1 + \dots + \lambda_n v_n + \lambda w = 0.$$

Debemos probar que $\lambda_i = 0$, $1 \leq i \leq n$, y $\lambda = 0$. Supongamos que $\lambda \neq 0$, entonces podemos dividir la ecuación por λ y haciendo pasaje de término obtenemos

$$w = \left(-\frac{\lambda_1}{\lambda}\right) v_1 + \dots + \left(-\frac{\lambda_n}{\lambda}\right) v_n.$$

Luego w estaría en el subespacio generado por S , lo cual contradice la hipótesis.

Por lo tanto $\lambda = 0$ y, en consecuencia

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0.$$

Como S es un conjunto linealmente independiente, todo $\lambda_i = 0$.

□

TEOREMA 2.3.5. Sea V espacio vectorial de dimensión finita n y S_0 un subconjunto LI de V . Entonces S_0 es finito y existen w_1, \dots, w_m vectores en V tal que $S_0 \cup \{w_1, \dots, w_m\}$ es una base de V .

DEMOSTRACIÓN. Se extiende S_0 a una base de V , como sigue. Si S_0 genera V , entonces S_0 es una base de V y está demostrado. Si S_0 no genera V , por el lema anterior se halla un vector w_1 en V tal que el conjunto $S_1 = S_0 \cup \{w_1\}$ es independiente. Si S_1 genera V , está demostrado. Si no, se aplica el lema para obtener un vector w_2 en V tal que el conjunto $S_2 = S_1 \cup \{w_2\} = S_0 \cup \{w_1, w_2\}$ es independiente. Si se continúa de este modo, entonces (y en no más de $\dim V$ de etapas) se llega a un conjunto

$$S_m = S_0 \cup \{w_1, \dots, w_m\}$$

que es independiente y que genera V (si no, continuamos), por lo tanto S_m es base de V . \square

Es decir, todo subconjunto LI de un espacio vectorial de dimensión finita se puede completar a una base.

COROLARIO 2.3.6. *Sea W es un subespacio de un espacio vectorial con de dimensión finita n y S_0 un subconjunto LI de W . Entonces, S_0 se puede completar a una base de W .*

DEMOSTRACIÓN. Como S_0 es un conjunto linealmente independiente de W , entonces S_0 es también un subconjunto linealmente independiente de V ; como V es de dimensión finita, S_0 no tiene más de n elementos y por lo tanto es finito.

Como W es un espacio vectorial, aplicando el teorema anterior completamos a una base de W . \square

COROLARIO 2.3.7. *Sea V espacio vectorial de dimensión finita y $V \neq \{0\}$, entonces $\dim V > 0$.*

DEMOSTRACIÓN. Como $V \neq \{0\}$, existe $v \in V$ con $v \neq 0$. Entonces, $S_0 = \{v\}$ es LI, pues $\lambda v = 0 \Rightarrow \lambda = 0$. Por el teorema anterior, S_0 se extiende a una base \mathcal{B} . Como $|\mathcal{B}| \geq |S_0| = 1$, tenemos que $\dim V > 0$. \square

COROLARIO 2.3.8. *Si W es un subespacio propio de un espacio vectorial de dimensión finita V , entonces W es de dimensión finita y $\dim W < \dim V$.*

DEMOSTRACIÓN. Si $W = \{0\}$, entonces $\dim W = 0$, como $W \subsetneq V$, tenemos que V es no nulo y por lo tanto $\dim W = 0 < \dim V$.

Si $W \neq \{0\}$, sea $w \in W$, $w \neq 0$. Por el teorema anterior existe una base \mathcal{B} de W que contiene a w . Como \mathcal{B} es un conjunto LI en W , lo es también en V y por teorema 2.3.1, $\dim W = |\mathcal{B}| \leq \dim V$.

Como W es un subespacio propio de V existe un vector v en V que no está en W . Agregando v a cualquier base de W se obtiene un subconjunto linealmente independiente de V (lema 2.3.4). Así, $\dim W < \dim V$. \square

Hemos visto que si V es un espacio de dimensión finita, entonces todo conjunto LI se puede extender a una base. Veremos ahora que dado un conjunto finito de generadores, existe un subconjunto que es una base.

TEOREMA 2.3.9. *Sea $V \neq 0$ espacio vectorial y S un conjunto finito de generadores de V , entonces existe un subconjunto \mathcal{B} de S que es una base.*

DEMOSTRACIÓN. Sea

$$C = \{|R| : R \subseteq S \wedge R \text{ es LI}\}.$$

Como V no es nulo y S genera V , $C \neq \emptyset$. C es un subconjunto no vacío de \mathbb{N} , acotado superiormente por $|S|$ y por lo tanto tiene máximo. Sea n el máximo de C y sea $\mathcal{B} \subseteq S$ tal que

$|\mathcal{B}| = n$ y \mathcal{B} es LI. Veremos que \mathcal{B} es una base. Para ello, como \mathcal{B} es LI, sólo falta ver que \mathcal{B} genera a V .

Supongamos que existe $v \in S$ tal que $v \notin \langle \mathcal{B} \rangle$. Por el lema 2.3.4, entonces $\mathcal{B} \cup \{v\}$ es LI y este subconjunto LI de S tiene $n+1$ elementos, lo cual contradice la maximalidad de n . Es claro entonces, que $v \in S \Rightarrow v \in \mathcal{B}$, es decir $S \subset \langle \mathcal{B} \rangle$. Como $S \subset \langle B \rangle$, entonces $V = \langle S \rangle \subset \langle B \rangle$, es decir $V = \langle B \rangle$. \square

TEOREMA 2.3.10. *Si W_1 , y W_2 son subespacios de dimensión finita de un espacio vectorial, entonces $W_1 + W_2$ es de dimensión finita y*

$$\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

DEMOSTRACIÓN. El conjunto $W_1 \cap W_2$ es un subespacio de W_1 y W_2 y por lo tanto un espacio vectorial de dimensión finita. Sea u_1, \dots, u_k una base de $W_1 \cap W_2$, por el teorema 2.3.5, existen v_1, \dots, v_n vectores en W_1 y w_1, \dots, w_m vectores en W_2 tal que

$$\{u_1, \dots, u_k, v_1, \dots, v_n\} \text{ es una base de } W_1,$$

y

$$\{u_1, \dots, u_k, w_1, \dots, w_m\} \text{ es una base de } W_2.$$

Es claro que, el subespacio $W_1 + W_2$ es generado por los vectores

$$u_1, \dots, u_k, v_1, \dots, v_n, w_1, \dots, w_m.$$

Veamos que estos vectores forman un conjunto independiente. En efecto, supóngase que

$$(35) \quad \sum \lambda_i u_i + \sum \gamma_i v_i + \sum \mu_i w_i = 0,$$

luego

$$\sum \mu_i w_i = -\sum \lambda_i u_i - \sum \gamma_i v_i.$$

Por lo tanto, $\sum \mu_i w_i \in (W_1 \cap W_2) + W_1 = W_1$. Es decir, $\sum \mu_i w_i \in W_2$ y $\sum \mu_i w_i \in W_1$, por lo tanto $\sum \mu_i w_i \in (W_1 \cap W_2)$, y entonces

$$\sum \mu_i w_i = \sum \alpha_i u_i \Rightarrow 0 = \sum \alpha_i u_i - \sum \mu_i w_i.$$

Como $\{u_1, \dots, u_k, w_1, \dots, w_m\}$ es una base y por lo tanto LI, tenemos que $0 = \alpha_i = \mu_j$, para todo i, j . Por lo tanto, por (35),

$$(36) \quad \sum \lambda_i u_i + \sum \gamma_i v_i = 0.$$

Como $\{u_1, \dots, u_k, v_1, \dots, v_n\}$ es una base de W_1 , tenemos que también $0 = \lambda_i = \gamma_j$ para todo i, j . Luego $0 = \lambda_i = \gamma_j = \mu_r$, para cualesquiera i, j, r y por lo tanto $u_1, \dots, u_k, v_1, \dots, v_n, w_1, \dots, w_m$ es LI y como generaban a $W_1 + W_2$ resultan ser una base de $W_1 + W_2$, por lo tanto $\dim(W_1 + W_2) = k + n + m$.

Finalmente,

$$\begin{aligned}\dim W_1 + \dim W_2 &= (k + n) + (k + m) \\ &= k + (k + n + m) \\ &= \dim(W_1 \cap W_2) + \dim(W_1 + W_2).\end{aligned}$$

□

2.4. Suma directa y complemento

En esta sección define la suma directa de subespacios y el complemento de un subespacio.

Si V_1, V_2, W dos subespacios del espacio vectorial V , entonces sabemos que

$$V_1 + V_2 = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}, \quad V_1 \cap V_2 = \{v : v \in V_1 \text{ y } v \in V_2\}$$

son subespacios vectoriales.

DEFINICIÓN 2.4.1. Sean V_1, V_2, W subespacios vectoriales del espacio vectorial V , entonces

$$W = V_1 \oplus V_2 \quad (W \text{ es suma directa de } V_1 \text{ y } V_2)$$

si $V_1 + V_2 = W$ y $V_1 \cap V_2 = 0$.

PROPOSICIÓN 2.4.1. Sea V espacio vectorial y V_1, V_2 subespacios vectoriales de V . Entonces, $V = V_1 \oplus V_2$ si y sólo si para todo vector $v \in V$ existe únicos $v_1 \in V_1, v_2 \in V_2$ tal que $v = v_1 + v_2$.

DEMOSTRACIÓN.

(\Rightarrow) Sea $v \in V$, como $V = V_1 + V_2$, existen $v_1 \in V_1, v_2 \in V_2$ tal que $v = v_1 + v_2$. Veamos que v_1 y v_2 son únicos. Sean $v'_1 \in V_1, v'_2 \in V_2$ tal que $v = v'_1 + v'_2$. Por lo tanto $v_1 + v_2 = v'_1 + v'_2$. Haciendo pasajes de término obtenemos

$$v_1 - v'_1 = v'_2 - v_2.$$

Sea $v_0 = v_1 - v'_1 = v'_2 - v_2$. Ahora bien, $v_1 - v'_1 \in V_1$, por lo tanto $v_0 = v_1 - v'_1 \in V_1$. Análogamente, $v'_2 - v_2 \in V_2$, por lo tanto $v_0 = v'_2 - v_2 \in V_2$. Es decir, $v_0 \in V_1 \cap V_2 = 0$, luego $v_0 = 0$, por lo tanto $v_1 = v'_1$ y $v_2 = v'_2$.

(\Leftarrow) Es claro que $V = V_1 + V_2$. Probemos que $V_1 \cap V_2 = 0$. Sea $v \in V_1 \cap V_2$. Por hipótesis, existe únicos $v_1 \in V_1, v_2 \in V_2$ tal que $v = v_1 + v_2$. Podemos escribir entonces

$$\begin{aligned}v &= v_1 + v_2, & v_1 \in V_1, & v_2 \in V_2 \\ v &= v + 0, & v \in V_1, & 0 \in V_2 \\ v &= 0 + v, & 0 \in V_1, & v \in V_2.\end{aligned}$$

Por la unicidad, resulta que $v_1 = v = 0$ y $v_2 = 0 = v$, es decir $v = 0$.

□

PROPOSICIÓN 2.4.2. Sea V espacio vectorial de dimensión finita y sean V_1, V_2 dos subespacios de V tal que $V = V_1 \oplus V_2$. Sea \mathcal{B}_1 base de V_1 y \mathcal{B}_2 base de V_2 , entonces $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ es base de V .

DEMOSTRACIÓN. Sea $\mathcal{B}_1 = \{u_1, \dots, u_r\}$ y $\mathcal{B}_2 = \{u_{r+1}, \dots, u_{r+s}\}$, debemos ver entonces que el conjunto $\mathcal{B} = \{u_1, \dots, u_{r+s}\}$ genera todo el espacio y es LI.

- Sea $v \in V$, como $V_1 + V_2 = V$, existen $v_1 \in V_1$ y $v_2 \in V_2$ tales que $v = v_1 + v_2$. Como \mathcal{B}_1 es base de V_1 , tenemos que $v_1 = a_1 u_1 + \dots + a_r u_r$, análogamente $v_2 = a_{r+1} u_{r+1} + \dots + a_{r+s} u_{r+s}$ y por lo tanto $v = a_1 u_1 + \dots + a_{r+s} u_{r+s}$. Es decir \mathcal{B} genera V .

- Si $a_1 u_1 + \dots + a_r u_r + a_{r+1} u_{r+1} + \dots + a_{r+s} u_{r+s} = 0$, entonces

$$a_1 u_1 + \dots + a_r u_r = -a_{r+1} u_{r+1} - \dots - a_{r+s} u_{r+s}.$$

Ahora bien, el término de la izquierda en la última igualdad pertenece a V_1 , mientras que el de la derecha pertenece a V_2 . Como $V_1 \cap V_2 = 0$, tenemos que

$$a_1 u_1 + \dots + a_r u_r = 0 = -a_{r+1} u_{r+1} - \dots - a_{r+s} u_{r+s}.$$

Como \mathcal{B}_1 es base de V_1 , $a_1 = \dots = a_r = 0$ y como \mathcal{B}_2 es base de V_2 , $a_{r+1} = \dots = a_{r+s} = 0$. Es decir \mathcal{B} es LI. \square

COROLARIO 2.4.3. Sea V espacio vectorial de dimensión finita y sean V_1, V_2 dos subespacios de V tal que $V = V_1 \oplus V_2$. Entonces $\dim(V) = \dim(V_1) + \dim(V_2)$.

DEMOSTRACIÓN. Sea $\mathcal{B}_1 = \{u_1, \dots, u_r\}$ base de V_1 y $\mathcal{B}_2 = \{u_{r+1}, \dots, u_{r+s}\}$ base de V_2 . Por la proposición anterior, $\mathcal{B} = \{u_1, \dots, u_{r+s}\}$ es base de V . Luego $\dim(V) = r + s = \dim(V_1) + \dim(V_2)$. \square

Se puede generalizar la noción de suma directa a varios subespacios.

DEFINICIÓN 2.4.2. Sean V_1, \dots, V_k, W subespacios vectoriales de un espacio vectorial V , entonces

$$W = V_1 \oplus V_2 \oplus \dots \oplus V_k$$

si $V_1 + V_2 + \dots + V_k = W$ y $V_j \cap (\sum_{i \neq j} V_i) = 0$. En este caso diremos que W es *suma directa* de V_1, \dots, V_k .

Esta definición se reduce a la de suma directa de dos subespacios cuando $k = 2$.

Observar que si definimos $W_j = \sum_{i \neq j} V_i$ ($j = 1, \dots, k$) entonces,

$$W = V_1 \oplus V_2 \oplus \dots \oplus V_k \text{ si y sólo si } W = V_j \oplus W_j \quad (j = 1, \dots, k).$$

DEFINICIÓN 2.4.3. Sea W un subespacio vectorial de un espacio vectorial V . Entonces un complemento de W es un subespacio U de V tal que $V = W \oplus U$.

PROPOSICIÓN 2.4.4. Sea V espacio vectorial de dimensión finita y sea W un subespacio de V . Sean \mathcal{B}_W una base de W y \mathcal{B}_V una base de V tal que $\mathcal{B}_W \subset \mathcal{B}_V$.

Sea

$$\mathcal{B}' = \mathcal{B}_V - \mathcal{B}_W = \{b \in \mathcal{B}_V \text{ tales que } b \notin \mathcal{B}_W\}.$$

Entonces $U = \langle \mathcal{B}' \rangle$ es un complemento de W y \mathcal{B}' es una base de U .

DEMOSTRACIÓN. Sea $\mathcal{B}_W = \{u_1, \dots, u_r\}$ y $\mathcal{B}_V = \{u_1, \dots, u_r, u_{r+1}, \dots, u_{r+s}\}$. Así, $\mathcal{B}' = \{u_{r+1}, \dots, u_{r+s}\}$. Como este conjunto es LI, entonces es una base del espacio $U = \langle \mathcal{B}' \rangle$ que genera. Por otro lado, como \mathcal{B}_V es base de V , entonces todo vector $v \in V$ puede escribirse como

$$v = a_1 u_1 + \dots + a_r u_r + a_{r+1} u_{r+1} + \dots + a_{r+s} u_{r+s},$$

para algunos $a_1, \dots, a_{r+s} \in \mathbb{K}$. Ahora, definimos $v_W = a_1 u_1 + \dots + a_r u_r$ y $v_U = a_{r+1} u_{r+1} + \dots + a_{r+s} u_{r+s}$, de manera tal que $v_W \in W$, $v_U \in U$ y $v = v_W + v_U$.

Finalmente, si $v \in W \cap U$, entonces existen $a_1, \dots, a_{r+s} \in \mathbb{K}$ tales que

$$v = a_1 u_1 + \dots + a_r u_r = a_{r+1} u_{r+1} + \dots + a_{r+s} u_{r+s}.$$

Pero esto determina que \mathcal{B}_V es LI (es una base). Luego, $W \cap U = \{0\}$ y por lo tanto $V = W \oplus U$. \square

2.5. Dimensiones de subespacios

Dada $A \in M_{m \times n}(\mathbb{K})$, ya hemos visto que las soluciones del sistema $AX = 0$ forman un subespacio vectorial. Sea R la MERF equivalente por filas a A y r la cantidad de filas no nulas de R . Ahora bien, cada fila no nula está asociada a una variable principal y las $n - r$ variables restantes son variables libres que generan todas las soluciones. El hecho de que tenemos $n - r$ variables libres no dice que hay $n - r$ vectores LI que generan W , y por lo tanto, $\dim W = n - r$. Esto lo veremos en ejemplos.

EJEMPLO 2.5.1. Encontrar una base del subespacio

$$W = \left\{ (x, y, z, w) \in \mathbb{R} : \begin{array}{rcl} x - y - 3z + w & = & 0 \\ y + 5z + 3w & = & 0 \end{array} \right\}.$$

SOLUCIÓN. W está definido implícitamente y usando el método de Gauss podemos describirlo paramétricamente, pues:

$$\begin{bmatrix} 1 & -1 & -3 & 1 \\ 0 & 1 & 5 & 3 \end{bmatrix} \xrightarrow{F_1 + F_2} \begin{bmatrix} 1 & 0 & 2 & 4 \\ 0 & 1 & 5 & 3 \end{bmatrix}.$$

Por lo tanto, el sistema de ecuaciones que define W es equivalente a

$$\begin{array}{rcl} x + 2z + 4w & = & 0 \\ y + 5z + 3w & = & 0, \end{array}$$

es decir

$$\begin{aligned}x &= -2z - 4w \\ y &= -5z - 3w,\end{aligned}$$

y por lo tanto

$$\begin{aligned}W &= \{(-2z - 4w, -5z - 3w, z, w) : z, w \in \mathbb{R}\} \\ &= \{(-2, -5, 1, 0)z + (-4, -3, 0, 1)w : z, w \in \mathbb{R}\} \\ &= \langle (-2, -5, 1, 0), (-4, -3, 0, 1) \rangle.\end{aligned}$$

Concluimos entonces que $(-2, -5, 1, 0), (-4, -3, 0, 1)$ es una base de W y, por lo tanto, su dimensión es 2. \square

TEOREMA 2.5.1. *Sean A matriz $m \times n$ con coeficientes en \mathbb{K} , P matriz $m \times m$ inversible y $B = PA$. Entonces el espacio fila de A es igual al espacio fila de B .*

DEMOSTRACIÓN. Sea $A = [a_{ij}]$, $P = [p_{ij}]$ y $B = [b_{ij}]$. Como $B = PA$, tenemos que la fila i de B es

$$\begin{aligned}(b_{i1}, \dots, b_{in}) &= (F_i(P) \cdot C_1(A), \dots, F_i(P) \cdot C_n(A)) \\ &= \left(\sum_{j=1}^m p_{ij} a_{j1}, \dots, \sum_{j=1}^m p_{ij} a_{jn} \right) \\ &= \sum_{j=1}^m p_{ij} (a_{j1}, \dots, a_{jn}).\end{aligned}$$

Luego, cada vector fila de B se puede obtener como combinación lineal de los vectores fila de A , y por lo tanto el espacio fila de B está incluido en el espacio fila de A .

Ahora bien, como P inversible, podemos multiplicar por P^{-1} a izquierda la fórmula $B = PA$, y obtenemos $P^{-1}B = P^{-1}PA = A$. Haciendo el mismo razonamiento que arriba concluimos que también el espacio fila de A está incluido en el espacio fila de B y por lo tanto son iguales. \square

COROLARIO 2.5.2. *Sean A matriz $m \times n$ y R la MERF equivalente por filas a A . Entonces, el espacio fila de A es igual al espacio fila de R y las filas no nulas de R forman una base del espacio fila de A .*

DEMOSTRACIÓN. $R = PA$, donde P es una matriz $m \times m$ inversible, luego, por el teorema anterior, el espacio fila de A es igual al espacio fila de R . Calculemos ahora cual es la dimensión del espacio fila de R . Veamos que filas no nulas de R son LI.

Recordemos que por definición de MERF cada fila no nula comienza con un 1 y en esa coordenada todas las demás filas tienen un 0, por lo tanto una combinación lineal no trivial resulta en un vector no nulo: si v es una fila no nula de R , con el 1 principal en la coordenada

i y $\lambda \neq 0$, entonces λv vale λ en la posición i y esta coordenada no puede ser anulada por la combinación de otras filas. \square

COROLARIO 2.5.3. Sean A matriz $n \times n$. Entonces, A es inversible si y sólo si las filas de A son una base de \mathbb{K}^n .

DEMOSTRACIÓN. Si A es inversible entonces la MERF de A es la identidad, por lo tanto el espacio fila de A genera \mathbb{K}^n .

Por otro lado, si el espacio fila de A genera \mathbb{K}^n , el espacio fila de la MERF es \mathbb{K}^n y por lo tanto la MERF de A es la identidad y en consecuencia A es inversible.

Hemos probado que A es inversible si y sólo si las n filas de A generan \mathbb{K}^n . Como $\dim \mathbb{K}^n = n$, todo conjunto de n generadores es una base. \square

El corolario 2.5.2 nos provee un método para encontrar una base de un subespacio de \mathbb{K}^n generado por m vectores: si $v_1, \dots, v_m \in \mathbb{K}^n$ y $W = \langle v_1, \dots, v_m \rangle$, consideramos la matriz

$$A = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

donde las filas son los vectores v_1, \dots, v_m . Luego calculamos R , la MERF equivalente por filas a A , y si R tiene r filas no nulas, las r filas no nulas son una base de W y, por consiguiente, $\dim W = r$.

EJEMPLO 2.5.2. Encontrar una base de $W = \langle (1, 0, 1), (1, -1, 0), (5, -3, 2) \rangle$.

SOLUCIÓN. Formemos la matriz cuyas filas son los vectores que generan W , es decir

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 5 & -3 & 2 \end{bmatrix}.$$

Entonces

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 5 & -3 & 2 \end{bmatrix} \xrightarrow[F_3 - 5F_1]{F_2 - F_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & -3 & -3 \end{bmatrix} \xrightarrow{-F_3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & -3 & -3 \end{bmatrix} \xrightarrow{F_3 - 3F_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Por lo tanto, $\dim W = 2$ y $(1, 0, 1), (0, 1, 1)$ es una base de W . \square

El método que nos provee el corolario 2.5.2 nos permite encontrar una base de un subespacio vectorial de \mathbb{K}^n a partir de un conjunto de generadores del subespacio. Como vimos en el teorema

2.3.9, en todo conjunto finito de generadores existe un subconjunto que es una base. El siguiente teorema nos permite encontrar uno de tales subconjuntos.

TEOREMA 2.5.4. *Sea v_1, \dots, v_r vectores en \mathbb{K}^n y $W = \langle v_1, \dots, v_r \rangle$. Sea A la matriz formada por las filas v_1, \dots, v_r y R una MRF equivalente por filas a A que se obtiene sin el uso de permutaciones de filas (ver observación 1.2.2). Si i_1, i_2, \dots, i_s son las filas no nulas de R , entonces $v_{i_1}, v_{i_2}, \dots, v_{i_s}$ es una base de W .*

DEMOSTRACIÓN. Se hará por inducción sobre r .

Si $r = 1$ es trivial ver que vale la afirmación.

Supongamos que tenemos el resultado probado para $1 \leq k < r$ (hipótesis inductiva). Sea $W' = \langle v_1, \dots, v_{r-1} \rangle$ y sea A' la matriz formada por las $r - 1$ filas v_1, \dots, v_{r-1} . Sea R' la MRF equivalente por filas a A' que se obtiene sin usar permutaciones de filas. Por hipótesis inductiva, si i_1, i_2, \dots, i_s son las filas no nulas de R' , entonces $v_{i_1}, v_{i_2}, \dots, v_{i_s}$ es una base de W' .

Sea

$$R_0 = \begin{bmatrix} R' \\ v_r \end{bmatrix}$$

donde R' es la MRF de A' .

Si $v_r \in W'$, entonces $v_{i_1}, v_{i_2}, \dots, v_{i_s}$ es una base de W y

$$R = \begin{bmatrix} R' \\ 0 \end{bmatrix}$$

es la MRF de A .

Si $v_r \notin W'$, entonces $v_{i_1}, v_{i_2}, \dots, v_{i_s}, v_r$ es una base de W (lema 2.3.4) y la MRF de A tiene la última fila no nula. \square

EJEMPLO 2.5.3. Sea $S = \{(1, 0, 1), (1, -1, 0), (5, -3, 2)\}$ y $W = \langle S \rangle$. Encontrar una base de W que sea un subconjunto de S .

SOLUCIÓN. Hemos visto en el ejemplo 2.5.2 que una MRF de A es

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

y que la misma se obtiene sin usar permutaciones. Esta matriz tiene las dos primeras filas no nulas, por lo tanto, $\{(1, 0, 1), (1, -1, 0)\}$ es una base de W . \square

Finalmente, terminaremos esta sección con un teorema que resume algunas equivalencias respecto a matrices inversibles.

TEOREMA 2.5.5. *Sea A matriz $n \times n$ con coeficientes en \mathbb{K} . Entonces son equivalentes*

(1) *A es inversible.*

- (2) A es equivalente por filas a I_n .
- (3) A es producto de matrices elementales.
- (4) El sistema $AX = Y$ tiene una única solución para toda matriz Y de orden $n \times 1$.
- (5) El sistema homogéneo $AX = 0$ tiene una única solución trivial.
- (6) $\det A \neq 0$.
- (7) Las filas de A son LI.
- (8) Las columnas de A son LI.

DEMOSTRACIÓN. Por teoremas 1.5.4 y 1.5.7, tenemos que $(1) \Leftrightarrow (2) \Leftrightarrow (3) \Leftrightarrow (4) \Leftrightarrow (5)$.

$(1) \Leftrightarrow (6)$. Por teorema 1.6.7.

$(1) \Leftrightarrow (7)$. Por corolario 2.5.3.

$(1) \Leftrightarrow (8)$. A inversible $\Leftrightarrow A^t$ inversible \Leftrightarrow las filas de A^t son LI \Leftrightarrow las columnas de A son LI. □

2.6. Coordenadas

Una de las características útiles de una base \mathcal{B} en un espacio vectorial V de dimensión n es que permite introducir coordenadas en V en forma análoga a las “coordenadas naturales”, x_i , de un vector $v = (x_1, \dots, x_n)$ en el espacio \mathbb{K}^n . En este esquema, las coordenadas de un vector v en V , respecto de la base \mathcal{B} , serán los escalares que sirven para expresar v como combinación lineal de los vectores de la base. En el caso de la base canónica e_1, \dots, e_n de \mathbb{K}^n tenemos

$$v = (x_1, \dots, x_n) = \sum_{i=1}^n x_i e_i.$$

por lo tanto x_i es la coordenada i -ésima de v respecto a la base canónica.

En forma análoga veremos que si v_1, \dots, v_n es una base de V , entonces existe una única forma de escribir

$$v = \sum_{i=1}^n x_i v_i,$$

y los valores x_i serán las coordenadas de v en la base dada.

DEFINICIÓN 2.6.1. Si V es un espacio vectorial de dimensión finita, una *base ordenada* de V es una sucesión finita de vectores linealmente independiente y que genera V .

La diferencia entre la definición de “base” y la de “base ordenada”, es que en la última es importante el orden de los vectores de la base. Si la sucesión v_1, \dots, v_n es una base ordenada de V , entonces el conjunto $\{v_1, \dots, v_n\}$ es una base de V . La base ordenada es el conjunto, juntamente con el orden dado. Se incurrirá en un pequeño abuso de notación y se escribirá

$$\mathcal{B} = \{v_1, \dots, v_n\}$$

diciendo que \mathcal{B} es una base ordenada de V .

PROPOSICIÓN 2.6.1. Sea V espacio vectorial de dimensión finita y sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ordenada de V . Entonces, para cada $v \in V$, existen únicos $x_1, \dots, x_n \in \mathbb{K}$ tales que

$$v = x_1 v_1 + \dots + x_n v_n.$$

DEMOSTRACIÓN. Como v_1, \dots, v_n generan V , es claro que existen $x_1, \dots, x_n \in \mathbb{K}$ tales que $v = x_1 v_1 + \dots + x_n v_n$. Sean $y_1, \dots, y_n \in \mathbb{K}$ tales que $v = y_1 v_1 + \dots + y_n v_n$. Veremos que $x_i = y_i$ para $1 \leq i \leq n$.

Como $v = \sum_{i=1}^n x_i v_i$ y $v = \sum_{i=1}^n y_i v_i$, restando miembro a miembro obtenemos

$$0 = \sum_{i=1}^n (x_i - y_i) v_i.$$

Ahora bien, v_1, \dots, v_n son LI, por lo tanto todos los coeficientes de la ecuación anterior son nulos, es decir $x_i - y_i = 0$ para $1 \leq i \leq n$ y entonces $x_i = y_i$ para $1 \leq i \leq n$. \square

La proposición anterior permite asociar a cada vector una n -upla: sea V espacio vectorial de dimensión finita y sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ordenada de V , si $v \in V$ y

$$v = x_1 v_1 + \dots + x_n v_n,$$

entonces x_i es la coordenada i -ésima de v y denotamos

$$v = (x_1, \dots, x_n)_{\mathcal{B}}.$$

También nos será útil describir a v como una matriz $n \times 1$: la matriz de v en la base \mathcal{B} es

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

EJEMPLO 2.6.1. Sea $\mathcal{B} = \{(1, -1), (2, 3)\}$ base ordenada de \mathbb{R}^2 . Encontrar las coordenadas de $(1, 0)$ y $(0, 1)$ en la base \mathcal{B} .

SOLUCIÓN. Debemos encontrar $x_1, x_2 \in \mathbb{R}$ tal que

$$(1, 0) = x_1(1, -1) + x_2(2, 3).$$

Es decir

$$x_1 + 2x_2 = 1$$

$$-x_1 + 3x_2 = 0.$$

Resolviendo el sistema de ecuaciones obtenemos $x_1 = \frac{3}{5}$ y $x_2 = \frac{1}{5}$, es decir

$$(1, 0) = \frac{3}{5}(1, -1) + \frac{1}{5}(2, 3) \quad \text{o equivalentemente} \quad (1, 0) = \left(\frac{3}{5}, \frac{1}{5}\right)_{\mathcal{B}}.$$

De forma análoga podemos ver que

$$(0, 1) = -\frac{2}{5}(1, -1) + \frac{1}{5}(2, 3) \quad \text{o equivalentemente} \quad (0, 1) = \left(-\frac{2}{5}, \frac{1}{5}\right)_{\mathcal{B}}.$$

□

PROPOSICIÓN 2.6.2. *Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ordenada de V un \mathbb{K} -espacio vectorial. Entonces*

- (1) $[v + w]_{\mathcal{B}} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}$, para $v, w \in V$,
- (2) $[\lambda v]_{\mathcal{B}} = \lambda[v]_{\mathcal{B}}$, para $\lambda \in \mathbb{K}$ y $v \in V$.

DEMOSTRACIÓN. (1) Si $v = x_1v_1 + \dots + x_nv_n$ y $w = y_1v_1 + \dots + y_nv_n$, entonces

$$v + w = (x_1 + y_1)v_1 + \dots + (x_n + y_n)v_n,$$

luego

$$[v + w]_{\mathcal{B}} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}.$$

(2) Si $v = x_1v_1 + \dots + x_nv_n$ y $\lambda \in \mathbb{K}$, entonces

$$\lambda v = (\lambda x_1)v_1 + \dots + (\lambda x_n)v_n,$$

luego

$$[\lambda v]_{\mathcal{B}} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \lambda[v]_{\mathcal{B}}.$$

□

TEOREMA 2.6.3. *Sea V espacio vectorial de dimensión n sobre el cuerpo \mathbb{K} , y sean $\mathcal{B} = \{v_1, \dots, v_n\}$ y $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ bases ordenadas de V . Entonces existe una única matriz P de orden $n \times n$, inversible, tal que para todo $v \in V$*

- (1) $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$ y
- (2) $[v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}}$.

Las columnas de P están dadas por

$$C_j = [v'_j]_{\mathcal{B}}, \quad 1 \leq j \leq n,$$

es decir, los coeficientes de la columna j de P son las coordenadas de v'_j en la base \mathcal{B} .

DEMOSTRACIÓN. (1) Tenemos

$$v = x_1 v_1 + \cdots + x_n v_n,$$

$$v = x'_1 v'_1 + \cdots + x'_n v'_n.$$

También podemos escribir cada v'_j en coordenadas respecto a \mathcal{B} :

$$v'_j = p_{1j} v_1 + \cdots + p_{nj} v_n, \quad 1 \leq j \leq n.$$

Luego,

$$\begin{aligned} \sum_{i=1}^n x_i v_i &= \sum_{j=1}^n x'_j v'_j \\ &= \sum_{j=1}^n x'_j \left(\sum_{i=1}^n p_{ij} v_i \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n x'_j p_{ij} v_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij} x'_j \right) v_i. \end{aligned}$$

Es decir, $x_i = \sum_{j=1}^n p_{ij} x'_j$. Por lo tanto, si $P = [p_{ij}]$ tenemos que

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix}.$$

Veamos que P es inversible: sea (x_1, \dots, x_n) solución del sistema de ecuaciones $PX = 0$. Eso implica que si $v = \sum x_i v'_i$, entonces $P[v]_{\mathcal{B}'} = 0$. Eso quiere decir que el vector escrito en la base \mathcal{B} tiene coordenadas 0, es decir $v = \sum 0 \cdot v_i = 0$ y entonces $v = \sum 0 \cdot v'_i$ por la unicidad de escritura. Por lo tanto, la única solución del sistema $PX = 0$ es 0 y en consecuencia P es inversible.

(2) Como P es inversible, multiplicamos por P^{-1} a izquierda la ecuación $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$ y obtenemos el resultado. \square

Observemos que, por simetría, los coeficientes de la columna j de P^{-1} son las coordenadas de v_j en la base \mathcal{B}' .

DEFINICIÓN 2.6.2. Sea V espacio vectorial de dimensión n sobre el cuerpo \mathbb{K} , y sean $\mathcal{B} = \{v_1, \dots, v_n\}$ y $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ bases ordenadas de V . Entonces la matriz

$$P = \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix}$$

con

$$C_j = [v'_j]_{\mathcal{B}}, \quad 1 \leq j \leq n,$$

es la matriz de *cambio de coordenadas* de la base \mathcal{B}' a la base \mathcal{B} .

EJEMPLO 2.6.2. $\mathcal{B} = \{v_1, v_2, v_3\} = \{(1, 0, -1), (1, 1, 1), (1, 0, 0)\}$ en \mathbb{R}^3 .

- (1) Probar que \mathcal{B} es una base de \mathbb{R}^3 , y
- (2) encontrar las coordenadas del vector $(2, 3, 5)$ respecto a la base \mathcal{B} .

SOLUCIÓN. (1) Sea A la matriz cuyas filas son los vectores de \mathcal{B} :

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

Entonces, calculando el determinante por desarrollo de la última fila, obtenemos

$$\det A = 1 \cdot (-1)^{3+1} \cdot \begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix} = 1 \cdot 1 \cdot 1 = 1.$$

Por lo tanto A es inversible y en consecuencia el espacio fila de A es \mathbb{R}^3 . Como \mathcal{B} es un conjunto de 3 generadores de \mathbb{R}^3 , \mathcal{B} es base.

- (2) Si x_1, x_2, x_3 son las coordenadas de $(2, 3, 5)$ respecto a la base \mathcal{B} , entonces

$$(2, 3, 5) = x_1(1, 0, -1) + x_2(1, 1, 1) + x_3(1, 0, 0)$$

y esto es un sistema de tres ecuaciones y tres incógnitas. Resolviendo este sistema encontramos la solución buscada.

Sin embargo, nos interesa conocer una solución que sirva para cualquier vector, no solo para $(2, 3, 5)$, es decir buscamos la matriz $P = [p_{ij}]$ de cambio de coordenadas de la base canónica a la base \mathcal{B} . En particular, esta matriz cumple

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}.$$

con $(x_1, x_2, x_3)_{\mathcal{B}} = (2, 3, 5)$.

Ahora bien, sabemos que P^{-1} es la matriz cuya columna j es la matriz de v_j en la base canónica (fácil de calcular) y luego calculamos su inversa que es P .

Por lo dicho en el párrafo anterior, P^{-1} es la matriz cuya columna 1 es $(1, 0, -1)$, la columna 2 es $(1, 1, 1)$ y la columna 3 es $(1, 0, 0)$, luego

$$P^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{bmatrix}.$$

Encontremos su inversa,

$$\begin{aligned}
 \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \end{array} \right] &\xrightarrow{F_3+F_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 1 \end{array} \right] \xrightarrow[F_3-2F_2]{F_1-F_2} \\
 &\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} \right] \xrightarrow{F_1-F_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} \right].
 \end{aligned}$$

Entonces

$$P = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}$$

y

$$\begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}.$$

Por lo tanto las coordenadas de $(2, 3, 5)$ respecto a la base \mathcal{B} son $(-2, 3, 1)_{\mathcal{B}}$ o, equivalentemente,

$$(2, 3, 5) = -2(1, 0, -1) + 3(1, 1, 1) + (1, 0, 0).$$

□

Transformaciones lineales

Las transformaciones lineales son las funciones con las que trabajaremos en álgebra lineal. Se trata de funciones entre espacios vectoriales que son compatibles con la estructura, es decir con la suma y el producto por escalares.

3.1. Transformaciones lineales

DEFINICIÓN 3.1.1. Sean V y W dos espacios vectoriales sobre el cuerpo \mathbb{K} . Una *transformación lineal* de V en W es una función $T : V \rightarrow W$ tal que

- (1) $T(v + v') = T(v) + T(v')$, para $v, v' \in V$,
- (2) $T(\lambda v) = \lambda T(v)$, para $v \in V$, $\lambda \in \mathbb{K}$.

OBSERVACIÓN 3.1.1. $T : V \rightarrow W$ es transformación lineal si y sólo si

- (a) $T(\lambda v + v') = \lambda T(v) + T(v')$, para $v, v' \in V$, $\lambda \in \mathbb{K}$.

Algunas veces usaremos esto último para comprobar si una aplicación de V en W es una transformación lineal.

EJEMPLO 3.1.1. Si V es cualquier espacio vectorial, la transformación identidad I , definida por $Iv = v$ ($v \in V$), es una transformación lineal de V en V . La transformación cero 0 , definida por $0v = 0$, es una transformación lineal de V en V .

EJEMPLO 3.1.2. Sea $T : \mathbb{K}^3 \rightarrow \mathbb{K}^2$ definida por

$$T(x_1, x_2, x_3) = (2x_1 - x_3, -x_1 + 3x_2 + x_3).$$

Entonces, T es una transformación lineal. La demostración la veremos en la observación que sigue a este ejemplo.

Observar que si

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 3 & 1 \end{bmatrix},$$

entonces

$$\begin{bmatrix} 2 & 0 & -1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_3 \\ -x_1 + 3x_2 + x_3 \end{bmatrix}.$$

Es decir, si \mathcal{C}_n es la base canónica de \mathbb{K}^n y $[x]_{\mathcal{C}_3}$ es la matriz de x en la base canónica, entonces

$$A \cdot [x]_{\mathcal{C}_3} = [T(x)]_{\mathcal{C}_2}.$$

OBSERVACIÓN 3.1.2. Sea $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$. En general si $T(x_1, \dots, x_n)$ en cada coordenada tiene una combinación lineal de los x_1, \dots, x_n , entonces T es una transformación lineal. Mas precisamente, si T está definida por

$$\begin{aligned} T(x_1, \dots, x_n) &= (a_{11}x_1 + \dots + a_{1n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n) \\ &= \left(\sum_{j=1}^n a_{1j}x_j, \dots, \sum_{j=1}^n a_{mj}x_j \right), \end{aligned}$$

con $a_{ij} \in \mathbb{K}$, entonces T es lineal.

DEMOSTRACIÓN. Sean $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{K}^n$ y $\lambda \in \mathbb{K}$, entonces

$$\begin{aligned} T(\lambda(x_1, \dots, x_n) + (y_1, \dots, y_n)) &= T(\lambda x_1 + y_1, \dots, \lambda x_n + y_n) \\ &= \left(\sum_{j=1}^n a_{1j}(\lambda x_j + y_j), \dots, \sum_{j=1}^n a_{mj}(\lambda x_j + y_j) \right) \\ &= \left(\lambda \sum_{j=1}^n a_{1j}x_j + \sum_{j=1}^n a_{1j}y_j, \dots, \lambda \sum_{j=1}^n a_{mj}x_j + \sum_{j=1}^n a_{mj}y_j \right) \\ &= \lambda \left(\sum_{j=1}^n a_{1j}x_j, \dots, \sum_{j=1}^n a_{mj}x_j \right) + \left(\sum_{j=1}^n a_{1j}y_j, \dots, \sum_{j=1}^n a_{mj}y_j \right) \\ &= \lambda T(x_1, \dots, x_n) + T(y_1, \dots, y_n). \end{aligned}$$

□

EJEMPLO 3.1.3 (Transformaciones de \mathbb{R}^2 en \mathbb{R}^2). Las rotaciones y reflexiones en \mathbb{R}^2 son transformaciones lineales. Sea $\theta \in \mathbb{R}$ tal que $0 \leq \theta \leq 2\pi$, definimos

$$\begin{aligned} R_\theta : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto (x \cos \theta - y \operatorname{sen} \theta, y \cos \theta + x \operatorname{sen} \theta) \end{aligned}$$

Observemos que si escribimos el vector (x, y) en coordenadas polares, es decir si

$$(x, y) = r(\cos \alpha, \operatorname{sen} \alpha), \quad r > 0, \quad 0 \leq \alpha < 2\pi,$$

entonces

$$\begin{aligned} R_\theta(x, y) &= R_\theta(r \cos \alpha, r \operatorname{sen} \alpha) \\ &= (r \cos \alpha \cos \theta - r \operatorname{sen} \alpha \operatorname{sen} \theta, r \operatorname{sen} \alpha \cos \theta + r \cos \alpha \operatorname{sen} \theta) \\ &= (r \cos(\alpha + \theta), r \operatorname{sen}(\alpha + \theta)) \\ &= r(\cos(\alpha + \theta), \operatorname{sen}(\alpha + \theta)). \end{aligned}$$

Por lo tanto $R_\theta(x, y)$ es el vector (x, y) rotado θ grados en sentido antihorario y en consecuencia R_θ es denominada la *rotación antihoraria en θ radianes*. No es difícil verificar que R_θ es una transformación lineal.

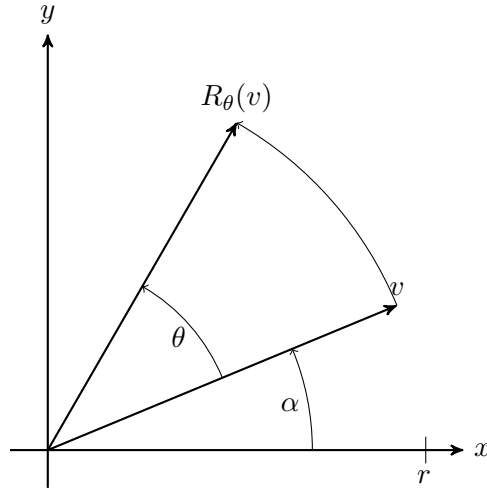


FIGURA 1. Rotación θ grados.

Otras transformaciones lineales importantes de \mathbb{R}^2 en \mathbb{R}^2 son

$$S_h(x, y) = (x, -y) \quad \text{y} \quad S_v(x, y) = (-x, y).$$

La primera es la reflexión en el eje x y la segunda la reflexión en el eje y . Las siguientes afirmaciones se comprueban algebraicamente en forma sencilla, pero nos podemos convencer de ellas por su interpretación geométrica:

$$\begin{aligned} R_\theta \circ R_\varphi &= R_{\theta+\varphi}, & (\text{rotar } \varphi \text{ y } \theta &= \text{rotar } \theta + \varphi) \\ R_{\pi/2} \circ S_h \circ R_{-\pi/2} &= S_v \end{aligned}$$

EJEMPLO 3.1.4. Sea $V = \mathbb{R}[x]$ el espacio vectorial de los polinomios con coeficientes reales. Definimos $D : V \rightarrow V$, por

$$D(P)(x) = P'(x), \quad x \in \mathbb{R}.$$

Observemos primero que la derivada de un polinomio es un polinomio, pues

$$(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0)' = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1.$$

Además D es lineal, pues $(f+g)' = f' + g'$ y $(\lambda f)' = \lambda f'$, para f, g funciones derivables y $\lambda \in \mathbb{R}$.

OBSERVACIÓN 3.1.3. Sean V y W dos espacios vectoriales sobre el cuerpo \mathbb{K} y $T : V \rightarrow W$ un transformación lineal. Entonces $T(0) = 0$

DEMOSTRACIÓN. $T(0) = T(0 + 0) = T(0) + T(0)$, por lo tanto

$$-T(0) + T(0) = -T(0) + T(0) + T(0) \Rightarrow 0 = 0 + T(0) \Rightarrow 0 = T(0).$$

□

OBSERVACIÓN 3.1.4. Las transformaciones lineales preservan combinaciones lineales, es decir si $T : V \rightarrow W$ es una transformación lineal, $v_1, \dots, v_k \in V$ y $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, entonces

$$T(\lambda_1 v_1 + \dots + \lambda_k v_k) = \lambda_1 T(v_1) + \dots + \lambda_k T(v_k).$$

Observar que el caso $k = 2$ se demuestra de la siguiente manera

$$T(\lambda_1 v_1 + \lambda_2 v_2) = T(\lambda_1 v_1) + T(\lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2).$$

El caso general se demuestra por inducción.

TEOREMA 3.1.1. Sean V un espacio vectorial de dimensión finita sobre el cuerpo \mathbb{K} y $\{v_1, \dots, v_n\}$ una base ordenada de V . Sean W un espacio vectorial sobre el mismo cuerpo y $\{w_1, \dots, w_n\}$, vectores cualesquiera de W . Entonces existe una única transformación lineal T de V en W tal que

$$T(v_j) = w_j, \quad j = 1, \dots, n.$$

DEMOSTRACIÓN. Recordemos que si $v \in V$, existen únicos $a_1, \dots, a_n \in \mathbb{K}$ (las coordenadas de v) tal que

$$v = a_1 v_1 + \dots + a_n v_n.$$

Luego para este vector v definimos

$$T(v) = a_1 w_1 + \dots + a_n w_n.$$

Entonces, T es una correspondencia bien definida que asocia a cada vector v de V un vector $T(v)$ de W . De la definición queda claro que $T(v_j) = w_j$ para cada j . Para ver que T es lineal, sea

$$w = b_1 v_1 + \dots + b_n v_n,$$

y sea $\lambda \in \mathbb{K}$. Ahora

$$\begin{aligned} \lambda v + w &= \lambda(a_1 v_1 + \dots + a_n v_n) + b_1 v_1 + \dots + b_n v_n \\ &= (\lambda a_1 + b_1) v_1 + \dots + (\lambda a_n + b_n) v_n \end{aligned}$$

con lo que, por definición

$$T(\lambda v + w) = (\lambda a_1 + b_1) w_1 + \dots + (\lambda a_n + b_n) w_n.$$

Por otra parte

$$\begin{aligned} \lambda T(v) + T(w) &= \lambda(a_1 w_1 + \dots + a_n w_n) + b_1 w_1 + \dots + b_n w_n \\ &= (\lambda a_1 + b_1) w_1 + \dots + (\lambda a_n + b_n) w_n, \end{aligned}$$

y así

$$T(\lambda v + w) = \lambda T(v) + T(w).$$

Finalmente, debemos probar la unicidad de T . Sea $S : V \rightarrow W$ transformación lineal tal que $S(v_j) = w_j$ para $1 \leq j \leq n$. Entonces, si $v \in V$ un vector arbitrario, $v = \sum_i a_i v_i$ y

$$S(v) = S\left(\sum_i a_i v_i\right) = \sum_i a_i S(v_i) = \sum_i a_i w_i = \sum_i a_i T(v_i) = T\left(\sum_i a_i v_i\right) = T(v)$$

□

El teorema 3.1.1 es muy elemental, pero por su importancia ha sido presentado detalladamente.

EJEMPLO 3.1.5. Usando el teorema 3.1.1, podemos demostrar la observación 3.1.2 de la siguiente manera: sea $\mathcal{C}_n = \{e_1, \dots, e_n\}$ es la base canónica de \mathbb{K}^n y sea $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ la única transformación lineal tal que

$$T(e_j) = (a_{1j}, \dots, a_{mj}), \quad j = 1, \dots, n$$

Entonces,

$$T(x_1, \dots, x_n) = (a_{11}x_1 + \dots + a_{1n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n).$$

es la transformación lineal resultante.

EJEMPLO 3.1.6. Los vectores

$$v_1 = (1, 2)$$

$$v_2 = (3, 4)$$

son linealmente independientes y, por tanto, forman una base de \mathbb{R}^2 . De acuerdo con el teorema 3.1.1, existe una única transformación lineal de \mathbb{R}^2 en \mathbb{R}^2 tal que

$$T(v_1) = (3, 2, 1)$$

$$T(v_2) = (6, 5, 4).$$

Para poder describir T respecto a las coordenadas canónicas debemos calcular $T(e_1)$ y $T(e_2)$, ahora bien,

$$(1, 0) = c_1(1, 2) + c_2(3, 4)$$

$$(0, 1) = c_3(1, 2) + c_4(3, 4)$$

y resolviendo este sistema de cuatro ecuaciones con cuatro incógnitas obtenemos

$$\begin{aligned} (1, 0) &= -2(1, 2) + (3, 4) \\ (0, 1) &= \frac{3}{2}(1, 2) - \frac{1}{2}(3, 4) \end{aligned}$$

Luego,

$$T(1, 0) = -2T(1, 2) + T(3, 4) = -2(3, 2, 1) + (6, 5, 4) = (0, 1, 2)$$

$$T(0, 1) = \frac{3}{2}T(1, 2) - \frac{1}{2}T(3, 4) = \frac{3}{2}(3, 2, 1) - \frac{1}{2}(6, 5, 4) = \left(\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}\right)$$

Entonces

$$T(x_1, x_2) = x_1(0, 1, 2) + x_2\left(\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}\right) = \left(\frac{3}{2}x_2, x_1 + \frac{1}{2}x_2, 2x_1 - \frac{1}{2}x_2\right)$$

3.2. Núcleo e imagen de una transformación lineal

DEFINICIÓN 3.2.1. Sean V, W espacios vectoriales sobre un cuerpo \mathbb{K} y sea $T : V \rightarrow W$ una transformación lineal. Definimos

$$\text{Im}(T) := \{w \in W : \text{existe } v \in V, \text{ tal que } T(v) = w\} = \{T(v) : v \in V\},$$

$$\text{Nu}(T) := \{v \in V : T(v) = 0\}.$$

A $\text{Im}(T)$ lo llamamos la *imagen* de T y a $\text{Nu}(T)$ el *núcleo* de T .

TEOREMA 3.2.1. Sean V, W espacios vectoriales sobre un cuerpo \mathbb{K} y sea $T : V \rightarrow W$ una transformación lineal; entonces $\text{Im}(T) \subset W$ y $\text{Nu}(T) \subset V$ son subespacios vectoriales.

DEMOSTRACIÓN. $\text{Im}(T) \neq \emptyset$, pues $0 = T(0) \in \text{Im}(T)$.

Si $T(v_1), T(v_2) \in \text{Im}(T)$ y $\lambda \in \mathbb{K}$, entonces $T(v_1) + T(v_2) = T(v_1 + v_2) \in \text{Im}(T)$ y $\lambda T(v_1) = T(\lambda v_1) \in \text{Im}(T)$.

$\text{Nu}(T) \neq \emptyset$ pues $T(0) = 0$ y por lo tanto $0 \in \text{Nu}(T)$.

Si $v, w \in V$ tales que $T(v) = 0$ y $T(w) = 0$, entonces, $T(v + w) = T(v) + T(w) = 0$. por lo tanto $v + w \in \text{Nu}(T)$. Si $\lambda \in \mathbb{K}$, entonces $T(\lambda v) = \lambda T(v) = \lambda \cdot 0 = 0$, luego $\lambda v \in \text{Nu}(T)$. \square

DEFINICIÓN 3.2.2. Sean V, W espacios vectoriales sobre un cuerpo \mathbb{K} y sea $T : V \rightarrow W$ una transformación lineal. Supongamos que V es de dimensión finita.

- (1) El *rango* de T es la dimensión de la imagen de T .
- (2) La *nulidad* de T es la dimensión del núcleo de T .

EJEMPLO 3.2.1. Sea $T : \mathbb{R}^3 \rightarrow \mathbb{R}$, definida

$$T(x, y, z) = x + 2y + 3z.$$

Encontrar una base del núcleo y de la imagen.

SOLUCIÓN. Es claro que como T no es 0, la imagen es todo \mathbb{R} (y por lo tanto cualquier $r \in \mathbb{R}, r \neq 0$ es base de la imagen).

Con respecto al núcleo, debemos encontrar una base del subespacio

$$\text{Nu}(T) = \{(x, y, z) : x + 2y + 3z = 0\}.$$

Como $x + 2y + 3z = 0 \Leftrightarrow x = -2y - 3z$, luego,

$$(37) \quad \text{Nu}(T) = \{(-2s - 3t, s, t) : s, t \in \mathbb{R}\}.$$

Ahora bien, $(-2s - 3t, s, t) = s(-2, 1, 0) + t(-3, 0, 1)$, por lo tanto

$$\text{Nu}(T) = \langle (-2, 1, 0), (-3, 0, 1) \rangle,$$

y como $(-2, 1, 0), (-3, 0, 1)$ son LI, tenemos que forman una base del núcleo. \square

La expresión (37), que depende de dos parámetros (s y t) que son independientes entre ellos, es llamada la *descripción paramétrica* del núcleo

EJEMPLO 3.2.2. Sea $T : \mathbb{R}^3 \rightarrow \mathbb{R}^4$, definida

$$T(x, y, z) = (x + y, x + 2y + z, 3y + 3z, 2x + 4y + 2z).$$

- (1) Describir $\text{Nu}(T)$ en forma paramétrica y dar una base.
- (2) Describir $\text{Im}(T)$ en forma paramétrica y dar una base.

SOLUCIÓN. Observemos que

$$\text{Nu}(T) = \{(x, y, z) \in \mathbb{R}^3 : (x + y, x + 2y + z, 3y + 3z, 2x + 4y + 2z) = (0, 0, 0, 0)\}.$$

Por lo tanto, describir el núcleo de T se reduce a encontrar las soluciones de una sistema de ecuaciones lineales homogéneo, y la matriz asociada a este sistema es.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 3 & 3 \\ 2 & 4 & 2 \end{bmatrix}$$

Por otro lado,

$$\text{Im}(T) = \{(y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : (x + y, x + 2y + z, 3y + 3z, 2x + 4y + 2z) = (y_1, y_2, y_3, y_4)\}$$

$$= \{(y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}\}.$$

Por lo tanto, y haciendo abuso de notación,

$$\text{Nu}(T) = \{v = (x, y, z) : A.v = 0\}$$

$$\text{Im}(T) = \{y = (y_1, y_2, y_3, y_4) : \text{tal que } \exists v \in \mathbb{R}^3, A.v = y\}$$

En ambos casos, la solución depende de resolver el sistema de ecuaciones cuya matriz asociada es A :

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & y_1 \\ 1 & 2 & 1 & y_2 \\ 0 & 3 & 3 & y_3 \\ 2 & 4 & 2 & y_4 \end{array} \right] \xrightarrow[F_4-2F_1]{F_2-F_1} \left[\begin{array}{ccc|c} 1 & 1 & 0 & y_1 \\ 0 & 1 & 1 & -y_1+y_2 \\ 0 & 3 & 3 & y_3 \\ 0 & 2 & 2 & -2y_1+y_4 \end{array} \right] \xrightarrow[F_4-2F_2]{F_1-F_2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 2y_1-y_2 \\ 0 & 1 & 1 & -y_1+y_2 \\ 0 & 0 & 0 & 3y_1-3y_2+y_3 \\ 0 & 0 & 0 & -2y_2+y_4 \end{array} \right],$$

es decir

$$(38) \quad T(x, y, z) = (y_1, y_2, y_3, y_4) \Leftrightarrow \begin{aligned} x - z &= 2y_1 - y_2 \\ y + z &= -y_1 + y_2 \\ 0 &= 3y_1 - 3y_2 + y_3 \\ 0 &= -2y_2 + y_4 \end{aligned}$$

Si hacemos $y_1 = y_2 = y_3 = y_4 = 0$, entonces las soluciones del sistema describen el núcleo de T , es decir

$$\text{Nu}(T) = \{(x, y, z) : x - z = 0, y + z = 0\} = \{(s, -s, s) : s \in \mathbb{R}\},$$

que es la forma paramétrica. Una base del núcleo de T es $(1, -1, 1)$.

Con respecto a la imagen de T tenemos

$$(y_1, y_2, y_3, y_4) \in \text{Im}(T) \Leftrightarrow \begin{aligned} \exists x, y, z \in \mathbb{R} \text{ tal que } 2y_1 - y_2 &= x - z \text{ y } -y_1 + y_2 = y + z; \text{ y} \\ 0 &= 3y_1 - 3y_2 + y_3 \text{ y } 0 = -2y_2 + y_4. \end{aligned}$$

Como variando los valores de x, y, z es posible obtener, para cualesquiera valores de y_1, y_2, y_3, y_4 , $2y_1 - y_2 = x - z$ y $-y_1 + y_2 = y + z$, tenemos

$$(39) \quad (y_1, y_2, y_3, y_4) \in \text{Im}(T) \Leftrightarrow 0 = 3y_1 - 3y_2 + y_3 \text{ y } 0 = -2y_2 + y_4.$$

Luego,

$$\text{Im}(T) = \{(y_1, y_2, y_3, y_4) : \text{tal que } 0 = 3y_1 - 3y_2 + y_3 \text{ y } 0 = -2y_2 + y_4\}.$$

Ahora bien,

$$0 = -2y_2 + y_4 \Leftrightarrow 2y_2 = y_4 \Leftrightarrow y_2 = \frac{1}{2}y_4.$$

Por otro lado

$$0 = 3y_1 - 3y_2 + y_3 \Leftrightarrow 3y_1 = 3y_2 - y_3 \Leftrightarrow y_1 = y_2 - \frac{1}{3}y_3 \Leftrightarrow y_1 = \frac{1}{2}y_4 - \frac{1}{3}y_3$$

Es decir, cambiando $y - 3$ por s e y_4 por t :

$$\text{Im}(T) = \{(-\frac{1}{3}s + \frac{1}{2}t, \frac{1}{2}t, s, t) : s, t \in \mathbb{R}\}.$$

Como $(-\frac{1}{3}s + \frac{1}{2}t, \frac{1}{2}t, s, t) = s(-\frac{1}{3}, 0, 1, 0) + t(\frac{1}{2}, \frac{1}{2}, 0, 1)$ el conjunto $\{(-\frac{1}{3}, 0, 1, 0), (\frac{1}{2}, \frac{1}{2}, 0, 1)\}$ es una base de $\text{Im}(T)$.

□

He aquí uno de los resultados más importantes del álgebra lineal.

TEOREMA 3.2.2. *Sean V, W espacios vectoriales sobre un cuerpo \mathbb{K} y sea $T : V \rightarrow W$ una transformación lineal. Supóngase que V es de dimensión finita. Entonces*

$$\text{rango}(T) + \text{nulidad}(T) = \dim V.$$

DEMOSTRACIÓN. Sean

$$n = \dim V$$

$$k = \dim(\text{Nu } T) = \text{nulidad}(T).$$

Entonces debemos probar que

$$n - k = \dim(\text{Im } T) = \text{rango}(T).$$

Sea $\{v_1, \dots, v_k\}$ una base de $\text{Nu } T$. Existen vectores $\{v_{k+1}, \dots, v_n\}$, en V tales que $\{v_1, \dots, v_n\}$ es una base de V . Para probar el teorema, demostraremos que $\{Tv_{k+1}, \dots, Tv_n\}$ es una base para la imagen de T .

(1) $\{Tv_{k+1}, \dots, Tv_n\}$ genera la imagen de T .

Si $w \in \text{Im}(T)$, entonces existe $v \in V$ tal que $T(v) = w$, como $\{v_1, \dots, v_n\}$ es base de V , existen $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, tal que $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, por lo tanto

$$\begin{aligned} w &= T(v) \\ &= \lambda_1 T(v_1) + \dots + \lambda_k T(v_k) + \lambda_{k+1} T(v_{k+1}) + \dots + \lambda_n T(v_n) \\ &= 0 + \dots + 0 + \lambda_{k+1} T(v_{k+1}) + \dots + \lambda_n T(v_n) \\ &= \lambda_{k+1} T(v_{k+1}) + \dots + \lambda_n T(v_n). \end{aligned}$$

Por lo tanto, $\{Tv_{k+1}, \dots, Tv_n\}$ genera la imagen de T .

(2) $\{Tv_{k+1}, \dots, Tv_n\}$ es un conjunto linealmente independiente.

Para ver que $\{Tv_{k+1}, \dots, Tv_n\}$ es linealmente independiente, supóngase que se tienen escalares μ_i tales que

$$\sum_{i=k+1}^n \mu_i T v_i = 0,$$

luego

$$0 = \sum_{i=k+1}^n \mu_i T v_i = T\left(\sum_{i=k+1}^n \mu_i v_i\right).$$

Por lo tanto $v = \sum_{i=k+1}^n \mu_i v_i \in \text{Nu}(T)$. Como $\{v_1, \dots, v_k\}$ es una base de $\text{Nu } T$, existen escalares λ_i tales que

$$v = \sum_{i=1}^k \lambda_i v_i,$$

es decir

$$\sum_{j=k+1}^n \mu_j v_j = \sum_{i=1}^k \lambda_i v_i.$$

Luego

$$\begin{aligned} 0 &= \sum_{i=1}^k \lambda_i v_i - \left(\sum_{j=k+1}^n \mu_j v_j \right) \\ &= \lambda_1 v_1 + \cdots + \lambda_k v_k - \mu_{k+1} v_{k+1} - \cdots - \mu_n v_n. \end{aligned}$$

Como $\{v_1, \dots, v_n\}$ es una base, y por lo tanto un conjunto LI, tenemos que $0 = \lambda_1 = \cdots = \lambda_k = \mu_{k+1} = \cdots = \mu_n$, y en particular $0 = \mu_{k+1} = \cdots = \mu_n$. Por lo tanto $\{Tv_{k+1}, \dots, Tv_n\}$ es un conjunto linealmente independiente. \square

Sea A una matriz $m \times n$ con coeficientes en \mathbb{K} . El *rango fila* de A es la dimensión del subespacio de \mathbb{K}^n generado por las filas de A , es decir la dimensión del espacio fila de A . El *rango columna* de A es la dimensión del subespacio de \mathbb{K}^m generado por las columnas de A . Un consecuencia importante del teorema 3.2.2 es el siguiente resultado.

TEOREMA 3.2.3. Si A es una matriz $m \times n$ con coeficientes en \mathbb{K} , entonces

$$\text{rango fila}(A) = \text{rango columna}(A).$$

DEMOSTRACIÓN. Sea T la transformación lineal

$$\begin{aligned} T: \mathbb{K}^{n \times 1} &\rightarrow \mathbb{K}^{m \times 1} \\ X &\mapsto AX. \end{aligned}$$

Observar que

$$\text{Nu}(T) = \{X \in \mathbb{K}^{n \times 1} : AX = 0\}.$$

Es decir $\text{Nu}(T)$ es el subespacio de soluciones del sistema homogéneo $AX = 0$. Ahora bien, si $k = \text{rango fila}(A)$, ya hemos dicho (capítulo 2, sección 2.5) que la dimensión del subespacio de soluciones del sistema homogéneo $AX = 0$ es $n - k$. Luego

$$(40) \quad \text{rango fila}(A) = \dim V - \text{nulidad}(T).$$

Por otro lado

$$\text{Im}(T) = \{AX : X \in \mathbb{K}^{n \times 1}\}.$$

Ahora bien,

$$AX = \begin{bmatrix} a_{11}x_1 + \cdots a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots a_{mn}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Es decir, que la imagen de T es el espacio generado por las columnas de A . Por tanto,

$$\text{rango}(T) = \text{rango columna}(A).$$

Por el teorema 3.2.2

$$\text{rango}(T) = \dim V - \text{nulidad}(T),$$

y por lo tanto

$$(41) \quad \text{rango columna } (A) = \dim V - \text{nulidad}(T).$$

Obviamente, las igualdades (40) y (41) implican

$$\text{rango fila } (A) = \text{rango columna } (A).$$

□

DEFINICIÓN 3.2.3. Si A es una matriz $m \times n$ con coeficientes en \mathbb{K} , entonces el *rango* de A es el rango fila de A (que es igual al rango columna).

3.3. Isomorfismos de espacios vectoriales

DEFINICIÓN 3.3.1. Sean V, W espacios vectoriales sobre un cuerpo \mathbb{K} y sea $T : V \rightarrow W$ una transformación lineal.

- (1) T es *epimorfismo* si T es suryectiva, es decir si $\text{Im}(T) = W$.
- (2) T es *monomorfismo* si T es inyectiva (o 1-1), es decir si dados $v_1, v_2 \in V$ tales que $T(v_1) = T(v_2)$, entonces $v_1 = v_2$.

OBSERVACIÓN 3.3.1. T es epimorfismo si y sólo si

$$T \text{ es lineal y } \forall w \in W, \exists v \in V \text{ tal que } T(v) = w.$$

Esto se deduce inmediatamente de la definiciones de función suryectiva y de $\text{Im}(T)$

T es monomorfismo si y sólo si

$$T \text{ es lineal y } \forall v_1, v_2 \in V : v_1 \neq v_2 \Rightarrow T(v_1) \neq T(v_2).$$

Esto se obtiene aplicando el contrarrecíproco a la definición de función inyectiva.

PROPOSICIÓN 3.3.1. Sea $T : V \rightarrow W$ una transformación lineal. Entonces T es monomorfismo si y sólo si $\text{Nu}(T) = 0$.

DEMOSTRACIÓN.

(\Rightarrow) Debemos ver que $\text{Nu}(T) = 0$, es decir que si $T(v) = 0$, entonces $v = 0$. Ahora bien, si $T(v) = 0$, como $T(0) = 0$, tenemos que $T(v) = T(0)$, y como T es inyectiva, implica que $v = 0$.

(\Leftarrow) Sean $v_1, v_2 \in V$ tal que $T(v_1) = T(v_2)$. Entonces

$$0 = T(v_1) - T(v_2) = T(v_1 - v_2).$$

Por lo tanto, $v_1 - v_2 \in \text{Nu}(T)$. Por hipótesis, tenemos que $v_1 - v_2 = 0$, es decir $v_1 = v_2$. □

OBSERVACIÓN 3.3.2. Sea $T : V \rightarrow W$ transformación lineal,

- (1) T es epimorfismo si y sólo si $\text{Im}(T) = W$ si y sólo si $\text{rango}(T) = \dim W$.
 (2) T es monomorfismo si y sólo si $\text{Nu}(T) = 0$ si y sólo si $\text{nulidad}(T) = 0$.

PROPOSICIÓN 3.3.2. *Sea $T : V \rightarrow W$ transformación lineal. Entonces,*

- (1) *T es monomorfismo si y sólo si T de un conjunto LI es LI.*
 (2) *T es epimorfismo si y sólo si T de un conjunto de generadores de V es un conjunto de generadores de W .*

DEMOSTRACIÓN.

(1) (\Rightarrow) Sea $\{v_1, \dots, v_n\}$ un conjunto LI en V y sean $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tales que

$$\lambda_1 T(v_1) + \dots + \lambda_n T(v_n) = 0,$$

entonces

$$0 = T(\lambda_1 v_1 + \dots + \lambda_n v_n).$$

Como T es inyectiva, por proposición 3.3.1,

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0,$$

lo cual implica que $\lambda_1, \dots, \lambda_n$ son todos nulos. Por lo tanto, $T(v_1), \dots, T(v_n)$ son LI.

(1) (\Leftarrow) Sea $v \in V$ tal que $T(v) = 0$. Veremos que eso implica que $v = 0$. Ahora bien, sea $\{v_1, \dots, v_n\}$ una base de V , entonces existen $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tales que

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n,$$

por lo tanto

$$0 = T(v) = T(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n).$$

Como $\{v_1, \dots, v_n\}$ es LI, por hipótesis, $\{T(v_1), \dots, T(v_n)\}$ es LI y, por lo tanto, $\lambda_1, \dots, \lambda_n$ son todos nulos. Luego $v = 0$. Es decir probamos que el núcleo de T es 0, luego por proposición 3.3.1, T es monomorfismo.

(1) $(\Leftarrow/\text{alternativa})$ Sea $v \in V$ tal que $T(v) = 0$. Si $v \neq 0$, entonces $\{v\}$ es un conjunto LI en V . Luego, $\{T(v)\}$ es un conjunto LI en W y por lo tanto $T(v) \neq 0$. Así, si $T(v) = 0$ entonces $v = 0$ y por lo tanto T es un monomorfismo.

(2) (\Rightarrow) Sea $\{v_1, \dots, v_n\}$ un conjunto de generadores de V y sea $w \in W$. Como T es epimorfismo, existe $v \in V$ tal que $T(v) = w$. Ahora bien,

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \text{ para algún } \lambda_1, \dots, \lambda_n \in \mathbb{K},$$

por lo tanto,

$$w = T(v) = T(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n).$$

Es decir, cualquier $w \in W$ se puede escribir como combinación lineal de los $T(v_1), \dots, T(v_n)$ y, por lo tanto, generan W .

(2) (\Leftarrow) Sea $\{v_1, \dots, v_n\}$ una base de V , por hipótesis $T(v_1), \dots, T(v_n)$ generan W , es decir dado cualquier $w \in W$, existen $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tales que

$$w = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n),$$

y por lo tanto $w = T(v)$, con

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

□

DEFINICIÓN 3.3.2. Si V y W son espacios vectoriales sobre el cuerpo \mathbb{K} , toda transformación lineal $T : V \rightarrow W$ suryectiva e inyectiva, se dice *isomorfismo* de V sobre W . Si existe un isomorfismo de V sobre W , se dice que V es *isomorfo* a W y se denota $V \cong W$.

Obsérvese que V es trivialmente isomorfo a V , ya que el operador identidad es un isomorfismo de V sobre V .

Recordemos que si una función $f : X \rightarrow Y$ es suryectiva e inyectiva, es decir biyectiva, existe su inversa, la cual también es biyectiva. La inversa se denota $f^{-1} : Y \rightarrow X$ y viene definida por

$$f^{-1}(y) = x \Leftrightarrow f(x) = y.$$

TEOREMA 3.3.3. Sea $T : V \rightarrow W$ un isomorfismo. Entonces $T^{-1} : W \rightarrow V$ es lineal y, por lo tanto, también es un isomorfismo.

DEMOSTRACIÓN.

(1) Sean $w_1, w_2 \in W$ y sean $v_1 = T^{-1}(w_1)$, $v_2 = T^{-1}(w_2)$. Por lo tanto $T(v_1) = w_1$ y $T(v_2) = w_2$. Ahora bien,

$$\begin{aligned} T^{-1}(w_1 + w_2) &= T^{-1}(T(v_1) + T(v_2)) = T^{-1}(T(v_1 + v_2)) = \\ &= (T^{-1} \circ T)(v_1 + v_2) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2). \end{aligned}$$

(2) Sean $w \in W$ y $\lambda \in \mathbb{K}$. Sea $v = T^{-1}(w)$, entonces

$$T^{-1}(\lambda w) = T^{-1}(\lambda T(v)) = T^{-1}(T(\lambda v)) = (T^{-1} \circ T)(\lambda v) = \lambda v = \lambda T^{-1}(w).$$

□

EJERCICIO 3.3.1. Sean V , W y Z espacios vectoriales sobre el cuerpo \mathbb{K} y sean $T : V \rightarrow W$, $S : W \rightarrow Z$ isomorfismos. Entonces,

(1) $S \circ T : V \rightarrow Z$ también es un isomorfismo y

(2) $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.

Como ya se ha dicho, V es isomorfo a V vía la identidad. Por el teorema anterior, si V es isomorfo a W , entonces W es isomorfo a V . Por el ejercicio anterior, si V es isomorfo a W y W es isomorfo a Z , entonces V es isomorfo a Z . En resumen, el isomorfismo es una relación de equivalencia sobre la clase de espacios vectoriales. Si existe un isomorfismo de V sobre W , se

dirá a veces que V y W son isomorfos, en vez de que V es isomorfo a W . Ello no será motivo de confusión porque V es isomorfo a W , si, y solo si, W es isomorfo a V .

TEOREMA 3.3.4. *Sean V, W espacios vectoriales de dimensión finita sobre \mathbb{K} tal que $\dim V = \dim W$. Sea $T : V \rightarrow W$ transformación lineal. Entonces, son equivalentes:*

- (a) T es un isomorfismo.
- (b) T es monomorfismo.
- (c) T es epimorfismo.
- (d) Si $\{v_1, \dots, v_n\}$ es una base de V , entonces $\{T(v_1), \dots, T(v_n)\}$ es una base de W .

DEMOSTRACIÓN. Sea $n = \dim V = \dim W$.

(a) \Rightarrow (b). Como T es isomorfismo, es biyectiva y por lo tanto inyectiva.

(b) \Rightarrow (c). T monomorfismo, entonces $\text{nulidad}(T) = 0$ (proposición 3.3.1). Luego, como $\text{rango}(T) + \text{nulidad}(T) = \dim V$, tenemos que $\text{rango}(T) = \dim V$. Como $\dim V = \dim W$, tenemos que $\dim \text{Im}(T) = \dim W$ y por lo tanto $\text{Im}(T) = W$. En consecuencia, T es suryectiva.

(c) \Rightarrow (a). T es suryectiva, entonces $\text{rango}(T) = n$, luego $\text{nulidad}(T) = 0$, por lo tanto $\text{Nu}(T) = 0$ y en consecuencia T es inyectiva. Como T es suryectiva e inyectiva es un isomorfismo.

Hasta aquí probamos que (a), (b) y (c) son equivalentes, luego si probamos que (a), (b) o (c) \Rightarrow (d) y que (d) \Rightarrow (a), (b) o (c), estaría probado el teorema.

(a) \Rightarrow (d). Sea $\{v_1, \dots, v_n\}$ una base de V , entonces $\{v_1, \dots, v_n\}$ es LI y genera V . Por proposición 3.3.2, tenemos que $\{T(v_1), \dots, T(v_n)\}$ es LI y genera W , por lo tanto $\{T(v_1), \dots, T(v_n)\}$ es una base de W .

(d) \Rightarrow (a). Como T de una base es una base, entonces T de un conjunto LI es un conjunto LI y T de un conjunto de generadores de V es un conjunto de generadores de W . Por lo tanto, por proposición 3.3.2, T es monomorfismo y epimorfismo, luego T es un isomorfismo. \square

COROLARIO 3.3.5. *Sean V, W espacios vectoriales de dimensión finita sobre \mathbb{K} tal que $\dim V = \dim W$. Entonces V y W son isomorfos.*

DEMOSTRACIÓN. Sea $\{v_1, \dots, v_n\}$ es una base de V y $\{w_1, \dots, w_n\}$ es una base de W . Por teorema 3.1.1 existe una única transformación lineal $T : V \rightarrow W$ tal que

$$T(v_i) = w_i, \quad i = 1, \dots, n.$$

Por el teorema anterior, T es un isomorfismo. \square

EJEMPLO 3.3.1. $\mathbb{K}_n[x] = \{a_0 + a_1x + \dots + a_{n-1}x^{n-1} : a_0, a_1, \dots, a_{n-1} \in \mathbb{K}\}$ es isomorfo a \mathbb{K}^n , esto es consecuencia inmediata del corolario anterior, pues ambos tienen dimensión n . Explícitamente, $1, x, \dots, x^{n-1}$ es base de $\mathbb{K}_n[x]$ y sea e_1, \dots, e_n la base canónica de \mathbb{K}^n , entonces

un isomorfismo de $\mathbb{K}_n[x]$ a \mathbb{K}^n viene dado por la única transformación lineal $T : \mathbb{K}_n[x] \rightarrow \mathbb{K}^n$ tal que

$$T(x^i) = e_{i+1}, \quad i = 0, \dots, n-1.$$

EJEMPLO 3.3.2. $M_{m \times n}(\mathbb{K})$ es isomorfo a \mathbb{K}^{mn} . El isomorfismo viene dado por $T : M_{m \times n}(\mathbb{K}) \rightarrow \mathbb{K}^{mn}$ tal que

$$T(E_{ij}) = e_{(i-1)n+j}, \quad i = 1, \dots, m, j = 1, \dots, n.$$

Por ejemplo, en el caso 2×2 ,

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &\mapsto (1, 0, 0, 0) & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} &\mapsto (0, 1, 0, 0) \\ \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} &\mapsto (0, 0, 1, 0) & \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} &\mapsto (0, 0, 0, 1). \end{aligned}$$

3.4. Álgebra de las transformaciones lineales

En el estudio de las transformaciones lineales de V en W es de fundamental importancia que el conjunto de estas transformaciones hereda una estructura natural de espacio vectorial. El conjunto de las transformaciones lineales de un espacio V en sí mismo tiene incluso una estructura algebraica mayor, pues la composición ordinaria de funciones da una “multiplicación” de tales transformaciones.

Observemos primero que si X conjunto y W espacio vectorial sobre el cuerpo \mathbb{K} , entonces

$$F(X, W) := \{f : X \rightarrow W\},$$

es decir el conjunto de funciones de X en W es un espacio vectorial sobre \mathbb{K} con la suma y el producto por escalares definido:

$$\begin{aligned} (f+g)(x) &= f(x) + g(x), & f, g &\in F(X, W), x \in X \\ (\lambda f)(x) &= \lambda f(x), & f &\in F(X, W), x \in X, \lambda \in \mathbb{K}. \end{aligned}$$

La demostración de esto es sencilla y se basa en el hecho que W es un espacio vectorial.

TEOREMA 3.4.1. Sean V y W espacios vectoriales sobre el cuerpo \mathbb{K} Sean $T, S : V \rightarrow W$ transformaciones y $\mu \in \mathbb{K}$. Entonces, $T + S$ y μT son transformaciones lineales de V en W .

DEMOSTRACIÓN. Sean $v, v' \in V$ y $\lambda \in \mathbb{K}$, entonces

$$\begin{aligned} (T + S)(\lambda v + v') &= T(\lambda v + v') + S(\lambda v + v') && \text{(definición de } T + S) \\ &= \lambda T(v) + T(v') + \lambda S(v) + S(v') && \text{(} T \text{ y } S \text{ lineales)} \\ &= \lambda(T(v) + S(v)) + T(v') + S(v') \\ &= \lambda((T + S)(v)) + (T + S)(v') && \text{(definición de } T + S) \\ &= \lambda(T + S)(v) + (T + S)(v') && \text{(definición de } \lambda(T + S)). \end{aligned}$$

que dice que $T + U$ es una transformación lineal. En forma análoga, si $\mu \in \mathbb{K}$,

$$\begin{aligned} (\mu T)(\lambda v + v') &= \mu T(\lambda v + v') && \text{(definición de } \mu T) \\ &= \mu \lambda T(v) + \mu T(v') && (T \text{ lineal}) \\ &= \lambda \mu T(v) + \mu T(v') \\ &= \lambda(\mu T)(v) + (\mu T)(v') && \text{(definición de } \mu T). \end{aligned}$$

que dice que μT es una transformación lineal. \square

COROLARIO 3.4.2. Sean V y W espacios vectoriales sobre el cuerpo \mathbb{K} . Entonces, el conjunto de transformaciones lineales de V en W es un subespacio vectorial de $F(V, W)$.

Se denotará $L(V, W)$ al espacio vectorial de las transformaciones lineales de V en W .

TEOREMA 3.4.3. Sean V , W y Z espacios vectoriales sobre el cuerpo \mathbb{K} . Sean $T : V \rightarrow W$ y $U : W \rightarrow Z$ transformaciones lineales. Entonces la función compuesta $U \circ T$ definida por $(U \circ T)(v) = U(T(v))$ es una transformación lineal de V en Z .

DEMOSTRACIÓN. Sean $v, v' \in V$ y $\lambda \in \mathbb{K}$, entonces

$$\begin{aligned} (U \circ T)(\lambda v + v') &= U(T(\lambda v + v')) && \text{(definición de composición)} \\ &= U(\lambda T(v) + T(v')) && (T \text{ lineal}) \\ &= \lambda U(T(v)) + U(T(v')) && (U \text{ lineal}) \\ &= \lambda(U \circ T)(v) + (U \circ T)(v') && \text{(definición de composición)}. \end{aligned}$$

\square

Para simplificar, a veces denotaremos la composición por yuxtaposición, es decir

$$U \circ T = UT.$$

En lo que sigue debemos interesarnos principalmente en transformaciones lineales de un espacio vectorial en sí mismo. Como se tendrá a menudo que escribir “ T es una transformación lineal de V en V ”, se dirá más bien: “ T es un operador lineal sobre V ”.

DEFINICIÓN 3.4.1. Si V es un espacio vectorial sobre el cuerpo \mathbb{K} , un *operador lineal* sobre V es una transformación lineal de V en V .

Cuando en el teorema 3.4.3, consideramos $V = W = Z$, tenemos que U y T son operadores lineales en el espacio V , y por lo tanto la composición UT es también un operador lineal sobre V . Así, el espacio $L(V, V)$ tiene una “multiplicación” definida por composición. En este caso el operador TU también está definido, y debe observarse que en general $UT \neq TU$, es decir, $UT - TU \neq 0$. Se ha de advertir de manera especial que si T es un operador lineal sobre V , entonces se puede componer T con T . Se usará para ello la notación $T^2 = TT$, y en general $T^n = T \cdots T$ (n veces) para $n = 1, 2, 3, \dots$. Si $T \neq 0$, se define $T^0 = I_V$, el operador identidad.

LEMA 3.4.4. Sea V un espacio vectorial sobre el cuerpo \mathbb{K} ; sean U , T y S operadores lineales sobre V y sea λ un elemento de \mathbb{K} . Denotemos I_V el operador identidad. Entonces

- (1) $U = I_V U = U I_V$,
- (2) $U(T + S) = UT + US$, $(T + S)U = TU + SU$,
- (3) $\lambda(UT) = (\lambda U)T = U(\lambda T)$.

i

DEMOSTRACIÓN. (1) es trivial.

Demostremos $U(T + S) = UT + US$ de (2) y todo lo demás se dejará como ejercicio. Sea $v \in V$, entonces

$$\begin{aligned}
 U(T + S)(v) &= U((T + S)(v)) && \text{(definición de composición)} \\
 &= U(T(v) + S(v)) && \text{(definición de } T + S) \\
 &= U(T(v)) + U(S(v)) && (U \text{ lineal}) \\
 &= UT(v) + US(v) && \text{(definición de composición)}.
 \end{aligned}$$

□

El contenido de este lema, y algunos otros resultados sobre composición de funciones de un conjunto en si mismo (como ser la asociatividad), dicen que el espacio vectorial $L(V, V)$, junto con la operación de composición, es lo que se conoce como una álgebra asociativa sobre \mathbb{K} , con identidad (ver https://es.wikipedia.org/wiki/Álgebra_asociativa).

3.5. Matriz de una transformación lineal

Sea V un espacio vectorial de dimensión n sobre el cuerpo \mathbb{K} , y sea W un espacio vectorial de dimensión m sobre \mathbb{K} . Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ordenada de V , y $\mathcal{B}' = \{w_1, \dots, w_m\}$ una base ordenada de W . Si T es cualquier transformación lineal de V en W , entonces T está determinada por su efecto sobre los vectores v_j , puesto que todo vector de V es combinación lineal de ellos. Cada uno de los n vectores Tv_j se expresa de manera única como combinación lineal

$$(42) \quad Tv_j = \sum_{i=1}^m a_{ij} w_i$$

de los w_i . Los escalares a_{1j}, \dots, a_{mj} son las coordenadas de Tv_j en la base ordenada \mathcal{B}' . Por consiguiente, la transformación T está determinada por los mn escalares a_{ij} mediante la expresión (42).

DEFINICIÓN 3.5.1. La matriz $m \times n$, A , definida por $[A]_{ij} = a_{ij}$, se llama *matriz de T respecto a las bases ordenadas \mathcal{B} y \mathcal{B}'* ; y se denota

$$[T]_{\mathcal{B}\mathcal{B}'} = A.$$

EJEMPLO 3.5.1. Sea $T : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ definida

$$T(x, y, z) = (2x + y, 3y, x + 4z, z).$$

Sean $B = \{e_1, e_2, e_3\}$ la base canónica de \mathbb{R}^3 y $B' = \{e_1, e_2, e_3, e_4\}$ la base canónica de \mathbb{R}^4 . Entonces

$$\begin{aligned} T(e_1) &= (2, 0, 1, 0) = 2e_1 + 0e_2 + e_3 + 0e_4 \\ T(e_2) &= (1, 3, 0, 0) = e_1 + 3e_2 + 0e_3 + 0e_4 \\ T(e_3) &= (0, 0, 4, 1) = 0e_1 + 0e_2 + 4e_3 + e_4 \end{aligned}$$

Por lo tanto

$$[T]_{\mathcal{B}\mathcal{B}'} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

Observar que si escribimos los vectores en coordenadas con respecto a las bases canónicas, tenemos que

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + y \\ 3y \\ x + 4z \\ z \end{bmatrix}$$

o más formalmente

$$[T]_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{B}'}.$$

PROPOSICIÓN 3.5.1. Sea V y W un espacios vectoriales de dimensión n y m respectivamente y sea $T : V \rightarrow W$ una transformación lineal. Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ordenada de V , y $\mathcal{B}' = \{w_1, \dots, w_m\}$ una base ordenada de W . Entonces

$$(43) \quad [T]_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{B}'}, \quad \forall v \in V.$$

DEMOSTRACIÓN. Si

$$Tv_j = \sum_{i=1}^m a_{ij}w_i$$

entonces $[T]_{ij} = a_{ij}$. Sea $v \in V$, entonces $v = x_1v_1 + \dots + x_nv_n$ con $x_i \in \mathbb{K}$, por lo tanto

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Ahora bien,

$$\begin{aligned} T(v) &= T\left(\sum_{j=1}^n x_j v_j\right) = \sum_{j=1}^n x_j T(v_j) = \sum_{j=1}^n \sum_{i=1}^m a_{ij} w_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}\right) w_i = \\ &= \left(\sum_{j=1}^n a_{1j}\right) w_1 + \left(\sum_{j=1}^n a_{2j}\right) w_2 + \cdots + \left(\sum_{j=1}^n a_{mj}\right) w_m \end{aligned}$$

y, por lo tanto,

$$(44) \quad [T(v)]_{\mathcal{B}'} = \begin{bmatrix} \sum_{j=1}^n a_{1j} \\ \sum_{j=1}^n a_{2j} \\ \vdots \\ \sum_{j=1}^n a_{mj} \end{bmatrix}.$$

Por otro lado,

$$(45) \quad [T]_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j \\ \sum_{j=1}^n a_{2j} x_j \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j \end{bmatrix}.$$

De las ecuaciones (44) y (45) se deduce la formula (43). \square

TEOREMA 3.5.2. *Sea V y W un espacios vectoriales de dimensión n y m respectivamente y $\mathcal{B} = \{v_1, \dots, v_n\}$ y $\mathcal{B}' = \{w_1, \dots, w_m\}$ dos bases ordenadas de V y W respectivamente. Entonces*

$$\kappa : L(V, W) \rightarrow M_{m \times n}(\mathbb{K})$$

definida

$$T \mapsto [T]_{\mathcal{B}\mathcal{B}'},$$

es un isomorfismos de espacios vectoriales.

DEMOSTRACIÓN. Primero probaremos que κ es lineal y luego que tiene inversa.

Sean $T, T' \in L(V, W)$ y $\lambda \in \mathbb{K}$, veamos que $\kappa(\lambda T + T') = \lambda \kappa(T) + \kappa(T')$, es decir

$$(46) \quad [\lambda T + T']_{\mathcal{B}\mathcal{B}'} = \lambda [T]_{\mathcal{B}\mathcal{B}'} + [T']_{\mathcal{B}\mathcal{B}'}.$$

Para $1 \leq j \leq n$, sean

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{y} \quad T'(v_j) = \sum_{i=1}^m a'_{ij} w_i,$$

es decir

$$[T]_{\mathcal{B}\mathcal{B}'} = [a_{ij}] \quad \text{y} \quad [T']_{\mathcal{B}\mathcal{B}'} = [a'_{ij}],$$

entonces

$$\begin{aligned}
 (\lambda T + T')(v_j) &= \lambda T(v_j) + T'(v_j) \\
 &= \lambda \sum_{i=1}^m a_{ij} w_i + \sum_{i=1}^m a'_{ij} w_i \\
 &= \sum_{i=1}^m (\lambda a_{ij} + a'_{ij}) w_i,
 \end{aligned}$$

por lo tanto

$$[\lambda T + T']_{\mathcal{B}\mathcal{B}'} = [\lambda a_{ij} + a'_{ij}] = \lambda [T]_{\mathcal{B}\mathcal{B}'} + [T']_{\mathcal{B}\mathcal{B}'}$$

y hemos probado (46) y, en consecuencia, κ es lineal.

Definamos ahora la inversa de κ : sea $A = [a_{ij}]$ matriz $m \times n$ y sea $T : V \rightarrow W$ la única transformación lineal que satisface, para $1 \leq j \leq n$, que

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i.$$

Es claro que esta aplicación tiene dominio en $M_{m \times n}(\mathbb{K})$ y su imagen está contenida en $L(V, W)$. Más aún, es muy sencillo comprobar que es la aplicación inversa a κ .

□

TEOREMA 3.5.3. Sean V, W y Z espacios vectoriales de dimensión finita sobre el cuerpo \mathbb{K} ; sean $T : V \rightarrow W$ y $U : W \rightarrow Z$ transformaciones lineales. Si $\mathcal{B}, \mathcal{B}'$ y \mathcal{B}'' son bases ordenadas de los espacios V, W y Z , respectivamente, entonces

$$(47) \quad [UT]_{\mathcal{B}\mathcal{B}''} = [U]_{\mathcal{B}'\mathcal{B}''} [T]_{\mathcal{B}\mathcal{B}'}.$$

DEMOSTRACIÓN. Sean

$$\mathcal{B} = \{v_1, \dots, v_n\}, \quad \mathcal{B}' = \{w_1, \dots, w_m\}, \quad \mathcal{B}'' = \{z_1, \dots, z_l\}$$

y

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i, \quad 1 \leq j \leq n; \quad U(w_i) = \sum_{k=1}^l b_{ki} z_k, \quad 1 \leq i \leq m.$$

Es decir

$$[T]_{\mathcal{B}\mathcal{B}'} = [a_{ij}] \quad \text{y} \quad [U]_{\mathcal{B}'\mathcal{B}''} = [b_{ij}].$$

Entonces

$$\begin{aligned}
 (UT)(v_j) &= U\left(\sum_{i=1}^m a_{ij}w_i\right) \\
 &= \sum_{i=1}^m a_{ij}U(w_i) \\
 &= \sum_{i=1}^m a_{ij} \sum_{k=1}^l b_{ki}z_k \\
 &= \sum_{k=1}^l \left(\sum_{i=1}^m b_{ki}a_{ij}\right)z_k.
 \end{aligned}$$

Luego el coeficiente kj de la matriz $[UT]_{\mathcal{B}\mathcal{B}'}$ es $\sum_{i=1}^m b_{ki}a_{ij}$ que es igual a la fila k de $[U]_{\mathcal{B}'\mathcal{B}'}$ por la columna j de $[T]_{\mathcal{B}\mathcal{B}'}$, en símbolos, si $A = [T]_{\mathcal{B}\mathcal{B}'}$, $B = [U]_{\mathcal{B}'\mathcal{B}'}$ y $C = [UT]_{\mathcal{B}\mathcal{B}'}$, entonces

$$[C]_{kj} = \sum_{i=1}^m b_{ki}a_{ij} = F_k(B)C_j(A) = [BA]_{kj}.$$

□

COROLARIO 3.5.4. Sean V espacio vectorial de dimensión finita, $\mathcal{B} = \{v_1, \dots, v_n\}$ base ordenada de V y $T, U : V \rightarrow V$ operadores lineales. Entonces

- (1) $[UT]_{\mathcal{B}} = [U]_{\mathcal{B}}[T]_{\mathcal{B}}$.
- (2) Si T es inversible, entonces $[T]_{\mathcal{B}}$ es una matriz inversible y

$$[T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}.$$

DEMOSTRACIÓN. (1). Es inmediato del teorema anterior tomado $\mathcal{B}' = \mathcal{B}'' = \mathcal{B}$.

(2). Denotemos I al operador identidad de V , entonces I se puede escribir $I = TT^{-1}$, luego

$$I_n = [I]_{\mathcal{B}} = [TT^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}[T^{-1}]_{\mathcal{B}}$$

y análogamente, $I = T^{-1}T$, luego

$$I_n = [I]_{\mathcal{B}} = [T^{-1}T]_{\mathcal{B}} = [T^{-1}]_{\mathcal{B}}[T]_{\mathcal{B}}.$$

Por lo tanto $[T]_{\mathcal{B}}^{-1} = [T^{-1}]_{\mathcal{B}}$.

□

TEOREMA 3.5.5. Sea V un espacio vectorial de dimensión finita sobre el cuerpo \mathbb{K} y sean

$$\mathcal{B} = \{v_1, \dots, v_n\}, \quad \mathcal{B}' = \{v'_1, \dots, v'_n\}$$

bases ordenadas de V . Supóngase que T es un operador lineal sobre V . Si $P = [P_1, \dots, P_n]$ es la matriz $n \times n$ de columnas $P_j = [v'_j]_{\mathcal{B}}$, entonces

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P.$$

De otra manera, si U es el operador lineal sobre V definido por $U(v_j) = v'_j$, $j = 1, \dots, n$, entonces

$$[T]_{\mathcal{B}'} = [U]_{\mathcal{B}}^{-1} [T]_{\mathcal{B}} [U]_{\mathcal{B}}.$$

DEMOSTRACIÓN. Denotemos I al operador identidad de V , entonces $T = IT$ y $T = TI$ y,

$$\begin{aligned} [T]_{\mathcal{B}'\mathcal{B}'} &= [IT]_{\mathcal{B}'\mathcal{B}'} \\ &= [I]_{\mathcal{B}\mathcal{B}'} [T]_{\mathcal{B}'\mathcal{B}} && \text{(teorema 3.5.3)} \\ &= [I]_{\mathcal{B}\mathcal{B}'} [TI]_{\mathcal{B}'\mathcal{B}} \\ &= [I]_{\mathcal{B}\mathcal{B}'} [T]_{\mathcal{B}\mathcal{B}} [I]_{\mathcal{B}'\mathcal{B}} && \text{(teorema 3.5.3),} \end{aligned}$$

Luego hemos probado

$$(48) \quad [T]_{\mathcal{B}'} = [I]_{\mathcal{B}\mathcal{B}'} [T]_{\mathcal{B}} [I]_{\mathcal{B}'\mathcal{B}}.$$

Solo resta comprobar que

$$(*) \quad [I]_{\mathcal{B}'\mathcal{B}} = [U]_{\mathcal{B}} \quad \text{y} \quad (**) \quad [I]_{\mathcal{B}\mathcal{B}'} = [U]_{\mathcal{B}}^{-1}.$$

Demostración de ().* Como \mathcal{B} es una base de V , entonces $v'_j = \sum_i a_{ij} v_i$, para $1 \leq j \leq n$, por lo tanto

$$I(v'_j) = v'_j = \sum_i a_{ij} v_i,$$

luego $[I]_{\mathcal{B}'\mathcal{B}} = [a_{ij}]$. Por otro lado,

$$U(v_j) = v'_j = \sum_i a_{ij} v_i,$$

luego $[U]_{\mathcal{B}} = [a_{ij}]$, lo que implica que $[I]_{\mathcal{B}'\mathcal{B}} = [U]_{\mathcal{B}}$.

*Demostración de (**).* Como la matriz identidad $I_{\mathcal{B}}$ e \mathcal{B}' , de nuevo aplicando teorema 3.5.3 tenemos

$$\begin{aligned} I_n &= I_{\mathcal{B}\mathcal{B}} = [I]_{\mathcal{B}'\mathcal{B}} [I]_{\mathcal{B}\mathcal{B}'}, \\ I_n &= I_{\mathcal{B}'\mathcal{B}'} = [I]_{\mathcal{B}\mathcal{B}'} [I]_{\mathcal{B}'\mathcal{B}}, \end{aligned}$$

luego $[I]_{\mathcal{B}\mathcal{B}'} = ([I]_{\mathcal{B}'\mathcal{B}})^{-1} = [U]_{\mathcal{B}}^{-1}$. □

La matriz P del teorema anterior es llamada la *matriz de cambio de base*. La fórmula (48) es importante por si misma y como $P = [I]_{\mathcal{B}'\mathcal{B}}$ nos da un método alternativo (muy similar) para calcular la matriz de cambio de base.

El teorema anterior nos permite definir el determinante de un operador lineal. Sea V un espacio vectorial de dimensión finita sobre el cuerpo \mathbb{K} y T un operador lineal sobre V . Sean \mathcal{B} , \mathcal{B}' bases ordenadas de V , entonces $[T]_{\mathcal{B}'} = P^{-1} [T]_{\mathcal{B}} P$, para P una matriz inversible. Por lo tanto,

$$\det([T]_{\mathcal{B}'}) = \det(P^{-1} [T]_{\mathcal{B}} P) = \det([T]_{\mathcal{B}} P P^{-1}) = \det([T]_{\mathcal{B}}).$$

Es decir, el determinante de la matriz de T en cualquier base siempre es igual.

DEFINICIÓN 3.5.2. Sea V un espacio vectorial de dimensión finita sobre el cuerpo \mathbb{K} y T un operador lineal sobre V . El *determinante de T* es el determinante de la matriz de T en alguna base de V .

3.6. Autovalores y autovectores

Sea V espacio vectorial de dimensión finita. Un operador lineal en V es *diagonalizable* si existe una base ordenada $\mathcal{B} = \{v_1, \dots, v_n\}$ de V y $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tal que

$$(49) \quad T(v_i) = \lambda_i v_i, \quad 1 \leq i \leq n.$$

En general, los operadores diagonalizables permiten hacer cálculos sobre ellos en forma sencilla, por ejemplo el núcleo del operador definido por (49) es $\text{Nu}(T) = \langle v_i : \lambda_i = 0 \rangle$ y su imagen es $\text{Im}(T) = \langle v_i : \lambda_i \neq 0 \rangle$ (veremos la demostración de estos resultados más adelante). Otra propiedad importante de los operadores diagonalizables es que la matriz de la transformación lineal en una base adecuada es diagonal (de allí viene el nombre de diagonalizable). En el caso del operador definido por (49) tenemos que

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

No todo operador lineal es diagonalizable y no es inmediato, ni sencillo, de la definición de un operador lineal decidir si es diagonalizable o no. En esta sección veremos herramientas para estudiar un operador lineal T y su posible diagonalización. La ecuación (49) sugiere se estudien los vectores que son transformados por T en múltiplos de sí mismos.

DEFINICIÓN 3.6.1. Sea V un espacio vectorial sobre el cuerpo \mathbb{K} y sea T un operador lineal sobre V . Un *valor propio* o *autovalor* de T es un escalar λ de \mathbb{K} tal que existe un vector no nulo $v \in V$ con $T(v) = \lambda v$. Si λ es un autovalor de T , entonces

- (1) cualquier $v \in V$ tal que $T(v) = \lambda v$ se llama un *vector propio* o *autovector* de T asociado al valor propio λ ;
- (2) la colección de todos los $v \in V$ tal que $T(v) = \lambda v$ se llama *espacio propio* o *autoespacio* asociado a λ .

Los valores propios se llaman también a menudo raíces características, eigenvalores, valores característicos o valores espectrales. Nosotros usaremos, preferentemente, “autovalores”.

Sea ahora $\lambda \in \mathbb{K}$, definimos

$$V_{\lambda} := \{v \in V : T v = \lambda v\}.$$

Observar que $V_\lambda \neq 0$ si y sólo si λ es autovalor.

TEOREMA 3.6.1. *Sea V un espacio vectorial y sea $T : V \rightarrow V$ una aplicación lineal. Sea $\lambda \in \mathbb{K}$ entonces, V_λ es subespacio de V .*

DEMOSTRACIÓN. Sean $v_1, v_2 \in V$ tales que $Tv_1 = \lambda v_1$ y $Tv_2 = \lambda v_2$. Entonces

$$T(v_1 + v_2) = T(v_1) + T(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2),$$

es decir si $v_1, v_2 \in V_\lambda$, probamos que $v_1 + v_2 \in V_\lambda$.

Sea ahora $c \in F$, entonces $T(cv_1) = cT(v_1) = c\lambda v_1 = \lambda(cv_1)$. Por lo tanto, si $v_1 \in V_\lambda$ y $c \in F$, probamos que $cv_1 \in V_\lambda$.

Esto termina de probar el teorema. \square

TEOREMA 3.6.2. *Sea V espacio vectorial y sea $T : V \rightarrow V$ una aplicación lineal. Sean v_1, \dots, v_m autovectores de T , con autovalores $\lambda_1, \dots, \lambda_m$ respectivamente. Suponga que estos autovalores son distintos entre si, esto es, $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces v_1, \dots, v_m son linealmente independientes.*

DEMOSTRACIÓN. Hagamos la demostración por inducción sobre m .

Caso base. Si $m = 1$, no hay nada que demostrar puesto que un vector no nulo es LI.

Paso inductivo. Supongamos que el enunciado es verdadero para el caso $m - 1$ con $m > 1$, (hipótesis inductiva o HI), y probemos entonces que esto implica que es cierto para m . Debemos ver que si

$$(*) \quad c_1 v_1 + c_2 v_2 + \dots + c_m v_m = 0$$

entonces $c_1 = \dots = c_m = 0$. Multipliquemos $(*)$ por λ_1 , obtenemos:

$$(**) \quad c_1 \lambda_1 v_1 + c_2 \lambda_1 v_2 + \dots + c_m \lambda_1 v_m = 0.$$

También apliquemos T a $(*)$ y obtenemos

$$(***) \quad c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \dots + c_m \lambda_m v_m = 0.$$

Ahora a $(**)$ le restamos $(***)$ y obtenemos:

$$(50) \quad c_2(\lambda_1 - \lambda_2)v_2 + \dots + c_m(\lambda_1 - \lambda_m)v_m = 0.$$

Como, por hipótesis inductiva, v_2, \dots, v_m son LI, tenemos que $c_i(\lambda_1 - \lambda_i) = 0$ para $i \geq 2$. Como $\lambda_1 - \lambda_i \neq 0$ para $i \geq 2$, obtenemos que $c_i = 0$ para $i \geq 2$. Por $(*)$ eso implica que $c_1 = 0$ y por lo tanto $c_i = 0$ para todo i . \square

COROLARIO 3.6.3. *Sea V espacio vectorial de dimensión n y sea $T : V \rightarrow V$ una aplicación lineal que tiene n autovectores v_1, \dots, v_n cuyos autovalores $\lambda_1, \dots, \lambda_n$ son distintos entre si. Entonces $\{v_1, \dots, v_n\}$ es una base de V .*

Recordemos que si T es una transformación lineal, el determinante de T se define como el determinante de la matriz de la transformación lineal en una base dada y que este determinante no depende de la base.

DEFINICIÓN 3.6.2. Sea $A \in M_n(\mathbb{K})$ el *polinomio característico* de A es $\chi_A(x) = \det(A - xI)$, donde x es una indeterminada.

Sea V espacio vectorial de dimensión finita y sea $T : V \rightarrow V$ lineal, el *polinomio característico* de T es $\chi_T(x) = \det(T - xId)$.

Si $\dim V = n$, entonces el polinomio característico de un operador T es un polinomio de grado n , más precisamente

$$\chi_T(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

Esto se puede demostrar fácilmente por inducción.

EJEMPLO 3.6.1. Sea $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ y su matriz en la base canónica es

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

entonces

$$\det \begin{bmatrix} a-x & b \\ c & d-x \end{bmatrix} = (a-x)(d-x) - bc = x^2 - (a+d)x + (ad-bc).$$

Es decir,

$$\chi_T(x) = x^2 - (a+d)x + (ad-bc).$$

En general, si $A = [a_{ij}]$ matriz $n \times n$, tenemos que

$$(51) \quad \chi_A(x) = \det(A - xI_n) = \det \begin{bmatrix} a_{11}-x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-x \end{bmatrix}$$

EJEMPLO 3.6.2. Sea

$$A = \begin{bmatrix} 10 & -10 & 6 \\ 8 & -8 & 6 \\ -5 & 5 & -3 \end{bmatrix},$$

entonces el polinomio característico de A es

$$\det \begin{bmatrix} 10-x & -10 & 6 \\ 8 & -8-x & 6 \\ -5 & 5 & -3-x \end{bmatrix} = -x^3 - x^2 + 6x.$$

Es posible factorizar esta expresión y obtenemos

$$\chi_A(x) = -x(x-2)(x+3).$$

PROPOSICIÓN 3.6.4. *Sea V espacio vectorial de dimensión n y sea $T : V \rightarrow V$ lineal. Entonces $\lambda \in \mathbb{K}$ es autovalor si y sólo si λ es raíz del polinomio característico.*

DEMOSTRACIÓN.

(\Rightarrow) Si λ es autovalor, entonces existe $v \in V$, no nulo, tal que $Tv = \lambda v$, luego

$$0 = Tv - \lambda v = Tv - \lambda Id v = (T - \lambda Id)v.$$

Por lo tanto, $T - \lambda Id$ no es inversible, lo cual implica que $0 = \det(T - \lambda Id) = \chi_T(\lambda)$. Es decir, λ es raíz del polinomio característico.

(\Leftarrow) Si λ es raíz del polinomio característico, es decir si $0 = \chi_T(\lambda) = \det(T - \lambda Id)$, entonces $T - \lambda Id$ no es una matriz inversible, por lo tanto su núcleo es no trivial. Es decir existe $v \in V$ tal que $(T - \lambda Id)v = 0$, luego $Tv = \lambda v$, por lo tanto v es autovector con autovalor λ . \square

Repetimos ahora algunos conceptos ya expresados al comienzo de la sección.

DEFINICIÓN 3.6.3. Sea V espacio vectorial de dimensión finita y sea $T : V \rightarrow V$ lineal. Diremos que T es *diagonalizable* si existe una base de V de autovectores de T .

En el caso que T sea una transformación lineal diagonalizable y $\mathcal{B} = \{v_1, \dots, v_n\}$ sea una base de autovectores con autovalores $\lambda_1, \dots, \lambda_n$, entonces

$$T(v_i) = \lambda_i v_i, \quad 1 \leq i \leq n,$$

y, por lo tanto, la matriz de T en la base \mathcal{B} es diagonal, más precisamente

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

EJEMPLO 3.6.3. Consideremos la transformación lineal de \mathbb{R}^3 en \mathbb{R}^3 definida por la matriz A del ejemplo 3.6.2, es decir (con abuso de notación incluido)

$$\begin{bmatrix} 10 & -10 & 6 \\ 8 & -8 & 6 \\ -5 & 5 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 10x - 10y + 6z \\ 8x - 8y + 6z \\ -5x + 5y - 3z \end{bmatrix}.$$

Ya vimos que el polinomio característico de esta aplicación es

$$\chi_A(x) = -x(x-2)(x+3).$$

Luego, por proposición 3.6.4, los autovalores de A son 0, 2 y -3 . Debido al corolario 3.6.3 existe una base de autovectores de A . Veamos cuales son. Si λ autovalor de A , para encontrar los autovectores con autovalor λ debemos resolver la ecuación $Av - \lambda v = 0$, en este caso sería

$$\begin{bmatrix} 10 - \lambda & -10 & 6 \\ 8 & -8 - \lambda & 6 \\ -5 & 5 & -3 - \lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

para $\lambda = 0, 2, -3$. Resolviendo estos tres sistemas, obtenemos que

$$V_0 = \{(y, y, 0) : y \in \mathbb{R}\}, \quad V_2 = \{(-2z, -z, z) : z \in \mathbb{R}\}, \quad V_{-3} = \{(-2z, -2z, z) : z \in \mathbb{R}\}.$$

Por lo tanto, $\{(1, 1, 0), (-2, -1, 1), (-2, -2, 1)\}$ es una base de autovectores de la transformación lineal.

PROPOSICIÓN 3.6.5. *Sea V espacio vectorial de dimensión n y sea $T : V \rightarrow V$ lineal tal que tiene una base de autovectores $\mathcal{B} = \{v_1, \dots, v_n\}$ con autovalores $\lambda_1, \dots, \lambda_n$. Entonces $\text{Nu}(T) = \langle v_i : \lambda_i = 0 \rangle$ e $\text{Im}(T) = \langle v_i : \lambda_i \neq 0 \rangle$.*

DEMOSTRACIÓN. Reordenemos la base de tal forma que $\lambda_i = 0$ para $1 \leq i \leq k$ y $\lambda_i \neq 0$ para $k < i \leq n$. Todo $v \in V$ se escribe en términos de la base como

$$v = x_1 v_1 + \dots + x_k v_k + x_{k+1} v_{k+1} + \dots + x_n v_n, \quad (x_i \in \mathbb{K}),$$

y entonces

$$(52) \quad T(v) = \lambda_{k+1} x_{k+1} v_{k+1} + \dots + \lambda_n x_n v_n.$$

Luego, $T(v) = 0$ si y sólo si $x_{k+1} = \dots = x_n = 0$, y esto se cumple si y solo si $v = x_1 v_1 + \dots + x_k v_k$, es decir $v \in \langle v_i : \lambda_i = 0 \rangle$. También es claro por la ecuación (52) que

$$\begin{aligned} \text{Im}(T) &= \{\lambda_{k+1} x_{k+1} v_{k+1} + \dots + \lambda_n x_n v_n : x_i \in \mathbb{K}\} \\ &= \{\mu_{k+1} v_{k+1} + \dots + \mu_n v_n : \mu_i \in \mathbb{K}\} \\ &= \langle v_i : \lambda_i \neq 0 \rangle. \end{aligned}$$

□

EJEMPLO 3.6.4. Sea $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ el operador definido por $T(x, y) = (y, x)$. Probar que T es diagonalizable y encontrar una base de autovectores.

DEMOSTRACIÓN. Por la proposición 3.6.4, los autovalores de T son las raíces del polinomio característico, es decir las raíces de

$$\det \begin{bmatrix} -x & 1 \\ 1 & -x \end{bmatrix} x^2 - 1 = (x - 1)(x + 1).$$

Luego los autovalores son 1 y -1 . Para hallar un autovector con autovalor 1 debemos resolver la ecuación $T(x, y) = (x, y)$. Ahora bien,

$$(x, y) = T(x, y) = (y, x),$$

luego $x = y$ y claramente $(1, 1)$ es autovector con autovalor 1.

Por otro lado $T(x, y) = -(x, y)$, implica que $(y, x) = -(x, y)$, es decir $y = -x$ y claramente podemos elegir $(1, -1)$ como autovector con autovalor -1 .

Luego $\mathcal{B} = \{(1, 1), (1, -1)\}$ es una base de \mathbb{R}^2 de autovectores de T . \square

No todas las matrices son diagonalizables, como veremos en el ejemplo a continuación.

EJEMPLO 3.6.5. Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el operador definido por $T(x, y) = (2x - y, x + 4y)$. Probar que T tiene un único autovalor λ cuyo autoespacio $V_\lambda = \{v \in \mathbb{R}^2 : Tv = \lambda v\}$ es de dimensión 1.

DEMOSTRACIÓN. La matriz de T en la base canónica es

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}.$$

Por la proposición 3.6.4, los autovalores de T son las raíces del polinomio característico, es decir las raíces de

$$\det \begin{bmatrix} 2-x & -1 \\ 1 & 4-x \end{bmatrix} = (2-x)(4-x) + 1 = x^2 - 6x + 9 = (x-3)^2.$$

Es decir el único autovalor posible es 3.

Debemos ver para que valores $(x, y) \in \mathbb{R}^2$ se satisface la ecuación

$$T(x, y) = 3(x, y).$$

tiene solución. Esta ecuación es equivalente a

$$\begin{aligned} (2x - y, x + 4y) &= (3x, 3y) && \Rightarrow \\ 2x - y = 3x &, \quad x + 4y = 3y && \Rightarrow \\ -y = x &, \quad x = -y && \Rightarrow \\ y &= -x \end{aligned}$$

Luego $V_3 = \{(x, -x) : x \in \mathbb{R}\}$ que es de dimensión 1. \square

PROPOSICIÓN 3.6.6. Sea T un operador lineal diagonalizable sobre un espacio vectorial V de dimensión finita. Sean $\lambda_1, \dots, \lambda_k$ los valores propios distintos de T . Entonces, el polinomio característico de T es

$$\chi_T(x) = (-1)^n (x - \lambda_1)^{d_1} \dots (x - \lambda_k)^{d_k}$$

con

$$d_i = \dim V_{\lambda_i},$$

para $i = 1, \dots, k$.

DEMOSTRACIÓN. T es un operador lineal diagonalizable y $\lambda_1, \dots, \lambda_k$ los valores propios distintos de T . Entonces existe una base ordenada \mathcal{B} con respecto a la cual T está representado por una matriz diagonal; es decir, los elementos de la diagonal son los escalares λ_j cada uno de los cuales se repite un cierto número de veces. Más específicamente, si v_{j1}, \dots, v_{jd_j} son los vectores en \mathcal{B} con autovalor λ_j ($1 \leq j \leq k$), reordenamos la base de tal forma que primero estén los autovectores con autovalor λ_1 , a continuación los de autovalor λ_2 , etc.:

$$\mathcal{B} = \{v_{11}, \dots, v_{1d_1}, \dots, v_{k1}, \dots, v_{kd_k}\}.$$

Ahora bien, si $v \in V$, entonces

$$\begin{aligned} v &= x_1 v_{11} + \dots + x_{d_1} v_{1d_1} + \dots + x_n v_{n1} + \dots + x_{nd_n} v_{nd_n} \\ &= v_1 + v_2 + \dots + v_k \end{aligned}$$

con $v_i = x_i v_{i1} + \dots + x_{d_i} v_{id_i} \in V_{\lambda_i}$. Luego

$$(53) \quad T(v) = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k$$

Veamos que $V_{\lambda_i} = \langle v_{i1}, \dots, v_{id_i} \rangle$ para $1 \leq i \leq k$. Es claro que $\langle v_{i1}, \dots, v_{id_i} \rangle \subset V_{\lambda_i}$. Probemos ahora que, $V_{\lambda_i} \subset \langle v_{i1}, \dots, v_{id_i} \rangle$: si $v \in V_{\lambda_i}$, entonces $T(v)$ es como en (53) y, por lo tanto, si $v_j \neq 0$ para $j \neq i$ entonces $T(v) \neq \lambda_j v$, lo que contradice la hipótesis. Es decir $v = v_i \in \langle v_{i1}, \dots, v_{id_i} \rangle$. Hemos probado que $V_{\lambda_i} = \langle v_{i1}, \dots, v_{id_i} \rangle$ y como v_{i1}, \dots, v_{id_i} son LI, entonces $\dim V_{\lambda_i} = d_i$.

Por otro lado, la matriz de T en la base \mathcal{B} tiene la forma

$$\begin{bmatrix} \lambda_1 I_1 & 0 & \dots & 0 \\ 0 & \lambda_2 I_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n I_n \end{bmatrix}$$

donde I_j es la matriz identidad $d_j \times d_j$. Luego, el polinomio característico de T es el producto

$$(\lambda_1 - x)^{d_1} \dots (\lambda_k - x)^{d_k} = (-1)^n (x - \lambda_1)^{d_1} \dots (x - \lambda_k)^{d_k}.$$

□

EJEMPLO 3.6.6. Sea T un operador lineal sobre \mathbb{R}^3 representado en la base ordenada canónica por la matriz

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

El polinomio característico de A es

$$\chi_A(x) = \det \begin{bmatrix} 5-x & -6 & -6 \\ -1 & 4-x & 2 \\ 3 & -6 & -4-x \end{bmatrix} = -x^3 + 5x^2 - 8x + 4 = -(x-2)^2(x-1).$$

¿Cuáles son las dimensiones de los espacios de los vectores propios asociados con los dos valores propios? Se deben resolver las ecuaciones asociadas a las matrices

$$A - 2I = \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix}$$

y

$$A - I = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}.$$

Las soluciones de estos sistemas son los autoespacios de autovalor 2 y 1 respectivamente. En el primer caso,

$$\begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} \xrightarrow[\begin{smallmatrix} F_3+3F_2 \end{smallmatrix}]{\begin{smallmatrix} F_1+3F_2 \\ F_3+3F_2 \end{smallmatrix}} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Luego, la solución del sistema asociado a $A - 2I$ es

$$V_2 = \{(2y + 2z, y, z) : y, z \in \mathbb{R}\} = \langle (2, 1, 0), (2, 0, 1) \rangle$$

cuya dimensión es 2.

Por otro lado,

$$\begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix} \xrightarrow[\begin{smallmatrix} F_3+3F_2 \end{smallmatrix}]{\begin{smallmatrix} F_1+4F_2 \\ F_3+3F_2 \end{smallmatrix}} \begin{bmatrix} 0 & 6 & 2 \\ -1 & 3 & 2 \\ 0 & 3 & 1 \end{bmatrix} \xrightarrow[\begin{smallmatrix} F_2-F_3 \end{smallmatrix}]{\begin{smallmatrix} F_1-2F_3 \\ F_2-F_3 \end{smallmatrix}} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 3 & 1 \end{bmatrix}.$$

Luego, la solución del sistema asociado a $A - I$ es

$$V_1 = \{(z, -\frac{1}{3}z, z) : z \in \mathbb{R}\} = \langle (1, -\frac{1}{3}, 1) \rangle.$$

Entonces, una base de autovectores de T podría ser

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

y en esa base la matriz de la transformación lineal es

$$[T]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

APÉNDICE A

Determinante

En el apéndice se harán las demostraciones de los resultados correspondientes a la sección de determinantes (sección 1.6).

A.1. Determinantes

Lo primero que veremos será la demostración del teorema 1.6.4. Los tres resultados de ese teorema los demostraremos en forma separada: serán los teoremas A.1.1, A.1.2 y A.1.4.

TEOREMA A.1.1. *Sea $A \in M_n(\mathbb{K})$ y sea $c \in \mathbb{K}$ y B la matriz que se obtiene de A multiplicando la fila r por c , es decir $A \xrightarrow{cF_r} B$, entonces $\det B = c \det A$.*

DEMOSTRACIÓN. Si multiplicamos la fila r por c obtenemos

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \vdots \\ ca_{r1} & ca_{r2} & \cdots & ca_{rn} \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \cdots & a_{nn} \end{bmatrix}.$$

Observemos que al hacer el desarrollo por la primera columna obtenemos

$$|B| = \sum_{i=1}^{r-1} a_{i1} C_{1i}^B + ca_{r1} C_{r1}^B + \sum_{i=r+1}^n a_{i1} C_{1i}^B.$$

Ahora bien, si $i \neq r$, la matriz $B(i|1)$ es la matriz $A(i|1)$ con una fila multiplicada por c , luego $|B(i|1)| = c|A(i|1)|$ y, en consecuencia $C_{1i}^B = c C_{1i}^A$. Además, $B(r|1) = A(r|1)$, luego $C_{r1}^B = C_{r1}^A$. Por lo tanto, reemplazando en la ecuación anterior C_{1i}^B por $c C_{1i}^A$ si $i \neq r$ y C_{r1}^B por C_{r1}^A , obtenemos

$$|B| = \sum_{i=1}^{r-1} a_{i1} c C_{1i}^A + ca_{r1} C_{r1}^A + \sum_{i=r+1}^n a_{i1} c C_{1i}^A = c|A|.$$

□

TEOREMA A.1.2. *Sea $A \in M_n(\mathbb{K})$ y sean $1 \leq r, s \leq n$. Sea B la matriz que se obtiene de A permutando la fila r con la fila s , es decir $A \xrightarrow{F_r \leftrightarrow F_s} B$, entonces $\det B = -\det A$.*

DEMOSTRACIÓN. Primero probaremos el teorema bajo el supuesto de que la fila 1 es permutada con la fila k , para $k > 1$. Esto será suficiente para probar el teorema, puesto que intercambiar las filas k y k_0 es equivalente a realizar tres permutaciones de filas: primero intercambiamos las filas 1 y k , luego las filas 1 y k_0 , y finalmente intercambiando las filas 1 y k . Cada permutación cambia el signo del determinante y al ser tres permutaciones, el intercambio de la fila k con la fila k_0 cambia el signo.

La prueba es por inducción en n . El caso base $n = 1$ es completamente trivial. (O, si lo prefiere, puede tomar $n = 2$ como el caso base, y el teorema es fácilmente probado usando la fórmula para el determinante de una matriz 2×2). Las definiciones de los determinantes de A y B son:

$$\det(A) = \sum_{i=1}^n a_{i1} C_{i1}^A \quad \text{y} \quad \det(B) = \sum_{i=1}^n b_{i1} C_{i1}^B.$$

Supongamos primero que $i \neq 1, k$. En este caso, está claro que $A(i|1)$ y $B(i|1)$ son iguales, excepto que dos filas se intercambian. Por lo tanto, por hipótesis inductiva $C_{i1}^A = -C_{i1}^B$. Ya que también $a_{i1} = b_{i1}$, tenemos entonces que

$$(54) \quad a_{i1} C_{i1}^A = -b_{i1} C_{i1}^B, \quad \text{para } i \neq 1, k.$$

Queda por considerar los términos $i = 1$ y $i = k$. Nosotros afirmamos que

$$(55) \quad -a_{k1} C_{k1}^A = b_{11} C_{11}^B \quad \text{y} \quad -a_{11} C_{11}^A = b_{k1} C_{k1}^B.$$

Si probamos esto, entonces

$$\begin{aligned} \det(A) &= \sum_{i=1}^n a_{i1} C_{i1}^A \\ &= a_{11} C_{11}^A + \sum_{i=2}^{k-1} a_{i1} C_{i1}^A + a_{k1} C_{k1}^A + \sum_{i=k+1}^n a_{i1} C_{i1}^A \quad \text{por (54) y (55)} \\ &= -b_{k1} C_{k1}^B - \sum_{i=2}^{k-1} b_{i1} C_{i1}^B - b_{11} C_{11}^B - \sum_{i=k+1}^n b_{i1} C_{i1}^B \\ &= -\sum_{i=1}^n b_{i1} C_{i1}^B = -\det(B). \end{aligned}$$

Luego el teorema está probado. Por lo tanto debemos probar (55). Por simetría, basta probar la primera identidad de (55), es decir que $a_{k1} C_{k1}^A = -b_{11} C_{11}^B$.

Para esto, primero debemos observar que $a_{k1} = b_{11}$, por lo tanto sólo hace falta probar que $-C_{k1}^A = C_{11}^B$. En segundo lugar, debemos tener en cuenta que $B(1|1)$ se obtiene de $A(k|1)$ reordenando las filas $1, 2, \dots, k-1$ de $A(k|1)$ en el orden $2, 3, \dots, k-1, 1$. Este reordenamiento puede hacerse permutando la fila 1 con la fila 2, luego permutando esa fila con la fila 3, etc.,

terminando con una permutación con la fila $k - 1$. Esto es un total de $k - 2$ permutaciones de fila. Así que, por hipótesis inductiva,

$$\det(B(1|1)) = (-1)^{k-2} \det(A(k|1)) = (-1)^k \det(A(k|1)) = -(-1)^{k+1} \det(A(k|1)),$$

es decir $C_{11}^B = -C_{k1}^A$. Esto completa la demostración del teorema. \square

OBSERVACIÓN A.1.1. Recordar que del resultado anterior se deduce fácilmente que si una matriz tiene dos filas iguales entonces su determinante es 0. Esto se debía a que, intercambiando las dos filas iguales obtenemos la misma matriz, pero calculando el determinante con el teorema anterior vemos que cambia de signo y el único número en \mathbb{K} que es igual a su opuesto es el 0.

LEMA A.1.3. Sean A, B, C matrices $n \times n$ tal que

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

y

$$C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{r1} + b_{r1} & a_{r2} + b_{r2} & \cdots & a_{rn} + b_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Es decir B es igual a A pero con la fila r cambiada y C es como A y B excepto en la fila r donde cada coeficiente es la suma de los de A y B correspondiente. Entonces $\det(C) = \det(A) + \det(B)$.

DEMOSTRACIÓN. Se hará por inducción en n . Para $n = 1$, del resultado se reduce a probar que $\det[a + b] = \det[a] + \det[b]$, lo cual es trivial, pues el determinante en matrices 1×1 es la identidad.

Primero consideremos el caso $r = 1$. En este caso tenemos que $A(1|1) = B(1|1) = C(1|1)$, pues en la única fila que difieren las matrices es en la primera. Además, si $i > 1$, $A(i|1)$, $B(i|1)$ y $C(i|1)$ son iguales, excepto que difieren en la primera fila donde los coeficientes de $C(i|1)$ son la suma de los de $A(i|1)$ y $B(i|1)$, entonces, por hipótesis inductiva, $\det C(i|1) = \det A(i|1) + \det B(i|1)$. Concluyendo, tenemos que

$$\begin{aligned} \det A(1|1) &= \det B(1|1) = \det C(1|1), \\ \det C(i|1) &= \det A(i|1) + \det B(i|1), \quad i > 1, \end{aligned}$$

lo cual implica que

$$\begin{aligned} C_{11}^C &= C_{11}^A = C_{11}^B, \\ C_{i1}^C &= C_{i1}^A + C_{i1}^B, \quad i > 1 \end{aligned} .$$

Luego

$$\begin{aligned} \det C &= (a_{11} + b_{11})C_{11}^C + \sum_{i=2}^n a_{i1}C_{i1}^C \\ &= a_{11}C_{11}^C + b_{11}C_{11}^C + \sum_{i=2}^n a_{i1}(C_{i1}^A + C_{i1}^B) \\ &= a_{11}C_{11}^A + b_{11}C_{11}^B + \sum_{i=2}^n a_{i1}(C_{i1}^A + C_{i1}^B) \\ &= a_{11}C_{11}^A + \sum_{i=2}^n a_{i1}C_{i1}^A + b_{11}C_{11}^B + \sum_{i=2}^n a_{i1}C_{i1}^B \\ &= \det A + \det B. \end{aligned}$$

El caso $r > 1$ se demuestra de manera similar o, si se prefiere, puede usarse el teorema [A.1.2](#), observando que la permutación entre la fila 1 y la fila r cambia el signo del determinante. \square

TEOREMA A.1.4. *Sea $A \in M_n(\mathbb{K})$. Sea $c \in \mathbb{K}$ y B la matriz que se obtiene de A sumando a la fila r la fila s multiplicada por c , es decir $A \xrightarrow{F_r + cF_s} B$, entonces $\det B = \det A$.*

DEMOSTRACIÓN. A y B difieren solo en la fila r , donde los coeficientes de B son los de A más c por los de la fila s . Luego si

$$A = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_r \\ \vdots \\ F_n \end{bmatrix}, \quad B = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_r + cF_s \\ \vdots \\ F_n \end{bmatrix}, \quad A' = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ cF_s \\ \vdots \\ F_n \end{bmatrix},$$

el lema anterior nos dice que

$$(56) \quad \det B = \det A + \det A'.$$

Ahora bien, por teorema [A.1.1](#),

$$\det A' = c \begin{vmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_s \\ \vdots \\ F_n \end{vmatrix},$$

y este último determinante es cero, debido a que la matriz tiene dos filas iguales. Luego, $\det B = \det A$. \square

A continuación veremos que el determinante del producto de matrices es el producto de los determinantes de las matrices.

TEOREMA A.1.5. *Sea $A \in M_n(\mathbb{K})$ y E una matriz elemental $n \times n$. Entonces*

$$(57) \quad \det(EA) = \det E \det A.$$

DEMOSTRACIÓN. (1) Si $c \neq 0$, y E es la matriz elemental que se obtiene de multiplicar por c la fila r de I_n , sabemos, por corolario [1.6.5](#), que tiene determinante igual a c . Por otro lado, el teorema [1.4.1](#) nos dice que EA es la matriz que se obtiene a partir de A multiplicando la fila r por c , y entonces, por teorema [1.6.4](#), $\det(EA) = c \det A = \det E \det A$.

(2) Si E es la matriz elemental que se obtiene de sumar a la fila r de I_n la fila s multiplicada por c , entonces $\det E = 1$. Por otro lado $\det(EA) = \det(A)$, por lo tanto $\det(EA) = \det(E) \det(A)$.

(3) Finalmente, si E es la matriz elemental que se obtiene de intercambiar la fila r por la fila s de I_n , entonces $\det E = -1$. Por otro lado $\det(EA) = -\det(A)$, por lo tanto $\det(EA) = \det(E) \det(A)$. \square

COROLARIO A.1.6. *Sea $A \in M_n(\mathbb{K})$ y E_1, \dots, E_k matrices elementales $n \times n$. Entonces*

$$\det(E_k E_{k-1} \dots E_1 A) = \det(E_k) \det(E_{k-1}) \dots \det(E_1) \det(A).$$

DEMOSTRACIÓN. Por la aplicación reiterada del teorema [A.1.5](#) tenemos,

$$\begin{aligned} \det(E_k E_{k-1} \dots E_1 A) &= \det(E_k) \det(E_{k-1} \dots E_1 A) \\ &= \det(E_k) \det(E_{k-1}) \det(E_{k-2} \dots E_1 A) \\ &\quad \vdots \\ &= \det(E_k) \det(E_{k-1}) \det(E_{k-2}) \dots \det(E_1) \det(A). \end{aligned}$$

\square

OBSERVACIÓN A.1.1. Sea $A \in M_n(\mathbb{K})$, entonces A es equivalente por filas a una matriz R que es MERF, por lo tanto R es equivalente por filas a A . Es decir, existen E_1, \dots, E_k matrices elementales, tal que

$$(58) \quad A = E_k E_{k-1} \dots E_1 R.$$

Ya vimos, en el teorema 1.5.4 que si A inversible $R = I_n$ y por lo tanto A es equivalente por filas a un producto de matrices elementales. Si A no es inversible, entonces R no es la identidad y por lo tanto tiene filas nulas.

TEOREMA A.1.7. Sean $A, B \in M_n(\mathbb{K})$, entonces

- (1) $\det(AB) = \det(A) \det(B)$.
- (2) A inversible si y solo si $\det(A) \neq 0$.

DEMOSTRACIÓN. Sea $A = E_k E_{k-1} \dots E_1 R$ con E_i elemental y R una MERF.

(1) Como $AB = E_k E_{k-1} \dots E_1 RB$, por corolario A.1.6 tenemos que $\det(AB) = \det(E_k) \dots \det(E_1) \det(RB)$.

Separaremos la demostración en dos casos: (a) A es inversible y (b) A no es inversible.

(a) Si A es inversible, entonces $R = I_n$ y $A = E_k \dots E_1$, por lo tanto

$$\det(AB) = \det(E_k) \dots \det(E_1) \det(B) = \det(A) \det(B).$$

(b) Si A no es inversible, entonces R no es la identidad, por lo tanto tiene la última fila nula y por lo tanto $\det(R) = 0$, y en consecuencia $\det(A) = 0$. Como R tiene la última fila nula, no es difícil ver que entonces RB tiene la última fila nula y por lo tanto $\det(RB) = 0$. Luego

$$\det(AB) = \det(E_k) \dots \det(E_1) \det(RB) = 0.$$

Como $\det(A) = 0$, tenemos también

$$\det(A) \det(B) = 0.$$

(2) (\Rightarrow) Como A inversible, $R = I_n$ y $\det A = \det(E_k) \dots \det(E_1)$ es no nulo, pues las matrices elementales tiene determinante no nulo.

(2) (\Leftarrow) Si $\det(A) \neq 0$, entonces $\det(R) \neq 0$, luego $R = I_n$ y A es producto de matrices elementales, por lo tanto inversible. \square

Haremos ahora la demostración del teorema 1.6.10.

TEOREMA A.1.8. Sea E matriz elemental, entonces E^t es matriz elemental del mismo tipo y $\det(E) = \det(E^t)$.

DEMOSTRACIÓN. Si $c \neq 0$ y E es la matriz elemental que se obtiene de multiplicar por c la fila r de I_n , es claro que $E^t = E$ y por lo tanto $\det(E) = \det(E^t)$.

Si E es la matriz elemental que se obtiene de sumar a la fila r de I_n la fila s multiplicada por $c \in \mathbb{K}$, entonces E^t es la matriz elemental que se obtiene de sumar a la fila s de I_n la fila r multiplicada por c . Luego, $\det(E) = \det(E^t) = 1$.

Finalmente, si E es la matriz elemental que se obtiene de intercambiar la fila r por la fila s de I_n , entonces $E^t = E$ y por lo tanto $\det(E) = \det(E^t)$. \square

TEOREMA A.1.9. Sea $A \in M_n(\mathbb{K})$, entonces $\det(A) = \det(A^t)$

DEMOSTRACIÓN. Si A es inversible, entonces $A = E_k E_{k-1} \dots E_1$ con E_i elemental, por lo tanto $\det(A) = \det(E_k) \det(E_{k-1}) \dots \det(E_1)$. Luego,

$$\det(A^t) = \det(E_1^t \dots E_k^t) = \det(E_1^t) \dots \det(E_k^t) = \det(E_1) \dots \det(E_k) = \det(A).$$

Si A no es inversible, entonces A^t no es inversible y en ese caso $\det(A) = \det(A^t) = 0$. \square

Finalmente, demostremos el teorema 1.6.14.

TEOREMA A.1.10. El determinante de una matriz A de orden $n \times n$ puede ser calculado por la expansión de los cofactores en cualquier columna o cualquier fila. Más específicamente,

(1) si usamos la expansión por la j -ésima columna, $1 \leq j \leq n$, tenemos

$$\begin{aligned} \det A &= \sum_{i=1}^n a_{ij} C_{ij} \\ &= a_{1j} C_{1j} + a_{2j} C_{2j} + \dots + a_{nj} C_{nj}. \end{aligned}$$

(2) si usamos la expansión por la i -ésima fila, $1 \leq i \leq n$, tenemos

$$\begin{aligned} \det A &= \sum_{j=1}^n a_{ij} C_{ij} \\ &= a_{i1} C_{i1} + a_{i2} C_{i2} + \dots + a_{in} C_{in}; \end{aligned}$$

DEMOSTRACIÓN. (1) Primero hagamos la demostración para $j = 2$, es decir para el desarrollo por la segunda columna. Escribamos A en función de sus columnas, es decir

$$A = \begin{bmatrix} C_1 & C_2 & C_3 & \dots & C_n \end{bmatrix},$$

donde C_k es la columna k de A . Sea $B = [b_{ij}]$ la matriz definida por

$$B = \begin{bmatrix} C_2 & C_1 & C_3 & \dots & C_n \end{bmatrix}.$$

Entonces, $\det(B) = -\det(A)$. Por otro lado, por la definición de determinante,

$$\begin{aligned}\det(B) &= \sum_{i=1}^n b_{i1} C_{i1}^B \\ &= \sum_{i=1}^n b_{i1} (-1)^{i+1} B(i|1) \\ &= \sum_{i=1}^n a_{i2} (-1)^{i+1} B(i|1).\end{aligned}$$

Ahora bien, es claro que $B(i|1) = A(i|2)$, por lo tanto

$$\det(B) = \sum_{i=1}^n a_{i2} (-1)^{i+1} A(i|2) = - \sum_{i=1}^n a_{i2} C_{i2}.$$

Es decir, $\det(A) = -\det(B) = \sum_{i=1}^n a_{i2} C_{i2}$.

El caso $j > 2$ se demuestra de forma similar: si B es la matriz

$$B = \begin{bmatrix} C_j & C_1 & C_2 & \cdots & C_{j-1} & C_{j+1} & \cdots & C_n \end{bmatrix}.$$

entonces $\det(B) = (-1)^{j-1} \det(A)$, pues son necesarios $j-1$ permutaciones para recuperar la matriz A (es decir, llevar la columna j a su lugar). Como $B(i|1) = A(i|j)$, desarrollando por la primera columna el determinante de B obtenemos el resultado.

(2) Observemos primero que $A^t(j|i) = A(i|j)^t$, por lo tanto, si calculamos $\det(A^t)$ por desarrollo por columna i , obtenemos

$$\begin{aligned}\det A &= \det(A^t) = \sum_{j=1}^n [A^t]_{ji} (-1)^{i+j} \det(A^t(j|i)) \\ &= \sum_{j=1}^n a_{ij} (-1)^{i+j} \det(A(i|j)^t) \\ &= \sum_{j=1}^n a_{ij} (-1)^{i+j} \det(A(i|j)).\end{aligned}$$

□

A.2. Regla de Cramer

Veremos ahora que la inversa de una matriz inversible se puede escribir en términos de determinantes de algunas matrices relacionadas.

TEOREMA A.2.1. *Sea A matriz $n \times n$, entonces*

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} A = \begin{bmatrix} 0 & \cdots & 0 & \det A & 0 & \cdots & 0 \end{bmatrix}.$$

$\uparrow i$

Es decir, la matriz fila formada por los cofactores correspondientes a la columna i multiplicada por la matriz A es igual a la matriz fila con valor $\det A$ en la posición i y 0 en las otras posiciones.

DEMOSTRACIÓN. Si C_j denota la matriz formada por la columna j de A debemos probar que

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = \sum_{k=1}^n a_{kj} C_{ki} = \begin{cases} \det(A) & \text{si } j = i \\ 0 & \text{si } j \neq i. \end{cases}$$

Ahora bien,

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_i = \sum_{j=1}^n C_{ji} a_{ji},$$

y esto último no es más que el cálculo del determinante por desarrollo de la columna i , es decir, es igual a $\det(A)$.

Para ver el caso $i \neq j$, primero observemos que si

$$B = \begin{bmatrix} C_1 & C_2 & \cdots & C_j & \cdots & C_j & \cdots & C_{n-1} & C_n \end{bmatrix},$$

$\uparrow i \qquad \qquad \uparrow j$

es decir, B es la matriz A donde reemplazamos la columna i por la columna j , entonces como B tiene dos columnas iguales, $\det(B) = 0$. Por lo tanto, si calculamos el determinante de B por el desarrollo en la columna i , obtenemos

$$(59) \qquad 0 = \det(B) = \sum_{k=1}^n a_{kj} C_{ki}.$$

Por otro lado,

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = \sum_{k=1}^n C_{ki} a_{kj},$$

luego, por la ecuación (59) tenemos que

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = 0$$

si $i \neq j$.

□

DEFINICIÓN A.2.1. Sea A matriz $n \times n$, la *matriz de cofactores* es la matriz cuyo coeficiente ij vale C_{ij} . La matriz de cofactores de A se denota $\text{cof}(A)$. La *matriz adjunta de A* es $\text{adj}(A) = \text{cof}(A)^t$.

TEOREMA A.2.2. Sea A matriz $n \times n$, entonces

$$\text{adj}(A) \cdot A = \det(A) I_n.$$

DEMOSTRACIÓN. Observar que la fila i de $\text{adj}(A)$ es $\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix}$. Por lo tanto, la fila i de $\text{adj}(A).A$ es

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} A,$$

que por el teorema A.2.1 es una matriz fila con el valor $\det A$ en la posición i y todos los demás coeficientes iguales a 0. Luego

$$\text{adj}(A).A = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} A = \begin{bmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det A \end{bmatrix} = \det(A)I_n$$

□

COROLARIO A.2.3. Si A es invertible, entonces

$$A^{-1} = \frac{1}{\det A} \text{adj } A.$$

DEMOSTRACIÓN.

$$\frac{1}{\det A} \text{adj } A.A = \frac{1}{\det A} \det A I_n = I_n.$$

□

TEOREMA A.2.4 (Regla de Cramer). Sea $AX = Y$ un sistema de ecuaciones tal que $A \in M_n(\mathbb{K})$ es invertible. Entonces, el sistema tiene una única solución (x_1, \dots, x_n) con

$$x_j = \frac{\det A_j}{\det A}, \quad j = 1, \dots, n,$$

donde A_j es la matriz $n \times n$ que se obtiene de A reemplazando la columna j de A por Y .

DEMOSTRACIÓN. Haremos la demostración para matrices 3×3 . La demostración en el caso general es completamente análoga.

Como A es invertible, existe A^{-1} y multiplicamos la ecuación a izquierda por A^{-1} y obtenemos que $A^{-1}AX = A^{-1}Y$, es decir $X = A^{-1}Y$ y esta es la única solución. Luego

$$\begin{aligned} A^{-1}Y &= \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ (*) \quad &= \frac{1}{\det A} \begin{bmatrix} y_1 C_{11} + y_2 C_{21} + y_3 C_{31} \\ y_1 C_{12} + y_2 C_{22} + y_3 C_{32} \\ y_1 C_{13} + y_2 C_{23} + y_3 C_{33} \end{bmatrix} \end{aligned}$$

Ahora bien, $y_1C_{11} + y_2C_{21} + y_3C_{31}$ es el cálculo de determinante por desarrollo de la primera columna de la matriz

$$\begin{bmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{32} & a_{33} \end{bmatrix},$$

y, de forma análoga, el segundo y tercer coeficiente de la matriz (*) son el determinante de las matrices 3×3 que se obtienen de A reemplazando la columna 2 y 3, respectivamente, de A por Y . Es decir

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A^{-1}Y = \frac{1}{\det A} \begin{bmatrix} \det A_1 \\ \det A_2 \\ \det A_3 \end{bmatrix} = \begin{bmatrix} \frac{\det A_1}{\det A} \\ \frac{\det A_2}{\det A} \\ \frac{\det A_3}{\det A} \end{bmatrix},$$

luego $x_j = \frac{\det A_j}{\det A}$ para $j = 1, 2, 3$. □

EJEMPLO A.2.1. Resolvamos usando la regla de Cramer el siguiente sistema:

$$x_1 + x_2 - x_3 = 6$$

$$3x_1 - 2x_2 + x_3 = -5$$

$$x_1 + 3x_2 - 2x_3 = 14.$$

La matriz asociada al sistema es

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 3 & -2 & 1 \\ 1 & 3 & -2 \end{bmatrix}.$$

Luego

$$A_1 = \begin{bmatrix} 6 & 1 & -1 \\ -5 & -2 & 1 \\ 14 & 3 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 6 & -1 \\ 3 & -5 & 1 \\ 1 & 14 & -2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 1 & 6 \\ 3 & -2 & -5 \\ 1 & 3 & 14 \end{bmatrix},$$

y

$$\det A = -3, \quad \det A_1 = -3, \quad \det A_2 = -9, \quad \det A_3 = 6.$$

Por lo tanto,

$$\begin{aligned} x_1 &= \frac{\det A_1}{\det A} = \frac{-3}{-3} = 1 \\ x_2 &= \frac{\det A_2}{\det A} = \frac{-9}{-3} = 3 \\ x_3 &= \frac{\det A_3}{\det A} = \frac{6}{-3} = -2. \end{aligned}$$

OBSERVACIÓN A.2.1. En general, la regla de Cramer no es utilizada para resolver sistemas de ecuaciones. El método de Gauss, además de ser mucho más rápido, tiene el beneficio de ser un método que abarca más casos y nos da más información que la regla de Cramer. Por ejemplo, el método de Gauss también funciona para matrices no inversibles, mientras que la regla de Cramer no. Más aún, el método de Gauss se puede utilizar para resolver sistemas con matrices no cuadradas (es decir, sistemas de ecuaciones con números no iguales de variables y ecuaciones), mientras que la regla de Cramer no.

Con respecto a la velocidad de resolución, para matrices $n \times n$ el método de Gauss requiere “alrededor de n^3 operaciones”, mientras que la regla de Cramer requiere “alrededor de n^4 operaciones”, haciendo el primer método n veces más rápido que el segundo. En lenguaje técnico, el método de Gauss tiene complejidad $O(n^3)$ y la regla de Cramer complejidad $O(n^4)$ (ver https://es.wikipedia.org/wiki/Eficiencia_Algorítmica)