Notas para el curso de Álgebra Lineal I (F)

Centro de Matemática
Facultad de Ciencias
Universidad de la República

Andrés Abella

31 de julio de 2018

Contenidos

Introducción

Estas son notas para el primer curso de álgebra lineal de las licenciaturas que brinda el Instituto de Física de la Facultad de Ciencias. Si bien tienen bastante detalle, están pensadas como un acompañamiento del curso, pero no como para que un alumno pueda leerlas y entenderlas sin asistir a clase. En este último caso, tendría que complementarlas con algún texto más básico.

Notaciones. La mayoría de las demostraciones que se realizan mediante un cálculo directo han sido omitidas y solo va a aparecer el símbolo $\square$ al final del enunciado. En los otros casos, va a haber alguna idea más o menos detallada de cómo probar las partes no triviales de la demostración o una referencia bibliográfica; esto último para pruebas que podrían ser omitidas en este primer curso. El fin de cada demostración lo señalaremos también con el símbolo $\square$. Para indicar en una prueba que hemos llegado a una contradicción, usaremos el símbolo $\lightning$. A veces usaremos el símbolo ":=", por ejemplo en $a(x, y) := (ax, ay)$, queriendo decir que la definición del producto $a(x, y)$ es (ax, ay).

Capítulo 1

Sistemas de ecuaciones

Como este tema se estudia en los cursos de enseñanza secundaria, lo trataremos como un repaso, poniendo énfasis en las técnicas de resolución y tratando muy ligeramente la teoría. La fundamentación teórica se va a desarrollar más adelante en la sección 5.3, después de estudiar transformaciones lineales.

1.1. Caso general

Un *sistema de ecuaciones* es un conjunto de ecuaciones de la forma

$$\left\{\begin{array}{rcl} a_{11}x_1+\cdots+a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1+\cdots+a_{mn}x_n & = & b_m \end{array}\right. . \tag{1.1}$$

Los números a_{ij}, con $i=1,\dots,m$, $j=1,\dots,n$, son los *coeficientes*, $b_1,\dots,b_m$ son los *resultados* y $x_1,\dots,x_n$ son las *variables* o *incógnitas*, del sistema. Si queremos explicitar que hay m ecuaciones y n incógnitas, diremos que es un sistema $m\times n$. Si son pocas las variables, se suele escribir x,y,z,t en vez de x_1,x_2,x_3,x_4.

Una *solución* del sistema (1.1) es una n-upla[1] de números $(\alpha_1,\dots,\alpha_n)$ tales que si en (1.1) sustituimos x_1 por α_1, x_2 por α_2 y así seguimos hasta x_n por α_n, entonces se verifican todas las igualdades del sistema. Un sistema se dice *compatible* si tiene solución, en caso contrario se dice *incompatible*. Un sistema compatible se dice *determinado* si tiene una única solución e *indeterminado* en caso contrario. Más adelante veremos que un sistema indeterminado siempre tiene infinitas soluciones (observación 5.3.1). Por *resolver* un sistema entendemos que es clasificarlo en incompatible, compatible determinado o indeterminado, y en estos últimos casos hallar todas sus soluciones.

Un sistema del tipo

$$\left\{\begin{array}{rcl} x+y+z & = & 2 \\ 2y+z & = & 4 \\ 3z & = & 6 \end{array}\right. ,$$

se dice que es un sistema *escalerizado*. Este sistema es muy fácil de resolver, puesto que de la última ecuación se deduce que es $z=2$, sustituyendo este valor de z en la segunda ecuación deducimos que es $y=1$, y finalmente sustituyendo estos valores de y y z en la primer ecuación deducimos que es $x=-1$. Luego el sistema es compatible determinado y su solución es $x=-1$, $y=1$ y $z=2$. Otros ejemplos de sistemas escalerizados son

$$\left\{\begin{array}{rcl} x+y+z & = & 2 \\ 2y+z & = & 4 \\ 0 & = & 2 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{rcl} x+y+z & = & 2 \\ 2y+z & = & 4 \\ 0 & = & 0 \end{array}\right. .$$

[1]Una *n-upla* es un conjunto ordenado de n elementos, pudiendo ser $n=1,2,3,\dots$.

Claramente el primer sistema es incompatible y el segundo es compatible indeterminado. Las soluciones de este último se pueden escribir de la forma, $x = y - 2$, $z = -2y + 4$ y la variable y queda libre. Otra forma de escribir las soluciones es

$$x = \lambda - 2, \quad y = \lambda, \quad z = -2\lambda + 4,$$

siendo λ un número real arbitrario. Este último sistema es compatible indeterminado con *un grado de libertad* (la variable y está libre, si le damos un valor, entonces x y z quedan determinadas). Más adelante veremos ejemplos de sistemas escalerizados en que la cantidad de variables y ecuaciones es distinta.

Dos sistemas se dicen *equivalentes* ($\sim$) si tienen las mismas soluciones[2].

Proposición 1.1.1. *Operaciones con las ecuaciones de un sistema que dan lugar a sistemas equivalentes.*

- *Intercambiar el orden de las variables (es decir, el orden de los sumandos).*
- *Intercambiar el orden de las ecuaciones.*
- *Sustituir una ecuación por un múltiplo no nulo de la misma.*
- *Sumarle a una ecuación un múltiplo de otra.* $\square$

El método para resolver sistemas de ecuaciones conocido como de *eliminación Gaussiana* o *escalerización*, consiste en aplicar reiteradamente las operaciones descritas en la proposición anterior, hasta obtener un sistema equivalente que esté en forma escalerizada y por lo tanto sea fácil de resolver. Mostraremos el método de escalerización con los siguientes ejemplos.

Ejemplo 1.1.2. Queremos resolver el siguiente sistema

$$\left\{\begin{array}{rcl} 2x + 3y + 5z & = & 9 \\ x + 2y + 3z & = & 5 \\ 3x + 4y + 5z & = & 5 \end{array}\right. .$$

Luego vamos a ir aplicando las propiedades de la proposición anterior hasta llevarlo a un sistema escalerizado.

$$\left\{\begin{array}{rcl} 2x + 3y + 5z & = & 9 \\ x + 2y + 3z & = & 5 \\ 3x + 4y + 5z & = & 5 \end{array}\right. \sim \left\{\begin{array}{rcl} x + 2y + 3z & = & 5 \\ 2x + 3y + 5z & = & 9 \\ 3x + 4y + 5z & = & 5 \end{array}\right. \sim \left\{\begin{array}{rcl} x + 2y + 3z & = & 5 \\ -y - z & = & -1 \\ 3x + 4y + 5z & = & 5 \end{array}\right.$$

$$\sim \left\{\begin{array}{rcl} x + 2y + 3z & = & 5 \\ -y - z & = & -1 \\ -2y - 4z & = & -10 \end{array}\right. \sim \left\{\begin{array}{rcl} x + 2y + 3z & = & 5 \\ y + z & = & 1 \\ 2y + 4z & = & 10 \end{array}\right. \sim \left\{\begin{array}{rcl} x + 2y + 3z & = & 5 \\ y + z & = & 1 \\ 2z & = & 8 \end{array}\right. .$$

En el primer paso, intercambiamos la primer ecuación con la segunda; en el segundo, a la segunda ecuación le sumamos la primera multiplicada por -2; en el tercero, a la tercera le sumamos la primera multiplicada por -3; en el cuarto, multiplicamos la segunda y tercer ecuación por -1; en el quinto, a la tercera ecuación le sumamos la segunda multiplicada por -2. Luego este sistema es compatible determinado, y su solución es $x = -1$, $y = -3$ y $z = 4$.

Ejemplo 1.1.3.

$$\left\{\begin{array}{rcl} x + 2y + 3z & = & -1 \\ 3x - y + 2z & = & 1 \\ -2x + 3y + z & = & 2 \end{array}\right. \sim \left\{\begin{array}{rcl} x + 2y + 3z & = & -1 \\ -7y - 7z & = & 4 \\ 7y + 7z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x + 2y + 3z & = & -1 \\ -7y - 7z & = & 4 \\ 0 & = & 4 \end{array}\right. .$$

En el primer paso, a la segunda ecuación le sumamos la primera multiplicada por -3 y a la tercera le sumamos la primera multiplicada por 2; en el segundo paso, sumamos la segunda ecuación a la tercera. Este sistema resultó ser incompatible, porque es claro que la última ecuación no tiene solución.

[2] Acá estamos sobreentendiendo que si uno de los sistemas es incompatible, entonces el otro también lo es.

Ejemplo 1.1.4.

$$\left\{\begin{array}{rcl} x+2y+3z & = & -1 \\ 3x-y+2z & = & 1 \\ -2x+3y+z & = & -2 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & -1 \\ -7y-7z & = & 4 \\ 7y+7z & = & -4 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & -1 \\ -7y-7z & = & 4 \\ 0 & = & 0 \end{array}\right. .$$

Observar que los coeficientes de este sistema son los mismos que los del sistema del ejemplo anterior, así que los pasos de reolución son los mismos (solo cambian los resultados). Las soluciones se pueden escribir de la forma $x=-z+\frac{1}{7}$, $y=-z-\frac{4}{7}$ y la variable z está libre. Luego es un sistema compatible indeterminado con un grado de libertad.

Ejemplo 1.1.5.

$$\left\{\begin{array}{rcl} x+y-z & = & 1 \\ -x-y+z & = & -1 \\ 2x+2y-2z & = & 2 \end{array}\right. \sim \left\{\begin{array}{rcl} x+y-z & = & 1 \\ 0 & = & 0 \\ 0 & = & 0 \end{array}\right. .$$

Acá a la segunda ecuación le sumamos la primera y a la tercera le sumamos la primera multiplicada por -2. Las soluciones se pueden escribir de la forma $z=x+y+1$ y las variables x e y están libres. Este es un ejemplo de un sistema compatible indeterminado con *dos grados de libertad*. Notar que en sistemas con más variables, pueden aparecer soluciones que tengan más grados de libertad.

Los anteriores son ejemplos de sistemas *cuadrados*, es decir que tienen tantas ecuaciones como variables. A continuación veremos un par de ejemplos de aplicación del método de escalerización para resolver sistemas que no son cuadrados.

Ejemplo 1.1.6. Un sistema 2×3:

$$\left\{\begin{array}{rcl} x+y+z & = & 2 \\ x+2y-z & = & 3 \end{array}\right. \sim \left\{\begin{array}{rcl} x+y+z & = & 2 \\ y-2z & = & 1 \end{array}\right. .$$

Acá restamos la primer ecuación a la segunda. Este sistema es compatible indeterminado y las soluciones se pueden escribir de la forma $x=1-3z$, $y=1+2z$ y la variable z está libre, luego tiene un grado de libertad.

Ejemplo 1.1.7. Un sistema 3×2:

$$\left\{\begin{array}{rcl} x+2y & = & 1 \\ 2x-y & = & 2 \\ -x+3y & = & -1 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y & = & 1 \\ -5y & = & 0 \\ 5y & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y & = & 1 \\ -5y & = & 0 \\ 0 & = & 0 \end{array}\right. .$$

Acá primero a la segunda ecuación le sumamos la primera multiplicada por -2, y a la tercera le sumamos la primera; luego a la tercera le sumamos la segunda. El sistema es compatible determinado y la solución es $x=1$ e $y=0$.

1.2. Sistemas homogéneos

Un sistema de ecuaciones se dice *homogéneo* si todas sus ecuaciones aparecen igualadas a cero, es decir si es de la forma

$$\left\{\begin{array}{rcl} a_{11}x_1+\cdots+a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{m1}x_1+\cdots+a_{mn}x_n & = & 0 \end{array}\right.$$

Un sistema homogéneo siempre admite la solución *trivial* $x_1=\cdots=x_n=0$, por lo que siempre es compatible. Luego es indeterminado en caso de admitir alguna solución no trivial, y es determinado si la única solución que admite es la trivial.

En los siguientes ejemplos estudiamos los sistemas homogéneos asociados a los sistemas de los ejemplos anteriores (es decir obtenidos cambiando la columna de resultados por una columna de ceros). Como las operaciones para resolverlos son las mismas que para los no homogéneos, simplemente escribiremos el camino que lleva a su resolución sin explicar más.

Ejemplo 1.2.1.

$$\left\{\begin{array}{rcl} 2x+3y+5z & = & 0 \\ x+2y+3z & = & 0 \\ 3x+4y+5z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ 2x+3y+5z & = & 0 \\ 3x+4y+5z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ -y-z & = & 0 \\ 3x+4y+5z & = & 0 \end{array}\right.$$

$$\sim \left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ -y-z & = & 0 \\ -2y-4z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ y+z & = & 0 \\ 2y+4z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ y+z & = & 0 \\ 2z & = & 0 \end{array}\right. .$$

El sistema es compatible determinado y su única solución es la trivial $x = y = z = 0$.

Ejemplo 1.2.2.

$$\left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ 3x-y+2z & = & 0 \\ -2x+3y+z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ -7y-7z & = & 0 \\ 7y+7z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y+3z & = & 0 \\ -7y-7z & = & 0 \\ 0 & = & 0 \end{array}\right. .$$

El sistema es compatible indeterminado. Una forma de escribir sus soluciones es $x = -z$, $y = -z$ y z libre, luego tiene un grado de libertad. Observar que tomando valores no nulos de z obtenemos soluciones no triviales del sistema, por ejemplo para $z = 1$ obtenemos que $x = y = -1$ y $z = 1$ es solución del sistema.

Ejemplo 1.2.3.

$$\left\{\begin{array}{rcl} x+y-z & = & 0 \\ -x-y+z & = & 0 \\ 2x+2y-2z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+y-z & = & 0 \\ 0 & = & 0 \\ 0 & = & 0 \end{array}\right. .$$

El sistema es compatible indeterminado con dos grados de libertad. Las soluciones las podemos escribir de la forma $z = x + y$, x e y libres.

Ejemplo 1.2.4. Un sistema 2×3:

$$\left\{\begin{array}{rcl} x+y+z & = & 0 \\ x+2y-z & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+y+z & = & 0 \\ y-2z & = & 0 \end{array}\right. .$$

El sistema es compatible indeterminado con un grado de libertad. Las soluciones las podemos escribir de la forma $x = -3z$, $y = 2z$ y z libre.

Ejemplo 1.2.5. Un sistema 3×2:

$$\left\{\begin{array}{rcl} x+2y & = & 0 \\ 2x-y & = & 0 \\ -x+3y & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y & = & 0 \\ -5y & = & 0 \\ 5y & = & 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+2y & = & 0 \\ -5y & = & 0 \\ 0 & = & 0 \end{array}\right. .$$

El sistema es compatible determinado y su única solución es la trivial $x = y = 0$.

Capítulo 2

Geometría

En este capítulo daremos una fundamentación matemática para los vectores que se estudian en los cursos básicos de física y veremos aplicaciones a problemas geométricos, en particular al estudio de rectas y planos en el espacio.

2.1. Vectores en el plano y el espacio

Plano.

Como es usual en los cursos de geometría analítica, mediante la introducción de un sistema de coordenadas identificamos el plano con el conjunto $\mathbb{R}^2 = \{(x, y) : \ x, y \in \mathbb{R}\}$.

A los elementos de $\mathbb{R}^2$ les llamamos *vectores* o *puntos* y para escribirlos usaremos en general las letras u, v, w. El *vector nulo* es $o = (0, 0)$, que coincide con el origen de coordenadas del plano. Cuando trabajamos con vectores, a los números los llamamos *escalares* y en general los escribiremos con las letras a, b, c.

Definimos el *producto de un escalar por un vector* y la *suma de vectores* mediante

$$a(x, y) := (ax, ay), \quad (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2),$$

para todo $a \in \mathbb{R}$, $(x, y), (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$. Estas definiciones coinciden con las de los cursos de física, cuando pensamos cada punto del plano como el extremo de un flecha (vector) que va del origen al punto. Si $v = (x, y)$ es un vector, entonces su *opuesto* $-v$ se define por $-v := (-x, -y)$.

Proposición 2.1.1. *Las operaciones recién definidas verifican las* propiedades de espacio vectorial[1]*:*

1. $u + v = v + u$, *para todo* $u, v \in \mathbb{R}^2$ *(conmutativa);*
2. $u + (v + w) = (u + v) + w$, *para todo* $u, v, w \in \mathbb{R}^2$ *(asociativa);*
3. $u + o = o + u = u$, *para todo* $u \in \mathbb{R}^2$ *(existencia de neutro);*
4. $u + (-u) = (-u) + u = o$, *para todo* $u \in \mathbb{R}^2$ *(existencia de opuesto);*
5. $1u = u$, *para todo* $u \in \mathbb{R}^2$*;*
6. $a(u + v) = au + av$, *para todo* $a \in \mathbb{R}$, $u, v \in \mathbb{R}^2$*;*
7. $(a + b)u = au + bu$, *para todo* $a, b \in \mathbb{R}$, $u \in \mathbb{R}^2$*;*
8. $a(bu) = (ab)u$, *para todo* $a, b \in \mathbb{R}$, $u \in \mathbb{R}^2$*.* □

Dos vectores u y v se dicen *colineales* o *paralelos* si existe un escalar a tal que $u = av$ o $v = au$. Notar que el vector nulo o es colineal con cualquier otro vector (es el único vector que verifica esta propiedad).

[1] El nombre "propiedades de espacio vectorial" va a quedar claro cuando estudiemos los espacios vectoriales.

Si $v = (x, y)$ es un vector, definimos su *norma* o *módulo* $\|v\|$ mediante

$$\|v\| := \sqrt{x^2 + y^2}.$$

Proposición 2.1.2. *La norma verifica las siguientes propiedades:*

1. $\|v\| \geq 0$, *para todo* $v \in \mathbb{R}^2$;
2. $\|v\| = 0$ *si y solo si* $v = o$;
3. $\|av\| = |a| \, \|v\|$, *para todo* $a \in \mathbb{R}$, $v \in \mathbb{R}^2$;
4. $\|u + v\| \leq \|u\| + \|v\|$, *para todo* $u, v \in \mathbb{R}^2$ *(*desigualdad triangular*).*

Dem. La prueba de las tres primeras afirmaciones es directa y queda como ejercicio. La prueba formal de la última afirmación se ve cuando se estudian los espacios vectoriales con producto interno, en el curso siguiente a este. En nuestro caso, nos conformamos con observar que si se hace un dibujo con los vectores u, v y $u + v$, se ve que la desigualdad triangular se deduce del hecho bien conocido de que en un triángulo, la longitud de un lado siempre es menor que la suma de las longitudes de los otros dos (o, equivalentemente, que la distancia más corta entre dos puntos es la línea recta). □

La *resta* de vectores se define por

$$(x_1, y_1) - (x_2, y_2) := (x_1 - x_2, y_1 - y_2),$$

para todo $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$. Notar que vale

$$w = u - v \quad \Leftrightarrow \quad w + v = u.$$

Mediante la resta podemos calcular la *distancia* $d(u, v)$ entre dos puntos $u, v \in \mathbb{R}^2$: si $u = (x_1, y_1)$ y $v = (x_2, y_2)$, entonces

$$d(u, v) := \|u - v\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

La *base canónica* del plano es el conjunto $\{\mathbf{i}, \mathbf{j}\}$, donde $\mathbf{i} = (1, 0)$ y $\mathbf{j} = (0, 1)$. Notar que vale

$$v = (x, y) \quad \Leftrightarrow \quad v = x\mathbf{i} + y\mathbf{j}.$$

Luego

$$x_1\mathbf{i} + y_1\mathbf{j} = x_2\mathbf{i} + y_2\mathbf{j} \quad \Leftrightarrow \quad x_1 = x_2 \text{ y } y_1 = y_2.$$

Un *versor* es un vector de norma 1. Los vectores $\mathbf{i}$ y $\mathbf{j}$ son versores. En general, si v es un vector no nulo, entonces

$$\frac{v}{\|v\|} := \frac{1}{\|v\|} v$$

es un versor colineal con v (y además tiene el mismo sentido).

Si u y v son dos vectores no nulos, su *producto escalar* $u \cdot v$ se define mediante

$$u \cdot v := \|u\| \|v\| \cos(\theta), \tag{2.1}$$

siendo $0 \leq \theta \leq \pi$ el ángulo[2] que forman u y v. Si $u = o$ o $v = o$, se define $u \cdot v = 0$. Usando la "fórmula del coseno" para triángulos, se deduce que en coordenadas vale

$$(x_1, y_1) \cdot (x_2, y_2) = x_1x_2 + y_1y_2,$$

para todo $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$. En términos de la base canónica, la fórmula anterior se escribe:

$$(x_1\mathbf{i} + y_1\mathbf{j}) \cdot (x_2\mathbf{i} + y_2\mathbf{j}) = x_1x_2 + y_1y_2.$$

Proposición 2.1.3. *El producto escalar verifica las siguientes propiedades:*

[2] Los ángulos los mediremos siempre en radianes.

1. $v \cdot v = \|v\|^2$, *para todo* $v \in \mathbb{R}^2$*;*

2. $u \cdot v = v \cdot u$, *para todo* $u, v \in \mathbb{R}^2$ *(conmutativa);*

3. $u \cdot (v + w) = u \cdot v + u \cdot w$, *para todo* $u, v, w \in \mathbb{R}^2$ *(distributiva);*

4. $(au) \cdot v = u \cdot (av) = a(u \cdot v)$, *para todo* $a \in \mathbb{R}$, $u, v \in \mathbb{R}^2$*;*

5. Desigualdad de Cauchy-Shwarz. *Para todo* $u, v \in \mathbb{R}^2$ *vale* $|u \cdot v| \leq \|u\| \, \|v\|$.
   Además vale $|u \cdot v| = \|u\| \, \|v\|$ *si y solo si* u *y* v *son colineales.*

Dem. La desigualdad de Cauchy-Shwarz se prueba tomando valor absoluto en la fórmula (2.1). □

Observación 2.1.4. El producto escalar sirve para calcular ángulos: si u y v son dos vectores no nulos, entonces el ángulo $0 \leq \theta \leq \pi$ que forman queda determinado por la fórmula

$$\cos(\theta) = \frac{u \cdot v}{\|u\| \|v\|}.$$

Decimos que dos vectores u y v son *ortogonales* y escribimos $u \perp v$ si verifican $u \cdot v = 0$. Los vectores $\mathbf{i}$ y $\mathbf{j}$ de la base canónica son ortogonales. Notar que el vector nulo o es ortogonal con cualquier otro vector, y es el único que verifica esto.

Si u es un vector no nulo y v es un vector cualquiera, entonces la *proyección* de v en la dirección de u, es el vector $\Pi_u(v)$ definido por

$$\Pi_u(v) = \frac{u \cdot v}{\|u\|^2} u.$$

La *componente de* v *en la dirección de* u es el escalar $\frac{u \cdot v}{\|u\|}$. Notar que vale

$$\Pi_u(v) = \frac{u \cdot v}{\|u\|} u_0,$$

siendo $u_0 := \frac{u}{\|u\|}$ un versor colineal con u. Luego $\|\Pi_u(v)\| = \left| \frac{u \cdot v}{\|u\|} \right| = \frac{|u \cdot v|}{\|u\|}$, es decir, la norma de la proyección coincide con el valor absoluto de la componente. Es un ejercicio verificar que $v - \Pi_u(v)$ es ortogonal a u, luego podemos descomponer el vector v en suma de dos vectores

$$v = \Pi_u(v) + \big(v - \Pi_u(v)\big),$$

en que $\Pi_u(v)$ es colineal con u y $v - \Pi_u(v)$ es ortogonal a u.

Ejemplo 2.1.5. Sean $u = (1, 1) = \mathbf{i} + \mathbf{j}$ y $v = (2, 3) = 2\mathbf{i} + 3\mathbf{j}$. Entonces

$$\Pi_u(v) = \frac{u \cdot v}{\|u\|^2} u = \frac{5}{2}(\mathbf{i} + \mathbf{j}) = \frac{5}{2}\mathbf{i} + \frac{5}{2}\mathbf{j} \qquad \text{y} \qquad v - \Pi_u(v) = 2\mathbf{i} + 3\mathbf{j} - \left(\frac{5}{2}\mathbf{i} + \frac{5}{2}\mathbf{j}\right) = -\frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}.$$

Luego $v = \left(\frac{5}{2}\mathbf{i} + \frac{5}{2}\mathbf{j}\right) + \left(-\frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}\right)$, con $\frac{5}{2}\mathbf{i} + \frac{5}{2}\mathbf{j}$ colineal con u y $-\frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}$ ortogonal a u.

Espacio

Todo lo que vimos en el plano se generaliza naturalmente al espacio, con solo agregar "una coordenada más". Introduciendo un sistema de coordenadas, al espacio lo identificamos con $\mathbb{R}^3 = \{(x, y, z) : \; x, y, z \in \mathbb{R}\}$. Como antes llamamos *vectores* o *puntos* a los elementos de $\mathbb{R}^3$ y *escalares* a los elementos de $\mathbb{R}$. Definimos el *producto de un escalar por un vector* y la *suma de vectores* mediante

$$a(x, y, z) := (ax, ay, az), \quad (x_1, y_1, z_1) + (x_2, y_2, z_2) := (x_1 + x_2, y_1 + y_2, z_1 + z_2),$$

para todo $a \in \mathbb{R}$, $(x, y, z), (x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$. Con estas operaciones, los vectores de $\mathbb{R}^3$ verifican las propiedades de espacio vectorial vistas en la proposición 2.1.1. La *base canónica* de $\mathbb{R}^3$ es el conjunto $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, donde $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$ y $\mathbf{k} = (0, 0, 1)$. Como antes es

$$(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

La *norma* es $\|(x,y,z)\| = \sqrt{x^2+y^2+z^2}$ y el producto escalar $u \cdot v = \|u\|\|v\|\cos(\theta)$ se calcula mediante

$$(x_1,y_1,z_1)\cdot(x_2,y_2,z_2) = x_1x_2+y_1y_2+z_1z_2 \quad \text{o} \quad (x_1\mathbf{i}+y_1\mathbf{j}+z_1\mathbf{k})\cdot(x_2\mathbf{i}+y_2\mathbf{j}+z_2\mathbf{k}) = x_1x_2+y_1y_2+z_1z_2. \tag{2.2}$$

La definición de versor, de vectores colineales y de vectores ortogonales es la misma que en $\mathbb{R}^2$ y se verifican las propiedades de las proposiciones 2.1.2 y 2.1.3. La *proyección* de un vector v en la dirección de un vector no nulo u, es el vector $\Pi_u(v)$ definido como en $\mathbb{R}^2$ por

$$\Pi_u(v) = \frac{u\cdot v}{\|u\|^2}u.$$

Luego $v = \Pi_u(v) + \big(v - \Pi_u(v)\big)$, en que $\Pi_u(v)$ es colineal con u y $v - \Pi_u(v)$ es ortogonal a u.

Ejemplo 2.1.6. Sean $u = (1,1,0) = \mathbf{i}+\mathbf{j}$ y $v = (2,1,2) = 2\mathbf{i}+\mathbf{j}+2\mathbf{k}$. Es $\|u\| = \sqrt{2}$, $\|v\| = 3$ y $u\cdot v = 3$. Luego

$$\cos(\theta) = \frac{3}{3\sqrt{2}} = \frac{1}{\sqrt{2}}$$

y por lo tanto el ángulo θ que forman u y v es $\pi/4$. La proyección de v en la dirección de u es

$$\Pi_u(v) = \frac{u\cdot v}{\|u\|^2}u = \frac{3}{2}(\mathbf{i}+\mathbf{j}) = \frac{3}{2}\mathbf{i} + \frac{3}{2}\mathbf{j},$$

y $v - \Pi_u(v) = \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + 2\mathbf{k}$. Luego v se puede escribir como una suma:

$$2\mathbf{i}+\mathbf{j}+2\mathbf{k} = \left(\frac{3}{2}\mathbf{i}+\frac{3}{2}\mathbf{j}\right) + \left(\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+2\mathbf{k}\right),$$

con $\frac{3}{2}\mathbf{i}+\frac{3}{2}\mathbf{j}$ colineal con $\mathbf{i}+\mathbf{j}$ y $\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+2\mathbf{k}$ perpendicular a $\mathbf{i}+\mathbf{j}$.

Rcordemos (ver [RH]) que si u y v son dos vectores no nulos, su *producto vectorial* es el vector $u \times v$ definido por lo siguiente.

- $\|u\times v\| = \|u\|\|v\|\operatorname{sen}(\theta)$, siendo $0 \le \theta \le \pi$ el ángulo entre u y v;
- la dirección de $u \times v$ es ortogonal al plano determinado por u y v;
- el sentido de $u \times v$ viene dado por la "regla de la mano derecha".

Si $u = o$ o $v = o$ ($o = (0,0,0)$ es el vector nulo de $\mathbb{R}^3$), se define $u \times v = o$.

Proposición 2.1.7. *El producto vectorial verifica las siguientes propiedades:*

1. $v \times v = -v \times u$, *para todo* $u, v \in \mathbb{R}^2$ *(anticonmutativa);*
2. $u \times (v+w) = u\times v + u \times w$, *para todo* $u,v,w \in \mathbb{R}^2$ *(distributiva);*
3. $(au)\times v = u \times (av) = a(u\times v)$, *para todo* $a \in \mathbb{R}$, $u, v \in \mathbb{R}^2$*;*
4. $u \times v = o$ *si y solo si* u *y* v *son colineales.* □

Observación 2.1.8. Las propiedades 1, 3 y 4 de la proposición 2.3 son fáciles de probar. La propiedad 2, no lo parece tanto, pero también asumiremos que se deduce de la definición de producto vectorial[3].

Si consideramos la base canónica de $\mathbb{R}^3$, las propiedades anteriores implican

$$\begin{gathered}\mathbf{i}\times\mathbf{j} = \mathbf{k}, \quad \mathbf{j}\times\mathbf{k} = \mathbf{i}, \quad \mathbf{k}\times\mathbf{i} = \mathbf{j},\\ \mathbf{j}\times\mathbf{i} = -\mathbf{k}, \quad \mathbf{k}\times\mathbf{j} = -\mathbf{i}, \quad \mathbf{i}\times\mathbf{k} = -\mathbf{j},\\ \mathbf{i}\times\mathbf{i} = \mathbf{j}\times\mathbf{j} = \mathbf{k}\times\mathbf{k} = o.\end{gathered}$$

[3]La prueba de esta afirmación, aparece como el ejercicio 24(b), en la página 62 de [RH].

Usando estas fórmulas y la proposición 2.1.7, se prueba que vale:

$$(x_1\mathbf{i} + y_1\mathbf{j} + z_1) \times (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}) = (y_1z_2 - z_1y_2)\mathbf{i} + (z_1x_2 - x_1z_2)\mathbf{j} + (x_1y_2 - y_1x_2)\mathbf{k}. \tag{2.3}$$

Esta fórmula escrita en coordenadas queda:

$$(x_1, y_1, z_1) \times (x_2, y_2, z_2) = (y_1z_2 - z_1y_2,\ z_1x_2 - x_1z_2,\ x_1y_2 - y_1x_2).$$

Una forma de recordar la fórmula (2.3), es observar que se obtiene formalmente como el desarrollo por la primer fila del siguiente determinante[4]

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}\mathbf{i} - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}\mathbf{j} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\mathbf{k} = (y_1z_2 - z_1y_2)\mathbf{i} + (z_1x_2 - x_1z_2)\mathbf{j} + (x_1y_2 - y_1x_2)\mathbf{k}.$$

Observación 2.1.9. Si tomamos la fórmula (2.3) como definición de producto vectorial, entonces todas las propiedades de la proposición 2.3 se prueban sin problema.

Aplicación 2.1.1. *Si u y v son dos vectores de $\mathbb{R}^3$ con $u \neq o$, entonces la componente de v en la dirección de u es $\frac{u \cdot v}{\|u\|}$ y la componente de v en la dirección ortogonal a u (en el plano de u y v, y dentro de él en el semiplano que contiene a v respecto a la recta determinada por o y u) es $\frac{\|u \times v\|}{\|u\|}$.*

Aplicación 2.1.2. *Si u y v son dos vectores no nulos, entonces el área del paralelogramo generado por u y v es $\|u \times v\|$.*

Definición 2.1.10. El *producto mixto* de tres vectores u, v, w es el escalar definido por $(u \times v) \cdot w$.

Usando las fórmulas (2.2) y (2.3) se deduce que el producto mixto se calcula mediante la fórmula siguiente. Si $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2)$ y $w = (x_3, y_3, z_3)$, entonces

$$(u \times v) \cdot w = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

Usando esta fórmula es fácil probar que vale $(u \times v) \cdot w = u \cdot (v \times w)$, para todo $u, v, w \in \mathbb{R}^3$.

Aplicación 2.1.3. *Si u, v y w son tres vectores no nulos, entonces el volumen del paralelepípedo generado por u, v y w es el valor absoluto del producto mixto de u, v y w, es decir $|(u \times v) \cdot w|$.*

2.2. Rectas y planos

Cuando trabajamos en geometría, a los elementos de $\mathbb{R}^3$ a veces es mejor pensarlos como puntos y otras veces como vectores flecha, esto último cuando lo que nos interesa es solo una dirección (es decir el punto $v \in \mathbb{R}^3$ lo pensamos como el extremo de una flecha que va del origen a v). En este sentido a veces usaremos las letras P, Q, R cuando los pensamos como puntos y u, v, w cuando los pensamos como vectores.

Recta que pasa por un punto y es paralela a una dirección. Sea r la recta que pasa por el punto $P = (a, b, c)$ y es paralela al vector no nulo $v = (v_1, v_2, v_3)$. Si $X = (x, y, z)$ es un punto arbitrario del espacio, entonces $X \in r$ si y solo si el vector $X - P$ es colineal con v, lo cual equivale a que exista $t \in \mathbb{R}$ tal que

$$X = P + tv \quad \Leftrightarrow \quad (x, y, z) = (a, b, c) + t(v_1, v_2, v_3). \tag{2.4}$$

Esta es la *ecuación vectorial* de r. Si desarrollamos la segunda igualdad obtenemos

$$\begin{cases} x = a + tv_1 \\ y = b + tv_2 \\ z = c + tv_3 \end{cases}. \tag{2.5}$$

[4]Para quienes no hayan estudiado determinantes, las fórmulas que necesitamos son las (3.2) y (3.3), en la sección 3.2.

Esta es la *ecuación paramétrica* de r y t es el *parámetro* de la ecuación. Si v_1, v_2 y v_3 son no nulos, entonces despejando t obtenemos

$$\frac{x-a}{v_1} = \frac{y-b}{v_2} = \frac{z-c}{v_3}. \tag{2.6}$$

Esta es la *ecuación cartesiana* de r. Notar que esta ecuación equivale a cualquiera de los siguientes sistemas

$$\begin{cases} \frac{x-a}{v_1} = \frac{y-b}{v_2} \\ \frac{y-b}{v_2} = \frac{z-c}{v_3} \end{cases} \sim \begin{cases} \frac{x-a}{v_1} = \frac{y-b}{v_2} \\ \frac{x-a}{v_1} = \frac{z-c}{v_3} \end{cases} \sim \begin{cases} \frac{x-a}{v_1} = \frac{z-c}{v_3} \\ \frac{y-b}{v_2} = \frac{z-c}{v_3} \end{cases}. \tag{2.7}$$

Ejemplo 2.2.1. Consideremos la recta r que pasa por el punto $P = (1, -1, 2)$ y es paralela al vector $v = (4, 3, 1)$. Las ecuaciones vectorial, paramétrica y cartesiana de r son, respectivamente

$$(x, y, z) = (1, -1, 2) + t(4, 3, 1); \qquad \begin{cases} x = 1 + 4t \\ y = -1 + 3t \\ z = 2 + t \end{cases}; \qquad \frac{x-1}{4} = \frac{y+1}{3} = z - 2.$$

La ecuación cartesiana se puede reescribir de la forma siguiente

$$\begin{cases} \frac{x-1}{4} = z - 2 \\ \frac{y+1}{3} = z - 2 \end{cases} \Leftrightarrow \begin{cases} x - 4z = -7 \\ y - 3z = -7 \end{cases}. \tag{2.8}$$

Para saber si un punto está en la recta, hay que ver si sus coordenadas verifican la ecuación de la misma. Por ejemplo, si $Q = (9, 5, 4)$, entonces $x = 9$, $y = 5$ y $z = 4$ es solución del sistema (2.8) y por lo tanto Q está en r.

Ejes coordenados. El eje Ox es una recta que pasa por origen $O = (0, 0, 0)$ y es paralela al vector $\mathbf{i} = (1, 0, 0)$. Luego sus ecuaciones vectorial y paramétrica son, respectivamente

$$(x, y, z) = (0, 0, 0) + t(1, 0, 0); \qquad \begin{cases} x = t \\ y = 0 \\ z = 0 \end{cases}.$$

En este caso la ecuación cartesiana de Ox es $\begin{cases} y = 0 \\ z = 0 \end{cases}$. Notar $Ox = \{(x, 0, 0) : \; x \in \mathbb{R}\}$.

En forma análoga se deduce que la ecuación cartesiana de Oy es $\begin{cases} x = 0 \\ z = 0 \end{cases}$ y la de Oz es $\begin{cases} x = 0 \\ y = 0 \end{cases}$.

Recta que pasa por dos puntos. Sea r la recta que pasa por dos puntos distintos P y Q. La recta r pasa por P y es paralela al vector $Q - P$, por lo tanto su ecuación vectorial es:

$$X = P + t(Q - P) \quad \Leftrightarrow \quad X = tQ + (1 - t)P.$$

Ejemplo 2.2.2. Si $P = (1, 2, 3)$ y $Q = (2, 2, 2)$, entonces $Q - P = (1, 0, -1)$ y por lo tanto las ecuaciones de la recta que pasa por P y Q son

$$(x, y, z) = (1, 2, 3) + t(1, 0, -1); \qquad \begin{cases} x = 1 + t \\ y = 2 \\ z = 3 - t \end{cases}; \qquad \begin{cases} x + z = 4 \\ y = 2 \end{cases}.$$

Plano que pasa por un punto y es perpendicular a una dirección. Sea Π el plano que pasa por el punto $P = (x_0, y_0, z_0)$ y es perpendicular al vector no nulo $n = (a, b, c)$. Un punto $X = (x, y, z)$ está en Π si y solo si el vector $X - P$ es ortogonal a n, es decir si X verifica

$$n \cdot (X - P) = 0 \quad \Leftrightarrow \quad a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad \Leftrightarrow \quad ax + by + cz = d,$$

siendo $d = ax_0 + by_0 + cz_0$. Esta última es la *ecuación cartesiana* del plano Π.

Ejemplo 2.2.3. Si Π es el plano que pasa por $P = (1, 2, 3)$ y es ortogonal a $n = (2, 4, -1)$, entonces su ecuación cartesiana se obtiene mediante

$$2(x-1) + 4(x-2) - (z-3) = 0 \qquad \Rightarrow \qquad 2x + 4y - z = 7.$$

Observación 2.2.4. Notar que la ecuación cartesiana del plano es una única ecuación $ax + by + cz = d$, a diferencia de la ecuación cartesiana de la recta $\frac{x-a}{v_1} = \frac{y-b}{v_2} = \frac{z-c}{v_3}$, que en realidad es un sistema de dos ecuaciones $\begin{cases} \frac{x-a}{v_1} = \frac{y-b}{v_2} \\ \frac{y-b}{v_2} = \frac{z-c}{v_3} \end{cases}$.

Planos coordenados. El plano horizontal Oxy es el plano que pasa por $O = (0, 0, 0)$ y es ortogonal a $\mathbf{k} = (0, 0, 1)$. Luego su ecuación cartesiana es $z = 0$. Análogamente se deduce que las ecuaciones cartesianas de los planos Oxz y Oyz son, respectivamente, $y = 0$ y $x = 0$.

Plano que pasa por un punto y es paralelo a dos vectores. Sea $P = (x_0, y_0, z_0)$ un punto, y $u = (u_1, u_2, u_3)$ y $v = (v_1, v_2, v_3)$ dos vectores no colineales. Consideremos primero el plano Π_0 que pasa por el origen O y es paralelo a u y a v. Un punto $X = (x, y, z)$ está en Π_0 si y solo existen $s, t \in \mathbb{R}$ tales que

$$X = su + tv \quad \Leftrightarrow \quad (x, y, z) = s(u_1, u_2, u_3) + t(v_1, v_2, v_3).$$

Esta es la *ecuación vectorial* de Π_0. Consideremos ahora el plano Π que pasa por P y es paralelo a u y a v. Un punto $X = (x, y, z)$ está en Π si y solo si $X - P \in \Pi_0$. Luego $X \in \Pi$ si y solo si existen $s, t \in \mathbb{R}$ tales que

$$X = P + su + tv \quad \Leftrightarrow \quad (x, y, z) = (x_0, y_0, z_0) + s(u_1, u_2, u_3) + t(v_1, v_2, v_3).$$

Esta es la *ecuación vectorial* de Π (notar que la ecuación vectorial de Π_0 es un caso particular de esta). Para obtener su ecuación cartesiana, observamos que Π pasa por el punto P y el vector $u \times v$ es ortogonal a Π, luego la ecuación de Π es $(X - P) \cdot (u \times v) = 0$. Notar que $(X - P) \cdot (u \times v)$ es un producto mixto, luego la ecuación cartesiana de Π se obtiene desarrollando la siguiente igualdad

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0. \tag{2.9}$$

Ejemplo 2.2.5. Sea Π el plano que pasa por $P = (2, 1, 2)$ y es paralelo a $u = (1, 2, 3)$ y $v = (2, -2, 1)$. La ecuación vectorial de Π es

$$(x, y, z) = (2, 1, 2) + s(1, 2, 3) + t(2, -2, 1) \quad \Rightarrow \quad \begin{cases} x = 2 + s + 2t \\ y = 1 + 2s - 2t \\ z = 2 + 3s + t \end{cases}.$$

Para obtener la ecuación paramétrica, desarrollamos

$$0 = \begin{vmatrix} x-2 & y-1 & z-2 \\ 1 & 2 & 3 \\ 2 & -2 & 1 \end{vmatrix} = (x-2)\begin{vmatrix} 2 & 3 \\ -2 & 1 \end{vmatrix} - (y-1)\begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} + (z-2)\begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix} = 8x + 5y - 6z - 9.$$

Luego la ecuación paramétrica de Π es $8x + 5y - 6z = 9$. Es un ejercicio el verificar que también se llega a la ecuación cartesiana despejando s y t en la ecuación vectorial.

Plano que pasa por tres puntos. Sean $P = (x_1, y_1, z_1)$, $Q = (x_2, y_2, z_2)$ y $R = (x_3, y_3, z_3)$, tres puntos no alineados. El plano Π que pasa por P, Q y R coincide con el plano que pasa por P y es paralelo a $Q - P$ y $R - P$, luego de la fórmula (2.9) deducimos que la ecuación cartesiana de Π es

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0.$$

Intersección de planos. Es bien conocido que la intersección de dos planos no paralelos da una recta. Veremos con un ejemplo cómo obtener su ecuación.

Consideremos la intersección r de los planos $\Pi:\ x+y-z=4$ y $\Pi':\ 2x-y-z=1$. Las coordenadas de los puntos de r tienen que verificar las ecuaciones anteriores, y por lo tanto son las soluciones del sistema

$$\begin{cases} x+y-z=4 \\ 2x-y-z=1 \end{cases}.$$

Vamos a operar con las ecuaciones del sistema para obtener uno equivalente en que cada ecuación tenga solo dos variables.

$$\begin{cases} x+y-z=4 \\ 2x-y-z=1 \end{cases} \sim \begin{cases} -3y+z=-7 \\ x-2y=-3 \end{cases} \sim \begin{cases} z=-7+3y \\ x=-3+2y \end{cases}.$$

En el primer paso de la equivalencia anterior, a la primer ecuación la multiplicamos por -2 y le sumamos la segunda, y a la segunda le restamos la primera. Observar que este sistema es compatible indeterminado con un grado de libertad. La solución se puede escribir de la forma

$$\begin{cases} x=-3+2t \\ y=t \\ z=-7+3t \end{cases},$$

siendo $t \in \mathbb{R}$ libre. Esta última es la ecuación paramétrica de r. Si nos interesa su ecuación vectorial o cartesiana, las podemos deducir de la ecuación paramétrica:

$$(x,y,z)=(-3,0,-7)+t(2,1,3) \quad \Leftrightarrow \quad \frac{x+3}{2}=y=\frac{z+7}{3}.$$

Observación 2.2.6. Notar que la ecuación cartesiana de la recta

$$\frac{x-a}{v_1}=\frac{y-b}{v_2}=\frac{z-c}{v_3} \quad \Leftrightarrow \quad \begin{cases} \frac{x-a}{v_1}=\frac{y-b}{v_2} \\ \frac{y-b}{v_2}=\frac{z-c}{v_3} \end{cases} \quad \Leftrightarrow \quad \begin{cases} v_2x-v_1y=v_2a-v_1b \\ v_3y-v_2z=v_3b-v_2c \end{cases},$$

en el fondo está dando la recta como intersección de dos planos. Por ejemplo, en el caso de la recta $r=\Pi\cap\Pi'$ del ejemplo anterior, obtenemos

$$r:\ \frac{x+3}{2}=y=\frac{z+7}{3} \quad \Leftrightarrow \quad \begin{cases} \frac{x+3}{2}=y \\ y=\frac{z+7}{3} \end{cases} \quad \Leftrightarrow \quad \begin{cases} x-2y=-3 \\ 3y-z=7 \end{cases}.$$

Luego r aparece descrita como la intersección de los planos de ecuaciones $x-2y=-3$ y $3y-z=7$.

Capítulo 3

Matrices y determinantes

Las matrices y los determinantes son herramientas básicaa para poder estudiar los espacios vectoriales y las transformaciones lineales, que estudiaremos en los capítulos siguientes. Los resultados que veremos sobre matrices son fáciles de probar, pero los de determinantes dan bastante más trabajo, así que omitiremos las pruebas más complejas (dando referencias para quien le interese estudiarlas).

3.1. Matrices

Una *matriz* es una tabla de números, por ejemplo

$$\begin{pmatrix} 2 & -3 & 1/5 \\ \sqrt{2} & 0 & -1 \end{pmatrix}$$

es un ejemplo de matriz 2×3 (tiene 2 filas y 3 columnas). En general una matriz 2×3 tiene la forma

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

con $a, b, \dots, f$ números reales arbitrarios. En general para matrices de tamaño relativamente grande es preferible usar subíndices en vez de letras distintas. En ese sentido la matriz anterior la podemos escribir

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

con $a_{ij} \in \mathbb{R}$, para todo $i = 1, 2$, $j = 1, 2, 3$. Notar que en la notación a_{ij}, el primer índice i indica la fila y el sgundo índice j indica la columna, a las que pertenece el elemento. Generalizando el ejemplo anterior, dados dos enteros positivos m y n, diremos que una *matriz* $m \times n$ es una tabla de números de la forma

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

en que los a_{ij} son números reales[1], para todo $i = 1, \dots, m$ y $j = 1, \dots, n$. Observar que una matriz $m \times n$ tiene m filas y n columnas. Los elementos a_{ij} se llaman los *coeficientes* o *entradas* de la matriz. En general una matriz $m \times n$ la escribiremos de la forma $A = (a_{ij})_{m \times n}$ y cuando no dé lugar a confusión usaremos la notación abreviada $A = (a_{ij})$. Si A es una matriz $m \times n$, diremos que el *tamaño* de A es $m \times n$. Al conjunto de todas las matrices de tamaño $m \times n$ lo escribiremos $M_{m \times n}$. Una *matriz columna* es una matriz de tamaño $m \times 1$ (para algún m) y una *matriz fila* es una matriz de tamaño $1 \times n$ (para algún n):

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \text{ matriz columna;} \qquad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix} \text{ matriz fila.}$$

[1]Los a_{ij} no tienen porqué ser necesariamente números reales, pueden ser números complejos o de otros tipos, pero en estas notas trabajaremos solo con números reales.

Definición 3.1.1. Dos matrices $A = (a_{ij})_{m\times n}$ y $B = (b_{ij})_{r\times s}$ son *iguales* y escribimos $A = B$ si $m = r$, $n = s$ y $a_{ij} = b_{ij}$, para todo $i = 1, \dots, m$ y $j = 1, \dots, n$.

Una matriz se dice *cuadrada* si tiene el mismo número de filas que de columnas, es decir si es de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Al espacio $M_{n\times n}$ de matrices cuadradas $n \times n$ lo escribimos M_n. Si $A = (a_{ij})$ es una matriz cuadrada $n \times n$, entonces la *diagonal principal* de A es la n-upla $(a_{11}, a_{22}, \dots, a_{nn})$. La matriz *identidad* I_n es la matriz de tamaño $n \times n$ que tiene el valor 1 en la diagonal principal y 0 en el resto, es decir que es de la forma:

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

Si introducimos el símbolo δ_{ij} llamado la *delta de Kronecker* definido por

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j \end{cases},$$

entonces $I_n = (\delta_{ij})_{n\times n}$. Cuando no de lugar a confusión, usaremos I en vez de I_n.

Matrices cuadradas especiales. Sea $A = (a_{ij})_{n\times n}$ una matriz cuadrada.

- A es *simétrica* si $a_{ij} = a_{ji}$, para todo $i, j = 1, \dots, n$ (es simétrica respecto a la diagonal principal).
- A es *antisimétrica* si $a_{ij} = -a_{ji}$, para todo $i, j = 1, \dots, n$.
- A es *triangular superior* si $a_{ij} = 0$, para todo $i > j$ (las entradas abajo de la diagonal principal son nulas).
- A es *triangular inferior* si $a_{ij} = 0$, para todo $i < j$ (las entradas arriba de la diagonal principal son nulas).
- A es *diagonal* si $a_{ij} = 0$, para todo $i \neq j$ (las entradas fuera de la diagonal principal son nulas).

Operaciones con matrices.

Trasposición. Si $A = (a_{ij})$ es una matriz de tamaño $m \times n$, su *traspuesta* es la matriz A^t de tamaño $n \times m$ obtenida intercambiando las filas con las columnas de A, es decir $A^t = (b_{ij})$ en que $b_{ij} = a_{ji}$, para todo $i = 1, \dots, n$ y $j = 1, \dots, m$. Observar que para cada par (m, n), la trasposición define una función $f : M_{m\times n} \to M_{n\times m}$ mediante $f(A) = A^t$, para todo $A \in M_{m\times n}$.

Observación 3.1.2. Valen los siguientes resultados.

- Para toda matriz A es $(A^t)^t = A$.
- Uma matriz A es simétrica si y solo si $A^t = A$.
- Uma matriz A es antisimétrica si y solo si[2] $A^t = -A$.
- Uma matriz A es triangular superior si y solo si A^t es triangular inferior.

[2] La matriz opuesta $-A$ está definida en el párrafo siguiente.

Suma y producto por escalar. Las operaciones de suma y producto por escalar que vimos en $\mathbb{R}^2$ y $\mathbb{R}^3$ se generalizan naturalmente a $\mathbb{R}^n = \{(x_1, \ldots, x_n) : \; x_1, \ldots, x_n \in \mathbb{R}\}$, n entero positivo arbitrario, definiendo

$$a(x_1, \ldots, x_n) := (ax_1, \ldots, ax_n), \quad (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

para todo $a \in \mathbb{R}$, $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$. A su vez estas operaciones las podemos generalizar a una suma y producto por escalar en $M_{m\times n}$ definiendo

$$c\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}, \quad c \in \mathbb{R};$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}.$$

La *matriz nula* es la matriz que tiene todas sus entradas nulas y la escribimos[3] O. Si $A = (a_{ij}) \in M_{m\times n}$, entonces su *opuesta* es la matriz $-A = (b_{ij}) \in M_{m\times n}$, definida por $b_{ij} := -a_{ij}$, para todo i, j. Por ejemplo, en matrices 2×3 es

$$O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \qquad -\begin{pmatrix} 2 & -3 & 0 \\ -5 & 1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 3 & 0 \\ 5 & -1 & -2 \end{pmatrix}.$$

Proposición 3.1.3. *Las matrices $m\times n$ con la suma y producto por escalar definidos anteriormente verifican las propiedades de espacio vectorial que vimos en la proposición 2.1.1.* □

Observación 3.1.4. Notar que una n-upla $(x_1, x_2, \ldots, x_n)$ es lo mismo que una matriz fila $1 \times n$, luego podemos identificar $\mathbb{R}^n$ con $M_{1\times n}$, y por lo tanto en $\mathbb{R}^n$ valen también las propiedades de espacio vectorial.

Producto de matrices. Si $A = (a_{ij}) \in M_{m\times n}$ y $B = (b_{ij}) \in M_{n\times p}$, entonces su *producto* es la matriz $AB = (c_{ij}) \in M_{m\times p}$, en que c_{ij} está definido por

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}, \tag{3.1}$$

para todo $i = 1, \ldots, m$ y $j = 1, \ldots, p$. Para recordar esta fórmula, es útil armar el siguiente diagrama

$$\begin{array}{cc} & \begin{pmatrix} \cdots & b_{ij} & \cdots \\ & \vdots & \\ \cdots & b_{nj} & \cdots \end{pmatrix} \\ \begin{pmatrix} \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{pmatrix} & \begin{pmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{pmatrix} \end{array}$$

El elemento c_{ij} se obtiene haciendo una especie de "producto escalar" del vector fila $(a_{i1}, \cdots, a_{in})$ y el vector columna $(b_{ij}, \cdots, b_{nj})$. Notar que para poder multiplicar una matriz A por otra B, es necesario que el número de columnas de A coincida con el de filas de B.

Ejemplo 3.1.5. Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$, entonces $AB = \begin{pmatrix} 5 & 4 & 5 & 0 \\ 11 & 10 & 14 & 0 \end{pmatrix}$.

[3] En general se usa 0 (el número cero) en vez de O (la letra o mayúscula), pero para evitar confusiones por ahora usaremos la letra O para la matriz nula. Observar además que hay una matriz nula para cada tamaño $m \times n$.

Observación 3.1.6. Si A y B son matrices cuadradas del mismo tamaño, entonces podemos realizar los productos AB y BA. Por ejemplo, si

$$A = \begin{pmatrix} 1 & 1 \\ -2 & -2 \end{pmatrix} \text{ y } B = \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix} \quad \Rightarrow \quad AB = \begin{pmatrix} 8 & 4 \\ -16 & -8 \end{pmatrix} \text{ y } BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Este ejemplo muestra varias particularidades del producto de matrices:

- es $AB \neq BA$, luego el producto de matrices no es conmutativo;
- es $BA = O$, pero $AB \neq O$;
- es $BA = O$, siendo $A \neq O$ y $B \neq O$; esto implica para el producto de matrices no vale la propiedad cancelativa.

Proposición 3.1.7. *Propiedades del producto de matrices.*

1. $A(BC) = (AB)C$, *para todo* $A \in M_{m\times n}$, $B \in M_{n\times p}$, $C \in M_{p\times q}$ *(asociativa).*
2. $A(B+C) = AB + AC$, *para todo* $A \in M_{m\times n}$, $B \in M_{n\times p}$, $C \in M_{n\times p}$ *(distributiva).*
3. $(A+B)C = AC + BC$, *para todo* $A \in M_{m\times n}$, $B \in M_{m\times n}$, $C \in M_{n\times p}$ *(distributiva).*
4. $AI_n = I_m A = A$ *y* $AO = OA = O$, *para todo* $A \in M_{m\times n}$, *siendo cada* O *una matriz nula de tamaño adecuado*[4]
5. $A(cB) = (cA)B = c(AB)$, *para todo* $A \in M_{m\times n}$, $B \in M_{n\times p}$ *y* $c \in \mathbb{R}$. □

Observación 3.1.8. 1. La propiedad asociativa permite escribir $ABC = (AB)C = A(BC)$ y en general escribir $A_1 \cdots A_n$ para un producto de n matrices (que se puedan multiplicar).

2. El único caso en que dadas dos matrices A y B, podemos calcular $A+B$ y AB, es cuando A y B son dos matrices cuadradas del mismo tamaño.
3. Como vimos anteriormente, la fórmula para el producto de matrices es complicada, pero para matrices diagonales queda simple

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & 0 & \cdots & 0 \\ 0 & a_{22}b_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn}b_{nn} \end{pmatrix}.$$

Definición 3.1.9. Si $A \in M_n$ y $k \in \mathbb{N}$, definimos la *potencia* k-ésima A^k de A mediante: $A^0 = I$ (la matriz identidad) y $A^{k+1} = A^k A$, para todo $k \in \mathbb{N}$. Es decir, $A^0 = I$, $A^1 = A$, $A^2 = AA$, $A^3 = AAA$, etc.

Observación 3.1.10. Es un ejercicio el probar que valen las siguientes propiedades

1. $(A+B)^t = A^t + B^t$ y $(cA)^t = cA^t$, para todo $A, B \in M_{m\times n}$ y $c \in \mathbb{R}$.
2. $(AB)^t = B^t A^t$, para todo $A \in M_{m\times n}$ y $B \in M_{n\times p}$.

Definición 3.1.11. Decimos que una matriz cuadrada $A \in M_n$ es *invertible* si existe una matriz $B \in M_n$ tal que $AB = BA = I$ (la matriz identidad).

Proposición 3.1.12. *Si* $A, B, C \in M_n$ *verifican* $AB = CA = I$, *entonces* $B = C$ *y por lo tanto* A *es invertible.* □

Corolario 3.1.13. *Si* $A \in M_n$ *es invertible, entonces la matriz* $B \in M_n$ *que verifica* $AB = BA = I$ *es única.* □

[4]Observar que en general en la igualdad $AO = OA = O$, las tres matrices O tienen distinto tamaño.

Por el corolario anterior, si la matriz $A \in M_n$ es invertible entonces existe una única matriz B que verifica $AB = BA = I$; esta matriz se llama la *inversa* de A y se escribe $B = A^{-1}$. Luego la inversa A^{-1} de A (en caso de existir) queda caracterizada por ser la única matriz que verifica $AA^{-1} = A^{-1}A = I$.

Proposición 3.1.14. *Sean $A, B, C \in M_n$, siendo A una matriz invertible. Entonces valen*

$$AB = C \quad \Leftrightarrow \quad B = A^{-1}C; \qquad BA = C \quad \Leftrightarrow \quad B = CA^{-1}. \quad \square$$

El siguiente resultado muestra que vale la propiedad cancelativa del producto de matrices, con la condición de que la matriz que multiplica a ambos lados sea invertible.

Corolario 3.1.15. *Sean $A, B, C \in M_n$.*

1. *Si $AB = O$ o $BA = O$ y A es invertible, entonces $B = O$.*
2. *Si $AB = AC$ o $BA = CA$ y A es invertible, entonces $B = C$.* $\square$

Corolario 3.1.16. *Si $A, B \in M_n$ verifican $AB = O$, siendo $A \neq O$ y $B \neq O$, entonces A y B no son invertibles.* $\square$

Ejemplo 3.1.17. En la observación 3.1.6 vimos que si $A = \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 1 \\ -2 & -2 \end{pmatrix}$, entonces $AB = O$. Esto implica que A y B no son invertibles.

Observación 3.1.18. Probar que una matriz es invertible usando la definición es posible, pero es muy trabajoso. Por ejemplo, para una matriz 4×4, implica a resolver un sistema de 16 ecuaciones con 16 incógnitas. Más adelante veremos que este problema se simplifica con el uso de los determinantes.

A continuación veremos algunos casos en que es fácil estudiar la invertibilidad.

1. Claramente la matriz identidad I es invertible y vale $I^{-1} = I$.
2. Si una matriz tiene una fila o columna formada solo por ceros, entonces no es invertible.
3. Una matriz diagonal es invertible si y solo si todas las entradas de la diagonal principal son no nulas. En ese caso es
$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} a_{11}^{-1} & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn}^{-1} \end{pmatrix}.$$

El siguiente resultado determina todas las matrices invertibles de tamaño 2×2.

Proposición 3.1.19. *Sea $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ una matriz arbitraria. Entonces A es invertible si y solo si $ad - bc \neq 0$. Si $ad - bc \neq 0$, entonces la inversa de A es*

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Dem. Si A es la matriz nula, el resultado es obvio. En caso contrario, si escribimos $\hat{A} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$, entonces vale $A\hat{A} = \hat{A}A = (ad - bc)I_2$. De esta relación se deduce la tesis. $\square$

Proposición 3.1.20. 1. *Si $A, B \in M_n$ son invertibles, entonces AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.*

2. *Si A es invertible, entonces A^{-1} es invertible y $\left(A^{-1}\right)^{-1} = A$.* $\square$

3.2. Determinantes

En este tema trabajaremos solo con matrices cuadradas.

Observar que en la proposición 3.1.19 probamos que una matriz $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ es invertible si y solo si $ad-bc \neq 0$. El número $ad-bc$ se llama el *determinante* de la matriz $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ y se escribe $\det(A) = ad-bc$. Luego la proposición 3.1.19 se puede reescribir diciendo que una matriz $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ es invertible si y solo si su determinante es no nulo, y en ese caso vale $A^{-1} = \frac{1}{\det(A)} \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$.

En general, el determinante de una matriz cuadrada $A \in M_n$ es un número que escribimos $\det(A)$ o $|A|$ y que para $n = 1, 2, 3$ se define por lo siguiente.

$\mathbf{n = 1.}$ Si $A = (a)$, entonces $\det(a) := a$.

$\mathbf{n = 2.}$ Si $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, entonces

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc. \tag{3.2}$$

$\mathbf{n = 3.}$ Si $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, entonces

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &:= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned} \tag{3.3}$$

Método de Sarrus. Es un método que <u>solo sirve</u> para calcular determinantes de matrices 3×3. Consiste en construir una tabla repitiendo las dos primeras filas de la matriz A, abajo de A:

$$\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}. \\ \hdashline a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array}$$

El determinante se obtiene sumando los productos de las diagonales que van de izquierda a derecha y de arriba hacia abajo, y luego restando los productos de las diagonales que van de derecha a izquierda y de arriba hacia abajo. El resultado queda

$$\det(A) = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21},$$

que claramente coincide con la fórmula (3.3).

Observación 3.2.1. El método de Sarrus no es el mejor método para calcular determinantes y solo lo explicamos acá por razones históricas y de cultura general. En la gran mayoría de los casos, el método que veremos a continuación es más rápido y eficiente (ver la observación 3.2.11 y los ejemplos siguientes).

La fórmula general para el determinante de una matriz $n \times n$ se define recursivamente, generalizando el método que usamos para definir la fórmula en el caso $n = 3$. Para la misma necesitamos introducir los cofactores, que definimos a continuación.

Consideremos $A = (a_{ij})$ una matriz $n \times n$. Para cada $i, j = 1, \ldots, n$, sea $\tilde{A}_{ij}$ la matriz $(n-1) \times (n-1)$ obtenida suprimiendo la fila i y la columna j de A. El *cofactor* correspondiente al lugar (i, j) de la matriz A es el número $\Delta_{ij} := (-1)^{i+j} \det\left(\tilde{A}_{ij}\right)$.

Por ejemplo, para una matriz $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, los cofactores Δ_{11}, Δ_{12} y Δ_{13}, son

$$\Delta_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32}, \quad \Delta_{12} = - \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = -a_{21}a_{33} + a_{23}a_{31},$$

$$\Delta_{11} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{21}a_{32} - a_{22}a_{31}.$$

Notar que la fórmula (3.3) se puede escribir usando cofactores

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\Delta_{11} + a_{12}\Delta_{12} + a_{13}\Delta_{13}.$$

Esta es la fórmula que vamos a generalizar. Si $A = (a_{ij}) \in M_n$, entonces el *determinante* de A se define en forma recursiva mediante

$$\det(A) := a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots + a_{1n}\Delta_{1n} = \sum_{j=1}^{n} a_{1j}\Delta_{1j}. \tag{3.4}$$

Si A es una matriz $n \times n$, entonces se dice que $\det(A)$ es un determinante de *orden* n. La fórmula anterior nos dice que para calcular un determinante de orden n, tenemos que calcular n determinantes de orden $n-1$. Luego, como sabemos calcular determinantes de orden 3, entonces podemos calcular determinantes de orden 4, sabiendo estos podemos calcular los de orden 5, y siguiendo así podemos calcular determinantes de orden n, para cualquier entero positivo n.

Ejemplo 3.2.2. El determinante de la matriz identidad I_n vale 1, para todo n. Esto se prueba fácil por inducción en n, usando la fórmula (3.4).

Observación 3.2.3. A partir de la fórmula (3.4) se puede obtener una fórmula explícita para el determinante de una matriz $n \times n$, que para el caso $n = 3$ es la fórmula (3.3). Para los lectores que tengan conocimientos de permutaciones, la fórmula del determinante de una matriz $A = (a_{ij}) \in M_n$ es

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sg}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots (\sigma) a_{n\sigma(n)}.$$

En la fórmula anterior, $\mathcal{S}_n$ es el conjunto de todas las permutaciones de n elementos (pensadas como biyecciones del conjunto $\{1, 2, \ldots, n\}$ en sí mismo) y $\operatorname{sg}(\sigma) = \pm 1$ es el signo de la permutación σ. La fórmula del determinante tiene $n!$ (factorial de n) sumandos; por ejemplo, para una matriz 5×5 tiene 120 sumandos. Como es de imaginar esta fórmula no es muy eficiente para el cálculo y tampoco tiene demasiada utilidad en la teoría, así que no vamos a seguir más por esta línea.

A continuación veremos algunas propiedades de los determinantes, que en particular sirven para simplificar los cálculos.

Proposición 3.2.4. *Propiedades del determinante.*

1. *Si la matriz B se obtiene intercambiando un par de filas de la matriz A, entonces* $\det(B) = -\det(A)$.

2. *Si la matriz B se obtiene multiplicando una fila cualquiera de una matriz A por una constante c, entonces* $\det(B) = c\det(A)$. *Es decir:*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ ca_{i1} & \cdots & ca_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

3. *El determinante es aditivo respecto a cada fila.*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + a'_{i1} & \cdots & a_{in} + a'_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a'_{i1} & \cdots & a'_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

4. *Si A tiene una fila i en la cual todos los elementos son nulos salvo el a_{ij}, entonces $\det(A) = a_{ij}\Delta_{ij}$.*

Dem. Las pruebas son por inducción en n, usando la fórmula (3.4) (ver en [F] el teorema 4.3, el lema de la página 202 y el teorema 4.5). □

Corolario 3.2.5. *Casos en que el determinante es nulo.*

1. *Si una matriz tiene una fila formada solo por ceros, entonces su determinante es nulo.*

2. *Si una matriz tiene dos filas iguales, entonces su determinante es nulo.* □

Corolario 3.2.6. *Si en una matriz le sumamos a una fila un múltiplo escalar de otra fila, entonces su determinante no varía.* □

Corolario 3.2.7. *Sea $A = (a_{ij}) \in M_n$. Para cada $i = 1, \ldots, n$ el determinante de A se puede calcular desarrollando a partir de la fila i:*

$$\det(A) = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in} = \sum_{j=1}^{n} a_{ij}\Delta_{ij}. \tag{3.5}$$

Dem. Se deduce de la parte (4) de la proposición 3.2.4, aplicada al siguiente desarrollo:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \quad \square$$

Proposición 3.2.8. *Para toda matriz cuadrada A vale $\det(A^t) = \det(A)$.*

Dem. Ver en [F] el teorema 4.8. □

Como la trasposición intercambia las filas con las columnas, entonces la proposición anterior implica que todas las propiedades que vimos para filas, valen también para columnas. Las resumimos en el siguiente corolario.

Corolario 3.2.9. *Propiedades del determinante respecto a las columnas.*

1. *Si la matriz B se obtiene intercambiando un par de columnas de la matriz A, entonces $\det(B) = -\det(A)$.*

2. *Si una matriz tiene una columna formada solo por ceros, entonces su determinante es nulo.*

3. *Si una matriz tiene dos columnas iguales, entonces su determinante es nulo.*

4. *Si la matriz B se obtiene multiplicando una columna cualquiera de una matriz A por una constante c, entonces $\det(B) = c\det(A)$. Es decir:*

$$\begin{vmatrix} a_{11} & \cdots & ca_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & ca_{nj} & \cdots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}.$$

5. *El determinante es aditivo respecto a cada columna.*

$$\begin{vmatrix} a_{11} & \cdots & a_{1j}+a'_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj}+a'_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a'_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a'_{nj} & \cdots & a_{nn} \end{vmatrix}.$$

6. *Si en una matriz le sumamos a una columna un múltiplo escalar de otra columna, entonces su determinante no varía.*

7. *Si A es tal que en una columna j todos los elementos son nulos salvo el a_{ij}, entonces* $\det(A) = a_{ij}\Delta_{ij}$.

8. *Sea $A = (a_{ij}) \in M_n$. Para cada $j = 1, \ldots, n$ el determinante de A se puede calcular desarrollando a partir de la columna j:*

$$\det(A) = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj} = \sum_{i=1}^{n} a_{ij}\Delta_{ij}. \quad \square$$

Corolario 3.2.10. *Si A es una matriz triangular superior o inferior, entonces su determinante se obtiene multiplicando los elementos de la diagonal principal. En particular esto vale para las matrices diagonales.* $\square$

Observación 3.2.11. La técnica para calcular determinantes consiste en aplicar reiteradamente el corolario 3.2.6 o la propiedad 6 del corolario 3.2.9, hasta llegar al determinante de una matriz que tiene solo un elemento no nulo en una fila o columna. Luego se reduce su orden mediante la propiedad 4 de la proposición 3.2.4 o la 7 del corolario 3.2.9 (usando también las otras propiedades para ir simplificando los cálculos). Reiterando este procedimiento, terminamos reduciendo el cálculo de un determinante de orden n al cálculo de un determinante de orden 2.

Ejemplos 3.2.12. A continuación veremos ejemplos de cálculo de determinantes. En lo que sigue, si por ejemplo escribimos $F_2 - F_1$ quiere decir que a la fila 2 le estamos restando la fila 1, análogamente $C_3 + 2C_4$ quiere decir que a la columna 3 le estamos sumando la columna 4 multiplicada por 2. Notar que en las sumas o restas anteriores siempre escribimos primero la fila o columna que es modificada. Solo vamos a aclarar, este tipo de operaciones, identificar las otras queda a cargo del lector.

1.

$$\begin{vmatrix} 1 & 5 & 7 \\ 0 & 3 & 1 \\ 2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 7 \\ 0 & 3 & 1 \\ 0 & -8 & -11 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ -8 & -11 \end{vmatrix} = -33 + 8 = -25.$$

El primer paso fue $F_3 - 2F_1$.

2.

$$\begin{vmatrix} 2 & 3 & 5 \\ 6 & 10 & 4 \\ 50 & 20 & 30 \end{vmatrix} = 2\begin{vmatrix} 2 & 3 & 5 \\ 3 & 5 & 2 \\ 50 & 20 & 30 \end{vmatrix} = 20\begin{vmatrix} 2 & 3 & 5 \\ 3 & 5 & 2 \\ 5 & 2 & 3 \end{vmatrix} = 20\begin{vmatrix} 2 & 3 & 5 \\ 1 & 2 & -3 \\ 5 & 2 & 3 \end{vmatrix} = 20\begin{vmatrix} 0 & -1 & 11 \\ 1 & 2 & -3 \\ 5 & 2 & 3 \end{vmatrix}$$

$$= 20\begin{vmatrix} 0 & -1 & 11 \\ 1 & 2 & -3 \\ 0 & -8 & 18 \end{vmatrix} = -20\begin{vmatrix} -1 & 11 \\ -8 & 18 \end{vmatrix} = -40\begin{vmatrix} -1 & 11 \\ -4 & 9 \end{vmatrix} = -1400.$$

El tercer paso es $F_2 - F_1$ (para obtener $a_{21} = 1$), el cuarto $F_1 - 2F_2$ y el quinto $F_3 - 5F_2$.

3.

$$\begin{vmatrix} 1 & -2 & 1 & 3 \\ 5 & 7 & 8 & 9 \\ 2 & 0 & -2 & 4 \\ 3 & 0 & -3 & 6 \end{vmatrix} = 6\begin{vmatrix} 1 & -2 & 1 & 3 \\ 5 & 7 & 8 & 9 \\ 1 & 0 & -1 & 2 \\ 1 & 0 & -1 & 2 \end{vmatrix} = 0.$$

El determinante final es nulo por tener dos filas iguales.

4.

$$\begin{vmatrix} 2 & 1 & 0 & 1 \\ -1 & 2 & 1 & 2 \\ 0 & -1 & 2 & 4 \\ 1 & 3 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 0 & 5 & 2 & 5 \\ -1 & 2 & 1 & 2 \\ 0 & -1 & 2 & 4 \\ 0 & 5 & 2 & 4 \end{vmatrix} = -(-1)\begin{vmatrix} 5 & 2 & 5 \\ -1 & 2 & 4 \\ 5 & 2 & 4 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ -1 & 2 & 4 \\ 5 & 2 & 4 \end{vmatrix} = \begin{vmatrix} -1 & 2 \\ 5 & 2 \end{vmatrix} = -12.$$

El primer paso es $F_1 + 2F_2$ y $F_4 + F_2$ (dos cambios al mismo tiempo), luego $F_1 - F_3$.

5.

$$\begin{vmatrix} 2 & 9 & 5 & 7 \\ 4 & 2 & 3 & 2 \\ 7 & 4 & 2 & 3 \\ 9 & 6 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 9 & 5 & 7 \\ 4 & 2 & 3 & 2 \\ 7 & 4 & 2 & 3 \\ 5 & 4 & 0 & 0 \end{vmatrix} = \begin{vmatrix} -7 & 9 & 5 & 7 \\ 2 & 2 & 3 & 2 \\ 3 & 4 & 2 & 3 \\ 1 & 4 & 0 & 0 \end{vmatrix} = \begin{vmatrix} -7 & 37 & 5 & 7 \\ 2 & -6 & 3 & 2 \\ 3 & -8 & 2 & 3 \\ 1 & 0 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 37 & 5 & 7 \\ -6 & 3 & 2 \\ -8 & 2 & 3 \end{vmatrix}$$

$$= -\begin{vmatrix} 37 & 5 & 2 \\ -6 & 3 & -1 \\ -8 & 2 & 1 \end{vmatrix} = -\begin{vmatrix} 25 & 11 & 0 \\ -6 & 3 & -1 \\ -14 & 5 & 0 \end{vmatrix} = -\begin{vmatrix} 25 & 11 \\ -14 & 5 \end{vmatrix} = 279.$$

El primer paso es $F_4 - F_2$, luego $C_1 - C_2$ (para obtener $a_{14} = 1$), luego $C_2 - 4C_1$, reducimos a orden 3, $C_3 - C_2$, $F_1 + 2F_2$ y $F_3 + F_2$ (dos cambios al mismo tiempo).

Observación 3.2.13. Hay que tener cuidado porque el determinante no se comporta bien respecto a las operaciones de suma y producto por escalar, es decir en general es

$$\det(A + B) \neq \det(A) + \det(B); \qquad \det(cA) \neq c\det(A),\ c \in \mathbb{R}.$$

Sin embargo el determinante preserva el producto de matrices, como lo muestra la siguiente proposición.

Proposición 3.2.14. *Para todo $A, B \in M_n$ vale* $\det(AB) = \det(A)\det(B)$.

Dem. Ver en [F] el teorema 4.7. □

Corolario 3.2.15. *Si $A \in M_n$ es invertible, entonces vale* $\det(A) \neq 0$ *y además* $\det\left(A^{-1}\right) = \frac{1}{\det(A)}$. □

Lema 3.2.16. *Sea $A \in M_n$. Si Δ_{ij} son los cofactores de A y definimos*

$$\hat{A} = \begin{pmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2n} \\ \vdots & \vdots & & \vdots \\ \Delta_{n1} & \Delta_{n2} & \cdots & \Delta_{nn} \end{pmatrix}^t = \begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix},$$

entonces vale $A\hat{A} = \hat{A}A = \det(A)I$, siendo $I \in M_n$ la mariz identidad.

Dem. Sea $A\hat{A} = (c_{ij})$. Entonces para todo $i, j = 1, \ldots, n$, es

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\Delta_{jk}.$$

Luego $c_{ii} = \sum_{k=1}^{n} a_{ik}\Delta_{ik} = \det(A)$ (desarrollo a partir de la fila i), para todo $i = 1, \ldots, n$. Por otro lado, si $i \neq j$, entonces $c_{ij} = \sum_{k=1}^{n} a_{ik}\Delta_{jk} = 0$, porque esta suma corresponde al desarrollo a partir de la fila i del determinante de una matriz que tiene dos filas iguales (la i y la j). En conclusión es $c_{ij} = \det(A)\delta_{ij}$ y por lo tanto $A\hat{A} = \det(A)I$. La prueba de $\hat{A}A = \det(A)I$ es análoga, usando desarrollos por columnas en vez de por filas. □

Teorema 3.2.17. *Una matriz $A \in M_n$ es invertible si y solo si* $\det(A) \neq 0$. *En caso afirmativo, la matriz inversa se obtiene mediante*

$$A^{-1} = \frac{1}{\det(A)}\begin{pmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2n} \\ \vdots & \vdots & & \vdots \\ \Delta_{n1} & \Delta_{n2} & \cdots & \Delta_{nn} \end{pmatrix}^t = \frac{1}{\det(A)}\begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix}.$$

Dem. El directo se deduce del corolario anterior. Para el recíproco, si $\det(A) \neq 0$, entonces del lema anterior se deduce que vale

$$A\left(\det(A)^{-1}\hat{A}\right) = \left(\det(A)^{-1}\hat{A}\right)A = I,$$

luego (por definición) la matriz A es invertible y su inversa es $A^{-1} = \det(A)^{-1}\hat{A}$. □

Observación 3.2.18. Notar que la fórmula para la inversa del teorema 3.1.19 es un caso particular de la fórmula anterior.

Corolario 3.2.19. *Si $A, B \in M_n$ verifican $AB = I$, entonces A y B son invertibles, y son inversas una de la otra.*

Dem. Si $AB = I$, entonces $\det(A)\det(B) = 1$ y por lo tanto A y B son invertibles. Usando esto es fácil probar que de $AB = I$ se deduce $B = A^{-1}$ y $A = B^{-1}$. □

Método para obtener la matriz inversa.

El teorema 3.2.17 nos permite saber si una matriz es invertible o no, y en caso de serlo nos da una fórmula para calcular su inversa. Lamentablemente la fórmula para la inversa implica cálculos muy largos, por lo que solo es práctica para matrices 2×2. Sin embargo existe un algoritmo simple para calcular la inversa que no necesita usar la fórmula anterior. Para explicar cómo es este algoritmo, necesitamos definir las operaciones elementales.

Definición 3.2.20. Sea $A \in M_{m\times n}$. Se llaman *operaciones elementales* en A a las siguientes:

1. Intercambiar dos filas o columnas de A.
2. Multiplicar una fila o columna de A por una constante no nula.
3. Sumarle a una fila o columna un múltiplo de otra fila o columna, respectivamente.

El algoritmo para hallar la inversa consiste en lo siguiente. Supongamos que tenemos una matriz $A \in M_n$ que sabemos es invertible y queremos hallar A^{-1}. Lo que hacemos es escribir la matriz A y a su derecha la matriz identidad $I \in M_n$. Luego, y esto es importante, elegimos si vamos a trabajar con columnas o con filas. Si por ejemplo decidimos trabajar con columnas, entonces solo vamos a poder seguir trabajando con columnas y no podemos trabajar con filas. De la misma forma, si empezamos trabajando con filas no podemos pasar a trabajar con columnas. Hecha la elección anterior, por ejemplo digamos que decidimos trabajar con columnas, entonces vamos a realizar operaciones elementales en las columnas de la matriz A, y cada vez que hagamos una operación en A la repetimos en la matriz I. La idea es realizar operaciones en A hasta transformarla en la matriz identidad I, cuando se llegó a este punto, la matriz I se transformó en A^{-1}.

A continuación veremos algunos ejemplos donde estudiamos la invertibilidad de matrices y calculamos las inversas. En lo que sigue usaremos las mismas notaciones de los ejemplos 3.2.12. Además, escribiremos $F_2 \leftrightarrow F_3$, para decir que intercambiamos la fila 2 con la fila 3, y análogamente para columnas.

Ejemplos 3.2.21. Sea $A = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$. Primero estudiamos su invertibilidad.

$$\begin{vmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{vmatrix} \underset{C_1 - C_3}{=} \begin{vmatrix} 1 & 3 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{vmatrix} \underset{F_2 \leftrightarrow F_3}{=} -\begin{vmatrix} 1 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{vmatrix} = 1.$$

Como $\det(A) \neq 0$, deducimos que A es invertible. Para calcular su inversa tenemos que elegir primero si trabajar con columnas o con filas. En este caso trabajaremos con columnas.

$$
\begin{array}{|ccc|ccc|c}
2 & 3 & 1 & 1 & 0 & 0 & \\
-1 & 0 & -1 & 0 & 1 & 0 & C_1 - C_3 \\
1 & 1 & 1 & 0 & 0 & 1 & \\
\hline
1 & 3 & 1 & 1 & 0 & 0 & \\
0 & 0 & -1 & 0 & 1 & 0 & C_2 - C_3 \\
0 & 1 & 1 & -1 & 0 & 1 & \\
\hline
1 & 2 & 1 & 1 & 0 & 0 & \\
0 & 1 & -1 & 0 & 1 & 0 & C_3 - C_1 \\
0 & 0 & 1 & -1 & -1 & 1 & \\
\hline
1 & 2 & 0 & 1 & 0 & -1 & \\
0 & 1 & -1 & 0 & 1 & 0 & C_2 - 2C_1 \\
0 & 0 & 1 & -1 & -1 & 2 & \\
\hline
1 & 0 & 0 & 1 & -2 & -1 & \\
0 & 1 & -1 & 0 & 1 & 0 & C_3 + C_2 \\
0 & 0 & 1 & -1 & 1 & 2 & \\
\hline
1 & 0 & 0 & 1 & -2 & -3 & \\
0 & 1 & 0 & 0 & 1 & 1 & \\
0 & 0 & 1 & -1 & 1 & 3 &
\end{array}
$$

Luego la inversa es $A^{-1} = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 1 \\ -1 & 1 & 3 \end{pmatrix}$.

Sea $A = \begin{pmatrix} 2 & 2 & 1 \\ -2 & -1 & -2 \\ 1 & 0 & 1 \end{pmatrix}$. Dejamos como ejercicio el probar que vale $\det(A) = 1$, luego A es invertible. Mostraremos cómo hallar su inversa operando con las filas. También iremos un poco más rápido, haciendo a veces más de una operación por paso.

$$
\begin{array}{|ccc|ccc|c}
2 & 2 & 1 & 1 & 0 & 0 & \\
-2 & -1 & -2 & 0 & 1 & 0 & F_1 \leftrightarrow F_3 \\
1 & 0 & 1 & 0 & 0 & 1 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
-2 & -1 & -2 & 0 & 1 & 0 & F_2 + 2F_1 \text{ y } F_3 - 2F_1 \\
2 & 2 & 1 & 1 & 0 & 0 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
0 & -1 & 0 & 0 & 1 & 2 & F_3 + 2F_2 \\
0 & 2 & -1 & 1 & 0 & -2 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
0 & -1 & 0 & 0 & 1 & 2 & (-1)F_2 \text{ y } (-1)F_3 \\
0 & 0 & -1 & 1 & 2 & 2 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
0 & 1 & 0 & 0 & -1 & -2 & F_1 - F_3 \\
0 & 0 & 1 & -1 & -2 & -2 & \\
\hline
1 & 0 & 0 & 1 & 2 & 3 & \\
0 & 1 & 0 & 0 & -1 & -2 & F_1 - F_3 \\
0 & 0 & 1 & -1 & -2 & -2 &
\end{array}
$$

Luego la inversa es $A^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ -1 & -2 & -2 \end{pmatrix}$.

Sea $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 3 \end{pmatrix}$. Dejamos como ejercicio el probar que vale $\det(A) = 6$, luego A es invertible. Trabajaremos con columnas.

$$\begin{array}{ccc|ccc|c}
3 & 1 & 1 & 1 & 0 & 0 & \\
2 & 3 & 2 & 0 & 1 & 0 & C_1 - C_3 \text{ y } C_2 - C_3 \\
3 & 3 & 3 & 0 & 0 & 1 & \\
\hline
2 & 0 & 1 & 1 & 0 & 0 & \\
0 & 1 & 2 & 0 & 1 & 0 & \frac{1}{2}C_1 \\
0 & 0 & 3 & -1 & -1 & 1 & \\
\hline
1 & 0 & 1 & 1/2 & 0 & 0 & \\
0 & 1 & 2 & 0 & 1 & 0 & C_3 - C_1 \\
0 & 0 & 3 & -1/2 & -1 & 1 & \\
\hline
1 & 0 & 0 & 1/2 & 0 & -1/2 & \\
0 & 1 & 2 & 0 & 1 & 0 & C_3 - 2C_2 \\
0 & 0 & 3 & -1/2 & -1 & 3/2 & \\
\hline
1 & 0 & 0 & 1/2 & 0 & -1/6 & \\
0 & 1 & 0 & 0 & 1 & -2/3 & \frac{1}{3}C_3 \\
0 & 0 & 1 & -1/2 & -1 & 7/6 & \\
\hline
\end{array}$$

Luego la inversa es $A^{-1} = \begin{pmatrix} 1/2 & 0 & -1/6 \\ 0 & 1 & -2/3 \\ -1/2 & -1 & 7/6 \end{pmatrix}$.

Observaciones 3.2.22.

1. Notar que en los ejemplos anteriores lo que hicimos primero fue transformar la matriz A en una matriz triangular (triangular superior en todos los casos, pero podría ser también triangular inferior) y luego llevarla a una forma diagonal. Es importante verificar siempre que la matriz hallada es realmente la matriz inversa (haciendo el producto), dado que es muy fácil cometer algún error de cuentas.

2. Conviene remarcar que cuando aplicamos el algoritmo anterior, tenemos que elegir trabajar con filas o columnas, pero no podemos mezclar. Es decir, si trabajamos con filas y columnas, cuando a la izquierda lleguemos a la matriz identidad, la matriz de la derecha no va a ser necesariamente la matriz inversa.

3. La justificación de este algoritmo requiere bastante trabajo y no la vamos a hacer (ver la sección 3.2 de [F]).

3.3. Método de Cramer

El método de Cramer es una técnica para resolver sistemas de ecuaciones cuadrados. Consideremos un sistema

$$\left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{n1}x_1 + \cdots + a_{nn}x_n & = & b_n \end{array} \right. . \tag{3.6}$$

Sean

$$\Delta = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}; \quad \Delta_i = \begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix}, \ i = 1, \ldots, n.$$

Notar que Δ_i es el determinante obtenido sustituyendo en Δ la columna i por la columna de resultados[5].

Teorema 3.3.1 (Cramer). *Consideremos el sistema* (3.6) *y los determinantes* $\Delta, \Delta_1, \ldots, \Delta_n$ *definidos arriba. Luego,*

[5] No confundir los determinantes Δ_i con los cofactores Δ_{ij}, que no tienen nada que ver aunque se escriben parecido.

1. *si* $\Delta \neq 0$*, entonces el sistema* (3.6) *es compatible determinado y la solución es* $x_i = \frac{\Delta_i}{\Delta}$*, para todo* $i = 1, \ldots, n$*;*

2. *si* $\Delta = 0$ *y existe algún* $i = 1, \ldots, n$ *tal que* $\Delta_i \neq 0$*, entonces el sistema* (3.6) *es incompatible;*

3. *si* $\Delta = \Delta_1 = \cdots = \Delta_n = 0$*, entonces el sistema* (3.6) *no es determinado; puede ser incompatible o compatible indeterminado.*

Dem. Esta prueba la veremos más adelante, en la sección 5.3. □

Observación 3.3.2. La aplicación del teorema anterior para resolver un sistema de ecuaciones es lo que se conoce como el *método de Cramer*. Cuando se aplica este método y resulta $\Delta = \Delta_1 = \cdots = \Delta_n = 0$, entonces el sistema es incompatible o compatible indeterminado, pero en general el método no permite distinguir entre estos dos casos y hay que aplicar escalerización para determinarlo. El único caso en que sí discrimina es cuando el sistema es homogéneo[6] (que siempre es compatible). En ese caso el método de Cramer nos dice que un sistema homogéneo cuadrado admite soluciones no triviales si y solo si $\Delta = 0$.

Ejemplos 3.3.3.

1. Consideremos el sistema $\begin{cases} 3x - y + z &= 2 \\ 2x + 3y + 2z &= 0 \\ 4x - y + 2z &= 2 \end{cases}$. Es

$$\Delta = \begin{vmatrix} 3 & -1 & 1 \\ 2 & 3 & 2 \\ 4 & -1 & 2 \end{vmatrix} = 6; \quad \Delta_1 = \begin{vmatrix} 2 & -1 & 1 \\ 0 & 3 & 2 \\ 2 & -1 & 2 \end{vmatrix} = 6; \quad \Delta_2 = \begin{vmatrix} 3 & 2 & 1 \\ 2 & 0 & 2 \\ 4 & 2 & 2 \end{vmatrix} = 0; \quad \Delta_3 = \begin{vmatrix} 3 & -1 & 2 \\ 2 & 3 & 0 \\ 4 & -1 & 2 \end{vmatrix} = -6.$$

   Como $\Delta = 6 \neq 0$, el sistema es compatible determinado. Usando la fórmula $x_i = \frac{\Delta_i}{\Delta}$, deducimos que la solución es $x = 1$, $y = 0$ y $z = -1$.

2. Consideremos el sistema $\begin{cases} x - y + 2z &= -1 \\ 3x + 2y - z &= 1 \\ 4x + y + z &= 1 \end{cases}$. Es

$$\Delta = \begin{vmatrix} 1 & -1 & 2 \\ 3 & 2 & -1 \\ 4 & 1 & 1 \end{vmatrix} = 0; \quad \Delta_1 = \begin{vmatrix} -1 & -1 & 2 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{vmatrix} = -3.$$

   Luego el sistema es incompatible.

3. Consideremos los siguientes sistemas $\begin{cases} x + y + z &= 1 \\ x + y + z &= 1 \\ x + y + z &= 1 \end{cases}$ y $\begin{cases} x + y + z &= 1 \\ x + y + z &= 0 \\ x + y + z &= -1 \end{cases}$. En ambos casos es $\Delta = \Delta_1 = \Delta_2 = \Delta_3 = 0$, pero es claro que el primer sistema es compatible indeterminado y el segundo es incompatible. Esto ejemplifica la última afirmación del método de Cramer.

Observación 3.3.4. El método de Cramer no es muy eficiente como forma de resolver un sistema de ecuaciones, ya que en general es más rápido escalerizar que calcular determinantes. Sin embargo tiene algunas ventajas para fines teóricos y cuando se quiere estudiar un sistema de ecuaciones que involucra parámetros, ya que "limpia" los cálculos. En la sección siguiente veremos también que el método de Cramer es útil para el estudio de la dependencia lineal en espacios vectoriales.

[6] Recordar que un sistema es homogéneo si todas sus ecuaciones aparecen igualadas a cero. Ver la sección 1.2.

Capítulo 4

Espacios vectoriales

Los espacios vectoriales generalizan el plano y el espacio, y permiten trabajar en mayores dimensiones. Es una estructura muy general, que tiene aplicaciones en la mayoría de las áreas de la matemática.

4.1. Definiciones y propiedades básicas

Un *espacio vectorial* es un conjunto V en el cual están definidas dos operaciones

$$\begin{array}{c} V \times V \xrightarrow{+} V \\ (u,v) \mapsto u+v \end{array}, \qquad \begin{array}{c} \mathbb{R} \times V \xrightarrow{\cdot} V \\ (a,v) \mapsto a \cdot v \end{array}$$

que llamamos *suma* y *producto por escalar*, que verifican las siguientes propiedades:

1. $u+v=v+u$, para todo $u,v \in V$ (conmutativa);
2. $u+(v+w)=(u+v)+w$, para todo $u,v,w \in V$ (asociativa);
3. Existe un elemento $o \in V$ tal que $v+o=o+v=v$, para todo $v \in V$ (existencia de neutro);
4. Para cada $v \in V$ existe un elemento $u \in V$, tal que $v+u=u+v=o$, para todo $v \in V$ (existencia de opuesto);
5. $1 \cdot u = u$, para todo $u \in V$;
6. $a \cdot (u+v) = a \cdot u + a \cdot v$, para todo $a \in \mathbb{R}$, $u,v \in V$;
7. $(a+b) \cdot v = a \cdot v + b \cdot v$, para todo $a,b \in \mathbb{R}$, $v \in V$;
8. $a \cdot (b \cdot v) = (ab) \cdot v$, para todo $a,b \in \mathbb{R}$, $v \in V$.

A los elementos de V les llamamos *vectores* y en general los escribimos con las letras $u,v,w,\dots$, mientras que a los de $\mathbb{R}$ les llamamos *escalares* y los escribimos con las letras $a,b,c,\dots$.

Ejemplos 4.1.1. 1. El plano $\mathbb{R}^2$, el espacio $\mathbb{R}^3$, en general el espacio $\mathbb{R}^n$, $n \geq 1$, y las matrices $M_{m\times n}$, son espacios vectoriales con las operaciones vistas anteriormente.

2. El conjunto de los números reales $\mathbb{R}$ es un espacio vectorial, con su suma y producto habituales.

3. Si $\emptyset \neq D \subset \mathbb{R}$ es un subconjunto, entonces el conjunto $\mathrm{Fun}(D,\mathbb{R})$ de las funciones que van de D a $\mathbb{R}$ es un espacio vectorial, definiendo la suma y producto por escalar en la forma usual: si $f,g : D \to \mathbb{R}$ son funciones y $a \in \mathbb{R}$, entonces $f+g : D \to \mathbb{R}$ y $a \cdot f : D \to \mathbb{R}$ son las funciones definidas por

$$(f+g)(x) := f(x)+g(x), \qquad (a \cdot f)(x) := af(x),$$

para todo $x \in \mathbb{R}$.

4. El conjunto de las sucesiones de números reales con la suma y producto por escalar usuales, tiene estructura de espacio vectorial. Recordar que si $x = (x_n)_{n\in\mathbb{N}}$ e $y = (y_n)_{n\in\mathbb{N}}$ son dos sucesiones de números reales y $a \in \mathbb{R}$, entonces las operaciones anteriores están definidas por $x + y := (x_n + y_n)_{n\in\mathbb{N}}$ y $a \cdot x := (ax_n)_{n\in\mathbb{N}}$.

Proposición 4.1.2. *En un espacio vectorial valen siempre las siguientes propiedades.*

1. *Existe un único elemento $o \in V$ que verifica la propiedad 3 y se le llama el* vector nulo *de V.*

2. *Para cada $v \in V$ existe un único elemento en V que verifica la propiedad 4 y se le llama el* opuesto *de v y lo escribimos $-v$. El opuesto queda caracterizado por verificar $v + (-v) = (-v) + v = o$.*

3. *Si $u, v, w \in V$ verifican $u + v = u + w$, entonces $v = w$ (cancelativa de la suma).*

4. *Para todo $v \in V$ vale $0 \cdot v = o$.*

5. *Para todo $a \in \mathbb{R}$ vale $a \cdot o = o$.*

6. *Si $a \in \mathbb{R}$ y $u, v \in V$ son tales que $a \cdot u = v$ con $a \neq 0$, entonces $u = a^{-1} \cdot v$.*

7. *Si $a \in \mathbb{R}$ y $v \in V$ verifican $a \cdot v = o$, entonces $a = 0$ o $v = o$.* □

Se define la *resta* de dos vectores u y v mediante $u - v := u + (-v)$.

Proposición 4.1.3. *La resta verifica las siguientes propiedades.*

1. *Vale $u = v - w$ si y solo si $u + w = v$, para todo $u, v, w \in V$.*

2. *Vale $-v = (-1) \cdot v$, para todo $v \in V$.*

3. *Vale $a \cdot (u - v) = a \cdot u - a \cdot v$, para todo $a \in \mathbb{R}$, $u, v \in V$.* □

Nota: Por comodidad de notación, de ahora en adelante al producto por escalar lo vamos a escribir av en vez de $a \cdot v$. También a menudo abreviaremos "espacio vectorial" escribiendo solo "espacio".

En lo que sigue V es siempre un espacio vectorial.

4.2. Subespacios

Definición 4.2.1. Un *subespacio* de V es un subconjunto $W \subset V$ que verifica las siguientes propiedades:

$$o \in W; \qquad w_1, w_2 \in W \Rightarrow w_1 + w_2 \in W; \qquad a \in \mathbb{R},\ w \in W \Rightarrow aw \in W.$$

La siguiente proposición da una forma más rápida de chequear la condición de subespacio.

Proposición 4.2.2. *Sea W un subconjunto de un espacio vectorial V. Entonces W es un subespacio de V si y solo si se verifica*

$$0 \in W; \qquad w_1, w_2 \in W,\ a \in \mathbb{R} \Rightarrow w_1 + aw_2 \in W. \quad \square$$

Ejemplos 4.2.3. 1. En todo espacio V, los subconjuntos $\{o\}$ y V son subespacios. Todo subespacio de V distinto de ellos se dice que es *propio*. El subespacio *trivial* es $\{o\}$.

2. En $\mathbb{R}^2$, toda recta que pasa por el origen, es decir todo conjunto de la forma $W = \{(x, y) \in \mathbb{R}^2 : ax + by = 0\}$ con a y b no simultáneamente nulos, es un subespacio.

3. En $\mathbb{R}^3$, toda recta y todo plano que pasen por el origen, son subespacios.

4. Las matrices triangulares superiores y las triangulares inferiores son subespacios del espacio de las matrices cuadradas.

5. Recordar que un *polinomio* es una función $p : \mathbb{R} \to \mathbb{R}$ de la forma $p(x) = a_0 + a_1x + \cdots + a_nx^n$, en que $a_0, a_1 \ldots, a_n$ son números reales fijos. Si p es un polinomio de la forma anterior con $a_n \neq 0$, entonces decimos que p tiene *grado* n. Al conjunto de todos los polinomios lo escribimos $\mathbb{R}[x]$ y al conjunto formado por el polinomio nulo y todos los polinomios de grado menor o igual a n lo escribimos $\mathbb{R}_n[x]$. Así $\mathbb{R}_n[x] \subset \mathbb{R}[x] \subset \mathrm{Fun}(\mathbb{R}, \mathbb{R})$ son subespacios, para todo $n = 0, 1, \ldots$.

6. Si I es un intervalo abierto de $\mathbb{R}$ y escribimos

$$C(I) = \{f : I \to \mathbb{R} : \; f \text{ es continua}\} \qquad \text{y} \qquad D(I) = \{f : I \to \mathbb{R} : \; f \text{ es derivable}\},$$

entonces $D(I) \subset C(I) \subset \mathrm{Fun}(I, \mathbb{R})$ son subespacios, para todo intervalo abierto I.

7. El espacio vectorial $\mathbb{R}$ no tiene subespacios propios.

8. Los siguientes subconjuntos de $\mathbb{R}^2$ **no son** subespacios:

$$W_1 = \{(x, y) : \; x + y = 1\}; \qquad W_2 = \{(x, y) : \; xy = 0\}; \qquad W_3 = \{(x, y) : \; x \geq 0 \text{ e } y \geq 0\}.$$

El interés en los subespacios viene dado por la siguiente proposición.

Proposición 4.2.4. *Si $W \subset V$ es un subespacio, entonces W es un espacio vectorial con las operaciones de V restringidas a W.* □

Luego todos los subespacios que vimos anteriormente son también ejemplos de espacios vectoriales.

Definición 4.2.5. Una *combinación lineal* de un conjunto finito $v_1, \ldots, v_n$ de elementos de V, es un vector de la forma $a_1v_1 + \cdots + a_nv_n$, siendo $a_1, \ldots, a_n$ escalares arbitrarios.

Dados $v_1, \ldots, v_n \in V$, consideremos el conjunto

$$[v_1, \ldots, v_n] := \{a_1v_1 + \cdots + a_nv_n : \; a_1, \ldots, a_n \in \mathbb{R}\}.$$

es decir, $[v_1, \ldots, v_n]$ es el conjunto de todas las combinaciones lineales de $v_1, \ldots, v_n$.

Proposición 4.2.6. *Para todo $v_1, \ldots, v_n \in V$ se cumple que $[v_1, \ldots, v_n]$ es un subespacio de V. Además, todo subespacio W que contenga a $v_1, \ldots, v_n$, contiene también a $[v_1, \ldots, v_n]$. Luego $[v_1, \ldots, v_n]$ es el "menor" subespacio de V que contiene a $v_1, \ldots, v_n$.* □

Definición 4.2.7. El conjunto $[v_1, \ldots, v_n]$ se llama el *subespacio generado* por los vectores $v_1, \ldots, v_n$ o también el subespacio generado por el conjunto $\{v_1, \ldots, v_n\}$.

Ejemplos 4.2.8.

1. Si $v \in \mathbb{R}^3$ es un vector no nulo, entonces $[v] = \{tv : \; t \in \mathbb{R}\}$ es la recta por el origen que pasa por v.

2. Si $u, v \in \mathbb{R}^3$ son dos vectores no colineales, entonces $[u, v] = \{su + tv : \; s, t \in \mathbb{R}\}$ es el plano por el origen que pasa por u y v (o paralelo a u y v).

4.3. Dependencia lineal

Definición 4.3.1. Sean $v_1, \ldots, v_n$ elementos de V. Decimos que $v_1, \ldots, v_n$ son *linealmente dependientes* y que el conjunto $\{v_1, \ldots, v_n\}$ es *linealmente dependiente* (LD) si existen escalares no todos nulos $a_1, \ldots, a_n$ tales que $a_1v_1 + \cdots + a_nv_n = o$. En caso contrario decimos que $v_1, \ldots, v_n$ son *linealmente independientes* y que el conjunto $\{v_1, \ldots, v_n\}$ es *linealmente independiente* (LI).

La combinación lineal *trivial* de $v_1, \ldots, v_n$ es

$$0v_1 + \cdots + 0v_n = o.$$

Luego $\{v_1, \ldots, v_n\}$ es LI si y solo si la única combinación lineal de $v_1, \ldots, v_n$ que da el vector nulo es la trivial, y es LD si y solo si existe una combinación lineal no trivial de $v_1, \ldots, v_n$ que da el vector nulo.

Observaciones 4.3.2. 1. Si en un conjunto $X = \{v_1, \dots, v_n\}$ se cumple que existe algún i tal que $v_i = o$, entonces X es LD.

2. El conjunto $\{v\}$ es LD si y solo si $v = o$.

3. El conjunto $\{u, v\}$ es LD si y solo si existe un escalar a tal que $u = av$ o $v = au$ (esto es un caso particular del siguiente resultado).

Proposición 4.3.3. *Un conjunto $\{v_1, \dots, v_n\}$ es LD ($n \geq 2$) si y solo si existe algún vector v_i que es combinación lineal de los restantes $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$.* □

Ejemplos 4.3.4. 1. Consideremos el conjunto $X = \{(-4, 1),\ (1, 1),\ (1, -1)\} \subset \mathbb{R}^2$. Para estudiar su dependencia lineal, tenemos que ver si la única solución de la ecuación

$$a(1, -1) + b(-4, 1) + c(1, 1) = (0, 0) \quad \Leftrightarrow \quad \left\{ \begin{array}{rcl} a - 4b + c & = & 0 \\ -a + b + c & = & 0 \end{array} \right.$$

es la trivial. Operando, obtenemos que la solución del sistema es $a = \frac{5}{2}b$, $c = \frac{3}{2}b$ y b queda libre. Luego el sistema es compatible indeterminado y por lo tanto X es LD. Observar que tomando $b = 2$ obtenemos $(a, b, c) = (5, 2, 3)$. Luego vale

$$5(1, -1) + 2(-4, 1) + 3(1, 1) = (0, 0).$$

A partir de esta fórmula vemos que podemos escribir cualquier vector de X como combinación de los restantes; por ejemplo

$$(1, -1) = \frac{-2}{5}(-4, 1) + \frac{-3}{5}(1, 1).$$

2. Sea $X = \{(1, 1, 1),\ (1, 2, 1),\ (2, 3, 3)\} \subset \mathbb{R}^3$. Para estudiar su dependencia lineal, planteamos la ecuación

$$a(1, 1, 1) + b(1, 2, 1) + c(2, 3, 3) = (0, 0, 0). \tag{4.1}$$

El conjunto X es LI si y solo si la única solución de (4.1) es la trivial $a = b = c = 0$. La ecuación (4.1) equivale al sistema homogéneo

$$\left\{ \begin{array}{rcl} a + b + 2c & = & 0 \\ a + 2b + 3c & = & 0 \\ a + b + 3c & = & 0 \end{array} \right. .$$

Calculando el determinante de la matriz asociada al sistema obtenemos $\begin{vmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 3 \end{vmatrix} = 1$. Luego el teorema de Cramer implica que la única solución que admite el sistema es la trivial, y por lo tanto X es LI.

3. Sea $X = \{(2, 1, 3),\ (-1, 3, 2),\ (1, 11, 12)\} \subset \mathbb{R}^3$. Razonando como en el caso anterior, planteamos la ecuación

$$a(2, 1, 3) + b(-1, 3, 2) + c(1, 11, 12) = (0, 0, 0), \tag{4.2}$$

y X es LI si y solo si la única solución de (4.2) es la trivial. La ecuación (4.2) equivale al sistema homogéneo

$$\left\{ \begin{array}{rcl} 2a - b + c & = & 0 \\ a + 3b + 11c & = & 0 \\ 3a + 2b + 12c & = & 0 \end{array} \right. .$$

Calculando el determinante de la matriz asociada al sistema obtenemos $\begin{vmatrix} 2 & -1 & 1 \\ 1 & 3 & 11 \\ 3 & 2 & 12 \end{vmatrix} = 0$. Luego el teorema de Cramer implica que el sistema admite alguna solución no trivial, y por lo tanto X es LD. Para obtener la combinación lineal no trivial hay que resolver el sistema. Realizando los cálculos se obtiene que una solución es

$$2(2, 1, 3) + 3(-1, 3, 2) - (1, 11, 12) = (0, 0, 0),$$

4. Sea $X = \left\{1,\ 1+x,\ 1+x+x^2\right\} \subset \mathbb{R}_2[x]$. Si escribimos la combinación lineal

$$a1 + b(1+x) + c\left(1+x+x^2\right) = 0,$$

entonces reordenando obtenemos

$$a+b+c+(b+c)x+cx^2 = 0 \quad \Rightarrow \quad \left\{\begin{array}{rcl} a+b+c & = & 0 \\ b+c & = & 0 \\ c & = & 0 \end{array}\right. .$$

Claramente la única solución del sistema es la trivial, por lo tanto X es LI.

5. Es fácil de probar que la base canónica $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ de $\mathbb{R}^3$ es un conjunto LI.

6. Generalizando el caso anterior, también es fácil de probar que el conjunto $\{e_1, e_2, \cdots, e_n\} \subset \mathbb{R}^n$ es LI, siendo
$$e_1 = (1,0,\cdots,0),\ e_2 = (0,1,\cdots,0),\ \cdots,\ e_n = (0,0,\cdots,1).$$

Proposición 4.3.5. *Sean $A = \{v_1,\cdots,v_n\}$ y $B = \{v_1,\cdots,v_n,v_{n+1},\cdots,v_m\}$, con $m \geq n$. Entonces:*

1. Si A es LD, entonces B es LD.

2. Si B es LI, entonces A es LI. □

El siguiente resultado nos da una condición para "agrandar" un conjunto LI y que siga siendo LI.

Proposición 4.3.6. *Si $\{v_1,\cdots,v_n\}$ es LI y $v_{n+1} \notin [v_1,\ldots,v_n]$, entonces $\{v_1,\cdots,v_n,v_{n+1}\}$ es LI.*

Dem. Sean $a_1,\ldots,a_{n+1}$ escalares tales que $a_1v_1 + \cdots + a_nv_n + a_{n+1}v_{n+1} = o$. Si fuese $a_{n+1} \neq 0$, entonces podríamos despejar v_{n+1} de la igualdad anterior y obtendríamos que v_{n+1} es combinación lineal de $v_1,\cdots,v_n$ en contra de nuestras hipótesis. Luego es $a_{n+1} = 0$ y por lo tanto vale $a_1v_1 + \cdots + a_nv_n = o$. Como $\{v_1,\cdots,v_n\}$ es LI, deducimos que es $a_1 = \cdots = a_n = 0$. Luego $a_1,\ldots,a_n,a_{n+1}$ son todos nulos. □

4.4. Generadores y bases

Definición 4.4.1. Sea $\mathcal{B} = \{v_1,\ldots,v_n\}$ un subconjunto de V. Decimos que $\mathcal{B}$ es un *conjunto generador* de V si todo vector de V se puede escribir como combinación lineal de $v_1,\ldots,v_n$, es decir si $[v_1,\ldots,v_n] = V$. Si además $\mathcal{B}$ verifica que todo vector de V se escribe en forma única como combinación lineal de $v_1,\ldots,v_n$, entonces decimos que $\mathcal{B}$ es una *base* de V.

Ejemplo 4.4.2. Consideremos los siguientes subconjuntos del espacio $\mathbb{R}^2$:

$$\mathcal{B}_1 = \{(2,-4),\ (-3,6)\},\quad \mathcal{B}_2 = \{(1,2)\},\quad \mathcal{B}_3 = \{(1,2),\ (1,1),\ (-3,5)\},\quad \mathcal{B}_4 = \{(1,2),\ (2,1)\}.$$

Los conjuntos $\mathcal{B}_1$ y $\mathcal{B}_2$ no son generadores, $\mathcal{B}_3$ es generador pero no es base y $\mathcal{B}_4$ es base.

Proposición 4.4.3. *Un conjunto $\mathcal{B}$ es una base de V si y solo si $\mathcal{B}$ es un conjunto generador y es LI.* □

Ejemplo 4.4.4. Consideremos el espacio vectorial $V = \left\{(x,y,z) \in \mathbb{R}^3 :\ 2x+y+z = 0\right\}$ (es un plano que pasa por el origen de $\mathbb{R}^3$). Si $(x,y,z) \in V$, entonces $2x+y+z = 0$ y por lo tanto $z = -2x-y$. Luego podemos escribir

$$(x,y,z) = (x,y,-2x-y) = (x,0,-2x) + (0,y,-y) = x(1,0,-2) + y(0,1,-1).$$

Esto nos dice que el conjunto $\mathcal{B} = \{(1,0,-2),\ (0,1,-1)\}$ es un generador de V (notar que $(1,0,-2)$ y $(0,1,-1)$ pertenecen a V, como se deduce tomando $(x,y) = (1,0)$ o $(x,y) = (0,1)$, respectivamente). Además es fácil de probar que $\mathcal{B}$ es LI, luego $\mathcal{B}$ es una base del plano V.

En algunos espacios en que trabajamos habitualmente, hay ciertas bases estándar a las que se les suele llamar la *base canónica* del espacio. En los siguientes ejemplos se describen estas bases.

Ejemplos 4.4.5. 1. La base canónica de $\mathbb{R}^n$ es el conjunto $\mathcal{B} = \{e_1, e_2, \cdots, e_n\}$ definido en el ejemplo 4.3.4.6. Notar que las bases canónicas $\{\mathbf{i}, \mathbf{j}\}$ de $\mathbb{R}^2$ y $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ de $\mathbb{R}^3$, son casos particulares de esta construcción.

2. La base canónica de $\mathbb{R}_n[x]$ es el conjunto $\{1, x, \ldots, x^n\}$.

3. La base canónica de $M_{m\times n}$ es el conjunto $\{e_{ij} :\ i = 1, \ldots, m,\ j = 1, \ldots, n\}$, donde e_{ij} es la matriz $m \times n$ que tiene todas sus entradas nulas, salvo la entrada (i, j) que vale 1. Por ejemplo, la base canónica de $M_2 = M_{2\times 2}$ es $\mathcal{B} = \{e_{11},\ e_{12},\ e_{21},\ e_{22}\}$, siendo

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Teorema 4.4.6. *En un espacio vectorial, la cantidad de elementos de un conjunto LI siempre es menor o igual que la cantidad de elementos de un conjunto generador.*

Dem. Sea $\mathcal{A} = \{v_1, \ldots, v_n\}$ un conjunto LI y $\mathcal{G} = \{w_1, \ldots, w_m\}$ un conjunto generador del espacio V. Queremos probar que es $n \leq m$.

Como $v_1 \in V$ y $\mathcal{G}$ es generador de V, entonces existen $a_1, \ldots, a_m \in \mathbb{R}$ tales que $v_1 = a_1w_1 + \cdots + a_mw_m$. Si fuesen $a_1 = \cdots = a_m = 0$, entonces sería $v_1 = o$, pero esto es imposible porque $v_1 \in \mathcal{A}$ y $\mathcal{A}$ es LI. Luego alguno de los escalares es no nulo. Eventualmente reordenando los elementos de $\mathcal{G}$, podemos suponer que es $a_1 \neq 0$. Luego

$$v_1 = a_1w_1 + a_2w_2 + \cdots + a_mw_m \quad \Rightarrow \quad w_1 = \frac{1}{a_1}v_1 + \left(\frac{-a_2}{a_1}\right)w_2 + \cdots + \left(\frac{-a_m}{a_1}\right)w_m.$$

Veremos que $\mathcal{G}_1 := \{v_1, w_2, \ldots, w_m\}$ es un conjunto generador de V. Sea $v \in V$ un vector arbitrario. Como $\mathcal{G}$ es un conjunto generador de V, entonces existen $b_1, \ldots, b_m \in \mathbb{R}$ tales que $v = b_1w_1 + \cdots + b_mw_m$. Luego

$$\begin{aligned} v &= b_1\left(\frac{1}{a_1}v_1 + \left(\frac{-a_2}{a_1}\right)w_2 + \cdots + \left(\frac{-a_m}{a_1}\right)w_m\right) + b_2w_2 + \cdots + b_mw_m \\ &= \frac{b_1}{a_1}v_1 + \left(b_2 - \frac{b_1a_2}{a_1}\right)w_2 + \cdots + \left(b_m - \frac{b_1a_m}{a_1}\right)w_m. \end{aligned}$$

Esto muestra que todo vector de V se puede escribir como combinación lineal de $\mathcal{G}_1$.

La prueba sigue por inducción. Partimos de que existe $k \geq 1$ tal que $\mathcal{G}_k := \{v_1, \ldots, v_k, w_{k+1}, \ldots, w_m\}$ es un conjunto generador de V. Como $v_{k+1} \in V$ y $\mathcal{G}_k$ es un conjunto generador, entonces existen $c_1, \cdots, c_m \in \mathbb{R}$ tales que

$$v_{k+1} = c_1v_1 + \cdots + c_kv_k + c_{k+1}w_{k+1} + \cdots + c_mw_m. \tag{4.3}$$

Si fuese $c_{k+1} = \cdots = c_m = 0$, entonces sería $v_{k+1} = c_1v_1 + \cdots + c_kv_k$, contradiciendo que $\mathcal{A}$ es LI. Luego existe algún i, con $k + 1 \leq i \leq m$ tal que $c_i \neq 0$. Como antes podemos suponer que es $c_{k+1} \neq 0$, y por lo tanto podemos despejar w_{k+1} de (4.3) obteniendo

$$w_{k+1} = \left(\frac{-c_1}{c_{k+1}}\right)v_1 + \cdots + \left(\frac{-c_k}{c_{k+1}}\right)v_k + \frac{1}{c_{k+1}}v_{k+1} + \left(\frac{-c_{k+2}}{c_{k+1}}\right)w_{k+2} + \cdots + \left(\frac{-c_m}{c_{k+1}}\right)w_m. \tag{4.4}$$

Veremos que $\mathcal{G}_{k+1} := \{v_1, \ldots, v_k, v_{k+1}, w_{k+2}, \ldots, w_m\}$ es un conjunto generador de V. La idea de la prueba es la misma que antes. Para simplificar la notación, la fórmula (4.4) la escribiremos

$$w_{k+1} = d_1v_1 + \cdots + d_kv_k + d_{k+1}v_{k+1} + d_{k+2}w_{k+2} + \cdots + d_mw_m.$$

Sea $v \in V$ un vector arbitrario. Como $\mathcal{G}_k$ es un conjunto generador de V, entonces existen $x_1, \ldots, x_m \in \mathbb{R}$ tales que $v = x_1v_1 + \cdots + x_kv_k + x_{k+1}w_{k+1} + \cdots + x_mw_m$. Luego

$$\begin{aligned} v &= x_1v_1 + \cdots + x_kv_k + x_{k+1}\left(d_1v_1 + \cdots + d_kv_k + d_{k+1}v_{k+1} + d_{k+2}w_{k+2} + \cdots + d_mw_m\right) + \cdots + x_mw_m \\ &= \left(x_1 + x_{k+1}d_1\right)v_1 + \cdots + \left(x_k + x_{k+1}d_k\right)v_k + x_{k+1}d_{k+1}v_{k+1} + \left(x_{k+1}d_{k+2} + x_{k+2}\right)w_{k+2} + \cdots \\ &\quad + \left(x_{k+1}d_m + x_m\right)w_m. \end{aligned}$$

Esto muestra que todo vector de V se puede escribir como combinación lineal de $\mathcal{G}_{k+1}$.

Luego probamos que para cada $k = 1, 2, \dots$ se cumple (eventualmente reordenando los elementos de $\mathcal{G}$) que $\mathcal{G}_k = \{v_1, \dots, v_k, w_{k+1}, \dots, w_m\}$ es un conjunto generador de V. Si fuese $n > m$, entonces podríamos sustituir todos los elementos de $\mathcal{G}$ por elementos de $\mathcal{A}$ para obtener que el conjunto $\mathcal{G}_m = \{v_1, \dots, v_m\}$ genera a V, siendo $\mathcal{A} = \{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$. Pero esto es imposible porque en ese caso el elemento $v_{m+1} \in \mathcal{A}$ sería combinación lineal de $v_1, \dots, v_m$, contradiciendo que $\mathcal{A}$ es LI. Luego necesariamente es $n \leq m$. $\square$

Corolario 4.4.7. *Si $\mathcal{B}_1 = \{v_1, \dots, v_n\}$ y $\mathcal{B}_2 = \{w_1, \dots, w_m\}$ son dos bases de V, entonces $n = m$.* $\square$

Definición 4.4.8. Si un espacio vectorial V admite una base $\mathcal{B} = \{v_1, \dots, v_n\}$, entonces diremos que V tiene *dimensión finita* y que la *dimensión* de V es n, lo cual escribimos $\dim(V) = n$. En caso contrario diremos que V es de *dimensión infinita.* Si V es el espacio trivial $V = \{o\}$, entonces definimos $\dim V = 0$.

En los ejemplos que siguen usamos las notaciones de los ejemplos 4.4.4 y 4.4.5.

Ejemplos 4.4.9. Espacios de dimensión finita.

1. El conjunto $\{e_1, e_2, \cdots, e_n\}$ es base de $\mathbb{R}^n$, luego $\dim(\mathbb{R}^n) = n$, para todo n.

2. El conjunto $\{1, \dots, x, x^n\}$ es base de $\mathbb{R}_n[x]$, luego $\dim(\mathbb{R}_n[x]) = n + 1$, para todo n.

3. El conjunto $\{e_{ij} : \ i = 1, \dots, m, \ j = 1, \dots, n\}$ es base de $M_{m\times n}$, luego $\dim(M_{m\times n}) = mn$, para todo m, n.

4. El conjunto $\mathcal{B} = \{(1, 0, -2), \ (0, 1, -1)\}$ es base del plano $V = \big\{(x, y, z) \in \mathbb{R}^3 : \ 2x + y + z = 0\big\}$, luego $\dim(V) = 2$.

Ejemplos 4.4.10. Ejemplos de espacios que tienen dimensión infinita son: los polinomios $\mathbb{R}[x]$, las funciones continuas $C(I)$ y las funciones derivables $D(I)$, I intervalo abierto de $\mathbb{R}$.

Nota: De ahora en adelante asumiremos siempre que los espacios con los que estamos trabajando son de dimensión finita[1].

Proposición 4.4.11. *En un espacio vectorial siempre vale que la cantidad de elementos de un conjunto LI es menor o igual que la dimensión del espacio y que la cantidad de elementos de un conjunto generador es mayor o igual a la dimensión del espacio.* $\square$

Proposición 4.4.12. *Sea V un espacio de dimensión n y $\mathcal{A} = \{v_1, \dots, v_m\}$ un subconjunto LI de V. Sabemos que es $m \leq n$.*

1. *Si $m < n$, entonces existen $v_{m+1}, \dots, v_n \in V$ tales que $\mathcal{B} = \{v_1, \dots, v_m, \ v_{m+1}, \dots, v_n\}$ es base de V.*

2. *Si $m = n$, entonces $\mathcal{A}$ es base de V.*

Dem. (Idea de la prueba.) Si $m < n$ entonces $\mathcal{A}$ no puede ser base de V, y por lo tanto no es un conjunto generador de V. Luego existe un vector $v_{m+1} \in V$ tal que $v_{m+1} \notin [v_1, \dots, v_m]$. Aplicando la proposición 4.3.6 deducimos que $\{v_1, \dots, v_m, \ v_{m+1}\}$ es LI. Si $m + 1 = n$, entonces ya terminamos. Si $m + 1 < n$ repetimos el razonamiento con $\{v_1, \dots, v_m, \ v_{m+1}\}$ en lugar de $\mathcal{A}$ y así seguimos. La misma idea sirve para probar la segunda parte de la tesis. $\square$

Proposición 4.4.13. *Si $\mathcal{G} = \{v_1, \dots, v_m, \ v_{m+1}\}$ es un conjunto generador de V y v_{m+1} es combinación lineal de $v_1, \dots, v_m$, entonces $\mathcal{G}' = \{v_1, \dots, v_m\}$ es también un conjunto generador de V.* $\square$

Proposición 4.4.14. *Sea V un espacio de dimensión n y $\mathcal{G}$ un conjunto generador de V con p elementos.*

1. *Si $p > n$, entonces existen n elementos $v_1, \dots, v_n$ de $\mathcal{G}$, tales que $\mathcal{G}' = \{v_1, \dots, v_n\}$ es base de V.*

2. *Si $p = n$, entonces $\mathcal{G}$ es base de V.*

[1] Va a ser claro que muchas de las definiciones que veremos se aplican también a los espacios de dimensión infinita, pero como los espacios que nos interesan son los de dimensión finita, ponemos esta hipótesis general para no tener que estarla aclarando en los resultados que solo tienen sentido o solo valen para dimensión finita.

Dem. Si $p > n$, entonces $\mathcal{G}$ es LD y por lo tanto tiene algún vector que es combinación lineal de los restantes. Aplicando la proposición anterior obtenemos un generador contenido en $\mathcal{G}$ con un elemento menos. Luego se sigue repitiendo el proceso. Con la misma idea se prueba la segunda parte. $\square$

Ejemplo 4.4.15. El conjunto $\mathcal{G} = \{(1,0),\,(2,0),\,(0,1)\}$ es un generador de $\mathbb{R}^2$. Los conjuntos $\{(1,0),\,(0,1)\}$ y $\{(2,0),\,(0,1)\}$ están contenidos en $\mathcal{G}$ y son bases de $\mathbb{R}^2$. Observar que el conjunto $\{(1,0),\,(2,0)\}$ está contenido en $\mathcal{G}$, pero no es base de $\mathbb{R}^2$.

Corolario 4.4.16. *Sea V un espacio de dimensión n y $\mathcal{A}$ un subconjunto de V que tiene n elementos. Entonces las siguientes afirmaciones son equivalentes:*

- *$\mathcal{A}$ es un conjunto generador de V;*
- *$\mathcal{A}$ es un conjunto LI;*
- *$\mathcal{A}$ es una base de V.* $\square$

Ejemplo 4.4.17. Consideremos el espacio $\mathbb{R}_2[x]$ y el conjunto $\mathcal{A} = \{1,\,1+x,\,1+x+x^2\} \subset \mathbb{R}_2[x]$. Es fácil de probar que $\mathcal{A}$ es LI, y como tiene 3 elementos y $\dim \mathbb{R}_2[x] = 3$, deducimos que $\mathcal{A}$ es base de $\mathbb{R}_2[x]$.

La siguiente proposición nos da un criterio simple para saber si un conjunto es base de $\mathbb{R}^n$.

Proposición 4.4.18. *Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ un subconjunto de $\mathbb{R}^n$ y sea Δ el determinante de la matriz $n \times n$ cuyas columnas son los vectores $v_1, \ldots, v_n$ escritos verticalmente. Si $\Delta \neq 0$, entonces $\mathcal{B}$ es base de $\mathbb{R}^n$.*

Dem. Sean $v_1 = (a_{11}, \ldots, a_{n1}),\ \cdots,\ v_n = (a_{1n}, \ldots, a_{nn})$ y $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$. Si $x_1, \ldots, x_n \in \mathbb{R}$ verifican $x_1 v_1 + \cdots + x_n v_n = o$, entonces es

$$x_1(a_{11}, \ldots, a_{n1}) + \cdots + x_n(a_{1n}, \ldots, a_{nn}) = (0, \ldots, 0) \quad \Leftrightarrow \quad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{n1}x_1 + \cdots + a_{nn}x_n & = & 0 \end{cases}.$$

Como es $\det(A) = \Delta \neq 0$, entonces el método de Cramer implica que el sistema anterior es compatible determinado y por lo tanto la única solución que admite es la trivial $x_1 = \cdots = x_n = 0$. Esto implica que $\mathcal{B}$ es LI, y como tiene n vectores, del corolario anterior se deduce que $\mathcal{B}$ es base de $\mathbb{R}^n$. $\square$

Proposición 4.4.19. *Si V es un espacio de dimensión finita y W es un subespacio de V, entonces W es un espacio de dimensión finita y* $\dim(W) \leq \dim(V)$.

Dem. Si W es el subespacio trivial, el resultado es obvio. Supongamos ahora $W \neq \{o\}$. Sea $\mathcal{A} = \{w_1, \ldots, w_m\}$ un subconjunto LI de W. Entonces $m \leq \dim(V)$. Si $\mathcal{A}$ es generador de W, entonces ya está probado. Si no, entonces existe $w_{m+1} \in W$ que no es combinación lineal de $w_1, \ldots, w_m$ y por lo tanto $\mathcal{A}_1 = \{w_1, \ldots, w_m, w_{m+1}\}$ es también un subconjunto LI de W. Luego seguimos iterando este procedimiento, y es claro que en algún momento se termina. $\square$

Subespacios de $\mathbb{R}^2$ y de $\mathbb{R}^3$. Lo que interesa es determinar los subespacios propios (los otros son obvios).

Subespacios propios de $\mathbb{R}^2$. Como $\dim\left(\mathbb{R}^2\right) = 2$, si W es un subespacio propio de $\mathbb{R}^2$, entonces $\dim(W) = 1$. Luego $W = \{tv :\ t \in \mathbb{R}\}$, con $o \neq v \in \mathbb{R}^2$, es una recta que pasa por el origen.

Subespacios propios de $\mathbb{R}^3$. Como $\dim\left(\mathbb{R}^3\right) = 3$, si W es un subespacio propio de $\mathbb{R}^3$, entonces $\dim(W) = 1, 2$. Si $\dim(W) = 1$, entonces $W = \{tv :\ t \in \mathbb{R}\}$, con $o \neq v \in \mathbb{R}^3$, es una recta que pasa por el origen; si $\dim(W) = 2$, entonces $W = \{tu + sv :\ t \in \mathbb{R}\}$, con $\{u,\,v\} \subset \mathbb{R}^3$ un conjunto LI, es un plano que pasa por el origen.

4.5. Suma directa

Sea V un espacio vectorial y W_1, W_2 dos subespacios de V. Definimos su *intersección* $W_1 \cap W_2$ y su *suma* $W_1 + W_2$, mediante

$$W_1 \cap W_2 := \{v \in V : \ v \in W_1 \text{ y } v \in W_2\}, \qquad W_1 + W_2 := \{w_1 + w_2 : \ w_1 \in W_1, \ w_2 \in W_2\}.$$

Proposición 4.5.1. *Si W_1, W_2 son dos subespacios de V, entonces $W_1 \cap W_2$ y $W_1 + W_2$ son también subespacios de V.* $\square$

Observación 4.5.2. Dados dos subespacios W_1, W_2, su intersección $W_1 \cap W_2$ es el mayor subespacio contenido en W_1 y W_2, mientras que $W_1 + W_2$ es el menor subespacio que contiene a W_1 y W_2.

Ejemplos 4.5.3. Consideremos el espacio $\mathbb{R}^3$.

1. Si $W_1 = \{(x, y, z) : \ z = 0\}$ (plano Oxy) y $W_2 = \{(x, y, z) : \ y = 0\}$ (plano Oxz), entonces

$$W_1 \cap W_2 = \{(x, y, z) : \ y = z = 0\} \text{ (eje Ox)}, \qquad W_1 + W_2 = \mathbb{R}^3.$$

2. Si $W_1 = \{(x, y, z) : \ y = z = 0\}$ (eje Ox) y $W_2 = \{(x, y, z) : \ x = z = 0\}$ (eje Oy), entonces

$$W_1 \cap W_2 = \{(0, 0, 0)\}, \qquad W_1 + W_2 = \{(x, y, z) : \ z = 0\} \text{ (plano Oxy)}.$$

Definición 4.5.4. Sean W_1, W_2 dos subespacios de un espacio V. Decimos que un subespacio U de V es *suma directa* de W_1 y W_2 y escribimos $U = W_1 \oplus W_2$ si se cumple $U = W_1 + W_2$ y $W_1 \cap W_2 = \{o\}$.

Ejemplo 4.5.5. Si consideramos los ejemplos 4.5.3, deducimos del primero que el espacio $\mathbb{R}^3$ es suma del plano Oxy con el plano Oxz, pero esta suma no es directa, y del segundo que el plano Oxy es suma directa de la recta Ox y la recta Oy.

Proposición 4.5.6. *Si W_1, W_2 son dos subespacios de V, entonces $V = W_1 \oplus W_2$ si y solo si todo vector de V se escribe en forma única como suma de un vector de W_1 con uno de W_2.* $\square$

Proposición 4.5.7. *Sean W_1, W_2 dos subespacios de V tales que $V = W_1 \oplus W_2$. Si $\mathcal{B}_1 = \{v_1, \dots, v_n\}$ y $\mathcal{B}_2 = \{w_1, \dots, w_m\}$ son bases respectivas de W_1 y W_2, entonces su unión*

$$\mathcal{B} = \{v_1, \dots, v_n, w_1, \dots, w_m\}$$

es una base de V. $\square$

Corolario 4.5.8. *Sean W_1, W_2 dos subespacios de V tales que $V = W_1 \oplus W_2$. Entonces*

$$\dim V = \dim W_1 + \dim W_2. \quad \square$$

Corolario 4.5.9. *Sean W_1, W_2 dos subespacios de V tales que $W_1 \cap W_2 = \{o\}$.*
Si $\dim V = \dim W_1 + \dim W_2$, entonces $V = W_1 \oplus W_2$. $\square$

Este último corolario nos da un método fácil para probar que un espacio es suma directa de dos subespacios dados.

Ejemplo 4.5.10. Consideremos el espacio $\mathbb{R}^3$ y los subespacios

$$W_1 = \{(x, y, z) : \ 2x + y + z = 0\} \qquad \text{y} \qquad W_2 = \{(x, y, z) : \ y = z = 0\}.$$

Si $(x, y, z) \in W_1 \cap W_2$, entonces vale

$$\begin{cases} 2x + y + z = 0 \\ \qquad y = z = 0 \end{cases} \quad \Rightarrow \quad x = y = z = 0 \quad \Rightarrow \quad (x, y, z) = (0, 0, 0).$$

Luego $W_1 \cap W_2 = \{(0, 0, 0)\}$. Por otro lado W_1 es un plano y W_2 una recta, así que es $\dim W_1 = 2$ y $\dim W_2 = 1$ y por lo tanto $\dim W_1 + \dim W_2 = 2 + 1 = 3 = \dim \mathbb{R}^3$; luego $\mathbb{R}^3 = W_1 \oplus W_2$.

Observación 4.5.11. Dados dos subespacios W_1, W_2 de V, no alcanza con que valga $\dim V = \dim W_1 + \dim W_2$ para que la suma sea directa, hay que probar que también vale $W_1 \cap W_2 = \{o\}$. Por ejemplo, si el espacio es $\mathbb{R}^3$ y consideramos los subespacios

$$W_1 = \{(x, y, z) : \ z = 0\} \qquad \text{y} \qquad W_2 = \{(x, y, z) : \ y = z = 0\},$$

entonces W_1 es un plano y W_2 una recta, y por lo tanto $\dim W_1 + \dim W_2 = 2 + 1 = 3 = \dim \mathbb{R}^3$. Pero W_2 está contenido en W_1, luego $W_1 + W_2 = W_1 \subsetneq \mathbb{R}^3$ y por lo tanto no puede ser $W_1 \oplus W_2 = \mathbb{R}^3$.

Observación 4.5.12. La suma directa de dos subespacios se generaliza a una cantidad finita cualquiera de subespacios (y también para infinitos subespacios), pero esta construcción es un poco más complicada que en el caso en que son solo dos y no la vamos a desarrollar en estas notas. Quien esté interesado en el tema, puede leerlo en las páginas 265-267 de [F].

4.6. Coordenadas y cambio de base

Definición 4.6.1. Sea V un espacio y $\mathcal{B} = \{v_1, \dots, v_n\}$ una base *ordenada* de V, es decir una base para la cual fijamos un orden de sus elementos. Sabemos que si $v \in V$, entonces existen únicos $x_1, \dots, x_n \in \mathbb{R}$ tales que $v = x_1 v_1 + \cdots + x_n v_n$. En este caso escribimos $\operatorname{coord}_{\mathcal{B}}(v) = (x_1, \dots, x_n)$ y decimos que la n-upla $(x_1, \dots, x_n) \in \mathbb{R}^n$ son las *coordenadas* del vector v en la base $\mathcal{B}$.

En lo que sigue de esta sección supondremos siempre que las bases son bases ordenadas.

Ejemplos 4.6.2.

1. El conjunto $\mathcal{B}_1 = \{(1, 1), \ (1, -1)\}$ es una base de $\mathbb{R}^2$. Si $(x, y) \in \mathbb{R}^2$, es

$$(x, y) = \frac{x+y}{2}(1, 1) + \frac{x-y}{2}(1, -1) \quad \Rightarrow \quad \operatorname{coord}_{\mathcal{B}_1}(x, y) = \left(\frac{x+y}{2}, \frac{x-y}{2}\right)$$

2. Si ahora consideramos la base $\mathcal{B}_2 = \{(1, -1), \ (1, 1)\}$ de $\mathbb{R}^2$. Si $(x, y) \in \mathbb{R}^2$, es

$$(x, y) = \frac{x-y}{2}(1, -1) + \frac{x+y}{2}(1, 1) \quad \Rightarrow \quad \operatorname{coord}_{\mathcal{B}_2}(x, y) = \left(\frac{x-y}{2}, \frac{x+y}{2}\right)$$

   Notar que $\mathcal{B}_1$ y $\mathcal{B}_2$ coinciden como conjuntos, pero no como conjuntos ordenados. Por eso es que las coordenadas en la base $\mathcal{B}_1$ son distintas que en la base $\mathcal{B}_2$.

3. Si ahora $\mathcal{B}$ es la base canónica $\{(1, 0), \ (0, 1)\}$ de $\mathbb{R}^2$, entonces

$$(x, y) = x(1, 0) + y(0, 1) \quad \Rightarrow \quad \operatorname{coord}_{\mathcal{B}}(x, y) = (x, y), \quad \forall (x, y) \in \mathbb{R}^2.$$

   Luego las coordenadas de un vector en la base canónica de $\mathbb{R}^2$ coinciden con el propio vector. Esta es la gran ventaja de trabajar con la base canónica frente a otra base.

4. Generalizando el caso anterior, es inmediato observar que si $\mathcal{B} = \{e_1, \dots, e_n\}$ es la base canónica de $\mathbb{R}^n$, entonces $\operatorname{coord}_{\mathcal{B}}(x_1, \dots, x_n) = (x_1, \dots, x_n)$, para todo $(x_1, \dots, x_n) \in \mathbb{R}^n$.

5. Si consideramos la base canónica $\mathcal{B} = \{1, \ x, \dots, x^n\}$ de $\mathbb{R}_n[x]$, entonces

$$\operatorname{coord}_{\mathcal{B}}\left(a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n\right) = (a_0, \ a_1, \ a_2, \ \cdots, \ a_n) \in \mathbb{R}^{n+1},$$

   para todo polinomio $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in \mathbb{R}_n[x]$.

A continuación veremos cómo se relacionan las cordenadas de un mismo vector en dos bases distintas.

Antes introducimos una notación. Si $v_1, \ldots, v_n$ son vectores de $\mathbb{R}^m$, entonces $[v_1|\cdots|v_n]$ es la matriz $m \times n$ obtenida escribiendo los vectores $v_1, \ldots, v_n$ en forma vertical. En este caso decimos que $v_1, \ldots, v_n$ los pensamos como *vectores columna*, es decir como matrices $m \times 1$. Por ejemplo, si $v_1 = (1, 2, 3)$ y $v_2 = (0, -4, 5)$, entonces

$$[v_1|v_2] = \begin{pmatrix} 1 & 0 \\ 2 & -4 \\ 3 & 5 \end{pmatrix}.$$

Sean $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{C} = \{w_1, \ldots, w_n\}$ dos bases de un mismo espacio V. Consideremos la matriz cuadrada ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} \in M_n$ definida por

$${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = [\mathrm{coord}_{\mathcal{C}}(v_1) \,|\, \mathrm{coord}_{\mathcal{C}}(v_2) \,|\cdots|\, \mathrm{coord}_{\mathcal{C}}(v_n)],$$

en que $\mathrm{coord}_{\mathcal{C}}(v_1)$, $\mathrm{coord}_{\mathcal{C}}(v_2), \cdots, \mathrm{coord}_{\mathcal{C}}(v_n)$ son las coordenadas en la base $\mathcal{C}$ de los vectores de la base $\mathcal{B}$ escritos en forma vertical; es decir, si

$$\begin{matrix} v_1 = a_{11}w_1 + a_{21}w_2 + \cdots + a_{n1}w_n \\ \vdots \\ v_n = a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{nn}w_n \end{matrix} \quad \Rightarrow \quad \begin{matrix} \mathrm{coord}_{\mathcal{C}}(v_1) = (a_{11}, a_{21}, \ldots, a_{n1}) \\ \vdots \\ \mathrm{coord}_{\mathcal{C}}(v_n) = (a_{1n}, a_{21}, \ldots, a_{nn}) \end{matrix} \quad \Rightarrow {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

La matriz ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$ se llama la *matriz de cambio de base*[2] de $\mathcal{B}$ a $\mathcal{C}$.

Proposición 4.6.3 (Fórmula de cambio de base)**.**
Sean $\mathcal{B}$ y $\mathcal{C}$ dos bases de un mismo espacio V. Si v es un vector de V, entonces las coordenadas de v en las bases $\mathcal{B}$ y $\mathcal{C}$ se relacionan mediante

$$\mathrm{coord}_{\mathcal{C}}(v) = {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}\, \mathrm{coord}_{\mathcal{B}}(v). \tag{4.5}$$

En la fórmula anterior se entiende que $\mathrm{coord}_{\mathcal{C}}(v)$ *y* $\mathrm{coord}_{\mathcal{B}}(v)$ *son vectores columna.* □

Ejemplo 4.6.4. Consideremos el espacio $\mathbb{R}^2$ y la base $\mathcal{B} = \{(5, 2),\ (2, 1)\}$. Si escribimos un vector genérico (x, y) en función de la base $\mathcal{B}$ obtenemos

$$(x, y) = (x - 2y)(5, 2) + (-2x + 5y)(2, 1) \quad \Rightarrow \quad \mathrm{coord}_{\mathcal{B}}(x, y) = (x - 2y,\ -2x + 5y).$$

Consideremos ahora la base $\mathcal{C} = \{(2, 1),\ (1, 1)\}$. Si escribimos los elementos de $\mathcal{B}$ como combinación lineal de la base $\mathcal{C}$ obtenemos

$$(5, 2) = 3(2, 1) - (1, 1), \quad (2, 1) = (2, 1) + 0(1, 1) \quad \Rightarrow \quad \mathrm{coord}_{\mathcal{C}}(5, 2) = (3, -1), \quad \mathrm{coord}_{\mathcal{C}}(2, 1) = (1, 0).$$

Luego la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

$${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = \big(\mathrm{coord}_{\mathcal{C}}(5, 2) \,|\, \mathrm{coord}_{\mathcal{C}}(2, 1)\big) = \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix}.$$

Usando la fórmula anterior, las coordenadas de un vector genérico (x, y) en la base $\mathcal{C}$ se obtienen aplicando la fórmula de cambio de base

$$\mathrm{coord}_{\mathcal{C}}(x, y) = {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}\, \mathrm{coord}_{\mathcal{B}}(x, y) = \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x - 2y \\ -2x + 5y \end{pmatrix} = \begin{pmatrix} x - y \\ -x + 2y \end{pmatrix}$$

Luego

$$\mathrm{coord}_{\mathcal{C}}(x, y) = (x - y,\ -x + 2y) \quad \Rightarrow \quad (x, y) = (x - y)(2, 1) + (-x + 2y)(1, 1).$$

De esa forma, conociendo las coordenadas de un vector en la base $\mathcal{B}$ y la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$, obtuvimos las coordenadas del vector en la base $\mathcal{C}$.

[2] Cuando estudiemos matrices asociadas a transformaciones lineales, va a quedar claro porqué es que a la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ la escribimos de la forma ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$.

Ahora veremos cómo se relacionan las matrices de cambio de base ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$ y ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}$. Para eso necesitamos el lema siguiente.

Lema 4.6.5. *Si una matriz cuadrada $A \in M_n$ verifica $Av = v$ para todo $v \in \mathbb{R}^n$ (v vector columna), entonces $A = I$.*

Dem. Dada $A \in M_{m\times n}$, es fácil de probar que la columna i-ésima de A es la matriz columna Ae_i, siendo $\{e_1, \dots, e_n\}$ la base canónica de $\mathbb{R}^n$. Entonces como vale $Ae_i = e_i$, para todo $i = 1, \dots, n$, deducimos

$$A = [Ae_1|\cdots|Ae_n] = [e_1|\cdots|e_n] = I. \quad \square$$

Proposición 4.6.6. *Sean $\mathcal{B}$ y $\mathcal{C}$ dos bases de un mismo espacio V. Entonces la matriz de cambio de base ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$ es la inversa de la matriz de cambio de base ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}$.*

Dem. Sea $v \in V$. Entonces

$$\mathrm{coord}_{\mathcal{B}}(v) = {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\, \mathrm{coord}_{\mathcal{C}}(v) = {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\ ({}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}\, \mathrm{coord}_{\mathcal{B}}(v)) = ({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\ {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}})\, \mathrm{coord}_{\mathcal{B}}(v).$$

Como $\{\mathrm{coord}_{\mathcal{B}}(v) : \ v \in V\} = \mathbb{R}^n$, entonces el lema 4.6.5 implica que es ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\ {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = I_n$. Luego se aplica el corolario 3.2.19. $\square$

Ejemplo 4.6.7. Consideremos el espacio $\mathbb{R}^2$ y la base $\mathcal{C} = \{(1, 1), (1, -1)\}$. Vanos a ver cómo escribir un vector genérico (x, y) como combinación lineal de la base $\mathcal{C}$, sin tener que resolver ningún sistema de ecuaciones. Sabemos que las coodenadas de (x, y) en la base canónica $\mathcal{B} = \{(1, 0), (0, 1)\}$ son $\mathrm{coord}_{\mathcal{B}}(x, y) = (x, y)$. Luego la matriz de cambio de base de $\mathcal{C}$ a $\mathcal{B}$ es

$${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} = \big(\mathrm{coord}_{\mathcal{B}}(1, 1) \,|\, \mathrm{coord}_{\mathcal{B}}(1, -1)\big) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Notar que la primer columna de la matriz ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}$ es el primer vector de la base $\mathcal{C}$ escrito en forma vertical y lo mismo sucede con la segunda columna. Entonces la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

$${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = ({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}})^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}.$$

Luego la fórmula de cambio de base (4.5) nos dice que para todo $(x, y) \in \mathbb{R}^2$ es

$$\mathrm{coord}_{\mathcal{C}}(x, y) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x-y}{2} \end{pmatrix},$$

es decir $\mathrm{coord}_{\mathcal{C}}(x, y) = \left(\frac{x+y}{2}, \frac{x-y}{2}\right)$ y por lo tanto vale

$$(x, y) = \frac{x + y}{2}(1, 1) + \frac{x - y}{2}(1, -1),$$

para todo $(x, y) \in \mathbb{R}^2$.

Ejemplo 4.6.8 (Giro de coordenadas)**.** Este ejemplo está pensado en las aplicaciones físicas del cambio de coordenadas, así que para los vectores usaremos una notación más al estilo de física.

Trabajaremos el plano $\mathbb{R}^2$. Sea $\mathcal{B} = \{\mathbf{i}, \mathbf{j}\}$ la base canónica de $\mathbb{R}^2$. Queremos saber cómo cambian las coordenadas cuando rotamos nuestro sistema de coordenadas un ángulo θ en sentido positivo (antihorario). El nuevo sistema de coordenadas va a tener una base $\mathcal{C} = \{\mathbf{i}_\theta, \mathbf{j}_\theta\}$, en que $\mathbf{i}_\theta$ y $\mathbf{j}_\theta$ son los rotados de $\mathbf{i}$ y $\mathbf{j}$, luego

$$\mathbf{i}_\theta = \cos\theta\, \mathbf{i} + \operatorname{sen}\theta\, \mathbf{j}, \quad \mathbf{j}_\theta = -\operatorname{sen}\theta\, \mathbf{i} + \cos\theta\, \mathbf{j}. \tag{4.6}$$

Si consideramos un vector arbitrario v, entonces v va a tener ciertas coordenadas (x, y) en la base $\mathcal{B}$ y otras coordenadas (X, Y) en la base $\mathcal{C}$, es decir

$$v = x\mathbf{i} + y\mathbf{j} \qquad \text{y} \qquad v = X\mathbf{i}_\theta + Y\mathbf{j}_\theta.$$

Lo que queremos saber es cómo es que se relacionan (x, y) con (X, Y). Con las notaciones de esta sección, es $\operatorname{coord}_{\mathcal{B}}(v) = (x, y)$ y $\operatorname{coord}_{\mathcal{C}}(v) = (X, Y)$. La matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ la deducimos usando las fórmulas (4.6):

$$_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \quad \Rightarrow \quad {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Luego la fórmula de cambio de base nos dice que vale

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \Rightarrow \quad \begin{cases} X = x\cos\theta + y\operatorname{sen}\theta \\ Y = -x\operatorname{sen}\theta + y\cos\theta \end{cases}.$$

Estas últimas fórmulas son las que estábamos buscando. Nos dicen que si la expresión de v en la base canónica $\mathcal{B}$ es $v = x\mathbf{i} + y\mathbf{j}$, entonces su expresión en la base $\mathcal{C}$ es

$$v = (x\cos\theta + y\operatorname{sen}\theta)\,\mathbf{i}_\theta + (-x\operatorname{sen}\theta + y\cos\theta)\,\mathbf{j}_\theta.$$

Capítulo 5

Transformaciones lineales

En esta sección estudiaremos las funciones entre espacios vectoriales que preservan la estructura lineal. Como antes, asumiremos siempre que los espacios vectoriales son de dimensión finita, aunque va a ser claro que varias de las definiciones y resultados valen en general.

5.1. Definiciones y propiedades básicas

Definición 5.1.1. Sean V y W dos espacios vectoriales. Una función $T : V \to W$ es una *transformación lineal* si verifica:

- $T(u + v) = T(u) + T(v)$, para todo $u, v \in V$;
- $T(av) = aT(v)$, para todo $a \in \mathbb{R}$ y $v \in V$.

La siguiente proposición da una forma rápida para chequear si una función entre dos espacios es una transformación lineal.

Proposición 5.1.2. *Sean V y W dos espacios vectoriales. Una función $T : V \to W$ es una transformación lineal si y solo si verifica:*

$$T(au + v) = aT(u) + T(v),$$

para todo $u, v \in V$ y $a \in \mathbb{R}$. □

Ejemplos 5.1.3. Ejemplos de transformaciones lineales.

1. La función *nula* $T : V \to W$ definida por $T(v) = 0$, para todo $v \in V$ (V y W son dos espacios vectoriales arbitrarios).

2. La función identidad $\mathrm{Id} : V \to V$ definida por $\mathrm{Id}(v) = v$, para todo $v \in V$ (V es un espacio vectorial arbitrario).

3. *Simetría respecto al eje Ox*. La función $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (x, -y)$, para todo $(x, y) \in \mathbb{R}^2$.

4. La función $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (x + y,\, x - y,\, 3y)$, para todo $(x, y) \in \mathbb{R}^2$.

5. La función $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T(x, y, z) = (x, y)$, para todo $(x, y, z) \in \mathbb{R}^3$, es la *proyección* de $\mathbb{R}^3$ sobre $\mathbb{R}^2$.

6. La función $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (x, y, 0)$, para todo $(x, y) \in \mathbb{R}^2$, es la *inclusión* de $\mathbb{R}^2$ en $\mathbb{R}^3$.

7. La función $T : M_{m\times n} \to M_{n\times m}$ definida por $T(A) = A^t$ (la traspuesta), para todo $A \in M_{m\times n}$.

Los siguientes son ejemplos de transformaciones lineales en espacios de dimensión infinita.

8. Si $I \subset \mathbb{R}$ es un intervalo abierto y $C^\infty(I) := \{f : I \to \mathbb{R} : \ f \text{ es derivable de todos los órdenes}\}$, entonces la función $T : C^\infty(I) \to C^\infty(I)$ definida por $T(f) = f'$ (la derivada) es una transformación lineal.

9. Si $[a, b]$ es un intervalo cerrado de $\mathbb{R}$ y consideramos $C([a, b]) := \{f : [a, b] \to \mathbb{R} : \ f \text{ es continua}\}$, entonces $T : C([a, b]) \to \mathbb{R}$ definida por $T(f) = \int_a^b f(x)\,dx$ es una transformación lineal.

La siguiente proposición resume las propiedades básicas de las transformaciones lineales.

Proposición 5.1.4. *Sea $T : V \to W$ una transformación lineal. Vale lo siguiente.*

1. *Si $a_1, \dots, a_n \in \mathbb{R}$ y $v_1, \dots, v_n \in V$ $(n = 1, 2, \dots)$, entonces*
$$T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n). \tag{5.1}$$

2. *$T(0) = 0$ (la imagen por T del vector nulo de V es el vector nulo de W).* □

Notar que de la fórmula (5.1) se deducen las siguientes igualdades
$$T(-v) = -T(v), \quad T(u - v) = T(u) - T(v),$$
para todo $u, v \in V$.

Observación 5.1.5. A cada matriz $A \in M_{m\times n}$ le podemos asociar una función $L_A : \mathbb{R}^n \to \mathbb{R}^m$, definida por $L_A(v) = Av$, para todo $v \in \mathbb{R}^n$. En la fórmula anterior se entiende que el vector v en Av está escrito como un vector columna (para poderlo multiplicar por la matriz A) y el resultado Av es también un vector columna; más formalmente Av es el producto de la matriz $A \in M_{m\times n}$ por la matriz columna $v \in M_{n\times 1}$ y su resultado es la matriz columna $Av \in M_{m\times 1}$. Explícitamente, si $A = (a_{ij})$ y $v = (x_1, \dots, x_n)$, entonces
$$L_A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix},$$
es decir
$$L_A(x_1, \dots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n,\ \dots,\ a_{m1}x_1 + \cdots + a_{mn}x_n),$$
para todo $(x_1, \dots, x_n) \in \mathbb{R}^n$. Por ejemplo, si $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 0 & -1 \end{pmatrix}$, entonces $L_A : \mathbb{R}^2 \to \mathbb{R}^3$ está definida por
$$L_A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + 3y \\ x + 2y \\ -y \end{pmatrix} \quad \Rightarrow \quad L_A(x, y) = (2x + 3y,\ x + 2y,\ -y).$$
Volviendo al caso general, notar que para toda matriz $A \in M_{m\times n}$ vale
$$L_A(au + v) = A(au + v) = A(au) + Av = a(Au) + Av = aL_A(u) + L_A(v),$$
para todo $a \in \mathbb{R}$, $u, v \in \mathbb{R}^n$; luego $L_A : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal.

El siguiente resultado muestra que toda transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ es de la forma L_A, para cierta matriz A.

Proposición 5.1.6. *Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal, entonces existen escalares $a_{ij} \in \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$ tales que*
$$T(x_1, \dots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n,\ \dots,\ a_{m1}x_1 + \cdots + a_{mn}x_n),$$
para todo $(x_1, \dots, x_n) \in \mathbb{R}^n$. Luego $T = L_A$, siendo $A \in M_{m\times n}$ definida por
$$A = [T(e_1)|\cdots|T(e_n)], \quad \{e_1, \dots, e_n\} \text{ base canónica de } \mathbb{R}^n.$$

Dem. Sean $T(e_1)=(a_{11},\dots,a_{m1}),\ \dots,\ T(e_n)=(a_{1n},\dots,a_{mn})$. Entonces

$$\begin{aligned} T(x_1,\dots,x_n) &= T(x_1e_1+\cdots+x_ne_n) = x_1T(e_1)+\cdots+x_nT(e_n) \\ &= x_1(a_{11},\dots,a_{m1})+\cdots+x_n(a_{1n},\dots,a_{mn}) \\ &= (a_{11}x_1+\cdots+a_{1n}x_n,\ \dots,\ a_{m1}x_1+\cdots+a_{mn}x_n). \end{aligned}$$

Esto prueba la primer afirmación. La segunda es inmediata. □

Observar que si $T=L_A$, entonces evaluando T en la base canónica obtenemos $A=[T(e_1)|\cdots|T(e_n)]$. Luego dada una transformación lineal $T:\mathbb{R}^n\to\mathbb{R}^m$, existe un única matriz A tal que $T=L_A$.

Ejemplo 5.1.7. Sea $T:\mathbb{R}^2\to\mathbb{R}^3$ la transformación lineal definida por $T(x,y)=(x+y,\ x-y,\ 3y)$. Es $T(1,0)=(1,1,0)$ y $T(0,1)=(1,-1,3)$, luego $T=L_A$ siendo $A=\begin{pmatrix}1&1\\1&-1\\0&3\end{pmatrix}$.

La proposición anterior describe las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$. Para espacios vectoriales arbitrarios tenemos el siguiente resultado.

Proposición 5.1.8. *Sean V y W dos espacios vectoriales. Sea $\mathcal{B}=\{v_1,\dots,v_n\}$ una base de V y $w_1,\dots,w_n$ vectores arbitrarios en W. Entonces existe una única transformación lineal $T:V\to W$ tal que $T(v_i)=w_i$, para todo $i=1,\dots,n$. Esta función está definida por*

$$T(x_1v_1+\cdots+x_nv_n)=x_1w_1+\cdots+x_nw_n, \tag{5.2}$$

para todo $x_1,\dots,x_n\in\mathbb{R}$.

Dem. Si existe una transformación lineal $T:V\to W$ que verifique $T(v_i)=w_i$, para todo i, entonces para todo $x_1,\dots,x_n\in\mathbb{R}$ es

$$T(x_1v_1+\cdots+x_nv_n)=x_1T(v_1)+\cdots+x_nT(v_n)=x_1w_1+\cdots+x_nw_n.$$

Esto prueba la unicidad.

Para probar la existencia, como $\mathcal{B}$ es base de V, entonces todo vector de V se escribe en forma única como combinación lineal de $v_1,\dots,v_n$. Luego tiene sentido definir una función $T:V\to W$ mediante la fórmula (5.2), y es fácil de probar que T definida de esta forma es una transformación lineal. □

Ejemplo 5.1.9. Nos preguntamos si existe una transformación lineal $T:\mathbb{R}^2\to\mathbb{R}^3$ tal que $T(1,1)=(2,3,3)$ y $T(1,-1)=(0,-1,1)$. Como $\mathcal{B}=\{(1,1),\ (1,-1)\}$ es una base de $\mathbb{R}^2$, entonces la proposición anterior nos dice que siempre existe esa transformación lineal y que es única. Para encontrarla tenemos que ver cómo escribir un vector arbitrario (x,y) de $\mathbb{R}^2$ como combinación lineal de la base $\mathcal{B}$, es decir hallar sus coordenadas en esa base. Esto lo vimos en el ejemplo 4.6.2.1: $\operatorname{coord}_{\mathcal{B}}(x,y)=\left(\frac{x+y}{2},\ \frac{x-y}{2}\right)$. Luego la transformación lineal T está definida por

$$T(x,y)=T\left(\frac{x+y}{2}(1,1)+\frac{x-y}{2}(1,-1)\right)=\frac{x+y}{2}(2,3,3)+\frac{x-y}{2}(0,1,-1)=(x+y,\ x+2y,\ 2x+y),$$

es decir, $T(x,y)=(x+y,\ x+2y,\ 2x+y)$, para todo $(x,y)\in\mathbb{R}^2$.

Observación 5.1.10. No siempre existe una transformación lineal como en el ejemplo anterior. Por ejemplo, si existiese una transformación lineal $T:\mathbb{R}^2\to\mathbb{R}^3$ tal que $T(1,0)=(2,3,3)$ y $T(2,0)=(0,-1,1)$, entonces valdría

$$(0,-1,1)=T(2,0)=2T(1,0)=2(2,3,3)=(4,6,6)\quad \lightning.$$

Por otro lado, si buscamos una transformación lineal $T:\mathbb{R}^2\to\mathbb{R}^3$ tal que $T(1,1)=(2,3,3)$, entonces existen infinitas transformaciones lineales que lo verifican. Alcanza con completar el vector $(1,1)$ a una base de $\mathbb{R}^2$ y aplicar la proposición 5.1.8 (el ejemplo anterior es un caso particular de esta construcción).

Recordar que si $f:X\to Y$ y $g:Y\to Z$ son dos funciones entre conjuntos, entonces su *compuesta* es la función $g\circ f:X\to Z$ definida por $(g\circ f)(x)=g\big(f(x)\big)$, para todo $x\in X$.

Proposición 5.1.11. *Si $T: V \to W$ y $S: W \to U$ son dos transformaciones lineales, entonces la función compuesta $S \circ T: V \to U$ es una transformación lineal.* □

Proposición 5.1.12. *Sean $A \in M_{m\times n}$ y $B \in M_{p\times m}$. Consideremos las transformaciones lineales correspondientes $L_A : \mathbb{R}^n \to \mathbb{R}^m$ y $L_B : \mathbb{R}^m \to \mathbb{R}^p$. Entonces la función compuesta $L_B \circ L_A : \mathbb{R}^n \to \mathbb{R}^p$ verifica $L_B \circ L_A = L_{BA}$.*

Dem. $(L_B \circ L_A)(v) = L_B\big(L_A(v)\big) = B(Av) = (BA)v = L_{BA}(v)$, para todo $v \in \mathbb{R}^n$. □

Observación 5.1.13. La proposición anterior es la razón por la cual el producto de matrices se define por la fórmula (3.1). Es decir el producto de matrices está definido para que valga $L_B \circ L_A = L_{BA}$.

5.2. El núcleo y la imagen.

Definición 5.2.1. Sea $T : V \to W$ una transformación lineal. El *núcleo* de T es el conjunto[1] $\operatorname{Ker}(T) \subset V$ y su *imagen* o *recorrido* es el conjunto $\operatorname{Im}(T) \subset W$, definidos por

$$\operatorname{Ker}(T) := \{v \in V : \ T(v) = o\}, \qquad \operatorname{Im}(T) := \{w \in W : \ \exists v \in V \text{ tal que } w = T(v)\}.$$

Proposición 5.2.2. *Si $T : V \to W$ es una transformación lineal, entonces $\operatorname{Ker}(T)$ es un subespacio de V y $\operatorname{Im}(T)$ es un subespacio de W.* □

Recordar que una función entre dos conjuntos $f : X \to Y$ se dice *inyectiva* si $x \neq x'$ implica $f(x) \neq f(x')$, *sobreyectiva* si para todo $y \in Y$ existe un $x \in X$ tal que $f(x) = y$, y *biyectiva* si es inyectiva y sobreyectiva.

La siguiente proposición relaciona el núcleo de una transformación lineal con su inyectividad.

Proposición 5.2.3. *Si $T : V \to W$ es una transformación lineal, entonces T es inyectiva si y solo si $\operatorname{Ker}(T) = \{o\}$.* □

Observación 5.2.4. Es claro que una transformación lineal $T : V \to W$ es sobreyectiva si y solo si $\operatorname{Im}(T) = W$. Luego conociendo el núcleo y la imagen de una transformación lineal T, sabemos si T es inyectiva, sobreyectiva o biyectiva.

Observación 5.2.5. Sea $T : V \to W$ una transformación lineal. Es fácil de probar usando la fórmula (5.1) que si $\mathcal{A} = \{v_1, \dots, v_n\}$ es un subconjunto LD de V, entonces $T(\mathcal{A}) = \{T(v_1), \dots, T(v_n)\}$ es un subconjunto LD de W. Por otro lado si consideramos la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T(x, y, z) = (x, y)$, entonces el conjunto $\mathcal{A} = \{(1, 2, 1), (1, 2, 3)\} \subset \mathbb{R}^3$ es LI, pero $T(\mathcal{A}) = \{(1, 2), (1, 2)\} \subset \mathbb{R}^2$ claramente es LD. Observar también que si consideramos la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (x, y, 0)$, entonces $\mathcal{G} = \{(1, 0), (0, 1), (1, 1)\}$ es un conjunto generador de $\mathbb{R}^2$, pero $T(\mathcal{G}) = \{(1, 0, 0), (0, 1, 0), (1, 1, 0)\}$ no es un conjunto generador de $\mathbb{R}^3$. Luego en general las transformaciones lineales no llevan conjuntos LI en conjuntos LI, ni conjuntos generadores en conjuntos generadores. Sin embargo la proposición siguiente muestra que si la transformación lineal verifica ciertas condiciones, entonces las propiedades anteriores sí se cumplen.

Proposición 5.2.6. *Sea $T : V \to W$ una transformación lineal y $\mathcal{A} = \{v_1, \dots, v_n\}$ un subconjunto de V.*

1. *Si T es inyectiva y $\mathcal{A}$ es un subconjunto LI de V, entonces $T(\mathcal{A})$ es un subconjunto LI de W.*
2. *Si $\mathcal{A}$ es un conjunto generador de V, entonces $T(\mathcal{A})$ es un conjunto generador de $\operatorname{Im}(T)$. Luego si T es sobreyectiva y $\mathcal{A}$ es un conjunto generador de V, entonces $T(\mathcal{A})$ es un conjunto generador de W.*
3. *Si T es biyectiva y $\mathcal{A}$ es una base de V, entonces $T(\mathcal{A})$ es una base de W.*

Dem. Observar que la tercer afirmación se deduce de las dos anteriores. Así que necesitamos probar solo las dos primeras.

Supongamos que T es inyectiva y $\mathcal{A} = \{v_1, \dots, v_n\}$ es un subconjunto LI de V. Queremos probar que $T(\mathcal{A}) = \{T(v_1), \dots, T(v_n)\}$ es un subconjunto LI de W. Sean $a_1, \dots, a_n \in \mathbb{R}$ tales que $a_1T(v_1) + \cdots + a_nT(v_n) = 0$. Entonces

$$T(a_1v_1 + \cdots + a_nv_n) = 0 \quad \Rightarrow \quad a_1v_1 + \cdots + a_nv_n \in \operatorname{Ker}(T) = \{0\} \quad \Rightarrow \quad a_1v_1 + \cdots + a_nv_n = 0.$$

Como $\mathcal{A}$ es LI, se concluye que es $a_1 = \cdots = a_n = 0$. Luego $T(\mathcal{A})$ es LI.

[1]La abreviación Ker viene de la palabra *kernel* que es el término que se usa en inglés para denominar el núcleo.

Veamos ahora la segunda afirmación. Sea $\mathcal{A} = \{v_1, \ldots, v_n\}$ un conjunto generador de V. Si $w \in \operatorname{Im}(T)$, entonces existe $v \in V$ tal que $w = T(v)$. Como $\mathcal{A}$ es un conjunto generador de V, entonces existen $a_1, \ldots, a_n \in \mathbb{R}$ tales que $v = a_1v_1 + \cdots + a_nv_n$. Luego

$$w = T(v) = T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n).$$

Esto muestra que todo elemento de $\operatorname{Im}(T)$ es combinación lineal de $T(\mathcal{A})$. $\square$

Observación 5.2.7. Si $T : V \to W$ una transformación lineal arbitraria y $\mathcal{B}$ es una base de V, entonces podemos afirmar que $T(\mathcal{B})$ es un conjunto generador de $\operatorname{Im}(T)$, pero no es necesariamente una base (ver el ejemplo siguiente).

Ejemplo 5.2.8. Consideremos la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$T(x, y, z) = (x + y + z,\ x + y - z,\ 2x + 2y + 2z,\ -x - y + z). \tag{5.3}$$

Para determinar el núcleo de T tenemos que resolver el sistema homogéneo

$$\left\{\begin{array}{rcl} x+y+z &=& 0\\ x+y-z &=& 0\\ 2x+2y+2z &=& 0\\ -x-y+z &=& 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+y+z &=& 0\\ x+y-z &=& 0\\ 0 &=& 0\\ 0 &=& 0 \end{array}\right. \sim \left\{\begin{array}{rcl} x+y+z &=& 0\\ 2z &=& 0\\ 0 &=& 0\\ 0 &=& 0 \end{array}\right. \Rightarrow \quad z = 0,\ y = -x.$$

Luego $\operatorname{Ker}(T) = \{(x, -x, 0) :\ x \in \mathbb{R}\} = [(1, -1, 0)]$ y es claro que $\{(1, -1, 0)\}$ es base de $\operatorname{Ker}(T)$.

Para saber si un vector $(a, b, c, d) \in \mathbb{R}^4$ está en la imagen de T, tenemos que averiguar si existe un vector $(x, y, z) \in \mathbb{R}^3$ tal que $(a, b, c, d) = T(x, y, z)$, es decir si el sistema

$$\left\{\begin{array}{rcl} x+y+z &=& a\\ x+y-z &=& b\\ 2x+2y+2z &=& c\\ -x-y+z &=& d \end{array}\right. \tag{5.4}$$

es compatible. Luego $\operatorname{Im}(T)$ es el conjunto de los vectores $(a, b, c, d) \in \mathbb{R}^4$ tales que el sistema de arriba es compatible. En nuestro caso, si a la tercer ecuación le sumamos la primera multiplicada por -2 y a la cuarta le sumamos la segunda, obtenemos que el sistema anterior es equivalente al siguiente

$$\left\{\begin{array}{rcl} x+y+z &=& a\\ x+y-z &=& b\\ 0 &=& c-2a\\ 0 &=& d+b \end{array}\right. .$$

Luego para que el sistema tenga solución, necesariamente debe ser $c = 2a$ y $d = -b$, y es claro que si esto sucede, entonces de las dos primeras ecuaciones siempre podemos despejar x, y, z (no en forma única). Luego el sistema (5.4) es compatible si y solo si $c = 2a$ y $d = -b$, por lo tanto la imagen de T es

$$\operatorname{Im}(T) = \left\{(x, y, z, t) \in \mathbb{R}^4 :\ z = 2x,\ t = -y\right\}. \tag{5.5}$$

Notar que determinar $\operatorname{Im}(T)$ trabajando con el sistema (5.4) fue bastante más complicado que para $\operatorname{Ker}(T)$. Una forma alternativa para hacerlo, consiste en aplicar la segunda parte de la proposición 5.2.6. Si consideramos $\mathcal{B} = \{(1, 0, 0),\ (0, 1, 0),\ (0, 0, 1)\}$ la base canónica de $\mathbb{R}^3$, entonces es

$$T(1, 0, 0) = (1, 1, 2, -1); \quad T(0, 1, 0) = (1, 1, 2, -1); \quad T(0, 0, 1) = (1, -1, 2, 1).$$

Luego el conjunto $T(\mathcal{B}) = \{(1, 1, 2, -1),\ (1, 1, 2, -1),\ (1, -1, 2, 1)\}$ es un generador de $\operatorname{Im}(T)$ y claramente es LD. Entonces $T(\mathcal{B})$ contiene una base de $\operatorname{Im}(T)$, que en este caso es $\{(1, 1, 2, -1),\ (1, -1, 2, 1)\}$. Luego la imagen de T es el subespacio generado por este conjunto, es decir

$$\operatorname{Im}(T) = [(1, 1, 2, -1),\ (1, -1, 2, 1)]. \tag{5.6}$$

Es un ejercicio el verificar que las dos formas de describir $\operatorname{Im}(T)$ dadas por (5.5) y (5.6), coinciden.

El resultado siguiente se utiliza con frecuencia.

Teorema 5.2.9. *Sea $T : V \to W$ una transformación lineal. Entonces vale*

$$\dim \operatorname{Ker}(T) + \dim \operatorname{Im}(T) = \dim V.$$

Dem. Sea $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ una base de $\operatorname{Ker} T$. Como $\{v_1, \ldots, v_n\} \subset V$ es LI, entonces existen $w_1, \ldots, w_m$ en V tales que $\mathcal{B}_2 = \{v_1, \ldots, v_n, w_1, \ldots, w_m\}$ es una base de V. Luego

$$\{T(v_1), \ldots, T(v_n), T(w_1), \ldots, T(w_m)\}$$

es un conjunto generador de $\operatorname{Im}(T)$. Pero $T(v_1) = \cdots = T(v_n) = 0$, luego $\mathcal{B}_3 = \{T(w_1), \ldots, T(w_m)\}$ es también un conjunto generador de $\operatorname{Im}(T)$. Veamos que $\mathcal{B}_3$ es LI. Sean $a_1 \ldots, a_m \in \mathbb{R}$ tales que $a_1 T(w_1) + \cdots + a_m T(w_m) = 0$. Luego

$$T(a_1 w_1 + \cdots + a_m w_m) = 0 \quad \Rightarrow \quad a_1 w_1 + \cdots + a_m w_m \in \operatorname{Ker}(T) = [v_1, \ldots, v_n].$$

Entonces existen $b_1 \ldots, b_n \in \mathbb{R}$ tales que

$$a_1 w_1 + \cdots + a_m w_m = b_1 v_1 + \cdots + b_n v_n \quad \Rightarrow \quad b_1 v_1 + \cdots + b_n v_n + (-a_1) w_1 + \cdots + (-a_m) w_m = 0.$$

Como $\mathcal{B}_2$ es base de V, deducimos que es $b_1 = \cdots = b_n = a_1 = \cdots = a_m = 0$. Luego $\mathcal{B}_3$ es LI y por lo tanto es base de $\operatorname{Im}(T)$. Entonces[2]

$$\dim \operatorname{Ker}(T) + \dim \operatorname{Im}(T) = \#\mathcal{B}_1 + \#\mathcal{B}_3 = n + m = \#\mathcal{B}_2 = \dim V. \quad \square$$

Observación 5.2.10. Dada una transformación lineal $T : V \to W$, la *nulidad* de T es la dimensión de su núcleo y el *rango* de T es la dimensión de su imagen. Luego el teorema anterior puede reescribirse diciendo que dada una transformación lineal, la dimensión de su dominio es igual a la suma de su nulidad con su rango.

Ejemplo 5.2.11. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T(x, y, z) = (x + y + 2z,\ x + z,\ y + z).$$

Razonando como en el ejercicio 5.2.8 deducimos que $\{(-1, -1, 1)\}$ es una base de $\operatorname{Ker}(T)$ y que el conjunto $\mathcal{G} = \{(1, 1, 0),\ (1, 0, 1),\ (2, 1, 1)\}$ es un generador de $\operatorname{Im}(T)$. Luego $\dim \operatorname{Ker}(T) = 1$ y como $\dim \mathbb{R}^3 = 3$, entonces el teorema anterior implica $\dim \operatorname{Im}(T) = 2$. Como el conjunto generador $\mathcal{G}$ de $\operatorname{Im}(T)$ tiene tres elementos, para obtener una base de $\operatorname{Im}(T)$ alcanza con tomar dos elementos de $\mathcal{G}$ que sean LI; en este caso cualquier par de esos elementos lo verifica, luego

$$\{(1, 1, 0),\ (1, 0, 1)\}, \quad \{(1, 1, 0),\ (2, 1, 1)\}, \quad \{(2, 1, 1),\ (1, 0, 1)\}$$

son bases de $\operatorname{Im}(T)$.

Definición 5.2.12. Un *isomorfismo* entre dos espacios vectoriales V y W es una transformación lineal biyectiva $T : V \to W$.

Ejemplo 5.2.13. Si la dimensión de un espacio vectorial V es n y $\mathcal{B}$ es una base de V, entonces la transformación lineal $T : V \to \mathbb{R}^n$ definida por $T(v) = \operatorname{coord}_{\mathcal{B}}(v)$, para todo $v \in V$, es un isomorfismo.

Recordar que si $f : X \to Y$ es una función biyectiva, entonces f establece una correspondencia uno a uno entre los elementos de X y los de Y, y por lo tanto nos permite definir su *función inversa* $f^{-1} : Y \to X$ mediante $f^{-1}(y) = x$ si y solo si $y = f(x)$. La función inversa de f queda caracterizada por ser (en caso de existir) la única función $f^{-1} : Y \to X$ que verifica

$$f \circ f^{-1} = \mathrm{id}_Y \quad \text{y} \quad f^{-1} \circ f = \mathrm{id}_X. \tag{5.7}$$

En particular esto implica que $f^{-1} : Y \to X$ es biyectiva y su inversa es f, es decir $\left(f^{-1}\right)^{-1} = f$.

[2]El símbolo # quiere decir la cantidad de elementos del conjunto.

Proposición 5.2.14. *Si $T : V \to W$ es un isomorfismo, entonces su función inversa $T^{-1} : W \to V$ es una transformación lineal y por lo tanto es también un isomorfismo.*

Dem. La función inversa $T^{-1} : W \to V$ está definida por $T^{-1}(w) = v$ si y solo si $w = T(v)$, para todo $v \in V,\ w \in W$. Sean $w_1, w_2 \in W$ y $a \in \mathbb{R}$. Consideremos $v_1 = T^{-1}(w_1)$ y $v_2 = T^{-1}(w_2)$. Como T es lineal, es

$$T(av_1 + v_2) = aT(v_1) + T(v_2) = aw_1 + w_2 \quad \Rightarrow \quad T^{-1}(aw_1 + w_2) = av_1 + v_2 = aT^{-1}(w_1) + T^{-1}(w_2). \quad \square$$

De la tercera parte de la proposición 5.2.6 se deduce la siguiente.

Proposición 5.2.15. *Si existe un isomorfismo $T : V \to W$, entonces $\dim V = \dim W$.* $\square$

Por la proposición anterior, solo puede existir un isomorfismo entre espacios que tengan la misma dimensión. El siguiente resultado muestra que si pasa esto último, entonces para probar que una transformación lineal entre estos es un isomorfismo, alcanza con probar que es inyectiva o sobreyectiva.

Teorema 5.2.16. *$T : V \to W$ una transformación lineal entre dos espacios que tienen la misma dimensión. Entonces las siguientes afirmaciones son equivalentes:*

1. *T es inyectiva,*
2. *T es sobreyectiva,*
3. *T es un isomorfismo.*

Dem. Es claro que si T es un isomorfismo, entonces T es inyectiva y biyectiva. Luego solo hay que probar que T es inyectiva si y solo si T es sobreyectiva.

Recordar que en el teorema 5.2.9 probamos que vale la siguiente fórmula

$$\dim \mathrm{Ker}\,(T) + \dim \mathrm{Im}(T) = \dim V. \tag{5.8}$$

Si T es inyectiva, entonces $\mathrm{Ker}\,(T) = \{0\}$ y usando (5.8) deducimos $\dim \mathrm{Im}(T) = \dim V$. Como por hipótesis es $\dim V = \dim W$, concluimos que T es sobreyectiva.

Si T es sobreyectiva, es $\mathrm{Im}(T) = W$ y por lo tanto $\dim \mathrm{Im}(T) = \dim W = \dim V$. Entonces usando de nuevo (5.8) deducimos $\dim \mathrm{Ker}\,(T) = 0$; luego T es inyectiva. $\square$

El siguiente resultado describe los isomorfismos de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Proposición 5.2.17.
1. *Si existe un isomorfismo $T : \mathbb{R}^n \to \mathbb{R}^m$, entonces $n = m$.*
2. *Si consideramos una transformación lineal $L_A : \mathbb{R}^n \to \mathbb{R}^n$, entonces L_A es un isomorfismo si y solo si la matriz A es invertible ($\det A \neq 0$). En ese caso vale $(L_A)^{-1} = L_{A^{-1}}$.*

Dem. La primera afirmación es inmediata a partir de la proposición 5.2.15. Consideremos la segunda. Si A es invertible, entonces la proposición 5.1.12 implica

$$L_A \circ L_{A^{-1}} = L_{A^{-1}} \circ L_A = L_I = \mathrm{Id},$$

luego L_A es biyectiva y su inversa es $L_{A^{-1}}$.

Supongamos ahora que $L_A : \mathbb{R}^n \to \mathbb{R}^n$ es biyectiva. Sabemos que su inversa $(L_A)^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ es también una transformación lineal y por lo tanto existe una matriz $B \in M_n$ tal que $(L_A)^{-1} = L_B$. Luego

$$\mathrm{Id} = L_A \circ (L_A)^{-1} = L_A \circ L_B = L_{AB}.$$

Así es $L_{AB} = \mathrm{Id}$ y por lo tanto $AB = I$ (usando el lema 4.6.5). Luego A es invertible y $B = A^{-1}$. $\square$

5.3. Aplicación a sistemas de ecuaciones

En esta sección aplicaremos técnicas de álgebra lineal para el estudio de sistemas de ecuaciones. Consideremos el sistema

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right. . \tag{5.9}$$

Si definimos

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{m\times n}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n\times 1}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in M_{m\times 1},$$

entonces el sistema (5.9) es equivalente a la ecuación matricial

$$AX = B. \tag{5.10}$$

La ventaja de (5.10) es que tiene una sola incógnita (matricial). Para el estudio de (5.9) conviene considerar el sistema *homogéneo* asociado:

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & 0 \end{array}\right. \tag{5.11}$$

y la correspondiente ecuación matricial

$$AX = O, \tag{5.12}$$

siendo O la matriz nula de tamaño $m \times 1$. Recordar que el sistema homogéneo siempre admite la solución trivial $x_1 = \cdots = x_n = 0$, por lo que siempre es compatible. Lo que interesa en este caso es saber si admite alguna solución no trivial, es decir si es indeterminado. Sea $\mathcal{E}_B$ el conjunto de soluciones de la ecuación $AX = B$ y $\mathcal{E}_O$ el conjunto de soluciones de la ecuación $AX = O$, es decir

$$\mathcal{E}_B = \{X \in \mathbb{R}^n : \ AX = B\}, \qquad \mathcal{E}_O = \{X \in \mathbb{R}^n : \ AX = O\},$$

Observar que si consideramos la transformación lineal $L_A : \mathbb{R}^n \to \mathbb{R}^m$ defnida por $L_A(X) = AX$, entonces $\mathcal{E}_0 = \operatorname{Ker}(L_A)$ y $\mathcal{E}_B = (L_A)^{-1}(\{B\})$ (el conjunto de las preimágenes de B por medio de la función L_A).

Observación 5.3.1. La igualdad $\mathcal{E}_0 = \operatorname{Ker}(L_A)$ implica que $\mathcal{E}_0$ es un subespacio de $\mathbb{R}^n$. Por lo tanto si $\mathcal{E}_0 \neq \{0\}$, entonces $\mathcal{E}_0$ es un conjunto infinito. Es decir, si el sistema homogéneo (5.11) es indeterminado (admite alguna solución no trivial), entonces tiene infinitas soluciones.

La siguiente proposición describe cómo se relacionan $\mathcal{E}_B$ y $\mathcal{E}_O$.

Proposición 5.3.2.

1. *Si $X_1, X_2 \in \mathcal{E}_B$, entonces $X_1 - X_2 \in \mathcal{E}_O$.*
2. *Si $X_1 \in \mathcal{E}_B$ y $X_0 \in \mathcal{E}_O$, entonces $X_1 + X_0 \in \mathcal{E}_B$.*
3. *Si $\mathcal{E}_B \neq \emptyset$ y $X_1 \in \mathcal{E}_B$, entonces $\mathcal{E}_B = \{X_1 + X : \ X \in \mathcal{E}_0\}$.* □

Observación 5.3.3. La primer parte de la proposición anterior muestra que si el sistema (5.9) admite más de una solución (es indeterminado), entonces el sistema homogéneo (5.11) admite alguna solución no trivial y por lo tanto tiene infinitas soluciones. Luego la tercer parte de la proposición anterior implica que el sistema (5.9) también tiene infinitas soluciones. Es decir, esto prueba que un sistema indeterminado (homogéneo o no) siempre tiene infinitas soluciones.

Notar que el sistema (5.9) puede ser compatible o no, pero en caso de ser compatible, entonces es determinado si y solo si la única solución que admite el sistema homogéneo asociado (5.11) es la trivial, y esto depende solo de la matriz A y no de la columna de resultados B.

En la proposición y corolario siguientes mantenemos las notaciones anteriores.

Proposición 5.3.4. *Si $n > m$, o $m = n$ y $\det A = 0$, entonces el sistema homogéneo* (5.11) *admite soluciones no triviales.*

Dem. Consideremos primero el caso $n > m$. Aplicando el teorema 5.2.9 obtenemos

$$\dim \operatorname{Ker}(L_A) + \dim \operatorname{Im}(L_A) = n.$$

Como $\operatorname{Im}(L_A)$ es un subespacio de $\mathbb{R}^m$, es

$$\dim \operatorname{Im}(L_A) \leq m < n \quad \Rightarrow \quad \dim \operatorname{Im}(L_A) < n \quad \Rightarrow \quad \dim \operatorname{Ker}(L_A) = n - \dim \operatorname{Im}(L_A) > 0.$$

Luego $\mathcal{E}_0 = \operatorname{Ker}(L_A) \neq \{0\}$.

Consideremos ahora el caso $m = n$ y $\det A = 0$. Como $\det A = 0$, entonces el corolario 5.2.17 implica que $L_A : \mathbb{R}^n \to \mathbb{R}^n$ no es un isomorfismo. Luego el teorema 5.2.16 implica que L_A no es un inyectiva y por lo tanto $\mathcal{E}_0 = \operatorname{Ker}(L_A) \neq \{0\}$. □

Corolario 5.3.5. *Si $n > m$ o $m = n$ y $\det A = 0$, entonces el sistema* (5.9) *es incompatible o compatible indeterminado.*

Observación 5.3.6. En palabras, el corolario anterior dice que si en un sistema de ecuaciones, el número de variables es mayor que el de ecuaciones o hay tantas ecuaciones como variables pero el determinante de la matriz del sistema es nulo, entonces el sistema es incompatible o compatible indeterminado.

A continuación veremos la prueba del método de Cramer, estudiado en la sección 3.3. Consideremos un sistema cuadrado

$$\left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{n1}x_1 + \cdots + a_{nn}x_n & = & b_n \end{array} \right. . \tag{5.13}$$

Sean

$$\Delta = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}; \quad \Delta_i = \begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix}, \ i = 1, \ldots, n.$$

Teorema 5.3.7 (Cramer).

1. *Si $\Delta \neq 0$, entonces el sistema* (5.13) *es compatible determinado y la solución es $x_i = \frac{\Delta_i}{\Delta}$, para todo $i = 1, \ldots, n$.*

2. *Si $\Delta = 0$ y existe algún $i = 1, \ldots, n$ tal que $\Delta_i \neq 0$, entonces el sistema* (5.13) *es incompatible.*

3. *Si $\Delta = \Delta_1 = \cdots = \Delta_n = 0$, entonces el sistema* (5.13) *no es determinado; puede ser incompatible o compatible indeterminado.*

Dem. Sean $A = (a_{ij})$ la matriz del sistema, $\hat{A} = (\Delta_{ij})^t$ la matriz definida en el lema 3.2.16 y $X = (x_1, \ldots, x_n)$. Recordar que vale $A\hat{A} = \hat{A}A = \Delta I$. Si multiplicamos la ecuación $AX = B$ por $\hat{A}$ obtenemos

$$AX = B \ \Rightarrow \ \hat{A}AX = \hat{A}B \quad \Leftrightarrow \quad \Delta X = \hat{A}B \quad \Leftrightarrow \quad \Delta \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \Delta_1 \\ \vdots \\ \Delta_n \end{pmatrix} \quad \Leftrightarrow \quad \Delta\, x_i = \Delta_i, \ \forall i = 1, \ldots, n.$$

La segunda afirmación de la tesis se deduce inmediatamente de la última igualdad. Observar que si $\Delta \neq 0$, entonces de $\hat{A}A = \Delta I_n$ se deduce que $\hat{A}$ es invertible y por lo tanto las ecuaciones $AX = B$ y $\hat{A}AX = \hat{A}B$ son equivalentes. Luego la primer afirmación se deduce también de la última igualdad.

La tercer afirmación se deduce de que vale $\Delta = 0$, y por lo tanto se aplica el corolario 5.3.5. □

5.4. Matriz asociada

Sea $T : V \to W$ una transformación lineal Si $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{C} = \{w_1, \ldots, w_m\}$ son bases ordenadas de V y W respectivamente, entonces la *matriz asociada* a T en las bases $\mathcal{B}$ y $\mathcal{C}$ es la matriz ${}_{\mathcal{C}}[T]_{\mathcal{B}} \in M_{m \times n}$ definida por

$${}_{\mathcal{C}}[T]_{\mathcal{B}} = \left[\operatorname{coord}_{\mathcal{C}}\big(T(v_1)\big)|\cdots|\operatorname{coord}_{\mathcal{C}}\big(T(v_n)\big)\right].$$

Explícitamente, si

$$\begin{aligned}\operatorname{coord}_{\mathcal{C}}\big(T(v_1)\big) &= a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m\\ &\ \ \vdots\\ \operatorname{coord}_{\mathcal{C}}\big(T(v_n)\big) &= a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m,\end{aligned}$$

entonces

$${}_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} a_{11} & \cdots & a_{1n}\\ a_{21} & \cdots & a_{2n}\\ \vdots & & \vdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix}.$$

Escribiendo un vector arbitrario v como combinación lineal de la base $\mathcal{B}$ y usando la linealidad de T, se deduce que vale

$$\operatorname{coord}_{\mathcal{C}}\big(T(v)\big) = {}_{\mathcal{C}}[T]_{\mathcal{B}}\operatorname{coord}_{\mathcal{B}}(v), \tag{5.14}$$

para todo $v \in V$. En la fórmula anterior, $\operatorname{coord}_{\mathcal{B}}(v)$ y $\operatorname{coord}_{\mathcal{C}}\big(T(v)\big)$ están pensados como vectores columna.

Observación 5.4.1. Si consideramos la transformación lineal identidad $\mathrm{Id} : V \to V$, y $\mathcal{B}$ y $\mathcal{C}$ son dos bases de V, entonces ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$ es la matriz de cambio de base que estudiamos en la sección 4.6.

Observación 5.4.2. Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal, $\mathcal{B} \subset \mathbb{R}^n$ y $\mathcal{C} \subset \mathbb{R}^m$ son las bases canónicas y $A = {}_{\mathcal{C}}[T]_{\mathcal{B}}$, entonces la fórmula (5.14) implica

$$T(v) = Av,$$

para todo $v \in V$. Esta es la fórmula de la proposición 5.1.6.

Por otro lado, si $T : V \to W$ es una transformación lineal, $\mathcal{B} \subset V$ y $\mathcal{C} \subset W$ son bases y $A = {}_{\mathcal{C}}[T]_{\mathcal{B}}$, entonces la fórmula (5.14) se puede escribir

$$\operatorname{coord}_{\mathcal{C}}\big(T(v)\big) = L_A\left(\operatorname{coord}_{\mathcal{B}}(v)\right),$$

para todo $v \in V$. Esta fórmula nos dice que si identificamos un vector con sus coordenadas en la base correspondiente ($v \in V$ con $\operatorname{coord}_{\mathcal{B}}(v)$ y $w \in W$ con $\operatorname{coord}_{\mathcal{C}}(w)$), entonces la transformación lineal $T : V \to W$ se identifica con la transformación lineal $L_A : \mathbb{R}^n \to \mathbb{R}^m$, siendo $A = {}_{\mathcal{C}}[T]_{\mathcal{B}}$.

Ejemplos 5.4.3.

1. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por
$$T(x, y) = (x + y, x + 2y, 2x + y). \tag{5.15}$$
Consideremos la base $\mathcal{B} = \{(1, 1), (1, -1)\}$ y la base canónica $\mathcal{C} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. Luego
$$T(1, 1) = (2, 3, 3) \quad \text{y} \quad T(1, -1) = (0, -1, 1) \quad \Rightarrow \quad {}_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} 2 & 0\\ 3 & -1\\ 3 & 1\end{pmatrix}.$$
Observar que en este caso el cálculo fue fácil porque al ser $\mathcal{C}$ la base canónica, es $\operatorname{coord}_{\mathcal{C}}(2, 3, 3) = (2, 3, 3)$ y $\operatorname{coord}_{\mathcal{C}}(0, -1, 1) = (0, -1, 1)$.

2. Consideremos la transformación lineal del ejercicio anterior $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por la fórmula (5.15). Sean las bases $\mathcal{B} = \{(1, 1), (1, -1)\}$ de $\mathbb{R}^2$ y $\mathcal{C} = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ de $\mathbb{R}^3$. Calculando las coordenadas de un vector genérico (x, y, z) en las base $\mathcal{C}$, obtenemos
$$(x, y, z) = z(1, 1, 1) + (y - z)(1, 1, 0) + (x - y)(1, 0, 0) \quad \Rightarrow \quad \operatorname{coord}_{\mathcal{C}}(x, y, z) = (z,\ y - z,\ x - y).$$
Luego
$$\operatorname{coord}_{\mathcal{C}}\big(T(1, 1)\big) = \operatorname{coord}_{\mathcal{C}}(2, 3, 3) = (3, 0, -1), \quad \operatorname{coord}_{\mathcal{C}}\big(T(1, -1)\big) = \operatorname{coord}_{\mathcal{C}}(0, -1, 1) = (1, -2, 1).$$
Entonces la matriz asociada es ${}_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 1\\ 0 & -2\\ -1 & 1\end{pmatrix}$.

3. Consideremos las bases $\mathcal{B} = \{(1,0),\ (1,1)\}$ de $\mathbb{R}^2$ y $\mathcal{C} = \{(1,1,1),\ (1,1,0),\ (1,0,0)\}$ de R^3. Supongamos que $T : \mathbb{R}^2 \to \mathbb{R}^3$ es una transformación lineal de la cual sabemos que su matriz asociada en estas bases es
$$_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
Nos interesa saber cómo obtener T. Lo primero es hallar las coordenadas de un vector genérico (x, y) en la base $\mathcal{B}$. En este caso es
$$\begin{aligned} (x,y) = (x-y)(1,0) + y(1,1) \quad &\Rightarrow \quad \operatorname{coord}_{\mathcal{B}}(x,y) = (x-y,\ y) \\ \Rightarrow \quad \operatorname{coord}_{\mathcal{C}}\big(T(x,y)\big) &= \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x-y \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \\ x-y \end{pmatrix} \\ \Rightarrow \quad T(x,y) = (x+y)(1,1,1) + y(1,1,0) &+ (x-y)(1,0,0) = (2x+y,\ x+2y,\ x+y). \end{aligned}$$

Luego la transformación lineal T está definida por $T(x,y) = (2x+y,\ x+2y,\ x+y)$, para todo $(x,y) \in \mathbb{R}^2$.

Proposición 5.4.4. *Sean $T : U \to V$ y $S : V \to W$ dos transformaciones lineales. Consideremos bases $\mathcal{B} \subset U$, $\mathcal{C} \subset V$ y $\mathcal{D} \subset W$. Entonces vale*
$$_{\mathcal{D}}[S \circ T]_{\mathcal{B}} = {}_{\mathcal{D}}[T]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[S]_{\mathcal{B}}.$$

Dem. Es análoga a la prueba de la proposición 5.1.12. □

El siguiente resultado muestra cómo se relacionan las matrices asociadas a una misma transformación lineal en bases distintas.

Corolario 5.4.5. *Sea $T : V \to W$ una transformación lineal. Sean $\mathcal{B}_1, \mathcal{B}_2$ bases de V y $\mathcal{C}_1, \mathcal{C}_2$ bases de W. Entonces*
$$_{\mathcal{C}_2}[T]_{\mathcal{B}_2} = {}_{\mathcal{C}_2}[\mathrm{Id}]_{\mathcal{C}_1} \cdot {}_{\mathcal{C}_1}[T]_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}[\mathrm{Id}]_{\mathcal{B}_2}\ ,$$
siendo $_{\mathcal{C}_2}[\mathrm{Id}]_{\mathcal{C}_1}$ y $_{\mathcal{B}_1}[\mathrm{Id}]_{\mathcal{B}_2}$ las matrices de cambio de base. □

Proposición 5.4.6. *Sea $T : V \to V$ una transformación lineal y $\mathcal{B}, \mathcal{C}$ bases de V. Entonces T es un isomorfismo si y solo si la matriz $_{\mathcal{C}}[T]_{\mathcal{B}}$ es invertible. En ese caso vale*
$$_{\mathcal{B}}\left[T^{-1}\right]_{\mathcal{C}} = \left({}_{\mathcal{C}}[T]_{\mathcal{B}}\right)^{-1}.$$

Dem. Recordar el corolario 5.2.17 y la observación 5.4.2. □

Observación 5.4.7. Notar que la proposición 4.6.6 es un caso particular de la proposición anterior.

Bibliografía

[F] Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence, *Linear Algebra*, Prentice Hall, 1999.

[RH] Robert Resnick, David Halliday, *Física. Parte I*, Compañía Editorial Continental S.A, 1975.