Álgebra Lineal y Geometría I

Centro de Matemática
Facultad de Ciencias
Universidad de la República

Andrés Abella

25 de junio de 2022

Índice

Notaciones. El fin de cada demostración lo señalaremos con el símbolo $\square$. Para indicar en una prueba que hemos llegado a una contradicción, usaremos el símbolo $\lightning$. A veces en las definiciones usaremos el símbolo ":=", por ejemplo en $a(x, y) := (ax, ay)$, queriendo decir que la definición del producto $a(x, y)$ es (ax, ay). El símbolo $\#A$ representa la cantidad de elementos de un conjunto A, llamado el *cardinal* de A.

1. Sistemas de ecuaciones

En este tema el énfasis va a estar en la parte práctica. La fundamentación teórica se verá más adelante en la sección 5.4, después de estudiar transformaciones lineales.

En general un *sistema de ecuaciones (lineales)* es un conjunto de ecuaciones, como son los siguientes

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

El primero es un ejemplo de un sistema 2×3, ya que tiene dos ecuaciones y tres incógnitas (o variables), el segundo es 3×2, el tercero es 2×2 y el cuarto es 3×3. En general un sistema puede tener una cantidad arbitraria de ecuaciones y de incógnitas. Una *solución* del sistema es una asignación de valores a las variables, de forma tal que se verifiquen todas las ecuaciones del sistema. Por ejemplo, si consideramos el primer sistema en (1), entonces una solución es $x = -1$, $y = 1$ y $z = -4$. Cuando nos referimos a *resolver* el sistema queremos decir hallar todas sus soluciones, si es que existen.

1.1. Sistemas 2×2

En esta sección veremos cómo resolver sistemas 2×2. Para eso tenemos tres métodos: *sustitución*, *igualación* y *reducción*. Mostraremos solo sustitución y reducción, ya que igualación es bastante parecido a sustitución y no tiene mayores ventajas. Lo ilustraremos mediante ejemplos. Consideremos

$$\left\{\begin{array}{rcl} 2x+y &=& 3 \\ 3x-5y &=& 11 \end{array}\right. .$$

Notar que en la primera ecuación es fácil despejar la segunda variable, obteniendo $y = 3-2x$. Si sustituimos este valor en la segunda ecuación, obtenemos una ecuación en una variable, la cual es fácil de resolver

$$3x - 5(3-2x) = 11 \quad \Rightarrow \quad 13x - 15 = 11 \quad \Rightarrow \quad 13x = 26 \quad \Rightarrow \quad x = 2.$$

Ahora el valor que obtuvimos de x lo sustituimos en $y = 3 - 2x$, obteniendo $y = 3 - 2(2) = -1$. Luego la solución es $x = 2$ e $y = -1$. Este es el método de *sustitución*.

El método anterior funciona bien cuando alguna de las variables es fácil de despejar, en otros casos es mejor el método de *reducción*, que veremos a continuación. Consideremos el siguiente sistema

$$\left\{\begin{array}{rcl} 7x+2y &=& 16 \\ 3x+5y &=& 11 \end{array}\right. . \tag{2}$$

Notar que, en general, si multiplicamos una ecuación por una constante no nula, entonces obtenemos otra ecuación que tiene las mismas soluciones. La técnica acá consiste en multiplicar cada ecuación por una constante adecuada y luego sumarlas, de forma tal que en el resultado final desaparezca una de las incógnitas. Por ejemplo, si en el sistema anterior queremos eliminar la incógnita x, entonces mutiplicando la primer ecuación por 3 y la segunda por -7, y después sumándolas, obtenemos

$$\left\{\begin{array}{rcl} 21x+6y &=& 48 \\ -21x+-35y &=& -77 \end{array}\right. \overset{(+)}{\Longrightarrow} \quad -29y = -29 \quad \Rightarrow \quad y = 1.$$

Si ahora sustituimos ese valor de y en la primera ecuación del sistema (2), obtenemos

$$7x + 2(1) = 16 \quad \Rightarrow \quad 7x + 2 = 16 \quad \Rightarrow \quad 7x = 14 \quad \Rightarrow \quad x = 2.$$

Luego la solución es $x = 2$ e $y = 1$. Lo mismo obtendríamos si sustituimos y en la segunda ecuación.
Otra forma de resolver el sistema consiste en eliminar la variable y en vez de la x. Para eso en (2) mutiplicamos la primer ecuación por 5 y la segunda por -2 y las sumamos:

$$\left\{\begin{array}{rcl} 35x+10y &=& 80 \\ -6x+-10y &=& -22 \end{array}\right. \quad\Rightarrow\quad 29x=58 \quad\Rightarrow\quad x=2.$$

Para obtener y, sustituimos el valor de x en cualquiera de las ecuaciones del sistema. Por ejemplo, sustituyendo en la primera obtenemos

$$7(2)+2y=16 \quad\Rightarrow\quad 14+2y=16 \quad\Rightarrow\quad 2y=2 \quad\Rightarrow\quad y=1,$$

así llegamos de nuevo a la solución $x = 2$ e $y = 1$.

Observación 1.1. Introduciendo un sistema de coordenadas podemos identificar el plano con el conjunto $\mathbb{R}^2 = \{(x, y) :\ x, y \in \mathbb{R}\}$, siendo $\mathbb{R}$ el conjunto de los números reales. Pensado $\mathbb{R}^2$ de esta forma, los puntos (x, y) que verifican una ecuación del tipo $ax + by = c$ (con a y b no nulos simultáneamente) son puntos que están alineados en una recta r; en ese caso decimos que $ax + by = c$ es *la ecuación* de r. Luego cuando queremos resolver un sistema del tipo

$$\left\{\begin{array}{rcl} ax+by &=& c \\ dx+ey &=& f \end{array}\right.$$

lo que estamos haciendo es buscar los puntos comunes a dos rectas, es decir hallar su intersección. Como dos rectas pueden ser secantes, paralelas (y distintas), o coincidentes, entonces el sistema anterior puede tener una solución, ninguna, o infinitas, respectivamente. Ejemplos de lo anterior son los siguientes, respectivamente.

$$\left\{\begin{array}{rcl} x+y &=& 1 \\ x-y &=& 1 \end{array}\right., \qquad \left\{\begin{array}{rcl} x+y &=& 1 \\ x+y &=& -1 \end{array}\right., \qquad \left\{\begin{array}{rcl} x+y &=& 1 \\ 2x+2y &=& 2 \end{array}\right..$$

Aplicación 1.2. Consideremos dos cuerdas sujetas en puntos distintos de un techo, que están unidas mediante un nudo a una tercer cuerda de la cual pende un objeto. Sabemos que una de estas dos cuerdas forma un ángulo de 30º con el techo y la otra forma uno de 45º. Asumimos que el peso de las cuerdas es despreciable y que no tienen estiramiento. Si nos concentramos en el nudo que une las tres cuerdas, tenemos que hay una fuerza $\overrightarrow{F_c}$ actuando hacia abajo (dada por el peso del objeto) que es contrarrestada por las fuerzas que actúan en las cuerdas que van al techo, que llamaremos $\overrightarrow{F_a}$ y $\overrightarrow{F_b}$. Lo que nos interesa es saber cómo se relacionan las magnitudes F_a, F_b y F_c de estas fuerzas. Como el objeto está en reposo, es $\overrightarrow{F_a}+\overrightarrow{F_b}+\overrightarrow{F_c} = \overrightarrow{0}$ (lo cual equivale a $\overrightarrow{F_a}+\overrightarrow{F_b} = -\overrightarrow{F_c}$). Notar que las tres fuerzas están en el plano que contiene a las tres cuerdas. Si ponemos un sistema de ejes en ese plano de forma tal que el origen esté en el nudo y fijamos dos versores[1] $\hat\imath$ y $\hat\jmath$ en las direcciones de los ejes, entonces obtenemos la figura 1. Descomponiendo $\overrightarrow{F_a}$, $\overrightarrow{F_b}$ y $\overrightarrow{F_c}$ en sus componentes según $\hat\imath$ y $\hat\jmath$ obtenemos

$$\begin{array}{l} \overrightarrow{F_a} = -F_a\cos(30^\circ)\hat\imath + F_a\,\mathrm{sen}(30^\circ)\hat\jmath \\ \overrightarrow{F_b} = F_b\cos(45^\circ)\hat\imath + F_b\,\mathrm{sen}(45^\circ)\hat\jmath \\ \overrightarrow{F_c} = -F_c\hat\jmath \end{array} \quad\Rightarrow\quad \begin{array}{l} \overrightarrow{F_a} = -\frac{\sqrt{3}}{2}F_a\,\hat\imath + \frac{1}{2}F_a\,\hat\jmath \\ \overrightarrow{F_b} = \frac{\sqrt{2}}{2}F_b\,\hat\imath + \frac{\sqrt{2}}{2}F_b\,\hat\jmath \\ \overrightarrow{F_c} = -F_c\,\hat\jmath \end{array}.$$

Imponiendo $\overrightarrow{F_a}+\overrightarrow{F_b}+\overrightarrow{F_b} = \overrightarrow{0}$ obtenemos

$$\left(-\frac{\sqrt{3}}{2}F_a + \frac{\sqrt{2}}{2}F_b\right)\hat\imath + \left(\frac{1}{2}F_a + \frac{\sqrt{2}}{2}F_b - F_c\right)\hat\jmath = \overrightarrow{0}.$$

[1] Recordar que un *versor* es un vector de módulo uno; en física para los versores se suele usar la notación $\hat v$ en vez de $\vec v$.

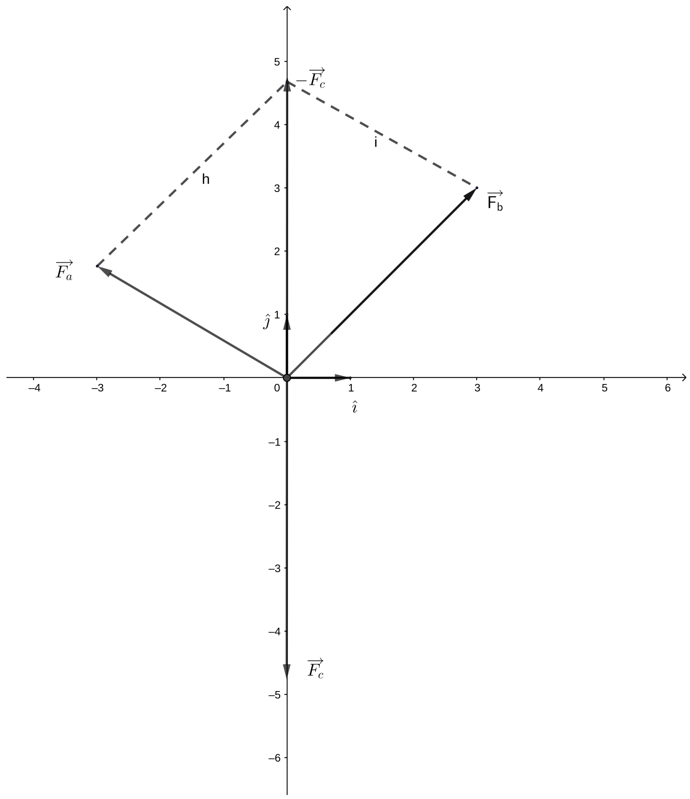

Figura 1: Diagrama de fuerzas

Para que un vector sea nulo sus componentes tienen que ser nulas, luego

$$\left\{ \begin{array}{r} -\frac{\sqrt{3}}{2}F_a + \frac{\sqrt{2}}{2}F_b = 0 \\ \frac{1}{2}F_a + \frac{\sqrt{2}}{2}F_b - F_c = 0 \end{array} \right. .$$

Este es un sistema de dos ecuaciones y tres incógnitas F_a, F_b y F_c. En este caso, si fijamos una de las incógnitas, podemos despejar las otras dos. Por ejemplo, en general lo que más interesa es conocer F_a y F_b sabiendo F_c. Para eso, podemos escribir el sistema anterior de la forma siguiente

$$\left\{ \begin{array}{r} -\frac{\sqrt{3}}{2}F_a + \frac{\sqrt{2}}{2}F_b = 0 \\ \frac{1}{2}F_a + \frac{\sqrt{2}}{2}F_b = F_c \end{array} \right.$$

y considerarlo como un sistema de dos ecuaciones con dos incógnitas. Resolviendo el sistema obtenemos

$$F_a = \frac{2}{1+\sqrt{3}}\,F_c, \quad F_b = \frac{\sqrt{6}}{1+\sqrt{3}}\,F_c \quad \Rightarrow \quad F_a \simeq 0,73\,F_c, \quad F_b \simeq 0,90\,F_c. \tag{3}$$

Luego F_a es un 73 % de F_c y F_b es un 90 % de F_c.

Las fórmulas (3) nos permiten obtener F_a y F_b a partir de F_c. Pero también podemos obtener cualquiera de estas fuerzas en función de las restantes. Por ejemplo, despejando F_c y F_b en (3) obtenemos

$$F_c = \frac{1+\sqrt{3}}{2}F_a, \quad F_b = \frac{\sqrt{6}}{2}F_a.$$

Queda como ejercicio obtener F_a y F_c en función de F_b.

1.2. Sistemas en general

En general un *sistema de ecuaciones lineales* es un conjunto de ecuaciones de la forma

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n & = & b_2 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right. . \tag{4}$$

Los números a_{ij}, con $i = 1, \ldots, m$, $j = 1, \ldots, n$, son los *coeficientes*, $b_1, \ldots, b_m$ son los *términos independientes* y $x_1, \ldots, x_n$ 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 hay pocas variables, se suele escribir x, y, z, t en vez de x_1, x_2, x_3, x_4.

Una *solución* del sistema (4) es una n-upla[2] de números $(\alpha_1, \ldots, \alpha_n)$ tales que si en (4) 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* si tiene más de una solución. Más adelante en la sección 5.4 veremos que un sistema indeterminado siempre tiene infinitas soluciones. 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 primera ecuación deducimos que es $x = -1$. Luego el sistema es compatible determinado y su solución es $x = -1$, $y = 1$ y $z = 2$.

En general, los sistemas de ecuaciones que más aparecen son los *cuadrados*, que son los que tienen tantas ecuaciones como variables. Nosotros estudiaremos principalmente este tipo de sistemas de ecuaciones y al final de la sección veremos un par de ejemplos en que la cantidad de variables y de ecuaciones es distinta.

Dos sistemas de ecuaciones se dicen *equivalentes* si tienen las mismas soluciones[3]. Usaremos el símbolo $\sim$ para indicar la equivalencia de sistemas. Vale el siguiente resultado.

Proposición 1.3. *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.*

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

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

Dem. Las tres primeras propiedades son evidentes, así que solo probaremos la última. Para simplificar la notación, haremos la prueba para el caso de sistemas 2×2, aunque la misma vale para sistemas $m \times n$ en general. Consideremos un sistema 2×2 genérico

$$\left\{ \begin{array}{rcl} ax + by & = & p \\ cx + dy & = & q \end{array} \right. . \tag{5}$$

y el sistema obtenido sumándole a la segunda ecuación, la primera multiplicada por una constante k:

$$\left\{ \begin{array}{rcl} ax + by & = & p \\ (c + ka)x + (d + kb)y & = & q + kp \end{array} \right. . \tag{6}$$

Si (x_0, y_0) es una solución del sistema (5), entonces se verifican las igualdades siguientes

$$\left\{ \begin{array}{rcl} ax_0 + by_0 & = & p \\ cx_0 + dy_0 & = & q \end{array} \right. .$$

Sumándole a la segunda igualdad la primera multiplicada por k, obtenemos

$$cx_0 + dy_0 + k(ax_0 + by_0) = q + kp \quad \Rightarrow \quad (c + ka)x_0 + (d + kb)y_0 = q + kp.$$

Luego vale

$$\left\{ \begin{array}{rcl} ax_0 + by_0 & = & p \\ (c + ka)x_0 + (d + kb)y_0 & = & q + kp \end{array} \right. .$$

Esto nos dice que toda solución del sistema (5) es también una solución del sistema (6). Pero si ahora en el sistema (6) le sumamos a la segunda ecuación la primera multiplicada por $-k$ volvemos a obtener el sistema (5). Luego el razonamiento anterior implica que toda solución del sistema (6) es también una solución del sistema (5). En definitiva, los dos sistemas tienen las mismas soluciones y por lo tanto son equivalentes. $\square$

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 mediante los siguientes ejemplos.

Ejemplo 1.4. Queremos resolver el siguiente sistema

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

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

$$\left\{ \begin{array}{rcl} 2x + 3y + 5z & = & 9 \\ x + 2y + 3z & = & 5 \\ 3x + 4y + 5z & = & 5 \end{array} \right. \sim \left\{ \begin{array}{rcl} x + 2y + 3z & = & 5 \\ 2x + 3y + 5z & = & 9 \\ 3x + 4y + 5z & = & 5 \end{array} \right. \sim \left\{ \begin{array}{rcl} x + 2y + 3z & = & 5 \\ -y - z & = & -1 \\ 3x + 4y + 5z & = & 5 \end{array} \right. \sim$$
$$\left\{ \begin{array}{rcl} x + 2y + 3z & = & 5 \\ -y - z & = & -1 \\ -2y - 4z & = & -10 \end{array} \right. \sim \left\{ \begin{array}{rcl} x + 2y + 3z & = & 5 \\ y + z & = & 1 \\ y + 2z & = & 5 \end{array} \right. \sim \left\{ \begin{array}{rcl} x + 2y + 3z & = & 5 \\ y + z & = & 1 \\ z & = & 4 \end{array} \right. .$$

En el primer paso, intercambiamos la primera 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 por -1 y la tercera por $-1/2$; en el quinto, a la tercera ecuación le sumamos la segunda multiplicada por -1. Al final obtuvimos un sistema escalerizado, cuya única solución es $x = -1$, $y = -3$ y $z = 4$. Luego el sistema original es compatible determinado.

Ejemplo 1.5.

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

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.

Ejemplo 1.6.

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

Observar que los coeficientes de este sistema son los mismos que los del sistema del ejemplo anterior, así que los pasos de resolución son los mismos (solo cambian los términos independientes). Como la última ecuación no aporta nada, la podemos eliminar obteniendo que el sistema original es equivalente al siguiente

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

La última ecuación tiene infinitas soluciones (es la ecuación de una recta en el plano y, z). Por ejemplo, si escribimos $z = \lambda$, con $\lambda \in \mathbb{R}$ arbitrario, entonces despejando y obtenemos que $z = \lambda$ e $y = -\lambda - \frac{4}{7}$ es una solución de la última ecuación. Sustituyendo estos valores en la primera y despejando obtenemos $x = -\lambda + \frac{1}{7}$. Luego para cada valor de $\lambda \in \mathbb{R}$ obtenemos que

$$x = -\lambda + \frac{1}{7}, \quad y = -\lambda - \frac{4}{7}, \quad z = \lambda$$

es una solución del sistema original. Este es un ejemplo de un sistema compatible indeterminado con *un grado de libertad*, dado que la variable z está libre y fijado este valor los otros quedan determinados. Notar que otra manera de escribir las soluciones es $x = -z + \frac{1}{7}$, $y = -z - \frac{4}{7}$ y la variable z está libre.

Ejemplo 1.7.

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

Acá a la segunda ecuación le sumamos la primera y a la tercera le sumamos la primera multiplicada por -2. Luego el sistema quedó equivalente a la ecuación $x+y-z=1$ que claramente tiene infinitas soluciones. Por ejemplo si tomamos $y = \lambda$ y $z = \mu$, entonces las soluciones son

$$x = -\lambda + \mu + 1, \quad y = \lambda, \quad z = \mu; \quad \lambda, \mu \in \mathbb{R}.$$

Otra forma de escribir las soluciones es $z = x + y + 1$ y las variables x e y están libres. Este es un ejemplo de un sistema compatible indeterminado con *dos grados de libertad*. En sistemas con más variables, pueden aparecer soluciones que tengan más grados de libertad.

A continuación veremos un par de ejemplos de aplicación del método de escalerización para resolver sistemas que no son cuadrados.

Ejemplo 1.8. 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 1.9. 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 $y = 0$.

1.3. 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 términos independientes por una columna de ceros). Como las operaciones para resolverlos son las mismas, solo describiremos el camino que lleva a su resolución.

Ejemplo 1.10.

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

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

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

Ejemplo 1.11.

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

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

Ejemplo 1.12.

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

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

Ejemplo 1.13. 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 se pueden escribir de la forma $x = -3z$, $y = 2z$, con z libre.

Ejemplo 1.14. 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 $x = y = 0$.

Ejemplo 1.15. En la aplicación 1.2 obtuvimos el siguiente sistema homogéneo

$$\left\{\begin{array}{r} -\frac{\sqrt{3}}{2}F_a + \frac{\sqrt{2}}{2}F_b = 0 \\ \frac{1}{2}F_a + \frac{\sqrt{2}}{2}F_b - F_c = 0 \end{array}\right. .$$

Cuando lo resolvimos obtuvimos que sus soluciones son $F_a = \frac{2}{1+\sqrt{3}}\,F_c$ y $F_b = \frac{\sqrt{6}}{1+\sqrt{3}}\,F_c$, con F_c libre. Luego este sistema es compatible indeterminado con un grado de libertad. Esto nos dice que sabiendo una de las fuerzas, esta determina a las otras dos.

2. Geometría

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

2.1. Vectores en el plano

En lo que sigue, mediante la introducción de un sistema de coordenadas ortogonales, identificamos el plano con el conjunto $\mathbb{R}^2 = \{(x, y) : \ x, y \in \mathbb{R}\}$. Así, si por ejemplo escribimos $P = (1, 2)$, estamos diciendo que $(1, 2)$ son las coordenadas del punto P respecto al sistema de referencia del plano. Al punto $O = (0, 0)$ se le llama el *origen*. En general usaremos las letras mayúsculas P, Q, R para referirnos a los puntos del plano.

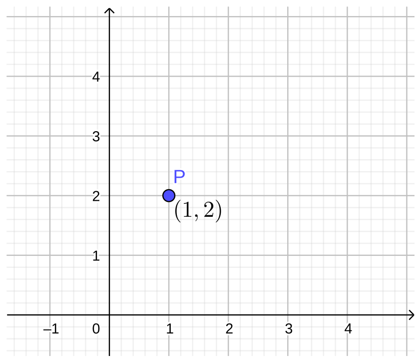

Figura 2: Punto

A cada punto P del plano, con $P \neq O$, le podemos hacer corresponder el vector $v = \overrightarrow{OP}$, que es la flecha que va de O a P; al origen O le hacemos corresponder el *vector nulo*. De esa forma obtenemos una correspondencia uno a uno entre los puntos y los vectores flecha; por eso a los puntos del plano también les llamamos *vectores*. Así, si por ejemplo escribimos el vector $v = (1, 2)$, estamos diciendo que es $v = \overrightarrow{OP}$, con $P = (1, 2)$. Cuando a los puntos los pensamos como vectores, usamos las letras minúsculas u, v, w; el vector nulo es $o = (0, 0)$.

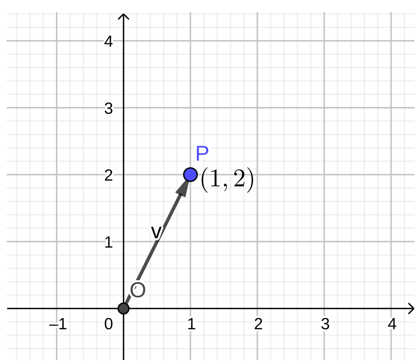

Figura 3: Vector

Al trabajar con puntos o vectores, a los números se les suele llamar *escalares* y en general los escribiremos con las letras a, b, c. En estas primeras secciones vamos a trabajar principalmente con vectores, pero después usaremos puntos y vectores.

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

$$a(x_0, y_0) := (ax_0, ay_0), \quad (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2),$$

para todo $a \in \mathbb{R}$, $(x_0, y_0), (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$. Si $v = (x, y)$, entonces su *opuesto* es $-v := (-x, -y)$.

La proposición siguiente muestra que la suma anterior coincide con la suma de vectores obtenida por la *regla del paralelogramo* (la que se aprende en los cursos de física).

Proposición 2.1. *Consideremos un paralelogramo de vértices O, P, R, Q (figura 4). Si escribimos $P = (x_1, y_1)$, $Q = (x_2, y_2)$ y $R = (x_3, y_3)$, entonces $x_3 = x_1 + x_2$ y $y_3 = y_1 + y_2$.*

Dem. Sea $H = (x_3, y_2)$ la intersección de la paralela al eje Ox por Q con la paralela al eje Oy por R, y sea $K = (x_1, 0)$ la intersección de la paralela al eje Oy por P con el eje Ox. Consideremos los triángulos OKP y QHR (ver figura 4). Los ángulos POK y RQH son iguales (son ángulos entre paralelas) y los ángulos PKO y RHQ son rectos, por lo que también son iguales. Como además la longitud del lado OP es igual a la del lado QR, concluimos que los triángulos OKP y QHR son congruentes[4]. Luego la longitud del lado OK coincide con la del QH y la del KP con la del HR. Si calculamos esas longitudes en función de las coordenadas de los puntos, obtenemos $x_1 = x_3 - x_2$ e $y_1 = y_3 - y_2$, lo cual implica $x_3 = x_1 + x_2$ e $y_3 = y_1 + y_2$. $\square$

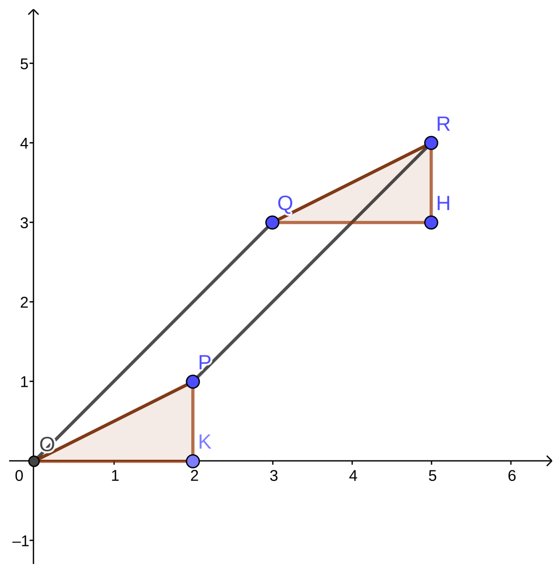

Figura 4: Suma de vectores

También es fácil de probar que el producto por un escalar por un vector coincide con el de física.

[4]Que sean *congruentes* es que tengan igual forma y tamaño.

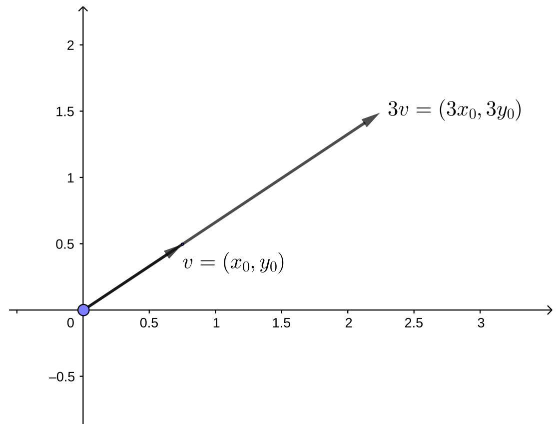

Figura 5: Producto de un escalar por un vector

Proposición 2.2. *Las operaciones recién definidas verifican las siguientes propiedades:*

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* $a, b \in \mathbb{R}$, $u \in \mathbb{R}^2$.

Dem. Probaremos solo la propiedad asociativa de la suma, dejando las otras pruebas como ejercicio. Sean $u = (x_1, y_1)$, $v = (x_2, y_2)$ y $w = (x_3, y_3)$. Entonces

$$u + (v + w) = (x_1, y_1) + \big((x_2, y_2) + (x_3, y_3)\big) = (x_1, y_1) + (x_2 + x_3, y_2 + y_3) = \big(x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)\big),$$
$$[u + v] + w = \big((x_1, y_1) + (x_2, y_2)\big) + (x_3, y_3) = (x_1 + x_2, y_1 + y_2) + (x_3, y_3) = \big((x_1 + x_2) + x_3, (y_1 + y_2) + y_3\big).$$

Como la suma de números reales es asociativa, entonces valen

$$x_1 + (x_2 + x_3) = (x_1 + x_2) + x_3 \quad \text{y} \quad y_1 + (y_2 + y_3) = (y_1 + y_2) + y_3.$$

Luego $u + (v + w) = (u + v) + w$. $\square$

Observaciones 2.3.

1. Notar que la prueba anterior de la propiedad asociativa es simple, sin embargo probarla usando la regla del paralelogramo parece bastante más complicado.
2. Las propiedades anteriores caracterizan a los *espacios vectoriales*, que veremos más adelante.

Dos vectores u y v se dicen *colineales* o *paralelos* si existe un escalar a tal que $u = av$ o $v = au$. Notar que si es $u = av$ con $a \neq 0$, entonces es $v = a^{-1}u$. Si u y v son no nulos y es $a > 0$ entonces u y v tienen el *mismo sentido* y tienen *sentidos opuestos* si $a < 0$.

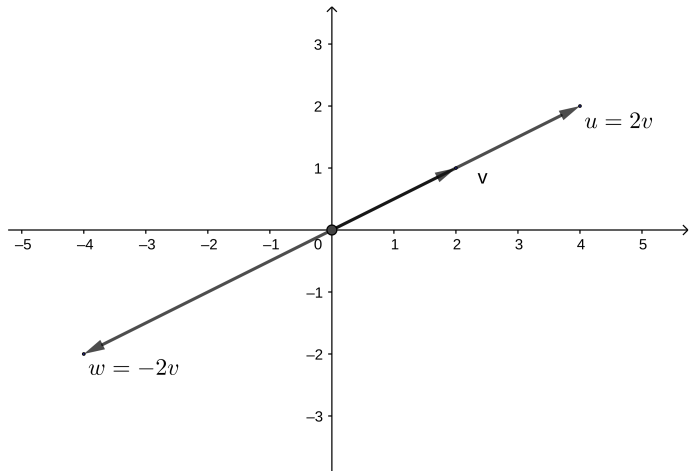

Figura 6: Vectores colineales

El vector nulo o es colineal con cualquier otro vector ($o = 0v$, para todo v); es el único vector que lo verifica.

Si $v = (x_0, y_0)$ es un vector, definimos su *norma* o *módulo* $\|v\|$ mediante $\|v\| := \sqrt{x_0^2 + y_0^2}$. La norma mide la longitud del vector (aplicar el *teorema de Pitágoras* al triángulo de vértices $(0,0)$, $(x_0, 0)$ y (x_0, y_0)).

Proposición 2.4. *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. Para la última, observar que si se hace un dibujo con los vectores u, v y $u + v$, entonces la desigualdad triangular se deduce 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). □

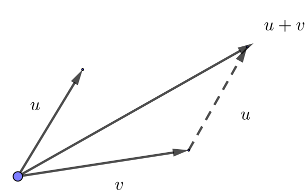

Figura 7: Desigualdad triangular

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 $w = u - v$ si y solo si $u = v + w$.

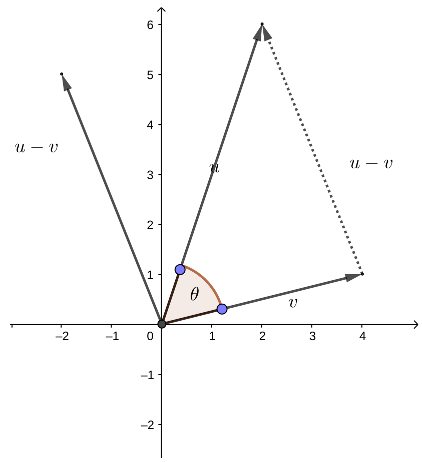

Figura 8: Resta de vectores

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 $\mathcal{B} = \{\hat{\imath}, \hat{\jmath}\}$, donde $\hat{\imath} = (1, 0)$ y $\hat{\jmath} = (0, 1)$. Notar que vale

$$v = (x, y) \quad \Leftrightarrow \quad v = x\hat{\imath} + y\hat{\jmath}.$$

Si es $v = x\hat{\imath} + y\hat{\jmath}$, entonces se suele decir que x e y son las *componentes* de v en las direcciones de $\hat{\imath}$ y $\hat{\jmath}$, respectivamente.

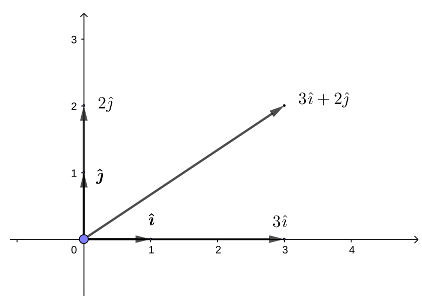

Figura 9: Base canónica

Notar que la expresión $v = x\hat{\imath} + y\hat{\jmath}$ es única, es decir, vale $x_1\hat{\imath} + y_1\hat{\jmath} = x_2\hat{\imath} + y_2\hat{\jmath}$ si y solo si $x_1 = x_2$ y $y_1 = y_2$.

Un *versor* es un vector de norma 1. Los vectores $\hat{\imath}$ y $\hat{\jmath}$ son versores. Los versores se suelen usar para expresar direcciones: si v es un vector no nulo, entonces $\frac{v}{\|v\|} := \frac{1}{\|v\|}v$ es un versor, que es colineal con v y tiene el mismo sentido que v.

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

$$u \cdot v := \|u\|\|v\| \cos(\theta), \tag{7}$$

siendo $0 \leq \theta \leq \pi$ el ángulo[5] que forman u y v (ver figura 8). Si u o v es nulo, se define $u \cdot v = 0$.

[5] Los ángulos los mediremos siempre en radianes.

Teorema 2.5. *Si* $u = (x_1, y_1)$ *y* $v = (x_2, y_2)$*, entonces*

$$u \cdot v = x_1 x_2 + y_1 y_2, \tag{8}$$

Dem. Es claro que la fórmula (8) vale cuando u o v es el vector nulo, así que de ahora en más supondremos que ninguno es nulo. Si aplicamos la *fórmula del coseno* al triángulo de lados u y v (figura 8), obtenemos

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\,\|v\|\,\cos(\theta) \quad \Rightarrow \quad \|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2u \cdot v.$$

Luego

$$\begin{aligned} u \cdot v &= \frac{1}{2}\left\{\|u\|^2 + \|v\|^2 - \|u - v\|^2\right\} \\ &= \frac{1}{2}\left\{x_1^2 + y_1^2 + x_2^2 + y_2^2 - \big((x_1 - x_2)^2 + (y_1 - y_2)^2\big)\right\} \\ &= \frac{1}{2}\left\{x_1^2 + y_1^2 + x_2^2 + y_2^2 - \big(x_1^2 - 2x_1x_2 + x_2^2 + y_1^2 - 2y_1y_2 + y_2^2\big)\right\} \\ &= x_1x_2 + y_1y_2. \quad \square \end{aligned}$$

Notar que en términos de la base canónica, la fórmula (8) se escribe:

$$(x_1\hat{\imath} + y_1\hat{\jmath}) \cdot (x_2\hat{\imath} + y_2\hat{\jmath}) = x_1x_2 + y_1y_2.$$

Proposición 2.6. *El producto escalar verifica las siguientes propiedades.*

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

Dem. Ejercicio (usar la fórmula (8)). $\square$

Observación 2.7. La propiedad conmutativa implica que las propiedades 2 y 3 valen también "del otro lado":

$$u \cdot (v + w) = u \cdot v + u \cdot w, \quad \forall u, v, w \in \mathbb{R}^2; \qquad u \cdot (av) = a(u \cdot v), \quad \forall a \in \mathbb{R},\ u, v \in \mathbb{R}^2.$$

Teorema 2.8 (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. Tomando valor absoluto en la fórmula (7) obtenemos

$$|u \cdot v| = \big|\|u\|\|v\|\cos(\theta)\big| = \|u\|\|v\|\,|\cos(\theta)| \leq \|u\|\|v\|.$$

Esto prueba la desigualdad de Cauchy-Shwarz. Además la fórmula de arriba implica que vale $|u \cdot v| = \|u\|\,\|v\|$ si y solo si

$$|\cos(\theta)| = 1 \quad \Leftrightarrow \quad \cos(\theta) = \pm 1 \quad \Leftrightarrow \quad \theta = 0 \text{ o } \theta = \pi.$$

Esto último ocurre si y solo si u y v son colineales. $\square$

Observación 2.9. 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

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

Si dos vectores u y v verifican $u \cdot v = 0$, entonces decimos que u y v son *ortogonales* y escribimos $u \perp v$; si u y v son no nulos, esto equivale a que formen un ángulo recto. Los vectores $\hat{\imath}$ y $\hat{\jmath}$ 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 ortogonal* de v en la dirección de u, es el vector $\Pi_u(v)$ definido por

$$\Pi_u(v) := \frac{u \cdot v}{\|u\|^2} u = \frac{u \cdot v}{\|u\|} u_0,$$

siendo $u_0 := \frac{u}{\|u\|}$ un versor colineal con u. La *componente de v en la dirección de u* es el escalar $\frac{u \cdot v}{\|u\|}$.
Si calculamos la norma de $\Pi_u(v)$, obtenemos

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

luego la norma de la proyección coincide con el valor absoluto de la componente. Por otro lado, si calculamos el producto escalar de $v - \Pi_u(v)$ por u, obtenemos

$$\big(v - \Pi_u(v)\big) \cdot u = v \cdot u - \Pi_u(v) \cdot u = v \cdot u - \left(\frac{u \cdot v}{\|u\|^2} u \right) \cdot u = v \cdot u - \left(\frac{u \cdot v}{\|u\|^2} \right) u \cdot u = v \cdot u - u \cdot v = 0,$$

Luego $v - \Pi_u(v)$ es ortogonal a u. Por lo tanto la proyección ortogonal nos permite descomponer 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.

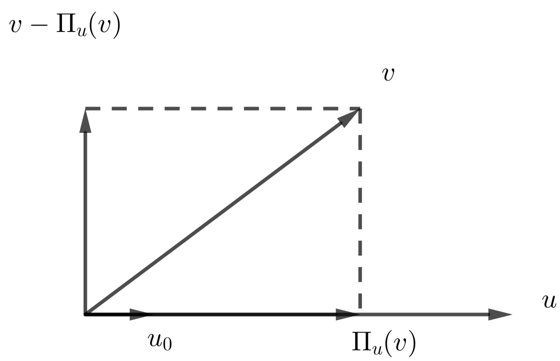

Figura 10: Proyección ortogonal

Ejemplo 2.10. Sean $u = \hat\imath + \hat\jmath$ y $v = 2\hat\imath + 3\hat\jmath$. Entonces

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

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

2.2. Vectores en el espacio

En forma análoga a como identificamos el plano con $\mathbb{R}^2$, al espacio lo identificamos con el conjunto $\mathbb{R}^3 = \{(x, y, z) : \ x, y, z \in \mathbb{R}\}$ mediante la elección de tres ejes perpendiculares que se cortan en un punto O, llamados Ox, Oy y Oz. 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 la proposición 2.2 (el vector nulo de $\mathbb{R}^3$ es $o = (0, 0, 0)$). La *base canónica* de $\mathbb{R}^3$ es el conjunto $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$, donde $\hat{\imath} = (1, 0, 0)$, $\hat{\jmath} = (0, 1, 0)$ y $\hat{k} = (0, 0, 1)$. En la figura siguiente el eje Ox está en rojo, el Oy en verde y el Oz en negro.

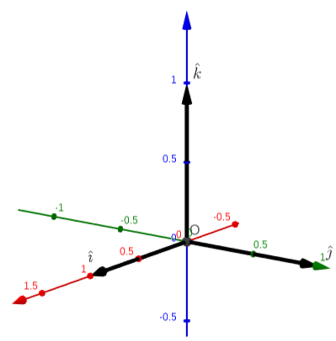

Figura 11: Ejes y base canónica en $\mathbb{R}^3$

Análogamente al caso del plano, es

$$(x, y, z) = x\hat{\imath} + y\hat{\jmath} + z\hat{k}.$$

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

$$(x_1, y_1, z_1) \cdot (x_2, y_2, z_2) = x_1x_2 + y_1y_2 + z_1z_2 \quad \text{o} \quad (x_1\hat{\imath} + y_1\hat{\jmath} + z_1\hat{k}) \cdot (x_2\hat{\imath} + y_2\hat{\jmath} + z_2\hat{k}) = x_1x_2 + y_1y_2 + z_1z_2. \quad (9)$$

La definición de versor, de vectores colineales y de vectores ortogonales es la misma que en $\mathbb{R}^2$ y todas la propiedades que vimos de los vectores en $\mathbb{R}^2$ (desigualdad triangular, Cauchy-Schwarz, etc.) valen también en $\mathbb{R}^3$. La *proyección ortogonal* de un vector v en la dirección de un vector no nulo u, es el vector $\Pi_u(v) = \frac{u \cdot v}{\|u\|^2} u$. Como antes vale $v = \Pi_u(v) + \big(v - \Pi_u(v)\big)$, en que $\Pi_u(v)$ es colineal con u y $v - \Pi_u(v)$ es ortogonal a u.

Ejemplo 2.11. Sean $u = \hat{\imath} + \hat{\jmath}$ y $v = 2\hat{\imath} + \hat{\jmath} + 2\hat{k}$. El coseno del ángulo θ que forman u y v es

$$\cos(\theta) = \frac{u \cdot v}{\|u\|\|v\|} = \frac{3}{\sqrt{2} \times 3} = \frac{1}{\sqrt{2}},$$

por lo tanto $\theta = \pi/4$. La proyección de v en la dirección de u es

$$\Pi_u(v) = \frac{u \cdot v}{\|u\|^2} u = \frac{3}{2}(\hat{\imath} + \hat{\jmath}) = \frac{3}{2}\hat{\imath} + \frac{3}{2}\hat{\jmath}.$$

Esto implica $v = \left(\frac{3}{2}\hat{\imath} + \frac{3}{2}\hat{\jmath}\right) + \left(\frac{1}{2}\hat{\imath} - \frac{1}{2}\hat{\jmath} + 2\hat{k}\right)$, con $\frac{3}{2}\hat{\imath} + \frac{3}{2}\hat{\jmath}$ colineal con v y $\frac{1}{2}\hat{\imath} - \frac{1}{2}\hat{\jmath} + 2\hat{k}$ perpendicular a v.

Observación 2.12. Es usual pensar el plano $\mathbb{R}^2$ dentro del espacio $\mathbb{R}^3$, identificando (x, y) con $(x, y, 0)$. En esa identificación al elemento $(1, 0) \in \mathbb{R}^2$ le corresponde el $(1, 0, 0) \in \mathbb{R}^3$, por lo cual tiene sentido escribir $\hat{\imath} = (1, 0)$ en $\mathbb{R}^2$ y también $\hat{\imath} = (1, 0, 0)$ en $\mathbb{R}^3$; lo mismo sucede con $\hat{\jmath} = (0, 1)$ y $\hat{\jmath} = (0, 1, 0)$. En ese sentido, si escribimos $v = x\hat{\imath} + y\hat{\jmath}$ y estamos en $\mathbb{R}^2$, entonces es $v = (x, y)$, pero si estamos en $\mathbb{R}^3$ es $v = (x, y, 0)$.

Rcordemos (ver [Resnick]) que si u y v son dos vectores no colineales, entonces su *producto vectorial* es el vector $u \times v$ definido por las siguientes propiedades.

- la norma de $u \times v$ verifica $\|u \times v\| = \|u\|\|v\| \operatorname{sen}(\theta)$, siendo $0 \leq \theta \leq \pi$ 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 y v son colineales, se define $u \times v = o$.

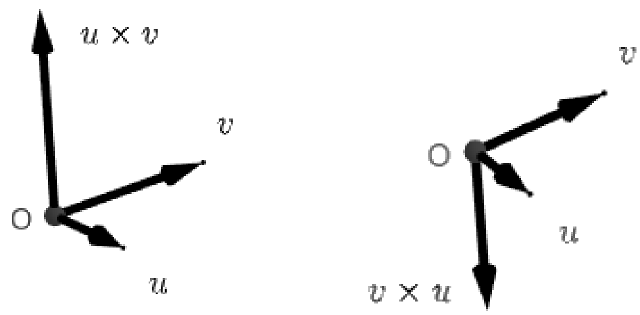

Figura 12: Producto vectorial

Proposición 2.13. *El producto vectorial verifica las siguientes propiedades.*

1. $v \times u = -v \times u$, *para todo* $u, v \in \mathbb{R}^3$ *(anticonmutativa);*
2. $(u + v) \times w = u \times w + v \times w$, *para todo* $u, v, w \in \mathbb{R}^3$ *(distributiva);*
3. $(au) \times v = a(u \times v)$, *para todo* $a \in \mathbb{R}$, $u, v \in \mathbb{R}^3$*;* □

Observaciones 2.14.

1. Las propiedades 1 y 3 de la proposición 10 son fáciles de probar. La propiedad 2 no lo parece tanto, pero también asumiremos que se deduce de la definición de producto vectorial[6].
2. La propiedad anticonmutativa implica que las propiedades 2 y 3 valen también "del otro lado":

$$u \times (v + w) = u \times v + u \times w, \quad \forall u, v, w \in \mathbb{R}^2; \qquad u \times (av) = a(u \times v), \quad \forall a \in \mathbb{R},\ u, v \in \mathbb{R}^2.$$

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

$$\hat{\imath} \times \hat{\jmath} = \hat{k}, \quad \hat{\jmath} \times \hat{k} = \hat{\imath}, \quad \hat{k} \times \hat{\imath} = \hat{\jmath},$$
$$\hat{\jmath} \times \hat{\imath} = -\hat{k}, \quad \hat{k} \times \hat{\jmath} = -\hat{\imath}, \quad \hat{\imath} \times \hat{k} = -\hat{\jmath},$$
$$\hat{\imath} \times \hat{\imath} = \hat{\jmath} \times \hat{\jmath} = \hat{k} \times \hat{k} = o.$$

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

Proposición 2.15. *Vale la fórmula siguiente*

$$(x_1\hat{\imath}+y_1\hat{\jmath}+z_1\hat{k})\times(x_2\hat{\imath}+y_2\hat{\jmath}+z_2\hat{k})=(y_1z_2-z_1y_2)\hat{\imath}+(z_1x_2-x_1z_2)\hat{\jmath}+(x_1y_2-y_1x_2)\hat{k}. \tag{10}$$

Esta fórmula escrita en coordenadas queda

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

Dem. La prueba consiste en desarrollar el producto, y usar las fórmulas anteriores y la proposición 2.13.

$$\begin{aligned}
&(x_1\hat{\imath}+y_1\hat{\jmath}+z_1\hat{k})\times(x_2\hat{\imath}+y_2\hat{\jmath}+z_2\hat{k})=\\
&=(x_1\hat{\imath})\times(x_2\hat{\imath}+y_2\hat{\jmath}+z_2\hat{k})+(y_1\hat{\jmath})\times(x_2\hat{\imath}+y_2\hat{\jmath}+z_2\hat{k})+(z_1\hat{k})\times(x_2\hat{\imath}+y_2\hat{\jmath}+z_2\hat{k})\\
&=x_1x_2\hat{\imath}\times\hat{\imath}+x_1y_2\hat{\imath}\times\hat{\jmath}+x_1z_2\hat{\imath}\times\hat{k}+y_1x_2\hat{\jmath}\times\hat{\imath}+y_1y_2\hat{\jmath}\times\hat{\jmath}+y_1z_2\hat{\jmath}\times\hat{k}+z_1x_2\hat{k}\times\hat{\imath}+z_1y_2\hat{k}\times\hat{\jmath}+z_1z_2\hat{k}\times\hat{k}\\
&=x_1y_2\hat{k}-x_1z_2\hat{\jmath}-y_1x_2\hat{k}+y_1z_2\hat{\imath}+z_1x_2\hat{\jmath}-z_1y_2\hat{\imath}=(y_1z_2-z_1y_2)\hat{\imath}+(z_1x_2-x_1z_2)\hat{\jmath}+(x_1y_2-y_1x_2)\hat{k}. \quad\square
\end{aligned}$$

Una manera de recordar la fórmula (10), es observar que se obtiene formalmente como el desarrollo por la primera fila del siguiente determinante[7]

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

Observación 2.16. Si tomamos la fórmula (10) como definición de producto vectorial, entonces todas las propiedades de la proposición 2.13 se prueban sin problema.

Proposición 2.17. *Sean u y v son dos vectores, con $u\neq o$. Entonces existen dos únicos vectores v_1 y v_2 tales que v_1 es colineal con u, v_2 es ortogonal a u y $v=v_1+v_2$. Valen $\|v_1\|=\frac{|u\cdot v|}{\|u\|}$ y $\|v_2\|=\frac{\|u\times v\|}{\|u\|}$.*

Dem. La descomposición $v=v_1+v_2$ sale de $v=\Pi_u(v)+\big(v-\Pi_u(v)\big)$, tomando $v_1=\Pi_u(v)$ y $v_2=v-\Pi_u(v)$. Es fácil de ver que esta descomposición es única (ver la figura 10).
La suma $v=v_1+v_2$ con v_1 y v_2 ortogonales, implica[8] $\|v_1\|=\|v\|\,|\cos(\theta)|$ y $\|v_2\|=\|v\|\operatorname{sen}(\theta)$, siendo θ el ángulo entre u y v. Luego

$$\|v_1\|=\|v\|\,|\cos(\theta)|=\frac{\|u\|\|v\|\,|\cos(\theta)|}{\|u\|}=\frac{|u\cdot v|}{\|u\|}\quad\text{y}\quad\|v_2\|=\|v\|\operatorname{sen}(\theta)=\frac{\|u\|\|v\|\operatorname{sen}(\theta)}{\|u\|}=\frac{\|u\times v\|}{\|u\|}.\quad\square$$

Corolario 2.18. *El área del paralelogramo generado por dos vectores u y v es $\|u\times v\|$.*

Dem. El área del paralelogramo es base por altura. Tomando $b=\|u\|$ como base, la altura es $h=\frac{\|u\times v\|}{\|u\|}$ (por la proposición anterior); luego el área es $bh=\|u\times v\|$. $\square$

[7] Para quienes no hayan estudiado determinantes, lo que necesitamos es la fórmula (16), en la sección 3.2.
[8] Notar que tenemos que tomar $|\cos(\theta)|$ porque puede ser $\frac{\pi}{2}<\theta\leq\pi$.

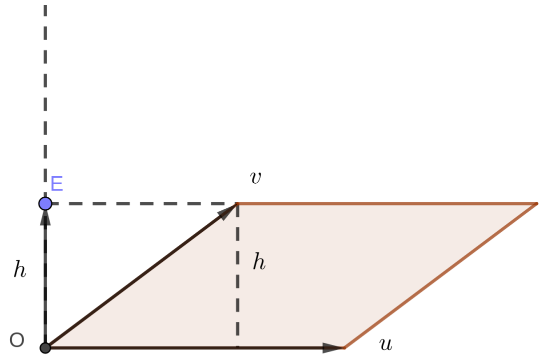

Figura 13: Paralelogramo

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

Proposición 2.19. *Si* $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2)$ *y* $w = (x_3, y_3, z_3)$, *entonces vale*

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

Dem. Desarrollando el determinante por su tercera fila obtenemos

$$\begin{aligned} \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} &= x_3 \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} - y_3 \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix} + z_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_3(y_1z_2 - z_1y_2) + y_3(z_1x_2 - x_1z_2) + z_3(x_1y_2 - y_1x_2) \\ &= \left(x_3\hat{\imath} + y_3\hat{\jmath} + z_3\hat{k}\right) \cdot \left((y_1z_2 - z_1y_2)\hat{\imath} + (z_1x_2 - x_1z_2)\hat{\jmath} + (x_1y_2 - y_1x_2)\hat{k}\right) = w \cdot (u \times v) \\ &= (u \times v) \cdot w. \quad \square \end{aligned}$$

Observación 2.20. Vale

$$(u \times v) \cdot w = u \cdot (v \times w),$$

para todo $u, v, w \in \mathbb{R}^3$. Esto se deduce del cálculo siguiente (mantenemos las notaciones anteriores).

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

En el cálculo con el determinante, primero intercambiamos las dos últimas filas y luego las dos primeras.

Proposición 2.21. *El volumen del paralelepípedo generado por tres vectores* u, v *y* w, *es el valor absoluto del producto mixto de* u, v *y* w, *es decir* $|(u \times v) \cdot w|$.

Dem. El volumen del paralelepípedo es el área de la base por la altura. Si tomamos la base como el paralelogramo generado por u y v, entonces su área es $a = \|u \times v\|$. Por otro lado $u \times v$ es un vector perpendicular al plano de u y v, luego la altura es (por la proposición anterior) $h = \frac{|(u \times v) \cdot w|}{\|u \times v\|}$. Entonces el volumen es $ah = \|u \times v\| \frac{|(u \times v) \cdot w|}{\|u \times v\|} = |(u \times v) \cdot w|$. $\square$

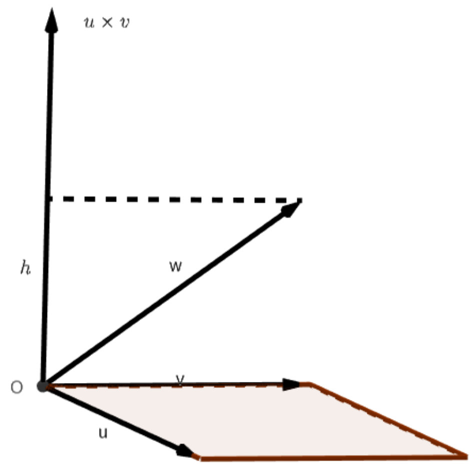

Figura 14: Paralelepípedo

2.3. Rectas y planos en el espacio

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 recordamos que 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 $P = (x_0, y_0, z_0)$ 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

$$X = P + tv \quad \Leftrightarrow \quad (x, y, z) = (x_0, y_0, z_0) + t(v_1, v_2, v_3). \tag{11}$$

Esta es la *ecuación vectorial* de r.

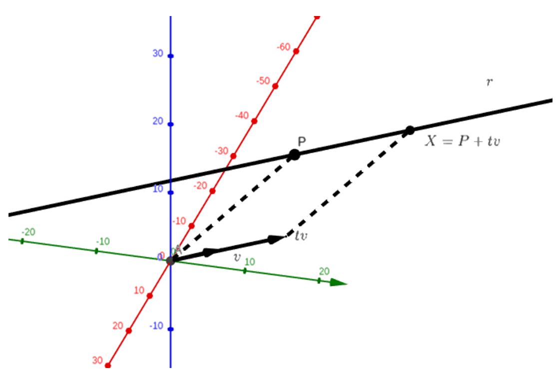

Figura 15: Recta

Si desarrollamos la segunda igualdad de (11) obtenemos

$$\begin{cases} x = x_0 + tv_1 \\ y = y_0 + tv_2 \\ z = z_0 + tv_3 \end{cases}, \quad t \in \mathbb{R}. \tag{12}$$

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

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3}. \tag{13}$$

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

$$\begin{cases} \frac{x-x_0}{v_1} = \frac{y-y_0}{v_2} \\ \frac{y-y_0}{v_2} = \frac{z-z_0}{v_3} \end{cases} \sim \begin{cases} \frac{x-x_0}{v_1} = \frac{y-y_0}{v_2} \\ \frac{x-x_0}{v_1} = \frac{z-z_0}{v_3} \end{cases} \sim \begin{cases} \frac{x-x_0}{v_1} = \frac{z-z_0}{v_3} \\ \frac{y-y_0}{v_2} = \frac{z-z_0}{v_3} \end{cases}.$$

Ejemplo 2.22. Sea r la recta 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

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

Para saber si un punto está en la recta, hay que ver si sus coordenadas verifican alguna de las ecuaciones anteriores. Por ejemplo, si $Q = (9, 5, 4)$, para ver que verifica la ecuación cartesiana hay que probar que vale

$$\frac{9-1}{4} = \frac{5+1}{3} = 4 - 2,$$

lo cual se verifica fácilmente; luego Q está en r.

Ejes coordenados. Si consideramos $Ox = \{(x, 0, 0) : \ x \in \mathbb{R}\}$, entonces un punto está en Ox si y solo si sus dos últimas coordenadas son nulas. Luego la ecuación cartesiana de Ox es $\begin{cases} y = 0 \\ z = 0 \end{cases}$. Análogamente la ecuación cartesiana de Oy es $\begin{cases} x = 0 \\ z = 0 \end{cases}$ y la de Oz es $\begin{cases} x = 0 \\ y = 0 \end{cases}$.

Rectas ortogonales y perpendiculares. Decimos que dos rectas son *ortogonales* si lo son sus direcciones; si además las rectas se cortan, entonces decimos que son *perpendiculares*. Por ejemplo los ejes coordenados son perpendiculares dos a dos, mientras que el eje Ox y la recta $\begin{cases} x = 0 \\ z = 1 \end{cases}$ son ortogonales pero no son perpendiculares.

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:

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

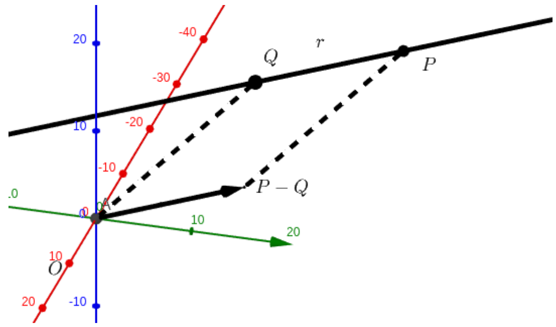

Figura 16: Recta por 2 puntos

Ejemplo 2.23. 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

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

Segmento de recta. El *segmento* de recta determinado por dos puntos P y Q es el conjunto

$$\overline{PQ} := \{tQ + (1 - t)P : \ 0 \le t \le 1\} = \{P + t(Q - P) : \ 0 \le t \le 1\}.$$

Notar que en el conjunto anterior, para $t = 0$ obtenemos el punto P y para $t = 1$ obtenemos el punto Q. Para $t = 1/2$ obtenemos el *punto medio* entre P y Q, dado por $M = \frac{1}{2}(P + Q)$.

Rectas en el plano. Lo que vimos de rectas en el espacio, se aplica también al caso de rectas en el plano. Para eso solo tenemos que omitir la tercera coordenada. Sea r la recta que pasa por el punto $P = (x_0, y_0)$ y es paralela al vector no nulo $v = (v_1, v_2)$. Las ecuaciones (11), (12) y (13) quedan, respectvamente

$$(x, y) = (x_0, y_0) + t(v_1, v_2); \qquad \begin{cases} x = x_0 + tv_1 \\ y = y_0 + tv_2 \end{cases} ; \qquad \frac{x - x_0}{v_1} = \frac{y - y_0}{v_2}.$$

En la última fórmula hay que asumir $v_1 \neq 0$ y $v_2 \neq 0$. Partiendo de esta última obtenemos

$$y - y_0 = m(x - x_0),$$

siendo $m = \frac{v_2}{v_1}$, que es la fórmula que su suele usar para la recta r que pasa por el punto $P = (x_0, y_0)$. En ese caso a m se le llama el *coeficiente angular* de r. Ese nombre viene de que es $m = \tan(\alpha)$, siendo α el ángulo que forma r con el eje Ox. De esta fórmula se obtiene

$$y = mx + n,$$

siendo $n = y_0 - mx_0$ (llamado la *ordenada en el origen* de r), que es una fórmula que permite describir todas las rectas del plano, menos las verticales. En general la *ecuación cartesiana* de una recta r tiene la forma $ax + by = c$, en que a y b no son simultáneamente nulos. Si es $b = 0$, entonces obtenemos $ax = c$ que es la ecuación de la recta vertical que pasa por el punto $(c/a, 0)$. Si es $b \neq 0$, entonces despejando y obtenemos una ecuación del tipo $y = mx + n$, que vimos anteriormente.

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.

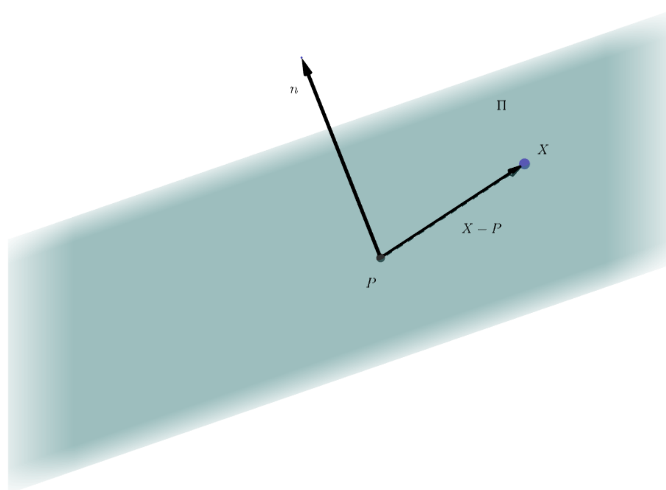

Figura 17: Plano por P ortogonal a n

Ejemplo 2.24. 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 2.25. 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 $\hat{k} = (0, 0, 1)$. 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

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0. \tag{14}$$

Ejemplo 2.26. 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

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

Para obtener la ecuación cartesiana, desarrollamos

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

Luego la ecuación cartesiana de Π es $8x + 5y - 6z = 9$. Es un ejercicio el verificar que se llega a la misma ecuación 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 (14) 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. 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

$$\left\{ \begin{array}{l} x + y - z = 4 \\ 2x - y - z = 1 \end{array} \right. .$$

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

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

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:

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

Observación 2.27. La ecuación cartesiana de la recta lo que hace es describirla como intersección de dos planos. Por ejemplo, si consideramos la recta de ecuación cartesiana $\frac{x-1}{4} = \frac{y+1}{3} = z - 2$, es

$$\frac{x-1}{4} = \frac{y+1}{3} = z - 2 \quad \Leftrightarrow \quad \begin{cases} \frac{x-1}{4} = z - 2 \\ \frac{y+1}{3} = z - 2 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x - 4z + 7 = 0 \\ y - 3z + 7 = 0 \end{cases}.$$

3. Matrices y determinantes

Las matrices y los determinantes son herramientas matemáticas básicas, que usaremos al estudiar espacios vectoriales y transformaciones lineales. Los resultados sobre matrices son fáciles de probar aunque engorrosos, pero los de determinantes dan bastante más trabajo, así que omitiremos la mayoría de las pruebas. Quien esté interesado en las mismas puede consultar [Friedberg] o cualquier otro texto sobre el tema.

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

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

siendo $a, b, \ldots, f$ números reales arbitrarios. Para escribir 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 segundo índice j indica la columna, a las que pertenece el elemento. Generalizando el caso 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[9], para todo $i = 1, \ldots, m$ y $j = 1, \ldots, n$. 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. Como la notación anterior es muy tediosa, en general usaremos $A = (a_{ij})_{m \times n}$, y cuando no dé lugar a confusión escribiremos simplemente $A = (a_{ij})$. Al conjunto de todas las matrices $m \times n$ lo escribiremos $M_{m \times n}$. Una *matriz columna* es una matriz $m \times 1$ (para algún m) y una *matriz fila* es una matriz $1 \times n$ (para algún n):

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

Análogamente a $\mathbb{R}^2$ y $\mathbb{R}^3$, definimos $\mathbb{R}^n := \{(x_1, x_2, \ldots, x_n) : \ x_1, \ldots, x_n \in \mathbb{R}\}$, el conjunto de todas las n-uplas de números reales, siendo $n = 2, 3, \ldots$. Una *n-upla* es un conjunto ordenado $(x_1, x_2, \ldots, x_n)$ de n

[9] 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.

elementos. Notar que dos n-uplas son iguales si y solo si contienen los mismos elementos y en el mismo orden;

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

Como una n-upla $(x_1, x_2, \ldots, x_n)$ es esencialmente lo mismo que una matriz fila $1 \times n$, entonces $\mathbb{R}^n$ se identifica naturalmente con $M_{1\times n}$. También se puede identificar $\mathbb{R}^n$ con el conjunto de las matrices columna $M_{n\times 1}$, escribiendo una n-upla en forma vertical

$$(x_1, x_2, \ldots, x_n) \quad \leftrightarrow \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Por la razón anterior a las matrices fila o columna se les suele llamar *vectores* (fila o columna), que es el nombre que se usa para los elementos de $\mathbb{R}^n$.

Decimos que dos matrices son *iguales* si tienen la misma cantidad de filas y de columnas, y coinciden elemento a elemento; es decir si $A = (a_{ij})_{m\times n}$ y $B = (b_{ij})_{r\times s}$, entonces $A = B$ si y solo si $m = r$, $n = s$ y $a_{ij} = b_{ij}$, para todo $i = 1, \ldots, m$ y $j = 1, \ldots, n$.

Matrices cuadradas. Una matriz se dice *cuadrada*[10] 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 conjunto $M_{n\times n}$ de matrices cuadradas $n \times n$ lo escribiremos M_n. Si $A = (a_{ij}) \in M_n$, entonces la *diagonal principal* de A es la n-upla $(a_{11}, a_{22}, \ldots, a_{nn})$. La matriz *identidad* I_n es la matriz $n \times n$ 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, escribiremos I en vez de I_n.

Si $A = (a_{ij})_{n\times n}$ es una matriz cuadrada, entonces damos las siguientes definiciones.

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

A continuación veremos distintas operaciones que se pueden realizar con matrices en general.

[10]En ete sentido a veces se dice que una matriz es *rectangular* cuando no es cuadrada.

Suma y producto por un escalar. Las operaciones de suma y producto por un escalar que vimos en $\mathbb{R}^2$ y $\mathbb{R}^3$ se generalizan naturalmente a $\mathbb{R}^n$, definiendo

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

para todo $a \in \mathbb{R}$, $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$. A su vez estas operaciones las podemos generalizar a una suma y producto por un 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*[11] $O \in M_{m\times n}$ es la matriz que tiene todas sus entradas nulas. Notar que decimos "la" matriz nula aunque en realidad hay una matriz nula $O \in M_{m\times n}$ para cada m, n. La *opuesta* de $A = (a_{ij}) \in M_{m\times n}$ es la matriz $-A = (b_{ij}) \in M_{m\times n}$ definida por $b_{ij} := -a_{ij}$, para todo i, j.

Ejemplo 3.1. En $M_{2\times 3}$ la matriz nula es $O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ y la opuesta de la matriz $A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 0 & 0 \end{pmatrix}$ es $-A = \begin{pmatrix} -1 & 2 & -3 \\ 0 & 0 & 0 \end{pmatrix}$.

Proposición 3.2. *Las matrices $m \times n$ con la suma y producto por un escalar definidos anteriormente verifican las siguientes propiedades:*

1. $A + B = B + A$, *para todo* $A, B \in M_{m\times n}$ *(conmutativa);*
2. $A + (B + C) = (A + B) + C$, *para todo* $A, B, C \in M_{m\times n}$ *(asociativa);*
3. $A + O = O + A = A$, *para todo* $A \in M_{m\times n}$ *(existencia de neutro);*
4. $A + (-A) = (-A) + A = O$, *para todo* $A, \in M_{m\times n}$ *(existencia de opuesto);*
5. $1A = A$, *para todo* $A \in M_{m\times n}$*;*
6. $a(B + C) = aB + aC$, *para todo* $a \in \mathbb{R}$, $B, C \in M_{m\times n}$*;*
7. $(a + b)C = aC + bC$, *para todo* $a, b \in \mathbb{R}$, $C \in M_{m\times n}$*;*
8. $a(bC) = (ab)C$, *para todo* $a, b \in \mathbb{R}$, $C \in M_{m\times n}$.

Dem. Ejercicio. □

Observación 3.3. Las matrices con la suma y producto por un escalar verifican las mismas propiedades que los vectores (proposición 2.2). Identificando $\mathbb{R}^n$ con $M_{1\times n}$, vemos que lo mismo sucede en $\mathbb{R}^n$.

La *resta* de matrices se define operando término a término. Si $A, B \in M_{m\times n}$, $A = (a_{ij})$ y $B = (b_{ij})$, entonces $A - B = (c_{ij}) \in M_{m\times n}$, en que $c_{ij} = a_{ij} - b_{ij}$, para todo i, j. Notar $A - B = A + (-B)$.

[11] En general para la matriz nula se usa 0 (el número cero) en vez de O (la letra o mayúscula), pero para evitar confusiones por ahora usaremos la letra O.

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

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}, \tag{15}$$

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 $(b_{ij}, \cdots, b_{nj})$.

$$c_{ij} = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} b_{ij} \\ \vdots \\ b_{nj} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

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 3.4. Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 2 \\ 1 & 0 \\ 5 & 3 \end{pmatrix}$, entonces su producto se calcula mediante

$$\begin{array}{c:c} & \begin{pmatrix} 0 & 2 \\ 1 & 0 \\ 5 & 3 \end{pmatrix} \\ \hdashline \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} & \begin{pmatrix} 1\times 0 + 2\times 1 + 3\times 5 & 1\times 2 + 2\times 0 + 3\times 3 \\ 4\times 0 + 5\times 1 + 6\times 5 & 4\times 2 + 5\times 0 + 6\times 3 \end{pmatrix} \end{array} \quad \Rightarrow \quad AB = \begin{pmatrix} 17 & 11 \\ 35 & 26 \end{pmatrix}.$$

Observación 3.5. Si $A, B \in M_n$, 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 último muestra que no vale la propiedad cancelativa: en general $AB = AC$ con $A \neq 0$, no implica $B = C$. Esto se debe a que $AB = AC$ equivale a $A(B - C) = O$, pero de esto no se deduce $B - C = O$.

Proposición 3.6. *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. $AI_n = I_m A = A$, *para todo* $A \in M_{m\times n}$.
5. $AO = OA = O$, *para todo* $A \in M_{m\times n}$, *siendo cada* O *la matriz nula*[12] *correspondiente.*
6. $A(cB) = (cA)B = c(AB)$, *para todo* $A \in M_{m\times n}$, $B \in M_{n\times p}$ *y* $c \in \mathbb{R}$.

Dem. Solo probaremos la primera propiedad, dejando las otras como ejercicio. Sean $A = (a_{ij})$, $B = (b_{ij})$ y $C = (c_{ij})$. Escribamos $BC = (d_{ij})$ y $A(BC) = (e_{ij})$. Entonces

$$d_{ij} = \sum_{h=1}^{p} b_{ih}c_{hj} \quad \Rightarrow \quad e_{ij} = \sum_{k=1}^{n} a_{ik}d_{kj} = \sum_{k=1}^{n} a_{ik}\sum_{h=1}^{p} b_{kh}c_{hj} = \sum_{k=1}^{n}\sum_{h=1}^{p} a_{ik}b_{kh}c_{hj}.$$

Por otro lado, si escribimos $AB = (f_{ij})$ y $(AB)C = (g_{ij})$, es

$$f_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \quad \Rightarrow \quad g_{ij} = \sum_{h=1}^{p} f_{ih}c_{hj} = \sum_{h=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kh}\right)c_{hj} = \sum_{h=1}^{p}\sum_{k=1}^{n} a_{ik}b_{kh}c_{hj}.$$

Comparando la sumas anteriores[13] vemos que vale $e_{ij} = g_{ij}$, para todo i, j; luego $A(BC) = (AB)C$. $\square$

Observaciones 3.7.

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 de las mismas dimensiones.
3. Si bien la fórmula para el producto es complicada, en el caso de matrices diagonales queda simple

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}\begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & 0 & \cdots & 0 \\ 0 & a_2b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_nb_n \end{pmatrix}.$$

Definición 3.8. Si $A \in M_n$ y $k \in \mathbb{N}$, entonces definimos la *potencia* k-ésima A^k de A mediante

$$A^0 = I \text{ (matriz identidad)}, \quad A^1 = A, \quad A^2 = AA, \quad A^3 = AAA, \quad \cdots$$

Calcular A^k para valores grandes de k en general es difícil, pero para matrices diagonales es fácil

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}^k = \begin{pmatrix} a_1^k & 0 & \cdots & 0 \\ 0 & a_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^k \end{pmatrix}, \quad \forall k = 0, 1, 2, \ldots$$

[12] En general en $AO = OA = O$, las tres matrices O tienen dimensiones distintas (solo coinciden cuando $m = n$).

[13] Notar que las sumas $\sum_{k=1}^{n}\sum_{h=1}^{p}\cdots$ y $\sum_{h=1}^{p}\sum_{k=1}^{n}\cdots$ coinciden, ya que una es una reordenación de la otra.

Trasposición. Si $A = (a_{ij})$ es una matriz $m \times n$, entonces su *traspuesta* A^t es la matriz $n \times m$ obtenida intercambiando las filas con las columnas de A, es decir $A^t = (b_{ij})$ en que $b_{ij} = a_{ji}$, para todo i, j.

Ejemplo 3.9. Si $A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, entonces $A^t = \begin{pmatrix} 1 & 0 \\ -2 & 0 \\ 3 & 0 \end{pmatrix}$.

Proposición 3.10. *Vale lo siguiente.*

- *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 $A^t = -A$.*
- *Uma matriz A es triangular superior si y solo si A^t es triangular inferior.*

Dem. Ejercicio. □

La siguiente proposición describe cómo se relaciona la trasposición con las otras operaciones.

Proposición 3.11. *Valen las siguientes propiedades*

$$(A + B)^t = A^t + B^t, \qquad (cA)^t = cA^t, \qquad (AB)^t = B^t A^t,$$

siendo, en cada caso, A y B matrices de dimesiones adecuadas para realizar las operaciones correspondientes y c una constante.

Dem. Solo probaremos la última propiedad, dejando las otras como ejercicio.
Sean $A = (a_{ij}) \in M_{m\times n}$ y $B = (b_{ij}) \in M_{n\times p}$. Entonces $AB = (c_{ij}) \in M_{m\times p}$, siendo $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$. Esto implica que $(AB)^t = (d_{ij}) \in M_{m\times p}$, siendo $d_{ij} = c_{ji} = \sum_{k=1}^n a_{jk}b_{ki}$. Por otro lado es $A^t = (\alpha_{ij}) \in M_{n\times m}$ y $B^t = (\beta_{ij}) \in M_{p\times n}$, siendo $\alpha_{ij} = a_{ji}$ y $\beta_{ij} = b_{ji}$. Luego $B^t A^t = (\gamma_{ij}) \in M_{m\times p}$, siendo

$$\gamma_{ij} = \sum_{k=1}^n \beta_{ik}\alpha_{kj} = \sum_{k=1}^n b_{ki}a_{jk} = \sum_{k=1}^n a_{jk}b_{ki} = d_{ij}, \quad \forall i, j \quad \Rightarrow \quad (AB)^t = B^t A^t. \quad \square$$

3.2. Determinantes

En esta sección las matrices son cuadradas.

El *determinante* de una matriz $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ es un número que escribimos indistintamente

$$\det(A) = |A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

y que para matrices 1×1, 2×2 y 3×3 se define por lo siguiente.

$$\det(a) := a,$$

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

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} := a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{16}$$

Desarrollando esta última fórmula obtenemos

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

Notar que para matrices 1×1 usamos $\det(a)$ en vez de $|a|$, para que no se confunda con el valor absoluto.

Observación 3.12. Una forma de recordar la fórmula para el caso 3×3 es el *método de Sarrus*. Consiste en construir una tabla repitiendo las dos primeras filas de la matriz debajo de ella misma

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

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 coincide con la fórmula obtenida anteriormente. El método de Sarrus no es la mejor forma de calcular determinantes y solo lo contamos acá por razones históricas y de cultura general. En la gran mayoría de los casos hay formas más rápidas y eficientes para calcular un determinante (ver la observación 3.25 y los ejemplos siguientes). Tener en cuenta que el método de Sarrus <u>solo sirve</u> para el caso 3×3.

A continuación veremos la definición del determinante de una matriz $n \times n$, para lo cual necesitamos introducir los cofactores. 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}, Δ_{13} y Δ_{21}, son

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

Una forma de saber si en la fórmula $\Delta_{ij} := (-1)^{i+j} \det\left(\tilde{A}_{ij}\right)$ el valor de $(-1)^{i+j}$ es $+1$ o -1, es pensar en el siguiente diagrama

$$\begin{vmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

En el vértice de arriba a la izquierda es $+$ y luego haces un cambio de signo cada vez que te mueves un lugar en forma horizontal o vertical.

Notar que la fórmula (16) 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 recursivamente mediante

$$\det(A) := a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots + a_{1n}\Delta_{1n}. \tag{17}$$

Si A es una matriz $n \times n$, entonces diremos 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 los de orden 4, sabiendo estos podemos calcular los de orden 5, etc.

Proposición 3.13. *El determinante de la matriz identidad I_n vale 1, para todo n.*

Dem. La prueba es por inducción en n. Para $n = 1$ es $I_1 = (1)$ y por lo tanto $\det(I_1) = \det(1) = 1$. Supongamos ahora, razonando por inducción, que vale $\det(I_{n-1}) = 1$ y queremos probar que vale $\det(I_n) = 1$. Recordar que es $I_n = (\delta_{ij})$, siendo δ_{ij} la delta de Kronecker. Aplicando la fórmula (17) obtenemos

$$\det(I_n) = \delta_{11}\Delta_{11} + \delta_{12}\Delta_{12} + \cdots + \delta_{1n}\Delta_{1n} = 1\Delta_{11} + 0\Delta_{12} + \cdots + 0\Delta_{1n} = \Delta_{11} = \det(I_{n-1}) = 1.$$

Luego $\det(I_n) = 1$, para todo n. □

Observación 3.14. Existe una fórmula para el determinante de una matriz $A = (a_{ij}) \in M_n$, que es la siguiente

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

siendo $\mathcal{S}_n$ 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. Esta fórmula tiene $n!$ (factorial de n) sumandos. Por ejemplo, para una matriz 10×10 tenemos 10!=3:628.800 sumandos. Como es de imaginar esta fórmula no es muy amigable para hacer cálculos y tampoco tiene demasiada utilidad en la teoría, así que no la usaremos.

La siguiente proposición resume las propiedades básicas del determinante.

Proposición 3.15. *1. Si una matriz B se obtiene de una matriz A intercambiando dos cualesquiera de sus columnas, entonces* $\det(B) = -\det(A)$.

2. Si en una matriz multiplicamos una columna por una constante, entonces su determinante también se multiplica por dicha constante. Es decir

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

3. El determinante es aditivo respecto a cada columna

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

Ejemplo 3.16. Consideremos $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Esta matriz se obtiene intercambiando las dos columnas de la matriz identidad $I_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, luego $\det(A) = -\det(I_2) = -1$. Consideremos ahora

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

Esta matriz se obtiene de la matriz identidad I_3 pasando la tercera columna al primer lugar, luego

$$\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = -\left(-\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \right) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

En el cálculo anterior, primero intercambiamos las dos primeras columnas y luego las dos últimas.

Usando las dos proposiciones anteriores se prueban todas las propiedades del determinante, aunque eso a veces requiere bastante esfuerzo. En particular se obtiene el siguiente resultado.

Proposición 3.17. *Para toda matriz A, vale* $\det(A^t) = \det(A)$. $\square$

Como la trasposición intercambia las filas con las columnas de la matriz, lo anterior implica que toda propiedad del determinante que vale para las columnas, vale también para las filas y recíprocamente. Luego alcanza con probar la propiedad en uno solo de estos casos, que es lo que haremos de ahora en adelante.

Proposición 3.18. *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.*

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11}a_{22} \cdots a_{nn}.$$

Dem. En el caso de una matriz triangular inferior, el resultado se obtiene aplicando la definición del determinante. Por ejemplo, para $n = 3$ es

$$\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\Delta_{11} + 0\Delta_{12} + 0\Delta_{13} = a_{11}\Delta_{11} = a_{11} \begin{vmatrix} a_{22} & 0 \\ a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}.$$

Notar que el razonamiento anterior vale para cualquier n. Como las matrices triangulares superiores son las traspuestas de las triangulares inferiores, entonces vale lo mismo para las triangulares superiores. Finalmente, observando que las matrices diagonales son un caso particular de matrices triangulares superiores o inferiores, deducimos que también vale para las matrices diagonales. □

De la proposición 3.15 se deduce el siguiente.

Corolario 3.19. *1. Si una matriz B se obtiene de una matriz A intercambiando dos cualesquiera de sus filas, entonces* $\det(B) = -\det(A)$.

2. Si en una matriz multiplicamos una fila por una constante, entonces su determinante también se multiplica por dicha constante. 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

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

Dado un conjunto de matrices columna $A_1, \dots, A_n$ (de las mismas dimensiones), decimos que una matriz columna A es *combinación lineal* de $A_1, \dots, A_n$ si existen escalares $c_1, \dots, c_n$ tales que $A = c_1A_1 + \cdots + c_nA_n$. Lo mismo se define para matrices fila.

Proposición 3.20. *Los siguientes tipos de matrices tienen determinante nulo.*

1. *Matriz con una columna formada solo por ceros.*

2. *Matriz con dos columnas iguales.*

3. *Matriz con una columna que es un múltiplo de otra columna.*

4. *Matriz con una columna que es combinación lineal de las otras columnas.*

Las afirmaciones anteriores valen también considerando las filas en lugar de las columnas.

Dem. La primera afirmación se prueba usando la parte 2 de la proposición 3.15:

$$\begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & 0 \cdot 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 \cdot 0 & \cdots & a_{nn} \end{vmatrix} = 0 \cdot \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = 0.$$

Consideremos la segunda afirmación. Sea A una matriz con dos columnas iguales y sea B la matriz obtenida intercambiando esas dos columnas. La parte 1 de la proposición 3.15 implica que es $\det(B) = -\det(A)$. Pero como esas dos columnas son iguales, es $B = A$ y por lo tanto $\det(B) = \det(A)$. Luego es $\det(A) = -\det(A)$, lo cual solo puede ocurrir cuando $\det(A) = 0$. La tercera afirmación se deduce de la segunda, usando la parte 2 de la proposición 3.15. Para la última afirmación vamos a escribir $A = (A_1|\cdots|A_n)$, siendo $A_1, \ldots, A_n$ las columnas de A. Para simplificar la notación, vamos a asumir que la segunda columna es combinación lineal de la primera y la tercera, luego suponemos que existen $c_1, c_3 \in \mathbb{R}$ tales que $A_2 = c_1A_1 + c_3A_3$. Entonces

$$\begin{aligned} \det(A) = (A_1|A_2|A_3|\cdots) &= \det(A_1|c_1A_1 + c_3A_3|A_3|\cdots) \\ &= \det(A_1|c_1A_1|A_3|\cdots) + \det(A_1|c_3A_3|A_3|\cdots) \\ &= c_1 \det(A_1|A_1|A_3|\cdots) + c_3 \det(A_1|A_3|A_3|\cdots) = 0 + 0 = 0. \end{aligned}$$

En el cálculo anterior aplicamos primero la parte 3 de la proposición 3.15 y luego la segunda afirmación. □

Ejemplos 3.21. Los siguientes son ejemplos de matrices cuyo determinante es nulo.

$$\begin{pmatrix} 1 & 2 & 0 \\ 5 & 7 & 0 \\ 4 & 3 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 5 & 7 & 6 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 2 \\ 5 & 7 & 10 \\ 2 & 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \\ 0 & 3 & 0 & 3 \\ 2 & 3 & 2 & 3 \end{pmatrix}.$$

La primera tiene una columna de ceros, la segunda tiene dos filas iguales, la tercera tiene una columna que es múltiplo de otra, en la cuarta, la última fila es suma de las dos anteriores.

Proposición 3.22. *Si en la fila i o en la columna j de A se cumple que todos los elementos salvo el a_{ij} son nulos, entonces* $\det(A) = a_{ij}\Delta_{ij}$*, siendo Δ_{ij} el cofactor correspondiente al lugar (i,j) de A. Es decir, vale*

$$\begin{vmatrix} a_{11} & \cdots & & a_{1,j} & & \cdots & a_{1n} \\ \vdots & & & \vdots & & & \vdots \\ 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & \cdots & & a_{n,j} & & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ & & 0 & & \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ & & 0 & & \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = a_{ij}\Delta_{ij}.$$

Dem. La prueba es el cálculo siguiente (para simplificar, escribimos asteriscos en lugar de los elementos)

$$
\begin{aligned}
\begin{vmatrix}
* & \cdots & * & * & * & \cdots & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
* & \cdots & * & * & * & \cdots & *
\end{vmatrix}
&= (-1)^{i-1}
\begin{vmatrix}
0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\
* & \cdots & * & * & * & \cdots & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\hdashline
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
* & \cdots & * & * & * & \cdots & *
\end{vmatrix} \\
&= (-1)^{i-1}(-1)^{j-1}
\left|\begin{array}{cccc:cc}
a_{ij} & 0 & \cdots & & \cdots & 0 \\
* & * & \cdots & & \cdots & * \\
\vdots & \vdots & & & & \vdots \\
\hdashline
\vdots & \vdots & & & & \vdots \\
* & * & \cdots & & \cdots & *
\end{array}\right|
= (-1)^{i+j} a_{ij}
\left|\begin{array}{cc:cc}
* & \cdots & \cdots & * \\
\vdots & & & \vdots \\
\hdashline
\vdots & & & \vdots \\
* & \cdots & \cdots & *
\end{array}\right| \\
&= a_{ij}\Delta_{ij}.
\end{aligned}
$$

En la cuenta anterior, en el primer paso llevamos la fila i al primer lugar, lo que implicó $i-1$ intercambios de filas, y luego llevamos la columna j al primer lugar, lo que implicó $j-1$ intercambios de columnas (las líneas discontinuas horizontal y vertical indican la fila y columna suprimidas). Finalmente, aplicamos la definición del determinante (fórmula (17)). □

Proposición 3.23. *Si en una matriz le sumamos a una fila un múltiplo de otra fila, o a una columna un múltiplo de otra columna, su determinante no varía.*

Dem. Sea A una matriz que escribimos $A = (A_1|\cdots|A_n)$, siendo $A_1,\dots,A_n$ las columnas de A. Sea B la matriz obtenida de A sumándole a la columna i la columna j multiplicada por una constante c, es decir

$$B = (A_1|\cdots|A_i + cA_j|\cdots|A_j|\cdots|A_n)$$

En lo de arriba, la columna i de B es $A_i + cA_j$. Luego aplicando las partes 3 de las proposiciones 3.15 y 3.20, obtenemos

$$
\begin{aligned}
\det(B) &= \det(A_1|\cdots|A_i + cA_j|\cdots|A_j|\cdots|A_n) \\
&= \det(A_1|\cdots|A_i|\cdots|A_j|\cdots|A_n) + \det(A_1|\cdots|cA_j|\cdots|A_j|\cdots|A_n) \\
&= \det(A) + 0 = \det(A). \quad \square
\end{aligned}
$$

Observación 3.24. Es fácil de probar que vale una versión un poco más general de la proposición anterior, que dice que si en una matriz le sumamos a una fila (columna) una combinación de las otras filas (columnas), su determinante no varía. Pero en la práctica lo que más se usa es la proposición anterior.

Observación 3.25. La técnica usual para calcular determinantes consiste en elegir una fila o columna, y aplicarle las veces que sea necesario la proposición anterior hasta llegar al determinante de una matriz que tiene solo un elemento no nulo en la fila o columna elegida. Luego se reduce su orden mediante la proposición 3.22 (usando también las otras propiedades para simplificar los cálculos). De esa forma pasamos de un determinante de orden n a uno de orden $n-1$. Aplicando reiteradamente este procedimiento terminamos reduciendo el cálculo de un determinante de orden n al de uno de orden 2.

Ejemplos 3.26. 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.

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

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), el tercero es $F_1 - F_3$.

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).

Puede haber casos en que el método anterior sea difícil de aplicar, por ejemplo cuando no nos damos cuenta de cómo "hacer ceros" mediante la proposición 3.23. En esos casos se puede usar el resultado siguiente, que si bien requiere más cálculos, tiene la ventaja de que siempre es aplicable.

Proposición 3.27. *Sea* $A = (a_{ij}) \in M_n$.

1. *Para cada* $i = 1, \dots, n$, *el determinante se puede calcular desarrollando a partir de la fila* i,
$$\det(A) = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}.$$

2. *Para cada* $j = 1, \dots, n$, *el determinante se puede calcular desarrollando a partir de la columna* j,
$$\det(A) = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj}.$$

Dem. Notar que la fila i de A se puede descomponer en suma de n filas:
$$(a_{i1}, a_{i2}, \cdots, a_{in}) = (a_{i1}, 0, \cdots, 0) + (0, a_{i2}, \cdots, 0) + \cdots + (0, 0, \cdots, a_{in}).$$
Luego aplicando reiteradamente la parte 3 del corolario 3.19 y luego la proposición 3.22, obtenemos
$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & & a_{nn} \end{vmatrix}$$
$$= a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}.$$
La prueba de la segunda afirmación es análoga. □

Observación 3.28. Respecto al producto de una matriz por un escalar, vale
$$\det(cA) = c^n \det(A), \quad \forall\, c \in \mathbb{R},\ A \in M_n.$$
Esta fórmula se obtiene aplicando la proposición 3.15 en cada columna (o fila) de la matriz cA. Hay que tener cuidado porque el determinante no se comporta bien respecto a la suma de matrices, es decir en general es $\det(A + B) \neq \det(A) + \det(B)$. Sin embargo, un resultado nada obvio pero que es verdadero, es que el determinante preserva el producto de matrices, es decir, vale lo siguiente.

Proposición 3.29. *Para todo* $A, B \in M_n$ *vale* $\det(AB) = \det(A)\det(B)$. □

3.3. Matriz inversa

Definición 3.30. Decimos que una matriz cuadrada $A \in M_n$ es *invertible* si existe una matriz $B \in M_n$ tal que $AB = BA = I$, siendo I la matriz identidad.

Proposición 3.31. *Si* A *es invertible, entonces la matriz* B *que verifica* $AB = BA = I$ *es única.*

Dem. Supongamos que existen dos matrices B, C tales que $AB = BA = I$ y $AC = CA = I$. Entonces
$$C = CI = C(AB) = (CA)B = IB = B, \quad \text{luego} \quad C = B. \quad \square$$

Observación 3.32. Por la proposición anterior, si una matriz A 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 de A (en caso de existir) queda caracterizada por ser la única matriz A^{-1} que verifica $AA^{-1} = A^{-1}A = I$.

Observación 3.33. La matriz identidad I es invertible y vale $I^{-1} = I$. Esto se deduce de $I \times I = I$.

Proposición 3.34.

1. *Si* A *es invertible, entonces* A^{-1} *es invertible y* $\left(A^{-1}\right)^{-1} = A$.
2. *Si* A *y* B *son invertibles, entonces su producto* AB *es invertible y* $(AB)^{-1} = B^{-1}A^{-1}$.

Dem. Si A es invertible, entonces existe su inversa A^{-1} que verifica $AA^{-1} = A^{-1}A = I$. Esta fórmula implica (por definición) que A^{-1} es invertible y que su inversa es A, es decir $\left(A^{-1}\right)^{-1} = A$.

Supongamos que A y B son invertibles, entonces

$$\begin{aligned}(AB)\left(B^{-1}A^{-1}\right) &= A\left(BB^{-1}\right)A^{-1} = AIA^{-1} = AA^{-1} = I,\\ \left(B^{-1}A^{-1}\right)(AB) &= B^{-1}\left(A^{-1}A\right)B = B^{-1}IB = B^{-1}B = I.\end{aligned}$$

Esto prueba simultánemente que AB es invertible y que $B^{-1}A^{-1}$ es la inversa de AB. □

En lo que sigue veremos de responder las preguntas siguientes. ¿Cómo saber si una matriz es invertible? Y si sabemos que es invertible, ¿cómo se halla su inversa?

Para matrices diagonales la respuesta es simple.

Proposición 3.35. *Una matriz diagonal es invertible si y solo si todas las entradas de la diagonal principal son no nulas. En ese caso*

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}^{-1} = \begin{pmatrix} a_1^{-1} & 0 & \cdots & 0 \\ 0 & a_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{-1} \end{pmatrix}. \tag{18}$$

Dem. Sea $A \in M_n$ la matriz diagonal. Si en A hay una entrada diagonal nula, por ejemplo la primera, entonces la primera fila de A es nula y por lo tanto para toda otra matriz $B \in M_n$ se cumple que la primera fila de AB es nula. Luego AB no puede ser la matriz identidad y por lo tanto A no es invertible.
Por otro lado si todas las entradas de la diagonal principal de A son no nulas, entonces es fácil de probar que la matriz que aparece a la derecha en (18) es la inversa de A (es solo multiplicarlas). □

Definición 3.36. Dada una matriz $A \in M_n$, la matriz *adjunta* de A es la matriz $\operatorname{adj}(A) \in M_n$ definida como la traspuesta de la matriz de los cofactores, es decir

$$\operatorname{adj}(A) := \begin{pmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2n} \\ \vdots & \vdots & & \vdots \\ \Delta_{n1} & \Delta_{n2} & \cdots & \Delta_{nn} \end{pmatrix}^t = \begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix}$$

siendo cada Δ_{ij} el cofactor correspondiente al lugar (i, j) de A.

Observación 3.37. La matriz adjunta anterior es la adjunta *clásica*. Cuando se estudian operadores en espacios con producto interno aparece otra matriz a la cual también también se llama la adjunta de A y se escribe A^*. No hay que confundirlas porque en general son distintas.

El interés en la matriz adjunta viene dado por el siguiente resultado.

Proposición 3.38. *Dada $A \in M_n$, su matriz adjunta $adj(A) \in M_n$ verifica*

$$A \cdot adj(A) = adj(A) \cdot A = \det(A)\, I.$$

Dem. Sea $\hat{A} := \operatorname{adj}(A)$. Tenemos que probar que vale $A\hat{A} = \hat{A}A = \det(A)\, I$. Si $A\hat{A} = (c_{ij})$, entonces $c_{ij} = \sum_{k=1}^{n} a_{ik}\Delta_{jk}$, para todo i, j. Luego $c_{ii} = \sum_{k=1}^{n} a_{ik}\Delta_{ik} = \det(A)$ (desarrollo a partir de la fila i), para todo i. 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 j del determinante de una matriz que tiene dos filas iguales (la i y la j):

$$\begin{array}{c} \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{array}\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = a_{i1}\Delta_{j1} + \cdots + a_{in}\Delta_{jn} \quad \Rightarrow \quad \sum_{k=1}^{n} a_{ik}\Delta_{jk} = 0, \quad \text{si } i \neq j..$$

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. $\square$

Teorema 3.39. *Una matriz $A \in M_n$ es invertible si y solo si* $\det(A) \neq 0$.
En caso afirmativo es $\det\left(A^{-1}\right) = \frac{1}{\det A}$ *y*

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

Dem. Si A es invertible, entonces

$$AA^{-1} = I \quad \Rightarrow \quad \det\left(AA^{-1}\right) = \det I \quad \Rightarrow \quad \det(A) \cdot \det\left(A^{-1}\right) = 1.$$

Esto implica $\det A \neq 0$ y $\det\left(A^{-1}\right) = \frac{1}{\det A}$.

Recíprocamente, supongamos que vale $\det(A) \neq 0$. Escribamos $\Delta = \det(A)$ y $\hat{A} = \mathrm{adj}(A)$. Sabemos que vale $A\hat{A} = \hat{A}A = \Delta I$, con $\Delta \neq 0$. Entonces multiplicando por Δ^{-1} obtenemos $A\left(\Delta^{-1}\hat{A}\right) = \left(\Delta^{-1}\hat{A}\right)A = I$. Luego (por definición) la matriz A es invertible y su inversa es $A^{-1} = \Delta^{-1}\hat{A}$. $\square$

Aplicando este teorema a matrices 2×2 obtenemos el sguiente resultado.

Corolario 3.40. *Una matriz $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ es invertible si y solo si $ad - bc \neq 0$. En es caso, es*

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

Método para obtener la matriz inversa. El teorema 3.39 nos permite saber si una matriz es invertible y nos da una fórmula para la inversa. El problema es que esta fórmula es poco práctica para matrices $n \times n$, con $n > 2$. Para hallar la inversa en el caso general existe un algoritmo que mostraremos a continuación. Para explicarlo, previamente necesitamos definir las operaciones elementales.

Definición 3.41. Las *operaciones elementales* en una matriz A son 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, entonces 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 hasta transformar la matriz A en la identidad I, cuando se llegó a ese 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 3.26. Además, escribiremos $F_2 \leftrightarrow F_3$, para decir que intercambiamos la fila 2 con la fila 3, y análogamente para columnas.

Ejemplos 3.42. Sea $A = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$. Vale $\det(A) = 1 \neq 0$, luego A es invertible. Para calcular su inversa tenemos que elegir primero si trabajar con columnas o con filas. En este caso trabajaremos con columnas.

2	3	1	1	0	0	
-1	0	-1	0	1	0	$C_1 - C_3$
1	1	1	0	0	1	
1	3	1	1	0	0	
0	0	-1	0	1	0	$C_2 - C_3$
0	1	1	-1	0	1	
1	2	1	1	0	0	
0	1	-1	0	1	0	$C_3 - C_1$
0	0	1	-1	-1	1	
1	2	0	1	0	-1	
0	1	-1	0	1	0	$C_2 - 2C_1$
0	0	1	-1	-1	2	
1	0	0	1	-2	-1	
0	1	-1	0	1	0	$C_3 + C_2$
0	0	1	-1	1	2	
1	0	0	1	-2	-3	
0	1	0	0	1	1	
0	0	1	-1	1	3	

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

Sea $A = \begin{pmatrix} 2 & 2 & 1 \\ -2 & -1 & -2 \\ 1 & 0 & 1 \end{pmatrix}$. Vale $\det(A) = 1$, luego A es invertible. Hallar su inversa operando con las filas.

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

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

Sea $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 3 \end{pmatrix}$. Vale $\det(A) = 6$, luego A es invertible. Trabajaremos con columnas.

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

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 3.43. 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 haber sido 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 equivocarse.

2. Conviene remarcar que cuando aplicamos el algoritmo anterior, tenemos que elegir entre trabajar con filas o con columnas, pero no podemos mezclar, porque 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.

Finalizamos esta sección con algunas propiedades.

Proposición 3.44. *Si $A, B \in M_n$ verifican $AB = I$, entonces A y B son invertibles y son inversas entre sí.*

Dem. Si $AB = I$, entonces $\det(A)\det(B) = 1$. Por lo tanto $\det(A) \neq 0$ y $\det(B) \neq 0$, lo cual implica que A y B son invertibles. Luego multiplicando por la izquierda por A^{-1} obtenemos

$$AB = I \quad \Rightarrow \quad A^{-1}(AB) = A^{-1}I \quad \Rightarrow \quad (A^{-1}A)B = A^{-1} \quad \Rightarrow \quad IB = A^{-1} \quad \Rightarrow \quad B = A^{-1}. \quad \square$$

Proposición 3.45. *Sea A una matriz invertible.*

1. *Si X y B son matrices tales que que $AX = B$, entonces $X = A^{-1}B$.*
2. *Si X y B son matrices tales que que $XA = B$, entonces $X = BA^{-1}$.*

Dem. Las dos fórmulas se prueban igual, así que probaremos solo la primera.

$$AX = B \quad \Rightarrow \quad A^{-1}(AX) = A^{-1}B \quad \Rightarrow \quad (A^{-1}A)X = A^{-1}B \quad \Rightarrow \quad IX = A^{-1}B \quad \Rightarrow \quad X = A^{-1}B. \quad \square$$

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 3.46. *Sean A, B, C matrices tales que A es invertible.*

1. *Si $AB = O$ o $BA = O$, entonces $B = O$ (la matriz nula).*
2. *Si $AB = AC$ o $BA = CA$, entonces $B = C$.*

Dem. Probaremos solo los primeros casos de cada ítem, los otros se prueban en forma análoga. Si $AB = O$, entonces $B = A^{-1}O = O$. Si $AB = AC$, entonces $B = A^{-1}(AC) = (A^{-1}A)C = IC = C$. $\square$

Observación 3.47. En la proposición y el corolario anteriores, la matriz A tiene que ser cuadrada, pero las otras matrices pueden ser cuadradas o rectangulares.

3.4. Método de Cramer

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

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ & \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= b_n \end{cases}. \tag{19}$$

Sean

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

Al número Δ se le llama el *determinante* del sistema. Notar que Δ_i es el determinante obtenido sustituyendo en Δ la columna i por la columna de resultados[14].

Teorema 3.48 (Cramer)**.** *Consideremos el sistema* (19) *y los determinantes $\Delta, \Delta_1, \ldots, \Delta_n$ definidos arriba.*

1. *Si $\Delta \neq 0$, entonces el sistema* (19) *es compatible determinado y la solución es $x_i = \frac{\Delta_i}{\Delta}$, para todo i.*
2. *Si $\Delta = 0$ y existe algún i tal que $\Delta_i \neq 0$, entonces el sistema* (19) *es incompatible.*
3. *Si $\Delta = \Delta_1 = \cdots = \Delta_n = 0$, entonces el sistema* (19) *no es compatible determinado; puede ser incompatible o compatible indeterminado.*

Dem. Probaremos ahora las dos primeras afirmaciones. La tercera se verá después de estudiar transformaciones lineales, en la observación 5.47.

Sean $A = (a_{ij})$ la matriz formada con los coeficientes del sistema. Notar que el sistema (19) equivale a la ecuación matricial $AX = B$, siendo $X = (x_1, \ldots, x_n)$ y $B = (b_1, \ldots, b_n)$ pensados como vectores columna.

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ & \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= b_n \end{cases} \quad \Leftrightarrow \quad \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

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

Consideremos $\hat{A} = (\Delta_{ij})^t$ la matriz adjunta de A. Si multiplicamos la ecuación $AX = B$ por $\hat{A}$ obtenemos $\hat{A}AX = \hat{A}B$. Usando que vale $\hat{A}A = \Delta I$ (proposición 3.38), obtenemos que esta última igualdad equivale a $\Delta X = \hat{A}B$. Si calculamos el producto $\hat{A}B$ obtenemos

$$\begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} b_1\Delta_{11} + b_2\Delta_{21} + \cdots + b_n\Delta_{n1} \\ b_1\Delta_{12} + b_2\Delta_{22} + \cdots + b_n\Delta_{n2} \\ \vdots \\ b_1\Delta_{1n} + b_2\Delta_{2n} + \cdots + b_n\Delta_{nn} \end{pmatrix} = \begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_n \end{pmatrix}$$

La última igualdad sale de desarrollar cada determinante Δ_i por la columna i. Luego obtuvimos

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

Si es $\Delta = 0$ y existe algún i_0 tal que $\Delta_{i_0} \neq 0$, entonces obtenemos una contradición en $\Delta\, x_{i_0} = \Delta_{i_0}$; luego en este caso la ecuación $AX = B$ no tiene solución.

Supongamos ahora que es $\Delta \neq 0$. Tomando determinantes en $\hat{A}A = \Delta I_n$ obtenemos

$$\det\big(\hat{A}A\big) = \det(\Delta I_n) \quad \Rightarrow \quad \det\big(\hat{A}\big)\Delta = \Delta^n \quad \Rightarrow \quad \det\big(\hat{A}\big) = \Delta^{n-1} \neq 0.$$

Luego $\hat{A}$ es invertible y por lo tanto las ecuaciones $AX = B$ y $\hat{A}AX = \hat{A}B$ son equivalentes. En resumen, obtuvimos

$$AX = B \quad \Leftrightarrow \quad \Delta\, x_i = \Delta_i,\ \forall i = 1, \dots, n. \tag{20}$$

Esto implica que si $\Delta \neq 0$, entonces $AX = B$ tiene una única solución que es $x_i = \frac{\Delta_i}{\Delta}$, para todo i. □

Observación 3.49. 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 el método no permite distinguir entre estos dos casos y hay que determinarlo de otra forma (en general se aplica escalerización). El único caso en que sí discrimina es cuando el sistema es homogéneo[15] (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 3.50. 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.

[15] Recordar que un sistema es homogéneo si todas sus ecuaciones aparecen igualadas a cero. Ver la sección 1.3.

3. Consideremos los siguientes sistemas

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

   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 ilustra la última afirmación del teorema de Cramer.

A continuación veremos la aplicación de Cramer a sistemas homogéneos.

Ejemplos 3.51. 1. Consideremos

$$\left\{\begin{array}{rcl} 2x+5y+27 & = & 0 \\ 3x+2y+5z & = & 0 \\ 5x+3y+7z & = & 0 \end{array}\right. .$$

   Su determinante es $\begin{vmatrix} 2 & 5 & 7 \\ 3 & 2 & 5 \\ 5 & 3 & 7 \end{vmatrix} = 11 \neq 0$. Luego la única solución que admite es la trivial $x = y = z = 0$.

2. Consideremos

$$\left\{\begin{array}{rcl} ax+y & = & 0 \\ x+ay & = & 0 \end{array}\right. .$$

   en que a es un parámetro. Nos interesa saber para qué valores de a el sistema admite soluciones no triviales.

   El determinante del sistema es $\Delta = \left|\begin{smallmatrix} a & 1 \\ 1 & a \end{smallmatrix}\right| = a^2 - 1$. Luego el sistema admite soluciones no triviales si y solo si $a^2 - 1 = 0$, lo cual equivale a $a = \pm 1$. Para hallar las soluciones tenemos que resolver los sistemas obtenidos sustituyendo a por 1 y por -1.

$$a = 1: \left\{\begin{array}{rcl} x+y & = & 0 \\ x+y & = & 0 \end{array}\right. ; \quad a = -1: \left\{\begin{array}{rcl} -x+y & = & 0 \\ x-y & = & 0 \end{array}\right. .$$

   Las soluciones para el caso $a = 1$ son $x = \lambda$, $y = -\lambda$ y para $a = -1$ son $x = y = \lambda$, siendo λ libre.

Observación 3.52. El método de Cramer no es eficiente para resolver sistemas de ecuaciones numéricos, ya que es más rápido escalerizar que calcular varios determinantes. Sin embargo tiene algunas ventajas para fines teóricos o cuando se quiere estudiar un sistema de ecuaciones que involucra parámetros, ya que “limpia” los cálculos (ver el ejemplo anterior). En la sección siguiente veremos también que es útil para el estudio de la dependencia lineal en espacios vectoriales.

4. Espacios vectoriales

Los espacios vectoriales son generalizaciones del espacio $\mathbb{R}^3$, que entre otras aplicaciones nos permiten trabajar en mayores dimensiones. Son una estructura muy general que aparece 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 un 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. $a \cdot (b \cdot v) = (ab) \cdot v$, para todo $a, b \in \mathbb{R}$, $v \in V$.
7. $a \cdot (u+v) = a \cdot u + a \cdot v$, para todo $a \in \mathbb{R}$, $u, v \in V$;
8. $(a+b) \cdot v = a \cdot v + b \cdot v$, 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 $u, v, w, \ldots$, mientras que a los de $\mathbb{R}$ les llamamos *escalares* y los escribimos con las letras $a, b, c, \ldots$.

Observación 4.1. La propiedad asociativa implica que podemos escribir $u+v+w := u+(v+w) = (u+v)+w$ y en general podemos considerar sumas de varios sumandos que escribiremos $\sum_{i=1}^{n} v_i = v_1 + \cdots + v_n$. Además la propiedad conmutativa implica que en una suma de vectores, podemos cambiar el orden de los sumandos sin afectar el resultado de la suma.

Ejemplos 4.2. Los siguientes son ejemplos de espacios vectoriales. En los dos primeros casos las operaciones son las habituales.

1. El espacio $\mathbb{R}^n$, $n \geq 1$. En particular el espacio $\mathbb{R}^3$, el plano $\mathbb{R}^2$ y la recta $\mathbb{R}$.
2. El conjunto $M_{m\times n}$ de las matrices con m filas y n columnas, $m, n \geq 1$.
3. El conjunto $\operatorname{Fun}(D, \mathbb{R})$ de las funciones que van de D a $\mathbb{R}$, siendo $D \neq \emptyset$ un subconjunto de $\mathbb{R}$. Las operaciones están definidas por: $(f+g)(x) := f(x)+g(x)$, $(a \cdot f)(x) := af(x)$, para todo $x \in \mathbb{R}$.
4. El conjunto de las sucesiones de números reales. 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 están definidas por $x+y := (x_n+y_n)_{n\in\mathbb{N}}$ y $a \cdot x := (ax_n)_{n\in\mathbb{N}}$.

Proposición 4.3. *Sea V un espacio vectorial. Valen las siguientes propiedades.*

1. *Existe un único elemento $o \in V$ que verifica la propiedad de existencia de neutro 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 de existencia de opuesto y se le llama el* opuesto *de v y lo escribimos $-v$. 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. *Si $a \in \mathbb{R}$ y $u, v \in V$ son tales que $a \cdot u = v$ con $a \neq 0$, entonces $u = a^{-1} \cdot v$.*

5. *Si $a \in \mathbb{R}$ y $v \in V$, entonces $a \cdot v = o$ si y solo si $a = 0$ o $v = o$.*

Dem.

1. Si un elemento o' verifica $v + o' = o' + v = v$, para todo $v \in V$, entonces $o = o + o' = o'$.

2. Sea $v \in V$. Si dos elementos u, w verifican $v + u = u + v = o$ y $v + w = w + v = o$, entonces $u = u + o = u + (v + w) = (u + v) + w = o + w = w$. Luego $u = w$.

3. $u + v = u + w \;\Rightarrow\; (-u) + (u + v) = (-u) + (u + w) \;\Rightarrow\; (-u + u) + v = (-u + u) + w \;\Rightarrow\; o + v = o + w \;\Rightarrow\; v = w$.

4. Como es $a \neq 0$, entonces existe el inverso de a. Luego
$$a \cdot u = v \quad \Rightarrow \quad a^{-1} \cdot (a \cdot u) = a^{-1} \cdot v \quad \Rightarrow \quad \left(a^{-1} a\right) \cdot u = a^{-1} \cdot v \quad \Rightarrow \quad 1 \cdot u = a^{-1} \cdot v \quad \Rightarrow \quad u = a^{-1} \cdot v.$$

5. Empezamos probando que si $a = 0$ o $v = o$, entonces $a \cdot v = o$. Para eso usamos la propiedad 3.
$$\begin{aligned} 0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v \quad &\Rightarrow \quad 0 \cdot v + 0 \cdot v = 0 \cdot v + o \quad \Rightarrow \quad 0 \cdot v = o, \\ a \cdot o = a \cdot (o + o) = a \cdot o + a \cdot o \quad &\Rightarrow \quad a \cdot o + a \cdot o = a \cdot o + o \quad \Rightarrow \quad a \cdot o = o. \end{aligned}$$
Supongamos $a \cdot v = o$. Si es $a = 0$ entonces ya está, en caso contrario (propiedad 4) es $v = a^{-1} \cdot o = o$. $\square$

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

Proposición 4.4. *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 $a \in \mathbb{R}$, $u, v \in V$.*

Dem. Probaremos solo la primera afirmación, dejando la prueba de las otras como ejercicio.

$$u = v - w \quad \Leftrightarrow \quad u = v + (-w) \quad \Leftrightarrow \quad u + w = \big(v + (-w)\big) + w \quad \Leftrightarrow \quad u + w = v + \big((-w) + w\big) \quad \Leftrightarrow \quad u + w = v + o \quad \Leftrightarrow \quad u + w = v. \quad \square$$

Nota: Por comodidad de notación, de ahora en adelante 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 4.5. 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 rápida de chequear la condición de subespacio.

Proposición 4.6. *Sea W un subconjunto de un espacio vectorial V. Entonces W es un subespacio de V si y solo si se verifica*

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

Dem. Ejercicio. □

Observación 4.7. Si W es un subespacio y $v_1, \dots, v_n \in W$, entonces $\sum_{i=1}^n a_i v_i \in W$, para todo $a_1, \dots, a_n \in \mathbb{R}$. Esto se prueba por inducción en n.

Ejemplos 4.8.

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

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 $p(x) = a_0 + a_1 x + \cdots + a_n x^n$, en que $a_0, a_1 \dots, a_n$ 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 $\mathbb{R}_n[x]$. Es claro que $\mathbb{R} \subset \mathbb{R}_1[x] \subset \mathbb{R}_2[x] \subset \cdots \subset \mathbb{R}[x] \subset \operatorname{Fun}(\mathbb{R}, \mathbb{R})$ son subespacios.

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

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

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

Notar que W_2 no es cerrado con la suma y que W_3 no es cerrado con el producto por escalares.

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

Proposición 4.9. *Si $W \subset V$ es un subespacio, entonces W es un espacio vectorial con las operaciones de V restringidas a W.* □

Dem. La propiedad de ser subespacio implica que W es cerrado por la suma y por el producto por escalares, y que contiene al vector nulo. Además, si $w \in W$, entonces $-w = (-1) \cdot w \in W$, luego también es cerrrado por opuestos. Las restantes propiedades de espacio vectorial se verifican trivialmente. □

Luego los ejemplos 4.8 son también ejemplos de espacios vectoriales.

Definición 4.10. 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 $v_1, \ldots, v_n \in V$, sea $[v_1, \ldots, v_n]$ el conjunto de todas las combinaciones lineales de $v_1, \ldots, v_n$:

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

Proposición 4.11. *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 $v_1, \ldots, v_n$.*

Dem. La prueba de que $[v_1, \ldots, v_n]$ es un subespacio se deduce de lo siguiente

$$\begin{aligned} o &= 0v_1 + \cdots 0v_n; \\ (a_1v_1 + \cdots + a_nv_n) + (b_1v_1 + \cdots + b_nv_n) &= (a_1 + b_1)v_1 + \cdots + (a_n + b_n)v_n; \\ c(a_1v_1 + \cdots + a_nv_n) &= (ca_1)v_1 + \cdots + (ca_n)v_n. \end{aligned}$$

Para la segunda afirmación, en la observación 4.7 vimos que si W es un subespacio y $v_1, \ldots, v_n \in W$, entonces $\sum_{i=1}^{n} a_iv_i \in W$, para todo $a_1, \ldots, a_n \in \mathbb{R}$. Esto implica $[v_1, \ldots, v_n] \subset W$. $\square$

Definición 4.12. Al subespacio $[v_1, \ldots, v_n]$ se le llama el *subespacio generado* por $v_1, \ldots, v_n$.

Ejemplos 4.13.

1. Si $o \neq v \in \mathbb{R}^3$, 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 4.14. Sean $v_1, \ldots, v_n \in V$. Decimos que el conjunto $\{v_1, \ldots, v_n\}$ es *linealmente dependiente* (LD) si existen $a_1, \ldots, a_n \in \mathbb{R}$ no todos nulos tales que $a_1v_1 + \cdots + a_nv_n = o$. En caso contrario decimos que el conjunto $\{v_1, \ldots, v_n\}$ es *linealmente independiente* (LI). A veces diremos simplemente que $v_1, \ldots, v_n$ son *linealmente dependientes* para referirnos a que $\{v_1, \ldots, v_n\}$ es LD y lo mismo para 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.

Proposición 4.15.

1. *Si en un conjunto $X = \{v_1, \ldots, v_n\}$ 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 $v = au$.*

Dem.

1. Como $v_i = o$, entonces $0v_1 + \cdots + 1v_i + \cdots + 0v_n = o$ es una combinación lineal no trivial que da el vector nulo.

2. Se deduce directamente de la última afirmación de la proposición 4.3.

3. Esto es un caso particular de la siguiente proposición. □

Proposición 4.16. *Un conjunto* $\{v_1, \ldots, v_n\}$ *es LD* ($n \geq 2$) *si y solo si existe algún vector* v_i *que es combinación lineal de los restantes* $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$.

Dem. Si $\{v_1, \ldots, v_n\}$ es LD, entonces existen $a_1, \ldots, a_n \in \mathbb{R}$ no todos nulos tales que $a_1 v_1 + \cdots + a_n v_n = o$. Luego existe algún i tal que $a_i \neq 0$, y por lo tanto despejando v_i obtenemos

$$v_i = \left(\frac{-a_1}{a_i}\right) v_1 + \cdots + \left(\frac{-a_{i-1}}{a_i}\right) v_{i-1} + \left(\frac{-a_{i+1}}{a_i}\right) v_{i+1} + \cdots + \left(\frac{-a_n}{a_i}\right) v_n.$$

Recíprocamente, si v_i es combinación lineal de los restantes vectores, entonces existen $b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n$ tales que $v_i = b_1 v_1 + \cdots + b_{i-1} v_{i-1} + b_{i+1} v_{i+1} + \cdots + b_n v_n$. Luego

$$b_1 v_1 + \cdots + b_{i-1} v_{i-1} + (-1) v_i + b_{i+1} v_{i+1} + \cdots + b_n v_n = o$$

es una combinación lineal no trivial que da el vector nulo. □

Ejemplos 4.17. 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

$$a(1, -1) + b(-4, 1) + c(1, 1) = (0, 0)$$

es la trivial. Operando, obtenemos que la ecuación anterior equivale al siguiente sistema

$$\left\{ \begin{array}{rcl} a - 4b + c & = & 0 \\ -a + b + c & = & 0 \end{array} \right. .$$

La solución de ese 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

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

Usando esta fórmula 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

$$a(1, 1, 1) + b(1, 2, 1) + c(2, 3, 3) = (0, 0, 0). \tag{21}$$

La ecuación (21) equivale al sistema homogéneo

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

El determinante del sistema es $\begin{vmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 3 \end{vmatrix} = 1$. 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
$$a(2,1,3) + b(-1,3,2) + c(1,11,12) = (0,0,0), \tag{22}$$
y X es LI si y solo si la única solución de (22) es la trivial. La ecuación (22) equivale al sistema homogéneo
$$\left\{\begin{array}{rcl} 2a - b + c &=& 0 \\ a + 3b + 11c &=& 0 \\ 3a + 2b + 12c &=& 0 \end{array}\right. .$$
El determinante del sistema es $\begin{vmatrix} 2 & -1 & 1 \\ 1 & 3 & 11 \\ 3 & 2 & 12 \end{vmatrix} = 0$. 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 $X = \left\{1,\ 1+x,\ 1+x+x^2\right\} \subset \mathbb{R}_2[x]$. Consideremos la combinación lineal
$$a1 + b(1+x) + c\left(1+x+x^2\right) = 0.$$
Operando 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 de este sistema es la trivial, por lo tanto X es LI.

5. Es fácil de probar que la base canónica $\left\{\hat{\imath}, \hat{\jmath}, \hat{k}\right\}$ 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 4.18. *Sean $A \subset B$ subconjuntos finitos de V.*

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

2. *Si B es LI, entonces A es LI.*

Dem. Supongamos $A = \{v_1, \cdots, v_n\}$ y $B = \{v_1, \cdots, v_n, v_{n+1}, \cdots, v_m\}$. Si A es LD, entonces existen $a_1, \ldots, a_n \in \mathbb{R}$ no todos nulos tales que $a_1v_1 + \cdots + a_nv_n = o$. Entonces
$$a_1v_1 + \cdots + a_nv_n + 0v_{n+1} + \cdots + 0v_m = o.$$
es una combinación linal no trivial, luego B es LD. La segunda parte se deduce de la primera: si B es LI, entonces B no es LD, entonces A no puede ser LD (por la primera parte), luego A es LI. $\square$

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

Proposición 4.19. *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 la 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. $\square$

4.4. Generadores y bases

Definición 4.20. Decimos que un conjunto $\{v_1, \ldots, v_n\} \subset V$ es un *generador*[16] 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 esa combinación lineal es única. entonces decimos que $\{v_1, \ldots, v_n\}$ es una *base* de V.

Observación 4.21. Si tenemos vectores $v_1, \ldots, v_n, v_{n+1}, \ldots, v_m$ y un vector w es combinación lineal de $v_1, \ldots, v_n$, entonces w es combinación lineal de $v_1, \ldots, v_n, v_{n+1}, \ldots, v_m$. Esto se debe a lo siguiente

$$w = a_1v_1 + \cdots + a_nv_n \ \Rightarrow\ w = a_1v_1 + \cdots + a_nv_n + 0v_{n+1} + \cdots + 0v_m.$$

Luego todo conjunto que contenga un generador de V, es también un generador de V.

Ejemplo 4.22. Consideremos los siguientes subconjuntos del espacio $\mathbb{R}^2$:

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

El conjunto $\mathcal{B}_1$ no es generador. Para verlo probaremos que $(1,1)$ no está en $[(1,2)]$:

$$(1,1) = a(1,2) \ \Rightarrow\ (1,1) = (a, 2a) \ \Rightarrow\ a = 1 \text{ y } a = 1/2 \ \lightning.$$

Respecto a $\mathcal{B}_2$, es $(2,-4) = 2(1,-2)$ y $(-3,6) = -2(1,-2)$. De esto se deduce $[(2,-4),\ (-3,6)] = [(1,-2)]$ y de la misma forma que antes se prueba que $\mathcal{B}_2$ no es generador. Probaremos que $\mathcal{B}_3$ es base.

$$(x,y) = a(1,2) + b(2,3) \ \Rightarrow\ \begin{cases} x = a + 2b \\ y = 2a + 3b \end{cases}.$$

Este sistema (en las variables a y b) es compatible determinado y su solución es $a = -3x + 2y$ y $b = 2x - y$. Esto prueba que $\mathcal{B}_3$ es base. Además obtuvimos cómo escribir un vector como combinación lineal de $\mathcal{B}_3$:

$$(x,y) = (-3x + 2y)(1,2) + (2x - y)(2,3), \quad \forall (x,y) \in \mathbb{R}^2.$$

Como $\mathcal{B}_4$ contiene a $\mathcal{B}_3$, entonces $\mathcal{B}_4$ es generador de $\mathbb{R}^2$ (por la observación 4.21). Además es

$$(3,5) = (1,2) + (2,3) \ \Rightarrow\ \begin{cases} (3,5) = 1(1,2) + 1(2,3) + 0(3,5) \\ (3,5) = 0(1,2) + 0(2,3) + 1(3,5) \end{cases}.$$

Esto prueba que $\mathcal{B}_4$ no es base. Luego $\mathcal{B}_4$ es generador pero no es base.

Proposición 4.23. *Un conjunto $\mathcal{B}$ es una base de V si y solo si $\mathcal{B}$ es un generador y es LI.*

Dem. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$. Supongamos que $\mathcal{B}$ es una base de V. Es claro que $\mathcal{B}$ es un generador de V, así que solo tenemos que probar que es LI. Sean $a_1, \ldots, a_n \in \mathbb{R}$ tales que $a_1v_1 + \cdots + a_nv_n = o$. Como también se cumple que $0v_1 + \cdots + 0v_n = o$, entonces la unicidad de la decomposición implica $a_1 = \cdots = a_n = 0$.

Asumimos ahora que $\mathcal{B}$ es un generador y que es LI. Sea $v \in V$. Supongamos que existen $a_1, \ldots, a_n \in \mathbb{R}$ y $b_1, \ldots, b_n \in \mathbb{R}$ tales que $v = a_1v_1 + \cdots + a_nv_n$ y $v = b_1v_1 + \cdots + b_nv_n$. Restando obtenemos $(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = o$. Como $\mathcal{B}$ es LI, esto implica $a_1 - b_1 = \cdots = a_n - b_n = 0$ y por lo tanto $a_1 = b_1,\ \ldots\ , a_n = b_n$. Luego probamos que todo vector del espacio se escribe de forma única como combinación lineal de $\mathcal{B}$. $\square$

[16]En otros países se le suele llamar un *conjunto generador* o *conjunto de generadores*, pero acá lo llamamos solo generador.

Ejemplo 4.24. Consideremos el espacio vectorial $V = \{(x, y, z) \in \mathbb{R}^3 : \ 2x + y + z = 0\}$ (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

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

Esto nos dice que $\mathcal{B} = \{(1, 0, -2), (0, 1, -1)\}$ 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 hay ciertas bases estándar a las que se les suele llamar la *base canónica* del espacio, que se describen a continuación.

Ejemplos 4.25.

1. La base canónica de $\mathbb{R}^n$ es el conjunto $\mathcal{B} = \{e_1, e_2, \cdots, e_n\}$ definido en el ejemplo 4.17.6. En particular la base canónica de $\mathbb{R}^2$ es $\{\hat{\imath}, \hat{\jmath}\} = \{e_1, e_2\}$ y la de $\mathbb{R}^3$ es $\{\hat{\imath}, \hat{\jmath}, \hat{k}\} = \{e_1, e_2, e_3\}$.
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 4.26. *En un espacio vectorial, la cantidad de elementos de un conjunto LI siempre es menor o igual que la cantidad de elementos de un generador.*

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

Como $v_1 \in V$ y G es generador de V, entonces existen $a_1, \ldots, a_m \in \mathbb{R}$ tales que $v_1 = a_1 w_1 + \cdots + a_m w_m$. 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 G, podemos suponer que es $a_1 \neq 0$. Luego

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

Veremos que $G_1 := \{v_1, w_2, \ldots, w_m\}$ es un generador de V. Sea $v \in V$ un vector arbitrario. Como G es un generador de V, entonces existen $b_1, \ldots, b_m \in \mathbb{R}$ tales que $v = b_1 w_1 + \cdots + b_m w_m$. Luego

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

Esto muestra que todo vector de V se puede escribir como combinación lineal de G_1.

La prueba sigue por inducción en k. Nuestra hipótesis inductiva es que existe $k \geq 1$ tal que el conjunto $G_k := \{v_1, \ldots, v_k, w_{k+1}, \ldots, w_m\}$ es un generador de V (acabamos de probar que esto vale para $k = 1$). Como $v_{k+1} \in V$ y G_k es un generador, entonces existen $c_1, \cdots, c_m \in \mathbb{R}$ tales que

$$v_{k+1} = c_1 v_1 + \cdots + c_k v_k + c_{k+1} w_{k+1} + \cdots + c_m w_m. \tag{23}$$

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 (23) obteniendo

$$w_{k+1} = \left(\frac{-c_1}{c_{k+1}}\right) v_1 + \cdots + \left(\frac{-c_k}{c_{k+1}}\right) v_k + \frac{1}{c_{k+1}} v_{k+1} + \left(\frac{-c_{k+2}}{c_{k+1}}\right) w_{k+2} + \cdots + \left(\frac{-c_m}{c_{k+1}}\right) w_m. \tag{24}$$

Probaremos que $G_{k+1} := \{v_1, \dots, v_k, v_{k+1}, w_{k+2}, \dots, w_m\}$ es un generador de V. La idea de la prueba es la misma que antes. Para simplificar la notación, la fórmula (24) 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 G_k es un generador de V, entonces existen $x_1, \dots, x_m \in \mathbb{R}$ 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 G_{k+1}.

Luego probamos que para cada $k = 1, 2, \dots$ se cumple (eventualmente reordenando los elementos de G) que $G_k = \{v_1, \dots, v_k, w_{k+1}, \dots, w_m\}$ es un generador de V. Si fuese $n > m$, entonces podríamos sustituir todos los elementos de G por elementos de $\mathcal{A}$ para obtener que el conjunto $G_m = \{v_1, \dots, v_m\}$ 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 $n \leq m$. □

Corolario 4.27. *Si $\mathcal{B}_1 = \{v_1, \dots, v_n\}$ y $\mathcal{B}_2 = \{w_1, \dots, w_m\}$ son dos bases de V, entonces $n = m$.*

Dem. Como $\mathcal{B}_1$ es LI y $\mathcal{B}_2$ es generador, entonces $n \leq m$. Intercambiando los roles de $\mathcal{B}_1$ y $\mathcal{B}_2$ obtenemos $m \leq n$. Luego $m = n$. □

Definición 4.28. 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; en ese caso escribimos $\dim(V) = n$. En caso contrario diremos que V es de *dimensión infinita*. Si $V = \{o\}$ (el espacio trivial), entonces definimos $\dim V = 0$.

En los ejemplos que siguen usamos las notaciones de los ejemplos 4.24 y 4.25.

Ejemplos 4.29. 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 $\{e_{ij} : 1 \leq i \leq m,\ 1 \leq j \leq n\}$ es base de $M_{m\times n}$, luego $\dim(M_{m\times n}) = mn$, para todo m, n.
4. El conjunto $\{(1,0,-2),\ (0,1,-1)\}$ es base del plano $V = \big\{(x,y,z) \in \mathbb{R}^3 :\ 2x+y+z=0\big\}$, luego $\dim(V) = 2$.

Proposición 4.30. *Si V es un espacio de dimensión finita n, entonces la cantidad de elementos de un conjunto LI es menor o igual que n y que la cantidad de elementos de un generador es mayor o igual a n.*

Dem. Esto es consecuencia directa del teorema 4.26. □

Ejemplos 4.31. El espacio de los polinomios $\mathbb{R}[x]$ tiene dimensión infinita. Esto se debe a que el conjunto $\{1,\ x,\ \cdots,\ x^n\}$ es LI, para todo n. Luego $\mathbb{R}[x]$ no puede tener dimensión infinita por la proposición anterior. Si consideramos el espacio de las funciones derivables $D(\mathbb{R})$, como $\mathbb{R}[x] \subset D(\mathbb{R})$ es un subespacio, deducimos que $D(\mathbb{R})$ también tiene dimensión infinita. Otros ejemplos de espacios de dimensón infinita son el de las funciones continuas $C(\mathbb{R})$, el de las funciones $\mathrm{Fun}(\mathbb{R})$ y el de las sucesiones.

Nota: De ahora en adelante asumiremos siempre que los espacios tienen dimensión finita.

El siguiente resultado muestra que todo conjunto LI se puede completar a una base.

Proposición 4.32. *Sea V un espacio de dimensión n y $\mathcal{A} = \{v_1, \ldots, v_m\} \subset V$ un subconjunto LI.*

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

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

Dem. Supongamos $m = n$. Si $\mathcal{A}$ no fuese base de V, entonces no sería generador y por lo tanto existiría un vector $v_{n+1} \in V$ tal que $v_{n+1} \notin [v_1, \ldots, v_n]$. Pero esto implicaría que $\{v_1, \ldots, v_n, v_{n+1}\}$ es LI (por la proposición 4.19), lo cual no es posible por la proposición anterior. Luego $\mathcal{A}$ es base de V.
Para la segunda parte la idea es la misma. Si $m < n$ entonces $\mathcal{A}$ no puede ser base de V, y por lo tanto no es un generador de V. Luego existe un vector $v_{m+1} \in V$ tal que $v_{m+1} \notin [v_1, \ldots, v_m]$ y por lo tanto $\{v_1, \ldots, v_m, v_{m+1}\}$ es LI. Si $m + 1 = n$, entonces ya terminamos (por la primera parte). Si $m + 1 < n$ repetimos el razonamiento con $\{v_1, \ldots, v_m, v_{m+1}\}$ en lugar de $\mathcal{A}$ y así seguimos hasta llegar a una base $\mathcal{B} = \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$. $\square$

Proposición 4.33. *Si $G = \{v_1, \ldots, v_m, v_{m+1}\}$ es un generador de V y v_{m+1} es combinación lineal de $v_1, \ldots, v_m$, entonces $G' = \{v_1, \ldots, v_m\}$ es también un generador de V.*

Dem. Sean $a_1, \ldots, a_m \in \mathbb{R}$ tales que $v_{m+1} = a_1 v_1 + \cdots + a_m v_m$. Sea $w \in V$. Como G es un generador de V, entonces existen $b_1, \ldots, b_{m+1} \in \mathbb{R}$ tales que $w = b_1 v_1 + \cdots + b_{m+1} v_{m+1}$. Luego

$$w = b_1 v_1 + \cdots + b_m v_m + b_{m+1}(a_1 v_1 + \cdots + a_m v_m) = (b_1 + b_{m+1} a_1) v_1 + \cdots + (b_m + b_{m+1} a_m) v_m.$$

Esto muestra que G' es un generador de V. $\square$

Proposición 4.34. *Sea V un espacio de dimensión n y G un generador de V con p elementos.*

1. *Si $p = n$, entonces G es base de V.*

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

Dem. Supongamos $p = n$. Si G no fuese base, entonces sería LD y por lo tanto existiría un vector de G que es combinación lineal de los restantes, Entonces sacando este vector de G obtendríamos un generador G' con menor cantidad de elementos que la dimensión del espacio, lo cual es imposible. Luego G es base.
Si $p > n$, entonces G es LD y por lo tanto tiene algún vector que es combinación lineal de los restantes. Aplicando la proposición anterior podemos sacarle ese elemento obteniendo un generador contenido en G con $p - 1$ elementos. Luego se sigue repitiendo el proceso hasta obtener un generador con n elementos, que por la primera parte sabemos es base. $\square$

Ejemplo 4.35. Consideremos los conjuntos

$$G = \{(1,0),\, (2,0),\, (0,1)\}, \quad \mathcal{A} = \{(1,0),\, (0,1)\}, \quad \mathcal{B} = \{(2,0),\, (0,1)\}, \quad \mathcal{C} = \{(1,0),\, (2,0)\}.$$

El conjunto G es un generador de $\mathbb{R}^2$ (contiene a la base canónica $\mathcal{A}$). Los conjuntos $\mathcal{A}$ y $\mathcal{B}$ están contenidos en G y son bases de $\mathbb{R}^2$. El conjunto $\mathcal{C}$ está contenido en G, pero no es base de $\mathbb{R}^2$.

De las proposiciones 4.32 y 4.34 se deduce el siguiente resultado.

Corolario 4.36. *Sea V un espacio de dimensión n y $\mathcal{A} \subset V$ un conjunto de n elementos. Entonces las siguientes afirmaciones son equivalentes:*

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

Ejemplo 4.37. 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 da un criterio simple para saber si un conjunto es base de $\mathbb{R}^n$.

Proposición 4.38. *Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ un subconjunto de $\mathbb{R}^n$ y Δ el determinante de la matriz cuyas columnas son los vectores $v_1, \dots, v_n$ escritos verticalmente. Entonces $\mathcal{B}$ es base de $\mathbb{R}^n$ si y solo si $\Delta \neq 0$.*

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

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

Sabemos por el método de Cramer que el sistema homogéneo anterior admite como única solución la trivial si y solo si $\Delta \neq 0$. Luego $\mathcal{B}$ es LI si y solo si $\Delta \neq 0$. La prueba se completa observando que como $\mathcal{B}$ tiene n elementos, entonces $\mathcal{B}$ es base de $\mathbb{R}^n$ si y solo si $\mathcal{B}$ es LI. $\square$

Proposición 4.39. *Si V es un espacio de dimensión finita y W es un subespacio de V, entonces W es un espacio de dimensión finita y* $\dim(W) \leq \dim(V)$. *Además, si* $\dim(W) = \dim(V)$, *entonces* $W = V$.

Dem. Si $W = \{o\}$ el resultado es obvio. Supongamos ahora $W \neq \{o\}$. Todo subconjunto de W que es LI en W, es también LI en V y por lo tanto tiene una cantidad de elementos menor o igual que $\dim(V)$. Sea $\mathcal{A} = \{w_1, \dots, w_m\}$ un subconjunto LI de W con $m \leq \dim(V)$ el máximo posible. Si $\mathcal{A}$ no fuese un generador de W, entonces existiría $w_{m+1} \in W$ que no es combinación lineal de $w_1, \dots, w_m$ y por lo tanto $\mathcal{A}_1 = \{w_1, \dots, w_m, w_{m+1}\}$ sería también un subconjunto LI de W, contradiciendo la maximalidad de m. Luego $\mathcal{A}$ es generador de W y por lo tanto es base de W. Esto a su vez implica $\dim(W) = m \leq \dim(V)$.

Si $\dim(W) = \dim(V) = n$ y $\mathcal{B} = \{w_1, \dots, w_n\}$ es una base de W, entonces $\mathcal{B}$ es un subconjunto LI de V con la misma cantidad de elementos que $\dim(V)$ y por lo tanto es base de V. Luego $V = W = [w_1, \dots, w_n]$. $\square$

En lo que sigue usaremos la proposición anterior para determinar los subespacios propios de $\mathbb{R}^2$ y de $\mathbb{R}^3$.

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 $W = [v]$, 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 $W = [v]$, con $o \neq v \in \mathbb{R}^3$, es una recta que pasa por el origen; si $\dim(W) = 2$, entonces $W = [u, v]$, 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 4.40. *Si W_1, W_2 son subespacios de V, entonces $W_1 \cap W_2$ y $W_1 + W_2$ son también subespacios de V.*

Dem. Ejercicio. $\square$

Observación 4.41. 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 4.42. 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 4.43. Sean W_1, W_2 dos subespacios de un espacio V. Decimos que un subespacio U 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 4.44. Si consideramos los ejemplos 4.42, el primero nos dice que el espacio $\mathbb{R}^3$ es suma del plano Oxy con el plano Oxz, pero esta suma no es directa, y el segundo, que el plano Oxy es suma directa de la recta Ox y la recta Oy.

Proposición 4.45. *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.*

Dem. Asumimos $V = W_1 \oplus W_2$. Sea $v \in V$ y supongamos que existen $w_1, w_1' \in W_1$ y $w_2, w_2' \in W_2$ tales que $v = w_1 + w_2 = w_1' + w_2'$. Restando obtenemos $w_1 - w_1' = w_2' - w_2$. Observar que es $w_1 - w_1' \in W_1$ y $w_2' - w_2 \in W_2$. Luego $w_1 - w_1' = w_2' - w_2 \in W_1 \cap W_2 = \{o\}$. Esto implica $w_1 - w_1' = w_2' - w_2 = o$ y por lo tanto $w_1 = w_1'$ y $w_2 = w_2'$.
Recíprocamente, supongamos ahora que todo vector de V se escribe en forma única como suma de un vector de W_1 con uno de W_2. Esto implica que es $V = W_1 + W_2$. Sea $v \in W_1 \cap W_2$. Considerando $o + v = v + o$ y usando la unicidad de la descomposición, se deduce que es $v = o$. Luego $W_1 \cap W_2 = \{o\}$. $\square$

Proposición 4.46. *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.*

Dem. Sea $v \in V$. Sabemos que existen $u_1 \in W_1$ y $u_2 \in W_2$ tales que $v = u_1 + u_2$. Usando que $\mathcal{B}_1$ y $\mathcal{B}_2$ son bases, tenemos que existen escalares a_i y b_j tales que $u_1 = \sum_{i=1}^n a_i v_i$ y $u_2 = \sum_{j=1}^m b_j w_j$. Luego $v = u_1 + u_2 = \sum_{i=1}^n a_i v_i + \sum_{j=1}^m b_j w_j$. Esto prueba que $\mathcal{B}$ es un generador de V.
Sean escalares c_i y d_j tales que $\sum_{i=1}^n c_i v_i + \sum_{j=1}^m d_j w_j = o$. Entonces

$$\sum_{i=1}^n c_i v_i = -\sum_{j=1}^m d_j w_j \in W_1 \cap W_2 = \{o\} \quad \Rightarrow \quad \sum_{i=1}^n c_i v_i = o \ \text{ y } \ \sum_{j=1}^m d_j w_j = o.$$

Como $\mathcal{B}_1$ y $\mathcal{B}_2$ son LI, esto implica que todos los escalares son cero. Esto prueba que $\mathcal{B}$ es LI. $\square$

Corolario 4.47. *Si $V = W_1 \oplus W_2$, entonces* $\dim V = \dim W_1 + \dim W_2$. $\square$

Corolario 4.48. *Sean W_1, W_2 dos subespacios tales que $W_1 \cap W_2 = \{o\}$. Si $\dim V = \dim W_1 + \dim W_2$, entonces $V = W_1 \oplus W_2$.*

Dem. Sea $U := W_1 + W_2$. Como es $W_1 \cap W_2 = \{o\}$, entonces $U = W_1 \oplus W_2$ y por lo tanto $\dim U = \dim W_1 + \dim W_2$. Como es $\dim V = \dim W_1 + \dim W_2$, obtenemos $\dim U = \dim V$ y por lo tanto $U = V$. Luego $W_1 \oplus W_2 = V$. $\square$

Ejemplo 4.49. 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 4.50. 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 4.51. La suma directa de dos subespacios se generaliza a una cantidad finita cualquiera de subespacios (y también para infinitos subespacios), pero esta construcción es más delicada y no la vamos a desarrollar. Quien esté interesado en el tema, puede leerlo en las páginas 265-267 de [Friedberg].

4.6. Coordenadas y cambio de base

Definición 4.52. Sea V un espacio y $\mathcal{B} = \{v_1, \ldots, v_n\}$ 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 $x_1, \ldots, x_n \in \mathbb{R}$ tales que $v = x_1 v_1 + \cdots + x_n v_n$. En este caso escribimos $\operatorname{coord}_{\mathcal{B}}(v) = (x_1, \ldots, x_n)$ y decimos que la n-upla $(x_1, \ldots, x_n) \in \mathbb{R}^n$ son las *coordenadas* del vector v en la base $\mathcal{B}$.

En lo que sigue supondremos siempre que las bases son bases ordenadas.

Ejemplos 4.53.

1. Consideremos las bases $\mathcal{B}_1 = \{(1, 1), (1, -1)\}$ y $\mathcal{B}_2 = \{(1, -1), (1, 1)\}$ de $\mathbb{R}^2$. Si $(x, y) \in \mathbb{R}^2$, es

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

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

   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 $\mathcal{B}_2$.

2. Si ahora consideramos la base canónica $\mathcal{B}=\{(1,0),\,(0,1)\}$ 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.

   Esto vale en general, si $\mathcal{B}=\{e_1,\dots,e_n\}$ es la base canónica de $\mathbb{R}^n$, entonces

$$\operatorname{coord}_{\mathcal{B}}(x_1,\dots,x_n)=(x_1,\dots,x_n), \quad \forall (x_1,\dots,x_n)\in\mathbb{R}^n.$$

3. Si consideramos la base canónica[17] $\mathcal{B}=\{1,\,x,\dots,x^n\}$ de $\mathbb{R}_n[x]$, entonces

$$\operatorname{coord}_{\mathcal{B}}\left(a_0+a_1x+a_2x^2+\cdots+a_nx^n\right)=(a_0,\,a_1,\,a_2,\,\cdots,\,a_n)\in\mathbb{R}^{n+1},$$

   para todo polinomio $a_0+a_1x+a_2x^2+\cdots+a_nx^n\in\mathbb{R}_n[x]$.

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

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

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

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

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

La matriz ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$ se llama la *matriz de cambio de base*[18] de $\mathcal{B}$ a $\mathcal{C}$.

Proposición 4.54 (Fórmula de cambio de base)**.**
Sean $\mathcal{B}$ y $\mathcal{C}$ dos bases de V. Si $v\in V$, entonces las coordenadas de v en las bases $\mathcal{B}$ y $\mathcal{C}$ se relacionan mediante

$$\operatorname{coord}_{\mathcal{C}}(v)={}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}\operatorname{coord}_{\mathcal{B}}(v). \tag{25}$$

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

Dem. Para simplificar la notación haremos la prueba en el caso $\dim V=2$, aunque la misma vale en general. Sea $v\in V$ y escribamos

$$\mathcal{B}=\{v_1,v_n\},\ \mathcal{C}=\{w_1,w_2\},\ {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}=(a_{ij}),\ \operatorname{coord}_{\mathcal{B}}(v)=(x_1,x_2),\ \operatorname{coord}_{\mathcal{C}}(v)=(y_1,y_2).$$

Esto quiere decir que vale

$$\begin{array}{l} v_1=a_{11}w_1+a_{21}w_2 \\ v_2=a_{12}w_1+a_{22}w_2 \end{array}; \quad v=x_1v_1+x_2v_2; \quad v=y_1w_1+y_2w_2. \tag{26}$$

[17] Acá hay que tener cuidado porque a veces se toma como base canónica de $\mathbb{R}_n[x]$ a $\{x^n,\,x^{n-1},\dots,1\}$. La conveniencia de una u otra depende de cómo escribamos los polinomios, por ejemplo $1+2x+3x^2$ o $3x^2+2x+1$.

[18] 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 ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$.

Luego

$$v = x_1 v_1 + x_2 v_2 = x_1 (a_{11} w_1 + a_{21} w_2) + x_2 (a_{12} w_1 + a_{22} w_2) = (x_1 a_{11} + x_2 a_{12}) w_1 + (x_1 a_{21} + x_2 a_{22}) w_2.$$

Luego comparando con la última igualdad de (26) obtenemos

$$(y_1, y_2) = (x_1 a_{11} + x_2 a_{12},\, x_1 a_{21} + x_2 a_{22}) \quad \Leftrightarrow \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Esta igualdad es la fórmula (25) en el caso $\dim V = 2$. □

Ejemplo 4.55. 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

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

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

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

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

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

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

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

Luego

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

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 $\mathcal{C}$.

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 siguiente resultado.

Lema 4.56. *Si una matriz $A \in M_n$ 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 $\{e_1, \ldots, e_n\}$ la base canónica de $\mathbb{R}^n$. Entonces como vale $Ae_i = e_i$, para todo $i = 1, \ldots, n$, deducimos

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

Proposición 4.57. *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

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

Como $\{\operatorname{coord}_{\mathcal{B}}(v) :\ v \in V\} = \mathbb{R}^n$, entonces la fórmula anterior implica

$$w = \left({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\, {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}\right) w, \quad \forall w \in \mathbb{R}^n.$$

Luego el lema 4.56 implica que es ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\, {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = I_n$. Esto a su vez implica que una es inversa de la otra. □

Ejemplo 4.58. Consideremos el espacio $\mathbb{R}^2$ y la base $\mathcal{B} = \{(1,1),\ (1,-1)\}$. Vanos a ver cómo escribir un vector genérico (x,y) como combinación lineal de la base $\mathcal{B}$, sin tener que resolver ningún sistema de ecuaciones. Sabemos que las coordenadas de (x,y) en la base canónica $\mathcal{C} = \{(1,0),\ (0,1)\}$ son $\operatorname{coord}_{\mathcal{C}}(x,y) = (x,y)$. Luego la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

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

Notar que la primer columna de la matriz ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}}$ 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 $\mathcal{C}$ a $\mathcal{B}$ es

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

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

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

es decir $\operatorname{coord}_{\mathcal{B}}(x,y) = \left(\frac{x+y}{2},\ \frac{x-y}{2}\right)$ 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 4.59 (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 $\mathcal{C} = \{\hat\imath,\ \hat\jmath\}$ 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 $\mathcal{B} = \{\hat\imath_\theta,\ \hat\jmath_\theta\}$, en que $\hat\imath_\theta$ y $\hat\jmath_\theta$ son los rotados de $\hat\imath$ y $\hat\jmath$, luego

$$\hat\imath_\theta = (\cos\theta)\hat\imath + (\operatorname{sen}\theta)\hat\jmath, \quad \hat\jmath_\theta = -(\operatorname{sen}\theta)\hat\imath + (\cos\theta)\hat\jmath. \tag{27}$$

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

$$v = x\hat\imath + y\hat\jmath \qquad \text{y} \qquad v = X\hat\imath_\theta + Y\hat\jmath_\theta.$$

Lo que queremos saber es cómo es que se relacionan (x,y) con (X,Y). La matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ se obtiene de las fórmulas (27):

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

La matriz de cambio de base de $\mathcal{C}$ a $\mathcal{B}$ se obtiene invirtiendo la matriz anterior

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

Luego aplicando la fórmula de cambio de base obtenemos

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

Esto nos dice que si la expresión de v en la base canónica es $v = x\hat{\imath} + y\hat{\jmath}$, entonces en la base $\mathcal{B}$ es

$$v = (x\cos\theta\, + y\,\mathrm{sen}\,\theta\,)\,\hat{\imath}_\theta + (-x\,\mathrm{sen}\,\theta\, + y\cos\theta\,)\,\hat{\jmath}_\theta.$$

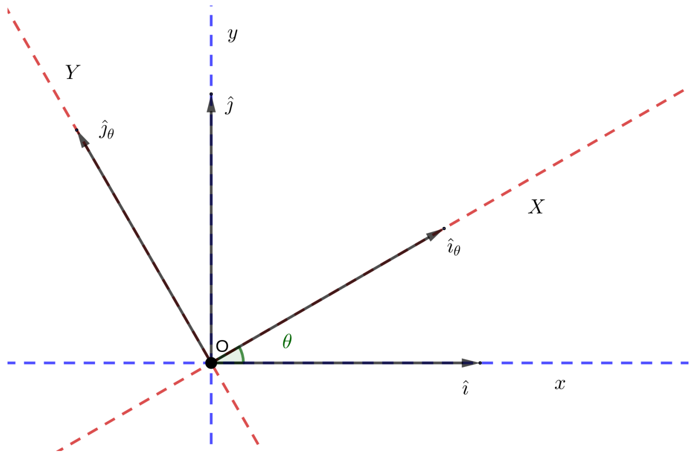

Figura 18: Rotación de ejes

5. Transformaciones lineales

En esta sección estudiaremos las funciones entre espacios vectoriales que preservan la estructura lineal. A menos de que explicitemos lo contrario, asumiremos siempre que los espacios tienen dimensión finita.

5.1. Definiciones y propiedades básicas

Definición 5.1. 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 alternativa de verificar si una función es una transformación lineal.

Proposición 5.2. *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 $a \in \mathbb{R}$.

Dem. Ejercicio. □

Ejemplos 5.3. Ejemplos de transformaciones lineales.

1. La función *nula* $T : V \to W$ definida por $T(v) = o$, para todo $v \in V$ (V y W 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. La *simetría respecto al eje Ox* es $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (x, -y)$, para todo $(x, y) \in \mathbb{R}^2$.
4. La *proyección* de $\mathbb{R}^3$ sobre $\mathbb{R}^2$ es $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T(x, y, z) = (x, y)$, para todo $(x, y, z) \in \mathbb{R}^3$.
5. 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.

6. Sea $I \subset \mathbb{R}$ un intervalo abierto y $C^\infty(I) := \{f : I \to \mathbb{R} : \ f \text{ es derivable de todos los órdenes}\}$. La función $T : C^\infty(I) \to C^\infty(I)$ definida por $T(f) = f'$ (la derivada) es una transformación lineal.
7. 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.

Proposición 5.4. *Sea $T : V \to W$ una transformación lineal. Vale lo siguiente.*

1. *Si $a_1, \dots, a_n \in \mathbb{R}$ y $v_1, \dots, v_n \in V$ ($n \geq 1$), entonces $T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n)$.*
2. *$T(o) = o$ (la imagen por T del vector nulo de V es el vector nulo de W).*
3. *Si $v \in V$, entonces $T(-v) = -T(v)$.*
4. *Si $u, v \in V$, entonces $T(u - v) = T(u) - T(v)$.*

Dem. La prueba de la primera afirmación se obtiene aplicando reiteradamente la primera propiedad de las transformaciones lineales y luego la segunda propiedad:

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

Las pruebas de las restantes afirmaciones son las siguientes

$$\begin{aligned}T(o)&=T(0\cdot o)=0\cdot T(o)=o,\\ T(-v)&=T\big((-1)\cdot v\big)=(-1)\cdot T(v)=-T(v),\\ T(u-v)&=T\big(u+(-v)\big)=T(u)+T(-v)=T(u)+\big(-T(v)\big)=T(u)-T(v).\quad\square\end{aligned}$$

Ejemplo 5.5. La función $T:\mathbb{R}^2\to\mathbb{R}^3$ definida por $T(x,y)=(2x+3y,\,3x-2y,\,x+y)$, es una transformación lineal. Para probarlo aplicamos la proposición 5.2.

$$\begin{aligned}T\big(a(x_1,y_1)+(x_2,y_2)\big)&=T(ax_1+x_2,\,ay_1+y_2)\\ &=\big(2(ax_1+x_2)+3(ay_1+y_2),\,3(ax_1+x_2)-2(ay_1+y_2),\,(ax_1+x_2)+(ay_1+y_2)\big)\\ &=\big(2ax_1+2x_2+3ay_1+3y_2,\,3ax_1+3x_2-2ay_1-2y_2,\,ax_1+x_2+ay_1+y_2\big)\\ &=\big(a(2x_1+3y_1)+2x_2+3y_2,\,a(3x_1-2y_1)+3x_2-2y_2,\,a(x_1+y_1)+x_2+y_2\big)\\ &=a\big(2x_1+3y_1,\,3x_1-2y_1,\,x_1+y_1\big)+\big(2x_2+3y_2,\,3x_2-2y_2),\,x_2+y_2\big)\\ &=aT(x_1,y_1)+T(x_2,y_2).\end{aligned}$$

Como esto vale para todo $a\in\mathbb{R}$ y $(x_1,y_1),(x_2,y_2)\in\mathbb{R}^2$, deducimos que T es lineal.

Observación 5.6. Veremos que toda transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^3$ es del tipo anterior. Consideremos una transformación lineal arbitraria $T:\mathbb{R}^2\to\mathbb{R}^3$. Sean $(a,b,c)=T(1,0)$ y $(d,e,f)=T(0,1)$. Entonces

$$T(x,y)=T\big(x(1,0)+y(0,1)\big)=xT(1,0)+yT(0,1)=x(a,b,c)+y(d,e,f)=(ax+dy,\,bx+ey,\,cx+fy).$$

Luego es $T(x,y)=(ax+dy,\,bx+ey,\,cx+fy)$. Esto que vimos vale en general.

Proposición 5.7. *Si $T:\mathbb{R}^n\to\mathbb{R}^m$ es una transformación lineal, entonces existen escalares a_{ij}, $1\le i\le m$ y $1\le j\le n$, tales que*

$$T(x_1,\dots,x_n)=(a_{11}x_1+\cdots+a_{1n}x_n,\ \dots,\ a_{m1}x_1+\cdots+a_{mn}x_n),\quad\forall(x_1,\dots,x_n)\in\mathbb{R}^n.$$

Además estos escalares están determinados por

$$T(e_1)=(a_{11},\,a_{21},\dots\,a_{m1}),\ T(e_2)=(a_{12},\,a_{22},\dots\,a_{m2}),\ \dots,\ T(e_n)=(a_{1n},\,a_{2n},\dots\,a_{mn}),$$

siendo $\{e_1,\dots,e_n\}$ la base canónica de $\mathbb{R}^n$. $\square$

Observación 5.8. 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 está escrito como un vector columna (para poderlo multiplicar por la matriz A). Es decir que 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 $Av\in M_{m\times 1}$. Vale

$$L_A(au+v)=A(au+v)=A(au)+Av=a(Au)+Av=aL_A(u)+L_A(v),\quad\forall a\in\mathbb{R},\ u,v\in\mathbb{R}^n.$$

Luego L_A es una transformación lineal. Más explícitamente, si escribimos $A=(a_{ij})$ y $v=(x_1,\dots,x_n)$, entonces

$$L_A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}a_{11}x_1+\cdots+a_{1n}x_n\\ \vdots\\ a_{m1}x_1+\cdots+a_{mn}x_n\end{pmatrix},$$

lo cual escrito en forma horizontal queda

$$L_A(x_1,\dots,x_n)=(a_{11}x_1+\cdots+a_{1n}x_n,\ \dots,\ a_{m1}x_1+\cdots+a_{mn}x_n),\quad\forall(x_1,\dots,x_n)\in\mathbb{R}^n.$$

Observando que la fórmula anterior coincide con la fórmula de la proposición 5.7, deducimos lo siguiente.

Proposición 5.9. *Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal, entonces existe una única matriz $A \in M_{m\times n}$ tal que $T = L_A$. Esta matriz es $A = [T(e_1)|\cdots|T(e_n)]$, siendo $\{e_1,\ldots,e_n\}$ la base canónica de $\mathbb{R}^n$.* □

Ejemplos 5.10. 1. 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

$$L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+3y \\ x+2y \\ -y \end{pmatrix} \quad \Rightarrow \quad L_A(x,y) = (2x+3y,\, x+2y,\, -y).$$

2. Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal definida por $T(x,y,z) = (x-y,\, 2x+3y+5z)$. Es $T(1,0,0) = (1,2)$, $T(0,1,0) = (-1,3)$ y $T(0,0,1) = (0,5)$. Luego es $T = L_A$, siendo $A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 3 & 5 \end{pmatrix}$.

Para espacios vectoriales en general tenemos el siguiente resultado.

Proposición 5.11. *Sean V y W dos espacios vectoriales. Si $\mathcal{B} = \{v_1,\ldots,v_n\}$ es una base de V y $w_1,\ldots,w_n$ son vectores arbitrarios de W, entonces existe una única transformación lineal $T : V \to W$ tal que $T(v_i) = w_i$, para todo $i = 1,\ldots,n$. Esta función está definida por*

$$T(x_1v_1 + \cdots + x_nv_n) = x_1w_1 + \cdots + x_nw_n, \tag{28}$$

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

Dem. Si existe una transformación lineal $T : V \to W$ que verifique $T(v_i) = w_i$, para todo i, entonces para todo $x_1,\ldots,x_n \in \mathbb{R}$ 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 $v_1,\ldots,v_n$. Luego tiene sentido definir una función $T : V \to W$ mediante la fórmula (28) y es fácil de probar que T definida de esa forma es una transformación lineal. □

Ejemplo 5.12. 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 4.53.1: $\operatorname{coord}_{\mathcal{B}}(x,y) = \left(\frac{x+y}{2},\, \frac{x-y}{2}\right)$. 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 $(x,y) \in \mathbb{R}^2$.

Ejemplo 5.13. Veremos que no existe una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(1,0) = (2,3,3)$ y $T(2,0) = (0,-1,1)$. Si existiese, entonces sería

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

Por otro lado, si buscamos una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ que solo verifique $T(1,1) = (2,3,3)$, 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 5.11 (el ejemplo anterior es un caso particular de esta construcción).

Terminamos esta parte considerando la composición de transformaciones lineales. 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 5.14. *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 también una transformación lineal.*

Dem. Sean $u, v \in V$ y $a \in \mathbb{R}$, entonces

$$(S \circ T)(au + v) = S\big(T(au + v)\big) = S\big(aT(u) + T(v)\big) = aS\big(T(u)\big) + S\big(T(v)\big) = a(S \circ T)(u) + (S \circ T)(v). \quad \square$$

Observación 5.15. Sean $A \in M_{m\times n}$ y $B \in M_{p\times m}$. Si 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 $L_B \circ L_A = L_{BA}$. Eso se debe al cálculo siguiente

$$(L_B \circ L_A)(v) = L_B\big(L_A(v)\big) = B(Av) = (BA)v = L_{BA}(v), \quad \forall v \in \mathbb{R}^n.$$

5.2. El núcleo y la imagen.

Definición 5.16. Sea $T : V \to W$ una transformación lineal. El *núcleo* de T es el conjunto[19] $\mathrm{Ker}(T) \subset V$ y su *imagen* o *recorrido* es el conjunto $\mathrm{Im}(T) \subset W$, definidos por

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

Proposición 5.17. *Si $T : V \to W$ es una transformación lineal, entonces* $\mathrm{Ker}(T)$ *es un subespacio de V y* $\mathrm{Im}(T)$ *es un subespacio de W.*

Dem. Probaremos solo que el núcleo es un subespacio, dejando la prueba de la imagen como ejercicio. Sabemos que vale $T(o) = o$, luego $o \in \mathrm{Ker}(T)$. Además, si $u, v \in \mathrm{Ker}(T)$ y $a \in \mathbb{R}$, entonces $T(au + v) = aT(u) + T(v) = ao + o = o$; luego $au + v \in \mathrm{Ker}(T)$. $\square$

Recordar que una función entre dos conjuntos $f : X \to Y$ se dice

- *inyectiva* si $x \neq x'$ implica $f(x) \neq f(x')$; esto equivale a $f(x) = f(x')$ implica $x = x'$;
- *sobreyectiva* si para todo $y \in Y$ existe $x \in X$ tal que $f(x) = y$;
- *biyectiva* si es inyectiva y sobreyectiva.

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

Proposición 5.18. *Una transformación lineal $T : V \to W$ es inyectiva si y solo si* $\mathrm{Ker}(T) = \{o\}$.

Dem. Supongamos que T es inyectiva. Si $v \neq o$, entonces $T(v) \neq T(o) = o$ y por lo tanto $v \notin \mathrm{Ker}(T)$. Luego $\mathrm{Ker}(T) = \{o\}$.

Supongamos $\mathrm{Ker}(T) = \{o\}$. Si $T(v_1) = T(v_2)$, entonces

$$T(v_1) - T(v_2) = o \quad \Rightarrow \quad T(v_1 - v_2) = o \quad \Rightarrow \quad v_1 - v_2 \in \mathrm{Ker}(T) = \{o\} \quad \Rightarrow \quad v_1 - v_2 = o \quad \Rightarrow \quad v_1 = v_2.$$

Luego T es inyectiva. $\square$

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

Observación 5.19. Es claro que una transformación lineal $T: V \to W$ es sobreyectiva si y solo si $\mathrm{Im}(T) = W$. Luego conociendo el núcleo y la imagen de T, podemos saber si T es inyectiva, sobreyectiva o biyectiva.

Proposición 5.20. *Las transformaciones lineales llevan conjuntos LD en conjuntos LD.*

Dem. Sea $T: V \to W$ es una transformación lineal y $\mathcal{A} = \{v_1, \dots, v_n\}$ un conjunto LD. Como $\mathcal{A}$ es LD, entonces existen $a_1, \dots, a_n \in \mathbb{R}$ no todos nulos tales que $a_1v_1 + \cdots + a_nv_n = o$, luego

$$T(a_1v_1 + \cdots + a_nv_n) = T(o) \quad \Rightarrow \quad a_1T(v_1) + \cdots + a_nT(v_n) = o.$$

Esto prueba que $T(\mathcal{A}) = \{T(v_1), \dots, T(v_n)\}$ es LD. $\square$

En general, las transformaciones lineales no llevan conjuntos LI en conjuntos LI, ni generadores en generadores. Por ejemplo, si consideramos $T: \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (x, 0)$, entonces $\mathcal{A} = \{(1, 0),\, (0, 1)\}$ es una base y por lo tanto es generador y LI, pero $T(\mathcal{A}) = \{(1, 0),\, (0, 0)\}$ no es generador ni es LI. Sin embargo, la proposición siguiente muestra que bajo ciertas condiciones las propiedades anteriores sí se cumplen.

Proposición 5.21. *Sea $T: V \to W$ una transformación lineal.*

1. *Si T es inyectiva, entonces T lleva conjuntos LI en conjuntos LI.*
2. *Si T es sobreyectiva, entonces T lleva generadores en generadores.*
3. *Si T es biyectiva, entonces T lleva conjuntos LI en conjuntos LI y generadores en generadores; luego T lleva bases en bases.*

Dem. La tercera afirmación se deduce de las dos anteriores. Así que necesitamos probar solo las dos primeras. Supongamos que T es inyectiva y que $\mathcal{A} = \{v_1, \dots, v_n\}$ es LI. Queremos probar que $T(\mathcal{A}) = \{T(v_1), \dots, T(v_n)\}$ es LI. Sean $a_1, \dots, a_n \in \mathbb{R}$ tales que $a_1T(v_1) + \cdots + a_nT(v_n) = o$. Entonces

$$\begin{aligned} &o = a_1T(v_1) + \cdots + a_nT(v_n) = T(a_1v_1 + \cdots + a_nv_n) \quad \Rightarrow \quad a_1v_1 + \cdots + a_nv_n \in \mathrm{Ker}(T) = \{o\} \\ &\quad \Rightarrow \quad a_1v_1 + \cdots + a_nv_n = o. \end{aligned}$$

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

Veamos ahora la segunda afirmación. Sea $\mathcal{A} = \{v_1, \dots, v_n\}$ un generador de V. Si $w \in W$, como T es sobreyectiva, entonces existe $v \in V$ tal que $w = T(v)$. Como $\mathcal{A}$ es un generador de V, entonces existen $a_1, \dots, a_n \in \mathbb{R}$ 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 W es combinación lineal de $T(\mathcal{A})$. $\square$

Observaciones 5.22. Sea $T: V \to W$ una transformación lineal arbitraria.

1. Si $\mathcal{A}$ es un generador de V, entonces $T(\mathcal{A})$ es un generador de $\mathrm{Im}(T)$. Esto se deduce de la proposición anterior, ya que $T: V \to \mathrm{Im}(T)$ es sobreyectiva.
2. Si $\mathcal{B}$ es una base de V, entonces $T(\mathcal{B})$ es un generador de $\mathrm{Im}(T)$, pero no es necesariamente una base (ver el ejemplo siguiente).

Ejemplo 5.23. Consideremos la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$T(x, y, z) = (x + y + z,\ x + y - z,\ 2x + 2y + 2z,\ -x - y + z). \tag{29}$$

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

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

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

$$\begin{cases} x+y+z &= a \\ x+y-z &= b \\ 2x+2y+2z &= c \\ -x-y+z &= d \end{cases} \tag{30}$$

es compatible. 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

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

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 (30) es compatible si y solo si $c = 2a$ y $d = -b$, por lo tanto la imagen de T es

$$\operatorname{Im}(T) = \left\{(x, y, z, t) \in \mathbb{R}^4 :\ z = 2x,\ t = -y\right\}. \tag{31}$$

Como vimos, determinar la imagen en la forma anterior es bastante más complicado que determinar el núcleo. Una alternativa es aplicar la observación anterior. 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

$$\operatorname{Im}(T) = [(1, 1, 2, -1), (1, -1, 2, 1)]. \tag{32}$$

Es un ejercicio el verificar que las dos formas de describir $\operatorname{Im}(T)$ dadas por (31) y (32), coinciden.

El resultado siguiente se utiliza con frecuencia.

Teorema 5.24. *Sea $T : V \to W$ una transformación lineal. Entonces*

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

Dem. Sea $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ una base de 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 generador de $\mathrm{Im}(T)$. Pero $T(v_1) = \cdots = T(v_n) = o$, luego $\mathcal{B}_3 = \{T(w_1), \ldots, T(w_m)\}$ es también un generador de $\mathrm{Im}(T)$. Veamos que $\mathcal{B}_3$ es LI. Sean $a_1, \ldots, a_m \in \mathbb{R}$ tales que $a_1T(w_1) + \cdots + a_mT(w_m) = o$. Luego

$$T(a_1w_1 + \cdots + a_mw_m) = o \quad \Rightarrow \quad a_1w_1 + \cdots + a_mw_m \in \mathrm{Ker}(T) = [v_1, \ldots, v_n].$$

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

$$a_1w_1 + \cdots + a_mw_m = b_1v_1 + \cdots + b_nv_n \quad \Rightarrow \quad b_1v_1 + \cdots + b_nv_n + (-a_1)w_1 + \cdots + (-a_m)w_m = o$$

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 $\mathrm{Im}(T)$. Entonces

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

Observación 5.25. Dada una transformación lineal $T : V \to W$, se llama *nulidad* de T a la dimensión de su núcleo y *rango* de T a la dimensión de su imagen. 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 5.26. Sea $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por

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

Razonando como en el ejercicio 5.23 deducimos que $\{(-1, -1, 1)\}$ es una base de $\mathrm{Ker}(T)$ y que el conjunto $G = \{(1, 1, 0, 1),\ (1, 0, 1, -1),\ (2, 1, 1, 0)\}$ es un generador de $\mathrm{Im}(T)$. Luego $\dim \mathrm{Ker}(T) = 1$ y como $\dim \mathbb{R}^3 = 3$, entonces el teorema anterior implica $\dim \mathrm{Im}(T) = 2$. Como el generador G de $\mathrm{Im}(T)$ tiene tres elementos, entonces G es LD. Para obtener una base de $\mathrm{Im}(T)$ alcanza con tomar dos elementos de G que sean LI; en este caso cualquier par de esos elementos lo verifica, luego

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

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

Corolario 5.27. *Sea $T : V \to W$ una transformación lineal.*

1. *Si* $\dim V > \dim W$*, entonces T no es inyectiva.*
2. *Si* $\dim V < \dim W$*, entonces T no es sobreyectiva.*

Dem. Si $\dim V > \dim W$, entonces

$$\dim \mathrm{Ker}(T) = \dim V - \dim \mathrm{Im}(T) > \dim W - \dim \mathrm{Im}(T) \geq 0 \quad \Rightarrow \quad \dim \mathrm{Ker}(T) > 0 \quad \Rightarrow \quad \mathrm{Ker}(T) \neq \{o\}.$$

Si $\dim V < \dim W$, entonces

$$\dim \mathrm{Im}(T) = \dim V - \dim \mathrm{Ker}(T) \leq \dim V < \dim W \quad \Rightarrow \quad \dim \mathrm{Im}(T) < \dim W. \quad \square$$

Isomorfismos. A las transformaciones lineales biyectivas se les llama *isomorfismos*. Si existe un isomorfismo entre dos espacios vectoriales V y W, entonces decimos que V y W son *isomorfos*. Desde el punto de vista del álgebra lineal, dos espacios isomorfos tienen las mismas propiedades y se suelen identificar.

Ejemplo 5.28. En general se identifica $\mathbb{R}^2$ con el plano $z = 0$ en $\mathbb{R}^3$. Esto se debe a que el mapa $T : \mathbb{R}^2 \to W = \{(x, y, z) \in \mathbb{R}^3 : \ z = 0\}$ definido por $T(x, y) = (x, y, 0)$, para todo $(x, y) \in \mathbb{R}^2$, es un isomorfismo.

Ejemplo 5.29. Todo espacio de dimensión n es isomorfo a $\mathbb{R}^n$. Esto se debe a que si $\dim V = n$ 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

$$f \circ f^{-1} = \mathrm{id}_Y \quad \text{y} \quad f^{-1} \circ f = \mathrm{id}_X. \tag{33}$$

De hecho es equivalente que f sea biyectiva con que f sea *invertible*, es decir, que exista una función $f^{-1} : Y \to X$ que verifique (33). Además, si f es biyectiva, entonces (33) implica que $f^{-1} : Y \to X$ es biyectiva y su inversa es f, es decir $\left(f^{-1}\right)^{-1} = f$.

Proposición 5.30. *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 también es 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 5.21 se deduce la siguiente.

Proposición 5.31. *Si existe un isomorfismo $T : V \to W$, entonces $\dim V = \dim W$.* $\square$

Por la proposición anterior, solo pueden existir isomorfismos 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 5.32. *Sea $T : V \to W$ 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. Eso sale de la fórmula

$$\dim \operatorname{Ker}(T) + \dim \operatorname{Im}(T) = \dim V. \tag{34}$$

Sabemos que T es inyectiva si y solo si $\operatorname{Ker}(T) = \{0\}$ y que T es sobreyectiva si y solo si $\operatorname{Im}(T) = W$. Además

$$\operatorname{Ker}(T) = \{0\} \Leftrightarrow \dim \operatorname{Ker}(T) = 0 \overset{(34)}{\Leftrightarrow} \dim \operatorname{Im}(T) = \dim V \Leftrightarrow \dim \operatorname{Im}(T) = \dim W \Leftrightarrow \operatorname{Im}(T) = W.$$

Luego T es inyectiva si y solo si es sobreyectiva. $\square$

El siguiente resultado describe los isomorfismos de $\mathbb{R}^n$ en $\mathbb{R}^m$. Notar que la proposición (5.31) implica que solo puede existir algún isomorfismo cuando es $n = m$.

Proposición 5.33. *Sea $A \in M_n$. Si consideramos la 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. En ese caso vale $(L_A)^{-1} = L_{A^{-1}}$.*

Dem. Supongamos que A es invertible. Entonces

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

Esto prueba que L_A es invertible (y por lo tanto es un isomorfismo) y que vale $(L_A)^{-1} = L_{A^{-1}}$. Supongamos ahora que L_A es un isomorfismo. Como $(L_A)^{-1}$ es una transformación lineal, entonces existe $B \in M_n$ tal que $(L_A)^{-1} = L_B$. Luego

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

Esto implica que A es invertible y $B = A^{-1}$. $\square$

5.3. Matriz asociada

En esta sección las bases siempre son ordenadas (como en la sección 4.6 de matrices de cambio de base).

Sea $T : V \to W$ una transformación lineal. Sean $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{C} = \{w_1, \ldots, w_m\}$ bases de V y W respectivamente. Consideramos las coordenadas en la base $\mathcal{C}$ de las imágenes de los elementos de $\mathcal{B}$ por medio de T:

$$\begin{array}{rcl} T(v_1) & = & a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m \\ \vdots & & \\ T(v_n) & = & a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m \end{array} \quad \Rightarrow \quad \begin{array}{rcl} \mathrm{coord}_{\mathcal{C}}\big(T(v_1)\big) & = & (a_{11},\, a_{21},\, \ldots,\, a_{m1}) \\ & \vdots & \\ \mathrm{coord}_{\mathcal{C}}\big(T(v_n)\big) & = & (a_{1n},\, a_{2n},\, \ldots,\, a_{mn}) \end{array}.$$

Con estos vectores escritos en columna construimos la matriz

$${}_{\mathcal{C}}[T]_{\mathcal{B}} := \left[\ \mathrm{coord}_{\mathcal{C}}\big(T(v_1)\big) \ \Big|\ \cdots \ \Big|\ \mathrm{coord}_{\mathcal{C}}\big(T(v_n)\big)\ \right] = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

La matriz ${}_{\mathcal{C}}[T]_{\mathcal{B}} \in M_{m\times n}$ se llama la *matriz asociada* a la transformación lineal T de la base $\mathcal{B}$ a la base $\mathcal{C}$.

Ejemplos 5.34.

1. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (x+y, x+2y, 2x+y)$. 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

   $$\begin{array}{rcl} T(1,1) & = & (2,3,3) \\ T(1,-1) & = & (0,-1,1) \end{array} \quad \Rightarrow \quad \begin{array}{rcl} \mathrm{coord}_{\mathcal{C}}\big(T(1,1)\big) & = & (2,3,3) \\ \mathrm{coord}_{\mathcal{C}}\big(T(1,-1)\big) & = & (0,-1,1) \end{array} \quad \Rightarrow \quad {}_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 3 & -1 \\ 3 & 1 \end{pmatrix}.$$

   Observar que en este caso el cálculo fue fácil porque $\mathcal{C}$ es la base canónica.

2. Consideremos la base $\mathcal{B}$ anterior y la base $\mathcal{D} = \{(1, 1, 1),\, (1, 1, 0),\, (1, 0, 0)\} \subset \mathbb{R}^3$. Calculando las coordenadas de un vector genérico (x, y, z) en la base $\mathcal{D}$, obtenemos

   $$(x, y, z) = z(1, 1, 1) + (y - z)(1, 1, 0) + (x - y)(1, 0, 0) \quad \Rightarrow \quad \mathrm{coord}_{\mathcal{D}}(x, y, z) = (z,\, y - z,\, x - y).$$

   Luego si T es la misma del ejercicio anterior, obtenemos

   $$\begin{array}{rcl} T(1,1) & = & (2,3,3) \\ T(1,-1) & = & (0,-1,1) \end{array} \quad \Rightarrow \quad \begin{array}{rcl} \mathrm{coord}_{\mathcal{D}}\big(T(1,1)\big) & = & (3,0,-1) \\ \mathrm{coord}_{\mathcal{D}}\big(T(1,-1)\big) & = & (1,-2,1) \end{array} \quad \Rightarrow \quad {}_{\mathcal{D}}[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 1 \\ 0 & -2 \\ -1 & 1 \end{pmatrix}.$$

3. Sea $V=\mathbb{R}_2[x]$ y $T:V\to V$ la derivación, es decir, $T\big(p(x)\big)=p'(x)$, para todo $p(x)\in\mathbb{R}_2[x]$. Vamos a hallar ${}_{\mathcal{B}}[T]_{\mathcal{B}}$, siendo $\mathcal{B}=\{1,\,x,\,x^2\}$. Es

$$\begin{array}{rcl} T(1) & = & 0 \\ T(x) & = & 1 \\ T(x^2) & = & 2x \end{array} \quad \Rightarrow \quad \begin{array}{rcl} \operatorname{coord}_{\mathcal{B}}\big(T(1)\big) & = & (0,0,0) \\ \operatorname{coord}_{\mathcal{B}}\big(T(x)\big) & = & (1,0,0) \\ \operatorname{coord}_{\mathcal{B}}\big(T(x^2)\big) & = & (0,2,0) \end{array} \quad \Rightarrow \quad {}_{\mathcal{B}}[T]_{\mathcal{B}}=\begin{pmatrix} 0 & 1 & 0\\ 0 & 0 & 2\\ 0 & 0 & 0\end{pmatrix}.$$

En la proposición y prueba siguientes mantenemos las notaciones previas a los ejemplos anteriores.

Proposición 5.35. *Sea $T:V\to V$ una transformación lineal y $\mathcal{B},\mathcal{C}$ bases de V. Entonces*

$$\operatorname{coord}_{\mathcal{C}}\big(T(v)\big)={}_{\mathcal{C}}[T]_{\mathcal{B}}\cdot\operatorname{coord}_{\mathcal{B}}(v), \tag{35}$$

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.*

Dem. Sea $v\in V$. Si $\operatorname{coord}_{\mathcal{B}}(v)=(x_1,\dots,x_n)$, entonces $v=x_1v_1+x_2v_2+\cdots+x_nv_n$ y por lo tanto

$$\begin{aligned} T(v)&=T\big(x_1v_1+\cdots+x_nv_n\big)\\ &=x_1T\big(v_1\big)+\cdots+x_nT\big(v_n\big)\\ &=x_1\left(a_{11}w_1+\cdots+a_{n1}w_n\right)+\cdots+x_n\left(a_{1n}w_1+\cdots+a_{nn}w_n\right)\\ &=\left(x_1a_{11}+\cdots+x_na_{1n}\right)w_1+\cdots+\left(x_1a_{n1}+\cdots+x_na_{nn}\right)w_n. \end{aligned}$$

Luego

$$\operatorname{coord}_{\mathcal{C}}\big(T(v)\big)=(x_1a_{11}+\cdots+x_na_{1n},\,\cdots,\,x_1a_{n1}+\cdots+x_na_{nn}).$$

Esta igualdad es la fórmula (35). □

Observaciones 5.36. 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. De hecho la prueba de la fórmula de cambio de base es un caso particular de la prueba anterior.

2. Sea $A\in M_{m,n}$. Si consideramos $T=L_A:\mathbb{R}^n\to\mathbb{R}^m$ y las bases canónicas respectivas $\mathcal{B}\subset\mathbb{R}^n$ y $\mathcal{C}\subset\mathbb{R}^m$, entonces ${}_{\mathcal{C}}[T]_{\mathcal{B}}=A$ y la fórmula (35) queda en $T(v)=A\cdot v$, para todo $v\in V$.

El siguiente resultado muestra que la ecuación (35) caracteriza la matriz asociada.

Proposición 5.37. *Sea $T:V\to W$ una transformación lineal, y $\mathcal{B}$ y $\mathcal{C}$ bases de V y W, respectivamente. Si una matriz A verifica*

$$\operatorname{coord}_{\mathcal{C}}\big(T(v)\big)=A\cdot\operatorname{coord}_{\mathcal{B}}(v), \tag{36}$$

para todo $v\in V$, entonces $A={}_{\mathcal{C}}[T]_{\mathcal{B}}$.

Dem. Sea $\mathcal{B}=\{v_1,\dots,v_n\}$. Aplicando (36) a los elementos de $\mathcal{B}$ obtenemos

$$\operatorname{coord}_{\mathcal{C}}\big(T(v_i)\big)=A\cdot\operatorname{coord}_{\mathcal{B}}(v_i)=A\cdot e_i,$$

para todo i, siendo $\{e_1,\dots,e_n\}$ la base canónica de $\mathbb{R}^n$. Como $A\cdot e_i$ nos da la columna i-ésima de A, concluimos

$$A=[A\cdot e_1|\cdots|A\cdot e_n]=\big[\operatorname{coord}_{\mathcal{C}}\big(T(v_1)\big)|\cdots|\operatorname{coord}_{\mathcal{C}}\big(T(v_n)\big)\big]={}_{\mathcal{C}}[T]_{\mathcal{B}}.\quad\square$$

Si conocemos las coordenadas de un vector en una base, entonces conocemos el vector. Luego la fórmula (35) nos dice que la matriz asociada determina la transformación lineal. El siguiente ejemplo muestra cómo obtener la transformación conociendo su matriz asociada.

Ejemplo 5.38. Consideremos las bases $\mathcal{B} = \{1,\ 1+x\}$ de $\mathbb{R}_1[x]$ y $\mathcal{C} = \{(1,1,1),\ (1,1,0),\ (1,0,0)\}$ de $\mathbb{R}^3$. Supongamos que $T : \mathbb{R}_1[x] \to \mathbb{R}^3$ es una transformación lineal de la cual conocemos su matriz asociada

$$_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Nos interesa recuperar T. Lo primero es hallar las coordenadas de un polinomio genérico en la base $\mathcal{B}$:

$$a + bx = (a-b)1 + b(1+x) \quad \Rightarrow \quad \mathrm{coord}_{\mathcal{B}}(a+bx) = (a-b,\ b).$$

Luego aplicando (35) obtenemos

$$\mathrm{coord}_{\mathcal{C}}\big(T(a+bx)\big) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a-b \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ b \\ a-b \end{pmatrix}.$$

Esto implica

$$T(a+bx) = (a+b)(1,1,1) + b(1,1,0) + (a-b)(1,0,0) = (2a+b,\ a+2b,\ a+b).$$

Luego T está definida por $T(a+bx) = (2a+b,\ a+2b,\ a+b)$, para todo $a + bx \in \mathbb{R}_1[x]$.

Proposición 5.39. *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*

$$_{\mathcal{D}}[S \circ T]_{\mathcal{B}} = {}_{\mathcal{D}}[S]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[T]_{\mathcal{B}}.$$

Dem. Sabemos $\mathrm{coord}_{\mathcal{C}}\big(T(v)\big) = {}_{\mathcal{C}}[T]_{\mathcal{B}} \cdot \mathrm{coord}_{\mathcal{B}}(v)$ y $\mathrm{coord}_{\mathcal{D}}\big(S(w)\big) = {}_{\mathcal{D}}[S]_{\mathcal{C}} \cdot \mathrm{coord}_{\mathcal{C}}(w)$ para todo $v \in V$ y $w \in W$. Luego

$$\begin{aligned}
\mathrm{coord}_{\mathcal{D}}\big((S \circ T)(v)\big) &= \mathrm{coord}_{\mathcal{D}}\Big(S\big(T(v)\big)\Big) = {}_{\mathcal{D}}[S]_{\mathcal{C}} \cdot \mathrm{coord}_{\mathcal{C}}\big(T(v)\big) = {}_{\mathcal{D}}[S]_{\mathcal{C}} \cdot \Big({}_{\mathcal{C}}[T]_{\mathcal{B}} \cdot \mathrm{coord}_{\mathcal{B}}(v)\Big) \\
&= \Big({}_{\mathcal{D}}[S]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[T]_{\mathcal{B}}\Big) \cdot \mathrm{coord}_{\mathcal{B}}(v).
\end{aligned}$$

Como esto vale para todo $v \in V$, entonces la proposición 5.37 implica ${}_{\mathcal{D}}[S]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[T]_{\mathcal{B}} = {}_{\mathcal{D}}[S \circ T]_{\mathcal{B}}$. □

El siguiente resultado muestra cómo se relacionan las matrices asociadas a una misma transformación lineal en bases distintas.

Corolario 5.40. *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.

Dem. Usando la proposición anterior es ${}_{\mathcal{C}_2}[\mathrm{Id}]_{\mathcal{C}_1} \cdot {}_{\mathcal{C}_1}[T]_{\mathcal{B}_1} \cdot {}_{\mathcal{B}_1}[\mathrm{Id}]_{\mathcal{B}_2} = {}_{\mathcal{C}_2}[\mathrm{Id} \circ T \circ \mathrm{Id}]_{\mathcal{B}_2} = {}_{\mathcal{C}_2}[T]_{\mathcal{B}_2}$. □

Observación 5.41. Es fácil de verificar que si consideramos la función identidad $\mathrm{Id} : V \to V$ y una base arbitraria $\mathcal{B}$ de V, entonces ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}} = I$ (la matriz identidad). Eso lo aplicaremos a continuación.

Proposición 5.42. *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. Si T es un isomorfismo, entonces existe $T^{-1} : V \to V$ tal que $T \circ T^{-1} = T^{-1} \circ T = \mathrm{Id}$. Esto implica

$$I = {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}} = {}_{\mathcal{B}}[T^{-1} \circ T]_{\mathcal{B}} = {}_{\mathcal{B}}\left[T^{-1}\right]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[T]_{\mathcal{B}}.$$

Eso muestra que la matriz ${}_{\mathcal{C}}[T]_{\mathcal{B}}$ es invertible y que vale $\left({}_{\mathcal{C}}[T]_{\mathcal{B}}\right)^{-1} = {}_{\mathcal{B}}\left[T^{-1}\right]_{\mathcal{C}}$.

Recíprocamente, supongamos ahora que la matriz $A = {}_{\mathcal{C}}[T]_{\mathcal{B}}$ es invertible. Si $v \in \mathrm{Ker}(T)$, entonces

$$A \cdot \mathrm{coord}_{\mathcal{B}}(v) = {}_{\mathcal{C}}[T]_{\mathcal{B}} \cdot \mathrm{coord}_{\mathcal{B}}(v) = \mathrm{coord}_{\mathcal{C}}\big(T(v)\big) = \mathrm{coord}_{\mathcal{C}}\big(o\big) = o.$$

Como A es invertible esto implica $\mathrm{coord}_{\mathcal{B}}(v) = o$, lo cual a su vez implica $v = o$. Esto prueba $\mathrm{Ker}(T) = \{o\}$ y por lo tanto $T : V \to V$ es inyectiva; luego T es un isomorfismo (teorema 5.32). □

Observación 5.43. La proposición 5.33 es un caso particular de la proposición anterior, tomando como $\mathcal{B}$ y $\mathcal{C}$ a la base canónica de $\mathbb{R}^n$. Sin embargo las pruebas que vimos son distintas, así que ambas tienen interés. Notar que de la proposición anterior también se deduce la proposición 4.57.

5.4. Aplicación a sistemas de ecuaciones

En esta sección veremos cómo los sistemas de ecuaciones (sección 1) se pueden estudiar usando álgebra lineal y completaremos la demostración del teorema de Cramer de la sección 3.4. Consideremos el sistema

$$\left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array} \right. . \tag{37}$$

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 (37) es equivalente a la ecuación matricial

$$AX = B.$$

La ventaja de la ecuación $AX = B$ es que tiene una sola incógnita (matricial). Para el estudio del sistema (37) conviene considerar el *sistema homogéneo asociado*:

$$\left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & 0 \end{array} \right. \tag{38}$$

y la correspondiente ecuación matricial

$$AX = O,$$

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

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

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

A continuación veremos cómo se relacionan las soluciones del sistema homogéneo (38) con las del sistema no homogéneo (37).

Proposición 5.44. *1. Si $X_1, X_2 \in \mathcal{S}_B$, entonces $X_1 - X_2 \in \mathcal{S}_O$.*

2. Si $X_1 \in \mathcal{S}_B$ y $X_0 \in \mathcal{S}_O$, entonces $X_1 + X_0 \in \mathcal{S}_B$.

3. Si $\mathcal{S}_B \neq \emptyset$ y $X_1 \in \mathcal{S}_B$, entonces $\mathcal{S}_B = \{X_1 + X : \ X \in \mathcal{S}_0\}$.

Dem.

1. $X_1, X_2 \in \mathcal{S}_B \ \Rightarrow \ A(X_1 - X_2) = AX_1 - AX_2 = B - B = O \ \Rightarrow \ X_1 - X_2 \in \mathcal{S}_O$.
2. $X_1 \in \mathcal{S}_B, \ X_O \in \mathcal{S}_O \ \Rightarrow \ A(X_1 + X_O) = AX_1 + AX_O = B + O = B \ \Rightarrow \ X_1 + X_O \in \mathcal{S}_B$.
3. Sea $X_1 \in \mathcal{S}_B$. La segunda parte implica $\{X_1 + X : \ X \in \mathcal{S}_0\} \subset \mathcal{S}_B$. Recíprocamente, si $Y \in \mathcal{S}_B$, entonces la primera parte implica $Y - X_1 \in \mathcal{S}_O$. Luego escribiendo $Y = X_1 + (Y - X_1)$ deducimos $Y \in \{X_1 + X : \ X \in \mathcal{S}_0\}$. Esto prueba la inclusión contraria y por lo tanto $\mathcal{S}_B = \{X_1 + X : \ X \in \mathcal{S}_0\}$. □

Observaciones 5.45. 1. La igualdad $\mathcal{S}_0 = \mathrm{Ker}(L_A)$ implica que $\mathcal{S}_0$ es un subespacio de $\mathbb{R}^n$. Por lo tanto si $\mathcal{S}_0 \neq \{0\}$, entonces $\mathcal{S}_0$ es un conjunto infinito. Es decir, si un sistema homogéneo admite alguna solución no trivial, entonces tiene infinitas soluciones.

2. La primera parte de la proposición anterior muestra que si un sistema no homogéneo es compatible indeterminado, entonces el sistema homogéneo asociado también lo es y por lo tanto tiene infinitas soluciones. Luego la tercera parte de la proposición anterior implica que el sistema no homogéneo también tiene infinitas soluciones. Es decir, lo anterior prueba que un sistema compatible indeterminado (homogéneo o no) siempre tiene infinitas soluciones.
3. Un sistema no homogéneo 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 es la trivial (por la parte 2 de la proposición anterior). Esto último depende solo de la matriz del sistema A y no de la columna de resultados B.

Proposición 5.46. *1. Si en un sistema homogéneo hay menos ecuaciones que incógnitas o hay la misma cantidad pero el determinante del sistema es nulo, entonces el sistema admite soluciones no triviales.*

2. Si en un sistema no homogéneo hay menos ecuaciones que incógnitas o hay la misma cantidad pero el determinante del sistema es nulo, entonces el sistema es incompatible o compatible indeterminado.

Dem. La segunda afirmación se deduce de la primera, así que probaremos solo la primera.

Supongamos que el sistema homogéneo tiene m ecuaciones y n incógnitas y que $A \in M_{m \times n}$ es la matriz del sistema. Si consideramos la transformación lineal asociada $L_A : \mathbb{R}^n \to \mathbb{R}^m$, entonces la tesis equivale a probar $\mathrm{Ker}(L_A) \neq \{0\}$.

Caso $n > m$. Consideremos la transformación lineal asociada $L_A : \mathbb{R}^n \to \mathbb{R}^m$. Como es $n > m$, entonces el corolario 5.27 implica que L_A no es inyectiva. Luego $\mathrm{Ker}(L_A) \neq \{0\}$.

Caso $n = m$ y $\det(A) = 0$. Como es $\det(A) = 0$, entonces $L_A : \mathbb{R}^n \to \mathbb{R}^n$ no es un isomorfismo (proposición 5.33) y por lo tanto L_A no es inyectiva (teorema 5.32). Luego $\mathrm{Ker}(L_A) \neq \{0\}$. $\square$

Observación 5.47. La proposición anterior termina la prueba del teorema de Cramer (teorema 3.48), dado que muestra que si es $\Delta = 0$, entonces el sistema es incompatible o compatible indeterminado.

Observación 5.48. Si en un sistema de ecuaciones hay más ecuaciones que incógnitas, entonces no podemos afirmar nada. Por ejemplo, si consideramos los siguientes sistemas

$$\left\{\begin{array}{rcl} x+y &=& 2 \\ x-y &=& 0 \\ 2x+y &=& 3 \end{array}\right. , \quad \left\{\begin{array}{rcl} x+y &=& 2 \\ x-y &=& 0 \\ 2x+y &=& 1 \end{array}\right. , \quad \left\{\begin{array}{rcl} x+y &=& 2 \\ 2x+2y &=& 4 \\ 3x+3y &=& 6 \end{array}\right. ,$$

entonces el primero es compatible determinado, el segundo es incompatible y el tercero es compatible indeterminado.

Referencias

[Friedberg] Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence, *Linear Algebra*, Prentice Hall, 1999.

[Resnick] Robert Resnick, David Halliday, *Física. Parte I*, Compañía Editorial Continental S.A, 1975.