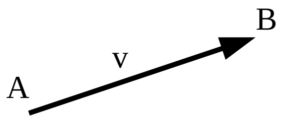
Capítulo 1

Álgebra vectorial

Un vector es una flecha que se utiliza para representar magnitudes físicas como fuerzas, velocidades, aceleraciones, etc.



Un vector consiste de un origen A y un extremo B que lo determinan completamente. Notaremos $v = \overrightarrow{AB}$.

Trabajaremos con vectores en el plano ($\mathbb{R}^2$) y en el espacio ($\mathbb{R}^3$).

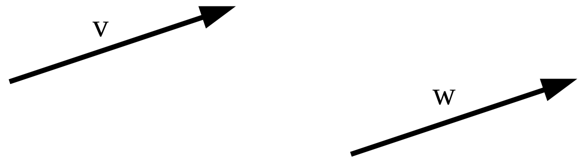
Recordemos que $\mathbb{R}^2 = \{(x, y) \,/\, x, y \in \mathbb{R}\}$. De manera similar $\mathbb{R}^3 = \{(x, y, z) \,/\, x, y, z \in \mathbb{R}\}$.

Un vector posee tres características importantes que detallamos a continuación:

- Dirección: Es la recta que contiene al vector.
- Sentido: Es el indicado por la flecha.
- Longitud: Es la medida del vector. Está dada por la distancia entre el origen A y el extremo B del vector.

Diremos que dos vectores son *equivalentes* si tienen direcciones paralelas, el mismo sentido y la misma longitud.

Los vectores v y w de la siguiente figura son equivalentes.



La siguiente proposición nos da un criterio para saber si dos vectores son equivalentes o no.

Proposición 1.1. *Sean $A = (x_1, y_1)$, $B = (x_2, y_2)$, $C = (x_3, y_3)$ y $D = (x_4, y_4)$ puntos de $\mathbb{R}^2$ con $A \neq B$ y $C \neq D$. Entonces los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son equivalentes si y sólo si $(x_2 - x_1, y_2 - y_1) = (x_4 - x_3, y_4 - y_3)$.*

Demostración. $\Rightarrow$) Supongamos que $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son vectores equivalentes y supongamos primero que $x_1 \neq x_2$. Entonces la recta que contiene al vector $\overrightarrow{AB}$ no es vertical. Como $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son equivalentes entonces las rectas que los contienen son paralelas y, por lo tanto, la recta que contiene al vector $\overrightarrow{CD}$ tampoco es vertical y tiene la misma pendiente que la recta que contiene al vector $\overrightarrow{AB}$. Si llamamos m a la pendiente de estas rectas tenemos que

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{y_4 - y_3}{x_4 - x_3}$$

de donde

$$y_2 - y_1 = m(x_2 - x_1) \ \text{ y } \ y_4 - y_3 = m(x_4 - x_3). \tag{1.1}$$

Como $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son equivalentes entonces $\operatorname{long}(\overrightarrow{AB}) = \operatorname{long}(\overrightarrow{CD})$ y luego se tiene que

$$\begin{aligned}
\mathrm{d(A,B)} &= \mathrm{d}(C, D) \\
\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} &= \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2} \\
(x_2 - x_1)^2 + (y_2 - y_1)^2 &= (x_4 - x_3)^2 + (y_4 - y_3)^2 \\
(x_2 - x_1)^2 + m^2(x_2 - x_1)^2 &= (x_4 - x_3)^2 + m^2(x_4 - x_3)^2 \\
(1 + m^2)(x_2 - x_1)^2 &= (1 + m^2)(x_4 - x_3)^2
\end{aligned}$$

y como $1 + m^2 \geq 1 > 0$ se tiene que $1 + m^2 \neq 0$ y entonces de la igualdad anterior obtenemos que

$$\begin{aligned}
(x_2 - x_1)^2 &= (x_4 - x_3)^2 \\
\sqrt{(x_2 - x_1)^2} &= \sqrt{(x_4 - x_3)^2} \\
|x_2 - x_1| &= |x_4 - x_3|
\end{aligned}$$

Luego

$$x_2 - x_1 = x_4 - x_3 \ \text{ o } \ x_2 - x_1 = -(x_4 - x_3). \tag{1.2}$$

Como los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son equivalentes entonces tienen el mismo sentido y luego

$$(x_2 > x_1 \text{ y } x_4 > x_3) \text{ o } (x_2 < x_1 \text{ y } x_4 < x_3)$$

(notar que en el primer caso ambos vectores apuntan hacia la derecha y en el segundo caso ambos vectores apuntan hacia la izquierda).

En ambos casos $x_2 - x_1$ y $x_4 - x_3$ tienen el mismo signo. Luego, de (1.2), obtenemos que $x_2 - x_1 = x_4 - x_3$. Entonces, de (1.1), se sigue que

$$y_2 - y_1 = m(x_2 - x_1) = m(x_4 - x_3) = y_4 - y_3.$$

Por lo tanto, $(x_2 - x_1, y_2 - y_1) = (x_4 - x_3, y_4 - y_3)$.

Resta analizar el caso $x_1 = x_2$ que queda como ejercicio para el lector.

$\Leftarrow$) Supongamos que $(x_2 - x_1, y_2 - y_1) = (x_4 - x_3, y_4 - y_3)$. Entonces

$$\mathrm{d}(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2} = \mathrm{d}(C, D)$$

y por lo tanto, los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ tienen la misma longitud.

Además, si $x_2 - x_1 \neq 0$ entonces $x_4 - x_3 \neq 0$ y, por lo tanto, las rectas que contienen a los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ no son verticales y tienen pendientes

$$\frac{y_2 - y_1}{x_2 - x_1} \ \text{ y } \ \frac{y_4 - y_3}{x_4 - x_3}$$

que claramente son iguales. Por lo tanto, los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ tienen direcciones paralelas. Y como $x_2 - x_1 = x_4 - x_3$ entonces $x_2 - x_1$ y $x_4 - x_3$ tienen el mismo signo. Luego

$$(x_2 > x_1 \text{ y } x_4 > x_3) \text{ o } (x_2 < x_1 \text{ y } x_4 < x_3)$$

y entonces $\overrightarrow{AB}$ y $\overrightarrow{CD}$ tienen el mismo sentido.

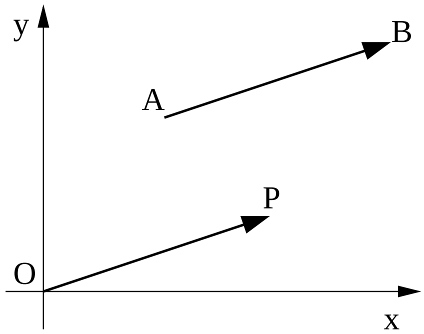
Si $x_2 - x_1 = 0$ entonces $x_4 - x_3 = 0$ y se obtiene que $x_1 = x_2$ y que $x_3 = x_4$. Por lo tanto, las rectas que contienen a los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son ambas verticales. Luego, son paralelas. Y como $y_2 - y_1 = y_4 - y_3$ entonces $y_2 - y_1$ e $y_4 - y_3$ tienen el mismo signo. Luego

$$(y_2 > y_1 \text{ y } y_4 > y_3) \text{ o } (y_2 < y_1 \text{ y } y_4 < y_3)$$

En el primer caso, los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ apuntan ambos hacia arriba y en el segundo caso apuntan ambos hacia abajo. Entonces $\overrightarrow{AB}$ y $\overrightarrow{CD}$ tienen el mismo sentido.

Por lo tanto, los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son equivalentes. □

Observación 1.2. Sean $A = (x_1, y_1)$ y $B = (x_2, y_2)$ puntos distintos de $\mathbb{R}^2$. Sea $O = (0, 0)$ y sea $P = (x_2 - x_1, y_2 - y_1)$. Entonces el vector $\overrightarrow{AB}$ es equivalente al vector $\overrightarrow{OP}$.



Notar que $\overrightarrow{OP}$ se obtiene trasladando el vector $\overrightarrow{AB}$ hasta el origen de coordenadas.

Diremos que un vector es nulo si su origen y su extremo coinciden. Notemos que un vector es nulo si y sólo si tiene longitud cero. Notemos también que para los vectores nulos no podemos definir dirección ni sentido pues no hay una única recta que contenga al vector.

Diremos entonces que todos los vectores nulos son equivalentes entre sí. Por ende, todo vector nulo es equivalente al vector $\overrightarrow{OO}$.

Ahora bien, por lo visto hasta aquí sabemos que todo vector del plano tiene uno equivalente a él que tiene su origen en el punto $O = (0, 0)$. Luego, podemos pensar a todos los vectores con su origen en el $(0, 0)$.

Además, dos vectores $\overrightarrow{OA}$ y $\overrightarrow{OB}$ de origen O son equivalentes si y sólo si $A = B$. La demostración de este hecho queda como ejercicio para el lector. Luego, podemos identificar a los vectores del plano con los puntos de $\mathbb{R}^2$ pues para dar un vector que empieza en $(0, 0)$ basta con dar el extremo de dicho vector.

Por lo tanto, a partir de ahora pensaremos a los vectores del plano como puntos de $\mathbb{R}^2$.

Una argumentación similar puede hacerse en $\mathbb{R}^3$ y, por lo tanto, podemos pensar a los vectores del espacio como puntos de $\mathbb{R}^3$.

Definición 1.3. Sea $n \in \mathbb{N}$. Definimos $\mathbb{R}^n = \overbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}^{n \text{ veces}}$, es decir,

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \,/\, x_i \in \mathbb{R} \;\; \forall i \in \{1, 2, \ldots, n\}\}.$$

Definimos en $\mathbb{R}^n$ la operación de *suma*, que notaremos $+$, por

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

Notaremos $0 = (0, 0, \ldots, 0) \in \mathbb{R}^n$.

Dado $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ definimos su opuesto por $-x = (-x_1, -x_2, \ldots, -x_n)$.

Si $x, y \in \mathbb{R}^n$ definimos la resta $x - y$ por $x - y = x + (-y)$.

Además, si $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ definimos $\alpha . x$ por

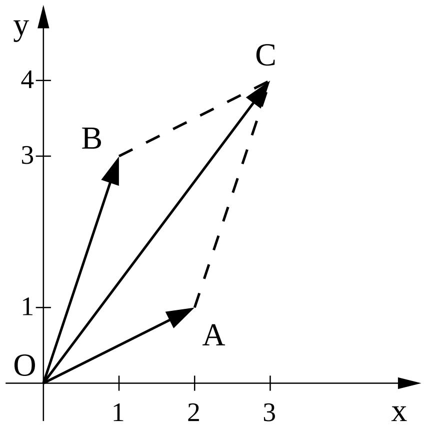
$$\alpha . x = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n).$$

Esta operación recibe el nombre de *producto por escalares*.

Interpretación geométrica en $\mathbb{R}^2$

Ejemplo 1.4. Sean $A = (2, 1)$, $B = (1, 3)$ y $C = A + B$.

Se tiene que $C = (2, 1) + (1, 3) = (3, 4)$. Sea $O = (0, 0)$. Grafiquemos los vectores $\overrightarrow{OA}$, $\overrightarrow{OB}$ y $\overrightarrow{OC}$.



Observamos que el segmento $\overline{OC}$ es la diagonal del paralelogramo que tiene como dos de sus lados a los segmentos $\overline{OA}$ y $\overline{OB}$.

Esto ocurre en general: si $A, B \in \mathbb{R}^2$ y A, B y $O = (0, 0)$ no están alineados entonces el paralelogramo que tiene como dos de sus lados a los segmentos $\overline{OA}$ y $\overline{OB}$ tiene como vértices a los puntos O, A, B y $A + B$.

En efecto, si $A = (a_1, a_2)$, $B = (b_1, b_2)$ y $C = A + B = (a_1 + b_1, a_2 + b_2)$ entonces los vectores $\overrightarrow{OA}$ y $\overrightarrow{BC}$ son equivalentes pues

$$(a_1 - 0, a_2 - 0) = (a_1, a_2) = (a_1 + b_1 - b_1, a_2 + b_2 - b_2)$$

y luego las rectas $\overleftrightarrow{OA}$ y $\overleftrightarrow{BC}$ son paralelas.

En forma similar, los vectores $\overrightarrow{OB}$ y $\overrightarrow{AC}$ son equivalentes pues

$$(b_1 - 0, b_2 - 0) = (b_1, b_2) = (a_1 + b_1 - a_1, a_2 + b_2 - a_2).$$

Luego las rectas $\overleftrightarrow{OB}$ y $\overleftrightarrow{AC}$ son paralelas.

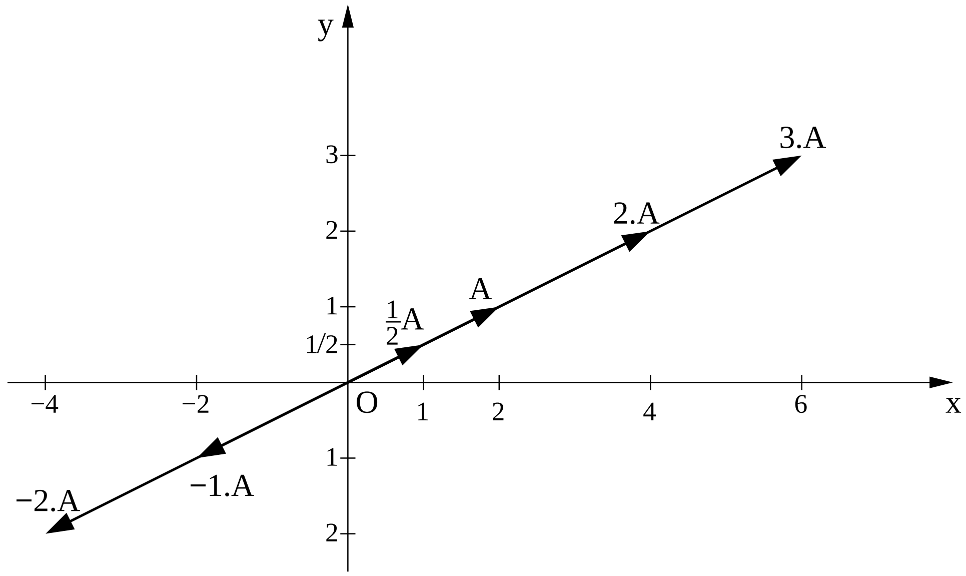
Por lo tanto, el cuadrilátero $OACB$ es un paralelogramo.

Este hecho se llama *regla del paralelogramo*.

Ejemplo 1.5. Sea $A = (2, 1)$. Calcular $2.A$, $3.A$, $\frac{1}{2}.A$, $-1.A$ y $-2.A$. Graficar.

Tenemos que

$$\begin{aligned} 2.A &= 2.(2, 1) = (4, 2) \\ 3.A &= 3.(2, 1) = (6, 3) \\ \tfrac{1}{2}.A &= \tfrac{1}{2}.(2, 1) = (1, \tfrac{1}{2}) \\ -1.A &= -1.(2, 1) = (-2, -1) \\ -2.A &= -2.(2, 1) = (-4, -2) \end{aligned}$$



Observamos que todos los múltiplos del vector $(2,1)$ tienen la misma dirección. Además, cuando multiplicamos el vector $(2,1)$ por un número positivo se conserva el sentido y cuando lo multiplicamos por un número negativo se invierte el sentido.

Observamos también que los vectores $2.(2,1)$ y $-2.(2,1)$ tienen el doble de longitud que el vector $(2,1)$, que el vector $3.(2,1)$ tiene el triple de longitud que el vector $(2,1)$ y que el vector $\frac{1}{2}.(2,1)$ tiene la mitad de longitud que el vector $(2,1)$.

Todo esto vale en general, como muestra la siguiente proposición.

Proposición 1.6. *Sea $A \in \mathbb{R}^2$ tal que $A \neq (0,0)$ y sea $\lambda \in \mathbb{R} - \{0\}$. Sea $B = \lambda.A$. Entonces*

(1) Los vectores $\overrightarrow{OA}$ y $\overrightarrow{OB}$ tienen la misma dirección.

(2) Si $\lambda > 0$ los vectores $\overrightarrow{OA}$ y $\overrightarrow{OB}$ tienen el mismo sentido y si $\lambda < 0$ los vectores $\overrightarrow{OA}$ y $\overrightarrow{OB}$ tienen sentidos opuestos.

(3) $d(O,B) = |\lambda| d(O,A)$

La demostración queda como ejercicio para el lector.

Veamos ahora qué propiedades tienen las operaciones que hemos definido.

Proposición 1.7. *Sea $n \in \mathbb{N}$, sean $x, y, z \in \mathbb{R}^n$ y sean $\alpha, \beta \in \mathbb{R}$. Entonces*

(1) $x + (y + z) = (x + y) + z$

(2) $x + y = y + x$

(3) $x + 0 = x \quad y \quad 0 + x = x$

(4) $x + (-x) = 0 \quad y \quad (-x) + x = 0$

(5) $\alpha.(x + y) = \alpha.x + \alpha.y$

(6) $(\alpha + \beta).x = \alpha.x + \beta.x$

(7) $\alpha.(\beta.x) = (\alpha\beta).x$

(8) $1.x = x$

Demostración.
(2) Supongamos que $x = (x_1, x_2, \dots, x_n)$ y que $y = (y_1, y_2, \dots, y_n)$. Entonces

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) = (y_1 + x_1, y_2 + x_2, \dots, y_n + x_n) = y + x.$$

(5) Utilizando las notaciones anteriores se tiene que

$$\begin{aligned}
\alpha.(x+y) &= \alpha.(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) = \\
&= (\alpha(x_1 + y_1), \alpha(x_2 + y_2), \dots, \alpha(x_n + y_n)) = \\
&= (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, \dots, \alpha x_n + \alpha y_n) = \\
&= (\alpha x_1, \alpha x_2, \dots, \alpha x_n) + (\alpha y_1, \alpha y_2, \dots, \alpha y_n) = \\
&= \alpha.x + \alpha.y.
\end{aligned}$$

Las propiedades (1), (3), (4), (6), (7) y (8) se deducen en forma similar utilizando las propiedades de las operaciones con números reales. Sus demostraciones quedan como ejercicio para el lector. □

Estas son ocho propiedades básicas y fundamentales como veremos más adelante. Otras propiedades de estas operaciones se enuncian en la siguiente proposición.

Proposición 1.8. *Sea $n \in \mathbb{N}$, sean $x, y \in \mathbb{R}^n$ y sea $\alpha \in \mathbb{R}$. Entonces*

(9) $0.x = 0$

(10) $\alpha.0 = 0$

(11) $(-1).x = -x$

(12) $-(x + y) = -x - y$

(13) $\alpha.(x - y) = \alpha.x - \alpha.y$

(14) $\alpha.x = 0$ *si y sólo si* $\alpha = 0$ *o* $x = 0$

Demostración. Las demostraciones de las propiedades (9), (10), (11) y (12) quedan como ejercicio para el lector. Demostraremos ahora las propiedades (13) y (14).
(13) Se tiene que

$\alpha.(x-y)$	$=$	$\alpha.(x + (-y)) =$	por definición de resta de vectores
	$=$	$\alpha.x + \alpha.(-y) =$	por propiedad (5)
	$=$	$\alpha.x + \alpha.((-1)y) =$	por propiedad (11)
	$=$	$\alpha.x + (\alpha(-1)).y =$	por propiedad (7)
	$=$	$\alpha.x + ((-1)\alpha).y =$	por propiedad conmutativa del producto en $\mathbb{R}$
	$=$	$\alpha.x + (-1)(\alpha.y) =$	por propiedad (7)
	$=$	$\alpha.x + (-(\alpha.y)) =$	por propiedad (11)
	$=$	$\alpha.x - \alpha.y$	por definición de resta de vectores

Por lo tanto $\alpha.(x - y) = \alpha.x - \alpha.y$.
(14) $\Rightarrow$) Si $\alpha = 0$ no hay nada que hacer pues el consecuente se cumple. Supongamos entonces que $\alpha \neq 0$. Escribamos $x = (x_1, x_2, \dots, x_n)$. Por hipótesis, $\alpha.x = 0$, es decir, $(\alpha x_1, \alpha x_2, \dots, \alpha x_n) = (0, 0, \dots, 0)$. Luego $\alpha x_i = 0$ para todo $i \in \{1, \dots, n\}$ y como $\alpha \neq 0$ obtenemos que $x_i = 0$ para todo $i \in \{1, \dots, n\}$. Por lo tanto, $x = 0$.
$\Leftarrow$) Si $\alpha = 0$ entonces $\alpha.x = 0.x = 0$ por propiedad (9). Si $x = 0$ entonces $\alpha.x = \alpha.0 = 0$ por propiedad (10). □

Definición 1.9. Sea $n \in \mathbb{N}$ y sea $x \in \mathbb{R}^n$. Definimos la *norma* de x por

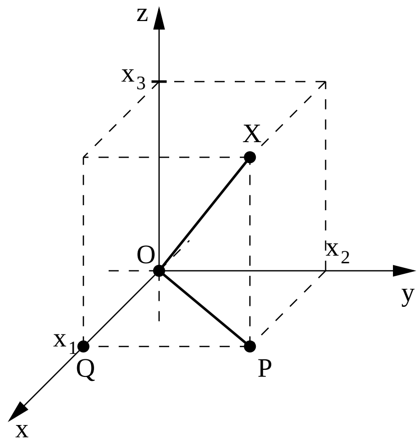
$$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

Ejemplos 1.10. (1) En $\mathbb{R}^2$, si $X = (x_1, x_2)$ se tiene que

$$\|X\| = \sqrt{x_1^2 + x_2^2} = \sqrt{(x_1 - 0)^2 + (x_2 - 0)^2} = \mathrm{d}((x_1, x_2), (0, 0)) = \mathrm{d}(X, O)$$

donde $O = (0, 0)$. Es decir, en $\mathbb{R}^2$ la norma del punto X es su distancia al origen. Por lo tanto, $\|X\|$ es la longitud del vector $\overrightarrow{OX}$.

(2) Analicemos lo que ocurre en $\mathbb{R}^3$. Supongamos que $X = (x_1, x_2, x_3)$. Sean $O = (0, 0, 0)$, $P = (x_1, x_2, 0)$ y $Q = (x_1, 0, 0)$.



Tenemos que $\mathrm{long}(\overline{OP}) = \mathrm{d}(O, P) = \sqrt{x_1^2 + x_2^2}$ por el teorema de Pitágoras pues el triángulo $\triangle OQP$ es rectángulo en Q. Ahora bien, como el triángulo $\triangle OPX$ es rectángulo en P se tiene que $m(\overline{OX})^2 = m(\overline{OP})^2 + m(\overline{PX})^2$ (donde $m(s)$ denota la medida del segmento s). Luego

$$m(\overline{OX})^2 = \left(\sqrt{x_1^2 + x_2^2}\right)^2 + x_3^2 = x_1^2 + x_2^2 + x_3^2.$$

Por lo tanto, $\mathrm{d}(O, X) = m(\overline{OX}) = \sqrt{x_1^2 + x_2^2 + x_3^2} = \|X\|$.

Veamos ahora algunas propiedades de la norma.

Proposición 1.11. *Sea $n \in \mathbb{N}$ y sea $x \in \mathbb{R}^n$. Entonces*

(1) $\|x\| \geq 0$

(2) $\|x\| = 0$ *si y sólo si* $x = 0$.

(3) Si $\alpha \in \mathbb{R}$ *entonces* $\|\alpha . x\| = |\alpha| . \|x\|$.

Demostración.

(1) Nada que hacer.

(2) $\Rightarrow$) Escribamos $x = (x_1, x_2, \ldots, x_n)$. Tenemos que $0 = \|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$. Entonces $x_1^2 + x_2^2 + \ldots + x_n^2 = 0$.

Por otra parte, $x_i^2 \geq 0$ para todo $i \in \{1, \ldots, n\}$. Luego, dado $j \in \{1, \ldots, n\}$ se tiene que

$$0 \leq x_j^2 \leq x_1^2 + x_2^2 + \ldots + x_n^2 = 0$$

de donde obtenemos que $x_j^2 = 0$, lo cual implica que $x_j = 0$.

Así, $x_j = 0$ para todo $j \in \{1, \ldots, n\}$ y por lo tanto $x = (0, 0, \ldots, 0) = 0$.

$\Leftarrow$) Trivial.

(3) Escribamos $x = (x_1, x_2, \ldots, x_n)$. Tenemos que

$$\begin{aligned} \|\alpha.x\| &= \|\alpha.x\| = \|\alpha.(x_1, x_2, \ldots, x_n)\| = \|(\alpha x_1, \alpha x_2, \ldots, \alpha x_n)\| = \\ &= \sqrt{(\alpha x_1)^2 + (\alpha x_2)^2 + \ldots + (\alpha x_n)^2} = \sqrt{\alpha^2 x_1^2 + \alpha^2 x_2^2 + \ldots + \alpha^2 x_n^2} = \\ &= \sqrt{\alpha^2 (x_1^2 + x_2^2 + \ldots + x_n^2)} = \sqrt{\alpha^2}\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = |\alpha|.\|x\| \end{aligned}$$

□

Ejemplo 1.12. Sea $A = (3, 4)$. Calcular la norma de A y hallar un punto $B \in \mathbb{R}^2$ tal que el vector $\overrightarrow{OB}$ tenga la misma dirección y el mismo sentido que el vector $\overrightarrow{OA}$ y tal que la longitud del vector $\overrightarrow{OB}$ sea 1.

Solución:

$$\|A\| = \|(3, 4)\| = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5.$$

Para conseguir un vector $\overrightarrow{OB}$ de la misma dirección que el vector $\overrightarrow{OA}$ necesitamos que $B = \alpha \, . \, A$ para algún $\alpha \in \mathbb{R}$. Además, para que $\overrightarrow{OB}$ tenga el mismo sentido que $\overrightarrow{OA}$ necesitamos que $\alpha > 0$.

Ahora bien, como la longitud del vector $\overrightarrow{OA}$ es $\|A\|$ que es 5, para obtener un vector de longitud 1 debemos multiplicarlo por $\frac{1}{5}$ o $-\frac{1}{5}$. En este caso, como queremos conservar el sentido, debemos multiplicarlo por $\frac{1}{5}$. Luego tomamos $B = \frac{1}{5}.A = \frac{1}{5}.(3, 4) = (\frac{3}{5}, \frac{4}{5})$.

Observación 1.13. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^n - \{0\}$. Entonces $\|A\| \neq 0$ y se tiene que

$$\left\|\frac{A}{\|A\|}\right\| = \left\|\frac{1}{\|A\|}.A\right\| = \left|\frac{1}{\|A\|}\right|.\|A\| = \frac{1}{\|A\|}.\|A\| = 1$$

Es decir, $\dfrac{A}{\|A\|}$ tiene norma 1.

Definición 1.14. Sea $n \in \mathbb{N}$.

- Un vector (o punto) $A \in \mathbb{R}^n$ se dice *unitario* si $\|A\| = 1$.
- *Normalizar* un vector no nulo significa hallar un vector unitario con la misma dirección y sentido que el vector dado. Para ello basta con dividir al vector dado por su longitud o norma.

Observación 1.15. Sean $A = (a_1, a_2)$ y $B = (b_1, b_2)$ puntos de $\mathbb{R}^2$. Entonces

$$\mathrm{d}(A, B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2} = \|(b_1 - a_1, b_2 - a_2)\| = \|B - A\|$$

Dejamos como ejercicio para el lector demostrar que lo mismo vale en $\mathbb{R}^3$.

Esto da lugar a la siguiente definición.

Definición 1.16. Sea $n \in \mathbb{N}$ y sean $A, B \in \mathbb{R}^n$. Se define la distancia entre A y B por $\mathrm{d}(A, B) = \|B - A\|$.

Notemos que $\|B - A\| = \|A - B\|$ ya que

$$\|B - A\| = \|(-1)(A - B)\| = |-1|\|A - B\| = \|A - B\|.$$

Producto interno

Definición 1.17. Sea $n \in \mathbb{N}$ y sean $x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ puntos de $\mathbb{R}^n$. Definimos el *producto interno* (o *producto escalar*) de x e y por

$$\langle x, y\rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

Notar que $\langle x, y\rangle \in \mathbb{R}$.

Proposición 1.18. *Sea* $n \in \mathbb{N}$. *Entonces*

(1) $\forall u, v, w \in \mathbb{R}^n$, $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.

(2) $\forall v, w \in \mathbb{R}^n$, $\forall \alpha \in \mathbb{R}$, $\langle \alpha v, w\rangle = \alpha\langle v, w\rangle$.

(3) $\forall u, v, w \in \mathbb{R}^n$, $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$.

(4) $\forall v, w \in \mathbb{R}^n$, $\forall \alpha \in \mathbb{R}$, $\langle v, \alpha w\rangle = \alpha\langle v, w\rangle$.

(5) $\forall v, w \in \mathbb{R}^n$, $\langle v, w\rangle = \langle w, v\rangle$.

(6) $\forall v \in \mathbb{R}^n$, $\langle v, v\rangle = \|v\|^2$.

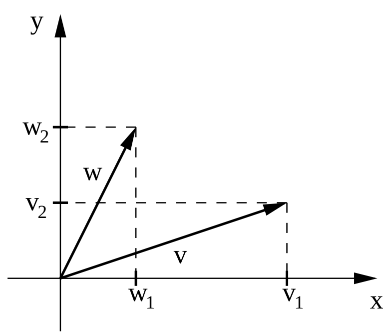
Demostración. (1) Sean $u, v, w \in \mathbb{R}^n$. Escribamos $u = (u_1, u_2, \ldots, u_n)$, $v = (v_1, v_2, \ldots, v_n)$ y $w = (w_1, w_2, \ldots, w_n)$. Entonces

$$\begin{array}{rcl}
\langle u + v, w\rangle & = & \langle (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n), (w_1, w_2, \ldots, w_n)\rangle = \\
& = & (u_1 + v_1)w_1 + (u_2 + v_2)w_2 + \ldots + (u_n + v_n)w_n = \\
& = & u_1 w_1 + v_1 w_1 + u_2 w_2 + v_2 w_2 + \ldots + u_n w_n + v_n w_n = \\
& = & (u_1 w_1 + u_2 w_2 + \ldots + u_n w_n) + (v_1 w_1 + v_2 w_2 + \ldots + v_n w_n) = \\
& = & \langle u, w\rangle + \langle v, w\rangle.
\end{array}$$

Las demostraciones de los ítems (2), (3), (4), (5) y (6) quedan como ejercicio para el lector. □

Ejemplo 1.19. Veamos cómo podemos usar el producto interno para calcular ángulos entre vectores.

En $\mathbb{R}^2$: Sean $v = (v_1, v_2)$ y $w = (w_1, w_2)$ vectores no nulos de $\mathbb{R}^2$. Supongamos que los vectores v y w se encuentran ubicados como muestra la siguiente figura:



Sea θ el ángulo que forman los vectores v y w y sean α el ángulo que forma el vector v con el semieje positivo de las x y β el ángulo que forma el vector w con el mismo semieje.

Entonces $\theta = \beta - \alpha$. Luego $\cos(\theta) = \cos(\beta - \alpha)$.

Es interesante mencionar que esta última igualdad vale siempre, independientemente de cómo estén ubicados los vectores v y w. Dejamos la demostración de este hecho como ejercicio para el lector.

Se tiene entonces que

$$\begin{aligned} \cos(\theta) &= \cos(\beta-\alpha) = \cos(\beta)\cos(\alpha) + \operatorname{sen}(\beta)\operatorname{sen}(\alpha) = \frac{w_1}{\|w\|}.\frac{v_1}{\|v\|} + \frac{w_2}{\|w\|}.\frac{v_2}{\|v\|} = \\ &= \frac{1}{\|v\|\|w\|}.(v_1w_1 + v_2w_2) = \frac{\langle v, w\rangle}{\|v\|\|w\|}. \end{aligned}$$

Luego

$$\cos(\theta) = \frac{\langle v, w\rangle}{\|v\|\|w\|}.$$

<u>En $\mathbb{R}^3$</u>: Sean $v = (v_1, v_2, v_3)$ y $w = (w_1, w_2, w_3)$ vectores no nulos de $\mathbb{R}^3$. Sea θ el ángulo que forman los vectores v y w. Por el teorema del coseno tenemos que

$$\|w - v\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos\theta \tag{1.3}$$

Ahora bien, notemos que

$$\begin{aligned} \|w - v\|^2 &= \langle w - v, w - v\rangle = \langle w, w - v\rangle - \langle v, w - v\rangle = \\ &= \langle w, w\rangle - \langle w, v\rangle - \langle v, w\rangle + \langle v, v\rangle = \|w\|^2 - 2\langle v, w\rangle + \|v\|^2. \end{aligned}$$

Por lo tanto, reemplazando en (1.3) se obtiene que

$$\|w\|^2 - 2\langle v, w\rangle + \|v\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos\theta$$

de donde

$$-2\langle v, w\rangle = -2\|v\|\|w\| \cos\theta.$$

Luego

$$\frac{-2\langle v, w\rangle}{-2\|v\|\|w\|} = \cos\theta$$

y obtenemos que

$$\cos(\theta) = \frac{\langle v, w\rangle}{\|v\|\|w\|}.$$

Observación 1.20. Sean v y w vectores no nulos de $\mathbb{R}^2$ o $\mathbb{R}^3$ y sea θ el ángulo que forman. Entonces v y w son perpendiculares si y sólo si $\cos\theta = 0$, lo cual, por lo visto antes, es equivalente a $\dfrac{\langle v, w\rangle}{\|v\|\|w\|} = 0$, que a su vez es equivalente a $\langle v, w\rangle = 0$.

Por lo tanto v y w son perpendiculares si y sólo si $\langle v, w\rangle = 0$.

La fórmula

$$\cos(\theta) = \frac{\langle v, w\rangle}{\|v\|\|w\|}$$

válida en $\mathbb{R}^2$ y $\mathbb{R}^3$ como vimos, sugiere que si $n \in \mathbb{N}$ y v y w son vectores de $\mathbb{R}^n$, podríamos definir el ángulo entre ellos a partir de la fórmula anterior. Pero para ello necesitamos que $\dfrac{\langle v, w\rangle}{\|v\|\|w\|}$ tenga siempre valores entre -1 y 1. Esto es cierto, como muestra el siguiente teorema.

Teorema 1.21 (Desigualdad de Cauchy-Schwartz)**.**
Sea $n \in \mathbb{N}$ y sean $v, w \in \mathbb{R}^n$. Entonces $|\langle v, w\rangle| \leq \|v\|\|w\|$.

Demostración. Si $w = 0$ entonces $\langle v, w\rangle = \langle v, 0\rangle = 0$ y $\|v\|\|w\| = \|v\|\|0\| = \|v\|.0 = 0$ y la desigualdad se satisface trivialmente.

Supongamos entonces que $w \neq 0$. Entonces $\|w\| \neq 0$. Sea $\lambda = \dfrac{\langle v, w\rangle}{\|w\|^2}$. Se tiene que

$$\begin{aligned}
0 \leq \|v - \lambda w\|^2 &= \langle v - \lambda w, v - \lambda w\rangle = \langle v, v\rangle + \langle v, -\lambda w\rangle + \langle -\lambda w, v\rangle + \langle -\lambda w, -\lambda w\rangle = \\
&= \langle v, v\rangle + (-\lambda)\langle v, w\rangle + (-\lambda)\langle w, v\rangle + (-\lambda)^2\langle w, w\rangle = \\
&= \langle v, v\rangle - 2\lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle = \|v\|^2 - 2\lambda\langle v, w\rangle + \lambda^2\|w\|^2 = \\
&= \|v\|^2 - 2\frac{\langle v, w\rangle}{\|w\|^2}\langle v, w\rangle + \left(\frac{\langle v, w\rangle}{\|w\|^2}\right)^2\|w\|^2 = \\
&= \|v\|^2 - 2\frac{\langle v, w\rangle^2}{\|w\|^2} + \frac{\langle v, w\rangle^2}{\|w\|^4}\|w\|^2 = \|v\|^2 - 2\frac{\langle v, w\rangle^2}{\|w\|^2} + \frac{\langle v, w\rangle^2}{\|w\|^2} = \\
&= \|v\|^2 - \frac{\langle v, w\rangle^2}{\|w\|^2}
\end{aligned}$$

Luego $0 \leq \|v\|^2 - \dfrac{\langle v, w\rangle^2}{\|w\|^2}$ y entonces $\dfrac{\langle v, w\rangle^2}{\|w\|^2} \leq \|v\|^2$.

De aquí se obtiene que

$$\langle v, w\rangle^2 \leq \|v\|^2\|w\|^2$$

y tomando raíz cuadrada tenemos que

$$\sqrt{\langle v, w\rangle^2} \leq \sqrt{\|v\|^2\|w\|^2}.$$

Luego,

$$|\langle v, w\rangle| \leq |\|v\||.|\|w\||$$

y como $\|v\| \geq 0$ y $\|w\| \geq 0$ se obtiene que $|\langle v, w\rangle| \leq \|v\|\|w\|$. □

Definición 1.22. Sea $n \in \mathbb{N}$ y sean $v, w \in \mathbb{R}^n - \{0\}$. Definimos el *ángulo* entre v y w como el único $\theta \in [0, \pi]$ tal que $\cos\theta = \dfrac{\langle v, w\rangle}{\|v\|\|w\|}$.

Notemos que

$$\left|\frac{\langle v, w\rangle}{\|v\|\|w\|}\right| = \frac{|\langle v, w\rangle|}{\|v\|\|w\|} \leq 1$$

por la desigualdad de Cauchy-Schwartz.

Luego

$$-1 \leq \frac{\langle v, w\rangle}{\|v\|\|w\|} \leq 1$$

y por lo tanto existe un único $\theta \in [0, \pi]$ tal que $\cos\theta = \dfrac{\langle v, w\rangle}{\|v\|\|w\|}$.

Así, el ángulo entre v y w está bien definido.

Definición 1.23. Sea $n \in \mathbb{N}$ y sean $v, w \in \mathbb{R}^n - \{0\}$. Decimos que v y w son *perpendiculares* (u *ortogonales*) si el ángulo entre ellos es $\frac{\pi}{2}$.

Observación 1.24. Dado que $\cos(\frac{\pi}{2}) = 0$ se tiene que dos vectores no nulos v, w de $\mathbb{R}^n$ son perpendiculares si y sólo si $\langle v, w\rangle = 0$.

Ejemplo 1.25. Hallar un vector de $\mathbb{R}^2$ perpendicular al vector $(2, 1)$ y que tenga longitud 1. ¿Es único? Graficar.

Solución:

Sea (a, b) el vector buscado. Como (a, b) debe ser perpendicular a $(2, 1)$ entonces

$$\langle (a, b), (2, 1) \rangle = 0.$$

Luego $2a + b = 0$, de donde

$$b = -2a \tag{1.4}$$

Además, queremos que el vector (a, b) tenga longitud 1. Luego

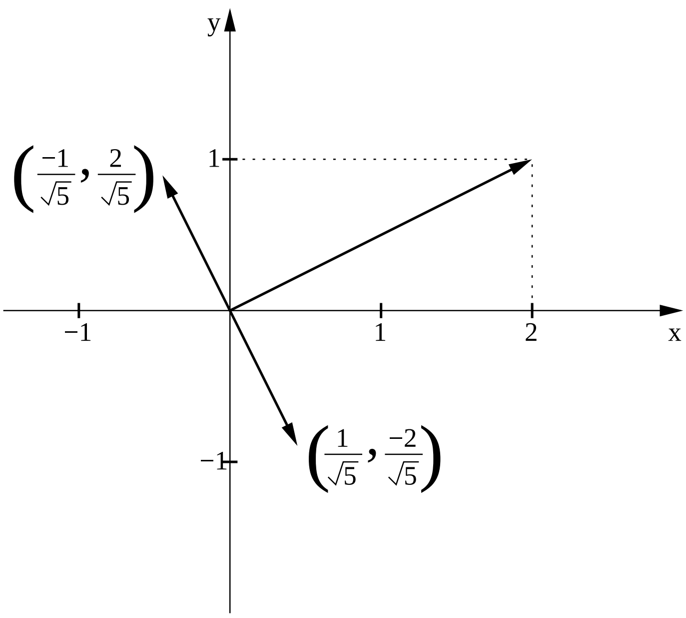
$$\begin{aligned} \|(a, b)\| &= 1 \\ \sqrt{a^2 + b^2} &= 1 \\ a^2 + b^2 &= 1 \end{aligned}$$

y reemplazando (1.4) se obtiene que

$$\begin{aligned} a^2 + (-2a)^2 &= 1 \\ a^2 + 4a^2 &= 1 \\ 5a^2 &= 1 \\ a^2 &= \frac{1}{5} \\ a &= \pm\frac{1}{\sqrt{5}} \end{aligned}$$

Si $a = \frac{1}{\sqrt{5}}$ entonces $b = -\frac{2}{\sqrt{5}}$ y $(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}})$ es un posible vector que cumple lo pedido.

También, si $a = -\frac{1}{\sqrt{5}}$ entonces $b = \frac{2}{\sqrt{5}}$ y el vector $(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$ también cumple lo pedido. Por lo tanto, no hay un único vector que cumpla lo pedido. De hecho, son dos.



Proposición 1.26 (Desigualdad triangular)**.** *Sea $n \in \mathbb{N}$ y sean $v, w \in \mathbb{R}^n$. Entonces $\|v + w\| \leq \|v\| + \|w\|$.*

Demostración. Tenemos que

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \leq \\ &\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2 \end{aligned}$$

donde la primer desigualdad vale pues $a \leq |a|$ para todo $a \in \mathbb{R}$ y la segunda vale por la desigualdad de Cauchy-Schwartz.

Entonces

$$\|v + w\|^2 \leq (\|v\| + \|w\|)^2.$$

Luego

$$\sqrt{\|v + w\|^2} \leq \sqrt{(\|v\| + \|w\|)^2}$$

y como las normas son no negativas se obtiene que

$$\|v + w\| \leq \|v\| + \|w\|.$$

□

Observación 1.27. Sea $n \in \mathbb{N}$ y sean $v, w \in \mathbb{R}^n$. Entonces $\|v - w\| \leq \|v\| + \|w\|$. En efecto,

$$\|v - w\| = \|v + (-w)\| \leq \|v\| + \|-w\| = \|v\| + \|(-1)w\| = \|v\| + |-1|\|w\| = \|v\| + \|w\|.$$

Proposición 1.28. *Sea $n \in \mathbb{N}$ y sean $v, w \in \mathbb{R}^n$. Entonces $|\|v\| - \|w\|| \leq \|v - w\|$.*

Demostración. Se tiene que

$$\|v\| = \|v - w + w\| \leq \|v - w\| + \|w\|$$

por la desigualdad triangular. Luego

$$\|v\| - \|w\| \leq \|v - w\|. \tag{1.5}$$

Por otra parte,

$$\|w\| = \|w - v + v\| \leq \|w - v\| + \|v\| = \|v - w\| + \|v\|$$

donde la desigualdad vale por la desigualdad triangular. Luego

$$\|w\| - \|v\| \leq \|v - w\|$$

y multiplicando por (-1) ambos lados de la desigualdad se obtiene que

$$-(\|w\| - \|v\|) \geq -\|v - w\|$$

es decir

$$\|v\| - \|w\| \geq -\|v - w\|. \tag{1.6}$$

De las ecuaciones (1.5) y (1.6) se obtiene que

$$-\|v - w\| \leq \|v\| - \|w\| \leq \|v - w\|.$$

y por lo tanto $|\|v\| - \|w\|| \leq \|v - w\|$. □

Enunciamos ahora una proposición con mucho contenido geométrico. En particular, el segundo ítem es una generalización del teorema de Pitágoras.

Proposición 1.29. *Sea $n \in \mathbb{N}$ y sean $A, B \in \mathbb{R}^n$. Entonces*

(1) $\|A - B\| = \|A + B\|$ si y sólo si $\langle A, B\rangle = 0$.

(2) $\|A + B\|^2 = \|A\|^2 + \|B\|^2$ si y sólo si $\langle A, B\rangle = 0$.

La demostración de esta proposición queda como ejercicio para el lector.

Producto vectorial

Definición 1.30. Sean $A = (a_1, a_2, a_3)$ y $B = (b_1, b_2, b_3)$ vectores de $\mathbb{R}^3$. Definimos el *producto vectorial* de A y B por $A \times B = (a_2b_3 - a_3b_2, -(a_1b_3 - a_3b_1), a_1b_2 - a_2b_1)$.

Ejemplo 1.31. Sean $A = (3, -1, 2)$ y $B = (-2, 1, 4)$. Calcular $A \times B$, $B \times A$ y $A \times A$.
Solución:

$$\begin{aligned} A \times B &= (-1.4 - 2.1\ ,\ -(3.4 - (-2).2)\ ,\ 3.1 - (-2).(-1)) = \\ &= (-4 - 2, -(12 + 4), 3 - 2) = (-6, -16, 1) \end{aligned}$$

$$\begin{aligned} B \times A &= (1.2 - (-1).4\ ,\ -((-2).2 - 3.4)\ ,\ (-2).(-1) - 3.1) = \\ &= (2 + 4, -(-4 - 12), 2 - 3) = (6, 16, -1) \end{aligned}$$

$$\begin{aligned} A \times A &= (-1.2 - (-1).2\ ,\ -(3.2 - 3.2)\ ,\ 3.(-1) - 3.(-1)) = \\ &= (0, 0, 0) \end{aligned}$$

Veamos algunas propiedades del producto vectorial.

Proposición 1.32. *Sean $A, B, C \in \mathbb{R}^3$ y sea $\alpha \in \mathbb{R}$. Entonces*

(1) $B \times A = -A \times B$

(2) $A \times (B + C) = A \times B + A \times C$

(3) $(B + C) \times A = B \times A + C \times A$

(4) $(\alpha.A) \times B = \alpha.(A \times B) = A \times (\alpha.B)$

(5) $A \times A = 0$

(6) $A \times B \perp A$ *y* $A \times B \perp B$

(7) Si $A \neq 0$ y $B \neq 0$ y θ es el ángulo entre los vectores A y B entonces

$$\|A \times B\| = \|A\|.\|B\|.\operatorname{sen}\theta\ .$$

Demostración.
(1) Tenemos que $A \times B = (a_2b_3 - a_3b_2, -(a_1b_3 - a_3b_1), a_1b_2 - a_2b_1)$. Por otra parte, $B \times A = (b_2a_3 - b_3a_2, -(b_1a_3 - b_3a_1), b_1a_2 - b_2a_1) = -A \times B$.
(6) Se tiene que

$$\begin{aligned} \langle A \times B, A\rangle &= \langle (a_2b_3 - a_3b_2, -(a_1b_3 - a_3b_1), a_1b_2 - a_2b_1), (a_1, a_2, a_3)\rangle = \\ &= (a_2b_3 - a_3b_2)a_1 - (a_1b_3 - a_3b_1)a_2 + (a_1b_2 - a_2b_1)a_3 = \\ &= a_1a_2b_3 - a_1a_3b_2 - a_1a_2b_3 + a_2a_3b_1 + a_1a_3b_2 - a_2a_3b_1 = 0 \end{aligned}$$

La igualdad $\langle A \times B, B\rangle = 0$ se demuestra en forma similar y queda como ejercicio para el lector.

(7) Se tiene que

$\|A \times B\|^2 = (a_2b_3 - a_3b_2)^2 + (a_1b_3 - a_3b_1)^2 + (a_1b_2 - a_2b_1)^2 =$
$= a_2^2b_3^2 - 2a_2b_3a_3b_2 + a_3^2b_2^2 + a_1^2b_3^2 - 2a_1b_3a_3b_1 + a_3^2b_1^2 + a_1^2b_2^2 - 2a_1b_2a_2b_1 + a_2^2b_1^2 =$
$= a_1^2b_1^2 + a_1^2b_2^2 + a_1^2b_3^2 + a_2^2b_1^2 + a_2^2b_2^2 + a_2^2b_3^2 + a_3^2b_1^2 + a_3^2b_2^2 + a_3^2b_3^2 - a_1^2b_1^2 - a_2^2b_2^2 - a_3^2b_3^2 +$
$-2a_1b_2a_2b_1 - 2a_1b_3a_3b_1 - 2a_2b_3a_3b_2 =$
$= a_1^2(b_1^2 + b_2^2 + b_3^2) + a_2^2(b_1^2 + b_2^2 + b_3^2) + a_3^2(b_1^2 + b_2^2 + b_3^2) + (-a_1^2b_1^2 - a_1b_2a_2b_1 +$
$-a_1b_3a_3b_1 - a_2^2b_2^2 - a_1b_2a_2b_1 - a_2b_3a_3b_2 - a_3^2b_3^2 - a_1b_3a_3b_1 - a_2b_3a_3b_2) =$
$= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) + (-1)(a_1b_1(a_1b_1 + a_2b_2 + a_3b_3) +$
$+a_2b_2(a_1b_1 + a_2b_2 + a_3b_3) + a_3b_3(a_1b_1 + a_2b_2 + a_3b_3)) =$
$= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) + (-1)(a_1b_1 + a_2b_2 + a_3b_3)(a_1b_1 + a_2b_2 + a_3b_3) =$
$= \|A\|^2.\|B\|^2 - \langle A, B\rangle^2 = \|A\|^2\|B\|^2 - (\|A\|\|B\| \cos\theta)^2 =$
$\|A\|^2\|B\|^2 - \|A\|^2\|B\|^2 \cos^2\theta = \|A\|^2\|B\|^2(1 - \cos^2\theta) = \|A\|^2\|B\|^2 \operatorname{sen}^2\theta.$

Entonces $\|A\times B\|^2 = \|A\|^2\|B\|^2 \operatorname{sen}^2\theta$ y como $\|A\times B\|$, $\|A\|$ y $\|B\|$ son mayores o iguales a cero, tomando raíz cuadrada a ambos miembros se obtiene que

$$\|A \times B\| = \|A\|\|B\||\operatorname{sen}\theta|.$$

Ahora bien, como $\theta \in [0, \pi]$ entonces $\operatorname{sen}\theta \geq 0$ y por lo tanto obtenemos que

$$\|A \times B\| = \|A\|\|B\| \operatorname{sen}\theta.$$

Las demostraciones de (2), (3), (4) y (5) se dejan como ejercicio para el lector. □

Observación 1.33. Si $A, B \in \mathbb{R}^3 - \{0\}$ son vectores con direcciones distintas, entonces hay una única dirección (recta) perpendicular a ambas. Como $A\times B \perp A$ y $A\times B \perp B$ entonces $A \times B$ debe tener esa dirección perpendicular a las dos primeras. Sabemos, además, que $\|A \times B\| = \|A\|\|B\| \operatorname{sen}\theta$, donde θ es el ángulo que forman los vectores A y B. Entonces la norma de $A \times B$ queda determinada conociendo A y B. Por lo tanto, para conocer el vector $A \times B$ falta saber su sentido. El sentido del vector $A \times B$ viene dado por la regla de la mano derecha.

Rectas y planos

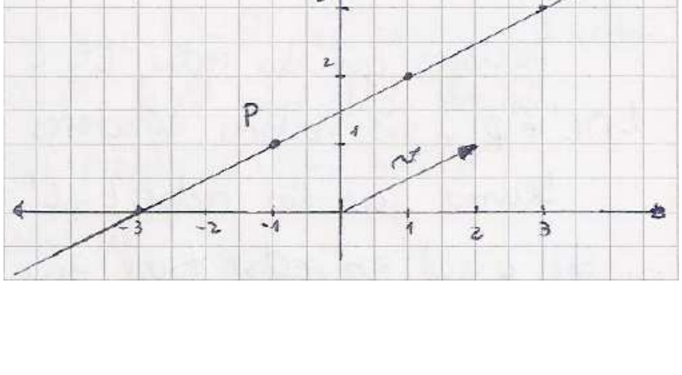
Sea $v \in \mathbb{R}^n - \{0\}$ y sea $P \in \mathbb{R}^n$. La *ecuación paramétrica* de la recta L paralela a la dirección de v que pasa por el punto P es

$$L : X = t.v + P,\ t \in \mathbb{R}.$$

Es decir, los puntos de la recta son los puntos de $\mathbb{R}^n$ de la forma $t.v + P$ donde $t \in \mathbb{R}$. El vector v se llama *vector director* de la recta L.

Ejemplo 1.34. Sea $v = (2, 1)$ y $P = (-1, 1)$. Sea L la recta paralela a la dirección de v y que pasa por el punto P. La ecuación paramétrica de la recta L es $L : X = t.(2, 1)+(-1, 1)$, $t \in \mathbb{R}$.

Grafiquemos la recta L.



Busquemos ahora el punto de la recta L que está más cerca del punto $O = (0,0)$. Para ello, notemos que si Q es el punto de L más cercano a O, entonces $\overline{QO} \perp L$. Luego, $\langle O - Q, v\rangle = 0$. Como $Q \in L$, entonces $Q = t.(2,1) + (-1,1)$ para algún $t \in \mathbb{R}$. Tenemos entonces que

$$\begin{array}{rcl} \langle O - Q, v\rangle & = & 0 \\ \langle (0,0) - (2t-1, t+1), (2,1)\rangle & = & 0 \\ \langle (-2t+1, -t-1), (2,1)\rangle & = & 0 \\ -4t + 2 - t - 1 & = & 0 \\ -5t + 1 & = & 0 \\ t & = & \dfrac{1}{5} \end{array}$$

Luego
$$Q = \frac{1}{5}(2,1) + (-1,1) = \left(-\frac{3}{5}, \frac{6}{5}\right).$$

Notemos que Q es efectivamente el punto de L que está más cerca de O, pues si X es un punto de L distinto de Q entonces el triángulo $\triangle OQX$ es rectángulo en Q, con lo cual $\mathrm{long}(\overline{OX}) > \mathrm{long}(\overline{OQ})$ puesto que $\overline{OX}$ es la hipotenusa del triángulo rectángulo $\triangle OQX$ y $\overline{OQ}$ es uno de sus catetos.

La distancia del punto O a la recta L es entonces

$$d(O,L) = d(O,Q) = \sqrt{\left(-\frac{3}{5}\right)^2 + \left(\frac{6}{5}\right)^2} = \sqrt{\frac{9}{25} + \frac{36}{25}} = \sqrt{\frac{45}{25}} = \frac{3}{5}\sqrt{5}.$$

Ejemplo 1.35. Sean $v, w \in \mathbb{R}^2$ con $v \neq (0,0)$. Sea L la recta que tiene a v como vector director y que pasa por el origen. Busquemos el punto z de la recta L tal que el segmento $\overline{zw}$ sea perpendicular a L.

Solución:

La ecuación paramétrica de la recta L es $L: \ X = t.v, \ t \in \mathbb{R}$. Por lo tanto, $z = t.v$ para algún $t \in \mathbb{R}$. Como queremos, además, que $\overline{zw}$ sea perpendicular a L entonces

$$\langle w - z, v\rangle = 0$$

de donde
$$\langle w - tv, v\rangle = 0$$

Luego

$$\begin{array}{rcl} \langle w, v\rangle - \langle tv, v\rangle & = & 0 \\ \langle w, v\rangle - t\langle v, v\rangle & = & 0 \\ \langle w, v\rangle & = & t\langle v, v\rangle \\ \langle w, v\rangle & = & t\|v\|^2 \end{array} \tag{1.7}$$

y como $v \neq 0$ entonces $\|v\| \neq 0$ y obtenemos que $t = \dfrac{\langle v, w\rangle}{\|v\|^2}$. Por lo tanto,

$$z = \frac{\langle v, w\rangle}{\|v\|^2}.v$$

El vector z se llama *proyección ortogonal de w sobre v* y se nota $p_v(w)$.

Esta misma definición se puede dar en $\mathbb{R}^n$:

Definición 1.36. Sean $v, w \in \mathbb{R}^n$ con $v \neq 0$. Definimos la *proyección ortogonal de w sobre v* por $p_v(w) = \dfrac{\langle v, w\rangle}{\|v\|^2}.v$.

Definición 1.37.

- Decimos que dos rectas son *paralelas* si sus vectores directores son paralelos, es decir, si uno de ellos es múltiplo del otro.
- Decimos que dos rectas L y L' son *perpendiculares* si $L \cap L' \neq \varnothing$ y sus vectores directores son perpendiculares.
- Decimos que dos rectas L y L' son *alabeadas* si $L \cap L' = \varnothing$ y L y L' no son paralelas.

Planos de $\mathbb{R}^3$

Sea $N \in \mathbb{R}^3 - \{0\}$ y sea $A \in \mathbb{R}^3$. La ecuación del plano perpendicular al vector N que pasa por el punto A es $\langle X - A, N\rangle = 0$. Si escribimos $X = (x, y, z)$ y $N = (a, b, c)$ entonces la ecuación queda:

$$\begin{aligned}\langle X - A, N\rangle &= 0\\ \langle X, N\rangle - \langle A, N\rangle &= 0\\ \langle X, N\rangle &= \langle A, N\rangle\\ ax + by + cz &= \langle A, N\rangle \qquad (1.8)\end{aligned}$$

Llamando $d = \langle A, N\rangle$, tenemos que $ax + by + cz = d$.
El vector N se llama vector *normal* al plano.

Ejemplo 1.38. Hallar la ecuación del plano perpendicular al vector $(0, 0, 1)$ y que pase por el punto $(1, 1, 2)$. Graficar.
Solución:

$$\begin{aligned}\langle X - A, N\rangle &= 0\\ \langle (x, y, z), (0, 0, 1)\rangle &= \langle (1, 1, 2), (0, 0, 1)\rangle\\ z &= 2\end{aligned}$$

Ejemplo 1.39. Sea $N = (1, -1, 3)$ y sea $A = (0, 2, 1)$. Sea Π el plano de vector normal N que pasa por el punto A. Sea $P = (1, 0, 1)$. Hallar el punto Q del plano Π que esté a menor distancia de P y calcular $d(P, \Pi)$.
Solución:
Buscamos $Q \in \Pi$ tal que $Q - P$ sea perpendicular a Π, es decir, buscamos $Q \in \Pi$ tal que $Q - P$ sea paralelo al vector N. Queremos entonces que $Q - P = kN$ para algún $k \in \mathbb{R}$. Luego,

$$Q = kN + P = k(1, -1, 3) + (1, 0, 1) = (k + 1, -k, 3k + 1).$$

Calculemos la ecuación del plano Π:

$$\begin{aligned}\langle X - A, N\rangle &= 0\\ \langle X, N\rangle &= \langle A, N\rangle\\ \langle (x, y, z), (1, -1, 3)\rangle &= \langle (0, 2, 1), (1, -1, 3)\rangle\\ x - y + 3z &= 1\end{aligned}$$

Como queremos que $Q \in \Pi$ planteamos

$$\begin{aligned} k+1-(-k)+3(3k+1) &= 1 \\ 11k+4 &= 1 \\ k &= -\frac{3}{11} \end{aligned}$$

Entonces $Q = (k+1, -k, 3k+1) = \left(-\frac{3}{11}+1, -(-\frac{3}{11}), 3(-\frac{3}{11})+1\right) = \left(\frac{8}{11}, \frac{3}{11}, \frac{2}{11}\right)$.

Así, $Q = \left(\frac{8}{11}, \frac{3}{11}, \frac{2}{11}\right)$ es el punto buscado.

Notemos que si X es otro punto del plano Π, distinto de Q, entonces el triángulo $\triangle PQX$ es rectángulo en Q, pues el segmento $\overline{PQ}$ es paralelo a N (pues $Q - P \parallel N$) y el segmento $\overline{QX}$ es perpendicular a N ya que X y Q son puntos del plano. Entonces $\mathrm{long}(\overline{PX}) > \mathrm{long}(\overline{PQ})$ ya que $\overline{PX}$ es la hipotenusa del triángulo rectángulo $\triangle PQX$ y $\overline{PQ}$ es uno de sus catetos.

La distancia del punto P al plano Π es entonces

$$d(P,\Pi) = d(P,Q) = \sqrt{\left(\frac{8}{11}-1\right)^2 + \left(\frac{3}{11}-0\right)^2 + \left(\frac{2}{11}-1\right)^2} = \sqrt{\frac{99}{11^2}} = \frac{\sqrt{99}}{11}.$$

Proposición 1.40. *Sea $N \in \mathbb{R}^3 - \{0\}$ y sea $A \in \mathbb{R}^3$. Sea Π el plano perpendicular a la dirección de N que contiene al punto A. Sea $P \in \mathbb{R}^3$. Entonces la distancia del punto P al plano Π es*

$$d(P,\Pi) = \frac{|\langle P-A, N\rangle|}{\|N\|}.$$

En particular, si $P = (x_0, y_0, z_0)$ y Π es el plano de ecuación $ax+by+cz=d$ entonces se obtiene que

$$d(P,\Pi) = \frac{|ax_0+by_0+cz_0-d|}{\sqrt{a^2+b^2+c^2}}.$$

Demostración. Buscamos $Q \in \Pi$ tal que $Q - P \parallel N$. Escribimos entonces $Q - P = kN$ para algún $k \in \mathbb{R}$, de donde $Q = kN + P$ y como queremos que $Q \in \Pi$, pedimos que $\langle Q - A, N\rangle = 0$. Luego

$$\begin{aligned} \langle k.N + P - A, N\rangle &= 0 \\ \langle k.N, N\rangle + \langle P - A, N\rangle &= 0 \\ k\langle N, N\rangle &= -\langle P-A, N\rangle \\ k &= -\frac{\langle P-A, N\rangle}{\|N\|^2} \end{aligned}$$

Entonces,

$$Q = k.N + P = -\frac{\langle P-A, N\rangle}{\|N\|^2}.N + P.$$

Veamos que Q es el punto de Π más cercano a P. Sea $X \in \Pi$ tal que $X \neq Q$. Afirmo que $P - Q \perp X - Q$. En efecto,

$$\begin{aligned} \langle P-Q, X-Q\rangle &= \langle k.N, X-Q\rangle = k.\langle N, X-Q\rangle = k(\langle N, X\rangle - \langle N, Q\rangle) = \\ &= k(\langle A, N\rangle - \langle A, N\rangle) = 0 \end{aligned}$$

donde la penúltima igualdad vale pues $X, Q \in \Pi$.

Por lo tanto, por el ítem (2) de la proposición 1.29 se tiene que

$$\|P-X\|^2 = \|(P-Q)-(X-Q)\|^2 = \|P-Q\|^2 + \|X-Q\|^2 > \|P-Q\|^2$$

donde la última desigualdad se cumple porque $\|X - Q\|^2 > 0$ ya que $X \neq Q$.

Entonces, $\|P - X\| > \|P - Q\|$. Es decir, $d(P, X) > d(P, Q)$. Así, Q es el punto de Π que minimiza la distancia a Π.

Calculemos ahora la distancia de P a Π.

$$\begin{aligned} d(P, \Pi) &= d(P, Q) = \|Q - P\| = \|k.N + P - P\| = \|k.N\| = |k|.\|N\| = \\ &= \left| -\frac{\langle P - A, N\rangle}{\|N\|^2} \right| .\|N\| = \frac{|\langle P - A, N\rangle|}{\|N\|^2}.\|N\| \end{aligned}$$

Así,

$$d(P, \Pi) = \frac{|\langle P - A, N\rangle|}{\|N\|^2}.\|N\| \ .$$

Observemos que si $P = (x_0, y_0, z_0), A = (a_1, a_2, a_3)$ y Π es el plano de ecuación

$$ax + by + cz = d$$

entonces podemos tomar $N = (a, b, c)$ y tenemos que $aa_1 + ba_2 + ca_3 = d$ pues $A \in \Pi$. Luego,

$$\begin{aligned} d(P, \Pi) &= \frac{|\langle P - A, N\rangle|}{\|N\|} = \frac{|\langle P, N\rangle - \langle A, N\rangle|}{\|N\|} = \\ &= \frac{|ax_0 + by_0 + cz_0 - (aa_1 + ba_2 + ca_3)|}{\sqrt{a^2 + b^2 + c^2}} = \\ &= \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}. \end{aligned}$$

$\square$

Capítulo 2

Sistemas de ecuaciones lineales y matrices

Definición 2.1. Sean $m, n \in \mathbb{N}$. Un *sistema de m ecuaciones lineales con n incógnitas* es un conjunto de m ecuaciones lineales en n variables, que escribiremos de la siguiente manera

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

donde $a_{ij} \in \mathbb{R}$ para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$, $b_i \in \mathbb{R}$ para todo $i \in \{1, 2, \dots, m\}$ y donde $x_1, x_2, \dots, x_n$ son las incógnitas.

Los números a_{ij} se llaman *coeficientes* del sistema.

Decimos que $z = (z_1, z_2, \dots, z_n) \in \mathbb{R}^n$ es solución el sistema anterior si al reemplazar cada x_i por z_i se satisfacen las m ecuaciones.

Si $b_i = 0$ para todo $i \in \{1, 2, \dots, m\}$ diremos que el sistema es *homogéneo.* En caso contrario decimos que el sistema es no homogéneo.

Veamos un ejemplo.

Ejemplo 2.2. Resolver el siguiente sistema de ecuaciones:

$$\left\{ \begin{array}{ccccccc} x & + & y & + & z & = & 2 \\ 2x & - & y & + & z & = & -3 \\ x & + & 2y & + & 3z & = & 1 \end{array} \right. \qquad \begin{array}{c} (1) \\ (2) \\ (3) \end{array}$$

Solución:
De (1) obtenemos que

$$z = 2 - x - y \qquad (4)$$

Reemplazando en (2) queda

$$\begin{array}{rcl} 2x - y + (2 - x - y) & = & -3 \\ x - 2y & = & -5 \\ -2y & = & -5 - x \\ y & = & \frac{5}{2} + \frac{1}{2}x \qquad (5) \end{array}$$

Reemplazando ahora la ecuación (5) en la ecuación (4) obtenemos

$$\begin{array}{rcll} z & = & 2 - x - (\frac{5}{2} + \frac{1}{2}x) & \\ z & = & -\frac{1}{2} - \frac{3}{2}x & \qquad (6) \end{array}$$

Y reemplazando (5) y (6) en la ecuación (3) queda

$$\begin{array}{rcl} x + 2(\frac{5}{2} + \frac{1}{2}x) + 3(-\frac{1}{2} - \frac{3}{2}x) & = & 1 \\ x + 5 + x - \frac{3}{2} - \frac{9}{2}x & = & 1 \\ -\frac{5}{2}x & = & 1 - 5 + \frac{3}{2} \\ -\frac{5}{2}x & = & -\frac{5}{2} \\ x & = & 1 \end{array}$$

Luego, utilizando las igualdades (5) y (6) se obtiene que

$$y = \frac{5}{2} + \frac{1}{2}.1 = \frac{5}{2} + \frac{1}{2} = 3$$

y

$$z = -\frac{1}{2} - \frac{3}{2}.1 = -\frac{1}{2} - \frac{3}{2} = -2$$

Por lo tanto, la solución del sistema es $(1, 3, -2)$ (es decir, $x = 1$, $y = 3$, $z = -2$).

Método de eliminación de Gauss

Definición 2.3. Decimos que dos sistemas de ecuaciones lineales son *equivalentes* si tienen las mismas soluciones.

Ejemplo 2.4. Los sistemas

$$\left\{ \begin{array}{rcrcl} x & + & 2y & = & 1 \\ x & + & y & = & 1 \end{array} \right. \qquad \text{y} \qquad \left\{ \begin{array}{rcrcl} x & + & 2y & = & 1 \\ & & y & = & 0 \end{array} \right.$$

son equivalentes. En efecto, es claro que el segundo sistema tiene solución única $(1, 0)$. Por otra parte, restando las dos ecuaciones del primer sistema se obtiene

$$\begin{array}{rcl} x + 2y - (x + y) & = & 1 - 1 \\ x + 2y - x - y & = & 0 \\ y & = & 0 \end{array}$$

y reemplazando esto en la segunda ecuación obtenemos $x = 1$. Por lo tanto, el primer sistema también tiene solución única $(1, 0)$.

Desarrollaremos ahora el método de eliminación de Gauss que nos permitirá llevar un sistema de ecuaciones lineales a otro equivalente a él y más fácil de resolver.

Supongamos que tenemos un sistema $\mathcal{S}$ de m ecuaciones lineales con n incógnitas:

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

Notaremos $x = (x_1, x_2, \ldots, x_n)$. Además, para $i \in \{1, 2, \ldots, m\}$, definimos

$$F_i(x) = F_i(x_1, x_2, \ldots, x_n) = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n.$$

También para $i \in \{1, 2, \ldots, m\}$ notaremos por E_i a la i-ésima ecuación y escribiremos

$$E_i : \quad F_i(x) = b_i.$$

Con estas notaciones definiremos ahora tres tipos de transformaciones que se pueden aplicar al sistema de ecuaciones lineales $\mathcal{S}$ para obtener otro sistema de ecuaciones lineales $\mathcal{S}'$ cuyas ecuaciones serán $E'_1, E'_2, \ldots, E'_m$.

(1) Intercambiar dos ecuaciones. Consiste en elegir $j, k \in \{1, 2, \ldots, m\}$ distintos y definir

$$\begin{array}{rcll} E'_j & = & E_k & \\ E'_k & = & E_j & \\ E'_i & = & E_i & \quad \text{para } i \in \{1, 2, \ldots, m\} - \{j, k\}. \end{array}$$

Notaremos $E_j \leftrightarrow E_k$.

Como ejemplo, notemos que si $j < k$ esta transformación sería algo así:

$$\left\{\begin{array}{ccccc} E_1 : & F_1(x) & = & b_1 \\ E_2 : & F_2(x) & = & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ E_{j-1} : & F_{j-1}(x) & = & b_{j-1} \\ E_j : & F_j(x) & = & b_j \\ E_{j+1} : & F_{j+1}(x) & = & b_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ E_{k-1} : & F_{k-1}(x) & = & b_{k-1} \\ E_k : & F_k(x) & = & b_k \\ E_{k+1} : & F_{k+1}(x) & = & b_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ E_m : & F_m(x) & = & b_m \end{array}\right. \quad \underset{E_j \leftrightarrow E_k}{\rightsquigarrow} \quad \left\{\begin{array}{cccc} E'_1 : & F_1(x) & = & b_1 \\ E'_2 : & F_2(x) & = & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ E'_{j-1} : & F_{j-1}(x) & = & b_{j-1} \\ E'_j = E_k : & F_k(x) & = & b_k \\ E'_{j+1} : & F_{j+1}(x) & = & b_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ E'_{k-1} : & F_{k-1}(x) & = & b_{k-1} \\ E'_k = E_j : & F_j(x) & = & b_j \\ E'_{k+1} : & F_{k+1}(x) & = & b_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ E'_m : & F_m(x) & = & b_m \end{array}\right.$$

(2) Multiplicar una ecuación por un número real no nulo. Consiste en elegir $j \in \{1, 2, \ldots, m\}$ y $\alpha \in \mathbb{R} - \{0\}$ y definir

$$\begin{array}{rl} E'_j : & \alpha F_j(x) = \alpha b_j \\ E'_i = E_i & \text{para } i \in \{1, 2, \ldots, m\} - \{j\}. \end{array}$$

Notaremos $E_j \leftarrow \alpha.E_j$.

La transformación sería algo así:

$$\left\{\begin{array}{cccc} E_1 : & F_1(x) & = & b_1 \\ E_2 : & F_2(x) & = & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ E_{j-1} : & F_{j-1}(x) & = & b_{j-1} \\ E_j : & F_j(x) & = & b_j \\ E_{j+1} : & F_{j+1}(x) & = & b_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ E_m : & F_m(x) & = & b_m \end{array}\right. \quad \underset{E_j \leftarrow \alpha.E_j}{\rightsquigarrow} \quad \left\{\begin{array}{cccc} E'_1 : & F_1(x) & = & b_1 \\ E'_2 : & F_2(x) & = & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ E'_{j-1} : & F_{j-1}(x) & = & b_{j-1} \\ E'_j : & \alpha F_j(x) & = & \alpha b_j \\ E'_{j+1} : & F_{j+1}(x) & = & b_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ E'_m : & F_m(x) & = & b_m \end{array}\right.$$

(3) Sumar a una ecuación un múltiplo de otra ecuación. Consiste en elegir $\alpha \in \mathbb{R}$ y $j, k \in \{1, 2, \dots, m\}$ distintos y definir

$$\begin{array}{rl} E'_j : & F_j(x) + \alpha F_k(x) = b_j + \alpha b_k \\ E'_i = E_i & \text{para } i \in \{1, 2, \dots, m\} - \{j\}. \end{array}$$

Notaremos $E_j \leftarrow E_j + \alpha . E_k$.

Como ejemplo, notemos que si $j < k$ esta transformación sería algo así:

$$\left\{ \begin{array}{rccc} E_1 : & F_1(x) & = & b_1 \\ E_2 : & F_2(x) & = & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ E_{j-1} : & F_{j-1}(x) & = & b_{j-1} \\ E_j : & F_j(x) & = & b_j \\ E_{j+1} : & F_{j+1}(x) & = & b_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ E_{k-1} : & F_{k-1}(x) & = & b_{k-1} \\ E_k : & F_k(x) & = & b_k \\ E_{k+1} : & F_{k+1}(x) & = & b_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ E_m : & F_m(x) & = & b_m \end{array} \right. \quad \rightsquigarrow \quad \left\{ \begin{array}{rccc} E'_1 : & F_1(x) & = & b_1 \\ E'_2 : & F_2(x) & = & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ E'_{j-1} : & F_{j-1}(x) & = & b_{j-1} \\ E'_j : & F_j(x) + \alpha F_k(x) & = & b_j + \alpha b_k \\ E'_{j+1} : & F_{j+1}(x) & = & b_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ E'_{k-1} : & F_{k-1}(x) & = & b_{k-1} \\ E'_k : & F_k(x) & = & b_k \\ E'_{k+1} : & F_{k+1}(x) & = & b_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ E'_m : & F_m(x) & = & b_m \end{array} \right.$$

Proposición 2.5. *Sea $\mathcal{S}$ un sistema de ecuaciones lineales. Entonces cualquiera de las transformaciones definidas antes aplicada a $\mathcal{S}$ da como resultado un sistema de ecuaciones lineales equivalente a $\mathcal{S}$.*

Demostración. Utilizaremos las notaciones introducidas anteriormente. Es claro que cada una de las tres transformaciones anteriores aplicadas a $\mathcal{S}$ dan como resultado un sistema de ecuaciones lineales. Llamemos $\mathcal{S}'$ al sistema obtenido. Debemos probar que el sistema de ecuaciones $\mathcal{S}'$ es equivalente a $\mathcal{S}$, es decir, que los sistemas $\mathcal{S}$ y $\mathcal{S}'$ tienen las mismas soluciones.

Para la transformación (1) esto es evidente, pues el orden en que damos las ecuaciones de un sistema no afecta a sus soluciones. En efecto, $x \in \mathbb{R}^n$ es solución de $\mathcal{S}$ si y sólo si $F_1(x) = b_1 \wedge F_2(x) = b_2 \wedge \dots \wedge F_m(x) = b_m$ y como la conjunción lógica es conmutativa y asociativa, esto es equivalente a que x sea solución de $\mathcal{S}'$.

Para la transformación (2), dado que la ecuación E_i coincide con la ecuación E'_i para todo $i \in \{1, 2, \dots, m\} - \{j\}$, para probar que el sistema de ecuaciones obtenido es equivalente al original, debemos probar que $\forall\, x \in \mathbb{R}^n$, $F_j(x) = b_j \Leftrightarrow \alpha F_j(x) = \alpha b_j$.

Sea $x \in \mathbb{R}^n$. Si $F_j(x) = b_j$ entonces, multiplicando por α, se obtiene que $\alpha F_j(x) = \alpha b_j$. Recíprocamente, si $\alpha F_j(x) = \alpha b_j$ entonces, como $\alpha \neq 0$, dividiendo por α se obtiene que $F_j(x) = b_j$. Por lo tanto, los sistemas $\mathcal{S}$ y $\mathcal{S}'$ resultan equivalentes.

Resta analizar qué ocurre al aplicar la transformación (3). Veamos que en este caso también vale que, para todo $x \in \mathbb{R}^n$, x es solución del sistema $\mathcal{S}$ si y sólo si es solución del sistema $\mathcal{S}'$.

Sea $x \in \mathbb{R}^n$. Supongamos primero que x es solución del sistema $\mathcal{S}$. Entonces x satisface las ecuaciones $E_1, E_2, \dots, E_m$. Luego, x satisface las ecuaciones E'_i para todo $i \in \{1, 2, \dots, m\} - \{j\}$ ya que la ecuación E'_i es la misma que la ecuación E_i para todo $i \in \{1, 2, \dots, m\} - \{j\}$. Veamos que x satisface también la ecuación E'_j. Como x satisface las ecuaciones E_j y E_k entonces $F_j(x) = b_j$ y $F_k(x) = b_k$. Luego

$$F_j(x) + \alpha F_k(x) = b_j + \alpha b_k,$$

es decir, x satisface la ecuación E'_j. Por lo tanto, x es solución de $\mathcal{S}'$.

Supongamos ahora que x es solución del sistema $\mathcal{S}'$. Entonces x satisface las ecuaciones $E'_1, E'_2, \ldots, E'_m$. Luego, x satisface las ecuaciones E_i para $i \in \{1, 2, \ldots, m\} - \{j\}$ ya que la ecuación E_i es la misma que la ecuación E'_i para todo $i \in \{1, 2, \ldots, m\} - \{j\}$. Veamos que x satisface también la ecuación E_j. Como x satisface las ecuaciones E'_j y E'_k entonces

$$F_j(x) + \alpha F_k(x) = b_j + \alpha b_k \qquad \text{y} \qquad F_k(x) = b_k.$$

Luego

$$F_j(x) + \alpha F_k(x) - \alpha F_k(x) = b_j + \alpha b_k - \alpha b_k,$$

es decir, $F_j(x) = b_j$. Así, x satisface la ecuación E_j. Por lo tanto, x es solución de $\mathcal{S}$.

Luego, los sistemas $\mathcal{S}$ y $\mathcal{S}'$ resultan equivalentes. □

Definición 2.6. Decimos que un sistema de ecuaciones lineales se encuentra en *forma escalonada* si cumple las siguientes condiciones:

(1) Las ecuaciones que tengan todos sus coeficientes nulos se agrupan en la parte inferior del sistema.

(2) Para cada ecuación que tenga algún coeficiente no nulo se cumple que si el primer coeficiente no nulo de dicha ecuación es el de la variable x_j entonces en las siguientes ecuaciones todos los coeficientes de las variables $x_1, x_2, \ldots, x_j$ son nulos.

Ejemplo 2.7. El siguiente sistema se encuentra en forma escalonada:

$$\left\{\begin{array}{rcrcrcrcrcr} 2x_1 & + & 3x_2 & - & 4x_3 & + & x_4 & - & x_5 & = & 1 \\ & & & & x_3 & - & x_4 & + & x_5 & = & 2 \\ & & & & & & 5x_4 & + & 2x_5 & = & 10 \\ & & & & & & & & 0 & = & 0 \end{array}\right.$$

Ejemplo 2.8. Mediante la aplicación de las transformaciones de la proposición anterior, llevar el siguiente sistema de ecuaciones lineales a uno equivalente a él que se encuentre en forma escalonada y luego resolverlo.

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & = & 5 \\ 2x_1 & & & + & 2x_3 & = & 4 \\ -3x_1 & - & 7x_2 & + & 10x_3 & = & 5 \end{array}\right.$$

Solución:

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & = & 5 \\ 2x_1 & & & + & 2x_3 & = & 4 \\ -3x_1 & - & 7x_2 & + & 10x_3 & = & 5 \end{array}\right. \underset{E_2 \leftarrow E_2 - 2E_1}{\rightsquigarrow}$$

$$\rightsquigarrow \left\{\begin{array}{rcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & = & 5 \\ & - & 6x_2 & + & 6x_3 & = & -6 \\ -3x_1 & - & 7x_2 & + & 10x_3 & = & 5 \end{array}\right. \underset{E_3 \leftarrow E_3 + 3E_1}{\rightsquigarrow}$$

$$\rightsquigarrow \left\{\begin{array}{rcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & = & 5 \\ & - & 6x_2 & + & 6x_3 & = & -6 \\ & & 2x_2 & + & 4x_3 & = & 20 \end{array}\right. \underset{E_2 \leftarrow -\frac{1}{6}E_2}{\rightsquigarrow}$$

$$\rightsquigarrow \left\{\begin{array}{rcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & = & 5 \\ & & x_2 & - & x_3 & = & 1 \\ & & 2x_2 & + & 4x_3 & = & 20 \end{array}\right. \underset{E_3 \leftarrow E_3 - 2E_2}{\rightsquigarrow}$$

$$\rightsquigarrow \begin{cases} x_1 + 3x_2 - 2x_3 = 5 & (1) \\ \quad\quad x_2 - x_3 = 1 & (2) \\ \quad\quad\quad 6x_3 = 18 & (3) \end{cases}$$

Este último sistema se encuentra en forma escalonada, es equivalente al original y es mucho más fácil de resolver que éste. Resolvámoslo.

De (3) obtenemos que $x_3 = 3$. Reemplazando en (2) nos queda que $x_2 - 3 = 1$ de donde obtenemos que $x_2 = 4$. Finalmente, reemplazando los valores de x_2 y x_3 en x_1 queda

$$\begin{aligned} x_1 + 3.\,4 - 2.\,3 &= 5 \\ x_1 + 6 &= 5 \\ x_1 &= -1 \end{aligned}$$

Por lo tanto, la solución del sistema es $(-1, 4, 3)$.

Este método que consiste en aplicar las transformaciones de la proposición anterior para llevar un sistema de ecuaciones lineales dado a otro equivalente a él que se encuentre en forma escalonada se llama *método de eliminación de Gauss*.

No es difícil demostrar que mediante la aplicación de estas tres transformaciones cualquier sistema de ecuaciones lineales puede llevarse a la forma escalonada. En efecto, se pueden seguir los siguientes pasos:

(1) Si en la primera columna (la de los x_1) hay algún coeficiente no nulo, intercambiar ecuaciones para lograr que la primera ecuación tenga coeficiente de x_1 no nulo.

(2) Multiplicar la primera ecuación resultante por el inverso multiplicativo del coeficiente de x_1 para lograr que el coeficiente de x_1 de la primera ecuación sea 1.

(3) Utilizar este coeficiente 1 y la tercer transformación para conseguir que todos los otros coeficientes de la primera columna sean cero.

(4) Si todos los coeficientes de la primera columna eran cero, tomar la siguiente columna y aplicar los pasos (1), (2), (3) y (4) para dicha columna.

(5) Una vez obtenido el coeficiente 1 en la primera ecuación y los ceros debajo, dejar esta primera ecuación fija y aplicar el procedimiento descripto en los pasos (1) a (5) a las ecuaciones que siguen.

Proposición 2.9. *Sea $\mathcal{S}$ un sistema de ecuaciones lineales. Entonces $\mathcal{S}$ cumple alguna de las tres proposiciones siguientes:*

- *$\mathcal{S}$ tiene solución única.*
- *$\mathcal{S}$ tiene infinitas soluciones.*
- *$\mathcal{S}$ no tiene solución.*

Demostración. Mediante las transformaciones definidas anteriormente podemos llevar el sistema dado a uno equivalente a él que se encuentre en forma escalonada. Sea $\mathcal{S}'$ este nuevo sistema.

Llamaremos ecuaciones absurdas a las ecuaciones del tipo

$$0x_1 + 0x_2 + \ldots + 0x_n = b \qquad \text{con } b \neq 0$$

Si en el sistema $\mathcal{S}'$ hay alguna ecuación absurda entonces dicho sistema no tiene soluciones. Como el sistema $\mathcal{S}$ es equivalente al sistema $\mathcal{S}'$ entonces el sistema $\mathcal{S}$ tampoco tiene soluciones y estamos en el tercer caso.

Supongamos ahora que en el sistema $\mathcal{S}'$ no hay ninguna ecuación absurda. Llamemos ecuación trivial a una ecuación del tipo

$$0x_1 + 0x_2 + \ldots + 0x_n = 0.$$

Como el sistema se encuentra en forma escalonada entonces en la segunda ecuación el coeficiente de la variable x_1 es cero ya que el primer coeficiente no nulo de la primera ecuación es el de la variable x_1 o el de una variable posterior. Por lo tanto, en la segunda ecuación, el primer coeficiente no nulo es el de la variable x_2 o el de una variable posterior. Y dado que el sistema se encuentra en forma escalonada, en la tercer ecuación los coeficientes de las variables x_1 y x_2 son 0.

Continuando con este razonamiento, tenemos que en la i-ésima ecuación los coeficientes de las variables $x_1, x_2, \ldots, x_{i-1}$ son cero.

Sea n la cantidad de variables del sistema $\mathcal{S}'$. Veamos que en $\mathcal{S}'$ no puede haber más de n ecuaciones no triviales. En efecto, si $j > n$ y en el sistema $\mathcal{S}'$ hay al menos j ecuaciones, entonces en la j-ésima ecuación de $\mathcal{S}'$ los coeficientes de las variables $x_1, x_2, \ldots, x_n$ deben ser nulos, y como estamos suponiendo que el sistema $\mathcal{S}'$ no tiene ecuaciones absurdas entonces la j-ésima ecuación debe ser una ecuación trivial. Por lo tanto, en $\mathcal{S}'$ hay a lo sumo n ecuaciones no triviales.

Si el sistema $\mathcal{S}'$ tiene exactamente n ecuaciones no triviales entonces para todo $i \in \{1, 2, \ldots, n\}$ debe ocurrir que el primer coeficiente no nulo de la i-ésima ecuación sea x_i, ya que, de no ser así, aplicando el razonamiento anterior obtendríamos que el sistema $\mathcal{S}'$ tendría menos de n ecuaciones no triviales.

Por lo tanto, la n-ésima ecuación de $\mathcal{S}'$ es de la forma $kx_n = b$ con $k, b \in \mathbb{R}$ y $k \neq 0$. Despejando x_n de esta ecuación obtenemos un único valor posible para x_n. Reemplazando este valor en la $(n-1)$-ésima ecuación y despejando x_{n-1} (ya que su coeficiente en esta ecuación es no nulo) obtenemos un único valor posible para x_{n-1}. Ahora reemplazando los valores obtenidos de x_{n-1} y x_n en la $(n-2)$-ésima ecuación y despejando obtenemos un único valor posible para x_{n-2}.

Continuando con este procedimiento obtenemos únicos valores posibles para todas las variables $x_1, x_2, \ldots, x_n$ que verifican las primeras n ecuaciones del sistema $\mathcal{S}'$. Y como las ecuaciones posteriores a la n-ésima son triviales, concluimos que los valores hallados forman la única solución del sistema $\mathcal{S}'$. Por lo tanto, el sistema $\mathcal{S}$ tiene solución única.

Supongamos ahora que el sistema $\mathcal{S}'$ tiene menos de n ecuaciones no triviales. Llamemos variables principales a las que acompañan a los primeros coeficientes no nulos de cada ecuación y variables libres a las que no son variables principales. Sea k la cantidad de ecuaciones no triviales que tiene el sistema $\mathcal{S}'$. Entonces $k < n$. Además, el sistema $\mathcal{S}'$ tendrá k variables principales (una por cada ecuación no trivial) y $n - k$ variables libres. Despejando como en el caso anterior desde la última ecuación hasta la primera, es posible despejar las variables principales en términos de las variables libres.

Ahora bien, como $k < n$ entonces $n - k > 0$. Como $n, k \in \mathbb{N}$ entonces $n - k \in \mathbb{Z}$ y como $n - k > 0$ entonces $n - k \geq 1$. Por lo tanto, hay al menos una variable libre. Para cada valor de la(s) variable(s) libre(s) habrá una solución del sistema $\mathcal{S}'$. Por lo tanto, el sistema $\mathcal{S}'$ tiene infinitas soluciones y entonces el sistema $\mathcal{S}$ tiene infinitas soluciones. $\square$

Definición 2.10. Sea $\mathcal{S}$ un sistema de ecuaciones lineales.

- Si el sistema $\mathcal{S}$ tiene solución única diremos que $\mathcal{S}$ es un *sistema compatible determinado.*

- Si el sistema $\mathcal{S}$ tiene infinitas soluciones diremos que $\mathcal{S}$ es un *sistema compatible indeterminado.*

- Si el sistema $\mathcal{S}$ no tiene solución diremos que $\mathcal{S}$ es un *sistema incompatible.*

Observación importante 2.11. Supongamos que tenemos un sistema de ecuaciones lineales en forma escalonada.

- Si hay alguna ecuación del tipo $0x_1 + 0x_2 + \ldots + 0x_n = b$ con $b \neq 0$ (ecuación absurda), entonces el sistema es incompatible.

- Si no hay ecuaciones absurdas y hay tantas ecuaciones no triviales como incógnitas, entonces el sistema es compatible determinado.

- Si no hay ecuaciones absurdas y hay más incógnitas que ecuaciones no triviales, entonces el sistema es compatible indeterminado.

Matrices

Si al aplicar el método de eliminación de Gauss para resolver sistemas de ecuaciones lineales tenemos el cuidado de mantener el orden de las variables y escribir todos los coeficientes, incluso los que son cero, entonces no parece necesario escribir las variables x_i para aplicar el proceso de escalonamiento del sistema.

Por ejemplo, el sistema

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & = & 5 \\ 2x_1 & & & + & 2x_3 & = & 4 \\ -3x_1 & - & 7x_2 & + & 10x_3 & = & 5 \end{array}\right.$$

podría escribirse en forma abreviada de la siguiente manera:

$$\left\{\begin{array}{rrr|r} 1 & 3 & -2 & 5 \\ 2 & 0 & 2 & 4 \\ -3 & -7 & 10 & 5 \end{array}\right.$$

Esto da lugar a la siguiente definición:

Definición 2.12. Sean $m, n \in \mathbb{N}$. Una *matriz de números reales de tamaño* $m \times n$ es un arreglo rectangular de números reales de m filas y n columnas. Al conjunto de matrices de números reales de tamaño $m \times n$ lo notaremos $\mathbb{R}^{m\times n}$.

Ejemplo 2.13. Sea $A = \begin{pmatrix} 2 & -1 & 0 \\ 3 & 5 & 1 \end{pmatrix}$. A es una matriz de 2×3. Luego $A \in \mathbb{R}^{2\times 3}$.

Notación 2.14. Si $A \in \mathbb{R}^{m\times n}$, $i \in \{1, 2, \ldots, m\}$ y $j \in \{1, 2, \ldots, n\}$, notaremos A_{ij} al número real que está en la intersección de la fila i y la columna j de la matriz. Diremos también que A_{ij} está en el *lugar* (i, j) de la matriz.

Los números A_{ij} se llaman *entradas* o *coeficientes* de la matriz.

Por ejemplo, si A es la matriz de arriba entonces $A_{13} = 0$, $A_{21} = 3$ y $A_{23} = 1$.

Notemos que dos matrices $A, B \in \mathbb{R}^{m\times n}$ son iguales si y sólo si $A_{ij} = B_{ij}$ para todos $i \in \{1, 2, \ldots, m\}$ y $j \in \{1, 2, \ldots, n\}$.

Definición 2.15. Sea $\mathcal{S}$ el siguiente sistema de m ecuaciones lineales con n incógnitas:

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

La *matriz asociada al sistema* $\mathcal{S}$ es la matriz

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

La *matriz ampliada asociada al sistema* $\mathcal{S}$ es

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

Observación 2.16. Las transformaciones que se aplican a un sistema de ecuaciones lineales se traducen en las siguientes transformaciones que se aplican sobre la matriz ampliada asociada al sistema:

(1) **Intercambiar dos filas.** Si intercambiamos la fila j con la fila k notaremos $F_j \leftrightarrow F_k$.

(2) **Multiplicar una fila por un número real distinto de cero.** Si multiplicamos la fila j por α notaremos $F_j \leftarrow \alpha F_j$.

(3) **Sumar a una fila un múltiplo de otra fila.** Si sumamos a la fila j el producto de α por la fila k notaremos $F_j \leftarrow F_j + \alpha F_k$.

A estas transformaciones las llamaremos *transformaciones permitidas*.

Definición 2.17. Sean $m, n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{m\times n}$ una matriz. Decimos que A es *escalonada* si cumple las dos condiciones siguientes:

(1) Las filas que tengan todos sus elementos iguales a cero se encuentran en la parte inferior de la matriz.

(2) Para cada fila que tenga alguna de sus entradas no nulas se cumple que si A_{kl} es el primer elemento no nulo de dicha fila entonces para todo $i > k$ y para todo $j \leq l$ se tiene que $A_{ij} = 0$.

Diremos que una matriz es *cuadrada* si tiene la misma cantidad de filas que de columnas. Si $n \in \mathbb{N}$ y A es una matriz cuadrada que tiene n filas y n columnas llamaremos *diagonal principal* de A al conjunto formado por los lugares $(1,1), (2,2), \dots, (n,n)$ de la matriz A.

Definición 2.18. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada.

- Decimos que A es una matriz *diagonal* si para todos $i, j \in \{1, 2, \dots, n\}$ con $i \neq j$ vale que $A_{ij} = 0$. Es decir, la matriz A es diagonal si sus entradas fuera de la diagonal principal son cero.

- Decimos que A es una matriz *triangular superior* si para todos $i, j \in \{1, 2, \dots, n\}$ con $i > j$ vale que $A_{ij} = 0$. Es decir, la matriz A es triangular superior si debajo de la diagonal principal todas las entradas son cero.
- Decimos que A es una matriz *triangular inferior* si para todos $i, j \in \{1, 2, \dots, n\}$ con $i < j$ vale que $A_{ij} = 0$. Es decir, la matriz A es triangular inferior si arriba de la diagonal principal todas las entradas son cero.

Definición 2.19. Sea $n \in \mathbb{N}$. La *matriz identidad de tamaño* $n \times n$ es la matriz $I_n \in \mathbb{R}^{n\times n}$ definida por

$$(I_n)_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

En ocasiones notaremos a la matriz identidad por I.

Ejemplo 2.20. La matriz identidad de tamaño 3×3 es la matriz

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Definición 2.21. Sean $m, n \in \mathbb{N}$. La *matriz cero* o *matriz nula* de tamaño $m \times n$ es la matriz $0 \in \mathbb{R}^{m\times n}$ definida por $0_{ij} = 0$ para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$.

Definiremos ahora operaciones con matrices.

Definición 2.22. Sean $m, n \in \mathbb{N}$ y sean $A, B \in \mathbb{R}^{m\times n}$. Definimos la *suma* de A y B como la matriz $A + B \in \mathbb{R}^{m\times n}$ definida por $(A + B)_{ij} = A_{ij} + B_{ij}$ para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$.

Definición 2.23. Sean $m, n \in \mathbb{N}$, sea $A \in \mathbb{R}^{m\times n}$ y sea $\alpha \in \mathbb{R}$. Definimos $\alpha.A \in \mathbb{R}^{m\times n}$ como la matriz definida por $(\alpha.A)_{ij} = \alpha A_{ij}$ para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$.

Ejemplo 2.24. Sean $A = \begin{pmatrix} 1 & -1 & 0 \\ 3 & 2 & -3 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & 0 \end{pmatrix}$. Calcular $A + B$ y $2.A$.

Solución:
$A + B = \begin{pmatrix} 2 & -1 & 4 \\ 5 & 3 & -3 \end{pmatrix}$
$2.A = \begin{pmatrix} 2 & -2 & 0 \\ 6 & 4 & -6 \end{pmatrix}$

Proposición 2.25. *Sean $m, n \in \mathbb{N}$ y sean $A, B, C \in \mathbb{R}^{m\times n}$ y $\alpha, \beta \in \mathbb{R}$. Entonces*

(1) $(A + B) + C = A + (B + C)$

(2) $A + B = B + A$

(3) $A + 0 = A$ *y* $0 + A = A$

(4) $A + (-1).A = 0$ *y* $(-1).A + A = 0.$

(5) $\alpha.(A + B) = \alpha.A + \alpha.B$

(6) $(\alpha + \beta).A = \alpha.A + \beta.A$

(7) $\alpha.(\beta.A) = (\alpha\beta).A$

(8) $1.A = A$

Demostración. (1) Debemos demostrar que para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$ vale que $((A + B) + C)_{ij} = (A + (B + C))_{ij}$.

Sean $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$. Entonces

$$\begin{aligned} ((A+B)+C)_{ij} &= (A+B)_{ij} + C_{ij} = (A_{ij} + B_{ij}) + C_{ij} = A_{ij} + (B_{ij} + C_{ij}) = \\ &= A_{ij} + (B+C)_{ij} = (A + (B+C))_{ij}. \end{aligned}$$

Por lo tanto, $(A + B) + C = A + (B + C)$.

Las demostraciones de los ítems (2) a (8) quedan como ejercicio para el lector. □

Definición 2.26. Sean $m, n, l \in \mathbb{N}$ y sean $A \in \mathbb{R}^{m\times n}$ y $B \in \mathbb{R}^{n\times l}$. Definimos el *producto* $A.B$ como la matriz de $\mathbb{R}^{m\times l}$ definida por

$$(A.B)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$$

para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$.

Notemos que sólo se pueden multiplicar dos matrices (según la definición anterior) cuando la cantidad de columnas de la primera matriz es igual a la cantidad de filas de la segunda matriz. En este caso, el resultado del producto es una matriz que tiene tantas filas como la primera matriz y tantas columnas como la segunda matriz.

Ejemplo 2.27. Sean

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & -2 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & -2 & -1 & 0 \\ 4 & 0 & 2 & -3 \\ 2 & 3 & 1 & -1 \end{pmatrix}.$$

Calcular $(A.B)_{11}$, $(A.B)_{13}$, $(A.B)_{23}$ y luego calcular $A.B$.

Solución:

Notemos que $A \in \mathbb{R}^{2\times 3}$ y $B \in \mathbb{R}^{3\times 4}$. Se tiene que

$$(A.B)_{11} = \sum_{k=1}^{3} A_{1k}B_{k1} = A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} = 2.1 + (-1).4 + 3.2 = 4.$$

$$(A.B)_{13} = \sum_{k=1}^{3} A_{1k}B_{k3} = A_{11}B_{13} + A_{12}B_{23} + A_{13}B_{33} = 2.(-1) + (-1).2 + 3.1 = -1.$$

$$(A.B)_{23} = \sum_{k=1}^{3} A_{2k}B_{k3} = A_{21}B_{13} + A_{22}B_{23} + A_{23}B_{33} = 1.(-1) + 0.2 + (-2).1 = -3.$$

Para calcular $A.B$ hacemos lo siguiente

$$\begin{array}{c|rrr|rrrr|c} & & & & 1 & -2 & -1 & 0 & \\ & & & & 4 & 0 & 2 & -3 & B \\ & & & & 2 & 3 & 1 & -1 & \\ \cline{2-8} A & 2 & -1 & 3 & 4 & 5 & -1 & 0 & \\ & 1 & 0 & -2 & -3 & -8 & -3 & 2 & \end{array}$$

Luego

$$A.B = \begin{pmatrix} 4 & 5 & -1 & 0 \\ -3 & -8 & -3 & 2 \end{pmatrix}.$$

Observación importante 2.28. El producto de matrices no es conmutativo. Puede ocurrir que $A.B$ pueda calcularse pero $B.A$ no, como ocurre en el ejemplo anterior ya que $B \in \mathbb{R}^{3\times 4}$ y $A \in \mathbb{R}^{2\times 3}$.

También puede ocurrir que los productos $A.B$ y $B.A$ puedan ambos calcularse, pero que tengan distinto tamaño. Por ejemplo, si A es una matriz de 2×3 y B es una matriz de 3×2 entonces $A.B$ es una matriz de 2×2 y $B.A$ es una matriz de 3×3.

Si A y B son matrices cuadradas del mismo tamaño, digamos $n\times n$, entonces tanto $A.B$ como $B.A$ se pueden calcular y tienen tamaño $n \times n$, pero puede ocurrir que $A.B \neq B.A$. Por ejemplo, si

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

entonces

$$A.B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad B.A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Luego $A.B \neq B.A$.

Notemos, además, que en este ejemplo $A.B = 0$ pero ni A ni B son la matriz 0.

Proposición 2.29. *Sean $m, n, r, l \in \mathbb{N}$.*

(1) Sea $A \in \mathbb{R}^{m\times n}$. Entonces $A.\,I_n = A$ y $I_m.\,A = A$ (recordemos que I_n es la matriz identidad de tamaño $n \times n$ y que I_m la matriz identidad de tamaño $m \times m$).

(2) Sea $A \in \mathbb{R}^{m\times n}$ y sean $B, C \in \mathbb{R}^{n\times l}$. Entonces $A.\,(B + C) = A.\,B + A.\,C$.

(3) Sean $B, C \in \mathbb{R}^{m\times n}$ y sea $A \in \mathbb{R}^{n\times l}$. Entonces $(B + C).\,A = B.\,A + C.\,A$.

(4) Sean $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times l}$ y $C \in \mathbb{R}^{l\times r}$. Entonces $(AB)C = A(BC)$.

(5) Sean $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times l}$ y $\alpha \in \mathbb{R}$. Entonces $(\alpha A)B = \alpha(AB) = A(\alpha B)$.

Demostración.

(1) Sean $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$. Entonces

$$(A.I_n)_{ij} = \sum_{k=1}^{n} A_{ik}(I_n)_{kj} = A_{ij}.1 = A_{ij}$$

ya que

$$(I_n)_{kj} = \begin{cases} 0 & \text{si } k \neq j \\ 1 & \text{si } k = j \end{cases}$$

Por lo tanto, $A.I_n = A$.

La demostración de que $I_m.A = A$ es similar y queda como ejercicio para el lector.

(4) Sean $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, r\}$. Entonces

$$((A.B).C)_{ij} = \sum_{k=1}^{l}(A.B)_{ik}C_{kj} = \sum_{k=1}^{l}\left(\sum_{h=1}^{n} A_{ih}.B_{hk}\right).C_{kj} = \sum_{k=1}^{l}\sum_{h=1}^{n} A_{ih}B_{hk}C_{kj}\ .$$

Por otra parte,

$$(A.(B.C))_{ij} = \sum_{k'=1}^{n} A_{ik'}(B.C)_{k'j} = \sum_{k'=1}^{n} A_{ik'}\left(\sum_{h'=1}^{l} B_{k'h'}.C_{h'j}\right) = \sum_{k'=1}^{n}\sum_{h'=1}^{l} A_{ik'}B_{k'h'}C_{h'j}\ .$$

Notemos que las expresiones de ambos renglones coinciden pues h' juega el rol de k y k' juega el rol de h.

Por lo tanto, $(A.B).C = A.(B.C)$.

Las demostraciones de los items (2), (3) y (5) quedan como ejercicio para el lector. $\square$

Definición 2.30. Sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Decimos que A es *inversible* si existe una matriz $B \in \mathbb{R}^{n\times n}$ tal que $A.B = I$ y $B.A = I$. En este caso, la matriz B se llama *matriz inversa de* A y se nota A^{-1}.

Observación 2.31. Sea $A \in \mathbb{R}^{n\times n}$. Si A es inversible entonces la inversa es única. En efecto, si las matrices B y C son inversas de A entonces se tiene que

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

Es decir, $B = C$.

Ejemplo 2.32. Hallar, si existe, la inversa de la matriz $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$.

Solución:

Buscamos una matriz $B \in \mathbb{R}^{2\times 2}$ tal que $AB = I$ y $BA = I$. Planteamos $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Entonces $A.B = \begin{pmatrix} 2a+5c & 2b+5d \\ a+3c & b+3d \end{pmatrix}$.

Queremos que $A.B = I$, entonces necesitamos que

$$\begin{cases} 2a+5c &= 1 \\ 2b+5d &= 0 \\ a+3c &= 0 \\ b+3d &= 1 \end{cases}$$

Tenemos entonces dos sistemas de dos ecuaciones con dos incógnitas cada uno:

$$\begin{cases} 2a + 5c = 1 \\ a + 3c = 0 \end{cases} \quad \text{y} \quad \begin{cases} 2b + 5d = 0 \\ b + 3d = 1 \end{cases}$$

Como ambos tienen la misma matriz asociada podemos resolverlos juntos:

$$\left(\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right) \underset{F_1 \leftrightarrow F_2}{\rightsquigarrow} \left(\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 2 & 5 & 1 & 0 \end{array}\right) \underset{F_2 \leftarrow F_2 - 2F_1}{\rightsquigarrow}$$

$$\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & -1 & 1 & -2 \end{array}\right) \underset{F_2 \leftarrow (-1)F_2}{\rightsquigarrow}$$

$$\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & 1 & -1 & 2 \end{array}\right) \underset{F_1 \leftarrow F_1 - 3F_2}{\rightsquigarrow} \left(\begin{array}{cc|cc} 1 & 0 & 3 & -5 \\ 0 & 1 & -1 & 2 \end{array}\right)$$

Luego, los sistemas originales son equivalentes a los sistemas

$$\begin{cases} a \quad = 3 \\ \quad c = -1 \end{cases} \quad \text{y} \quad \begin{cases} b \quad = -5 \\ \quad d = 2 \end{cases} \qquad \text{respectivamente.}$$

Así obtenemos $B = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$.

Dejamos como ejercicio para el lector verificar que $A.B = I$ y que $B.A = I$.

Luego, A es inversible y $A^{-1} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$.

Definición 2.33. Sean $m, n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{m\times n}$. Definimos la *matriz transpuesta* de A como la matriz $A^t \in \mathbb{R}^{n\times m}$ definida por $(A^t)_{ij} = A_{ji}$ para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$. Es decir, la matriz transpuesta de A se obtiene poniendo las filas de A como columnas (y por ende, las columnas quedan como filas).

Ejemplo 2.34. Sea $A = \begin{pmatrix} 1 & -2 & 5 \\ 0 & 3 & 4 \end{pmatrix}$. Entonces $A^t = \begin{pmatrix} 1 & 0 \\ -2 & 3 \\ 5 & 4 \end{pmatrix}$.

Definición 2.35. Sean $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Decimos que A es una matriz *simétrica* si $A^t = A$, es decir, si $A_{ji} = A_{ij}$ para todos $i, j \in \{1, 2, \ldots, n\}$.

Ejemplo 2.36. Sea
$$A = \begin{pmatrix} 1 & 0 & 4 \\ 0 & -3 & 5 \\ 4 & 5 & -2 \end{pmatrix} .$$
La matriz A es una matriz simétrica.

Proposición 2.37. *Sean $m, n, l \in \mathbb{N}$. Entonces:*

(1) $\forall A \in \mathbb{R}^{m\times n}$, $(A^t)^t = A$.

(2) $\forall A, B \in \mathbb{R}^{m\times n}$, $(A + B)^t = A^t + B^t$.

(3) $\forall A \in \mathbb{R}^{m\times n}$, $\forall \alpha \in \mathbb{R}$, $(\alpha.A)^t = \alpha.A^t$.

(4) $\forall A \in \mathbb{R}^{m\times n}$, $\forall B \in \mathbb{R}^{n\times l}$, $(A.B)^t = B^t.A^t$.

(5) Si $A \in \mathbb{R}^{n\times n}$ es inversible entonces A^t es inversible y $(A^t)^{-1} = (A^{-1})^t$.

Demostración.
(4) Notemos que $A.B \in \mathbb{R}^{m\times l}$. Luego $(A.B)^t \in \mathbb{R}^{l\times m}$. Por otra parte, $B^t \in \mathbb{R}^{l\times n}$ y $A^t \in \mathbb{R}^{n\times m}$. Luego, es posible hacer el producto $B^t.A^t$ y se obtiene que $B^t.A^t \in \mathbb{R}^{l\times m}$. Así, $(A.B)^t$ y $B^t.A^t$ tienen el mismo tamaño.

Debemos probar ahora que para todos $i \in \{1, 2, \ldots, m\}$ y $j \in \{1, 2, \ldots, n\}$ vale que $((A.B)^t)_{ij} = (B^t.A^t)_{ij}$. Sean $i \in \{1, 2, \ldots, m\}$ y $j \in \{1, 2, \ldots, n\}$. Se tiene que
$$((A.B)^t)_{ij} = (A.B)_{ji} = \sum_{k=1}^{n} A_{jk}B_{ki} = \sum_{k=1}^{n} (A^t)_{kj}(B^t)_{ik} = \sum_{k=1}^{n} (B^t)_{ik}(A^t)_{kj} = (B^t.A^t)_{ij}$$
Por lo tanto, $(A.B)^t = B^t.A^t$. □

(5) Aplicando la propiedad (4) tenemos que
$$A^t.(A^{-1})^t = (A^{-1}.A)^t = I^t = I.$$
Similarmente
$$(A^{-1})^t.A^t = (A.A^{-1})^t = I^t = I.$$
Por lo tanto A^t es inversible y su inversa es $(A^{-1})^t$.

Las demostraciones de los ítems (1), (2) y (3) quedan como ejercicio para el lector.

Observación importante 2.38. Consideremos un sistema de m ecuaciones lineales con n incógnitas
$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{cases}$$

Sea A su matriz asociada y sea $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$.

Sea $x = (x_1, x_2, \ldots, x_n)$. Podemos pensar a x como un vector de $\mathbb{R}^n$ o también como una matriz $\begin{pmatrix} x_1 & x_2 & \ldots & x_m \end{pmatrix} \in \mathbb{R}^{1\times n}$.

Notemos entonces que el sistema de ecuaciones lineales anterior puede escribirse en la forma $A.x^t = b$.

A partir de la observacion anterior prodremos deducir fácilmente algunas propiedades de las soluciones de un sistema de ecuaciones lineales homogéneo así como también las relaciones entre las soluciones de un sistema no homogéneo y las de sus sistema homogéneo asociado. Estas propiedades se detallan en la siguiente proposición.

Proposición 2.39. *Sean $m, n \in \mathbb{N}$ y sean $A \in \mathbb{R}^{m\times n}$ y $b \in \mathbb{R}^{m\times 1}$. Sea S_b el conjunto de soluciones del sistema $A.x^t = b$, es decir, $S_b = \{x \in \mathbb{R}^n \,/\, Ax^t = b\}$. Sea S_0 el conjunto de soluciones del sistema homogéneo asociado al anterior, es decir, $S_0 = \{x \in \mathbb{R}^n \,/\, Ax^t = 0\}$. Entonces*

(1)
- $0 \in S_0$
- *Si $x, y \in S_0$ entonces $x + y \in S_0$.*
- *Si $x \in S_0$ y $\alpha \in \mathbb{R}$ entonces $\alpha x \in S_0$.*

(2) Si $x, y \in S_b$ entonces $x - y \in S_0$.

(3) Si $s \in S_b$ entonces $S_b = \{y \in \mathbb{R}^n \,/\, y = x + s$ para algún $x \in S_0\}$.

Demostración.

(1) Como $A.0^t = 0$, se tiene que $0 \in S_0$.

Si $x, y \in S_0$ entonces $A.x^t = 0$ y $A.y^t = 0$. Luego

$$A.(x + y)^t = A.(x^t + y^t) = A.x^t + A.y^t = 0 + 0 = 0.$$

Por lo tanto, $x + y \in S_0$.

Si $x \in S_0$ y $\alpha \in \mathbb{R}$ entonces

$$A.(\alpha x)^t = A.(\alpha x^t) = \alpha(A.x^t) = \alpha.0 = 0.$$

Por lo tanto, $\alpha x \in S_0$.

(2) Si $x, y \in S_b$ entonces $A.x^t = b$ y $A.y^t = b$. Luego

$$A.(x - y)^t = A.(x^t - y^t) = A.x^t - A.y^t = b - b = 0.$$

Por lo tanto, $x - y \in S_0$.

(3) Sea $s \in S_b$. Queremos probar que $S_b = \{y \in \mathbb{R}^n \,/\, y = x + s$ para algún $x \in S_0\}$. Veamos la doble inclusión.

($\supseteq$) Sea $y \in \mathbb{R}^n$ tal que $y = x + s$ para cierto $x \in S_0$ entonces

$$A.y^t = A.(x + s)^t = A.(x^t + s^t) = A.x^t + A.s^t = 0 + b = b.$$

Por lo tanto, $y \in S_b$.

($\subseteq$) Sea $y \in S_b$. Entonces, por el item (2), $y - s \in S_0$. Llamemos $x = y - s$. Entonces $y = x + s$ con $x \in S_0$. $\square$

Observación 2.40. El tercer ítem de la proposición anterior dice que si tenemos una solución s del sistema de ecuaciones $A.x^t = b$ entonces todas las soluciones de dicho sistema se obtienen sumando s a las soluciones del sistema homogéneo asociado.

Proposición 2.41. *Sea $n \in \mathbb{N}$ y sean $A, B \in \mathbb{R}^{n\times n}$. Si A y B son inversibles entonces $A.B$ es inversible y $(A.B)^{-1} = B^{-1}.A^{-1}$.*

Demostración. Se tiene que

$$(A.B).(B^{-1}.A^{-1}) = (A.(B.B^{-1})).A^{-1} = (A.I).A^{-1} = A.A^{-1} = I$$

y

$$(B^{-1}.A^{-1}).(A.B) = (B^{-1}.(A^{-1}.A)).B = (B^{-1}.I).B = B^{-1}.B = I$$

Por lo tanto, $A.B$ es inversible y $(A.B)^{-1} = B^{-1}.A^{-1}$. □

Definición 2.42. Sea $n \in \mathbb{N}$. Las *matrices elementales* de tamaño $n \times n$ son las matrices de $\mathbb{R}^{n\times n}$ que se obtienen de aplicar una de las transformaciones permitidas a la matriz identidad.

Tenemos entonces matrices elementales de tres tipos:

- *Matrices elementales de transposición.* Son las obtenidas intercambiando dos filas de la matriz identidad. Si $j, k \in \{1, 2, \dots, n\}$ son tales que $j \neq k$, notaremos $E(T, j, k)$ a la matriz que se obtiene de intercambiar las filas j y k de la matriz identidad.
- *Matrices elementales de multiplicación.* Son las obtenidas multiplicando una fila de la matriz identidad por un número real no nulo. Si $j \in \{1, 2, \dots, n\}$ y $\alpha \in \mathbb{R} - \{0\}$, notaremos $E(M, j, \alpha)$ a la matriz que se obtiene de multiplicar las fila j de la matriz identidad por α.
- *Matrices elementales de combinación.* Son las obtenidas sumando a una fila de la matriz identidad un múltiplo de otra de sus filas. Si $j, k \in \{1, 2, \dots, n\}$ son tales que $j \neq k$ y $\alpha \in \mathbb{R}$, notaremos $E(C, j, k, \alpha)$ a la matriz que se obtiene de aplicar la transformación $F_j \leftarrow F_j + \alpha F_k$ a la matriz identidad.

Ejemplo 2.43. Las siguientes son matrices elementales de $\mathbb{R}^{5\times 5}$:

$$E(T, 2, 5) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$E(M, 4, 6) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$E(C, 3, 2, 4) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Proposición 2.44. *Sean $m, n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{m\times n}$. Sea A' la matriz que se obtiene aplicando a A una de las transformaciones permitidas y sea E la matriz elemental que se obtiene aplicando la misma transformación a la matriz I_m. Entonces $E.A = A'$.*

Demostración. Analizaremos cada uno de los tres casos correspondientes a cada uno de los tipos de transformaciones permitidas.

Supongamos primero que la transformación aplicada a la matriz A es $F_j \leftrightarrow F_k$ para ciertos $j, k \in \{1, 2, \dots, m\}$ con $j < k$.

Tenemos que

$$
\begin{array}{rcccccccc|cccc}
 & & & & & & & & & a_{11} & a_{12} & \dots & a_{1n} \\
 & & & & & & & & & a_{21} & a_{22} & \dots & a_{2n} \\
 & & & & & & & & & \vdots & \vdots & \vdots & \vdots \\
 & & & & & & & & & a_{m1} & a_{m2} & \dots & a_{mn} \\
\cline{2-13}
 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1n} \\
 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 & a_{21} & a_{22} & \dots & a_{2n} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Fila } j \ \rightarrow & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 & a_{k1} & a_{k2} & \dots & a_{kn} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Fila } k \ \rightarrow & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 & a_{j1} & a_{j2} & \dots & a_{jn} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 & a_{m1} & a_{m2} & \dots & a_{mn} \\
\multicolumn{1}{r}{} & & & & \uparrow & & \uparrow & & \multicolumn{1}{c}{} & & & & \\
\multicolumn{1}{r}{} & & & & \text{Col. } j & & \text{Col. } k & & \multicolumn{1}{c}{} & & & &
\end{array}
$$

Por lo tanto, en este caso $E.A = A'$

Supongamos ahora que la transformación aplicada a la matriz A es $F_j \leftrightarrow \alpha F_j$ para ciertos $j \in \{1, 2, \dots, m\}$ y $\alpha \in \mathbb{R} - \{0\}$.

Tenemos que

$$
\begin{array}{rcccccc|cccc}
 & & & & & & & a_{11} & a_{12} & \dots & a_{1n} \\
 & & & & & & & a_{21} & a_{22} & \dots & a_{2n} \\
 & & & & & & & \vdots & \vdots & \vdots & \vdots \\
 & & & & & & & a_{m1} & a_{m2} & \dots & a_{mn} \\
\cline{2-11}
 & 1 & 0 & \dots & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1n} \\
 & 0 & 1 & \dots & 0 & \dots & 0 & a_{21} & a_{22} & \dots & a_{2n} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Fila } j \ \rightarrow & 0 & 0 & \dots & \alpha & \dots & 0 & \alpha a_{j1} & \alpha a_{j2} & \dots & \alpha a_{jn} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & 0 & 0 & \dots & 0 & \dots & 1 & a_{m1} & a_{m2} & \dots & a_{mn} \\
\multicolumn{1}{r}{} & & & & \uparrow & & \multicolumn{1}{c}{} & & & & \\
\multicolumn{1}{r}{} & & & & \text{Col. } j & & \multicolumn{1}{c}{} & & & &
\end{array}
$$

Por lo tanto, en este caso también vale que $E.A = A'$

Por último, supongamos que la transformación aplicada a la matriz A es $F_j \leftrightarrow F_j + \alpha F_k$ para ciertos $\alpha \in \mathbb{R}$ y $j, k \in \{1, 2, \dots, m\}$ con $j \neq k$.

Tenemos que

$$
\begin{array}{r cccccccc|cccc}
 & & & & & & & & & a_{11} & a_{12} & \dots & a_{1n} \\
 & & & & & & & & & a_{21} & a_{22} & \dots & a_{2n} \\
 & & & & & & & & & \vdots & \vdots & \vdots & \vdots \\
 & & & & & & & & & a_{m1} & a_{m2} & \dots & a_{mn} \\
\hline
 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1n} \\
 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 & a_{21} & a_{22} & \dots & a_{2n} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Fila } j \rightarrow & 0 & 0 & \dots & 1 & \dots & \alpha & \dots & 0 & a_{j1}+\alpha a_{k1} & a_{j2}+\alpha a_{k2} & \dots & a_{jn}+\alpha a_{kn} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Fila } k \rightarrow & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 & a_{k1} & a_{k2} & \dots & a_{kn} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 & a_{m1} & a_{m2} & \dots & a_{mn} \\
\multicolumn{1}{r}{} & & & & \uparrow & & \uparrow & & \multicolumn{1}{c}{} & & & & \\
\multicolumn{1}{r}{} & & & & \text{Col. } j & & \text{Col. } k & & \multicolumn{1}{c}{} & & & &
\end{array}
$$

Por lo tanto, en este caso también vale que $E.A = A'$. □

Proposición 2.45. *Las matrices elementales son inversibles y sus inversas son también matrices elementales.*

Demostración. Aplicando la proposición anterior para $A = E(T, j, k)$ tenemos que $E(T, j, k) \,.\, E(T, j, k)$ es la matriz que se obtiene de intercambiar las filas j y k de la matriz $E(T, j, k)$. Por lo tanto, $E(T, j, k).E(T, j, k) = I$. Luego, $E(T, j, k)$ es inversible y su inversa es ella misma.

Por otra parte, y también por la proposición anterior, si $\alpha \neq 0$ entonces la matriz $E(M, j, \alpha).E(M, j, \frac{1}{\alpha})$ es la que se obtiene de multiplicar la fila j de la matriz $E(M, j, \frac{1}{\alpha})$ por α. Por lo tanto, $E(M, j, \alpha).E(M, j, \frac{1}{\alpha}) = I$.

Similarmente, la matriz $E(M, j, \frac{1}{\alpha}).E(M, j, \alpha)$ es la que se obtiene de multiplicar la fila j de la matriz $E(M, j, \alpha)$ por $\frac{1}{\alpha}$. Por lo tanto, $E(M, j, \frac{1}{\alpha}).E(M, j, \alpha) = I$.

Luego, $E(M, j, \alpha)$ es inversible y su inversa es $E(M, j, \frac{1}{\alpha})$.

Por otra parte, y aplicando nuevamente la proposición anterior, tenemos que la matriz $E(M, j, k, \alpha).E(M, j, k, -\alpha)$ es la matriz que se obtiene de aplicar la transformación $F_j \leftarrow F_j + \alpha F_k$ a la matriz $E(M, j, k, -\alpha)$. Y como la matriz $E(M, j, k, -\alpha)$ es la que se obtiene de aplicar la transformación $F_j \leftarrow F_j - \alpha F_k$ a la matriz identidad obtenemos que la matriz $E(M, j, k, \alpha).E(M, j, k, -\alpha)$ se obtiene de aplicar a la matriz identidad primero la transformación $F_j \leftarrow F_j - \alpha F_k$ y luego la transformación $F_j \leftarrow F_j + \alpha F_k$. Luego $E(M, j, k, \alpha).E(M, j, k, -\alpha) = I$.

De manera similar se prueba que $E(M, j, k, -\alpha) \,.\, E(M, j, k, \alpha) = I$. Por lo tanto, $E(M, j, k, \alpha)$ es inversible y su inversa es $E(M, j, k, -\alpha)$. □

Teorema 2.46. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n \times n}$. Son equivalentes:*

(1) A es inversible.

(2) Para todo $b \in \mathbb{R}^{n \times 1}$ el sistema $Ax^t = b$ tiene solución única.

(3) El sistema $Ax^t = 0$ tiene solución única $x = 0$.

(4) A es producto de matrices elementales.

Demostración.

(1)⇒(2) Sea $b \in \mathbb{R}^{n\times 1}$. Veamos primero que el sistema $Ax^t = b$ tiene solución.

Sea $z = (A^{-1}.b)^t$. Notemos que $A^{-1}.b \in \mathbb{R}^{n\times 1}$ y luego $(A^{-1}.b)^t \in \mathbb{R}^{1\times n}$. Entonces podemos ver a $z = (A^{-1}.b)^t$ como un vector de $\mathbb{R}^n$.

Tenemos entonces que

$$A.z^t = A.((A^{-1}.b)^t)^t = A.(A^{-1}.b) = (A.A^{-1}).b) = I.b = b\ .$$

Por lo tanto, z es solución del sistema.

Veamos ahora que z es la única solución del sistema. Sea y una solución del sistema $Ax^t = b$. Entonces

$$Ay^t = b.$$

Luego

$$A^{-1}(Ay^t) = A^{-1}b$$

de donde se obtiene que

$$(A^{-1}A)y^t = A^{-1}b\ ,$$

es decir,

$$(I.y^t) = A^{-1}b.$$

Luego

$$y^t = A^{-1}b$$

de donde obtenemos que

$$y = (y^t)^t = (A^{-1}b)^t = z.$$

Así, $y = z$ y por lo tanto z es la única solución del sistema.

(2)⇒(3) Trivial.

(3)⇒(4) Como el sistema $Ax^t = 0$ tiene solución única y $A \in \mathbb{R}^{n\times n}$ entonces al escalonar la matriz A quedan n ecuaciones no triviales (tantas como incógnitas). Por lo tanto, por lo hecho en la demostración de la proposición 2.9 sabemos que el primer coeficiente no nulo de la j-ésima ecuación en el sistema escalonado obtenido debe ser el de la variable x_j. Dicho de otra forma, al escalonar la matriz queda una matriz triangular superior con entradas no nulas en la diagonal principal.

Aplicando las transformaciones permitidas a esta matriz escalonada podemos llevar esta matriz a otra que tenga unos en la diagonal. Luego, podemos seguir aplicando transformaciones permitidas para llevar esta nueva matriz a otra cuyas entradas de arriba de la diagonal principal sean todas cero. De este modo, aplicando transformaciones permitidas hemos llevado la matriz A a la matriz identidad.

Por la proposición 2.44 sabemos que cada una de estas transformaciones permitidas se puede obtener multiplicando a la matriz a izquierda por una matriz elemental. Luego, existen matrices elementales $E_1, E_2, \ldots, E_r$ tales que

$$E_r. \ldots .E_2.E_1.A = I\ .$$

Y como, por la proposición anterior, las matrices elementales son inversibles, de la igualdad anterior obtenemos que

$$\begin{aligned}
(E_r)^{-1}.(E_r.E_{r-1}.\dots.E_2.E_1.A) &= (E_r)^{-1}.I \\
((E_r)^{-1}.E_r).(E_{r-1}.\dots.E_2.E_1.A) &= (E_r)^{-1} \\
I.E_{r-1}.\dots.E_2.E_1.A &= (E_r)^{-1} \\
E_{r-1}.\dots.E_2.E_1.A &= (E_r)^{-1} \\
(E_{r-1})^{-1}.E_{r-1}.\dots.E_2.E_1.A &= (E_{r-1})^{-1}.(E_r)^{-1} \\
E_{r-2}.\dots.E_2.E_1.A &= (E_{r-1})^{-1}.(E_r)^{-1} \\
\vdots \quad & \vdots \quad \quad \vdots \\
A &= (E_1)^{-1}.(E_2)^{-1}.\dots.(E_{r-1})^{-1}.(E_r)^{-1}
\end{aligned}$$

Y como, por la proposición anterior, las inversas de matrices elementales son también matrices elementales, obtenemos que A es producto de matrices elementales.

(4)$\Rightarrow$(1) Por la proposición 2.41, el producto de dos matrices inversibles es inversible. Por lo tanto, el producto de una cantidad finita de matrices inversibles es también inversible.

Ahora bien, como por hipótesis A es un producto de matrices elementales y las matrices elementales son inversibles por la proposición anterior, entonces A es un producto de matrices inversibles y luego A es inversible. □

Determinantes

Definición 2.47. Sea $n \in \mathbb{N}$. Una *permutación* del conjunto $\{1, 2, \dots, n\}$ es una función biyectiva $f : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$.

Notación 2.48. Al conjunto de permutaciones del conjunto $\{1, 2, \dots, n\}$ lo notaremos S_n. Y si $f \in S_n$ notaremos a f por $(f(1)\ f(2)\ f(3)\ \dots\ f(n))$

Ejemplo 2.49. Sea $f : \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}$ definida por $f(1) = 3$, $f(2) = 1$, $f(3) = 2$ y $f(4) = 4$. Entonces $f \in S_4$ y podemos notar a f por (3 1 2 4).

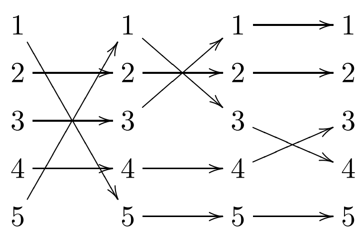
Ejemplo 2.50.
Las permutaciones del conjunto $\{1, 2\}$ son (1 2) y (2 1).
Las permutaciones del conjunto $\{1, 2, 3\}$ son (1 2 3), (1 3 2), (2 1 3), (2 3 1), (3 1 2) y (3 2 1).

Definición 2.51. Sea $n \in \mathbb{N}$. Una *transposición* de S_n es una permutación f de S_n tal que $f(j) \neq j$ para exactamente dos valores de $j \in \{1, 2, \dots, n\}$.

Ejemplo 2.52. (1 5 3 4 2) es una transposición de S_5.

Proposición 2.53. *Sea $n \in \mathbb{N}$. Toda permutación de S_n puede escribirse como composición de transposiciones.*

Ejemplo 2.54. Escribamos a la permutación (5 2 1 3 4) como composición de transposiciones:



Proposición 2.55. *Sea $n \in \mathbb{N}$ y sea $f \in S_n$. Supongamos que $f = \sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_r$ y que $f = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_s$ donde $\sigma_1, \sigma_2, \ldots \sigma_r, \tau_1, \tau_2, \ldots, \tau_s$ son transposiciones. Entonces r y s tienen la misma paridad.*

Definición 2.56. Sea $n \in \mathbb{N}$ y sea $f \in S_n$. Supongamos que $f = \sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_r$ donde $\sigma_1, \sigma_2, \ldots \sigma_r$ son transposiciones. Definimos el *signo* de f por $\operatorname{sgn}(f) = (-1)^r$.

Es decir, el signo de f es 1 si f se escribe como la composición de una cantidad par de transposiciones y -1 si f se escribe como la composición de una cantidad impar de transposiciones. La proposición anterior nos dice que el signo de una permutación está bien definido pues no dependerá de la descomposición de f elegida.

Definición 2.57. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Definimos el *determinante de A* como el número real

$$\det(A) = \sum_{f \in S_n} \operatorname{sgn}(f) A_{1f(1)} A_{2f(2)} \ldots A_{nf(n)}$$

Ejemplo 2.58.
(1) Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Ya hemos visto que en S_2 hay dos permutaciones: (1 2) y (2 1). Claramente la permutación (1 2) es la permutación identidad que tiene signo 1 y la permutación (2 1) es una transposición y por lo tanto tiene signo -1. Entonces

$$\det(A) = 1.A_{11}.A_{22} + (-1).A_{12}.A_{21} = ad - bc.$$

(2) Sea $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$.

Ya hemos visto que en S_3 hay seis permutaciones: (1 2 3), (1 3 2), (2 1 3), (2 3 1), (3 1 2) y (3 2 1). Analicemos el signo de cada una de ellas.

La permutación (1 2 3) es la identidad y por lo tanto tiene signo 1.

Las permutaciones (1 3 2), (2 1 3) y (3 2 1) son transposiciones y por lo tanto tienen signo -1.

La permutación (2 3 1) se puede escribir como composición de las transposiciones (2 1 3) y (3 2 1) y por lo tanto tiene signo 1.

La permutación (3 1 2) se puede escribir como composición de las transposiciones (3 2 1) y (2 1 3) y por lo tanto tiene signo 1.

Luego

$$\begin{aligned} \det(A) &= 1.A_{11}.A_{22}.A_{33} + (-1).A_{11}.A_{23}.A_{32} + (-1).A_{12}.A_{21}.A_{33} + \\ &+ 1.A_{12}.A_{23}.A_{31} + 1.A_{13}.A_{21}.A_{32} + (-1).A_{13}.A_{22}.A_{31} = \\ &= aei - afh - bdi + bfg + cdh - ceg\ . \end{aligned}$$

Veremos ahora una regla para calcular determinantes sin utilizar permutaciones.

Notación 2.59. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Sean $i, j \in \{1, 2, \ldots, n\}$. Notaremos $A_{(i|j)}$ a la matriz que se obtiene a partir de la matriz A eliminando su fila i y su columna j.

Ejemplo 2.60. Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Entonces $A_{(2|1)} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}$.

Proposición 2.61. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Sea $k \in \{1, 2, \dots, n\}$. Entonces*

$$\det(A) = \sum_{j=1}^{n} (-1)^{k+j} A_{kj} \det(A_{(k|j)})$$

y

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} A_{ik} \det(A_{(i|k)})$$

La primera fórmula se conoce con el nombre de *desarrollo del determinante por la fila* k y la segunda se conoce con el nombre de *desarrollo del determinante por la columna* k.

Ejemplo 2.62. Sea $A = \begin{pmatrix} 1 & -2 & 2 \\ 3 & 0 & 5 \\ 1 & 3 & 4 \end{pmatrix}$. Calcular $\det(A)$.

Solución:

Haciendo el desarrollo del determinante por la primera fila queda

$$\begin{aligned} \det(A) &= (-1)^{1+1}.1.\det\begin{pmatrix} 0 & 5 \\ 3 & 4 \end{pmatrix} + (-1)^{1+2}.(-2).\det\begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix} + \\ &+ (-1)^{1+3}.2.\det\begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix} = \\ &= -15 + 2.7 + 2.9 = -15 + 14 + 18 = 17. \end{aligned}$$

Por lo tanto, $\det(A) = 17$.

Si desarrollamos el determinante por la segunda columna queda

$$\begin{aligned} \det(A) &= (-1)^{1+2}.(-2).\det\begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix} + (-1)^{2+2}.0.\det\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} + \\ &+ (-1)^{3+2}.3.\det\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} = \\ &= 2.7 + 0 + (-3).(-1) = 14 + 3 = 17. \end{aligned}$$

Y obtenemos nuevamente que $\det(A) = 17$.

Ejemplo 2.63. Sea $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 10 \end{pmatrix}$. Calcular $\det(A)$.

Solución:

Desarrollando el determinante por la primera fila obtenemos que

$$\det(A) = (-1)^{1+1}.1.\det\begin{pmatrix} 3 & 0 & 0 \\ 5 & 6 & 0 \\ 8 & 9 & 10 \end{pmatrix} + 0 + 0 + 0 = 1.\det\begin{pmatrix} 3 & 0 & 0 \\ 5 & 6 & 0 \\ 8 & 9 & 10 \end{pmatrix}$$

Y desarrollando este último determinante por la primera fila se obtiene que

$$\begin{aligned} \det(A) &= 1.\det\begin{pmatrix} 3 & 0 & 0 \\ 5 & 6 & 0 \\ 8 & 9 & 10 \end{pmatrix} = 1.\left((-1)^{1+1}.3.\det\begin{pmatrix} 6 & 0 \\ 9 & 10 \end{pmatrix} + 0 + 0\right) = \\ &= 1.3.\det\begin{pmatrix} 6 & 0 \\ 9 & 10 \end{pmatrix} = 1.3.(6.10 - 9.0) = 1.3.6.10 = 180\ . \end{aligned}$$

Aplicando el desarrollo del determinante por filas o columnas no es difícil demostrar la siguiente proposición.

Proposición 2.64. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$.*

(1) Si A tiene una fila o una columna de ceros entonces $\det(A) = 0$.

(2) Si A es una matriz triangular superior o inferior entonces $\det(A)$ *es el producto de los elementos de la diagonal.*

(3) $\det(A^t) = \det(A)$.

Proposición 2.65. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$.*

(1) Si A' es la matriz que se obtiene aplicando a A la transformación $F_j \leftrightarrow F_k$ entonces $\det(A') = -\det(A)$.

(2) Si A' es la matriz que se obtiene aplicando a A la transformación $F_j \to \alpha F_j$ entonces $\det(A') = \alpha \det(A)$.

(3) Si A' es la matriz que se obtiene aplicando a A la transformación $F_j \to F_j + \alpha F_k$ entonces $\det(A') = \det(A)$.

Ejemplo 2.66. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Supongamos que $\det(A) = 3$.

Entonces

$\det \begin{pmatrix} c & d \\ a & b \end{pmatrix} = -3.$

$\det \begin{pmatrix} a & b \\ 5c & 5d \end{pmatrix} = 5.3 = 15.$

$\det \begin{pmatrix} a & b \\ c+3a & d+3b \end{pmatrix} = 3.$

Corolario 2.67. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Si la matriz A tiene dos filas iguales entonces* $\det(A) = 0$.

Proposición 2.68. *Sea $n \in \mathbb{N}$, sea $A \in \mathbb{R}^{n\times n}$ y sea $\alpha \in \mathbb{R}$. Entonces* $\det(\alpha A) = \alpha^n.\det(A)$.

Proposición 2.69. *Sea $n \in \mathbb{N}$ y sean $A, B \in \mathbb{R}^{n\times n}$. Entonces* $\det(A.B) = \det(A).\det(B)$.

Definición 2.70. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Sean $i, j \in \{1, 2, \ldots, n\}$.

- El *cofactor correspondiente al lugar (i, j) de la matriz A* es el número real

$$C_{ij} = (-1)^{i+j} \det(A_{(i|j)}) \ .$$

- La *matriz de los cofactores de A* es la matriz $\mathrm{cof}(A) \in \mathbb{R}^{n\times n}$ definida por

$$(\mathrm{cof}(A))_{ij} = C_{ij} \ .$$

- La *matriz adjunta de A* es la matriz $\mathrm{adj}(A) \in \mathbb{R}^{n\times n}$ definida por

$$\mathrm{adj}(A) = (\mathrm{cof}(A))^t \ .$$

Ejemplo 2.71. Sea $A = \begin{pmatrix} 2 & -3 & 5 \\ -1 & 0 & 4 \\ -2 & 1 & 3 \end{pmatrix}$.

Entonces

$$\mathrm{cof}(A) = \begin{pmatrix} -4 & -5 & -1 \\ 14 & 16 & 4 \\ -12 & -13 & -3 \end{pmatrix}$$

y luego

$$\mathrm{adj}(A) = \begin{pmatrix} -4 & 14 & -12 \\ -5 & 16 & -13 \\ -1 & 4 & -3 \end{pmatrix}$$

Observemos que $A.\mathrm{adj}(A) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ y que $\mathrm{adj}(A).A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

Además,

$$\det(A) = (-1)^{2+1}.(-1).(-14) + 0 + (-1)^{2+3}.4.(-4) = -14 + 16 = 2 \ .$$

Este ejemplo muestra algo que vale en general, como vemos en la siguiente proposición.

Proposición 2.72. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Entonces*

$$A.\,\mathrm{adj}(A) = \det(A).\,I = \mathrm{adj}(A).\,A \ .$$

Demostración. Para demostrar que $A.\,\mathrm{adj}(A) = \det(A).\,I$ debemos probar que para todos $i, j \in \{1, 2, \ldots, n\}$ se cumple que

$$(A.\,\mathrm{adj}(A))_{ij} = \begin{cases} \det(A) & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Sean $i, j \in \{1, 2, \ldots, n\}$. Supongamos primero que $i = j$. Entonces

$$\begin{aligned} (A.\,\mathrm{adj}(A))_{ij} &= (A.\,\mathrm{adj}(A))_{jj} = \sum_{k=1}^{n} A_{jk}(\mathrm{adj}(A))_{kj} = \sum_{k=1}^{n} A_{jk}(\mathrm{cof}(A))_{jk} = \\ &= \sum_{k=1}^{n} A_{jk}(-1)^{j+k}\det(A_{(j|k)}) = \det(A) \ . \end{aligned}$$

Analicemos ahora el caso en que $i \neq j$. Sea B la matriz que se obtiene reemplazando la fila j de A por la fila i de A. Es decir, tanto la fila i como la fila j de B coinciden con la fila i de A y el resto de las filas de B son iguales a las correspondientes filas de A.

Entonces para todo $k \in \{1, 2, \ldots, n\}$ se tiene que $B_{jk} = A_{ik}$ y que las matrices $B_{(j|k)}$ y $A_{(j|k)}$ son iguales. Luego

$$\begin{aligned} (A.\,\mathrm{adj}(A))_{ij} &= \sum_{k=1}^{n} A_{ik}(\mathrm{adj}(A))_{kj} = \sum_{k=1}^{n} A_{ik}(\mathrm{cof}(A))_{jk} = \\ &= \sum_{k=1}^{n} A_{ik}(-1)^{j+k}\det(A_{(j|k)}) = \sum_{k=1}^{n} B_{jk}(-1)^{j+k}\det(B_{(j|k)}) = \\ &= \det(B) = 0 \end{aligned}$$

ya que la matriz B tiene dos filas iguales.

Por lo tanto, $A.\,\mathrm{adj}(A) = \det(A).\,I$.

Análogamente se demuestra que $\mathrm{adj}(A).\,A = \det(A).\,I$. □

Teorema 2.73. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Entonces A es inversible si y sólo si $\det(A) \neq 0$. Además, en este caso se tiene que $A^{-1} = \frac{1}{\det(A)}.\mathrm{adj}(A)$.*

Demostración. ($\Rightarrow$) Supongamos que A es inversible. Entonces $A.A^{-1} = I$. Luego

$$\det(A.A^{-1}) = \det(I)$$

entonces por 2.69 tenemos que

$$\det(A).\det(A^{-1}) = 1 \ .$$

Por lo tanto, $\det(A) \neq 0$.

($\Leftarrow$) Supongamos que $\det(A) \neq 0$. Sea $B = \frac{1}{\det(A)}.\mathrm{adj}(A)$. Se tiene que

$$A.B = A.\left(\frac{1}{\det(A)}.\mathrm{adj}(A)\right) = \frac{1}{\det(A)}.(A.\mathrm{adj}(A)) = \frac{1}{\det(A)}.(\det(A).I) = I \ .$$

En forma similar

$$B.A = \left(\frac{1}{\det(A)}.\mathrm{adj}(A)\right).A = \frac{1}{\det(A)}.(\mathrm{adj}(A).A) = \frac{1}{\det(A)}.(\det(A).I) = I \ .$$

Por lo tanto, la matriz A es inversible.

Además, obtenemos que si $\det(A) \neq 0$ entonces $A^{-1} = \frac{1}{\det(A)}.\mathrm{adj}(A)$. □

De los teoremas 2.46 y 2.73 obtenemos el siguiente corolario.

Corolario 2.74. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Son equivalentes:*

(1) $\det(A) \neq 0$.

(2) Para todo $b \in \mathbb{R}^{n\times 1}$ el sistema $Ax^t = b$ tiene solución única.

Ejemplo 2.75. Sea $a \in \mathbb{R}$. Consideremos el siguiente sistema de ecuaciones:

$$\left\{\begin{array}{ccccc} x & + & 3y & = & a \\ 3x & + & a^2y & = & 9 \end{array}\right.$$

Hallar los valores de a para los cuales el sistema anterior tiene solución única, aquellos para los que tiene infinitas soluciones y aquellos para los que no tiene solución.

Solución:

La matriz asociada al sistema es

$$A = \begin{pmatrix} 1 & 3 \\ 3 & a^2 \end{pmatrix}$$

y su determinante es $\det(A) = a^2 - 9$.

Sabemos que el sistema tendrá solución única si y sólo si $\det(A) \neq 0$. Observamos que

$$\det(A) = 0 \Leftrightarrow a^2 - 9 = 0 \Leftrightarrow a = 3 \text{ o } a = -3 \ .$$

Por lo tanto, si $a \in \mathbb{R} - \{3, -3\}$ entonces el sistema tiene solución única.

Además, si $a = 3$ o $a = -3$ entonces $\det(A) = 0$ y por lo tanto el sistema no tiene solución única. Luego, si $a = 3$ o $a = -3$ el sistema puede ser compatible indeterminado o incompatible. Analicemos qué ocurre para estos valores de a.

Si $a = 3$ entonces la matriz ampliada asociada al sistema es

$$\left(\begin{array}{cc|c} 1 & 3 & 3 \\ 3 & 9 & 9 \end{array}\right)$$

y aplicando el método de eliminación de Gauss queda que

$$\left(\begin{array}{cc|c} 1 & 3 & 3 \\ 3 & 9 & 9 \end{array}\right) \xrightarrow[F_2 \leftarrow F_2 - 3F_1]{} \left(\begin{array}{cc|c} 1 & 3 & 3 \\ 0 & 0 & 0 \end{array}\right)$$

El sistema asociado a esta última matriz ampliada se encuentra en forma escalonada y tiene sólo una ecuación no trivial y dos incógnitas. Por lo tanto, si $a = 3$ el sistema tiene infinitas soluciones.

Si $a = -3$ entonces la matriz ampliada asociada al sistema es

$$\left(\begin{array}{cc|c} 1 & 3 & -3 \\ 3 & 9 & 9 \end{array}\right)$$

y aplicando el método de eliminación de Gauss queda que

$$\left(\begin{array}{cc|c} 1 & 3 & -3 \\ 3 & 9 & 9 \end{array}\right) \xrightarrow[F_2 \leftarrow F_2 - 3F_1]{} \left(\begin{array}{cc|c} 1 & 3 & 3 \\ 0 & 0 & 18 \end{array}\right)$$

El sistema asociado a esta última matriz ampliada se encuentra en forma escalonada y tiene una ecuación absurda. Por lo tanto, si $a = -3$ el sistema no tiene solución.

Capítulo 3

Espacios vectoriales

Definición 3.1. Un *espacio vectorial sobre* $\mathbb{R}$ (o $\mathbb{R}$-*espacio vectorial*) es un conjunto $\mathbb{V}$ junto con dos operaciones $+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ y $. : \mathbb{R} \times \mathbb{V} \to \mathbb{V}$ (llamadas *suma* y *producto por escalares*, respectivamente) que cumplen las siguientes propiedades:

(1) $\forall\, u, v, w \in \mathbb{V},\ (u + v) + w = u + (v + w)$.

(2) $\forall\, v, w \in \mathbb{V},\ v + w = w + v$.

(3) $\exists\ 0 \in \mathbb{V} /\ \forall\, v \in \mathbb{V},\ 0 + v = v$ y $v + 0 = v$.

(4) $\forall\, v \in \mathbb{V},\ \exists - v \in \mathbb{V} /\ (-v) + v = 0$ y $v + (-v) = 0$.

(5) $\forall\, v \in \mathbb{V},\ 1.v = v$.

(6) $\forall\, \alpha \in \mathbb{R},\ \forall\, v, w \in \mathbb{V},\ \alpha.(v + w) = \alpha.v + \alpha.w$.

(7) $\forall\, \alpha, \beta \in \mathbb{R},\ \forall\, v \in \mathbb{V},\ (\alpha + \beta).v = \alpha.v + \beta.v$.

(8) $\forall\, \alpha, \beta \in \mathbb{R},\ \forall\, v \in \mathbb{V},\ \alpha.(\beta.v) = (\alpha.\beta).v$.

Los elementos de $\mathbb{V}$ se llaman *vectores*.

Notación 3.2. Notaremos $u - v = u + (-v)$.

Ejemplos 3.3.

(1) Sea $n \in \mathbb{N}$. Entonces $\mathbb{R}^n$ con las operaciones usuales, es un $\mathbb{R}$-espacio vectorial.

(2) Sea $n, m \in \mathbb{N}$. Entonces $\mathbb{R}^{m \times n}$ con las operaciones usuales, es un $\mathbb{R}$-espacio vectorial.

(3) El conjunto $\mathbb{R}[x]$ de polinomios con coeficientes reales con las operaciones usuales, es un $\mathbb{R}$-espacio vectorial.

(4) El conjunto de polinomios de grado menor o igual a n, con las operaciones usuales es un $\mathbb{R}$-espacio vectorial.

Observación 3.4. Trabajaremos con espacios vectoriales sobre $\mathbb{R}$ por lo cual diremos simplemente que $\mathbb{V}$ es un espacio vectorial para indicar que es un $\mathbb{R}$-espacio vectorial.

Proposición 3.5. *Sea* $\mathbb{V}$ *un espacio vectorial. Entonces*

(1) $0.v = 0$.

(2) $\alpha.0 = 0$.

(3) $(-1).v = -v$.

(4) $\forall\, v, w \in \mathbb{V},\ -(v + w) = (-v) + (-w)$.

(5) $\forall\, \alpha \in \mathbb{R}, \forall\, v, w \in \mathbb{V},\ \alpha.(v - w) = \alpha.v - \alpha.w$.

(6) Si $\alpha \in \mathbb{R}, v \in \mathbb{V}$*, entonces* $\alpha.v = 0$ *si y sólo si* $\alpha = 0$ *o* $v = 0$.

Demostración.

(1) Sea $v \in \mathbb{V}$. Tenemos que

$$0.v = (0 + 0).v = 0.v + 0.v\ .$$

Luego,

$$0.v + (-(0.v)) = (0.v + 0.v) + (-(0.v)) = 0.v + (0.v + (-(0.v)).$$

y entonces $0 = 0.v + 0$ de donde se obtiene que $0 = 0.v$.

(2) Tenemos que

$$\alpha.0 = \alpha.(0 + 0) = \alpha.0 + \alpha.0.$$

Luego,

$$\alpha.0 + (-(\alpha.0)) = (\alpha.0 + \alpha.0) + (-(\alpha.0)).$$

Entonces

$$0 = \alpha.0 + (\alpha.0 + (-(\alpha.0))) = \alpha.0 + 0$$

y por lo tanto

$$0 = \alpha.0.$$

(3) Notemos que $(-1).v + v = 0$ ya que

$$(-1).v + v = (-1).v + 1.v = ((-1) + 1).v = 0.v = 0$$

Entonces,

$$(-1).v = (-1).v + 0 = (-1).v + (v + (-v)) = ((-1).v + v) + (-v) = 0 + (-v) = -v.$$

Por lo tanto, $(-1).v = -v$.

Las demostracioes de los ítems (4), (5) y (6) quedan como ejercicio para el lector. $\square$

Definición 3.6. Sea $\mathbb{V}$ un espacio vectorial y sea $S \subseteq \mathbb{V}$ un subconjunto. Decimos que S es un *subespacio vectorial de* $\mathbb{V}$ (o simplemente un *subespacio de* $\mathbb{V}$) si S con las operaciones de $+$ y $.$ heredadas de $\mathbb{V}$ es un espacio vectorial.

Ejemplo 3.7. El conjunto de polinomios de grado menor o igual a 3 es un subespacio de $\mathbb{R}[x]$.

Proposición 3.8. *Sea* $\mathbb{V}$ *un espacio vectorial y sea* $S \subseteq \mathbb{V}$*. Entonces* S *es un subespacio de* $\mathbb{V}$ *si y sólo si cumple las siguientes 3 condiciones:*

(i) $0 \in S$.

(ii) $\forall\, v, w \in S,\ v + w \in S$.

(iii) $\forall\, \alpha \in \mathbb{R},\ \forall\, v \in S,\ \alpha.v \in S$.

Demostración. ($\Rightarrow$) De la definición de subespacio se obtiene que las operaciones de suma y producto por escalares de $\mathbb{V}$ inducen operaciones $+$ y $.$ en S, es decir, $+ : S \times S \to S$ y $. : \mathbb{R} \times S \to S$. Por lo tanto, para todos $v, w \in S$ vale que $v + w \in S$ y para todos $\alpha \in \mathbb{R}$ y $v \in S$ vale que $\alpha.v \in S$. Así, se verifican las condiciones (ii) y (iii).

Por otra parte, como S es un espacio vectorial, entonces la suma de vectores de S (que es la suma heredada de $\mathbb{V}$) debe tener elemento neutro. En particular S es no vacío. Entonces existe $s \in S$. Luego, $(-1).s \in S$ y entonces $0 = s + (-1).s \in S$. Por lo tanto, la condición (i) también se verifica.

($\Leftarrow$) Supongamos que se cumplen las 3 condiciones. Por *(ii)* y *(iii)* las operaciones suma y producto por escalar de $\mathbb{V}$ inducen operaciones $+ : S \times S \to S$ y $. : \mathbb{R} \times S \to S$.

Debemos probar que se cumplen las propiedades 1 a 8 de la definición 3.1.

(1) Se hereda de $\mathbb{V}$ (como la suma en $\mathbb{V}$ es asociativa, en S también lo es).

(2) Se hereda de $\mathbb{V}$.

(3) Por la condición *(i)*, $0 \in S$, luego 0 es el elemento neutro de la suma.

(4) Si $v \in S$ entonces $-v = (-1).v \in S$ por la condición *(iii)*.

(5), (6), (7) y (8) se heredan de $\mathbb{V}$. □

Ejemplo 3.9. Sea $S = \{(x, y) \in \mathbb{R}^2 / y = 2.x\}$. Demostrar que S es un subespacio de $\mathbb{R}^2$.

Solución: Veamos que se cumplen las condiciones (i), (ii) y (iii) de la proposición 3.8.

(i) $(0, 0) \in S$ porque $0 = 2.0$.

(ii) Sean $v = (v_1, v_2)$ y $w = (w_1, w_2) \in S$. Veamos que $v + w = (v_1 + w_1, v_2 + w_2) \in S$. Como $v \in S$ entonces $v_2 = 2v_1$ y como $w \in S$ entonces $w_2 = 2w_1$. Luego tenemos que
$$v_2 + w_2 = 2v_1 + 2w_1 = 2(v_1 + w_1).$$
Entonces, $v + w \in S$.

(iii) Sea $\alpha \in \mathbb{R}$ y $v = (v_1, v_2) \in S$. Veamos que $\alpha.v = (\alpha.v_1, \alpha.v_2) \in S$. Como $v \in S$ entonces $v_2 = 2v_1$. Luego tenemos que $\alpha.v_2 = \alpha.(2.v_1) = 2.(\alpha.v_1)$. Entonces, $\alpha.v \in S$.

Por lo tanto, S es un subespacio de $\mathbb{R}^2$.

Definición 3.10. Sea $\mathbb{V}$ un espacio vectorial. Sean $v_1, v_2, \ldots, v_n \in \mathbb{V}$ una *combinación lineal de los vectores* $v_1, v_2, \ldots, v_n$ es una expresión del tipo $\alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n$ donde $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$.

Ejemplo 3.11. Escribir $(2, 6)$ como combinación lineal de los vectores $(1, 0)$ y $(3, 2)$.

Solución: Queremos hallar números reales α y β tales que
$$(2, 6) = \alpha.(1, 0) + \beta.(3, 2)$$
Queremos entonces que $(2, 6) = (\alpha + 3\beta, 2\beta)$. Luego,
$$\begin{cases} \alpha + 3\beta = 2 \\ \quad\;\; 2\beta = 6 \end{cases}$$
Entonces $\beta = 3$ y luego $\alpha = -7$. Por lo tanto,
$$(2, 6) = -7.(1, 0) + 3.(3, 2)\ .$$

Definición 3.12. Sea $\mathbb{V}$ un espacio vectorial. Sean $v_1, v_2, \ldots, v_n \in \mathbb{V}$. Decimos que el conjunto $\{v_1, v_2, \ldots, v_n\}$ *genera* $\mathbb{V}$ (o que es *un conjunto de generadores de* $\mathbb{V}$) si todo vector de $\mathbb{V}$ se puede escribir como combinación lineal de los vectores $v_1, v_2, \ldots, v_n$.

Ejemplos 3.13.

(1) El conjunto $\{(1,0,0),(0,1,0),(0,0,1)\}$ genera $\mathbb{R}^3$. En efecto, si $(x,y,z)\in\mathbb{R}^3$ entonces

$$(x,y,z) = x.(1,0,0) + y.(0,1,0) + z.(0,0,1) \quad .$$

Por lo tanto, todo vector de $\mathbb{R}^3$ se puede escribir como combinación lineal de $(1,0,0)$, $(0,1,0)$ y $(0,0,1)$.

(2) El conjunto $(1,-3),(-2,6)$ no genera $\mathbb{R}^2$.

Veamos que el vector $(1,0)$ no es combinación lineal de $(1,-3)$ y $(-2,6)$. En efecto, si lo fuera existirían $\alpha,\beta\in\mathbb{R}$, tales que $(1,0)=\alpha\,.(1,-3)+\beta\,.(-2,6)$, es decir, $(1,0)=(\alpha-2\beta,-3\alpha+6\beta)$. Luego,

$$\begin{cases} \alpha - 2\beta = 1 \\ -3\alpha + 6\beta = 0 \end{cases}$$

Aplicando el método de eliminación de Gauss obtenemos

$$\left(\begin{array}{cc|c} 1 & -2 & 1 \\ -3 & 6 & 0 \end{array}\right) \underset{F_2\leftarrow F_2+3.F_1}{\rightsquigarrow} \left(\begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 0 & 3 \end{array}\right)$$

Entonces, el sistema no tiene solución.

Luego, no existen $\alpha,\beta\in\mathbb{R}$, tales que $(1,0)=\alpha.(1,-3)+\beta.(-2,6)$. Entonces $(1,0)$ no es combinación lineal de $(1,-3)$ y $(-2,6)$.

Por lo tanto, el conjunto $\{(1,-3),(-2,6)\}$ no genera $\mathbb{R}^2$.

Definición 3.14. Sea $\mathbb{V}$ un espacio vectorial y sean $v_1,v_2,\ldots,v_n\in\mathbb{V}$. El *subespacio generado por los vectores* $v_1,v_2,\ldots,v_n$ es el subespacio $\mathrm{Gen}(\{v_1,v_2,\ldots,v_n\})$ definido por

$$\mathrm{Gen}(\{v_1,v_2,\ldots,v_n\}) = \{\alpha_1.v_1+\alpha_2.v_2+\ldots+\alpha_n.v_n \,/\, \alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{R}\}.$$

También notaremos $\mathrm{Gen}(\{v_1,v_2,\ldots,v_n\}) = \langle v_1,v_2,\ldots,v_n\rangle$.

Ejercicio: Demostrar que si $v_1,v_2,\ldots,v_n\in\mathbb{V}$ entonces $\mathrm{Gen}(\{v_1,v_2,\ldots,v_n\})$ es efectivamente un subespacio de V.

Ejemplo 3.15. Sea $S=\{(3r+t,2r,t-r,r,t)/r,t\in\mathbb{R}\}$. Hallar un conjunto de generadores para S.

Solución: Notemos que

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

Entonces, todo elemento de S es combinación lineal de $(3,2,-1,1,0)$ y $(1,0,1,0,1)$. Por lo tanto,

$$S = \mathrm{Gen}(\{(3,2,-1,1,0),(1,0,1,0,1)\}).$$

Definición 3.16. Sea $\mathbb{V}$ un espacio vectorial. Decimos que $\mathbb{V}$ es *finitamente generado* si existen $v_1,v_2,\ldots,v_n\in\mathbb{V}$ tales que el conjunto $\{v_1,v_2,\ldots,v_n\}$ genera $\mathbb{V}$.

Definición 3.17. Sea $\mathbb{V}$ un espacio vectorial y sean $v_1,v_2,\ldots,v_n\in\mathbb{V}$. Decimos que el conjunto $\{v_1,v_2,\ldots,v_n\}$ es *linealmente independiente* si se cumple que

$$\forall\,\alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{R},\ \alpha_1.v_1+\alpha_2.v_2+\ldots+\alpha_n.v_n=0 \ \Rightarrow\ \alpha_1=0,\alpha_2=0,\ldots,\alpha_n=0.$$

Diremos que el conjunto $\{v_1,v_2,\ldots,v_n\}$ es *linealmente dependiente* si no es linealmente independiente.

Ejemplo 3.18. Decidir si el conjunto $\{(1,2),(4,8)\}$ es linealmente independiente o linealmente dependiente.
Solución: Observamos que $-4.(1,2)+(4,8)=(0,0)$. Luego el conjunto anterior es linealmente dependiente.

Si no encontramos los escalares a ojo, planteamos

$$\alpha.(1,2)+\beta.(4,8)=(0,0)$$

Luego,

$$(\alpha+4.\beta,2.\alpha+8.\beta)=(0,0)$$

y se obtiene un sistema de ecuaciones lineales:

$$\begin{cases} \alpha+4\beta=0 \\ 2\alpha+8\beta=0 \end{cases}$$

Resolvemos el sistema anterior por el método de eliminación de Gauss:

$$\left(\begin{array}{cc|c} 1 & 4 & 0 \\ 2 & 8 & 0 \end{array}\right) \underset{F_2 \leftarrow F_2-2.F_1}{\rightsquigarrow} \left(\begin{array}{cc|c} 1 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Obtenemos que la solución general del sistema es $(-4\beta,\beta),\ \beta\in\mathbb{R}$. Y por ejemplo, haciendo $\beta=1$ obtenemos que $\alpha=-4$ y luego

$$-4.(1,2)+(4,8)=(0,0)$$

como ya habíamos observado antes.

Observación 3.19. Si $v\neq 0$ entonces el conjunto $\{v\}$ es linealmente independiente.

En efecto, si $\alpha.v=0$, como $v\neq 0$ entonces, por el ítem (6) de 3.5 obtenemos que $\alpha=0$.

Observación 3.20. Si un conjunto de vectores contiene al vector 0, entonces es linealmente dependiente. En efecto, si tenemos $\{v_1,v_2,\ldots,v_n,0\}$ entonces el conjunto es linealmente dependiente, pues

$$0.v_1+0.v_2+\ldots+0.v_n+1.0=0.$$

Proposición 3.21. *Sea $\mathbb{V}$ un espacio vectorial y sean $v_1,v_2,\ldots,v_n\in\mathbb{V}$. Entonces el conjunto $\{v_1,v_2,\ldots,v_n\}$ es linealmente dependiente si y sólo si uno de los vectores $v_1,\ldots,v_n$ es combinación lineal de los otros.*

Demostración.

($\Rightarrow$) Supongamos que el conjunto $\{v_1,v_2,\ldots,v_n\}$ es linealmente dependiente. Entonces existen $\alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{R}$ no todos nulos tales que $\alpha_1.v_1+\alpha_2.v_2+\ldots+\alpha_n.v_n=0$.

Luego existe $k\in\{1,2,\ldots,n\}$ tal que $\alpha_k\neq 0$. Entonces

$$\alpha_k.v_k=-\alpha_1.v_1-\alpha_2.v_2-\ldots-\alpha_{k-1}.v_{k-1}-\alpha_{k+1}.v_{k+1}-\ldots-\alpha_n.v_n$$

y luego

$$v_k=-\frac{\alpha_1}{\alpha_k}.v_1-\frac{\alpha_2}{\alpha_k}.v_2-\ldots-\frac{\alpha_{k-1}}{\alpha_k}.v_{k-1}-\frac{\alpha_{k+1}}{\alpha_k}.v_{k+1}-\ldots-\frac{\alpha_n}{\alpha_k}.v_n$$

($\Leftarrow$) Podemos suponer, sin pérdida de generalidad, que v_n es combinación lineal de $v_1,v_2,\ldots,v_{n-1}$. Entonces, existen $\alpha_1,\alpha_2,\ldots,\alpha_{n-1}\in\mathbb{R}$ tales que

$$v_n=\alpha_1.v_1+\alpha_2.v_2+\ldots+\alpha_{n-1}.v_{n-1}$$

Luego,
$$0 = \alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_{n-1}.v_{n-1} + (-1).v_n$$
Entonces, 0 es combinación lineal de $v_1, v_2, \ldots, v_n$ con escalares no todos nulos. Por lo tanto, el conjunto $\{v_1, v_2, \ldots, v_n\}$ es linealmente dependiente. $\square$

Proposición 3.22. *Sea $\mathbb{V}$ un espacio vectorial y sean $v_1, v_2, \ldots, v_n \in \mathbb{V}$ tales que $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente. Sea $w \in \mathbb{V}$. Si $\{v_1, v_2, \ldots, v_n, w\}$ es linealmente dependiente, entonces w es combinación lineal de $v_1, v_2, \ldots, v_n$.*

Demostración. Como el conjunto $\{v_1, v_2, \ldots, v_n, w\}$ es linealmente dependiente, existen $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_{n+1} \in \mathbb{R}$, no todos nulos, tales que
$$\alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n + \alpha_{n+1}.w = 0.$$

Si $\alpha_{n+1} = 0$ entonces queda que $\alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n = 0$ donde los números reales $\alpha_1, \alpha_2, \ldots, \alpha_n$ no son todos nulos. Luego $\{v_1, v_2, \ldots, v_n\}$ es linealmente dependiente, lo que es absurdo por hipótesis. Por lo tanto, $\alpha_{n+1} \neq 0$.

Ahora bien, teníamos que
$$\alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n + \alpha_{n+1}.w = 0$$
luego
$$\alpha_{n+1}.w = -\alpha_1.v_1 - \alpha_2.v_2 - \ldots - \alpha_n.v_n$$
y como $\alpha_{n+1} \neq 0$, obtenemos que
$$w = -\frac{\alpha_1}{\alpha_{n+1}}.v_1 - \frac{\alpha_2}{\alpha_{n+1}}.v_2 - \ldots - \frac{\alpha_n}{\alpha_{n+1}}.v_n$$
y entonces w es combinación lineal de $v_1, v_2, \ldots, v_n$. $\square$

Proposición 3.23. *Sea $\mathbb{V}$ un espacio vectorial y sean $v_1, v_2, \ldots, v_n \in \mathbb{V}$ tales que $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente. Sea $w \in \mathbb{V}$ tal que w es combinación lineal de $v_1, v_2, \ldots, v_n$. Entonces w se puede escribir de una única manera como combinación lineal de $v_1, v_2, \ldots, v_n$.*

Demostración. Supongamos que existen $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n \in \mathbb{R}$ tales que
$$w = \alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n$$
y
$$w = \beta_1.v_1 + \beta_2.v_2 + \ldots + \beta_n.v_n$$
Entonces,
$$\alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n = \beta_1.v_1 + \beta_2.v_2 + \ldots + \beta_n.v_n\ .$$
Luego,
$$(\alpha_1 - \beta_1).v_1 + (\alpha_2 - \beta_2).v_2 + \ldots + (\alpha_n - \beta_n).v_n = 0$$
y como $\{v_1, v_2, \ldots, v_n\}$ es un conjunto linealmente independiente obtenemos que
$$\alpha_1 - \beta_1 = 0,\ \alpha_2 - \beta_2 = 0,\ \ldots\ ,\ \alpha_n - \beta_n = 0\ .$$
Por lo tanto,
$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2,\ \ldots\ ,\ \alpha_n = \beta_n\ .$$
$\square$

Teorema 3.24. *Sea $\mathbb{V}$ un espacio vectorial. Sean $v_1, v_2, \ldots, v_n \in \mathbb{V}$ tales que el conjunto $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente y sean $w_1, w_2, \ldots, w_m \in \mathbb{V}$ tales que el conjunto $\{w_1, w_2, \ldots, w_m\}$ genera $\mathbb{V}$. Entonces $n \leqslant m$.*

Demostración. Como $\{w_1, w_2, \ldots, w_m\}$ genera $\mathbb{V}$, entonces existen $\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1m}$, $\alpha_{21}, \alpha_{22}, \ldots, \alpha_{2m}, \ldots, \alpha_{n1}, \alpha_{n2}, \ldots, \alpha_{nm} \in \mathbb{R}$ tales que

$$\begin{array}{ccccccccc} v_1 & = & \alpha_{11}w_1 & + & \alpha_{12}w_2 & + & \ldots & + & \alpha_{1m}w_m \\ v_2 & = & \alpha_{21}w_1 & + & \alpha_{22}w_2 & + & \ldots & + & \alpha_{2m}w_m \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ v_n & = & \alpha_{n1}w_1 & + & \alpha_{n2}w_2 & + & \ldots & + & \alpha_{nm}w_m \end{array}$$

Sean $\gamma_1, \gamma_2, \ldots, \gamma_n \in \mathbb{R}$. Se tiene que

$$\begin{aligned} \gamma_1 v_1 + \gamma_2 v_2 + \ldots + \gamma_n v_n &= \gamma_1(\alpha_{11}w_1 + \ldots + \alpha_{1m}w_m) + \gamma_2(\alpha_{21}w_1 + \ldots + \alpha_{2m}w_m) + \\ &\quad + \ldots + \gamma_n(\alpha_{n1}w_1 + \ldots + \alpha_{nm}w_m) = \\ &= (\gamma_1\alpha_{11} + \gamma_2\alpha_{21} + \ldots + \gamma_n\alpha_{n1})w_1 + \ldots + \\ &\quad + (\gamma_1\alpha_{1m} + \gamma_2\alpha_{2m} + \ldots + \gamma_n\alpha_{nm})w_m \end{aligned}$$

Por lo tanto, si

$$\begin{array}{cccccccccc} \gamma_1\alpha_{11} & + & \gamma_2\alpha_{21} & + & \ldots & + & \gamma_n\alpha_{n1} & = & 0 & \text{y} \\ \gamma_1\alpha_{12} & + & \gamma_2\alpha_{22} & + & \ldots & + & \gamma_n\alpha_{n2} & = & 0 & \text{y} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots & \text{y} \\ \gamma_1\alpha_{1m} & + & \gamma_2\alpha_{2m} & + & \ldots & + & \gamma_n\alpha_{nm} & = & 0 & \end{array}$$

entonces

$$\gamma_1 v_1 + \gamma_2 v_2 + \ldots + \gamma_n v_n = 0$$

y como $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente, se obtendría que

$$\gamma_1 = 0,\ \gamma_2 = 0,\ \ldots\ ,\ \gamma_n = 0.$$

Por lo tanto, el sistema

$$\left\{ \begin{array}{ccccccccc} \gamma_1\alpha_{11} & + & \gamma_2\alpha_{21} & + & \ldots & + & \gamma_n\alpha_{n1} & = & 0 \\ \gamma_1\alpha_{12} & + & \gamma_2\alpha_{22} & + & \ldots & + & \gamma_n\alpha_{n2} & = & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ \gamma_1\alpha_{1m} & + & \gamma_2\alpha_{2m} & + & \ldots & + & \gamma_n\alpha_{nm} & = & 0 \end{array} \right.$$

tiene solución única $\gamma_1 = 0,\ \gamma_2 = 0,\ \ldots,\ \gamma_n = 0$.

Ahora bien, si $m < n$, como el sistema anterior es homogéneo y tiene m ecuaciones y n incógnitas, entonces luego de escalonarlo quedaría sin ecuaciones absurdas y con menos ecuaciones no triviales que incógnitas y, por ende, tendría infinitas soluciones, lo cual es absurdo. Por lo tanto, debe ocurrir que $m \geqslant n$. $\square$

Ejemplo 3.25. Sabemos que $\{(1,0,0), (0,1,0), (0,0,1)\}$ genera $\mathbb{R}^3$. Entonces todo conjunto linealmente independiente de $\mathbb{R}^3$ debe tener, a lo sumo, 3 elementos.

Por lo tanto, un conjunto de 4 vectores en $\mathbb{R}^3$ no puede ser linealmente independiente.

Definición 3.26. Sea $\mathbb{V}$ un espacio vectorial. Una *base de* $\mathbb{V}$ es un conjunto (ordenado) de vectores de $\mathbb{V}$ que es linealmente independiente y que genera $\mathbb{V}$.

Ejemplos 3.27.
(1) Vimos que $\{(1,0,0),(0,1,0),(0,0,1)\}$ genera $\mathbb{R}^3$. Además, es claro que este conjunto es linealmente independiente. Por lo tanto, $\{(1,0,0),(0,1,0),(0,0,1)\}$ es base de $\mathbb{R}^3$. Este conjunto llama *base canónica* de $\mathbb{R}^3$.

El conjunto $\{(0,1,0),(1,0,0),(0,0,1)\}$ también es linealmente independiente y también genera $\mathbb{R}^3$ por lo tanto, también es base de $\mathbb{R}^3$. Pero es una base distinta a la anterior, porque cuando trabajamos con bases debemos tener en cuenta el orden de los elementos.
(2) El conjunto $\{(1,0),(0,1)\}$ es la *base canónica de* $\mathbb{R}^2$.
(3) El conjunto $\{(1,0,0,\dots,0)\ ,\ (0,1,0,\dots,0)\ ,\ \dots\ ,\ (0,0,\dots,0,1)\}$ es la *base canónica de* $\mathbb{R}^n$.
(4) El conjunto $\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}$ es una base de $\mathbb{R}^{2\times 2}$.

Observación 3.28. Sea $\mathbb{V}$ un espacio vectorial y sea B una base de $\mathbb{V}$. Entonces, por 3.23, tenemos que todo vector de $\mathbb{V}$ se puede escribir de una única manera como combinación lineal de los elementos de B.

Teorema 3.29. *Sea $\mathbb{V}$ un espacio vectorial finitamente generado. Entonces $\mathbb{V}$ admite una base. Más aún, si $A=\{v_1,v_2,\dots,v_n\}$ es un conjunto linealmente independiente de vectores de $\mathbb{V}$ y $G=\{w_1,w_2,\dots,w_m\}$ es un conjunto generador de $\mathbb{V}$ entonces existe un subconjunto $H\subseteq G$ tal que $A\cup H$ es una base de $\mathbb{V}$.*

Demostración. Definamos un conjunto B_1 de vectores de $\mathbb{V}$ de la siguiente manera. Si w_1 es combinación lineal de los elementos de A, definimos $B_1=A$. Si w_1 no es combinación lineal de los elementos de A, definimos $B_1=A\cup\{w_1\}$.

Notemos que, en cualquier caso, B_1 es linealmente independiente. En efecto, si w_1 es combinación lineal de los elementos de A entonces $B_1=A$ y claramente B_1 es linealmente independiente pues A lo es por hipótesis. Y si w_1 no es combinación lineal de los elementos de A entonces $B_1=A\cup\{w_1\}$ y si B_1 fuera linealmente dependiente, como A es linealmente independiente, entonces, por la proposición 3.22, w_1 sería combinación lineal de los elementos de A, lo que es absurdo. Por lo tanto, B_1 es linealmente independiente.

Definimos ahora B_2 de la siguiente manera:

$$B_2=\begin{cases}B_1 & \text{si } w_2 \text{ es combinación lineal de los elementos de } B_1,\\ B_1\cup\{w_2\} & \text{si no}\end{cases}$$

En forma similar a lo hecho con B_1, se puede probar que el conjunto B_2 es linealmente independiente.

Inductivamente, supongamos que tenemos definido B_k para cierto $k\leq m-1$ y que B_k es linealmente independiente. Definimos B_{k+1} por

$$B_{k+1}=\begin{cases}B_k & \text{si } w_{k+1} \text{ es combinación lineal de los elementos de } B_k,\\ B_k\cup\{w_{k+1}\} & \text{si no}\end{cases}$$

Como antes, B_{k+1} resulta linealmente independiente.

Tenemos entonces definidos $B_1,B_2,\dots,B_m$. Sea $B=B_m$. Es claro que $B=A\cup H$, para algún subconjunto $H\subseteq G$.

Veamos que B es base de $\mathbb{V}$. Por construcción, B es linealmente independiente. Veamos que B genera $\mathbb{V}$.

Afirmamos que los vectores $w_1,w_2,\dots,w_m$ son combinación lineal de los elementos de B. En efecto, sea $i\in\{1,2,\dots,m\}$. Si $w_i\in B$, claramente w_i es combinación lineal de los elementos de B. Si $w_i\notin B$, entonces w_i es combinación lineal de los elementos de B_{i-1}, y como $B_{i-1} \subseteq B$ entonces w_i es combinación lineal de los elementos de B. Por lo tanto, w_i es combinación lineal de los elementos de B, $\forall i \in \{1, 2, \ldots, m\}$.

Ahora, dado $v \in \mathbb{V}$, sabemos que v es combinación lineal de $w_1, w_2, \ldots, w_m$, pues el conjunto $\{w_1, w_2, \ldots, w_m\}$ genera $\mathbb{V}$, y como cada w_i es combinación lineal de los elementos de B, entonces v será combinación lineal de los elementos de B.

Por lo tanto, B genera $\mathbb{V}$. Luego, B es base de $\mathbb{V}$ □

Corolario 3.30. *Sea $\mathbb{V}$ un espacio vectorial finitamente generado.*

(1) Si $\{v_1, v_2, \ldots, v_n\}$ es un conjunto linealmente independiente de vectores de $\mathbb{V}$, entonces $\{v_1, v_2, \ldots, v_n\}$ se puede extender a una base de $\mathbb{V}$.

(2) Si el conjunto $\{w_1, w_2, \ldots, w_n\}$ genera $\mathbb{V}$, entonces de $\{w_1, w_2, \ldots, w_n\}$ se puede extraer una base de $\mathbb{V}$.

Demostración.

(1) Elijo un conjunto $\{w_1, w_2, \ldots, w_n\}$ que genere a $\mathbb{V}$ (tal conjunto existe porque $\mathbb{V}$ está finitamente generado). Aplicando el teorema 3.29 con $A = \{v_1, v_2, \ldots, v_n\}$ como conjunto linealmente independiente y $G = \{w_1, w_2, \ldots, w_m\}$ como conjunto generador obtenemos que existe $H \subseteq G$ tal que $A \cup H$ es una base de $\mathbb{V}$.

Por lo tanto, hemos extendido A a una base de $\mathbb{V}$.

(2) Sea $G = \{w_1, w_2, \ldots, w_m\}$ un conjunto generador para $\mathbb{V}$. Tomando $A = \varnothing$ (que es un conjunto linealmente independiente) y aplicando el teorema 3.29 obtenemos que existe $H \subseteq G$ tal que $A \cup H$ es base de $\mathbb{V}$. Pero $A \cup H = \varnothing \cup H = H$. Es decir, H es base de $\mathbb{V}$.

Por lo tanto, del conjunto G dado hemos extraído un conjunto H que es base de $\mathbb{V}$. □

Proposición 3.31. *Sea $\mathbb{V}$ un espacio vectorial finitamente generado. Sean B y B' bases de $\mathbb{V}$. Entonces $\#B = \#B'$.*

Demostración. Tenemos que B es un conjunto linealmente independiente y B' es un conjunto generador para $\mathbb{V}$. Luego, por el teorema 3.24, se obtiene que $\#B \leqslant \#B'$.

Por otra parte, B' es un conjunto linealmente independiente y B es un conjunto generador para $\mathbb{V}$. Luego, aplicando nuevamente el teorema 3.24 se obtiene que $\#B' \leqslant \#B$.

Por lo tanto, $\#B = \#B'$. □

Definición 3.32. Sea $\mathbb{V}$ un espacio vectorial finitamente generado. Definimos la *dimensión* de $\mathbb{V}$ como la cantidad de elementos de alguna base de $\mathbb{V}$ y la notaremos $\dim(\mathbb{V})$.

Si $\mathbb{V}$ es finitamente generado también diremos que $\mathbb{V}$ *tiene dimensión finita*.

Ejemplos 3.33.

(1) $\dim(\mathbb{R}^3) = 3$.

(2) Si $n \in \mathbb{N}$ entonces $\dim(\mathbb{R}^n) = n$.

(3) $\dim(\mathbb{R}^{2\times 2}) = 4$.

(4) Si $m, n \in \mathbb{N}$ entonces $\dim(\mathbb{R}^{m\times n}) = m.n$.

Proposición 3.34. *Sea $\mathbb{V}$ un espacio vectorial de dimensión finita. Sea $n = \dim(\mathbb{V})$.*

(1) Si A es un conjunto linealmente independiente de $\mathbb{V}$ con n elementos entonces A es base de $\mathbb{V}$.

(2) Si G es un conjunto generador de $\mathbb{V}$ con n elementos entonces G es base de $\mathbb{V}$.

Demostración.

(1) Como A es un conjunto linealmente independiente, entonces, por el corolario 3.30, A se puede extender a una base de $\mathbb{V}$. Es decir, existe una base B de $\mathbb{V}$ tal que $A \subseteq B$.

Pero $\#A = n = \dim(\mathbb{V}) = \#B$ y como $A \subseteq B$ obtenemos que $A = B$. Por lo tanto, A es base de $\mathbb{V}$.

(2) Como G es un conjunto generador de $\mathbb{V}$, entonces, por el corolario 3.30, de G se puede extraer una base de $\mathbb{V}$. Es decir, existe una base B de $\mathbb{V}$ tal que $B \subseteq G$.

Pero $\#G = n = \dim(\mathbb{V}) = \#B$ y como $G \supseteq B$ obtenemos que $G = B$. Por lo tanto, G es base de $\mathbb{V}$. □

Coordenadas

Definición 3.35. Sea $\mathbb{V}$ un espacio vectorial y sea $B = \{v_1, v_2, \ldots, v_n\}$ una base de $\mathbb{V}$. Sea $w \in \mathbb{V}$. Sabemos que existen únicos $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tales que

$$w = \alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n \ .$$

Los números reales $\alpha_1, \alpha_2, \ldots, \alpha_n$ se llaman *coordenadas* de w en la base B. Más precisamente, diremos que el *vector de coordenadas* de w en la base B es el vector $[w]_B \in \mathbb{R}^n$ definido por

$$[w]_B = (\alpha_1, \alpha_2, \ldots, \alpha_n) \ .$$

Ejemplo 3.36. Sea $\mathbb{V} = \mathbb{R}^2$, sea $B = \{(1,2),(3,4)\}$ y sea $w = (5,6)$. Hallar las coordenadas de w en la base B.

Solución: Escribimos $(5,6)$ como combinación lineal de $(1,2)$ y $(3,4)$.

$$(5,6) = \alpha.(1,2) + \beta.(3,4)$$

$$(5,6) = (\alpha + 3\beta, 2.\alpha + 4\beta)$$

Entonces,

$$\begin{cases} \alpha + 3\beta = 5 \\ 2\alpha + 4\beta = 6 \end{cases}$$

Ahora resolvemos el sistema por el método de eliminación de Gauss:

$$\left(\begin{array}{cc|c} 1 & 3 & 5 \\ 2 & 4 & 6 \end{array}\right) \xrightarrow[F_2 \leftarrow F_2 - 2.F_1]{} \left(\begin{array}{cc|c} 1 & 3 & 5 \\ 0 & -2 & -4 \end{array}\right)$$

Entonces obtenemos

$$\begin{cases} \alpha + 3\beta = 5 \\ \quad\ -2\beta = -4 \end{cases}$$

Luego $\beta = 2$ y reemplazando en la primer ecuación y despejando α obtenemos que $\alpha = -1$.

Por lo tanto, $[w]_B = (-1,2)$.

Notemos que si $B' = \{(3,4),(1,2)\}$ entonces $[w]_{B'} = (2,-1)$.

Ejemplo 3.37. Consideremos la misma base B del ejemplo anterior y sea $u \in \mathbb{R}^2$ tal que $[u]_B = (3,-1)$. Calcular u.

Solución: Como $[u]_B = (3,-1)$ entonces $u = 3.(1,2) - 1.(3,4) = (3,6) + (-3,-4) = (0,2)$. Luego, $u = (0,2)$.

Suma e intersección de subespacios

Definición 3.38. Sea $\mathbb{V}$ un espacio vectorial. Sean S y T subespacios de $\mathbb{V}$. Definimos la *suma de* S y T por $S+T=\{s+t \ /s\in S,\ t\in T\}$.

Proposición 3.39. *Sea $\mathbb{V}$ un espacio vectorial y sean S y T subespacios de $\mathbb{V}$. Entonces $S+T$ es un subespacio de $\mathbb{V}$. Además, $S+T\supseteq S$, $S+T\supseteq T$ y $S+T$ es el menor subespacio de $\mathbb{V}$ que contiene a S y a T.*

Demostración. Veamos que $S+T$ es un subespacio de $\mathbb{V}$.

(i) $0\in S+T$ porque $0\in S$ y $0\in T$ y $0=0+0$.

(ii) Sean $v,w\in S+T$. Debemos probar que $v+w\in S+T$.

Como $v,w\in S+T$, entonces existen $s_1,s_2\in S$ y $t_1,t_2\in T$, tales que $v=s_1+t_1$ y $w=s_2+t_2$. Luego, $v+w=s_1+t_1+s_2+t_2=(s_1+s_2)+(t_1+t_2)$.

Ahora bien, como $s_1,s_2\in S$ y S es un subespacio de $\mathbb{V}$, entonces $s_1+s_2\in S$ y como $t_1,t_2\in T$ y T es un subespacio de $\mathbb{V}$, entonces $t_1+t_2\in T$. Por lo tanto, tenemos que

$$v+w=(s_1+s_2)+(t_1+t_2)\in S+T.$$

(iii) Sean $v\in S+T$ y $\alpha\in\mathbb{R}$. Debemos probar que $\alpha.v\in S+T$.

Como $v\in S+T$, entonces existen $s_1\in S,\ t_1\in T$ tales que $v=s_1+t_1$. Luego, se tiene que $\alpha.v=\alpha.(s_1+t_1)=\alpha.s_1+\alpha.t_1$

Ahora bien, como $s_1\in S$ y $\alpha\in\mathbb{R}$ y S es un subespacio de $\mathbb{V}$, entonces $\alpha.s_1\in S$ y como $t_1\in T$ y T es un subespacio de $\mathbb{V}$, entonces $\alpha.t_1\in T$. Por lo tanto, tenemos que

$$\alpha.v=\alpha.s_1+\alpha.t_1\in S+T.$$

Luego $S+T$ es un subespacio de $\mathbb{V}$.

Veamos ahora que $S\subseteq S+T$. Sea $s\in S$. Como T es un subespacio, entonces $0\in T$. Luego, $s=s+0\in S+T$.

Análogamente, $T\subseteq S+T$.

Por último, veamos que $S+T$ es el menor subespacio de $\mathbb{V}$ que contiene a S y a T. Debemos probar que si W es un subespacio de $\mathbb{V}$ que contiene a S y a T, entonces $S+T\subseteq W$. Sea entonces W un subespacio de $\mathbb{V}$ que contiene a S y a T. Sea $v\in S+T$. Entonces existen $s_1\in S$ y $t_1\in T$, tales que $v=s_1+t_1$. Ahora bien, como $s_1\in S$ y $S\subseteq W$, entonces $s_1\in W$. Además, como $t_1\in T$ y $T\subseteq W$, entonces $t_1\in W$. Luego, $v=s_1+t_1\in W$, pues W es un subespacio de $\mathbb{V}$.

Por lo tanto, $S+T\subseteq W$. □

Proposición 3.40. *Sea $\mathbb{V}$ un espacio vectorial y sean S y T subespacios de $\mathbb{V}$. Sea G_S un conjunto de generadores de S y sea G_T un conjunto de generadores de T. Sea $G=G_S\cup G_T$. Entonces G genera $S+T$.*

La demostración de esta proposición queda como ejercicio para el lector.

Ejemplo 3.41. Si $S=\text{Gen}(\{(1,3,1)\})$ y $T=\text{Gen}(\{(1,2,0),(0,1,1)\})$. Entonces

$$S+T=\text{Gen}(\{(1,3,1),(1,2,0),(0,1,1)\})$$

Proposición 3.42. *Sea $\mathbb{V}$ un espacio vectorial y sean S y T subespacios de $\mathbb{V}$. Entonces $S\cap T$ es un subespacio de $\mathbb{V}$.*

La demostración de esta proposición queda como ejercicio para el lector.

El siguiente resultado se conoce con el nombre de teorema de las dimensiones para espacios vectoriales.

Teorema 3.43. *Sea $\mathbb{V}$ un espacio vectorial de dimensión finita y sean S y T subespacios de $\mathbb{V}$. Entonces*

$$\dim(S+T) = \dim(S) + \dim(T) - \dim(S \cap T).$$

Demostración. Sean $n = \dim(S)$, $m = \dim(T)$ y $r = \dim(S \cap T)$. Sea $B_1 = \{v_1, v_2, \ldots, v_r\}$ una base de $S \cap T$. Como B_1 es base de $S \cap T$ y $S \cap T \subseteq S$, entonces B_1 es un conjunto linealmente independiente de vectores de S. Por lo tanto, B_1 se puede extender a una base de S, es decir, existen $w_1, w_2, \ldots, w_{n-r} \in S$ tales que $\{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{n-r}\}$ es base de S. Sea $B_2 = \{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{n-r}\}$.

Similarmente, como B_1 es base de $S \cap T$ y $S \cap T \subseteq T$, entonces B_1 es un conjunto linealmente independiente de vectores de T. Por lo tanto, B_1 se puede extender a una base de T, es decir, existen $w'_1, w'_2, \ldots, w'_{m-r} \in T$ tales que $\{v_1, v_2, \ldots, v_r, w'_1, w'_2, \ldots, w'_{m-r}\}$ es base de T. Sea $B_3 = \{v_1, v_2, \ldots, v_r, w'_1, w'_2, \ldots, w'_{m-r}\}$.

Sea $B = B_2 \cup B_3 = \{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{n-r}, w'_1, w'_2, \ldots, w'_{m-r}\}$. Veamos que B es base de $S+T$.

Como B_2 genera S y B_3 genera T entonces, por 3.40, $B = B_2 \cup B_3$ genera $S+T$.

Veamos ahora que B es linealmente independiente. Sean $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_{n-r}$, $\gamma_1, \gamma_2, \ldots, \gamma_{m-r} \in \mathbb{R}$ tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_r v_r + \beta_1 w_1 + \beta_2 w_2 + \ldots + \beta_{n-r} w_{n-r} + \gamma_1 w'_1 + \gamma_2 w'_2 + \ldots + \gamma_{m-r} w'_{m-r} = 0.$$

Entonces,

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_r v_r + \beta_1 w_1 + \beta_2 w_2 + \ldots + \beta_{n-r} w_{n-r} = -\gamma_1 w'_1 - \gamma_2 w'_2 - \ldots - \gamma_{m-r} w'_{m-r} \quad (*)$$

Claramente el lado izquierdo de esta igualdad es un elemento de S por ser una combinación lineal de elementos de S y el lado derecho es un elemento de T por ser una combinación lineal de elementos de T.

Luego $-\gamma_1 w'_1 - \gamma_2 w'_2 - \ldots - \gamma_{m-r} w'_{m-r} \in S \cap T$. Entonces existen $\delta_1, \delta_2, \ldots, \delta_r \in \mathbb{R}$ tales que

$$-\gamma_1 w'_1 - \gamma_2 w'_2 - \ldots - \gamma_{m-r} w'_{m-r} = \delta_1 v_1 + \delta_2 v_2 + \ldots + \delta_r v_r$$

Luego,

$$0 = \delta_1 v_1 + \delta_2 v_2 + \ldots + \delta_r v_r + \gamma_1 w'_1 + \gamma_2 w'_2 + \ldots + \gamma_{m-r} w'_{m-r} .$$

y como $B_3 = \{v_1, v_2, \ldots, v_r, w'_1, w'_2, \ldots, w'_{m-r}\}$ es base de T, entonces B_3 es un conjunto linealmente independiente y luego de la igualdad anterior concluimos que

$$\delta_1 = 0, \delta_2 = 0, \ldots, \delta_r = 0, \gamma_1 = 0, \gamma_2 = 0, \ldots, \gamma_{m-r} = 0 .$$

Entonces, reemplazando en $(*)$, queda

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_r v_r + \beta_1 w_1 + \beta_2 w_2 + \ldots + \beta_{n-r} w_{n-r} = 0$$

y como $B_2 = \{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{m-r}\}$ es base de S, entonces es un conjunto linealmente independiente y luego de la igualdad anterior concluimos que

$$\alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_r = 0, \beta_1 = 0, \beta_2 = 0, \ldots, \beta_{n-r} = 0 .$$

Entonces B es linealmente independiente.

Por lo tanto, B es base de $S+T$.

Entonces obtenemos que

$$\dim(S+T) = \#B = r + (n-r) + (m-r) = n + m - r = \dim(S) + \dim(T) - \dim(S \cap T).$$

□

Definición 3.44. Sea $\mathbb{V}$ un espacio vectorial y sean S y T subespacios de $\mathbb{V}$. Sea $W = S + T$. Si $S \cap T = \{0\}$, decimos que S y T están en *suma directa* o que W *es la suma directa de* S *y* T. Lo notaremos, $W = S \oplus T$.

Observación 3.45. Si $W = S \oplus T$, entonces $\dim(W) = \dim(S) + \dim(T)$.

Ejemplos 3.46.
(1) Sean S y T dos subespacios distintos de $\mathbb{R}^2$ de dimensión 1, es decir, S y T son dos rectas no coincidentes que pasan por el origen.

Entonces, $S + T = \mathbb{R}^2$, $S \cap T = \{0\}$ y

$$\dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T) = 1 + 1 - 0 = 2.$$

(2) Sean S y T subespacios de $\mathbb{R}^3$ tales que $\dim(S) = 2$, $\dim(T) = 1$ y $\dim(S \cap T) = 0$, es decir, S es un plano que pasa por el origen y T una recta que pasa por el origen no contenida en el plano.

Entonces, $S \cap T = \{0\}$ y

$$\dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T) = 2 + 1 - 0 = 3.$$

Luego, como $S + T \subseteq \mathbb{R}^3$ y $\dim(S + T) = \dim(\mathbb{R}^3)$ concluimos que $S + T = \mathbb{R}^3$.

(3) Sean S y T subespacios de $\mathbb{R}^3$ de dimensión 2, es decir S y T son dos planos que pasan por el origen, entonces se tiene que:

$$3 \geqslant \dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T) = 2 + 2 - \dim(S \cap T).$$

de donde se obtiene que $\dim(S \cap T) \geqslant 2 + 2 - 3 = 1$ y por lo tanto podemos deducir que la intersección entre dos planos en $\mathbb{R}^3$ es al menos una recta.

Además, $S \cap T \subseteq S$, entonces $\dim(S \cap T) \leqslant \dim(S) = 2$, de donde podemos deducir que, la intersección entre dos planos en $\mathbb{R}^3$ es como mucho un plano.

Capítulo 4

Transformaciones lineales

Definición 4.1. Sean $\mathbb{V}, \mathbb{W}$ espacios vectoriales. Una *transformación lineal de* $\mathbb{V}$ *en* $\mathbb{W}$ es una función $f : \mathbb{V} \to \mathbb{W}$ tal que

- $\forall v, w \in \mathbb{V}, f(v + w) = f(v) + f(w)$.
- $\forall v \in \mathbb{V},\ \forall \alpha \in \mathbb{R}, f(\alpha.v) = \alpha.f(v)$.

Observación 4.2. Notar que si $f : \mathbb{V} \to \mathbb{W}$ es una transformación lineal entonces $f(0) = f(0.0) = 0.f(0) = 0$.

Ejemplos 4.3.
(1) Sea $f : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $f(x, y) = (x - y, 2x + y, 3y)$. Determinar si f es una transformación lineal.

Solución: Demostremos primero que $\forall v, w \in \mathbb{R}^2, f(v+w) = f(v)+f(w)$. Sean $v = (v_1, v_2)$ y $w = (w_1, w_2)$ vectores de $\mathbb{R}^2$. Se tiene que

$$\begin{aligned}
f(v + w) &= f(v_1 + w_1, v_2 + w_2) = \\
&= ((v_1 + w_1) - (v_2 + w_2), 2(v_1 + w_1) + (v_2 + w_2), 3(v_2 + w_2)) = \\
&= (v_1 + w_1 - v_2 - w_2, 2v_1 + 2w_1 + v_2 + w_2, 3v_2 + 3w_2) = \\
&= (v_1 - v_2, 2v_1 + v_2, 3v_2) + (w_1 - w_2, 2w_1 + w_2, 3w_2) = \\
&= f(v_1, v_2) + f(w_1, w_2) = \\
&= f(v) + f(w).
\end{aligned}$$

Demostremos ahora que $\forall v \in \mathbb{R}^2,\ \forall \alpha \in \mathbb{R}, f(\alpha.v) = \alpha.f(v)$. Sean $v = (v_1, v_2) \in \mathbb{R}^2$ y $\alpha \in \mathbb{R}$. Se tiene que

$$\begin{aligned}
f(\alpha.v) &= f(\alpha.v_1, \alpha.v_2) = \\
&= (\alpha.v_1 - \alpha.v_2, 2.\alpha.v_1 + \alpha.v_2, 3.\alpha.v_2) = \\
&= (\alpha.(v_1 - v_2), \alpha.(2v_1 + v_2), \alpha.(3v_2)) = \\
&= \alpha.(v_1 - v_2, 2v_1 + v_2, 3v_2) = \\
&= \alpha.f(v_1, v_2) = \\
&= \alpha.f(v).
\end{aligned}$$

Por lo tanto, f es una transformación lineal.

(2) Sea $f : \mathbb{R}^3 \to \mathbb{R}^{2\times 2}$ definida por $f(x, y, z) = \begin{bmatrix} x & y \\ z & 1 \end{bmatrix}$. Determinar si f es una transformación lineal.

Solución: Si $v = (1, 0, 0)$ y $w = (0, 0, 1)$ entonces $f(v + w) = f(1, 0, 1) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ y

$$f(v) + f(w) = f(1, 0, 0) + f(0, 0, 1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = f(v + w).$$

Por lo tanto, f no es transformación lineal.

(3) Sea $A \in \mathbb{R}^{m\times n}$. Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ definida por $f(x) = (A \,.\, x^t)^t$. Entonces f es una transformación lineal. La demostración de este hecho queda como ejercicio para el lector.

Proposición 4.4. *Sean $\mathbb{V}$, $\mathbb{V}'$ y $\mathbb{V}''$ espacios vectoriales. Sean $f : \mathbb{V} \to \mathbb{V}'$ y $g : \mathbb{V}' \to \mathbb{V}''$ transformaciones lineales. Entonces $g \circ f$ es una transformación lineal.*

Demostración. Veamos primero que $\forall v, w \in \mathbb{V}, (g \circ f)(v + w) = (g \circ f)(v) + (g \circ f)(w)$.

Sean $v, w \in \mathbb{V}$. Como f y g son transformaciones lineales obtenemos que

$$(g\circ f)(v+w) = g(f(v+w)) = g(f(v)+f(w)) = g(f(v))+g(f(w)) = (g\circ f)(v)+(g\circ f)(w) \;.$$

Veamos ahora que $\forall v \in \mathbb{V},\ \forall \alpha \in \mathbb{R}, (g \circ f)(\alpha.v) = \alpha.(g \circ f)(v)$. Sean $v \in \mathbb{V}$ y $\alpha \in \mathbb{R}$. Entonces

$$(g \circ f)(\alpha.v) = g(f(\alpha.v)) = g(\alpha.f(v)) = \alpha.g(f(v)) = \alpha.(g \circ f)(v).$$

Por lo tanto, $g \circ f$ es una transformación lineal. □

Proposición 4.5. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales. Sea $\{v_1, v_2, \ldots, v_n\}$ una base de $\mathbb{V}$ y sean $w_1, w_2, \ldots, w_n \in \mathbb{W}$. Entonces existe una única transformación lineal $f : \mathbb{V} \to \mathbb{W}$ tal que $f(v_i) = w_i$ para todo $i \in \{1, 2, \ldots, n\}$.*

Demostración. Definimos $f : \mathbb{V} \to \mathbb{W}$ de la siguiente manera. Dado $v \in \mathbb{V}$, sabemos que existen únicos $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tales que $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$. Definimos $f(v) = \alpha_1 w_1 + \alpha_2 w_2 + \ldots + \alpha_n w_n$.

Veamos que f es una transformación lineal. En primer lugar demostremos que para todos $v, v' \in V$ vale que $f(v + v') = f(v) + f(v')$.

Sean $v, v' \in \mathbb{V}$. Entonces existen únicos $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha'_1, \alpha'_2, \ldots, \alpha'_n \in \mathbb{R}$ tales que $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$ y $v' = \alpha'_1 v_1 + \alpha'_2 v_2 + \ldots + \alpha'_n v_n$. Luego

$$\begin{aligned} v + v' &= (\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) + (\alpha'_1 v_1 + \alpha'_2 v_2 + \ldots + \alpha'_n v_n) = \\ &= (\alpha_1 + \alpha'_1)v_1 + (\alpha_2 + \alpha'_2)v_2 + \ldots + (\alpha_n + \alpha'_n)v_n \end{aligned}$$

y por lo tanto,

$$\begin{aligned} f(v + v') &= (\alpha_1 + \alpha'_1)w_1 + (\alpha_2 + \alpha'_2)w_2 + \ldots + (\alpha_n + \alpha'_n)w_n = \\ &= (\alpha_1 w_1 + \alpha_2 w_2 + \ldots + \alpha_n w_n) + (\alpha'_1 w_1 + \alpha'_2 w_2 + \ldots + \alpha'_n w_n) = \\ &= f(v) + f(v') \;. \end{aligned}$$

Similarmente se demuestra que si $k \in \mathbb{R}$ y $v \in \mathbb{V}$ entonces $f(kv) = kf(v)$. La demostración de este hecho queda como ejercicio para el lector.

Por lo tanto, f es una transformación lineal.

Veamos ahora que $f(v_i) = w_i$ para todo $i \in \{1, 2, \ldots, n\}$. Sea $i \in \{1, 2, \ldots, n\}$. Entonces

$$v_i = 0v_1 + 0v_2 + \ldots + 0v_{i-1} + 1v_i + 0v_{i+1} + \ldots + 0v_n$$

y luego

$$f(v_i) = 0w_1 + 0w_2 + \ldots + 0w_{i-1} + 1w_i + 0w_{i+1} + \ldots + 0w_n = w_i$$

como queríamos demostrar.

Resta probar que una tal f es única. Supongamos que $g : \mathbb{V} \to \mathbb{W}$ es una transformación lineal tal que $g(v_i) = w_i$ para todo $i \in \{1, 2, \ldots, n\}$. Debemos demostrar que $g = f$, es decir que, $g(v) = f(v)$ para todo $v \in \mathbb{V}$.

Sea $v \in \mathbb{V}$. Entonces existen únicos $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tales que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \ .$$

Luego

$$\begin{aligned} g(v) &= g(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = \\ &= g(\alpha_1 v_1) + g(\alpha_2 v_2) + \ldots + g(\alpha_n v_n) = && \text{(pues } g \text{ es transformación lineal)} \\ &= \alpha_1 g(v_1) + \alpha_2 g(v_2) + \ldots + \alpha_n g(v_n) = && \text{(pues } g \text{ es transformación lineal)} \\ &= \alpha_1 w_1 + \alpha_2 w_2 + \ldots + \alpha_n w_n = && \text{(pues } g(v_i) = w_i \text{ para todo } i\text{)} \\ &= f(v) && \text{(por definición de } f\text{)} \end{aligned}$$

Por lo tanto, $g = f$. □

Definición 4.6. Sean $\mathbb{V}, \mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal.

- Definimos el *núcleo* de f por $\mathrm{Nu}(f) = \{v \in \mathbb{V} \,/\, f(v) = 0\}$.
- Definimos la *imagen* de f por $\mathrm{Im}(f) = \{w \in \mathbb{W} \,/\, \exists v \in \mathbb{V} \text{ tal que } f(v) = w\}$.

Proposición 4.7. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Entonces $\mathrm{Nu}f$ es un subespacio de $\mathbb{V}$ y $\mathrm{Im}f$ es un subespacio de $\mathbb{W}$.*

Demostración. Veamos primero que $\mathrm{Nu}(f)$ es un subespacio de $\mathbb{V}$.

(i) Como $f(0) = 0$ por ser f una transformación lineal, entonces $0 \in \mathrm{Nu}(f)$.

(ii) Sean $v, w \in \mathrm{Nu}(f)$. Entonces

$$f(v + w) = f(v) + f(w) = 0 + 0 = 0$$

donde la primera igualdad vale porque f es una transformación lineal y la segunda vale porque $v, w \in \mathrm{Nu}(f)$. Luego, $v + w \in \mathrm{Nu}(f)$.

(iii) Sean $v \in \mathbb{V}$ y $\alpha \in \mathbb{R}$. Entonces

$$f(\alpha.v) = \alpha.f(v) = \alpha.0 = 0$$

donde la primera igualdad vale porque f es una transformación lineal y la segunda vale porque $v \in \mathrm{Nu}(f)$. Luego, $\alpha.v \in \mathrm{Nu}(f)$.

Por lo tanto, $\mathrm{Nu}(f)$ es un subespacio de $\mathbb{V}$.

Veamos ahora que $\mathrm{Im}(f)$ es un subespacio de $\mathbb{W}$.

(i) Como $0 = f(0)$ por ser f una transformación lineal, entonces $0 \in \mathrm{Im}(f)$.

(ii) Sean ahora $w_1, w_2 \in \mathrm{Im}(f)$. Entonces existen $v_1, v_2 \in \mathbb{V}$ tales que $f(v_1) = w_1$ y $f(v_2) = w_2$. Entonces, como f es transformación lineal, tenemos que

$$f(v_1 + v_2) = f(v_1) + f(v_2) = w_1 + w_2 \ .$$

Luego, $w_1 + w_2 \in \mathrm{Im}(f)$.

(iii) Sean $w \in Im(f)$ y $\alpha \in \mathbb{R}$. Entonces, existe $v \in \mathbb{V}$ tal que $f(v) = w$. Luego,

$$f(\alpha.v) = \alpha.f(v) = \alpha.w \ .$$

Entonces, $\alpha.w \in \mathrm{Im}(f)$.

Por lo tanto, $\mathrm{Im}(f)$ es un subespacio de $\mathbb{W}$. □

Proposición 4.8. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea $G = \{v_1, v_2, \ldots, v_n\}$ un conjunto generador de $\mathbb{V}$. Entonces el conjunto $f(G) = \{f(v_1), f(v_2), \ldots, f(v_n)\}$ genera* $\mathrm{Im} f$.

Demostración. Sea $w \in \mathrm{Im}(f)$. Entonces, existe $v \in \mathbb{V}$ tal que $f(v) = w$. Como G es un conjunto generador de $\mathbb{V}$ y $v \in \mathbb{V}$, entonces existen $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tales que $v = \alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n$. Luego,

$$w = f(v) = f(\alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n) = \alpha_1.f(v_1) + \alpha_2.f(v_2) + \ldots + \alpha_n.f(v_n).$$

Así, w es combinación lineal de $f(v_1), f(v_2), \ldots, f(v_n)$.

Por lo tanto, $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ genera $\mathrm{Im} f$. □

Ejemplo 4.9. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $f(x, y, z) = (0, x - y, z)$. Hallar bases de $\mathrm{Nu}(f)$ e $\mathrm{Im}(f)$.

Solución: $\mathrm{Nu}(f) = \{(x, y, z) \in \mathbb{R}^3 \,/\, f(x, y, z) = (0, 0, 0)\}$.

Planteamos

$$f(x, y, z) = (0, 0, 0)$$

$$(0, x - y, z) = (0, 0, 0)$$

y obtenemos

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

El conjunto solución de este sistema es $\{(y, y, 0) \,/\, y \in \mathbb{R}\}$ y como $(y, y, 0) = y.(1, 1, 0)$ para todo $y \in \mathbb{R}$ se obtiene que $\mathrm{Nu}(f) = \mathrm{Gen}(\{(1, 1, 0)\})$.

Ahora bien, como $\{(1, 1, 0)\}$ es un conjunto linealmente independiente (por estar formado por un único vector no nulo), entonces $\{(1, 1, 0)\}$ es base de $\mathrm{Nu}(f)$.

Ahora busquemos una base para $\mathrm{Im}(f)$. Como $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ genera $\mathbb{R}^3$ entonces, por la proposición 4.8, tenemos que $\{f(1, 0, 0), f(0, 1, 0), f(0, 0, 1)\}$ genera $\mathrm{Im}(f)$. Ahora bien,

$$\begin{array}{rcl} f(1,0,0) & = & (0,1,0) \\ f(0,1,0) & = & (0,-1,0) \\ f(0,0,1) & = & (0,0,1) \end{array}$$

Luego, $\mathrm{Im}(f) = \mathrm{Gen}(\{(0, 1, 0), (0, -1, 0), (0, 0, 1)\})$.

Ahora extraemos una base de este conjunto de generadores. Como

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

entonces

$$\mathrm{Gen}(\{(0, 1, 0), (0, -1, 0), (0, 0, 1)\}) = \mathrm{Gen}(\{(0, 1, 0), (0, 0, 1)\}) \ .$$

y por lo tanto el conjunto $\{(0, 1, 0), (0, 0, 1)\}$ genera $\mathrm{Im}(f)$. Además, es claro que este conjunto es linealmente independiente.

Por lo tanto, $\{(0, 1, 0), (0, 0, 1)\}$ es base de $\mathrm{Im}(f)$.

Definición 4.10. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal.

- Decimos que f es un *monomorfismo* si $\mathrm{Nu}(f) = \{0\}$.
- Decimos que f es un *epimorfismo* si $\mathrm{Im}(f) = \mathbb{W}$.

- Decimos que f es un *isomorfismo* si f es monomorfismo y epimorfismo.

Proposición 4.11. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Entonces:*

(1) f es un monomorfismo si y sólo si f es inyectiva.

(2) f es un epimorfismo si y sólo si f es sobreyectiva.

(3) f es un isomorfismo si y sólo si f es biyectiva.

Demostración.
(1) ($\Rightarrow$) Supongamos que f es un monomorfismo. Sean $v_1, v_2 \in \mathbb{V}$ tales que $f(v_1) = f(v_2)$. Debemos probar que $v_1 = v_2$. Tenemos que

$$f(v_1 - v_2) = f(v_1 + (-1).v_2) = f(v_1) + (-1).f(v_2) = f(v_1) - f(v_2) = 0 \ .$$

Luego, $v_1 - v_2 \in \mathrm{Nu}(f)$ y como $\mathrm{Nu}(f) = \{0\}$ (por ser f un monomorfismo), entonces $v_1 - v_2 = 0$. Luego $v_1 = v_2$.

Por lo tanto, f es inyectiva.

($\Leftarrow$) Supongamos que f es inyectiva. Debemos probar que f es monomorfismo, es decir, debemos probar que $\mathrm{Nu}(f) = \{0\}$. Veamos la doble inclusión.

($\supseteq$) Por ser $\mathrm{Nu}(f)$ un subespacio, $0 \in \mathrm{Nu}(f)$ entonces $\{0\} \subseteq \mathrm{Nu}(f)$.

($\subseteq$) Sea $v \in \mathrm{Nu}(f)$. Entonces, $f(v) = 0 = f(0)$. Luego, $f(v) = f(0)$ y como f es inyectiva, entonces $v = 0$. Luego $v \in \{0\}$.

Por lo tanto, $\mathrm{Nu}(f) = \{0\}$ y entonces f es monomorfismo.

(2) Trivial.

(3) f es isomorfismo $\Leftrightarrow$ f es monomorfismo y epimorfismo $\Leftrightarrow$ f es inyectiva y sobreyectiva $\Leftrightarrow$ f es biyectiva. □

Proposición 4.12. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Si f es un isomorfismo entonces $f^{-1} : \mathbb{W} \to \mathbb{V}$ es una transformación lineal.*

Demostración. Demostremos primero que $f^{-1}(w_1 + w_2) = f^{-1}(w_1) + f^{-1}(w_2)$ para todos $w_1, w_2 \in \mathbb{W}$. Sean $w_1, w_2 \in \mathbb{W}$. Como f es un isomorfismo, existen únicos $v_1, v_2 \in \mathbb{V}$ tales que $f(v_1) = w_1$ y $f(v_2) = w_2$. Luego, $f^{-1}(w_1) = v_1$ y $f^{-1}(w_2) = v_2$.

Por otra parte, $f(v_1 + v_2) = f(v_1) + f(v_2) = w_1 + w_2$, entonces $f^{-1}(w_1 + w_2) = v_1 + v_2$. Luego,

$$f^{-1}(w_1 + w_2) = v_1 + v_2 = f^{-1}(w_1) + f^{-1}(w_2) \ .$$

Veamos ahora que $f^{-1}(\alpha.w) = \alpha.f^{-1}(w)$ para todos $w \in \mathbb{W}$ y $\alpha \in \mathbb{R}$. Sean $w \in \mathbb{W}$ y $\alpha \in \mathbb{R}$. Como f es un isomorfismo, existe un único $v \in \mathbb{V}$ tal que $f(v) = w$. Entonces $f^{-1}(w) = v$.

Por otra parte, $f(\alpha.v) = \alpha.f(v) = \alpha.w$, entonces $f^{-1}(\alpha.w) = \alpha.v$. Luego,

$$f^{-1}(\alpha.w) = \alpha.v = \alpha.f^{-1}(w) \ .$$

Por lo tanto, f^{-1} es una transformación lineal. □

Proposición 4.13. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea $\{v_1, v_2, \ldots, v_n\}$ un conjunto linealmente independiente de vectores de $\mathbb{V}$. Si f es un monomorfismo entonces $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ es linealmente independiente.*

Demostración. Sean $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tales que $\alpha_1.f(v_1)+\alpha_2.f(v_2)+\ldots+\alpha_n.f(v_n)=0$. Como f es transformación lineal tenemos que $f(\alpha_1.v_1+\alpha_2.v_2+\ldots+\alpha_n.v_n)=0$. Y como f es monomorfismo, obtenemos que $\alpha_1.v_1+\alpha_2.v_2+\ldots+\alpha_n.v_n=0$. Ahora, como $\{v_1, v_2, \ldots, v_n\}$ es un conjunto linealmente independiente, se obtiene que $\alpha_1=0, \alpha_2=0, \ldots, \alpha_n=0$.

Por lo tanto, el conjunto $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ es linealmente independiente. □

Proposición 4.14. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea $\{v_1, v_2, \ldots, v_n\}$ una base de $\mathbb{V}$. Si f es un isomorfismo entonces $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ es una base de $\mathbb{W}$.*

Demostración. Como f es un isomorfismo, entonces f es un monomorfismo y como el conjunto $\{v_1, v_2, \ldots, v_n\}$ es base de $\mathbb{V}$, entonces es linealmente independiente. Luego, por la proposición anterior, tenemos que $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ es linealmente independiente.

Veamos ahora que $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ genera $\mathbb{W}$. Como $\{v_1, v_2, \ldots, v_n\}$ genera $\mathbb{V}$ por ser base de $\mathbb{V}$, entonces por la proposición 4.8 obtenemos que $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ genera $\mathrm{Im}(f)$. Ahora bien, como f es isomorfismo, entonces f es un epimorfismo y luego $\mathrm{Im}(f)=\mathbb{W}$. Entonces $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ genera $\mathbb{W}$.

Por lo tanto, como el conjunto $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ es linealmente independiente y genera $\mathbb{W}$, concluimos que $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ es base de $\mathbb{W}$. □

El siguiente teorema se conoce con el nombre de teorema de las dimensiones para transformaciones lineales.

Teorema 4.15. *Sea $\mathbb{V}$ un espacio vectorial de dimensión finita y sea $\mathbb{W}$ otro espacio vectorial. Sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Entonces*

$$\dim(\mathbb{V}) = \dim(\mathrm{Nu} f) + \dim(\mathrm{Im} f) \ .$$

Demostración. Sean $r=\dim(\mathrm{Nu}(f))$ y $n=\dim(\mathbb{V})$. Sea $B_1=\{v_1, v_2, \ldots, v_r\}$ una base de $\mathrm{Nu}(f)$. Como $\mathrm{Nu}(f) \subseteq \mathbb{V}$ y B_1 es base de $\mathrm{Nu}(f)$ entonces B_1 es un conjunto linealmente independiente de vectores de $\mathbb{V}$. Entonces, por 3.30, B_1 se puede extender a una base de $\mathbb{V}$, es decir, existen $v_{r+1}, \ldots, v_n \in \mathbb{V}$ tales que $B=\{v_1, v_2, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ es base de $\mathbb{V}$.

Como B genera $\mathbb{V}$, tenemos que $f(B)=\{f(v_1), f(v_2), \ldots, f(v_r), f(v_{r+1}), \ldots, f(v_n)\}$ genera $\mathrm{Im}(f)$ por la proposición 4.8. Ahora bien, como $v_1, v_2, \ldots, v_r \in \mathrm{Nu}(f)$ entonces $f(v_1)=0, f(v_2)=0, \ldots, f(v_r)=0$ y luego

$$\mathrm{Im}(f) = \mathrm{Gen}(\{f(v_1), f(v_2), \ldots, f(v_r), f(v_{r+1}), \ldots, f(v_n)\}) = \mathrm{Gen}(\{f(v_{r+1}), \ldots, f(v_n)\}).$$

Veamos ahora que el conjunto $\{f(v_{r+1}), \ldots, f(v_n)\}$ es linealmente independiente. Sean $\alpha_{r+1}, \ldots, \alpha_n \in \mathbb{R}$ tales que

$$\alpha_{r+1} . f(v_{r+1}) + \ldots + \alpha_n . f(v_n) = 0 \ .$$

Como f es una transformación lineal, entonces $f(\alpha_{r+1} . v_{r+1} + \ldots + \alpha_n . v_n) = 0$. Luego, $\alpha_{r+1} . v_{r+1} + \ldots + \alpha_n . v_n \in \mathrm{Nu}(f)$. Y como B_1 es base de $\mathrm{Nu}(f)$ obtenemos que existen $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}$ tales que

$$\alpha_{r+1} . v_{r+1} + \ldots + \alpha_n . v_n = \alpha_1 . v_1 + \ldots + \alpha_r . v_r \ .$$

Entonces,

$$-\alpha_1 . v_1 - \ldots - \alpha_r . v_r + \alpha_{r+1} . v_{r+1} + \ldots + \alpha_n . v_n = 0$$

y como $B = \{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ es base de $\mathbb{V}$, entonces B es linealmente independiente y luego de la igualdad de arriba obtenemos que

$$\alpha_1 = 0, \ldots, \alpha_r = 0, \alpha_{r+1} = 0, \ldots, \alpha_n = 0$$

Entonces, el conjunto $\{f(v_{r+1}), \ldots, f(v_n)\}$ es linealmente independiente.

Por lo tanto, $\{f(v_{r+1}), \ldots, f(v_n)\}$ es base de $\mathrm{Im}(f)$.

Luego, $\dim(\mathrm{Im}(f)) = n - r$ y obtenemos que

$$\dim(\mathrm{Nu}(f)) + \dim(\mathrm{Im}(f)) = r + (n - r) = n = \dim(\mathbb{V}) \ .$$

□

Matriz de una transformación lineal

Definición 4.16. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sean $n = \dim(\mathbb{V})$ y $m = \dim(\mathbb{W})$. Sea $B = \{v_1, v_2, \ldots, v_n\}$ una base de $\mathbb{V}$ y sea B' una base de $\mathbb{W}$. La *matriz de la transformación lineal f en las bases B y B'* es la matriz $M_{BB'}(f) \in \mathbb{R}^{m\times n}$ cuyas columnas son los vectores $[f(v_1)]_B, [f(v_2)]_B, \ldots, [f(v_n)]_B$.

Ejemplo 4.17. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal definida por

$$f(x, y, z) = (x - y, y - z).$$

Sea $B = \{(1, 1, 1), (0, 1, 0), (0, 0, 2)\}$ base de $\mathbb{R}^3$ y sea $B' = \{(3, 0), (0, 1)\}$ base de $\mathbb{R}^2$. Calcular $M_{BB'}(f)$.

Solución: Tenemos que

$$\begin{array}{rcrcrcr} f(1,1,1) & = & (0,0) & = & 0.(3,0) & + & 0\ .(0,1) \\ f(0,1,0) & = & (-1,1) & = & -\frac{1}{3}.(3,0) & + & 1\ .(0,1) \\ f(0,0,2) & = & (0,-2) & = & 0.(3,0) & + & (-2).(0,1) \end{array}$$

Luego,

$$\begin{array}{rcl} [f(1,1,1)]_{B'} & = & (0,0) \\ [f(0,1,0)]_{B'} & = & (-\frac{1}{3},1) \\ [f(0,0,2)]_{B'} & = & (0,-2) \end{array}$$

Entonces,

$$M_{BB'}(f) = \begin{pmatrix} 0 & -\frac{1}{3} & 0 \\ 0 & 1 & -2 \end{pmatrix} .$$

La siguiente proposición nos muestra para qué sirve la matriz de una transformación lineal.

Proposición 4.18. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea B una base de $\mathbb{V}$ y sea B' una base de $\mathbb{W}$. Entonces, para todo $v \in \mathbb{V}$ vale que*

$$M_{BB'}(f).([v]_B)^t = ([f(v)]_{B'})^t.$$

Demostración. Sean $B = \{v_1, v_2, \ldots, v_n\}$ y $B' = \{w_1, w_2, \ldots, w_m\}$. Sea $v \in \mathbb{V}$. Entonces existen $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tales que $v = \alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n$. Luego,

$$f(v) = f(\alpha_1.v_1 + \alpha_2.v_2 + \ldots + \alpha_n.v_n) = \alpha_1.f(v_1) + \alpha_2.f(v_2) + \ldots + \alpha_n.f(v_n) \qquad (*)$$

Sea $A = M_{BB'}(f)$. Luego, $[f(v_i)]_{B'} = (A_{1i}, A_{2i}, \dots, A_{mi})$ para todo $i \in \{1, 2, \dots, n\}$. Entonces, $f(v_i) = A_{1i}w_1 + A_{2i}w_2 + \dots + A_{mi}w_m$, $\forall i \in \{1, 2, \dots, n\}$. Luego, de $(*)$ obtenemos que

$$\begin{aligned} f(v) &= \alpha_1(A_{11}w_1 + A_{21}w_2 + \dots + A_{m1}w_m) + \alpha_2(A_{12}w_1 + A_{22}w_2 + \dots + A_{m2}w_m) + \\ &\quad + \dots + \alpha_n(A_{1n}w_1 + A_{2n}w_2 + \dots + A_{mn}w_m) = \\ &= (\alpha_1 A_{11} + \alpha_2 A_{12} + \dots + \alpha_n A_{1n})w_1 + (\alpha_1 A_{21} + \alpha_2 A_{22} + \dots + \alpha_n A_{2n})w_2 + \\ &\quad + \dots + (\alpha_1 A_{m1} + \alpha_2 A_{m2} + \dots + \alpha_n A_{mn})w_m \ . \end{aligned}$$

Para cada $j \in \{1, 2, \dots, m\}$ sea $b_j = \alpha_1 A_{j1} + \alpha_2 A_{j2} + \dots + \alpha_n A_{jn}$. Entonces,

$$[f(v)]_{B'} = (b_1, b_2, \dots, b_m).$$

Luego,

$$([f(v)]_{B'})^t = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = A. \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = M_{BB'}(f).([v]_B)^t \ .$$

□

Ejemplo 4.19. Sea f la transformación lineal del ejemplo anterior. Usemos la matriz hallada en ese ejemplo y la proposición 4.18 para calcular $f(1, 3, 3)$.

Llamemos $v = (1, 3, 3)$. Tenemos que

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

Luego, $[v]_{B'} = (1, 2, 1)$. Entonces

$$([f(v)]_{B'})^t = M_{BB'}(f).([v]_B)^t = \begin{pmatrix} 0 & -\frac{1}{3} & 0 \\ 0 & 1 & -2 \end{pmatrix} . \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} \\ 0 \end{pmatrix}.$$

Luego, $[f(v)]_{B'} = (-\frac{2}{3}, 0)$. Entonces,

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

Por lo tanto, $f(v) = (-2, 0)$.

Lema 4.20. *Sean* $\mathbb{V}$ *y* $\mathbb{W}$ *espacios vectoriales y sea* $f : \mathbb{V} \to \mathbb{W}$ *una transformación lineal. Sea* $n = \dim(\mathbb{V})$ *y sea* $m = \dim(\mathbb{W})$. *Sea* B *una base de* $\mathbb{V}$ *y sea* B' *una base de* $\mathbb{W}$. *Sea* $A \in \mathbb{R}^{m \times n}$ *tal que para todo* $v \in \mathbb{V}$ *vale que* $A([v]_B)^t = ([f(v)]_{B'})^t$. *Entonces* $A = M_{BB'}(f)$.

Demostración. Por definición $M_{BB'}(f) \in \mathbb{R}^{m \times n}$ entonces las matrices A y $M_{BB'}(f)$ tienen el mismo tamaño. Veamos que para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$ vale que $A_{ij} = (M_{BB'}(f))_{ij}$.

Sean $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$. Sea e_j el j-ésimo vector de la base canónica de $\mathbb{R}^n$ y sea v_j el j-ésimo vector de la base B. Se tiene que

$$\begin{aligned} A_{ij} &= (A.(e_j)^t)_{i1} = (A.([v_j]_B)^t)_{i1} = (([f(v)]_{B'})^t)_{i1} = (M_{BB'}(f).([v_j]_B)^t)_{i1} = \\ &= (M_{BB'}(f).(e_j)^t)_{i1} = (M_{BB'}(f))_{ij} \end{aligned}$$

Luego, $A_{ij} = (M_{BB'}(f))_{ij}$ para todos $i \in \{1, 2, \dots, m\}$ y $j \in \{1, 2, \dots, n\}$.

Por lo tanto, $A = M_{BB'}(f)$. □

Proposición 4.21. *Sean* $\mathbb{V}$*,* $\mathbb{V}'$ *y* $\mathbb{V}''$ *espacios vectoriales y sean* $f : \mathbb{V} \to \mathbb{V}'$ *y* $g : \mathbb{V}' \to \mathbb{V}''$ *transformaciones lineales. Sean* B*,* B' *y* B'' *bases de* $\mathbb{V}$*,* $\mathbb{V}'$ *y* $\mathbb{V}''$ *respectivamente. Entonces*

$$M_{BB''}(g \circ f) = M_{B'B''}(g).M_{BB'}(f) \ .$$

Demostración. Notemos que si $n = \dim(\mathbb{V})$, $m = \dim(\mathbb{V}')$ y $l = \dim(\mathbb{V}'')$ entonces $M_{B'B''}(g) \in \mathbb{R}^{l\times m}$ y $M_{BB'}(f) \in \mathbb{R}^{m\times n}$ y luego $M_{B'B''}(g).M_{BB'}(f) \in \mathbb{R}^{l\times n}$.

Sea $v \in \mathbb{V}$. Se tiene que

$$M_{B'B''}(g).M_{BB'}(f).([v]_B)^t = M_{B'B''}(g).([f(v)]_{B'})^t = ([g(f(v))]_{B''})^t = ([(g\circ f)(v)]_{B''})^t \ .$$

Luego $M_{B'B''}(g).M_{BB'}(f).([v]_B)^t = ([(g\circ f)(v)]_{B''})^t$ para todo $v \in \mathbb{V}$.

Por lo tanto, aplicando el lema previo obtenemos que

$$M_{B'B''}(g).M_{BB'}(f) = M_{BB''}(g\circ f) \ .$$

□

Definición 4.22. Sea $\mathbb{V}$ un espacio vectorial y sean B y B' bases de $\mathbb{V}$. Definimos la *matriz de cambio de base de* B *a* B' como la matriz $C_{BB'} = M_{BB'}(\mathrm{Id}_{\mathbb{V}})$.

Observación 4.23. Notemos que si $v \in \mathbb{V}$ entonces $C_{BB'}.([v]_B)^t = ([v]_{B'})^t$ ya que

$$C_{BB'}.([v]_B)^t = M_{BB'}(\mathrm{Id}_{\mathbb{V}}).([v]_B)^t = ([\mathrm{Id}_{\mathbb{V}}(v)]_{B'})^t = ([v]_{B'})^t$$

Ejemplo 4.24. Sean $B = \{(1,2),(2,1)\}$ y $B' = \{(1,1),(0,1)\}$ bases de $\mathbb{R}^2$ y sea $v \in \mathbb{R}^2$ tal que $[v]_B = (1,-1)$. Calcular la matriz $C_{BB'}$ y utilizarla para calcular $[v]_{B'}$.

Solución: Tenemos que

$$\begin{array}{rclclcl} \mathrm{Id}(1,2) & = & (1,2) & = & 1.(1,1) & + & 1.(0,1) \ . \\ \mathrm{Id}(2,1) & = & (2,1) & = & 2.(1,1) & + & (-1).(0,1) \ . \end{array}$$

Entonces

$$\begin{array}{rcl} [\mathrm{Id}(1,2)]_{B'} & = & (1,1) \\ [\mathrm{Id}(2,1)]_{B'} & = & (2,-1) \end{array}$$

y luego

$$C_{BB'} = M_{BB'}(\mathrm{Id}) = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \ .$$

Ahora calculemos $[v]_{B'}$:

$$([v]_{B'})^t = C_{BB'}.([v]_B)^t = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} . \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \ .$$

Luego, $[v]_{B'} = (-1,2)$.

Ejemplo 4.25. Sea $B = \{(3,1),(-1,2)\}$ base de $\mathbb{R}^2$. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal tal que $M_{EB}(f) = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$. Calcular $M_{EE}(f)$ y hallar una fórmula para f.

Solución: Consideremos el siguiente diagrama:

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "f"] \arrow[rr, bend left, "f"] & \mathbb{R}^2 \arrow[r, "\mathrm{Id}"] & \mathbb{R}^2 \\ E & B & E \end{tikzcd}$$

Luego
$$M_{EE}(f) = M_{EE}(\mathrm{Id} \circ f) = M_{BE}(\mathrm{Id}).M_{EB}(f) = C_{BE}.M_{EB}(f)\ .$$
Calculemos $C_{BE} = M_{BE}(\mathrm{Id})$.
$$\begin{array}{rcrcrcr} \mathrm{Id}(3,1) & = & (3,1) & = & 3.(1,0) & + & 1.(0,1)\ . \\ \mathrm{Id}(-1,2) & = & (-1,2) & = & -1.(1,0) & + & 2.(0,1)\ . \end{array}$$
Entonces
$$\begin{array}{rcr} [\mathrm{Id}(3,1)]_E & = & (3,1) \\ [\mathrm{Id}(-1,2)]_E & = & (-1,2) \end{array}$$
y luego
$$C_{BE} = M_{BE}(\mathrm{Id}) = \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix}\ .$$
Entonces
$$M_{EE}(f) = C_{BE}.M_{EB}(f) = \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix} . \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 5 \\ 2 & 4 \end{pmatrix}\ .$$

Hallemos ahora una fórmula para f. Sea $v = (x,y) \in \mathbb{R}^2$. Tenemos que
$$([f(x,y)]_E)^t = ([f(v)]_E)^t = M_{EE}(f).([v]_E)^t = \begin{pmatrix} -1 & 5 \\ 2 & 4 \end{pmatrix} . \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x+5y \\ 2x+4y \end{pmatrix}\ .$$
Luego $[f(x,y)]_E = (-x+5y, 2x+4y)$ y por lo tanto $f(x,y) = (-x+5y, 2x+4y)$.

Notación 4.26. En ocasiones notaremos $M_{BB}(f)$ por $M_B(f)$ y $M_{EE}(f)$ por $M(f)$.

Proposición 4.27. *Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y sea $f : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea B una base de $\mathbb{V}$ y sea B' una base de $\mathbb{W}$. Si f es un isomorfismo entonces $M_{B'B}(f^{-1}) = (M_{BB'}(f))^{-1}$.*

Demostración. Consideremos el siguiente diagrama:
$$\begin{tikzcd} \mathbb{V} \arrow[r, bend left, "f"] & \mathbb{W} \arrow[l, bend left, "f^{-1}"] \end{tikzcd}$$
$$B \qquad B'$$
Tenemos que
$$M_{B'B}(f^{-1}).M_{BB'}(f) = M_{BB}(f^{-1} \circ f) = M_{BB}(\mathrm{Id}) = I$$
y
$$M_{BB'}(f).M_{B'B}(f^{-1}) = M_{B'B'}(f \circ f^{-1}) = M_{B'B'}(\mathrm{Id}) = I\ .$$
Por lo tanto, $M_{B'B}(f^{-1}) = (M_{BB'}(f))^{-1}$. □

Observación 4.28. Sea $\mathbb{V}$ un espacio vectorial y sean B y B' bases de $\mathbb{V}$. Entonces
$$C_{BB'} = (C_{B'B})^{-1}$$
ya que $C_{BB'} = M_{BB'}(\mathrm{Id}) = M_{BB'}(\mathrm{Id}^{-1}) = (M_{B'B}(\mathrm{Id}))^{-1} = (C_{B'B})^{-1}$.

Ejemplo 4.29. Calcular la inversa del isomorfismo $f : \mathbb{R}^2 \to \mathbb{R}^2$ definido por

$$f(x, y) = (x + 3y, 2x + 4y) \ .$$

Solución: Calculemos primero $M_{EE}(f)$. Tenemos que

$$\begin{array}{rcccccc} f(1,0) & = & (1,2) & = & 1.(1,0) & + & 2.(0,1) \ . \\ f(0,1) & = & (3,4) & = & 3.(1,0) & + & 4.(0,1) \ . \end{array}$$

Entonces

$$\begin{array}{rcl} [f(1,0)]_E & = & (1,2) \\ [f(0,1)]_E & = & (3,4) \end{array}$$

y luego

$$M_{EE}(f) = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \ .$$

Aplicando ahora la proposición anterior obtenemos que

$$\begin{array}{rcl} M_{EE}(f^{-1}) & = & (M_{EE}(f))^{-1} = \dfrac{1}{\det(M_{EE}(f))}.\mathrm{adj}(M_{EE}(f)) = \dfrac{1}{-2}.\begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix} = \\ & = & \begin{pmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} \ . \end{array}$$

Entonces si $(a, b) \in \mathbb{R}^2$ tenemos que

$$([f^{-1}(a,b)]_E)^t = M_{EE}(f^{-1}).([(a,b)]_E)^t = \begin{pmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}.\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -2a + \frac{3}{2}b \\ a - \frac{1}{2}b \end{pmatrix}.$$

Luego $[f^{-1}(a,b)]_E = (-2a + \frac{3}{2}b, a - \frac{1}{2}b)$ y por lo tanto $f^{-1}(a,b) = (-2a + \frac{3}{2}b, a - \frac{1}{2}b)$.

Ejemplo 4.30. Sea $\mathbb{V}$ un espacio vectorial y sean

$$B = \{v_1, v_2, v_3\} \ \text{ y } \ B' = \{v_1 + v_3, -2v_1 + v_2 - v_3, -v_1 + v_2 - v_3\}$$

bases de $\mathbb{V}$. Sea $f : \mathbb{V} \to \mathbb{V}$ una transformación lineal tal que

$$M_{B'B'}(f) = \begin{pmatrix} 0 & 1 & -1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \ .$$

(a) Hallar $M_{BB}(f)$.

(b) Hallar bases de $\mathrm{Nu}(f)$ y de $\mathrm{Im}(f)$.

Solución:
(a) Consideremos el siguiente diagrama:

$$\begin{tikzcd} \mathbb{V} \arrow[r, "\mathrm{Id}"] \arrow[rrr, bend left, "f"] & \mathbb{V} \arrow[r, "f"] & \mathbb{V} \arrow[r, "\mathrm{Id}"] & \mathbb{V} \\ B & B' & B' & B \end{tikzcd}$$

Luego

$$M_{BB}(f) = M_{BB}(\mathrm{Id} \circ f \circ \mathrm{Id}) = M_{B'B}(\mathrm{Id}).M_{B'B'}(f).M_{BB'}(\mathrm{Id}) = C_{B'B}.M_{B'B'}(f).C_{BB'} \ .$$

Calculemos $C_{B'B} = M_{B'B}(\mathrm{Id})$.

$$\begin{array}{rcccccccc}
\mathrm{Id}(v_1+v_3) & = & v_1+v_3 & = & 1.v_1 & + & 0.v_2 & + & 1.v_3 \\
\mathrm{Id}(-2v_1+v_2-v_3) & = & -2v_1+v_2-v_3 & = & (-2).v_1 & + & 1.v_2 & + & (-1).v_3 \\
\mathrm{Id}(-v_1+v_2-v_3) & = & -v_1+v_2-v_3 & = & (-1).v_1 & + & 1.v_2 & + & (-1).v_3\ .
\end{array}$$

Entonces

$$\begin{array}{rcl}
[\mathrm{Id}(v_1+v_3)]_B & = & (1,0,1) \\
[\mathrm{Id}(-2v_1+v_2-v_3)]_B & = & (-2,1,-1) \\
[\mathrm{Id}(-v_1+v_2-v_3)]_B & = & (-1,1,-1)
\end{array}$$

y luego

$$C_{B'B} = M_{B'B}(\mathrm{Id}) = \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}.$$

Calculemos ahora $C_{BB'}$. Por 4.28 tenemos que $C_{BB'} = (C_{B'B})^{-1}$. Luego

$$\begin{aligned}
C_{BB'} &= (C_{B'B})^{-1} = \frac{1}{\det(C_{B'B})}.\mathrm{adj}(C_{B'B}) = \frac{1}{-1}.\begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 1 \end{pmatrix} = \\
&= \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}
\end{aligned}$$

Entonces

$$\begin{aligned}
M_{BB}(f) &= C_{B'B}.M_{B'B'}(f).C_{BB'} = \\
&= \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}.\begin{pmatrix} 0 & 1 & -1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}.\begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix} = \\
&= \begin{pmatrix} -7 & -1 & 12 \\ 3 & 0 & -6 \\ -5 & -1 & 8 \end{pmatrix}.
\end{aligned}$$

(b) Hallemos primero una base de $\mathrm{Im}(f)$. Como $B = \{v_1, v_2, v_3\}$ es base de $\mathbb{V}$ entonces B genera $\mathbb{V}$ y luego, por 4.8, tenemos que $f(B) = \{f(v_1), f(v_2), f(v_3)\}$ genera $\mathrm{Im}(f)$. Calculemos $f(v_1)$, $f(v_2)$ y $f(v_3)$. Tenemos que

$$([f(v_1)]_B)^t = M_{BB}(f).([v_1]_B)^t = \begin{pmatrix} -7 & -1 & 12 \\ 3 & 0 & -6 \\ -5 & -1 & 8 \end{pmatrix}.\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -7 \\ 3 \\ -5 \end{pmatrix}$$

Luego $[f(v_1)]_B = (-7,3,-5)$. y entonces $f(v_1) = -7v_1 + 3v_2 - 5v_3$.

Por otra parte,

$$([f(v_2)]_B)^t = M_{BB}(f).([v_2]_B)^t = \begin{pmatrix} -7 & -1 & 12 \\ 3 & 0 & -6 \\ -5 & -1 & 8 \end{pmatrix}.\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$$

Luego $[f(v_2)]_B = (-1,0,-1)$. y entonces $f(v_2) = -v_1 - v_3$.

También,

$$([f(v_3)]_B)^t = M_{BB}(f).([v_3]_B)^t = \begin{pmatrix} -7 & -1 & 12 \\ 3 & 0 & -6 \\ -5 & -1 & 8 \end{pmatrix}.\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 12 \\ -6 \\ 8 \end{pmatrix}$$

Luego $[f(v_1)]_B = (12, -6, 8)$. y entonces $f(v_3) = 12v_1 - 6v_2 + 8v_3$.

Por lo tanto, $\mathrm{Im}(f) = \mathrm{Gen}\{-7v_1 + 3v_2 - 5v_3 \ , \ -v_1 - v_3 \ , \ 12v_1 - 6v_2 + 8v_3\}$. Ahora bien, notemos que

$$12v_1 - 6v_2 + 8v_3 = (-2).(-7v_1 + 3v_2 - 5v_3) + 2.(-v_1 - v_3)$$

Luego,

$$\begin{aligned} \mathrm{Im}(f) &= \mathrm{Gen}\{-7v_1 + 3v_2 - 5v_3 \ , \ -v_1 - v_3 \ , \ 12v_1 - 6v_2 + 8v_3\} = \\ &= \mathrm{Gen}\{-7v_1 + 3v_2 - 5v_3 \ , \ -v_1 - v_3\} \ . \end{aligned}$$

Veamos si el conjunto $\{-7v_1 + 3v_2 - 5v_3 \ , \ -v_1 - v_3\}$ es linealmente independiente. Sean $\alpha, \beta \in \mathbb{R}$ tales que

$$\alpha(-7v_1 + 3v_2 - 5v_3) + \beta(-v_1 - v_3) = 0 \ .$$

Entonces

$$(-7\alpha - \beta)v_1 + 3\alpha v_2 + (-5\alpha - \beta)v_3 = 0$$

y como $B = \{v_1, v_2, v_3\}$ es base de $\mathbb{V}$ entonces es un conjunto linealmente independiente y de la igualdad de arriba obtenemos que

$$\begin{cases} -7\alpha - \beta &= 0 \\ 3\alpha &= 0 \\ -5\alpha - \beta &= 0 \end{cases}$$

de donde se obtiene claramente que $\alpha = 0$ y $\beta = 0$.

Por lo tanto, el conjunto $\{-7v_1 + 3v_2 - 5v_3 \ , \ -v_1 - v_3\}$ es linealmente independiente y entonces es base de $\mathrm{Im}(f)$.

Hallemos ahora una base de $\mathrm{Nu}(f)$. Para ello veamos qué vectores v de $\mathbb{V}$ pertenecen a $\mathrm{Nu}(f)$. Sea $v \in \mathbb{V}$. Entonces existen $\alpha, \beta, \gamma \in \mathbb{R}$ tales que $v = \alpha v_1 + \beta v_2 + \gamma v_3$. Se tiene que $v \in \mathrm{Nu}(f)$ si y sólo si $f(v) = 0$, lo cual ocurre si y sólo si $[f(v)]_B = 0$. Ahora bien, tenemos que

$$([f(v)]_B)^t = M_{BB}(f).([v]_B)^t = \begin{pmatrix} -7 & -1 & 12 \\ 3 & 0 & -6 \\ -5 & -1 & 8 \end{pmatrix} . \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} -7\alpha - \beta + 12\gamma \\ 3\alpha - 6\gamma \\ -5\alpha - \beta + 8\gamma \end{pmatrix} .$$

Luego

$$[f(v)]_B = (-7\alpha - \beta + 12\gamma \ , \ 3\alpha - 6\gamma \ , \ -5\alpha - \beta + 8\gamma) \ .$$

Por lo tanto, $v \in \mathrm{Nu}(f)$ si y sólo si $(-7\alpha - \beta + 12\gamma \ , \ 3\alpha - 6\gamma \ , \ -5\alpha - \beta + 8\gamma) = (0, 0, 0)$.

Resolvemos entonces el sistema

$$\begin{cases} -7\alpha - \beta + 12\gamma = 0 \\ 3\alpha - 6\gamma = 0 \\ -5\alpha - \beta + 8\gamma = 0 \end{cases}$$

Aplicamos el método de eliminación de Gauss:

$$\left(\begin{array}{rrr|r} -7 & -1 & 12 & 0 \\ 3 & 0 & -6 & 0 \\ -5 & -1 & 8 & 0 \end{array}\right) \underset{F_1 \leftrightarrow F_2}{\rightsquigarrow} \left(\begin{array}{rrr|r} 3 & 0 & -6 & 0 \\ -7 & -1 & 12 & 0 \\ -5 & -1 & 8 & 0 \end{array}\right) \underset{F_1 \leftarrow \frac{1}{3}F_1}{\rightsquigarrow}$$

$$\rightsquigarrow \left(\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ -7 & -1 & 12 & 0 \\ -5 & -1 & 8 & 0 \end{array}\right) \underset{F_2 \leftarrow F_2 + 7F_1}{\rightsquigarrow}$$

$$\rightsquigarrow \left(\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ 0 & -1 & -2 & 0 \\ -5 & -1 & 8 & 0 \end{array}\right) \underset{F_3 \leftarrow F_3 + 5F_1}{\rightsquigarrow}$$

$$\rightsquigarrow \left(\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right) \underset{F_3 \leftarrow F_3 - F_2}{\rightsquigarrow} \left(\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Obtenemos entonces el siguiente sistema

$$\begin{cases} \alpha \qquad\quad - 2\gamma = 0 \\ \quad - \beta - 2\gamma = 0 \\ \qquad\qquad\quad 0 = 0 \end{cases}$$

Este sistema está en forma escalonada y tiene dos ecuaciones no triviales y tres incógnitas. Luego el sistema tiene infinitas soluciones. Además, α y β son variables principales y γ es variable libre. Luego γ puede tomar cualquier valor real. Si $\gamma = t$ con $t \in \mathbb{R}$ obtenemos de la segunda ecuación que $\beta = -2t$ y de la primera ecuación que $\alpha = 2t$. Por lo tanto, la solución general del sistema es $(2t, -2t, t)$, $t \in \mathbb{R}$.

Esto nos dice que el vector $v = \alpha v_1 + \beta v_2 + \gamma v_3$ pertenece a $\mathrm{Nu}(f)$ si y sólo si existe $t \in \mathbb{R}$ tal que $\alpha = 2t$, $\beta = -2t$ y $\gamma = t$. Luego, $v \in \mathrm{Nu}(f)$ si y sólo si $v = 2tv_1 - 2tv_2 + tv_3$ para algún $t \in \mathbb{R}$. Y como

$$2tv_1 - 2tv_2 + tv_3 = t(2v_1 - 2v_2 + v_3)$$

para cualquier $t \in \mathbb{R}$ obtenemos que $v \in \mathrm{Nu}(f)$ si y sólo si $v = t(2v_1 - 2v_2 + v_3)$ para algún $t \in \mathbb{R}$, es decir, $v \in \mathrm{Nu}(f)$ si y sólo si $v \in \mathrm{Gen}\{2v_1 - 2v_2 + v_3\}$. Por lo tanto,

$$\mathrm{Nu}(f) = \mathrm{Gen}\{2v_1 - 2v_2 + v_3\}\ .$$

Veamos que el conjunto $\{2v_1 - 2v_2 + v_3\}$ es linealmente independiente. Como este conjunto está formado por un único vector, por 3.19 basta probar que este vector es no nulo, lo cual es cierto pues el conjunto $B = \{v_1, v_2, v_3\}$ es linealmente independiente (por ser base de $\mathbb{V}$) y los escalares de la combinación lineal $2v_1 - 2v_2 + v_3$ no son todos nulos. Así, el conjunto $\{2v_1 - 2v_2 + v_3\}$ es linealmente independiente y, por lo tanto, es base de $\mathrm{Nu}(f)$.

Capítulo 5

Autovalores y autovectores

Definición 5.1. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Sea $v \in \mathbb{R}^n - \{0\}$. Decimos que v es *autovector* de A si existe $\lambda \in \mathbb{R}$ tal que $Av^t = \lambda v^t$. En este caso diremos que λ es un *autovalor* de A y que v es un *autovector asociado al autovalor* λ o que v es un *autovector de autovalor* λ.

Definición 5.2. Sea $\mathbb{V}$ un espacio vectorial y sea $f : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Sea $v \in \mathbb{V} - \{0\}$. Decimos que v es *autovector* de f si existe $\lambda \in \mathbb{R}$ tal que $f(v) = \lambda v$. En este caso diremos que λ es un *autovalor* de f y que v es un *autovector asociado al autovalor* λ o que v es un *autovector de autovalor* λ.

Ejemplo 5.3. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$f(x, y, z) = (x + y + z, y, 2x + 2z) \ .$$

Observemos que $f(1, 0, 2) = (3, 0, 6) = 3.(1, 0, 2)$. Luego $(1, 0, 2)$ es un autovector de f de autovalor 3.

Además, notemos que $f(5, 0, 10) = (15, 0, 30) = 3\,.(5, 0, 10)$. Luego $(5, 0, 10)$ es otro autovector de f de autovalor 3.

Por otra parte, $f(0, 1, 0) = (0, 1, 0) = 1.(0, 1, 0)$. Luego $(0, 1, 0)$ es un autovector de f de autovalor 1.

También se tiene que $f(1, 0, -1) = (0, 0, 0) = 0\,.(1, 0, -1)$. Luego $(1, 0, -1)$ es un autovector de f de autovalor 0.

Observación 5.4. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Sea $f_A : \mathbb{R}^n \to \mathbb{R}^n$ la transformación lineal definida por $f_A(x) = (Ax^t)^t$. Entonces v es autovector de A de autovalor λ si y sólo si v es autovector de f_A de autovalor λ. En efecto,

$$f_A(v) = \lambda v \ \Leftrightarrow \ (Av^t)^t = \lambda v \ \Leftrightarrow \ Av^t = (\lambda v)^t \ \Leftrightarrow \ Av^t = \lambda v^t$$

Proposición 5.5. *Sea* $\mathbb{V}$ *un espacio vectorial y sea* $f : \mathbb{V} \to \mathbb{V}$ *una transformación lineal. Sea* B *una base de* $\mathbb{V}$*. Entonces* v *es autovector de* f *de autovalor* λ *si y sólo si* $[v]_B$ *es autovector de* $M_{BB}(f)$ *de autovalor* λ*.*

Demostración. Si v es autovector de f de autovalor λ entonces $f(v) = \lambda v$ y luego

$$M_{BB}([v]_B)^t = ([f(v)]_B)^t = ([\lambda v]_B)^t = (\lambda[v]_B)^t = \lambda([v]_B)^t \ .$$

Así, $[v]_B$ es autovector de $M_{BB}(f)$ de autovalor λ.

Recíprocamente, si $[v]_B$ es autovector de $M_{BB}(f)$ de autovalor λ entonces

$$M_{BB}([v]_B)^t = \lambda([v]_B)^t \ .$$

Luego,

$$[f(v)]_B = (M_{BB}([v]_B)^t)^t = (\lambda([v]_B)^t)^t = \lambda[v]_B = [\lambda v]_B$$

y entonces $f(v) = \lambda v$. Así, v es autovector de f de autovalor λ. □

Proposición 5.6. *Sea $\mathbb{V}$ un espacio vectorial y sea $f : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Sea $\lambda \in \mathbb{R}$ y sea $S_\lambda = \{v \in \mathbb{V} \,/\, f(v) = \lambda v\}$. Entonces S_λ es un subespacio de $\mathbb{V}$.*

La demostración de este hecho queda como ejercicio para el lector. El conjunto S_λ definido como en la proposición anterior se llama *subespacio asociado al autovalor λ de f*. Notemos que S_λ está formado por el vector 0 y por todos los autovectores de f de autovalor λ.

Proposición 5.7. *Sea $n \in \mathbb{N}$, sea $A \in \mathbb{R}^{n\times n}$ y sea $\lambda \in \mathbb{R}$. Entonces λ es autovalor de A si y sólo si* $\det(A - \lambda I) = 0$.

Demostración. Tenemos que λ es autovalor de A si y sólo si existe $v \in \mathbb{R}^n - \{0\}$ tal que $Av^t = \lambda v^t$. Ahora bien,

$$Av^t = \lambda v^t \;\Leftrightarrow\; Av^t = \lambda I v^t \;\Leftrightarrow\; Av^t - \lambda I v^t = 0 \;\Leftrightarrow\; (A - \lambda I)v^t = 0$$

y luego, λ es autovalor de A si y sólo si existe $v \in \mathbb{R}^n - \{0\}$ tal que $(A - \lambda I)v^t = 0$.

Por lo tanto, λ es autovalor de A si y sólo si el sistema $(A - \lambda I)x^t = 0$ admite alguna solución no trivial y puesto que este sistema es homogéneo, esto último ocurre si y sólo si $\det(A - \lambda I) = 0$. □

Este resultado da lugar a la siguiente definición.

Definición 5.8. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Definimos el *polinomio característico de A* por $\chi_A(x) = \det(A - xI)$.

Observación 5.9.

(1) Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Sea $\lambda \in \mathbb{R}$. Entonces λ es autovalor de A si y sólo si λ es raíz del polinomio característico de A.

(2) Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Entonces $\chi_A(x)$ es un polinomio de grado n.

(3) Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n\times n}$. Entonces A tiene a lo sumo n autovalores distintos.

Ejemplo 5.10. Sea $A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}$. Hallar los autovalores y autovectores de A.

Solución: Tenemos que

$$\begin{aligned}
\chi_A(x) &= \det(A - xI) = \det\left(\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} - x\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) = \\
&= \det\begin{pmatrix} 1-x & 0 & 2 \\ 1 & 1-x & 0 \\ 1 & 0 & 2-x \end{pmatrix} = (-1)^{2+2}.(1-x).\det\begin{pmatrix} 1-x & 2 \\ 1 & 2-x \end{pmatrix} = \\
&= (1-x).((1-x)(2-x) - 2) = (1-x)(2 - x - 2x + x^2 - 2) = \\
&= (1-x)(-3x + x^2)\ .
\end{aligned}$$

Entonces $\chi_A(x) = 0$ si y sólo si $1 - x = 0$ o $-3x + x^2 = 0$. Resolviendo estas ecuaciones obtenemos que las raíces del polinomio característico son 1, 0 y 3. Por lo tanto, los autovalores de A son 1, 0 y 3.

Hallemos ahora los autovectores correspondientes a cada uno de los autovalores anteriores. Comencemos con el autovalor 3. Un vector $v \in \mathbb{R}^3 - \{0\}$ es autovector de A de autovalor 3 si y sólo si $Av^t = 3v^t$ o, equivalentemente, si y sólo si $(A - 3I)v^t = 0$. Si planteamos $v = (x, y, z)$ nos queda

$$\begin{pmatrix} 1-3 & 0 & 2 \\ 1 & 1-3 & 0 \\ 1 & 0 & 2-3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y obtenemos el sistema

$$\begin{cases} -2x & & + \; 2z & = 0 \\ x & - \; 2y & & = 0 \\ x & & - \; z & = 0 \end{cases}$$

Resolvemos este sistema aplicando el método de eliminación de Gauss.

$$\left(\begin{array}{ccc|c} -2 & 0 & 2 & 0 \\ 1 & -2 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right) \underset{F_1 \leftrightarrow F_3}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 1 & -2 & 0 & 0 \\ -2 & 0 & 2 & 0 \end{array}\right) \underset{F_2 \leftarrow F_2 - F_1}{\rightsquigarrow}$$

$$\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & -2 & 1 & 0 \\ -2 & 0 & 2 & 0 \end{array}\right) \underset{F_3 \leftarrow F_3 + 2F_1}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Obtenemos entonces el siguiente sistema

$$\begin{cases} x & & - \; z & = 0 \\ & - \; 2y & + \; z & = 0 \\ & & 0 & = 0 \end{cases}$$

Este sistema está en forma escalonada y tiene dos ecuaciones no triviales y tres incógnitas. Luego el sistema tiene infinitas soluciones. Además, x e y son variables principales y z es variable libre. Luego z puede tomar cualquier valor real. Si $z = t$ con $t \in \mathbb{R}$ obtenemos de la segunda ecuación que $y = \frac{1}{2}t$ y de la primera ecuación que $x = t$. Entonces la solución general del sistema es $(t, \frac{1}{2}t, t)$, $t \in \mathbb{R}$.

Por lo tanto, los autovectores de A de autovalor 3 son los vectores de la forma $(t, \frac{1}{2}t, t)$ con $t \in \mathbb{R} - \{0\}$.

El cálculo de los autovectores de autovalores 1 y 0 se deja como ejercicio para el lector. Además, dejamos para el lector comparar este ejemplo con el ejemplo 5.3.

Proposición 5.11. *Sea $\mathbb{V}$ un espacio vectorial y sea $f : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Sea $r \in \mathbb{N}$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_r$ autovalores distintos de f. Sean $v_1, v_2, \ldots, v_r$ autovectores de f asociados a los autovalores $\lambda_1, \lambda_2, \ldots, \lambda_r$ respectivamente. Entonces el conjunto $\{v_1, v_2, \ldots, v_r\}$ es linealmente independiente.*

Demostración. Si $r = 1$ el resultado es trivial pues como los autovectores son no nulos entonces el conjunto $\{v_1\}$ es linealmente independiente.

Demostremos ahora el caso $r = 2$. Debemos probar que el conjunto $\{v_1, v_2\}$ es linealmente independiente. Sean $\alpha_1, \alpha_2 \in \mathbb{R}$ tales que $\alpha_1 v_1 + \alpha_2 v_2 = 0$. Entonces

$$f(\alpha_1 v_1 + \alpha_2 v_2) = f(0)$$

y por ser f una transformación lineal obtenemos que

$$\alpha_1 f(v_1) + \alpha_2 f(v_2) = 0 \; .$$

Ahora como v_1 es autovector de f de autovalor λ_1 y v_2 es autovector de f de autovalor λ_2 tenemos que $f(v_1) = \lambda_1 v_1$ y $f(v_2) = \lambda_2 v_2$. Reemplazando esto en la igualdad de arriba obtenemos que

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 = 0 \; .$$

Ahora bien, teníamos que $\alpha_1 v_1 + \alpha_2 v_2 = 0$, luego, multiplicando ambos miembros de esta igualdad por λ_1 obtenemos que

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_1 v_2 = 0$$

y restando miembro a miembro esta igualdad con la obtenida previamente queda que

$$(\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_1 v_2) - (\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2) = 0$$

Luego,

$$\alpha_2(\lambda_1 - \lambda_2) v_2 = 0 \; .$$

Ahora bien, como v_2 es un autovector, entonces $v_2 \neq 0$ y luego, por el ítem (6) de 3.5 tenemos que $\alpha_2(\lambda_1 - \lambda_2) = 0$. Pero $\lambda_1 \neq \lambda_2$, entonces $\lambda_1 - \lambda_2 \neq 0$ y luego $\alpha_2 = 0$.

Ahora, como $\alpha_1 v_1 + \alpha_2 v_2 = 0$, obtenemos que $\alpha_1 v_1 = 0$ y como v_1 es un autovector, entonces $v_1 \neq 0$ y aplicando nuevamente el ítem (6) de 3.5 obtenemos que $\alpha_1 = 0$.

Hemos demostrado así que $\alpha_1 = 0$ y $\alpha_2 = 0$. Por lo tanto, el conjunto $\{v_1, v_2\}$ es linealmente independiente.

El caso general puede demostrarse por inducción con un razonamiento similar al del caso $r = 2$. Esta demostración queda como ejercicio para el lector. □

Definición 5.12. Sea $\mathbb{V}$ un espacio vectorial y sea $f : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Decimos que f es *diagonalizable* si existe una base B de $\mathbb{V}$ tal que la matriz $M_{BB}(f)$ es diagonal.

Definición 5.13. Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n \times n}$. Decimos que A es *diagonalizable* si existe una matriz inversible $C \in \mathbb{R}^{n \times n}$ tal que $C^{-1}AC$ es diagonal.

Notemos que si una matriz es diagonal entonces es diagonalizable.

La siguiente proposición nos muestra que las definiciones de matriz diagonalizable y transformación lineal diagonalizable están relacionadas.

Proposición 5.14. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n \times n}$. Sea $f_A : \mathbb{R}^n \to \mathbb{R}^n$ la transformación lineal definida por $f_A(x) = (Ax^t)^t$. Entonces A es diagonalizable si y sólo si f_A es diagonalizable.*

Demostración.
($\Leftarrow$) Supongamos que f_A es diagonalizable. Entonces existe una base B de $\mathbb{R}^n$ tal que $M_{BB}(f_A)$ es diagonal. Ahora bien,

$$M_{BB}(f_A) = C_{EB}.M_{EE}(f_A).C_{BE}$$

y $C_{EB} = C_{BE}^{-1}$. Además, $M_{EE}(f_A) = A$ pues si $\{e_1, e_2, \ldots, e_n\}$ es la base canónica de $\mathbb{R}^n$ entonces para todo $j \in \{1, 2, \ldots, n\}$ se tiene que $(f_A(e_j))^t = Ae_j^t$ y luego, para todo j, la j-ésima columna de $M_{EE}(f_A)$ es igual a la j-ésima columna de A.

Por lo tanto, queda que

$$M_{BB}(f_A) = C_{BE}^{-1}.A.C_{BE}$$

y como $M_{BB}(f_A)$ es diagonal concluimos que A es diagonalizable.
($\Rightarrow$) Supongamos que A es diagonalizable. Entonces existe una matriz inversible $C \in \mathbb{R}^{n\times n}$ tal que $C^{-1}AC$ es diagonal. Para cada $j \in \{1, 2, \ldots, n\}$ sea v_j el vector definido por la j-ésima columna de la matriz C, es decir, $v_j = (C_{1j}, C_{2j}, \ldots, C_{nj})$.

Sea $B = \{v_1, v_2, \ldots, v_n\}$. Veamos que B es base de $\mathbb{R}^n$. Supongamos que el conjunto B es linealmente dependiente. Entonces, por 3.21 tenemos que uno de los vectores $v_1, v_2, \ldots, v_n$ es combinación lineal de los otros. Luego, una de las filas de la matriz C^t es combinación lineal de las otras. Entonces restándole a esa fila múltiplos de las otras filas podemos llevar la matriz C^t a una matriz D que tenga una fila de ceros y luego, por 2.64 y 2.65, tenemos que

$$\det(C) = \det(C^t) = \det(D) = 0$$

lo cual es absurdo pues C es inversible. Por lo tanto, el conjunto B es linealmente independiente.

Entonces, como $\dim(\mathbb{R}^n) = n$, por 3.34 obtenemos que B es base de $\mathbb{R}^n$.

Ahora bien, notemos que $C_{BE} = C$ pues la matriz C_{BE} tiene como columnas los vectores de la base B. Además, como observamos en la demostración de la otra implicación $M_{EE}(f_A) = A$. Luego,

$$M_{BB}(f_A) = C_{EB}.M_{EE}(f_A).C_{BE} = C_{BE}^{-1}.M_{EE}(f_A).C_{BE} = C^{-1}.A.C$$

que es diagonal.

Por lo tanto, la transformación lineal f_A es diagonalizable. $\square$

Proposición 5.15. *Sea $\mathbb{V}$ un espacio vectorial y sea $f : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Entonces f es diagonalizable si y sólo si existe una base B de $\mathbb{V}$ formada por autovectores de f.*

Demostración.
($\Rightarrow$) Supongamos que f es diagonalizable. Entonces existe una base $B = \{v_1, v_2, \ldots, v_n\}$ de $\mathbb{V}$ tal que la matriz $M_{BB}(f)$ es diagonal. Para cada $i \in \{1, 2, \ldots, n\}$ sea $\lambda_i = (M_{BB}(f))_{ii}$. Sea $j \in \{1, 2, \ldots, n\}$. Se tiene que

$$([f(v_j)]_B)^t = M_{BB}(f).([v_j]_B)^t = M_{BB}(f).(e_j)^t$$

donde e_j es el j-ésimo vector de la base canónica de $\mathbb{R}^n$. Ahora bien, $M_{BB}(f).(e_j)^t$ coincide con la j-ésima columna de $M_{BB}(f)$ que tiene λ_j en la posición j y ceros en el resto de los lugares (pues $M_{BB}(f)$ es una matriz diagonal). Luego $[f(v_j)]_B = \lambda_j e_j$ y entonces

$$f(v_j) = 0v_1 + \ldots + 0v_{j-1} + \lambda_j v_j + 0v_{j+1} + \ldots + 0v_n = \lambda_j v_j \ .$$

Por lo tanto, v_j es autovector de f de autovalor λ_j (notar que $v_j \neq 0$ pues $v_j \in B$ y B es linealmente independiente).

Así, B es una base formada por autovectores de f.
($\Leftarrow$) Supongamos que existe una base $B = \{v_1, v_2, \ldots, v_n\}$ de $\mathbb{V}$ formada por autovectores de f. Entonces existen $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ tales que $f(v_i) = \lambda_i v_i$ para todo $i \in \{1, 2, \ldots, n\}$. Sea $j \in \{1, 2, \ldots, n\}$. Entonces

$$[f(v_j)]_B = [\lambda_j v_j]_B = \lambda_j e_j$$

donde e_j es el j-ésimo vector de la base canónica de $\mathbb{R}^n$. Luego, como la matriz $M_{BB}(f)$ tiene como j-ésima columna al vector $[f(v_j)]_B = \lambda_j e_j$ es claro que $M_{BB}(f)$ es una matriz diagonal. $\square$

De 5.4, 5.14 y 5.15 obtenemos el siguiente corolario.

Corolario 5.16. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n \times n}$. Entonces A es diagonalizable si y sólo si existe una base B de $\mathbb{R}^n$ formada por autovectores de A.*

Demostración. Por 5.14 tenemos que A es diagonalizable si y sólo si la transformación lineal f_A es diagonalizable, y por 5.15 esto último ocurre si y sólo si existe una base B de $\mathbb{R}^n$ formada por autovectores de f_A. Ahora bien, esto último es equivalente a que exista una base B de $\mathbb{R}^n$ formada por autovectores de A ya que, por 5.4, un vector de $\mathbb{R}^n$ es autovector de f_A si y sólo si es autovector de A. $\square$

Las siguiente proposición da una condición bajo la cual podemos asegurar que una transformación lineal sea diagonalizable. Su demostración queda como ejercicio para el lector.

Proposición 5.17. *Sea $\mathbb{V}$ un espacio vectorial de dimensión n y sea $f : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Si f tiene n autovalores distintos entonces f es diagonalizable.*

El análogo de la proposición anterior para matrices es el siguiente.

Proposición 5.18. *Sea $n \in \mathbb{N}$ y sea $A \in \mathbb{R}^{n \times n}$. Si A tiene n autovalores distintos entonces A es diagonalizable.*

La demostración de este hecho también se deja como ejercicio para el lector.