Notas para el curso de Álgebra Lineal 1

Andrés Abella

29 de junio de 2023

Introducción

Estas son notas para el curso de Álgebra Lineal 1 de la Licenciatura en Matemática de la Facultad de Ciencias de la Universidad de la República. A continuación veremos una breve descripción de cada capítulo.

Capítulo 1. Sistemas de ecuaciones. Este tema se estudia en los cursos de enseñanza secundaria, por lo que aparece tratado como un repaso, con énfasis en la parte operativa. La teoría y otras técnicas de resolución se verán al finalizar el capítulo 3.

Capítulo 2. Se estudian las propiedades de los vectores en el espacio y se ven aplicaciones a rectas y planos. Estos temas tienen interés por sus aplicaciones geométricas, pero además importan porque los espacios vectoriales, que son el tema central de estas notas, son el resultado de generalizar estas propiedades.

Capítulo 3. Comienza con el estudio de las matrices y sus operaciones; suma, producto, etc. Luego se introducen las operaciones elementales y la equivalencia de matrices, lo cual juega un rol central en este capítulo. Después se introducen los determinantes y se prueban sus propiedades. Finalmente se ven aplicaciones a sistemas de ecuaciones.

Capítulo 4. Se introducen los espacios vectoriales, con énfasis en los espacios dimensión finita. Se definen los subespacios y se estudian las operaciones entre subespacios. Luego se introducen los conceptos de conjunto linealmente independiente, conjunto linealmente dependiente, generador y base, y se estudian sus propiedades. Finalmente se ven las definiciones necesarias para trabajar en dimensión infinita, aunque sin profundizar en el tema.

Capítulo 5. Se estudian de las transformaciones lineales entre espacios vectoriales, concentrándonos en las propiedades de los espacios de dimensión finita. Luego se estudia el rango de matrices y de transformaciones lineales. El capítulo finaliza con una breve sección sobre espacio dual.

Apéndice A. Contiene un repaso de conceptos básicos de teoría de conjuntos, funciones y relaciones de equivalencia. La idea es que sirva como referencia para los capítulos anteriores.

Convenciones y notaciones. El símbolo $\lightning$ indica que hemos llegado a una contradicción y $\square$ indica que hemos llegado al final de una demostración. Si X es un conjunto finito, entonces $\#X$ indica la cantidad de elementos de X, a la cual llamamos el *cardinal* de X.

Al conjunto de los números reales lo escribiremos $\mathbb{R}$, al de los racionales $\mathbb{Q}$ al de los enteros $\mathbb{Z}$ y al de los naturales $\mathbb{N} = \{0, 1, 2, \dots\}$. Para que las notas tengan un carácter lo más universal posible, en general trabajaremos sobre un cuerpo arbitrario $\Bbbk$, salvo en los temas de geometría en los cuales el cuerpo será el de los números reales $\mathbb{R}$. Sin embargo, dado que estas notas son para un curso del primer semestre de la licenciatura, los ejemplos van a ser siempre en los reales, así que si se prefiere se puede pensar que el cuerpo $\Bbbk$ es $\mathbb{R}$, aunque lo que veremos vale para cualquier cuerpo[1], como ser el de los racionales, los complejos, los enteros módulo un primo, etc.

[1] Hay algunas pocas propiedades que no valen en cualquier cuerpo, eso será explicitado al tratarlo.

Índice

Capítulo 1

Sistemas de ecuaciones

En este tema el énfasis va a estar en la parte práctica, la fundamentación teórica se verá en capítulos posteriores. En lo que sigue asumiremos que estamos trabajando con números reales, aunque en general lo que veremos vale también para otros cuerpos numéricos, como los números racionales o los complejos.

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

$$\left\{ \begin{array}{rcl} 2x+y-z & = & 3 \\ 3x+2y+z & = & -5 \end{array} \right. , \quad \left\{ \begin{array}{rcl} x+y & = & 1 \\ 2x+3y & = & 0 \\ -x+3y & = & -2 \end{array} \right. , \quad \left\{ \begin{array}{rcl} 2x+y & = & 1 \\ x+3y & = & 2 \end{array} \right. , \quad \left\{ \begin{array}{rcl} x+y+z & = & 0 \\ 2x+3y-2z & = & 0 \\ -x+3y-z & = & 0 \end{array} \right. . \tag{1.1}$$

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 (1.1), 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

$$\left\{ \begin{array}{rcl} 7x+2y & = & 16 \\ 3x+5y & = & 11 \end{array} \right. . \tag{1.2}$$

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

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

Si ahora sustituimos ese valor de y en la primera ecuación del sistema (1.2), 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 (1.2) mutiplicamos la primera 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 1.1.1. 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.

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

1.2. Sistemas en general

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

$$\left\{\begin{array}{rcl} a_{11}x_1+\cdots+a_{1n}x_n & = & b_1 \\ a_{21}x_1+\cdots+a_{2n}x_n & = & b_2 \\ & \vdots & \\ a_{m1}x_1+\cdots+a_{mn}x_n & = & b_m \end{array}\right.. \tag{1.3}$$

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

Una *solución* de un sistema de ecuaciones, es una asignación de valores a las variables de forma tal que se verifiquen todas las ecuaciones del sistema Si un sistema admite alguna solución, entonces se dice que es *compatible* y 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 veremos que un sistema indeterminado siempre tiene infinitas soluciones. Cuando nos referimos a *resolver* un sistema, entendemos que es clasificarlo en incompatible, compatible determinado o compatible 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[1]. Usaremos el símbolo $\sim$ para indicar la equivalencia de sistemas. Vale el siguiente resultado. Para simplificar la exposición omitiremos la prueba, aunque es bastante simple y se puede hacer como ejercicio.

Proposición 1.2.1. *Operaciones con las ecuaciones de un sistema que dan lugar a sistemas equivalentes.*

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

El método para resolver sistemas de ecuaciones conocido como de *eliminación Gaussiana* o *de 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 1.2.2. Queremos resolver el siguiente sistema

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

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

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 1.2.3.

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

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

Ejemplo 1.2.4.

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

Observar que los coeficientes de este sistema son los mismos que los del sistema del ejemplo anterior, así que los pasos de resolución son los mismos (solo cambian los términos independientes). Como la última ecuación no aporta nada, entonces 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 las soluciones del sistema se pueden escribir de la forma

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

Notar que para cada valor que le demos a λ vamos a obtener una solución, por ejemplo para $\lambda=0$ obtenemos la solución $x=\frac{1}{7}$, $y=-\frac{4}{7}$, $z=0$, para $\lambda=1$ obtenemos $x=\frac{6}{7}$, $y=-\frac{11}{7}$, $z=1$, etc. Otra manera de escribir las soluciones es $x=-z+\frac{1}{7}$, $y=-z-\frac{4}{7}$ y la variable z está libre. Este es un ejemplo de un sistema compatible indeterminado con *un grado de libertad*, dado que la variable z está libre y fijado su valor, los otros quedan determinados.

Ejemplo 1.2.5.

$$\begin{cases} x+y-z &= 1 \\ -x-y+z &= -1 \\ 2x+2y-2z &= 2 \end{cases} \sim \begin{cases} x+y-z &= 1 \\ 0 &= 0 \\ 0 &= 0 \end{cases}.$$

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 1.2.6. Un sistema 2×3:

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

Acá restamos la primera 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 1.2.7. Un sistema 3×2:

$$\begin{cases} x+2y &= 1 \\ 2x-y &= 2 \\ -x+3y &= -1 \end{cases} \sim \begin{cases} x+2y &= 1 \\ -5y &= 0 \\ 5y &= 0 \end{cases} \sim \begin{cases} x+2y &= 1 \\ -5y &= 0 \\ 0 &= 0 \end{cases}.$$

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

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n &= 0 \end{cases}$$

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 1.3.1.

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

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

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

Ejemplo 1.3.2.

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

El sistema es compatible indeterminado. Una forma de escribir sus soluciones es $x = -z$, $y = -z$, 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 1.3.3.

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

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

Ejemplo 1.3.4. Un sistema 2×3:

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

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

Ejemplo 1.3.5. Un sistema 3×2:

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

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

Capítulo 2

Geometría

En este capítulo daremos una fundamentación matemática para los vectores que se estudian en física y veremos aplicaciones geométricas, en particular al estudio de rectas y planos. No vamos a ser muy formales y a menudo usaremos argumentos geométricos que no justificaremos. Al final del capítulo se incluye una sección donde se ve cómo se puede formalizar todo lo anterior.

2.1. Vectores

Mediante la introducción de un sistema de coordenadas ortogonales, se puede identificar el plano con el conjunto $\mathbb{R}^2 = \{(x, y) : \ x, y \in \mathbb{R}\}$. Así, cuando por ejemplo se escribe "el punto $P = (1, 2)$", se quiere decir que $(1, 2)$ son las coordenadas del punto P respecto al sistema de referencia.

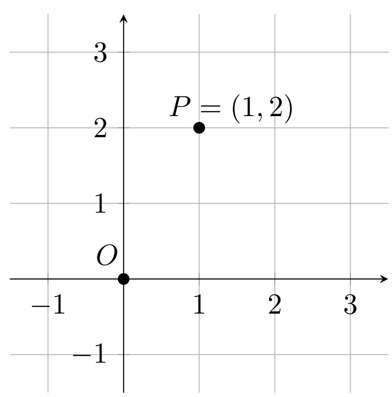

Figura 2.1: Punto

Lo mismo se puede hacer con el espacio, identificándolo con el conjunto $\mathbb{R}^3 = \{(x, y, z) : \ x, y, z \in \mathbb{R}\}$. En general usaremos letras mayúsculas P, Q, R, etc., para los puntos del espacio, siendo $O = (0, 0, 0)$ el *origen* del sistema. En el caso del espacio, el sistema de coordenadas tiene tres ejes ortogonales, que llamamos Ox, Oy y Oz. Al plano que contiene a los ejes Ox y Oy le llamamos el plano Oxy; en forma análoga se definen los planos Oxz y Oyz.

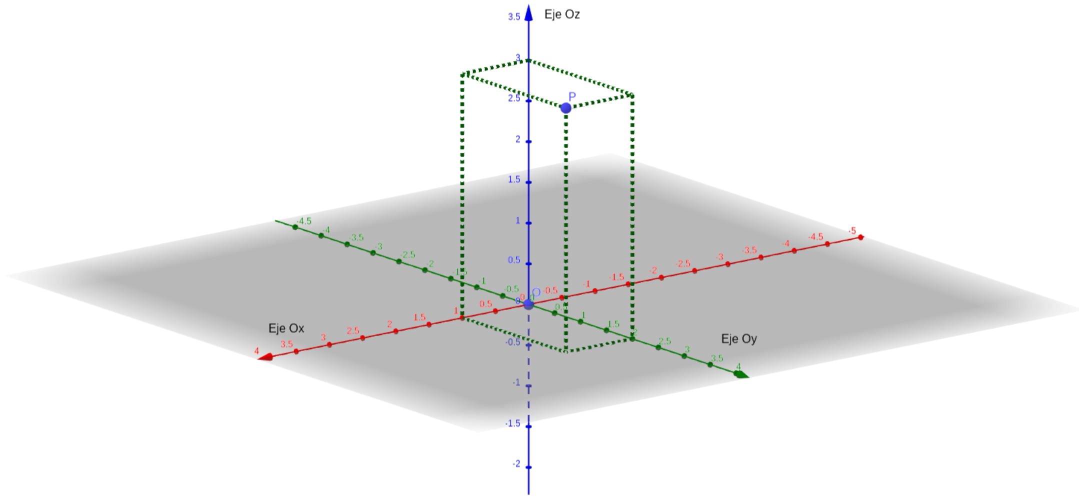

Figura 2.2: El punto $P = (1, 2, 3)$

A cada punto P del espacio, con $P \neq O$, le podemos hacer corresponder el vector $v = \overrightarrow{OP}$, que es la flecha que va de O a P. Si además al origen O le hacemos corresponder el vector nulo, entonces obtenemos una correspondencia uno a uno entre los puntos del espacio y los vectores que salen del origen; por esa razón a los puntos del espacio también les llamamos *vectores*. Así, cuando por ejemplo nos referimos al vector $v = (1, 2, 3)$, estamos diciendo que es $v = \overrightarrow{OP}$, con $P = (1, 2, 3)$. En general, cuando a los puntos los pensemos como vectores, usaremos las letras minúsculas u, v, w; en particular al vector nulo lo escribiremos $o = (0, 0, 0)$.

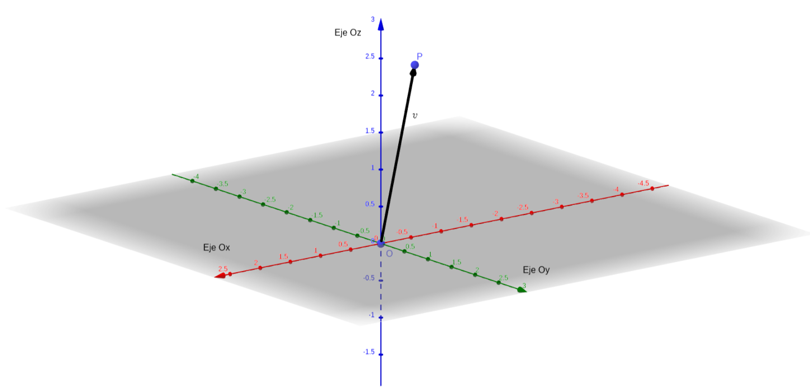

Figura 2.3: Correspondencia punto-vector

Al trabajar con puntos o vectores, a los números se les suele llamar *escalares* y en general los escribiremos con las letras minúsculas a, b, c. En esta sección vamos a trabajar principalmente con vectores, pero al trabajar con rectas y planos usaremos puntos y vectores.

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

$$a(x_0, y_0, z_0) := (ax_0, ay_0, az_0), \quad (x_1, y_1, z_1) + (x_2, y_2, z_2) := (x_1 + x_2, y_1 + y_2, z_1 + z_2),$$

para todo $a \in \mathbb{R}$, $(x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$.

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 2.1.1. *Si los puntos $O = (0, 0, 0)$, $P = (x_1, y_1, z_1)$, $Q = (x_2, y_2, z_2)$ y $R = (x_3, y_3, z_3)$ son los vértices de un paralelogramo como en la figura 2.4, entonces $x_3 = x_1 + x_2$, $y_3 = y_1 + y_2$ y $z_3 = z_1 + z_2$.*

Dem.

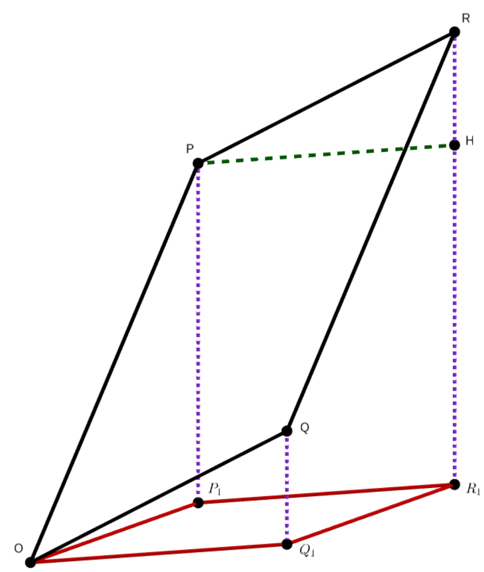

Figura 2.4: Suma de vectores

Para entender la prueba conviene seguir el razonamiento en la figura 2.4. Vamos a probar que vale $z_3 = z_1 + z_2$, las otras igualdades se prueban en forma análoga.

Sea $\mathcal{P}$ el paralelogramo de vértices O, P, R, Q. Proyectamos ortogonalmente $\mathcal{P}$ sobre el plano Oxy, obteniendo los puntos P_1, R_1, Q_1. La altura de los puntos P, Q, R corresponde a sus coordenadas en el eje Oz, por lo cual lo que tenemos que probar es que vale $d(R_1, R) = d(P_1, P) + d(Q_1, Q)$, siendo $d(R_1, R)$ la distancia entre R_1 y R, etc. Sea H la intersección de la recta $\overline{R_1R}$ con el plano horizontal que pasa por P.

Consideremos el cuadrilátero $\mathcal{C}$ de vértices P, H, R_1, P_1. Las rectas $\overline{P_1P}$ y $\overline{R_1H}$ son verticales (paralelas al eje Oz) mientras que las rectas $\overline{P_1R_1}$ y $\overline{PH}$ están en planos horizontales (paralelos al plano Oxy).

Esto implica que los cuatro ángulos del cuadrilátero son rectos[1] y por lo tanto $\mathcal{C}$ es un rectángulo. Luego $d(P_1, P) = d(R_1, H)$.

Consideremos ahora $\mathcal{T}_1$ el triángulo de vértices P, H, R y $\mathcal{T}_2$ el triángulo de vértices O, Q, Q_1. Nosotros ya vimos que el ángulo $\widehat{PHR}$ es recto y por el mismo argumento el ángulo $\widehat{OQ_1Q}$ también es recto. Por otro lado, las rectas $\overline{OQ}$ y $\overline{PR}$ son paralelas (por ser lados opuestos del paralelogramo $\mathcal{P}$) y $\overline{QQ_1}$ y $\overline{RH}$ son paralelas (por ser ambas paralelas al eje Oz), luego los ángulos $\widehat{OQQ_1}$ y $\widehat{PRH}$ coinciden. Como además vale $d(OQ) = d(PR)$ (por ser lados opuestos del paralelogramo $\mathcal{P}$), conluimos que los triángulos $\mathcal{T}_1$ y $\mathcal{T}_2$ son congruentes. Eso a su vez implica $d(Q_1, Q) = d(H, R)$.

Juntando lo anterior, obtenemos $z_3 = d(R_1, R) = d(R_1, H) + d(H, R) = d(P_1, P) + d(Q_1, Q) = z_1 + z_2$. $\square$

Es fácil de probar, y lo dejamos como ejercicio, que el producto de un escalar por un vector coincide con el definido en física (ver la figura 2.5).

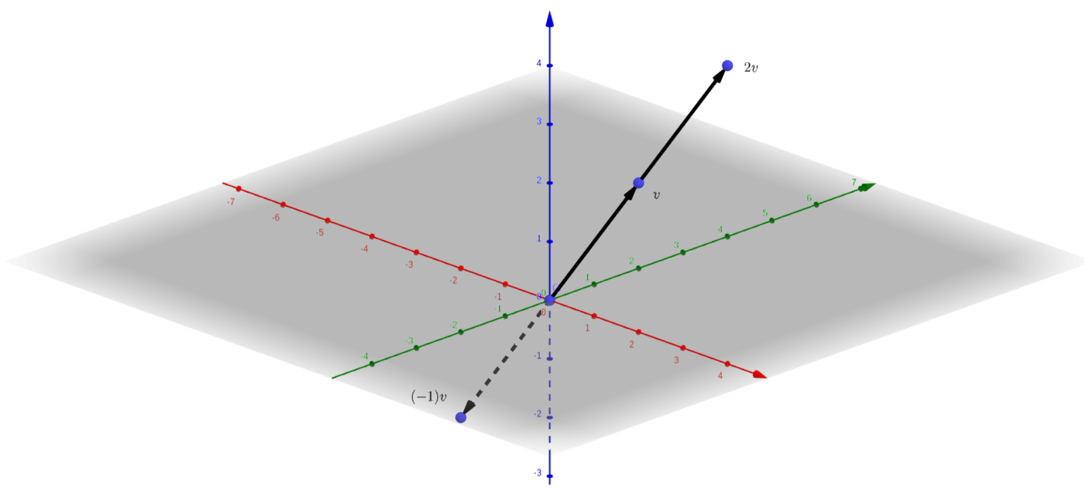

Figura 2.5: Producto de un escalar por un vector

Si $v = (x, y, z)$, entonces definimos su *opuesto* mediante $-v := (-x, -y, -z)$.

Proposición 2.1.2. *Las operaciones recién definidas verifican las siguientes propiedades:*

1. $u + v = v + u$, *para todo* $u, v \in \mathbb{R}^3$ *(conmutativa);*
2. $u + (v + w) = (u + v) + w$, *para todo* $u, v, w \in \mathbb{R}^3$ *(asociativa);*
3. $u + o = o + u = u$, *para todo* $u \in \mathbb{R}^3$ *(existencia de neutro);*
4. $u + (-u) = (-u) + u = o$, *para todo* $u \in \mathbb{R}^3$ *(existencia de opuesto);*

[1] Recordar que si una recta r es perpendicular a un plano Π, entonces r es perpendicular a toda recta contenida en Π que pase por la intersección de r con Π.

5. $1u = u$, *para todo* $u \in \mathbb{R}^3$*;*

6. $a(u+v) = au + av$, *para todo* $a \in \mathbb{R}$*,* $u, v \in \mathbb{R}^3$*;*

7. $(a+b)u = au + bu$, *para todo* $a, b \in \mathbb{R}$*,* $u \in \mathbb{R}^3$*;*

8. $a(bu) = (ab)u$, *para todo* $a, b \in \mathbb{R}$*,* $u \in \mathbb{R}^3$*.*

Dem. Las pruebas son rutinarias, así que demostraremos solo la propiedad asociativa, dejando las otras como ejercicio. Sean $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2)$ y $w = (x_3, y_3, z_3)$. Entonces

$$\begin{aligned} u + (v + w) &= (x_1, y_1, z_1) + \big((x_2, y_2, z_2) + (x_3, y_3, z_3)\big) = (x_1, y_1, z_1) + (x_2 + x_3, y_2 + y_3, z_2 + z_3) \\ &= \big(x_1 + (x_2 + x_3), y_1 + (y_2 + y_3), z_1 + (z_2 + z_3)\big), \\ (u + v) + w &= \big((x_1, y_1, z_1) + (x_2, y_2, z_2)\big) + (x_3, y_3, z_3) = (x_1 + x_2, y_1 + y_2, z_1 + z_2) + (x_3, y_3, z_3) \\ &= \big((x_1 + x_2) + x_3, (y_1 + y_2) + y_3, (z_1 + z_2) + z_3\big). \end{aligned}$$

Como la suma de números reales es asociativa, valen

$$x_1 + (x_2 + x_3) = (x_1 + x_2) + x_3, \quad y_1 + (y_2 + y_3) = (y_1 + y_2) + y_3 \quad \text{y} \quad z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3.$$

Luego $u + (v + w) = (u + v) + w$. □

Observaciones 2.1.3. 1. Notar que la prueba anterior de la propiedad asociativa es simple, sin embargo probarla geométricamente usando la regla del paralelogramo parece bastante más complicado.

2. Las propiedades anteriores caracterizan a los *espacios vectoriales*, que veremos más adelante.

Diremos que dos vectores u y v son *colineales* o *paralelos* si existe un escalar a tal que $u = av$ o $v = au$.

Observaciones 2.1.4. 1. Si es $u = av$ con $a \neq 0$, entonces es $v = a^{-1}u$.

2. Si u y v son no nulos y $u = av$, entonces u y v tienen el mismo sentido cuando $a > 0$ y tienen sentidos opuestos cuando $a < 0$ (figura 2.5).

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

Norma. Dado un vector $v = (x_0, y_0, z_0)$, definimos su *norma*[2] mediante $\|v\| := \sqrt{x_0^2 + y_0^2 + z_0^2}$. La norma mide la longitud del vector v (es decir, la distancia del origen O al punto P, si $v = \overrightarrow{OP}$). Para verlo, notar que si aplicamos el teorema de Pitágoras al triángulo de vértices $(0,0,0)$, $(x_0, 0, 0)$ y $(x_0, y_0, 0)$, obtenemos que la distancia del origen O al punto $(x_0, y_0, 0)$ es $\sqrt{x_0^2 + y_0^2}$. Si ahora aplicamos el teorema de Pitágoras al triángulo de vértices $(0,0,0)$, $(x_0, y_0, 0)$ y (x_0, y_0, z_0), obtenemos la fórmula de arriba.

Proposición 2.1.5. *La norma verifica las siguientes propiedades:*

1. $\|v\| \geq 0$, *para todo* $v \in \mathbb{R}^3$*;*

2. $\|v\| = 0$ *si y solo si* $v = o$*;*

3. $\|av\| = |a|\,\|v\|$, *para todo* $a \in \mathbb{R}$*,* $v \in \mathbb{R}^3$*;*

4. $\|u + v\| \leq \|u\| + \|v\|$, *para todo* $u, v \in \mathbb{R}^3$ *(*desigualdad triangular*).*

[2]En física a la norma se le suele llamar el *módulo*.

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 $u+v$ (figura 2.6), 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). □

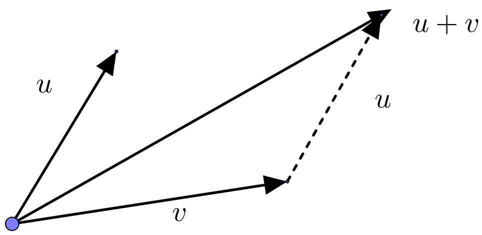

Figura 2.6: Desigualdad triangular

La *resta* de vectores se define por

$$(x_1, y_1, z_1) - (x_2, y_2, z_2) := (x_1 - x_2, y_1 - y_2, z_1 - z_2),$$

para todo $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$. Notar que vale $w = u - v$ si y solo si $u = v + w$.

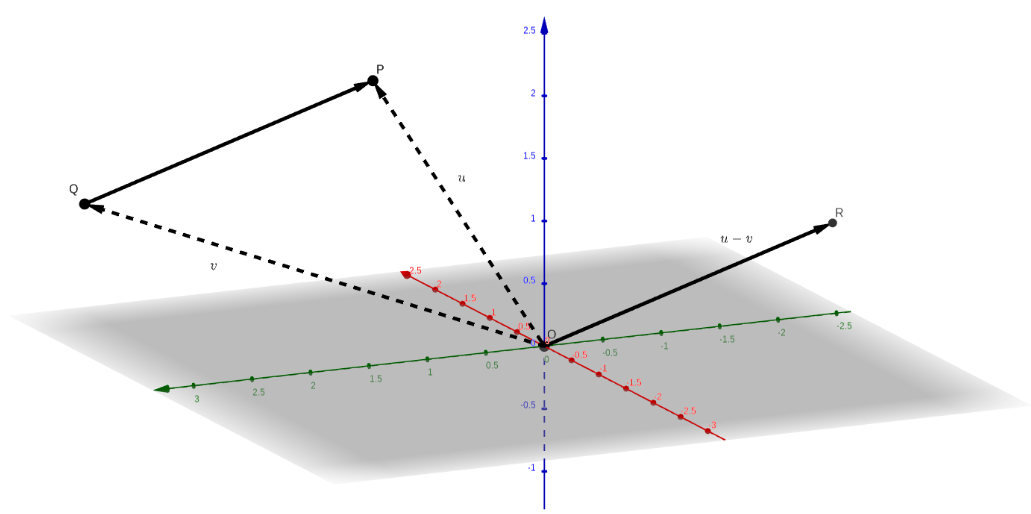

Figura 2.7: Resta de vectores

Observación 2.1.6. Si $u = \overrightarrow{OP}$ y $v = \overrightarrow{OQ}$, entonces $u - v = \overrightarrow{OR}$, siendo R como en la figura 2.7. Luego la distancia entre los puntos P y Q coincide con las norma del vector $u - v$. Luego la *distancia* entre $u = (x_1, y_1, z_1)$ y $v = (x_2, y_2, z_2)$, se calcula mediante

$$d(u, v) := \|u - v\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

La *base canónica* del espacio $\mathbb{R}^3$ es el conjunto $\mathcal{B} = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, donde $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$ y $\mathbf{k} = (0, 0, 1)$.

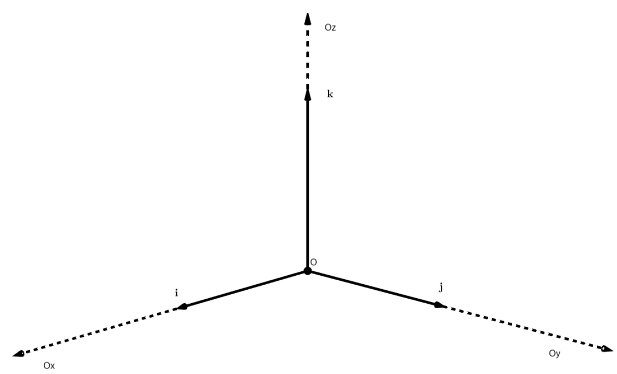

Figura 2.8: Ejes y base canónica en $\mathbb{R}^3$

Notar que vale[3]

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

Un *versor* es un vector de norma 1. Por ejemplo, los vectores de la base canónica $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ son versores. Los versores se suelen usar para expresar direcciones. Notar que 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.

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

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

siendo $0 \le \theta \le \pi$ el ángulo[4] que forman u y v. Si u o v es nulo, se define $u \cdot v = 0$.

Teorema 2.1.7. *Si* $u = (x_1, y_1, z_1)$ *y* $v = (x_2, y_2, z_2)$, *entonces*

$$u \cdot v = x_1x_2 + y_1y_2 + z_1z_2, \tag{2.2}$$

Dem. Es claro que la fórmula (2.2) 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 2.7), 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

$$\begin{aligned} u \cdot v &= \frac{1}{2}\left\{\|u\|^2 + \|v\|^2 - \|u - v\|^2\right\} \\ &= \frac{1}{2}\left\{x_1^2 + y_1^2 + z_1^2 + x_2^2 + y_2^2 + z_2^2 - \left((x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2\right)\right\} \\ &= \frac{1}{2}\left\{x_1^2 + y_1^2 + z_1^2 + x_2^2 + y_2^2 + z_2^2 - \left(x_1^2 - 2x_1x_2 + x_2^2 + y_1^2 - 2y_1y_2 + y_2^2 + z_1^2 - 2z_1z_2 + z_2^2\right)\right\} \\ &= x_1x_2 + y_1y_2 + z_1z_2. \quad \square \end{aligned}$$

Notar que en términos de la base canónica, la fórmula (2.2) se escribe:

$$(x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{j}) \cdot (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{j}) = x_1x_2 + y_1y_2 + z_1z_2.$$

[3]Si $v = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, entonces en física se suele decir que x, y, z son las *componentes* de v en las direcciones de $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$.

[4]Los ángulos los mediremos siempre en radianes.

Proposición 2.1.8. *El producto escalar verifica las siguientes propiedades.*

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

Dem. Ejercicio (usar la fórmula (2.2)). □

Observación 2.1.9. La propiedad conmutativa implica que las propiedades 2 y 3 valen también "del otro lado":

$$u \cdot (v + w) = u \cdot v + u \cdot w, \quad \forall u, v, w \in \mathbb{R}^3; \qquad u \cdot (av) = a(u \cdot v), \quad \forall a \in \mathbb{R},\ u, v \in \mathbb{R}^3.$$

Teorema 2.1.10 (Desigualdad de Cauchy-Shwarz)**.** *Para todo* $u, v \in \mathbb{R}^3$ *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 (2.1) 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. □

Observación 2.1.11. El producto escalar sirve para calcular ángulos: si u y v son dos vectores no nulos, entonces el ángulo θ $(0 \leq \theta \leq \pi)$ que forman u y v queda determinado por $\cos(\theta) = \frac{u \cdot v}{\|u\|\|v\|}$.

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 **i**, **j** y **k** de la base canónica son ortogonales dos a dos. Notar que el vector nulo es ortogonal a cualquier otro vector del espacio; es el único vector que verifica esta propiedad[5].

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 $\frac{u \cdot v}{\|u\|}$. Notar que vale $\frac{u \cdot v}{\|u\|} = \|v\| \cos(\theta)$, siendo θ el ángulo entre v y u.

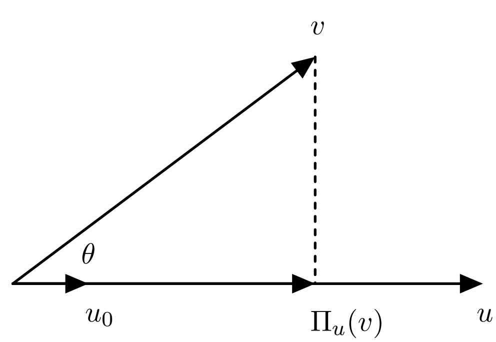

Figura 2.9: Proyección ortogonal

[5]Esto se debe al cálculo siguiente: si $v \cdot u = 0,\ \forall u \ \Rightarrow\ v \cdot v = 0 \ \Rightarrow\ \|v\|^2 = 0 \ \Rightarrow\ \|v\| = 0 \ \Rightarrow\ v = o$.

Ejemplo 2.1.12. Sean $u = \mathbf{i} + \mathbf{j}$ y $v = 2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$. 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

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

El plano. Las fórmulas que vimos para el espacio valen también para el plano $\mathbb{R}^2 = \{(x, y) : \ x, y \in \mathbb{R}\}$, "omitiendo la coordenada z". Por ejemplo la suma de dos vectores, el producto de un escalar por un vector, el producto escalar de dos vectores y la norma de un vector, están definidos por

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \quad a(x, y) = (ax, ay), \quad (x_1, y_1) \cdot (x_2, y_2) = x_1 x_2 + y_1 y_2,$$
$$\|(x_1, y_1)\| = \sqrt{x_1^2 + y_1^2}.$$

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 $\mathbf{i} = (1, 0)$ en $\mathbb{R}^2$ y también $\mathbf{i} = (1, 0, 0)$ en $\mathbb{R}^3$; lo mismo sucede con $\mathbf{j} = (0, 1)$ y $\mathbf{j} = (0, 1, 0)$. En ese sentido, si escribimos $v = x\mathbf{i} + y\mathbf{j}$ y estamos en $\mathbb{R}^2$, entonces es $v = (x, y)$, pero si estamos en $\mathbb{R}^3$ es $v = (x, y, 0)$.

Producto vectorial. 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 $0 < \theta < \pi$ 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*, que dice que si uno cierra esa mano de u hacia v, entonces el pulgar apunta en el sentido de $u \times v$.

Si u y v son colineales, se define $u \times v = o$ (el vector nulo).

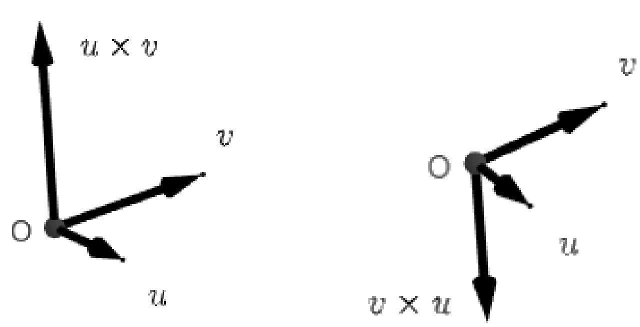

Figura 2.10: Producto vectorial

Proposición 2.1.13. *El producto vectorial verifica las siguientes propiedades.*

1. $v \times u = -u \times v$, *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 $a \in \mathbb{R}$, $u, v \in \mathbb{R}^3$;* $\square$

Observaciones 2.1.14. 1. Las propiedades 1 y 3 de la proposición 2.3 son fáciles de probar. La propiedad 2 no lo parece tanto, pero también asumiremos que se deduce de la definición de producto vectorial.

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 realizamos los productos vectoriales entre los elementos de la base canónica de $\mathbb{R}^3$, obtenemos

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

Proposición 2.1.15. *Vale la fórmula siguiente*

$$(x_1, y_1, z_1) \times (x_2, y_2, z_2) = (y_1 z_2 - z_1 y_2,\ z_1 x_2 - x_1 z_2,\ x_1 y_2 - y_1 x_2). \tag{2.3}$$

Dem. La prueba consiste en desarrollar el producto escribiendo los vectores en función de la base canónica, y usar las fórmulas anteriores y la proposición 2.1.13.

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

Observación 2.1.16. Si tomamos la fórmula (2.3) como definición de producto vectorial, entonces todas las propiedades de la proposición 2.1.13 se prueban fácilmente.

Observación 2.1.17. Una manera de recordar la fórmula (2.3) es utilizando los determinantes. El *determinante* es un número que se asocia a un cuadro de números. Más adelante estudiaremos el tema en detalle, pero ahora solo necesitamos los determinantes de orden 2 y 3, que definimos a continuación.

El determinante de orden 2 está definido por

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

y el determinante de orden 3 está definido por

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} := a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

Por ejemplo,

$$\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = 2 \times 5 - 3 \times 4 = 10 - 12 = -2,$$

$$\begin{vmatrix} 2 & 3 & -2 \\ 4 & 5 & 1 \\ -1 & 0 & 2 \end{vmatrix} = 2 \begin{vmatrix} 5 & 1 \\ 0 & 2 \end{vmatrix} - 3 \begin{vmatrix} 4 & 1 \\ -1 & 2 \end{vmatrix} + (-2) \begin{vmatrix} 4 & 5 \\ -1 & 0 \end{vmatrix} = 2 \times 10 - 3 \times 9 - 2 \times 5 = -14.$$

Con estas notaciones, si calculamos el siguiente "determinante" obtenemos[6]

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

Esta última fórmula coincide con el lado derecho de (2.3) y por lo tanto

$$(x_1, y_1, z_1) \times (x_2, y_2, z_2) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}. \tag{2.4}$$

Ejemplo 2.1.18. Vamos a calcular el producto vectorial de $(1, 2, 3)$ por $(4, 5, 6)$.

$$(1, 2, 3) \times (4, 5, 6) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \mathbf{k} = -3\mathbf{i} + 6\mathbf{j} - 3\mathbf{k} = (-3, 6, -3).$$

Luego $(1, 2, 3) \times (4, 5, 6) = (-3, 6, -3)$.

Proposición 2.1.19. *Si u y v son dos vectores con u no nulo, 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 $v = v_1 + v_2$.*

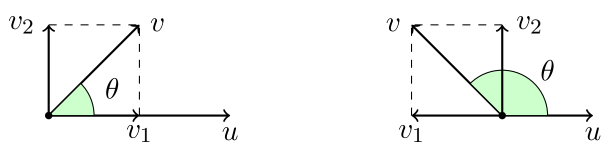

Figura 2.11: Descomposición $v = v_1 + v_2$

Además $\|v_1\| = \frac{|u \cdot v|}{\|u\|}$ *y* $\|v_2\| = \frac{\|u \times v\|}{\|u\|}$.

Dem. Empezamos probando la unicidad. Supongamos que existen v_1 y v_2 que verifican la tesis. Como v_1 es colineal con u, entonces existe $a \in \mathbb{R}$ tal que $v_1 = au$. Calculando la proyección ortogonal de v sobre u, obtenemos

$$\Pi_u(v) = \frac{u \cdot v}{\|u\|^2} u = \frac{u \cdot (v_1 + v_2)}{\|u\|^2} u = \frac{u \cdot v_1 + u \cdot v_2}{\|u\|^2} u = \frac{u \cdot (au) + 0}{\|u\|^2} u = \frac{a\|u\|^2}{\|u\|^2} u = au = v_1.$$

Entonces $v_1 = \Pi_u(v)$ y por lo tanto $v_2 = v - v_1 = v - \Pi_u(v)$. Por lo tanto, si existen v_1 y v_2 que verifican la tesis, entonces $v_1 = \Pi_u(v)$ y $v_2 = v - \Pi_u(v)$. Esto prueba la unicidad.

Recíprocamente, si definimos $v_1 := \Pi_u(v)$ y $v_2 := v - \Pi_u(v)$, entonces es claro que v_1 es colineal con u y que $v = v_1 + v_2$. Además, 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 $v_2 = v - \Pi_u(v)$ es ortogonal a u. Esto termina la prueba de la existencia.

[6] Escribimos "determinante" entre comillas, porque lo que aparece en (2.4) no es en realidad un determinante, dado que involucra escalares y vectores. En realidad la fórmula (2.4) es solo una manera de recordar la definición del producto vectorial.

Ahora calcularemos $\|v_1\|$ y $\|v_2\|$. La suma $v = v_1 + v_2$ con v_1 y v_2 ortogonales, implica (ver figura 2.11) que es $\|v_1\| = \|v\| \, |\cos(\theta)|$ y $\|v_2\| = \|v\| \operatorname{sen}(\theta)$, siendo $0 \leq \theta \leq \pi$ el ángulo entre u y v. Notar que en $\|v_1\|$ tenemos que tomar $|\cos(\theta)|$, porque puede ser $\frac{\pi}{2} < \theta \leq \pi$. Luego

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

Ejemplo 2.1.20. Sean $u = \mathbf{i} + \mathbf{j}$ y $v = 2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$. En el ejemplo 2.1.12 vimos que la proyección de v en la dirección de u es $\Pi_u(v) = \frac{3}{2}\mathbf{i} + \frac{3}{2}\mathbf{j}$. Restando obtenemos $v - \Pi_u(v) = \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + 2\mathbf{k}$. Luego tenemos la descomposición

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

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

Corolario 2.1.21. *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 $bh = \|u \times v\|$. $\square$

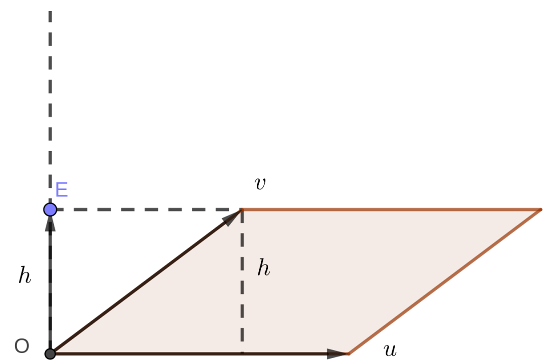

Figura 2.12: Paralelogramo

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

Proposición 2.1.22. *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. Aplicando las fórmulas del producto vectorial y escalar ((2.3) y (2.2), respectivamente), obtenemos

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

Proposición 2.1.23. *Vale*
$$(u \times v) \cdot w = u \cdot (v \times w),$$
para todo $u, v, w \in \mathbb{R}^3$.

Dem. Sean $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2)$ y $w = (x_3, y_3, z_3)$, entonces

$$\begin{aligned} u \cdot (v \times w) &= (x_1, y_1, z_1) \cdot \big((x_2, y_2, z_2) \times (x_3, y_3, z_3)\big) \\ &= (x_1, y_1, z_1) \cdot (y_2z_3 - z_2y_3,\ z_2x_3 - x_2z_3,\ x_2y_3 - y_2x_3) \\ &= x_1(y_2z_3 - z_2y_3) + y_1(z_2x_3 - x_2z_3) + z_1(x_2y_3 - y_2x_3) = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}. \end{aligned}$$

Luego usando la fórmula de la proposición anterior obtenemos la tesis. □

Observación 2.1.24. Para quienes sepan del tema, la fórmula anterior también se puede probar usando propiedades de los determinantes, pero no es realmente necesario.

Proposición 2.1.25. *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 2.1.19) $h = \frac{|(u\times v)\cdot w|}{\|u\times v\|}$. Entonces el volumen es $ah = \|u \times v\| \frac{|(u\times v)\cdot w|}{\|u\times v\|} = |(u \times v) \cdot w|$. □

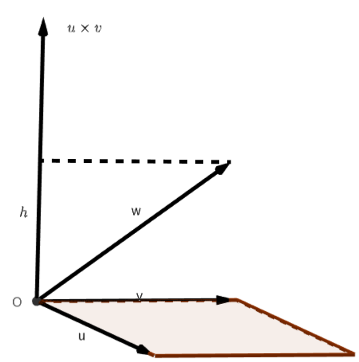

Figura 2.13: Paralelepípedo

2.2. Rectas y planos

Cuando trabajamos en geometría, a los elementos de $\mathbb{R}^3$ a veces es mejor pensarlos como puntos y otras veces como vectores flecha. Recordar 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

$$X = P + tv \quad \Leftrightarrow \quad (x, y, z) = (x_0, y_0, z_0) + t(v_1, v_2, v_3). \tag{2.5}$$

Esta es la *ecuación vectorial* de r.

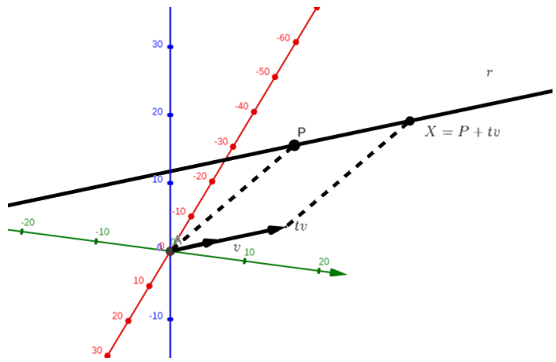

Figura 2.14: Recta

Si desarrollamos la segunda igualdad de (2.5) obtenemos

$$\begin{cases} x = x_0 + tv_1 \\ y = y_0 + tv_2 \\ z = z_0 + tv_3 \end{cases}, \quad t \in \mathbb{R}. \tag{2.6}$$

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

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3}. \tag{2.7}$$

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 2.2.1. 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:

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

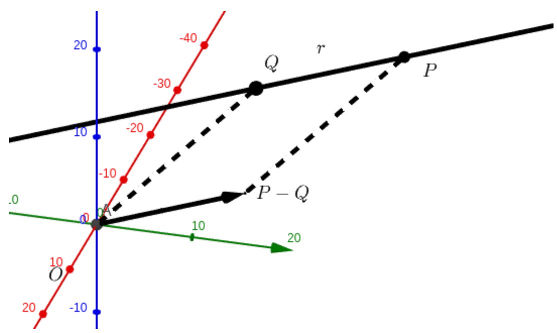

Figura 2.15: Recta por 2 puntos

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

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

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

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

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 $M = \frac{1}{2}(P + Q)$.

Distancia entre un punto y una recta. Queremos hallar la distancia[7] entre un punto P y una recta r de ecuación vectorial $X = Q + tu$. Sea H el pie de la perpendicular a r que pasa por P.

[7] La *distancia* entre dos subconjuntos de $\mathbb{R}^3$ se define como el ínfimo de las distancias entre los puntos de los conjuntos; en nuestro caso coincide con la distancia entre P y el pie de la perpendicular a r que pasa por P.

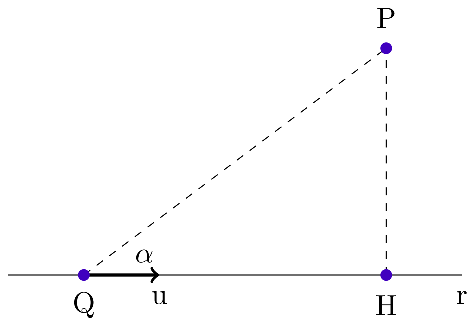

Figura 2.16: Distancia entre punto y recta

Entonces aplicando la proposición 2.1.19 (con $v = \overrightarrow{QP}$) obtenemos

$$d(P,r) = d(P,H) = \|P - H\| = \frac{\|(P-Q)\times u\|}{\|u\|}$$

Luego la distancia del punto P a la recta r es $d(P,r) = \frac{\|(P-Q)\times u\|}{\|u\|}$.

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 (2.5), (2.6) y (2.7) quedan, respectivamente

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

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 ortogonal a una dirección. Sea Π el plano que pasa por el punto $P = (x_0, y_0, z_0)$ y es ortogonal 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

$$n \cdot (X - P) = 0.$$

Escribiendo la fórmula anterior en coordenadas obtenemos

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

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

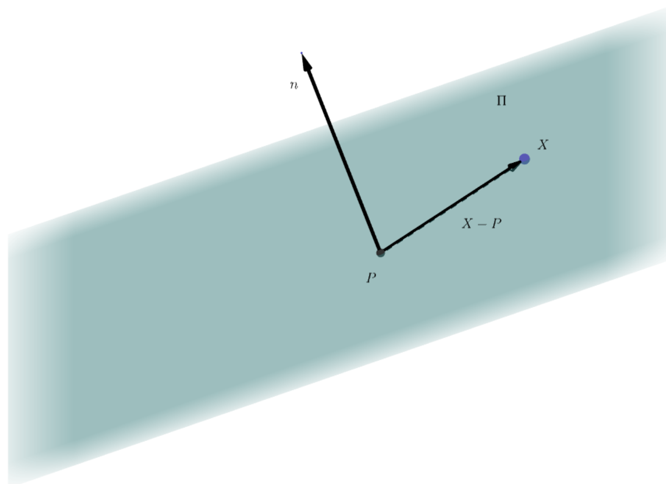

Figura 2.17: Plano por P ortogonal a n

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

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

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

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

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

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

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

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

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0. \tag{2.8}$$

Ejemplo 2.2.4. Sea Π el plano que pasa por $P=(2,1,2)$ y es paralelo a $u=(1,2,3)$ y $v=(2,-2,1)$. La ecuación vectorial de Π es

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

Para obtener la ecuación 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 (2.8) 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

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

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

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

En el primer paso de la equivalencia anterior, a la primera 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:

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

Observación 2.2.5. 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

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

Distancia entre un punto y un plano. Queremos hallar la distancia entre un punto P y un plano Π. Supongamos primero que Π es el plano pasa por un punto Q y es ortogonal a un vector n. Sea H el pie de la perpendicular a Π que pasa por P. En el plano determinado por Q, H y P tenemos la figura siguiente

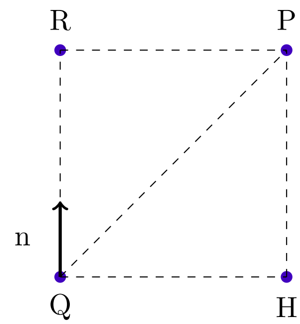

Figura 2.18: Distancia entre punto y plano

Entonces

$$d(P,\Pi) = d(P,H) = d(R,Q) = \|\Pi_n(P-Q)\| = \frac{|(P-Q)\cdot n|}{\|n\|}.$$

Supongamos ahora que Π tiene ecuación cartesiana $ax+by+cz=d$ y P tiene coordenadas (x_0,y_0,z_0). Sabemos que $n=(a,b,c)$ es ortogonal a Π. Sea $Q=(x_1,y_1,z_1)$ un punto cualquiera de Π. Entonces

$$(P-Q)\cdot n = P\cdot n - Q\cdot n = ax_0+by_0+cz_0-(ax_1+by_1+cz_1) = ax_0+by_0+cz_0-d.$$

Luego si $P=(x_0,y_0,z_0)$ y la ecuación de Π es $ax+by+cz=d$, entonces la distancia entre P y Π es

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

2.3. Formalización.

En las secciones anteriores estuvimos trabajando con los puntos y los vectores desde un punto de vista intuitivo. Lo que sigue está pensado para quienes quieran saber cómo se puede formalizar todo lo anterior. Es bastante abstracto y no es necesario leerlo si no hay interés.

Partimos de $\mathbb{R}^3$ con las operaciones de suma y producto por un escalar definidas anteriormente. Definimos una relación $\sim$ en $\mathbb{R}^3\times\mathbb{R}^3 = \{(P,Q):\ P,Q\in\mathbb{R}^3\}$, mediante

$$(P,Q)\sim(R,S) \quad\Leftrightarrow\quad Q-P=S-R.$$

Es fácil de probar que $\sim$ es una relación de equivalencia en $\mathbb{R}^3\times\mathbb{R}^3$. A la clase de equivalencia del par (P,Q) la escribimos $\overrightarrow{PQ}$ y decimos que es el *vector* que va de P a Q. Así

$$\overrightarrow{PQ}=\overrightarrow{RS} \quad\Leftrightarrow\quad (P,Q)\sim(R,S) \quad\Leftrightarrow\quad Q-P=S-R.$$

Al conjunto de todos los vectores lo escribiremos $\mathcal{V}=\left\{\overrightarrow{PQ}:\ P,Q\in\mathbb{R}^3\right\}$.

Nos interesa definir una suma de vectores y un producto por escalares. Dejamos como ejercicio el verificar que si $(P,Q)\sim(P',Q')$ y $(R,S)\sim(R',S')$, entonces $(P+R,Q+S)\sim(P'+R',Q'+S')$. Esto implica que tiene sentido definir

$$\overrightarrow{PQ}+\overrightarrow{RS} := \overrightarrow{(P+R)(Q+S)}.$$

En forma análoga se prueba que tiene sentido definir $a \cdot \overrightarrow{PQ} := \overrightarrow{(aP)(aQ)}$, para todo $a \in \mathbb{R}$. Notar que los vectores que acabamos de definir corresponden a la idea de "vectores libres", es decir sin un punto de aplicación. Eso nos permite sumar $\overrightarrow{PQ}$ y $\overrightarrow{RS}$, aunque sean $P \neq R$.

Ahora vamos a ver cómo formalizamos la idea de que "podemos identificar los vectores con los puntos". Para eso, observar que si $\overrightarrow{PQ} = \overrightarrow{RS}$, entonces $Q - P = S - R$, y por lo tanto tiene sentido definir $\psi : \mathcal{V} \to \mathbb{R}^3$ mediante $\psi\left(\overrightarrow{PQ}\right) = Q - P$. Consideremos $\varphi : \mathbb{R}^3 \to \mathcal{V}$ definida por $\varphi(P) = \overrightarrow{OP}$ (siendo O el origen de $\mathbb{R}^3$). Veamos que las funciones $\varphi : \mathbb{R}^3 \to \mathcal{V}$ y $\psi : \mathcal{V} \to \mathbb{R}^3$ son inversas entre sí.

$$\begin{aligned}
\psi(\varphi(P)) &= \psi\left(\overrightarrow{OP}\right) = P - O = P;\\
\varphi\left(\psi\left(\overrightarrow{PQ}\right)\right) &= \varphi(Q-P) = \overrightarrow{O(Q-P)} \stackrel{?}{=} \overrightarrow{PQ} \ \Leftrightarrow\ (Q-P) - O \stackrel{?}{=} Q - P \ \Leftrightarrow\ Q - P = Q - P.
\end{aligned}$$

Luego $\varphi : \mathbb{R}^3 \to \mathcal{V}$ y $\psi : \mathcal{V} \to \mathbb{R}^3$ son inversas y por lo tanto establecen una correspondencia uno a uno entre los vectores y los puntos de $\mathbb{R}^3$. Además

$$\begin{aligned}
\varphi(P+Q) &= \overrightarrow{O(P+Q)} = \overrightarrow{(O+O)(P+Q)} = \overrightarrow{OP} + \overrightarrow{OQ} = \varphi(P) + \varphi(Q),\\
\varphi(aP) &= \overrightarrow{O(aP)} = \overrightarrow{(aO)(aP)} = a\overrightarrow{OP} = a\varphi(P).
\end{aligned}$$

Luego esta correspondencia preserva la suma y el producto por escalares. En particular esto implica que $\mathcal{V}$ con esas operaciones verifica todas las propiedades que vimos anteriormente.

Cuando trabajamos con rectas y planos a menudo necesitamos sumar puntos y vectores, lo cual ahora no tiene mucho sentido porque son elementos que no están en un mismo conjunto. Para arreglarlo definimos una "suma" entre puntos y vectores, mediante $P \boldsymbol{+} \overrightarrow{QR} := P + \psi\left(\overrightarrow{QR}\right) = P + Q - R$. Notar que esta nueva operación $\boldsymbol{+}$ le asocia a un punto y a un vector un punto, es decir, es una función $\mathbb{R}^3 \times \mathcal{V} \xrightarrow{\boldsymbol{+}} \mathbb{R}^3$. Cambiando un poco de notación, si escribimos a los vectores mediante $\vec{u}, \vec{v}, \ldots$ y al vector nulo $\vec{o} = \overrightarrow{OO}$, entonces esta suma verifica

$$P \boldsymbol{+} \vec{o} = P, \quad (P \boldsymbol{+} \vec{u}) \boldsymbol{+} \vec{v} = P \boldsymbol{+} (\vec{u} + \vec{v}), \qquad \forall P \in \mathbb{R}^3,\ \vec{u}, \vec{v} \in \mathcal{V}.$$

Esas dos propiedades son las que su usan habitualmente. Notar que con esta definición vale $Q \boldsymbol{+} \vec{v} = P$ si y solo si $\vec{v} = \overrightarrow{PQ}$. Además, vale $P \boldsymbol{+} \overrightarrow{OQ} = P + Q$. Luego si identificamos los vectores con los puntos mediante $Q \leftrightarrow \overrightarrow{OQ}$, entonces la suma $\boldsymbol{+}$ coincide con la suma usual de $\mathbb{R}^3$. Con estas definiciones se puede hacer una construcción bien formal de todo lo que vimos en este capítulo.

Capítulo 3

Matrices y determinantes

Los temas de este capítulo tienen aplicaciones en diversas áreas de la matemática, en particular en las transformaciones lineales que veremos más adelante. El final del capítulo contiene aplicaciones a sistemas de ecuaciones. En lo que sigue $\Bbbk$ es un cuerpo arbitrario.

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 una matriz 2×3 (tiene 2 filas y 3 columnas)[1]. En general una matriz 2×3 tiene la forma

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

con $a, b, \ldots, f$ números arbitrarios. Para matrices de tamaño relativamente grande es preferible usar subíndices en vez de letras distintas. Por ejemplo, la matriz anterior se escribe

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

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

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

en la cual los a_{ij} son elementos de $\Bbbk$, 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. En general una matriz $m \times n$ la escribiremos de la forma $A = (a_{ij})_{m \times n}$ y cuando no dé lugar a confusión usaremos la notación abreviada $A = (a_{ij})$. Si A es una matriz $m \times n$, entonces a diremos que $m \times n$ es el *tamaño* de A. Al conjunto de las matrices $m \times n$ con coeficientes en $\Bbbk$ lo escribiremos $M_{m \times n}(\Bbbk)$ o simplemente $M_{m \times n}$.

[1] Notar que 2×3 es solo una notación y no se refiere a un producto.

Observación 3.1.1. Si queremos formalizar la definición de matriz, podemos hacer lo siguiente. Para cada entero positivo n, consideremos el conjunto $I_n = \{1, 2, \ldots, n\}$ y para cada par de enteros positivos m, n consideremos el producto cartesiano $I_m \times I_n$. Notar que dada una matriz $A = (a_{ij})_{m\times n}$, podemos definir una función $A : I_m \times I_n \to \Bbbk$ mediante $A(i, j) = a_{ij}$, para todo $(i, j) \in I_m \times I_n$. Recíprocamente, dada una función $A : I_m \times I_n \to \Bbbk$, le podemos asociar una matriz $A = (a_{ij})_{m\times n}$, definiendo $a_{ij} = A(i, j)$, para todo (i, j). Luego tenemos una correspondencia uno a uno entre matrices y funciones, y por lo tanto podemos definir las matrices $m \times n$ como las funciones de $I_m \times I_n$ en $\Bbbk$. Como esta formalización no aporta nada nuevo, en lo que sigue continuaremos pensando las matrices como tablas de números y no como funciones.

Una *matriz columna* es una matriz $m \times 1$ y una *matriz fila* es una matriz $1 \times n$:

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

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

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

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

Para simplificar, al conjunto $M_{n\times n}(\Bbbk)$ de matrices cuadradas $n \times n$ lo escribiremos $M_n(\Bbbk)$ o M_n. Si $A \in M_n$, entonces diremos que A es una matriz cuadrada de *orden* n. Si $A = (a_{ij})$ es una matriz cuadrada de orden 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 cuadrada de orden n que tiene el valor 1 en la diagonal principal y 0 en el resto

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

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 simplemente I en vez de I_n.

Hay ciertos tipos de matrices cuadradas que aparecen con frecuencia y reciben nombres particulares. Si $A = (a_{ij})$ es una matriz cuadrada, entonces decimos que A es

- *simétrica* si $a_{ij} = a_{ji}$, para todo i, j (es simétrica respecto a la diagonal principal).
- *antisimétrica* si $a_{ij} = -a_{ji}$, para todo i, j.
- *triangular superior* si $a_{ij} = 0$, para todo $i > j$ (las entradas abajo de la diagonal principal son nulas).

- *triangular inferior* si $a_{ij} = 0$, para todo $i < j$ (las entradas arriba de la diagonal principal son nulas).
- *diagonal* si $a_{ij} = 0$, para todo $i \neq j$ (las entradas fuera de la diagonal principal son nulas).

Observación 3.1.2. En las matrices simétricas y en las antisimétricas importa el cuerpo $\mathbb{k}$, dado que en un cuerpo arbitrario puede ser $a = -a$, para todo $a \in \mathbb{k}$ (por ejemplo en el cuerpo $\mathbb{k} = \{0, 1\}$ formado por solo dos elementos). En ese caso las matrices antisimétricas coinciden con las simétricas y aparecen complicaciones. Por esa razón en los ejemplos con matrices simétricas o antisimétricas siempre impondremos que el cuerpo sea $\mathbb{R}$, aunque lo mismo valdría en cuerpos donde no pase lo antes mencionado (por ejemplo, en $\mathbb{Q}$ o $\mathbb{C}$). En particular, notar que si $A = (a_{ij}) \in M_n(\mathbb{R})$ es antisimétrica, entonces vale $a_{ii} = 0$, para todo $i = 1, \dots, n$.

A continuación veremos operaciones que se pueden realizar con matrices y sus propiedades.

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 $\mathbb{k}^n = \{(x_1, \dots, x_n) : \ x_1, \dots, x_n \in \mathbb{k}\}$, n entero positivo arbitrario, 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 $a \in \mathbb{k}$, $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{k}^n$. A su vez estas operaciones las podemos generalizar a una suma y producto por un escalar en $M_{m\times n}$ definiendo

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

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

En notación abreviada, si $A = (a_{ij})$ y $B = (b_{ij})$ entonces $cA = (d_{ij})$ y $A + B = (e_{ij})$, siendo $d_{ij} := ca_{ij}$ y $e_{ij} := a_{ij} + b_{ij}$, para todo i, j. La *matriz nula* de tamaño $m \times n$ es la matriz que tiene todas sus entradas nulas. Usamos el símbolo 0 para denotar a la matriz nula[2]. Si $A = (a_{ij}) \in M_{m\times n}$, entonces su *opuesta* es la matriz $-A = (b_{ij}) \in M_{m\times n}$, definida por $b_{ij} := -a_{ij}$, para todo i, j. Por ejemplo, en matrices 2×3 es

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

La prueba de la siguiente proposición es directa y queda como ejercicio.

Proposición 3.1.3. *La suma y el producto por un escalar 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. $A + 0 = 0 + A = A$, *para todo* $A \in M_{m\times n}$ *(existencia de neutro).*

[2]El uso del símbolo 0 (el del número cero) para la matriz nula es un poco confuso al principio, pero es lo usual y resulta cómodo en su uso. Observar que hay una matriz nula para cada tamaño $m \times n$.

4. $A + (-A) = (-A) + A = 0$, *para todo* $A \in M_{m\times n}$ *(existencia de opuesto).*

5. $1u = A$, *para todo* $A \in M_{m\times n}$.

6. $c(A + B) = cA + cB$, *para todo* $c \in \Bbbk$, $A, B \in M_{m\times n}$.

7. $(b + c)A = bA + cA$, *para todo* $b, c \in \Bbbk$, $A \in M_{m\times n}$.

8. $b(cA) = (bc)A$, *para todo* $b, c \in \Bbbk$, $A \in M_{m\times n}$. □

Observaciones 3.1.4. 1. Las propiedades anteriores son las mismas que verifica $\mathbb{R}^3$ con la suma y el producto por escalares.

2. Una n-upla $(x_1, x_2, \ldots, x_n)$ es esencialmente lo mismo que una matriz fila $1 \times n$, luego podemos identificar $\Bbbk^n$ con $M_{1\times n}(\Bbbk)$, y por lo tanto en $\Bbbk^n$ valen también las propiedades anteriores.

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 la cual el coeficiente c_{ij} está definido por

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

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

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

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

Ejemplo 3.1.5. Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 2 \\ 1 & 0 \\ 5 & 3 \end{pmatrix}$, entonces su producto se calcula mediante

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

El siguiente ejemplo muestra varias particularidades del producto de matrices.

Ejemplo 3.1.6. Si consideramos las matrices $A = \begin{pmatrix} 1 & 1 \\ -2 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix}$ y $C = \begin{pmatrix} 3 & 0 \\ 5 & 4 \end{pmatrix}$, entonces

$$AB = AC = \begin{pmatrix} 8 & 4 \\ -16 & -8 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Notar que

- es $AB \neq BA$, luego el producto de matrices no es conmutativo;
- es $BA = 0$, pero $AB \neq 0$;
- es $BA = 0$, siendo $B \neq 0$ y $A \neq 0$.
- es $AB = AC$, con $A \neq 0$ y $B \neq C$. Luego no vale la propiedad cancelativa del producto.

Proposición 3.1.7. *Propiedades del producto de matrices.*

1. $A(BC) = (AB)C$, *para todo* $A \in M_{m\times n}$, $B \in M_{n\times p}$, $C \in M_{p\times q}$ *(asociativa).*
2. $A(B + C) = AB + AC$, *para todo* $A \in M_{m\times n}$, $B \in M_{n\times p}$, $C \in M_{n\times p}$ *(distributiva).*
3. $(A + B)C = AC + BC$, *para todo* $A \in M_{m\times n}$, $B \in M_{m\times n}$, $C \in M_{n\times p}$ *(distributiva).*
4. $AI_n = I_m A = A$, *para todo* $A \in M_{m\times n}$.
5. $A0 = 0A = 0$, *para todo* $A \in M_{m\times n}$, *siendo cada* 0 *una matriz nula de tamaño adecuado.*
6. $A(cB) = (cA)B = c(AB)$, *para todo* $A \in M_{m\times n}$, $B \in M_{n\times p}$ *y* $c \in \Bbbk$.

Dem. Todas estas propiedades se prueban de forma similar, así que solo demostraremos la propiedad asociativa dejando las otras pruebas como ejercicio.

Sean $A = (a_{ij}) \in M_{m\times n}$, $B = (b_{ij}) \in M_{n\times p}$ y $C = (c_{ij}) \in M_{p\times q}$. Escribamos

$$AB = (x_{ij}) \in M_{m\times p}, \quad (AB)C = (y_{ij}) \in M_{m\times q}, \quad BC = (z_{ij}) \in M_{n\times q}, \quad A(BC) = (t_{ij}) \in M_{m\times q}.$$

Luego aplicando reiteradamente la fórmula (3.1) obtenemos

$$\begin{aligned} y_{ij} &= \sum_{h=1}^{p} x_{ih} c_{hj} = \sum_{h=1}^{p} \left(\sum_{k=1}^{n} a_{ik} b_{kh} \right) c_{hj} = \sum_{h=1}^{p} \sum_{k=1}^{n} a_{ik} b_{kh} c_{hj}, \\ t_{ij} &= \sum_{k=1}^{p} a_{ik} z_{kj} = \sum_{k=1}^{p} a_{ik} \left(\sum_{h=1}^{p} b_{kh} c_{hj} \right) = \sum_{k=1}^{p} \sum_{h=1}^{p} a_{ik} b_{kh} c_{hj}. \end{aligned}$$

Como el orden en que sumamos no altera el resultado, podemos intercambiar las sumatorias $\sum_{k=1}^{p}$ y $\sum_{h=1}^{p}$ y por lo tanto es $y_{ij} = t_{ij}$, para todo i, j, es decir $(AB)C = A(BC)$. $\square$

Observaciones 3.1.8.

1. La propiedad asociativa permite escribir $ABC = (AB)C = A(BC)$ y en general escribir $A_1 \cdots A_n$ para un producto de n matrices (que se puedan multiplicar).
2. El único caso en que dadas dos matrices A y B, podemos calcular $A + B$ y AB, es cuando A y B son dos matrices cuadradas del mismo orden.
3. El producto de matrices diagonales es particularmente simple

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

   Notar que en la fórmula anterior usamos un punto $(\cdot)$ para el producto, porque aquí queda más claro que la yuxtaposición. Esa notación la seguiremos usando cuando resulte conveniente.

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

La siguiente proposición describe la relación entre la trasposición y las operaciones anteriores.

Proposición 3.1.9. *Valen las siguientes propiedes.*

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

Dem. Probaremos solo la propiedad del producto, dejando la prueba de las otras como ejercicio.

Sean $A = (a_{ij}) \in M_{m\times n}$ y $B = (b_{ij}) \in M_{n\times p}$. Entonces $A^t = (\alpha_{ij}) \in M_{n\times m}$ y $B^t = (\beta_{ij}) \in M_{p\times n}$, siendo $\alpha_{ij} = a_{ji}$ y $\beta_{ij} = b_{ji}$, para todo i, j. Luego

$$AB = (c_{ij}), \quad c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}; \quad B^t A^t = (\eta_{ij}), \quad \eta_{ij} = \sum_{k=1}^{n} \beta_{ik}\alpha_{kj} = \sum_{k=1}^{n} b_{ki} a_{jk}.$$

Entonces

$$(AB)^t = (\gamma_{ij}), \quad \gamma_{ij} = c_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki} = \sum_{k=1}^{n} b_{ki} a_{jk} = \eta_{ij}, \quad \forall i, j \quad \Rightarrow \quad (AB)^t = B^t A^t. \quad \square$$

Los conceptos que veremos a continuación se definen solo para matrices cuadradas.

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

Las potencias de matrices verifican algunas de las propiedades de las potencias de los números reales, por ejemplo, es fácil de probar que valen $A^n A^m = A^{n+m}$ y $(A^n)^m = A^{nm}$, para todo $n, m \in \mathbb{N}$. Las propiedades que involucran matrices distintas no suelen ser válidas, por ejemplo, en general es $(AB)^n \neq A^n B^n$, aunque si A y B conmutan, entonces vale $(AB)^n = A^n B^n$ para todo n.

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

Proposición 3.1.10. *Sea* $A \in M_n$*. Si existen* $B, C \in M_n$ *tales que* $AB = CA = I$*, entonces* $B = C$ *y por lo tanto* A *es invertible.*

Dem. La tesis se obtiene del cálculo siguiente: $C = CI = C(AB) = (CA)B = IB = B$. $\square$

Corolario 3.1.11. *Si* $A \in M_n$ *es invertible, entonces la matriz* $B \in M_n$ *que verifica* $AB = BA = I$ *es única.*

Dem. Si dos matrices $B, C \in M_n$ verifican $AB = BA = I$ y $AC = CA = I$, entonces es $AB = CA = I$ y se aplica la proposición anterior. $\square$

Por el corolario anterior, si $A \in M_n$ es invertible entonces existe una única $B \in M_n$ tal que $AB = BA = I$; esta matriz B 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 que verifica $AA^{-1} = A^{-1}A = I$.

Ejemplos 3.1.12. 1. La matriz identidad I verifica $I^2 = I$, luego I es invertible y $I^{-1} = I$.

2. La matriz nula $0 \in M_n$ verifica $0B = B0 = 0$, para toda $B \in M_n$. Luego 0 no puede ser invertible.

Proposición 3.1.13. *1. Si $A, B \in M_n$ son invertibles, entonces AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.*

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

Dem. Si $A, B \in M_n$ son invertibles, entonces

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

Esto prueba la primera afirmación. Para la segunda, la matriz A^{-1} verifica $AA^{-1} = A^{-1}A = I$; luego (por definición) A^{-1} es invertible y su inversa es A. $\square$

Observación 3.1.14. Aplicando reiteradamente la primera parte de la proposición anterior, deducimos que si $A_1, \ldots, A_k \in M_n$ son invertibles, entonces su producto $A_1 \cdots A_k$ es invertible y $(A_1 \cdots A_k)^{-1} = A_k^{-1} \cdots A_1^{-1}$.

Proposición 3.1.15. *Supongamos que $A \in M_n$ es invertible.*

1. *Si $B, C \in M_{n\times p}$, entonces $AB = C$ si y solo si $B = A^{-1}C$.*

2. *Si $B, C \in M_{p\times n}$, entonces $BA = C$ si y solo si $B = CA^{-1}$.*

Dem. Para la primera equivalencia, multiplicando por la izquierda obtenemos

$$\begin{aligned}&\text{si } AB = C \quad\Rightarrow\quad A^{-1}AB = A^{-1}C \quad\Rightarrow\quad IB = A^{-1}C \quad\Rightarrow B = A^{-1}C,\\ &\text{si } B = A^{-1}C \quad\Rightarrow\quad AB = AA^{-1}C \quad\Rightarrow\quad AB = IC \quad\Rightarrow\quad AB = C.\end{aligned}$$

La prueba de la segunda equivalencia es análoga y queda como ejercicio. $\square$

Como consecuencia de lo anterior, veremos que vale la propiedad cancelativa del producto, con la condición de que la matriz que multiplica a ambos lados sea invertible.

Corolario 3.1.16. *Supongamos que $A \in M_n$ es invertible.*

1. *Si $AB = 0$ o $BA = 0$, entonces $B = 0$.*

2. *Si $AB = AC$ o $BA = CA$, entonces $B = C$.*

En lo anterior se entiende que en cada caso B y C tienen tamaños adecuados para poderse multiplicar por A.

Dem. Si $AB = 0$, entonces $B = A^{-1}0 = 0$. Si $AB = C$, entonces $B = A^{-1}AC = IC = C$. Los otros casos se prueban igual. $\square$

El siguiente resultado permite determinar si una matriz de orden 2 es invertible y, en caso afirmativo, hallar su inversa.

Proposición 3.1.17. *Una matriz* $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ *es invertible si y solo si* $ad - bc \neq 0$. *En caso afirmativo, la inversa de* A *es*

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

Dem. Para la matriz nula el resultado es obvio, así que supondremos $A \neq 0$. Definimos $\Delta = ad - bc$ y $\tilde{A} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. Operando obtenemos $A\tilde{A} = \tilde{A}A = \Delta I$, con A y $\tilde{A}$ no nulas. Si $\Delta = 0$, entonces es $A\tilde{A} = \tilde{A}A = 0$ con $\tilde{A} \neq 0$; luego el corolario 3.1.16 implica que A no es invertible. Si $\Delta \neq 0$, entonces de $A\tilde{A} = \tilde{A}A = \Delta I$, deducimos

$$A\left(\Delta^{-1}\tilde{A}\right) = \left(\Delta^{-1}\tilde{A}\right)A = I.$$

Luego A es invertible (por definición) y su inversa es $\Delta^{-1}\tilde{A}$. $\square$

A continuación veremos un par de casos en los que es fácil determinar la invertibilidad.

Proposición 3.1.18.

1. *Si una matriz tiene una fila o columna formada solo por ceros, entonces no es invertible.*
2. *Una matriz diagonal es invertible si y solo si todas las entradas de la diagonal principal son no nulas. En ese caso es*

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} a_{11}^{-1} & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn}^{-1} \end{pmatrix}. \tag{3.2}$$

Dem. Sean $A, B \in M_n$. Si A tiene una fila de ceros, entonces AB también tiene una fila de ceros y si B tiene una columna de ceros, entonces AB también tiene una columna de ceros; luego en ninguno de estos casos puede ser $AB = I$.
Consideremos ahora la segunda afirmación. Sea A una matriz diagonal. Si A tiene una entrada diagonal nula, entonces A tiene una fila y una columna nulas, y por la parte anterior sabemos que A no es invertible. Si todas las entradas diagonales de A son no nulas, entonces usando la fórmula para el producto de matrices diagonales se ve que la matriz de la derecha en (3.2) es la inversa de A. $\square$

La traza. Si $A = (a_{ij}) \in M_n$, entonces definimos su *traza* mediante $\operatorname{tr}(A) := \sum_{i=1}^n a_{ii}$ (la suma de los elementos de la diagonal principal). La siguiente proposición describe propiedades de la traza.

Proposición 3.1.19. *Sean* $A, B \in M_n$ *y* $c \in \Bbbk$. *Valen las siguientes propiedades.*

1. $\operatorname{tr}\left(A^t\right) = \operatorname{tr}(A)$.
2. $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$ *y* $\operatorname{tr}(cA) = c \cdot \operatorname{tr}(A)$.
3. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.
4. *Si* $P \in M_n$ *es invertible, entonces* $\operatorname{tr}\left(PAP^{-1}\right) = \operatorname{tr}(A)$.

Dem. La prueba de las dos primeras propiedades es simple y queda como ejercicio. Consideremos la tercera. Sean $A = (a_{ij})$ y $B = (b_{ij})$. Entonces $AB = (c_{ij})$ y $BA = (d_{ij})$, siendo $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$ y $d_{ij} = \sum_{k=1}^n b_{ik}a_{kj}$, para todo i, j. Luego

$$\operatorname{tr}(AB) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n\sum_{k=1}^n a_{ik}b_{ki} \quad \text{y} \quad \operatorname{tr}(BA) = \sum_{i=1}^n d_{ii} = \sum_{i=1}^n\sum_{k=1}^n b_{ik}a_{ki} = \sum_{i=1}^n\sum_{k=1}^n a_{ki}b_{ik}.$$

Si en la última sumatoria se intercambian i con k y se invierte el orden de sumación, se obtiene $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. Para probar la cuarta propiedad, escribimos $Q = AP^{-1}$ y aplicamos la tercera propiedad

$$\operatorname{tr}\big(PAP^{-1}\big) = \operatorname{tr}(PQ) = tr(QP) = tr\big(AP^{-1}P\big) = \operatorname{tr}(A). \quad \square$$

Observación 3.1.20. La propiedad 3 nos dice que siempre vale $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, aun cuando $AB \neq BA$.

3.2. Operaciones elementales

Se llaman *operaciones elementales* en una matriz a las siguientes.

- Operación de *tipo I*: intercambiar dos filas o columnas.
- Operación de *tipo II*: multiplicar una fila o columna por una constante no nula.
- Operación de *tipo III*: sumarle a una fila o columna un múltiplo de otra fila o columna, respectivamente.

Se llaman *matrices elementales* de tipo I, II o III a las matrices que se obtienen aplicando las operaciones elementales correspondientes a la matriz identidad.

$$\text{Tipo I: } \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}; \quad \text{Tipo II: } \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & p & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \quad 0 \neq p \in \Bbbk;$$

$$\text{Tipo III: } \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & q & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & q & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \quad q \in \Bbbk.$$

Proposición 3.2.1. *Realizar una operación elemental de tipo I, II o III en las filas (columnas) de una matriz A, equivale a multiplicar a A por la izquierda (derecha) con una matriz elemental del mismo tipo.*

Dem. La prueba es directa y queda como ejercicio. $\square$

Ejemplo 3.2.2. Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \\ 4 & 5 & 6 \end{pmatrix}$. Consideremos $E_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, matriz elemental de tipo I y $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, matriz elemental de tipo III. Si realizamos los productos E_1A y AE_2, obtenemos

$$E_1A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 1 & 1 & 1 \\ 4 & 5 & 6 \end{pmatrix}, \quad AE_2 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & 0 \\ 4 & 5 & 1 \end{pmatrix}.$$

Luego multiplicar del lado izquierdo por E_1 equivale a intercambiar las filas 2 y 3, mientras que multiplicar del lado derecho por E_2 equivale a restarle a la columna 3 la columna 2.

Observación 3.2.3. Notar que la traspuesta de una matriz elemental, es una matriz elemental del mismo tipo. Además, si realizar una operación elemental en las columnas de A corresponde a multiplicar por la derecha a A con una cierta matriz elemental E, entonces realizar la misma operación en las filas de A corresponde a multiplicar por la izquierda a A con E^t.

Proposición 3.2.4. *La inversa de una matriz elemental es una matriz elemental del mismo tipo.*

Dem. Usaremos las siguientes notaciones para las matrices elementales. Llamaremos M_{ij} a la matriz correspondiente a intercambiar la fila i con la fila j de la matriz identidad, $P_i(p)$ a la correspondiente a multiplicar la fila i por el escalar $p \neq 0$ y $Q_{ij}(q)$ a la correspondiente a sumarle a la fila j la fila i multiplicada por $q \in \mathbb{k}$. Veremos que esas matrices son invertibles y que sus inversas verifican:

$$M_{ij}^{-1} = M_{ij}, \quad P_i(p)^{-1} = P_i\left(p^{-1}\right), \quad Q_{ij}(q)^{-1} = Q_{ij}(-q).$$

Para matrices de tipo I, aplicando la proposición 3.2.1 obtenemos que el producto $M_{ij}M_{ij}$ lo que hace es intercambiar en M_{ij} la fila i con la fila j y por lo tanto $M_{ij}M_{ij} = I$; luego M_{ij} es invertible y $M_{ij}^{-1} = M_{ij}$. La prueba para las matrices de los otros tipos es similar y queda como ejercicio. $\square$

Equivalencia de matrices. Sean $A, B \in M_{m\times n}$. Decimos que A es *equivalente* a B y escribimos $A \sim B$, si A se puede obtener realizando una serie de operaciones elementales en las filas y columnas de B. De la proposición 3.2.1 se deduce que $A \sim B$ si y solo si existen matrices elementales $E_1, \dots, E_k \in M_m$ y $F_1, \dots, F_h \in M_n$ tales que

$$A = E_k \cdots E_1 B F_1 \cdots F_h. \tag{3.3}$$

Proposición 3.2.5. *La equivalencia de matrices es una relación de equivalencia en $M_{m\times n}$, es decir verifica*

$$A \sim A, \qquad A \sim B \;\Rightarrow\; B \sim A, \qquad A \sim B \text{ y } B \sim C \;\Rightarrow\; A \sim C.$$

Dem. *Propiedad refleja.* Sea $A \in M_{m\times n}$. Notar que la matriz identidad es una matriz elemental de tipo II, luego de $I_m A I_n = A$ se deduce $A \sim A$.

Propiedad simétrica. Sean $A, B \in M_{m\times n}$ tales que $A \sim B$. Entonces existen matrices elementales $E_1, \dots, E_k$ en M_m y $F_1, \dots, F_h$ en M_n tales que $A = E_k \cdots E_1 B F_1 \cdots F_h$. Como las matrices elementales son invertibles, entonces teniendo en cuenta la observación 3.1.14, obtenemos

$$B = (E_k \cdots E_1)^{-1} A (F_1 \cdots F_h)^{-1} = E_1^{-1} \cdots E_k^{-1} A F_h^{-1} \cdots F_1^{-1},$$

Como la inversa de una matriz elemental es también una matriz elemental, deducimos $B \sim A$.

Propiedad transitiva. Sean $A, B, C \in M_{m\times n}$ tales que $A \sim B$ y $B \sim C$. Entonces existen matrices elementales tales que

$$A = E_k \cdots E_1 B F_1 \cdots F_h, \quad B = F_l \cdots F_1 C G_1 \cdots G_m.$$

Esto implica

$$A = E_k \cdots E_1 (F_l \cdots F_1 C G_1 \cdots G_m) F_1 \cdots F_h = (E_k \cdots E_1 F_l \cdots F_1)\, C\, (G_1 \cdots G_m F_1 \cdots F_h),$$

luego $A \sim C$. $\square$

Observación 3.2.6. Si $A \in M_{m\times n}$ se obtiene realizando operaciones elementales solo en las filas de B, entonces A es equivalente a B, dado que si es $A = E_k \cdots E_1 B$, entonces vale $A = E_k \cdots E_1 B I_n$, y la matriz identidad I_n es una matriz elemental de tipo II. Lo mismo sucede si A se obtiene realizando operaciones elementales solo en las columnas de B.

Proposición 3.2.7. *Si $A, B \in M_n$ son equivalentes, entonces A es invertible si y solo si B es invertible.*

Dem. Empezaremos probando que si $A \sim B$ y B es invertible, entonces A es invertible. Como es $A \sim B$, entonces existen matrices elementales $E_1, \dots, E_k, F_1, \dots, F_h \in M_n$ tales que

$$A = E_k \cdots E_1 B F_1 \cdots F_h. \tag{3.4}$$

Como las matrices elementales son invertibles, y el producto de matrices invertibles es también una matriz invertible, entonces (3.4) implica que A es invertible. Recíprocamente, si $A \sim B$ y A es invertible, entonces es $B \sim A$ y A es invertible. Luego por lo recién probado deducimos que B es invertible. $\square$

Una pregunta natural es, ¿cuál es la forma más simple que podemos obtener realizando operaciones elementales en las filas y columnas de una matriz arbitraria? Notar que la única matriz equivalente a la matriz nula es ella misma, así que el caso interesante es cuando la matriz no es nula.

Sean $m, n \in \mathbb{N}$, con $m, n \geq 1$, para cada $r \in \mathbb{N}$ tal que $1 \leq r \leq \text{mín}\{m, n\}$, definimos $\Phi_r^{m,n} = \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right) \in M_{m\times n}$, siendo $I_r \in M_r$ la matriz identidad. En la definición anterior entendemos que es $\Phi_m^{m,n} = \left(\begin{smallmatrix} I_m & 0 \end{smallmatrix}\right)$ si $n > m$, $\Phi_n^{m,n} = \left(\begin{smallmatrix} I_n \\ 0 \end{smallmatrix}\right)$ si $m > n$ y $\Phi_n^{n,n} = I_n$. Por ejemplo,

$$\Phi_1^{2,2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ \Phi_3^{3,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3,\ \Phi_2^{2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\ \Phi_1^{3,2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix},\ \Phi_2^{3,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

El siguiente teorema responde la pregunta anterior.

Teorema 3.2.8. *Si $A \in M_{m\times n}$ y $A \neq 0$, entonces existe r, $1 \leq r \leq min\{m, n\}$, tal que $A \sim \Phi_r^{m,n}$.*

Antes de ver la prueba del teorema veremos algunos ejemplos que ayudan a entenderla.

Ejemplos 3.2.9. Mostraremos en casos concretos cómo obtener las matrices $\Phi_r^{m,n}$ del teorema 3.2.8.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 7 & 13 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 8 \\ 3 & 1 & 13 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 2 \\ 3 & 1 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 3 & 1 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$
$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \Phi_3^{3,3} = I_3.$$

En la primera operación elemental, le sumamos a la columna 2 la columna 1 multiplicada por -2; en la segunda, le sumamos a la columna 3 la columna 1 multiplicada por -3; en la tercera, le sumamos a la fila 2 la fila 1 multiplicada por -2; en la cuarta, le sumamos a la fila 2 la fila 1 multiplicada por -3; en la quinta, le sumamos a la columna 3 la columna 2 multiplicada por -2; en la sexta, le sumamos a la fila 3 la fila 2 multiplicada por -1; en la séptima, multiplicamos la fila 3 (o la columna 3) por $1/2$.

$$\begin{pmatrix} 2 & 4 & 4 \\ 2 & 3 & 4 \\ 3 & 9 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 4 \\ 1 & 3 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \Phi_2^{3,3}.$$

Ahora fuimos un poco más rápido, haciendo más de una operación por paso. En el primer paso multiplicamos la primera fila por $1/2$ y la tercera por $1/3$; en el segundo le sumamos a la segunda columna la primera multiplicada por -2 y le sumamos a la tercera columna la primera multiplicada por -2; en el tercero le sumamos a la segunda fila la primera multiplicada por -2 y le sumamos a la tercera fila la primera multiplicada por -1; en el cuarto le sumamos a la tercera fila la segunda; en el último paso multiplicamos la segunda fila (o columna) por -1.

$$\begin{pmatrix}1&1&0&1\\2&3&2&1\\3&4&1&4\end{pmatrix} \sim \begin{pmatrix}1&0&0&0\\2&1&2&-1\\3&1&1&1\end{pmatrix} \sim \begin{pmatrix}1&0&0&0\\0&1&2&-1\\0&1&1&1\end{pmatrix} \sim \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&1&-1&2\end{pmatrix} \sim \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&2\end{pmatrix}$$
$$\sim \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\end{pmatrix} \sim \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\end{pmatrix} = \Phi_3^{3,4}.$$

En este caso hicimos al principio operaciones de tipo I en las columnas para "hacer ceros" en la primera fila, luego hicimos lo mismo en las filas para hacer ceros en la primera columna y después seguimos razonando como en los casos anteriores.

La prueba del teorema requiere del lema siguiente.

Lema 3.2.10. *Si $A \in M_{m\times n}$ es no nula ($m, n \geq 2$), entonces existe $B \in M_{(m-1)\times(n-1)}$ tal que*[3] $A \sim \left(\begin{smallmatrix}1&0\\0&B\end{smallmatrix}\right)$.

Dem. Sea $A = (a_{ij})$. Como es $A \neq 0$, entonces existen k, l tales que $a_{kl} \neq 0$. Si intercambiamos la columna 1 con la columna l y luego intercambiamos la fila 1 con la fila k, obtenemos una matriz cuyo coeficiente en el lugar $(1,1)$ es a_{kl}. Si ahora multiplicamos la primera fila o columna de esa matriz por a_{kl}^{-1}, obtenemos una matriz $C = (c_{ij})$, tal que $c_{11} = 1$ y $A \sim C$. Ahora es claro que siendo $c_{11} = 1$, haciendo operaciones elementales de tipo III se otiene una matriz como en la tesis. $\square$

Observación 3.2.11. Con las notaciones anteriores, dadas $B, C \in M_{m\times n}$, si $B \sim C$, entonces $\left(\begin{smallmatrix}1&0\\0&B\end{smallmatrix}\right) \sim \left(\begin{smallmatrix}1&0\\0&C\end{smallmatrix}\right)$. Esto se debe a que si C se obtiene haciendo ciertas operaciones elementales en las filas y columnas de B, entonces $\left(\begin{smallmatrix}1&0\\0&C\end{smallmatrix}\right)$ se obtiene haciendo las mismas operaciones en las filas y columnas correspondientes de $\left(\begin{smallmatrix}1&0\\0&B\end{smallmatrix}\right)$. Notar que estas operaciones en $\left(\begin{smallmatrix}1&0\\0&B\end{smallmatrix}\right)$ no afectan su primera fila ni su primera columna, dado que son operaciones en las otras filas y columnas.

A continuación veremos la edmostración del teorema 3.2.8. Haremos la prueba en etapas, considerando primero el caso $m = n$ y luego los casos $m \geq n$ y $m \leq n$.

I. Caso $\mathbf{m = n}$. La prueba es por inducción en n. Para $n = 1$, es $A = (a)$, con $a \neq 0$, luego multiplicando A por a^{-1} (operación elemental de tipo II) obtenemos $A \sim (1) = \Phi_1^{1,1}$.

Supongamos ahora que vale la tesis para matrices cuadradas de orden n y sea $A \neq 0$ una matriz cuadrada de orden $n+1$. Aplicando el lema 3.2.10 obtenemos que existe $B \in M_n$ tal que $A \sim \left(\begin{smallmatrix}1&0\\0&B\end{smallmatrix}\right)$. Si es $B = 0$, entonces $A \sim \Phi_1^{n+1,n+1}$. Si es $B \neq 0$, entonces aplicando la hipótesis de inducción a B obtenemos que existe r, $1 \leq r \leq n$, tal que $B \sim \Phi_r^{n,n}$. Luego $A \sim \left(\begin{smallmatrix}1&0\\0&B\end{smallmatrix}\right)$ y $B \sim \left(\begin{smallmatrix}I_r&0\\0&0\end{smallmatrix}\right)$. Teniendo en cuenta la observación anterior, obtenemos

$$A \sim \begin{pmatrix}1&0&0\\0&I_r&0\\0&0&0\end{pmatrix} = \begin{pmatrix}I_{r+1}&0\\0&0\end{pmatrix} = \Phi_{r+1}^{n+1,n+1}.$$

II. Caso $\mathbf{m \geq n}$. La prueba es por inducción en m. El caso $m = n$ es lo que probamos en el caso anterior. Supongamos ahora que vale la tesis para matrices de tamaño $m \times n$ con $m \geq n$ y sea $A \neq 0$ una matriz de tamaño $(m+1)\times n$. Consideremos $A_1 \in M_{m\times n}$ la matriz formada por las primeras m filas de A.

[3] Al escribir $\left(\begin{smallmatrix}1&0\\0&B\end{smallmatrix}\right)$ queremos decir que la fila de arriba está formada por un 1 en el primer lugar y luego el resto son ceros. Lo mismo acurre con la primera columna.

Si $A_1 = 0$, entonces A es una matriz de la forma $A = \binom{0}{B}$, siendo $B = (b_1, \dots, b_n)$ una matriz fila no nula. Razonando como en el lema 3.2.10, obtenemos $B \sim (1, 0, \dots, 0)$. Luego

$$A = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ b_1 & \cdots & b_n \end{pmatrix} \sim \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

entonces $A \sim \Phi_1^{m+1,n}$ y se cumple la tesis.

Supongamos ahora $A_1 \neq 0$. Entonces aplicando a A_1 la hipótesis de inducción obtenemos que existe r, $1 \leq r \leq \min\{m, n\}$, tal que $A_1 \sim \Phi_r^{m,n}$. Esto implica que existen $c_1, \dots, c_n \in \Bbbk$ tales que $A \sim C$, siendo

$$C = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ c_1 & \cdots & c_r & c_{r+1} & \cdots & c_n \end{pmatrix}.$$

Como en las primeras r filas de C hay unos en la diagonal, entonces realizando operaciones elementales de tipo III en las filas obtenemos que A es equivalente a la matriz D obtenida poniendo $c_1 = \cdots = c_r = 0$ en la matriz C.

Si $c_{r+1} = \cdots = c_n = 0$, entonces $A \sim \Phi_r^{m+1,n}$. Si existe algún $i \in \{r+1, \dots, n\}$ tal que $c_i \neq 0$, entonces la última fila de D es equivalente a la matriz $F = (0, \dots, 0, 1, 0, \dots, 0) \in M_{1 \times n}$ (el 1 está en el lugar $r+1$). Esto implica que A es equivalente a la matriz E obtenida cambiando la última fila de D por la fila F. Ahora intercambiando en E la última fila con la fila $r+1$ (operación elemental de tipo I), obtenemos $A \sim \Phi_{r+1}^{m+1,n}$.

III. Caso $\mathbf{m \leq n}$. Trasponiendo la matriz $A \in M_{m\times n}$, obtenemos $A^t \in M_{n \times m}$ con $n \geq m$ y $A^t \neq 0$. Luego aplicando a A^t la parte anterior, obtenemos que existen matrices elementales $E_1, \dots, E_k, F_1, \dots, F_h$ tales que $A^t = E_k \cdots E_1 \Phi_r^{n,m} F_1 \cdots F_h$, para cierto r tal que $1 \leq r \leq \min\{m, n\}$. Entonces

$$A = (E_k \cdots E_1 \Phi_r^{n,m} F_1 \cdots F_h)^t = F_h^t \cdots F_1^t (\Phi_r^{n,m})^t E_1^t \cdots E_k^t = F_h^t \cdots F_1^t \Phi_r^{m,n} E_1^t \cdots E_k^t.$$

Como la traspuesta de una matriz elemental vuele a ser una matriz elemental, concluimos $A \sim \Phi_r^{m,n}$. $\square$

Nota: La relación de equivalencia $\sim$ parte al conjunto $M_{m\times n}$ en clases de equivalencia. El teorema anterior nos dice que si $h = \min\{m, n\}$, entonces hay a lo más hay $h+1$ clases, que corresponden a la clase de la matriz nula (que tiene un solo elemento) y a las clases de las matrices $\Phi_r^{m,n}$, para $r = 1, \dots, h$. Más adelante, después de ver transformaciones lineales, probaremos que si $r \neq s$, entonces $\Phi_r^{m,n}$ y $\Phi_s^{m,n}$ no son equivalentes, por lo que hay exactamente $h+1$ clases de equivalencia.

El teorema anterior tiene aplicaciones importantes. La primera es la siguiente.

Teorema 3.2.12. *Una matriz es invertible si y solo si es producto de matrices elementales.*

Dem. Para que una matriz $A \in M_n$ sea invertible, tiene que ser no nula. Entonces sabemos que es $A \sim \Phi_r^{n,n}$, para cierto $1 \leq r \leq n$ y por lo tanto A es invertible si y solo si $\Phi_r^{n,n}$ es invertible (proposición 3.2.7). Si es $r = n$, entonces $\Phi_n^{n,n} = I_n$ es invertible, por otro lado, si es $r < n$, entonces en $\Phi_r^{n,n}$ hay una fila de ceros y por lo tanto $\Phi_r^{n,n}$ no es invertible. Luego $\Phi_r^{n,n}$ es invertible si y solo si $r = n$ y por lo tanto A es invertible si y solo si $A \sim \Phi_n^{n,n} = I_n$. Esto último equivale a que existan matrices elementales $E_1, \dots, E_k, F_1, \dots, F_h \in M_n$ tales que $A = E_k \cdots E_1 I_n F_1 \cdots F_h$, es decir, $A = E_k \cdots E_1 F_1 \cdots F_h$. $\square$

Corolario 3.2.13. *Dos matrices $A, B \in M_{m\times n}$ son equivalentes si y solo si existen matrices invertibles $U \in M_m$ y $V \in M_n$ tales que $A = UBV$.*

Dem. Si es $A \sim B$, entonces existen matrices elementales $E_1, \dots, E_k \in M_m$ y $F_1, \dots, F_h \in M_n$ tales que $A = E_k \cdots E_1 B F_1 \cdots F_h$. Como las matrices elementales son invertibles y la invertibilidad se preserva por el producto, obtenemos $A = UBV$, siendo $U := E_k \cdots E_1$ y $V := F_1 \cdots F_h$ matrices invertibles.
Recíprocamente, si es $A = UBV$, con U y V matrices invertibles, entonces aplicando el teorema 3.2.12 obtenemos que existen matrices elementales $E_1, \dots, E_k \in M_m$ y $F_1, \dots, F_h \in M_n$ tales que $U = E_k \cdots E_1$ y $V = F_1 \cdots F_h$. Luego $A = UBV = E_k \cdots E_1 B F_1 \cdots F_h$, lo cual implica (por definición) $A \sim B$. □

Nota 3.2.14. En la sección siguiente veremos una condición necesaria y suficiente para que una matriz sea invertible y luego mostraremos cómo usar las transformaciones elementales para calcular la inversa.

3.3. Determinantes

En esta sección trabajaremos solo con matrices cuadradas.

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

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

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

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

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &:= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned} \tag{3.5}$$

Observación 3.3.1. Hay una manera de recordar la fórmula (3.5) llamada el *método de Sarrus.* Consiste en construir una tabla repitiendo abajo las dos primeras filas de la matriz

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

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

$$a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21},$$

que claramente coincide con la fórmula (3.5). El método de Sarrus solo sirve para matrices 3×3 y no recomendamos su aplicación. El método general que veremos más adelante es mucho más eficiente.

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

Consideremos $A = (a_{ij})$ una matriz de orden n. Para cada $i, j = 1, \ldots, n$, sea $\tilde{A}_{ij}$ la matriz de orden $n-1$ 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 $\Delta_{ij}(A) := (-1)^{i+j} \det\left(\tilde{A}_{ij}\right)$. En lo que sigue, para simplifcar la notación, cuando no genere ambigüedad escribiremos Δ_{ij} en vez de $\Delta_{ij}(A)$.

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

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

Notar que la fórmula (3.5) se puede escribir usando los cofactores:

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

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

$$\det(A) := \sum_{j=1}^{n} a_{1j}\Delta_{1j} = a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots + a_{1n}\Delta_{1n}. \tag{3.6}$$

Esta fórmula escrita más explícitamente queda

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix}.$$

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

Ejemplo 3.3.2. Veremos cómo calcular un determinante de orden 4.

$$\begin{vmatrix} 2 & 3 & 4 & 5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix} = 2 \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} - 3 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + 4 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix} - 5 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$
$$= 2 \times 0 - 3 \times \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} + 4 \times \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} - 5 \times \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -5.$$

En general calcular determinantes requiere bastante esfuerzo, pero para las matrices diagonales es fácil.

Proposición 3.3.3. *El determinante de una matriz es diagonal es el producto de los elementos de la diagonal principal.*

Dem. Sea $A \in M_n$ una matriz diagonal. La prueba por inducción en n. El caso $n = 1$ es trivial. Supongamos ahora que vale la tesis para matrices diagonales de orden $n-1$, entonces

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{vmatrix} + 0 + \cdots + 0 = a_{11}(a_{22} \cdots a_{nn}) = a_{11}a_{22} \cdots a_{nn}. \quad \square$$

Aplicación 3.3.1. *Las matrices identidad y nula son matrices diagonales, luego aplicando la proposición anterior deducimos que el determinante de la matriz identidad vale 1 y el de la matriz nula vale 0.*

Observación 3.3.4. Existe una fórmula explícita para el determinante de una matriz $n \times n$, que en el caso $n = 3$ es la fórmula (3.5). Para quienes tengan conocimientos de permutaciones, la fórmula del determinante de una matriz $A = (a_{ij}) \in M_n$ es

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

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

Observación 3.3.5. Al operar con determinantes hay que tener cuidado, porque el determinante no se comporta bien respecto a las operaciones de suma y producto por un escalar, es decir en general es

$$\det(A + B) \neq \det(A) + \det(B); \qquad \det(cA) \neq c \det(A), \ c \in \Bbbk.$$

A continuación veremos algunas propiedades de los determinantes, que en particular sirven para simplificar los cálculos. En general las primeras pruebas van a ser por inducción, usando la fórmula (3.6).

Proposición 3.3.6. *Si la matriz B se obtiene multiplicando una fila de una matriz A por una constante c, entonces* $\det(B) = c \cdot \det(A)$, *es decir*

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

Dem. Sea $A = (a_{ij}) \in M_n$. Si $n = 1$, entonces es $\det(a_{11}) = a_{11}$, luego $\det(ca_{11}) = ca_{11} = c \det(a_{11})$. Sea ahora $n \geq 2$. Supongamos, razonando inductivamente, que se cumple la tesis para matrices de orden $n - 1$. Sea i la fila de la matriz $A = (a_{ij}) \in M_n$ que aparece multiplicada por la constante c. Si $i = 1$, es

$$\begin{vmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = ca_{11}\Delta_{11}(A) + \cdots + ca_{1n}\Delta_{1n}(A) = c\big(a_{11}\Delta_{11}(A) + \cdots + a_{1n}\Delta_{1n}(A)\big) = c \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Supongamos ahora $i \geq 2$. Aplicando la definición del determinante y la hipótesis inductiva, obtenemos

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ ca_{i1} & \cdots & ca_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ ca_{i2} & \cdots & ca_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ ca_{i1} & \cdots & ca_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix}$$

$$= a_{11} c \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{i2} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} c \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix}$$

$$= c \left(a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{i2} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix} \right) = c \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}. \qquad \square$$

Corolario 3.3.7. *Si una matriz tiene una fila formada solo por ceros, entonces su determinante es nulo.*

Dem. Tomar $c = 0$ en la proposición anterior. $\square$

El siguiente resultado se prueba por inducción de la misma forma que el anterior.

Proposición 3.3.8. *El determinante es aditivo respecto a cada fila, es decir, para cada $i = 1, \ldots, n$, vale*

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

Dem. La prueba es similar a la anterior. Si $n = 1$, entonces el resultado es obvio. Supongamos ahora $n \geq 2$. Para $i = 1$, es

$$\begin{vmatrix} b_{11} + c_{11} & \cdots & b_{1n} + c_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = (b_{11} + c_{11})\Delta_{11} + \cdots + (b_{1n} + c_{1n})\Delta_{1n}$$

$$= b_{11}\Delta_{11} + \cdots + b_{1n}\Delta_{1n} + c_{11}\Delta_{11} + \cdots + c_{1n}\Delta_{1n}$$

$$= \begin{vmatrix} b_{11} & \cdots & b_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} c_{11} & \cdots & c_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Ahora consideremos $i \geq 2$. Aplicando la definición del determinante y razonando inductivamente, obtenemos

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{i1}+c_{i1} & \cdots & b_{in}+c_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ b_{i2}+c_{i2} & \cdots & b_{in}+c_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ b_{i1}+c_{i1} & \cdots & b_{i,n-1}+c_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix}$$

$$= a_{11} \left(\begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ b_{i2} & \cdots & b_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ c_{i2} & \cdots & c_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} \right) + \cdots + (-1)^{1+n} a_{1n} \left(\begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix} + \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ c_{i1} & \cdots & c_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix} \right)$$

$$= a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ b_{i2} & \cdots & b_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix} +$$

$$a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ c_{i2} & \cdots & c_{in} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} a_{21} & \cdots & a_{2,n-1} \\ \vdots & & \vdots \\ c_{i1} & \cdots & c_{i,n-1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,n-1} \end{vmatrix}$$

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

La prueba del siguiente teorema requiere esfuerzo, pero el resultado es importante. Todo lo que demostraremos en el resto de esta sección se apoya en el mismo.

Teorema 3.3.9. *Si una matriz $A = (a_{ij}) \in M_n$ $(n \geq 2)$ tiene una fila i en la cual para algún j es $a_{ij} = 1$ y el resto de los elementos de la fila son nulos, entonces* $\det(A) = \Delta_{ij}$.

Dem. En el caso $i = 1$ tenemos

$$\det(A) = 0 \cdot \Delta_{11} + \cdots + 0 \cdot \Delta_{1,j-1} + 1 \cdot \Delta_{1j} + 0 \cdot \Delta_{1,j+1} + \cdots + 0 \cdot \Delta_{1n} = \Delta_{1j}.$$

Luego la tesis vale para $i = 1$. De ahora en adelante asumiremos $i \geq 2$.

La prueba es por inducción en n. Para $n = 2$, es $i = 2$. Si $A = \left(\begin{smallmatrix} a & b \\ 1 & 0 \end{smallmatrix}\right)$, entonces $\det(A) = -b = \Delta_{21}$ y si $A = \left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$, entonces $\det(A) = a = \Delta_{22}$. Luego la tesis vale para $n = 2$. Supongamos ahora que la matriz A

tiene orden n y que la tesis vale para matrices de orden $n-1$. La matriz $A \in M_n$ tiene la forma

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & a_{2,j} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

Sea $B \in M_{n-1}$ la submatriz de A obtenida eliminando la fila i y la columna j

$$B = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Lo que tenemos que probar es que vale $\det(A) = (-1)^{i+j} \det(B)$.

Aplicando la definición del determinante, obtenemos

$$\det(A) = a_{11}\Delta_{11}(A) + \cdots + a_{1,j-1}\Delta_{1,j-1}(A) + a_{1j}\Delta_{1j}(A) + a_{1,j+1}\Delta_{1,j+1}(A) + \cdots + a_{1n}\Delta_{1n}(A).$$

En el determinante que aparece en $\Delta_{1j}(A)$ la fila i-ésima está formada solo por ceros, luego $\Delta_{1j}(A) = 0$. Entonces aplicando la definición del determinante para B obtenemos

$$\begin{aligned} \det(A) &= a_{11}\Delta_{11}(A) + \cdots + a_{1,j-1}\Delta_{1,j-1}(A) + a_{1,j+1}\Delta_{1,j+1}(A) + \cdots + a_{1n}\Delta_{1n}(A), \\ \det(B) &= a_{11}\Delta_{11}(B) + \cdots + a_{1,j-1}\Delta_{1,j-1}(B) + a_{1,j+1}\Delta_{1,j}(B) + \cdots + a_{1n}\Delta_{1,n-1}(B). \end{aligned}$$

Notar que en la primera fila de B, en el lugar j está el coeficiente $a_{1,j+1}$, en el lugar $j+1$ está $a_{1,j+2}$ y así seguimos hasta llegar al lugar $n-1$ donde está $a_{1,n}$. Esa es la razón por la cual en $\det(B)$ hay una diferencia entre los subíndices de los primeros $j-1$ sumandos y los siguientes.

A continuación probaremos que vale

$$\Delta_{1k}(A) = \begin{cases} (-1)^{i+j}\Delta_{1k}(B), & 1 \leq k \leq j-1 \\ (-1)^{i+j}\Delta_{1,k-1}(B), & j+1 \leq k \leq n \end{cases}. \tag{3.7}$$

Luego es claro que esta fórmula junto con las dos de arriba, implican $\det(A) = (-1)^{i+j} \det(B)$.

La idea de la prueba es la siguiente. Empezamos describiendo[4] los cofactores $\Delta_{1k}(B)$ involucrados en (3.7).

$$\Delta_{1k}(B) = (-1)^{k+1}\begin{vmatrix} a_{21} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix},\ 1 \leq k \leq j-1,$$

$$\Delta_{1,k-1}(B) = (-1)^{k}\begin{vmatrix} a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,n} \end{vmatrix},\ j+1 \leq k \leq n-1.$$

Tener en cuenta que los cofactores anteriores se obtienen eliminando la i-ésima fila de A, lo cual no estamos explicitando. Ahora calculamos los cofactores de A, aplicando la hipótesis inductiva. Para $1 \leq k \leq j-1$, es

$$\begin{aligned}\Delta_{1k}(A) &= (-1)^{k+1}\begin{vmatrix} a_{21} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2,j-1} & a_{2j} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,k-1} & a_{i-1,k+1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,k-1} & a_{i+1,k+1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} \\ &= (-1)^{k+1}(-1)^{i-1+j-1}\begin{vmatrix} a_{21} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} \\ &= (-1)^{i+j}\Delta_{1k}(B).\end{aligned}$$

Respecto al cálculo anterior, en el primer determinante el coeficiente 1 se encuentra en la fila $i-1$ (porque habíamos eliminado la primera fila) y en la columna $j-1$ (porque habíamos eliminado la columna k, y es $k < j$), de ahí el factor $(-1)^{i-1+j-1}$. Notar también que el segundo determinante se obtiene eliminando la i-ésima fila y la j-ésima columna del primero, pero no estamos explicitando la fila eliminada. Ahora consideramos el caso $j+1 \leq k \leq n$.

$$\begin{aligned}\Delta_{1k}(A) &= (-1)^{1+k}\begin{vmatrix} a_{21} & \cdots & a_{2,j-1} & a_{2,j} & a_{2,j+1} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,k-1} & a_{i-1,k+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,k-1} & a_{i+1,k+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,n} \end{vmatrix} \\ &= (-1)^{1+k}(-1)^{i+j-1}\begin{vmatrix} a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,n} \end{vmatrix} \\ &= (-1)^{i+j}\Delta_{1,k-1}(B).\end{aligned}$$

Este caso difiere del anterior en que en el primer determinante, el coeficiente 1 se encuentra en la fila $i-1$ y la columna j, de ahí el factor $(-1)^{i-1+j}$. Con esta última igualdad completamos la prueba. □

[4]Si nos ponemos muy estrictos, los casos $\Delta_{1,j-1}(B)$ y $\Delta_{1,j}(B)$ hay que discutirlos por separado, porque no quedan bien descritos por las fórmulas que presentamos. En ambos casos también vale (3.7) y se pueden hacer las pruebas como ejercicio.

Corolario 3.3.10. *Si una matriz cuadrada A tiene una fila i en la cual todos los elementos son nulos salvo eventualmente el a_{ij}, entonces $\det(A) = a_{ij}\Delta_{ij}$.*

Dem. La prueba consiste en aplicar primero la proposición 3.3.6 en la fila i, y luego el teorema anterior:

$$\begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \\ 0 & \dots & 0 & a_{ij} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix} = a_{ij} \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix} = a_{ij}\Delta_{ij}. \quad \square$$

Ejemplo 3.3.11. Veamos un par de determinantes en los cuales se puede aplicar el corolario anterior.

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 7 & 0 \\ 4 & 5 & 6 \end{vmatrix} = 7 \times \Delta_{22} = 7 \times \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -84, \qquad \begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & 9 \\ 4 & 5 & 6 \end{vmatrix} = 9 \times \Delta_{23} = 9 \times \left(- \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \right) = 27.$$

Proposición 3.3.12. *El determinante de una matriz triangular (superior o inferior) es el producto de los elementos de la diagonal principal.*

Dem. Sea $A \in M_n$ una matriz triangular superior. Vamos a probar el resultado por inducción en n. Si $n = 1$, el resultado es obvio. Supongamos ahora $n > 1$ y que el resultado vale para $n-1$. Entonces aplicando el corolario 3.3.10 en la última fila y la hipótesis inductiva, obtenemos

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{nn} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ 0 & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} \end{vmatrix} = a_{nn} \cdot (a_{11} \cdots a_{n-1,n-1}) = a_{11} \cdots a_{nn}.$$

Si $A \in M_n$ es triangular inferior, entonces la prueba es esencialmente la misma, pero aplicando el corolario 3.3.10 en la primera fila:

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n-1,2} & \cdots & a_{n-1,n-1} & 0 \\ a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix} = a_{11} \cdot (a_{22} \cdots a_{nn}) = a_{11} \cdots a_{nn}.$$

$\square$

Proposición 3.3.13. *Sea $A = (a_{ij}) \in M_n$. Para cada $i = 1, \dots, n$, el determinante de A se puede calcular desarrollando a partir de la fila i:*

$$\det(A) = \sum_{j=1}^{n} a_{ij}\Delta_{ij} = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}. \tag{3.8}$$

Dem. La i-ésima fila de A se puede descomponer en

$$(a_{i1}, a_{i2}, \dots, a_{in}) = (a_{i1}, 0, \dots, 0) + (0, a_{i2}, 0, \dots, 0) + \cdots + (0, 0, \dots, 0, a_{in}).$$

Luego, aplicando reiteradamente la proposición 3.3.8 en la fila i, obtenemos

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

Finalmente la tesis se obtiene aplicando el corolario 3.3.10 en cada sumando de la fórmula de arriba. □

Proposición 3.3.14. *Si una matriz tiene dos filas iguales, entonces su determinante es nulo.*

Dem. Sea $A = (a_{ij})$ una matriz $n \times n$. En el caso $n = 2$ es

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{vmatrix} = a_{11}a_{12} - a_{11}a_{12} = 0.$$

Supongamos ahora que es $n \geq 3$. Razonando inductivamente, supongamos que vale la tesis para matrices de orden $n-1$ y que tenemos una matriz $A = (a_{ij}) \in M_n$ en la cual las filas i y j son iguales. Sea k una fila de A distinta de i y de j. Entonces desarrollando el determinante de A a partir de la fila k obtenemos

$$\det(A) = a_{k1}\Delta_{k1} + a_{k2}\Delta_{k2} + \cdots + a_{kn}\Delta_{kn}.$$

Notar que cada cofactor Δ_{kl} es nulo ($1 \leq l \leq n$), por ser $(-1)^{k+l}$ multiplicado por un determinante de una matriz de orden $n-1$ que tiene dos filas iguales. Luego $\det(A) = 0$. □

Corolario 3.3.15. *Si una matriz tiene una fila que es múltiplo de otra, entonces su determinante es nulo.*

Dem. Aplicar la proposición 3.3.6 y luego la proposición 3.3.14. □

Proposición 3.3.16. *Si la matriz B se obtiene de la matriz A intercambiando dos de sus filas, entonces $\det(B) = -\det(A)$.*

Dem. Supongamos que B se obtiene de $A = (a_{ij})$ intercambiando las filas i y j. Entonces

$$0 = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+a_{j1} & \cdots & a_{in}+a_{jn} \\ \vdots & & \vdots \\ a_{i1}+a_{j1} & \cdots & a_{in}+a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+a_{j1} & \cdots & a_{in}+a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+a_{j1} & \cdots & a_{in}+a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = 0 + \det(B) + \det(A) + 0.$$

Luego $\det(B) + \det(A) = 0$, lo cual implica la tesis. □

Proposición 3.3.17. *Si en una matriz le sumamos a una fila un múltiplo de otra fila, entonces su determinante no varía.*

Dem. Supongamos que a la fila i le sumamos un múltiplo de la fila j, entonces

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+ca_{j1} & \cdots & a_{in}+ca_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ ca_{j1} & \cdots & ca_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}. \quad \square$$

A continuación estudiaremos el comportamiento el determinante frente al producto y la trasposición. Para ello necesitamos conocer los determinantes de las matrices elementales.

Proposición 3.3.18. *Sea E una matriz elemental.*

1. *Si E es de tipo I, entonces* $\det(E) = -1$.
2. *Si E es de tipo II y k es la constante que aparece en la diagonal de E, entonces* $\det(E) = k$.
3. *Si E es de tipo III, entonces* $\det(E) = 1$.

Dem. Si E es de tipo I, entonces E se obtiene intercambiando dos filas de la matriz identidad I, luego $\det(E) = -\det(I) = -1$. Las matrices de tipo II son matrices diagonales y las de tipo III son matrices triangulares, luego en cada caso su determinante es el producto de los elementos de la diagonal principal. $\square$

Teorema 3.3.19. *Para todo $A, B \in M_n$, vale* $\det(AB) = \det(A)\det(B)$.

Dem. Haremos la prueba en etapas, discutiendo según la matriz A.

I) A es una matriz elemental. Discutimos según el tipo de la matriz elemental (recordar la proposición 3.3.18). Si A es de tipo I, entonces AB es una matriz que se obtiene de B intercambiando dos de sus filas, luego

$$\det(AB) = -\det(B) = (-1)\det(B) = \det(A)\det(B).$$

Si A es de tipo II y $0 \neq k$ es la constante que aparece en la diagonal de A, entonces AB es una matriz que se obtiene de B multiplicando una de sus filas por k, luego

$$\det(AB) = k\det(B) = \det(A)\det(B).$$

Si A es de tipo III, entonces AB es una matriz que se obtiene de B sumándole a una fila un múltiplo escalar de otra fila. Como esto no afecta el determinante, deducimos

$$\det(AB) = \det(B) = 1\det(B) = \det(A)\det(B).$$

II) A es una matriz invertible. Como A es invertible, entonces el teorema 3.2.12 nos dice que existen matrices elementales $E_1, \dots, E_l \in M_n$ tales que $A = E_1 \cdots E_l$. Luego necesitamos probar que vale

$$\det(E_1E_2\cdots E_lB) = \det(E_1E_2\cdots E_l)\det(B).$$

La prueba es por inducción en la cantidad de factores l. Notar que el caso $l=1$ es la parte anterior. A partir de ahora asumimos $l>1$ y que el resultado vale para $l-1$. Entonces

$$\begin{aligned}\det(E_1E_2\cdots E_lB) &= \det\big(E_1(E_2\cdots E_lB)\big) = \det(E_1)\det(E_2\cdots E_lB) = \det(E_1)\det(E_2\cdots E_l)\det(B) \\ &= \det\big(E_1(E_2\cdots E_l)\big)\det(B) = \det(E_1E_2\cdots E_l)\det(B).\end{aligned}$$

En la segunda igualdad aplicamos la parte I, en la tercera la hipótesis inductiva y en la cuarta aplicamos de nuevo la parte I.

III) A es una matriz no invertible. Empezaremos probando que vale $\det(AB)=0$, para toda $B\in M_n$. Si $A=0$, entonces $AB=0$ y por lo tanto $\det(AB)=0$. Si $A\neq 0$, entonces existe $1\le r\le n$ y matrices invertibles $U,V\in M_n$ tales que $A=U\Phi_r^{n,n}V$ (teorema 3.2.8). Como A no es invertible, entonces es $r<n$ y por lo tanto la matriz $\Phi_r^{n,n}$ tiene sus últimas $n-r$ filas nulas. Esto implica que la matriz $\Phi_r^{n,n}VB$ también tiene sus últimas $n-r$ filas nulas y por lo tanto $\det(\Phi_r^{n,n}VB)=0$. Como U es invertible, entonces aplicando la parte II obtenemos

$$\det(AB)=\det(U\Phi_r^{n,n}VB)=\det(U)\det(\Phi_r^{n,n}VB)=\det(U)\cdot 0=0.$$

Luego hemos probado que si A no es invertible, entonces $\det(AB)=0$, para toda $B\in M_n$. Tomando $B=I$ en $\det(AB)=0$, obtenemos $\det(A)=0$, lo cual implica $\det(A)\det(B)=0$. Por lo tanto cuando A no es invertible, vale $\det(AB)=\det(A)\det(B)=0$, para toda $B\in M_n$. □

Teorema 3.3.20. *Una matriz $A\in M_n$ es invertible si y solo si $\det(A)\neq 0$. En caso afirmativo,*

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

Dem. Si $A\in M_n$ no es invertible, entonces en la prueba de la parte III del teorema anterior vimos que es $\det(A)=0$. Si $A\in M_n$ es invertible, entonces existe su inversa A^{-1} que verifica $AA^{-1}=I$. Luego

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

Esto implica $\det(A)\neq 0$ y $\det\big(A^{-1}\big)=\det(A)^{-1}$. □

El siguiente resultado muestra que si $A,B\in M_n$ verifican $AB=I$, entonces también verifican $BA=I$.

Corolario 3.3.21. *Si $A,B\in M_n$ verifican $AB=I$, entonces A y B son inversas entre sí.*

Dem.

$$AB=I \quad\Rightarrow\quad \det(AB)=\det(I) \quad\Rightarrow\quad \det(A)\det(B)=1 \quad\Rightarrow\quad \det(A)\neq 0 \text{ y } \det(B)\neq 0.$$

Entonces A y B son invertibles. Luego multiplicando por A^{-1} en $AB=I$ obtenemos $B=A^{-1}$. □

Observación 3.3.22. Es un ejercicio el probar que si $A\in M_n$ es invertible, entonces su traspuesta A^t también lo es y vale $\big(A^t\big)^{-1}=\big(A^{-1}\big)^t$. Aplicando esto deducimos que si A^t es invertible, entonces $A=\big(A^t\big)^t$ también lo es. Es decir, A es invertible si y solo si A^t es invertible.

Teorema 3.3.23. *Para toda matriz cuadrada A, vale $\det\big(A^t\big)=\det(A)$.*

Dem. Si A no es invertible, entonces A^t tampoco lo es y por lo tanto $\det\big(A^t\big)=\det(A)=0$.
Supongamos ahora que A es invertible. Entonces existen matrices elementales $E_1,\dots,E_k\in M_n$ tales que $A=E_1\cdots E_k$ y por lo tanto $A^t=E_k^t\cdots E_1^t$. Entonces aplicando reiteradamente el teorema 3.3.19, obtenemos

$$\det(A)=\det(E_1,\dots,E_k)=\det(E_1)\cdots\det(E_k);\quad \det\big(A^t\big)=\det\big(E_k^t\cdots E_1^t\big)=\det\big(E_k^t\big)\cdots\det\big(E_1^t\big).$$

Para completar la demostración, alcanza con probar que vale $\det\big(E^t\big)=\det(E)$ para toda matriz elemental E, lo cual es inmediato (recordar la observación 3.2.3 y la proposición 3.3.18). □

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

Teorema 3.3.24. *Propiedades del determinante respecto a las columnas.*

1. *Si una matriz B se obtiene intercambiando un par de columnas de A, entonces* $\det(B) = -\det(A)$.
2. *Si una matriz tiene una columna formada solo por ceros o tiene una columna que es múltiplo de otra columna, entonces su determinante es nulo.*
3. *Si B se obtiene multiplicando una columna de A por una constante c, entonces* $\det(B) = c\det(A)$
$$\begin{vmatrix} a_{11} & \cdots & ca_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & ca_{nj} & \cdots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}.$$
4. *El determinante es aditivo respecto a cada columna*
$$\begin{vmatrix} a_{11} & \cdots & b_{1j}+c_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_{nj}+c_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & b_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_{nj} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & c_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & c_{nj} & \cdots & a_{nn} \end{vmatrix}.$$
5. *Si en una matriz le sumamos a una columna un múltiplo escalar de otra columna, entonces su determinante no varía.*
6. *Si A es tal que en una columna j todos los elementos son nulos salvo el a_{ij}, entonces* $\det(A) = a_{ij}\Delta_{ij}$.
7. *Sea $A = (a_{ij}) \in M_n$. Para cada $j = 1, \ldots, n$, el determinante de A se puede calcular desarrollando a partir de la columna j:*
$$\det(A) = \sum_{i=1}^{n} a_{ij}\Delta_{ij} = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj}. \quad \square$$

Observación 3.3.25. La técnica para calcular determinantes consiste en aplicar reiteradamente la propiedad 5 del teorema 3.3.24 (o la correspondiente en filas), hasta llegar al determinante de una matriz que tiene solo un elemento no nulo en una fila o columna. Luego se reduce su orden mediante la propiedad 6 del teorema 3.3.24 (o la correspondiente en filas), usando también las otras propiedades para ir simplificando los cálculos. Reiterando este procedimiento se reduce el cálculo de un determinante de orden n, al de uno de orden 2.

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

Ejemplo 3.3.26.

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

El primer paso fue $F_3 - 2F_1$.

Ejemplo 3.3.27.

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

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

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

Ejemplo 3.3.28.

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

La primera igualdad es porque en la fila 3 todos los elementos son mútiplos de 2 y en la 4 todos son múltiplos de 3. El determinante final es nulo por tener dos filas iguales.

Ejemplo 3.3.29.

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

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

Ejemplo 3.3.30.

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

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

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).

Método para hallar la matriz inversa. Sabemos que una matriz es invertible si y solo si su determinante es distinto de cero. A continuación veremos un método para hallar su inversa.

Supongamos que $A \in M_n$ es invertible. Sabemos que existen matrices elementales $E_1, \dots, E_s$ tales que $A^{-1} = E_1 \cdots E_s$. Por otro lado, si $E_1, \dots, E_r$ son matrices elementales, entonces $A^{-1}A = I$ implica $A^{-1}(AE_1 \cdots E_r) = E_1 \cdots E_r$. Luego si elegimos $E_1, \dots, E_r$ de forma tal que $AE_1 \cdots E_r = I$, entonces $A^{-1} = E_1 \cdots E_r$. Observar que multiplicar a A por la derecha por matrices elementales equivale a hacer las operaciones elementales correspondientes en las columnas de A, por lo tanto si realizando ciertas operaciones elementales en las columnas de una matriz A llegamos a la matriz identidad ($AE_1 \cdots E_r = I$),

entonces realizando las mismas operaciones elementales en las columnas de la matriz identidad I obtenemos $A^{-1} = E_1 \cdots E_r = IE_1 \cdots E_r$. De la misma forma, operar con las filas equivale a multiplicar por la izquierda, y por lo tanto partiendo de la otra igualdad $AA^{-1} = I$, se obtiene un algoritmo análogo operando con filas en vez de columnas.
Notar que cuando operamos con columnas partimos de $A^{-1}A = I$, pero cuando operamos con las filas partimos de $AA^{-1} = I$. Por esa razón para hallar la inversa no podemos realizar operaciones elementales con filas y columnas al mismo tiempo.

A continuación veremos algunos ejemplos donde estudiamos la invertibilidad de matrices y calculamos las inversas. 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. Si escribimos $F_2 \leftrightarrow F_3$, quiere decir que intercambiamos la fila 2 con la fila 3, y análogamente para columnas.

Ejemplos 3.3.31. Sea $A = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$. Primero estudiamos su invertibilidad.

$$\begin{vmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix} = 1.$$

En el primer paso, a la primera columna le restamos la tercera. Como es $\det(A) = 1 \neq 0$, entonces A es invertible. Para calcular su inversa tenemos que elegir primero si operar con columnas o con filas. En este caso elegimos las columnas.

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

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

Sea $A = \begin{pmatrix} 2 & -2 & 1 \\ 2 & -1 & 0 \\ 1 & -2 & 1 \end{pmatrix}$. Dejamos como ejercicio verificar $\det(A) = -1$, luego A es invertible. Hallaremos su inversa operando con las filas. Esta vez iremos más rápido, haciendo a veces más de una operación por paso.

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

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

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

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

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

Observación 3.3.32. Notar que en los ejemplos anteriores lo que hicimos primero fue transformar la matriz A en una matriz triangular y luego llevarla a una forma diagonal. Es importante verificar que la matriz hallada es realmente la matriz inversa (haciendo el producto), dado que es fácil cometer errores de cálculo.

Observación 3.3.33. Conviene remarcar que hay dos situaciones distintas.

1. Si queremos investigar si una matriz es invertible, entonces necesitamos ver si su determinante es cero o no, y para calcularlo podemos usar operaciones elementales en sus filas <u>y/o</u> columnas.

2. Si ya sabemos que una matriz es invertible y queremos hallar su inversa, entonces aplicamos el algoritmo de arriba, pero tenemos que elegir trabajar con filas <u>o</u> columnas, y no podemos mezclar.

Fórmula general de la matriz inversa. Para probar la fórmula necesitamos el siguiente resultado, el cual tiene interés en sí mismo.

Lema 3.3.34. *Si $A \in M_n$ y definimos*

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

entonces $A\hat{A} = \hat{A}A = \det(A)I$.

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

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

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

Teorema 3.3.35. *Si una matriz $A \in M_n$ es invertible, entonces su matriz inversa se obtiene mediante*

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

Dem. Usando el lema anterior obtenemos $A\hat{A} = \hat{A}A = \det(A)I$. Como A es invertible, es $\det(A) \neq 0$, luego vale $A\left(\det(A)^{-1}\hat{A}\right) = \left(\det(A)^{-1}\hat{A}\right)A = I$, lo cual implica $A^{-1} = \det(A)^{-1}\hat{A}$. □

Observación 3.3.36. Lo interesante de este teorema es que nos da una descripción explícita de la inversa, pero como método para invertir una matriz, el algoritmo que vimos antes es mucho más rápido (salvo para matrices 2×2).

3.4. Aplicación a sistemas de ecuaciones

Consideremos un sistema de ecuaciones

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right. . \tag{3.9}$$

Si definimos

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

entonces el sistema (3.9) es equivalente a la ecuación matricial

$$AX = B. \tag{3.10}$$

La ventaja de (3.10) es que tiene una sola incógnita (matricial). Para el estudio de (3.9) conviene considerar el sistema homogéneo asociado:

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & 0 \end{array}\right. \tag{3.11}$$

y la correspondiente ecuación matricial

$$AX = 0. \tag{3.12}$$

Recordar que el sistema homogéneo siempre admite la solución trivial $x_1 = \cdots = x_n = 0$, por lo que siempre es compatible. Lo que interesa en este caso es saber si admite alguna solución no trivial, es decir si es indeterminado. Sea $\mathcal{S}_B$ el conjunto de soluciones de la ecuación $AX = B$ y $\mathcal{S}_0$ el conjunto de soluciones de la ecuación $AX = 0$, es decir

$$\mathcal{S}_B = \{X \in \Bbbk^n : \; AX = B\}, \qquad \mathcal{S}_0 = \{X \in \Bbbk^n : \; AX = 0\},$$

Notar que en lo anterior estamos identificando $M_{n\times 1}(\Bbbk)$ con $\Bbbk^n$.

Proposición 3.4.1. *Propiedades de $\mathcal{S}_0$ y $\mathcal{S}_B$.*

1. *Si $X_1 \in \mathcal{S}_0$ y $c \in \Bbbk$, entonces $cX_1 \in \mathcal{S}_0$.*
2. *Si $X_1, X_2 \in \mathcal{S}_B$, entonces $X_1 - X_2 \in \mathcal{S}_0$.*
3. *Si $X_1 \in \mathcal{S}_B$ y $X_0 \in \mathcal{S}_0$, entonces $X_1 + X_0 \in \mathcal{S}_B$.*
4. *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. $X_1 \in \mathcal{S}_0 \;\Rightarrow\; A(cX_1) = cAX_1 = c0 = 0 \;\Rightarrow\; cX_1 \in \mathcal{S}_0$.
2. $X_1, X_2 \in \mathcal{S}_B \;\Rightarrow\; A(X_1 - X_2) = AX_1 - AX_2 = B - B = 0 \;\Rightarrow\; X_1 - X_2 \in \mathcal{S}_0$.
3. $X_1 \in \mathcal{S}_B$ y $X_0 \in \mathcal{S}_0 \;\Rightarrow\; A(X_1 + X_0) = AX_1 + AX_0 = B + 0 = B \;\Rightarrow\; X_1 + X_0 \in \mathcal{S}_B$.

4. La parte 3 implica $\{X_1 + X : \ X \in \mathcal{S}_0\} \subset \mathcal{S}_B$. Si $X_2 \in \mathcal{S}_B$, entonces $X_2 = X_1 + (X_2 - X_1)$, con $X_2 - X_1 \in \mathcal{S}_0$, luego $\mathcal{S}_B \subset \{X_1 + X : \ X \in \mathcal{S}_0\}$. Esto termina la prueba. $\square$

Observaciones 3.4.2. Respecto a la proposición anterior.

1. El sistema puede ser compatible o no, pero en caso de ser compatible, entonces la parte 4 nos dice que es determinado si y solo si la única solución que admite el sistema homogéneo asociado es la trivial. Esto depende solo de la matriz A y no de la columna de resultados B.

   Los comentarios siguientes se refieren a cuando el cuerpo $\Bbbk$ es infinito (como es el caso de $\Bbbk = \mathbb{R}$).

2. La parte 1 nos dice que si el sistema homogéneo (3.11) admite alguna solución no trivial, entonces tiene infinitas soluciones.

3. La parte 2 muestra que si el sistema no homogéneo admite más de una solución (es indeterminado), entonces el sistema homogéneo asociado admite alguna solución no trivial y por lo tanto tiene infinitas soluciones. Luego la parte 4 implica que el sistema no homogéneo también tiene infinitas soluciones. Es decir, un sistema compatible indeterminado (homogéneo o no) siempre tiene infinitas soluciones.

Recordar que llamamos *sistemas cuadrados* a los que tienen tantas ecuaciones como variables. En ese caso llamamos *determinante del sistema* al determinante de la matriz formada por sus coeficientes (la matriz A de (3.10).).

Proposición 3.4.3. *Si en un sistema homogéneo hay más incógnitas que ecuaciones o hay la misma cantidad pero el determinante del sistema es cero, entonces el sistema admite alguna solución no trivial.*

Dem. Mantenemos las notaciones anteriores. Si $A = 0$, el resultado es obvio. Si $A \neq 0$, entonces sabemos que existen matrices invertibles $U \in M_m$ y $V \in M_n$ tales que $A = U\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)V$, siendo $1 \leq r \leq \min\{m, n\}$. Si hay más incógnitas que ecuaciones, es decir, si $n > m$, entonces es $r \leq \min\{m, n\} = m < n$. Por otro lado, si es $m = n$ y $\det(A) = 0$, entonces $A \in M_n$ no es invertible y por lo tanto es $r < n$ (si fuese $r = n$, valdría $A = UV$ y por lo tanto A sería invertible). Luego en cualquiera de los dos casos es $A = U\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)V$, con $1 \leq r < n$. Sea

$$Y_1 = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \Bbbk^n.$$

en el cual hay r ceros y $n - r$ unos. Definimos $Y_0 = V^{-1}Y_1$, de forma tal que $Y_1 = VY_0$. Entonces

$$AY_0 = U\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)VY_0 = U\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)Y_1 = B \cdot 0 = 0.$$

Además notar que $Y_1 \neq 0$ implica $Y_0 \neq 0$. Luego Y_0 es una solución no trivial del sistema. $\square$

Teniendo en cuenta la parte 1 de las observaciones 3.4.2, deducimos lo siguiente.

Corolario 3.4.4. *Si en un sistema hay más incógnitas que ecuaciones o hay la misma cantidad pero el determinante del sistema es cero, entonces el sistema es incompatible o compatible indeterminado.* $\square$

El siguiente ejemplo muestra que se pueden dar cualquiera de las dos opciones anteriores.

Ejemplo 3.4.5. Consideremos los siguientes sistemas

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

En los dos primeros casos el determinante del sistema es cero, pero el primero es compatible indeterminado y el segundo es incompatible. En los casos tercero y cuarto hay más incógnitas que ecuaciones, pero el tercero es indeterminado y el cuarto es incompatible.

Proposición 3.4.6. *Si en un sistema hay la misma cantidad de ecuaciones que de incógnitas, entonces el sistema es compatible determinado si y solo si su determinante es distinto de cero.*

Dem. Mantenemos las notaciones anteriores. En el corolario anterior vimos que si $\det A = 0$, entonces el sistema es incompatible o compatible indeterminado. Por otro lado, si $\det(A) \neq 0$, entonces A es invertible y por lo tanto $AX = B$ implica $X = A^{-1}B$; luego en este caso el sistema es compatible determinado. $\square$

Corolario 3.4.7. *Si en un sistema homogéneo hay la misma cantidad de ecuaciones que de incógnitas, entonces el sistema admite alguna solución no trivial si y solo si su determinante es cero.* $\square$

Los siguientes ejemplos muestran que si en un sistema hay menos incógnitas que ecuaciones, entonces no se puede anticipar nada.

Ejemplos 3.4.8. Consideremos los siguientes sistemas

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

Respecto a los dos primeros sistemas (homogénos), el primero es compatible determinado y el segundo es compatible indeterminado. El tercer sistema es compatible determinado, el cuarto es incompatible y el quinto es compatible indeterminado.

Capítulo 4

Espacios vectoriales

Los espacios vectoriales son conjuntos en los cuales están definidas unas operaciones de suma y producto por un escalar, con propiedades que generalizan las que verifican los vectores en el plano y el espacio. Es una estructura muy general, que tiene aplicaciones en la mayoría de las áreas de la matemática. En este capítulo trabajaremos sobre un cuerpo arbitrario $\Bbbk$, aunque los ejemplos van a ser en $\mathbb{R}$.

4.1. Definiciones y propiedades básicas

Un $\Bbbk$-*espacio vectorial* es un conjunto V en el cual están definidas dos operaciones

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

que llamamos *suma* y *producto por un escalar*, que verifican los siguientes axiomas:

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 $v + o = v$, para todo $v \in V$ (existencia de neutro);
4. Para cada $v \in V$ existe un elemento $u \in V$, tal que $v + u = o$, para todo $v \in V$ (existencia de opuesto);
5. $1 \cdot u = u$, para todo $u \in V$;
6. $a \cdot (u + v) = a \cdot u + a \cdot v$, para todo $a \in \Bbbk$, $u, v \in V$;
7. $(a + b) \cdot v = a \cdot v + b \cdot v$, para todo $a, b \in \Bbbk$, $v \in V$;
8. $a \cdot (b \cdot v) = (ab) \cdot v$, para todo $a, b \in \Bbbk$, $v \in V$.

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 $\Bbbk$ les llamamos *escalares* y los escribimos con las letras $a, b, c, \ldots$. Cuando no sea necesario explicitar el cuerpo $\Bbbk$, diremos simplemente que V es un espacio vectorial.

Observaciones 4.1.1.

1. La propiedad conmutativa implica que en el axioma 3 vale $v + o = o + v = v$ y en el axioma 4 vale $v + u = u + v = o$.
2. La propiedad asociativa permite sacar paréntesis y escribir $u + v + w = u + (v + w) = (u + v) + w$, para todo $u, v, w \in V$. Lo mismo sucede cuando hay más sumandos.

Ejemplos 4.1.2. Ejemplos de espacios vectoriales.

1. El plano $\mathbb{R}^2$ y el espacio $\mathbb{R}^3$.

2. El conjunto $\mathbb{k}^n$ $(n \geq 1)$, con las operaciones usuales. En particular el cuerpo $\mathbb{k}$ es un $\mathbb{k}$-espacio vectorial, con la suma y producto usuales.

3. El conjunto de las matrices $M_{m\times n}(\mathbb{k})$ $(m, n \geq 1)$, con las operaciones usuales.

Proposición 4.1.3. *Si V es un espacio vectorial, entonces valen las siguientes propiedades.*

1. *Existe un único elemento $o \in V$ que verifica el axioma 3; se llama el* vector nulo *de V.*

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

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

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

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

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

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

Dem.

1. Si $o' \in V$ verifica $v + o' = v$, para todo $v \in V$, entonces tomando $v = o$ obtenemos $o + o' = o$. Pero aplicando la propiedad 3 a $v = o'$ obtenemos $o' + o = o'$. Luego $o' = o' + o = o + o' = o$.

2. Sea $v \in V$. Supongamos que existen $u, w \in V$ tales que $v + u = o$ y $v + w = o$. Entonces
$$w = o + w = (v + u) + w = (u + v) + w = u + (v + w) = u + o = u.$$
Luego hemos probado la unicidad del opuesto.

3. La propiedad cancelativa se obtiene del cálculo siguiente
$$u+v = u+w \Rightarrow -u+(u+v) = -u+(u+w) \Rightarrow (-u+u)+v = (-u+u)+w \Rightarrow o+v = o+w \Rightarrow v = w.$$

4. Observar que valen
$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v \quad \text{y} \quad 0 \cdot v = 0 \cdot v + o \quad \Rightarrow \quad 0 \cdot v + 0 \cdot v = 0 \cdot v + o,$$
luego la propiedad cancelativa implica $0 \cdot v = o$.

5. Esta propiedad se prueba en forma análoga a la anterior:
$$a \cdot o = a \cdot (o + o) = a \cdot o + a \cdot o \quad \Rightarrow \quad o + a \cdot o = a \cdot o + a \cdot o \quad \Rightarrow \quad o = a \cdot o.$$

6. La prueba consiste en multiplicar en ambos lados de la igualdad por a^{-1}
$$v = a \cdot u \quad \Rightarrow \quad a^{-1} \cdot v = a^{-1} \cdot (a \cdot u) = \left(a^{-1} a\right) \cdot u = 1 \cdot u = u \quad \Rightarrow \quad a^{-1} \cdot v = u.$$

7. Esta propiedad se deduce de la anterior: si $a \neq 0$, entonces $a \cdot u = o$ implica $u = a^{-1} \cdot o = o$. $\square$

Observación 4.1.4. Vale $-v = (-1) \cdot v$, para todo $v \in V$. Esto se debe al cálculo siguiente

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = \big(1 + (-1)\big) \cdot v = 0 \cdot v = o.$$

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

Proposición 4.1.5. *La resta verifica las propiedades siguientes.*

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

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

Dem. Ejercicio. □

Si V es un espacio vectorial y D es un conjunto arbitrario (no vacío), entonces podemos considerar el conjunto $\mathcal{F}(D, V)$ de las funciones de D en V. Si $f, g : D \to V$ son funciones y $a \in \Bbbk$, entonces definimos la suma y el producto por un escalar $f + g : D \to V$ y $a \cdot f : D \to V$ mediante

$$(f + g)(x) := f(x) + g(x), \quad (a \cdot f)(x) := a \cdot f(x), \quad \forall x \in D. \tag{4.1}$$

Proposición 4.1.6. *El conjunto $\mathcal{F}(D, V)$ con las operaciones anteriores es un espacio vectorial.*

Dem. Tenemos que probar que en $\mathcal{F}(D, V)$ se verifican los 8 axiomas de espacio vectorial. Probaremos la existencia de neutro y de opuesto. Consideremos la *función nula* $\theta : D \to V$ definida por $\theta(x) = o$, para todo $x \in D$. Si $f : D \to V$ es una función arbitraria, entonces

$$(f + \theta)(x) = f(x) + \theta(x) = f(x) + o = f(x),\ \forall x \in D \quad \Rightarrow \quad f + \theta = f.$$

Esto prueba que la función nula es el neutro de la suma. Dada $f : D \to V$, definimos su *opuesta* $-f : D \to V$ mediante $(-f)(x) := -f(x)$, para todo $x \in D$. Luego

$$\big(f + (-f)\big)(x) = f(x) + (-f)(x) = f(x) + \big(-f(x)\big) = o,\ \forall x \in D \quad \Rightarrow \quad f + (-f) = \theta.$$

La prueba de que en $\mathcal{F}(D, V)$ se verifican los restantes axiomas es rutinaria y queda como ejercicio. □

Ejemplo 4.1.7. Como $\mathbb{R}$ es un espacio vectorial, si D es un conjunto arbitrario, entonces $\mathcal{F}(D, \mathbb{R})$ es un espacio vectorial. En particular, conjuntos del tipo $\mathcal{F}(\mathbb{R}, \mathbb{R})$, $\mathcal{F}([0, 1], \mathbb{R})$, etc., son espacios vectoriales.

Sucesiones. Una *sucesión* en $\Bbbk$ es una función de $\mathbb{N}$ en $\Bbbk$. Si $a : \mathbb{N} \to \Bbbk$ es una sucesión, entonces lo usual es escribir a_n en vez de $a(n)$ y referirnos a la sucesión mediante $(a_0, a_1, \dots)$, o $(a_n)_{n \in \mathbb{N}}$ o simplemente (a_n). Como $\Bbbk$ es un $\Bbbk$-espacio vectorial, entonces el conjunto de las sucesiones en $\Bbbk$ con la suma y producto por un escalar definidos en (4.1), forman un $\Bbbk$-espacio vectorial. Notar que si $a = (a_n)$ y $b = (b_n)$ son dos sucesiones y $c \in \Bbbk$, entonces la suma y el producto por un escalar están definidos por $a + b := (a_n + b_n)$ y $c \cdot a := (ca_n)$.

Nota: Por comodidad de notación, de ahora en adelante al producto por un escalar lo vamos a escribir av en vez de $a \cdot v$ y al vector nulo lo escribiremos 0 (cero) en vez de o. Esto último puede ser un poco confuso al principio, pero es la notación habitual y resulta cómoda de usar. También a menudo abreviaremos “espacio vectorial” escribiendo solo “espacio”.

En lo que sigue de este capítulo, V denotará un espacio vectorial arbitrario.

4.2. Subespacios

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

$$0 \in W; \quad w_1, w_2 \in W \;\Rightarrow\; w_1 + w_2 \in W; \quad a \in \Bbbk,\, w \in W \;\Rightarrow\; aw \in W.$$

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

Proposición 4.2.1. *Sea W un subconjunto de un espacio V. Entonces W es un subespacio de V si y solo si se verifica*

$$0 \in W; \quad w_1, w_2 \in W,\, a \in \Bbbk \;\Rightarrow\; w_1 + aw_2 \in W.$$

Dem. Ejercicio. □

Ejemplos 4.2.2.

1. En todo espacio V, los subconjuntos $\{0\}$ y V son subespacios. Al subespacio $\{0\}$ 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. En el espacio de las matrices cuadradas, el conjunto formado por las matrices triangulares superiores es un subespacio. Lo mismo sucede con las matrices triangulares inferiores.
5. 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 $D(I) \subset C(I) \subset \mathcal{F}(I, \mathbb{R})$ son subespacios.
6. Los siguientes subconjuntos de $\mathbb{R}^2$ **no son** subespacios

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

   Notar que cada uno de ellos verifica solo dos de las tres condiciones de subespacio.

Observación 4.2.3. Los únicos subespacios del cuerpo $\Bbbk$ pensado como $\Bbbk$-espacio son $\Bbbk$ y $\{0\}$. Esto se debe a que si $W \subset \Bbbk$ es un subespacio tal que $W \neq \{0\}$, entonces existe $0 \neq a \in W$, luego $1 = a^{-1}a \in W$ y por lo tanto $b = b1 \in W$, para todo $b \in \Bbbk$. Esto implica $W = \Bbbk$.

El interés en los subespacios viene dado por la siguiente proposición. La prueba es directa y queda como ejercicio.

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

Polinomios. En la enseñanza secundaria se suele definir un *polinomio* como una función $p : \mathbb{R} \to \mathbb{R}$ de la forma

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n, \quad \forall x \in \mathbb{R},$$

para ciertos $a_0, a_1 \ldots, a_n \in \mathbb{R}$ y $n \in \mathbb{N}$. Si escribimos $\mathbb{R}[x]$ al conjunto de todos los polinomios, entonces $\mathbb{R}[x]$ está contenido en el espacio $\mathcal{F}(\mathbb{R}, \mathbb{R})$ y es claro que $\mathbb{R}[x]$ es un subespacio de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

La definición anterior funciona bien para los números reales y se puede extender naturalmente para todo cuerpo $\Bbbk$, pero lo que estamos llamando polinomio es en realidad una *función polinómica*. Intuitivamente, un *polinomio* con coeficientes en $\Bbbk$ es una expresión del tipo $a_0+a_1x+\cdots+a_nx^n$, en la cual $a_0, a_1 \ldots, a_n \in \Bbbk$ y la letra x es una *indeterminada*, es decir, un símbolo que se puede cambiar por cualquier otro. Uno opera con los polinomios de la manera usual, sumando coeficiente a coeficiente, etc.

Una manera de formalizar esta idea, es considerar un polinomio como una "sucesión finita", es decir identificar el polinomio con la sucesión de sus coeficientes y luego completar con ceros

$$a_0 + a_1x + \cdots + a_nx^n \;\leftrightarrow\; (a_0, a_1, \ldots, a_n, 0, 0, \ldots).$$

En ese sentido, si V es el espacio de las sucesiones en $\Bbbk$, entonces los polinomios se pueden pensar como el subconjunto

$$W = \{(a_n)_{n\in\mathbb{N}} \in V :\ \exists k \in \mathbb{N} \text{ tal que } a_n = 0,\ \forall n \geq k\}.$$

Es claro que la sucesión nula está en W y que si $(a_n) \in W$ y $c \in \Bbbk$, entonces $c(a_n) \in W$. Además, si $(a_n), (b_n) \in W$ y $k, h \in \mathbb{N}$ son tales que $a_n = 0$, para todo $n \geq k$ y $b_n = 0$, para todo $n \geq h$, entonces $a_n+b_n = 0$, para todo $n \geq \max\{h, k\}$. Luego W es un subespacio de V y por lo tanto es un espacio vectorial. El *producto* de polinomios corresponde a un producto definido en W mediante lo siguiente: si $\alpha = (a_n)$ y $\beta = (b_n)$, entonces $\alpha\beta = (c_n)$, siendo

$$c_n := a_0b_n + a_1b_{n-1} + \cdots a_{n-1}b_1 + a_nb_0, \quad \forall n \in \mathbb{N}.$$

La función $f : \Bbbk \to W$ definida por $f(a) := (a, 0, 0, \ldots)$ es inyectiva y preserva la suma y el producto, es decir

$$(a, 0, 0, \ldots) + (b, 0, 0, \ldots) = (a + b, 0, 0, \ldots), \quad (a, 0, 0, \ldots) \cdot (b, 0, 0, \ldots) = (a \cdot b, 0, 0, \ldots).$$

Luego podemos pensar $\Bbbk \subset W$ identificando $\Bbbk$ con $f(\Bbbk) \subset W$. En base a lo anterior, escribiremos $a = (a, 0, 0, \ldots)$. Notar que con esta identificación vale

$$a(b_0, \ldots, b_n, 0, 0, \ldots) = (ab_0, \ldots, ab_n, 0, 0, \ldots).$$

Además, si llamamos $x = (0, 1, 0, 0, \cdots)$, entonces usando la fórmula del producto obtenemos

$$x = (0, 1, 0, 0, \cdots), \quad x^2 = (0, 0, 1, 0, 0, \cdots), \quad x^3 = (0, 0, 0, 1, 0, 0, \cdots), \quad \ldots$$

Lo anterior implica

$$\begin{aligned}(a_0, a_1, \ldots, a_n, 0, 0, \ldots) &= (a_0, 0, 0, \ldots) + (0, a_1, 0, 0, \ldots) + \cdots + (0, \ldots, 0, a_n, 0, 0, \ldots)\\ &= (a_0, 0, 0, \ldots) + a_1(0, 1, 0, 0, \ldots) + \cdots + a_n(0, \ldots, 0, 1, 0, 0, \ldots)\\ &= a_0 + a_1x + \cdots + a_nx^n.\end{aligned}$$

Luego $(a_0, a_1, \ldots, a_n, 0, 0, \ldots) = a_0 + a_1x + \cdots + a_nx^n$ y así retomamos la notación usual de los polinomios.

Al conjunto de los polinomios con coeficientes en $\Bbbk$ lo escribimos $\Bbbk[x]$. El *polinomio nulo* es $p(x) = 0$ y corresponde a la sucesión nula $(0, 0, \ldots)$. Si $p(x) = a_0 + a_1x + \cdots + a_nx^n$ con $a_n \neq 0$, entonces decimos que el polinomio $p(x)$ tiene *grado* n. El polinomio nulo se dice que tiene grado $-\infty$ y convendremos en que tiene grado menor que n, para todo $n \in \mathbb{N}$. Al conjunto de todos los polinomios de grado menor o igual a n lo escribimos $\Bbbk_n[x]$. Notar que $\Bbbk_n[x]$ contiene al polinomio nulo y es cerrado bajo la suma y el producto por escalares, así que $\Bbbk_n[x]$ es un subespacio de $\Bbbk[x]$, para todo $n \in \mathbb{N}$.

Observación 4.2.5. En contextos algebraicos conviene pensar los polinomios como objetos más que como funciones. No hay problemas en pensar los polinomios como funciones cuando el cuerpo es $\mathbb{R}$, pero aparecen complicaciones al trabajar con cuerpos finitos (lo cual no haremos).

Combinaciones lineales y subespacios generados por conjuntos. Una *combinación lineal* de un conjunto de vectores $v_1, \dots, v_n \in V$ es un vector de la forma $w = a_1v_1 + \cdots + a_nv_n$, siendo $a_1, \dots, a_n$ escalares arbitrarios. Consideremos el conjunto formado por todas las combinaciones lineales de $v_1, \dots, v_n$,

$$[v_1, \dots, v_n] := \{a_1v_1 + \cdots + a_nv_n : \ a_1, \dots, a_n \in \Bbbk\}.$$

Proposición 4.2.6. *Para todo $v_1, \dots, v_n \in V$, el conjunto $[v_1, \dots, v_n]$ es un subespacio de V.*

Dem. El vector nulo 0 verifica $0 = 0v_1 + \cdots + 0v_n$, luego $0 \in [v_1, \dots, v_n]$. Sean $u, v \in [v_1, \dots, v_n]$ y $c \in \Bbbk$. Sabemos que existen escalares $a_i, b_j \in \Bbbk$ tales que $u = a_1v_1 + \cdots + a_nv_n$ y $v = b_1v_1 + \cdots + b_nv_n$, entonces

$$u + cv = (a_1v_1 + \cdots + a_nv_n) + c(b_1v_1 + \cdots + b_nv_n) = (a_1 + cb_1)v_1 + \cdots + (a_n + cb_n)v_n \in [v_1, \dots, v_n]. \quad \square$$

Observación 4.2.7. Es fácil de probar por inducción, que si un subespacio W contiene a $v_1, \dots, v_n$, entonces W contiene también a toda combinación lineal de $v_1, \dots, v_n$ y por lo tanto $[v_1, \dots, v_n] \subset W$. Luego $[v_1, \dots, v_n]$ es el "menor" subespacio de V que contiene a $v_1, \dots, v_n$.

Al conjunto $[v_1, \dots, v_n]$ se le llama el *subespacio generado* por $v_1, \dots, v_n$.

Ejemplos 4.2.8.

1. Si v es un vector no nulo de $\mathbb{R}^2$ o $\mathbb{R}^3$, 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).

Finalizando, definimos $[\emptyset] := \{0\}$, el cual es el menor subespacio que contiene al conjunto vacío $\emptyset$. Así tenemos definido $[X]$, para todo subconjunto finito X de V.

4.3. Dependencia lineal

Sean $v_1, \dots, v_n \in V$. Decimos que el conjunto $\{v_1, \dots, v_n\}$ es *linealmente dependiente* (LD) si existen escalares no todos nulos $a_1, \dots, a_n$ tales que $a_1v_1 + \cdots + a_nv_n = 0$. En caso contrario decimos que $\{v_1, \dots, v_n\}$ es *linealmente independiente* (LI). También se dice que los vectores $v_1, \dots, v_n \in V$ son *linealmente dependientes* o *linealmente independientes*, respectivamente.

La *combinación lineal trivial* de $v_1, \dots, v_n$ es $0v_1 + \cdots + 0v_n = 0$. Luego $\{v_1, \dots, v_n\}$ es LI si y solo si la única forma de escribir el vector nulo como combinación lineal de $v_1, \dots, v_n$ es la trivial, y es LD si y solo si el vector nulo se puede escribir como una combinación lineal no trivial de $v_1, \dots, v_n$.

Ejemplos 4.3.1.

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

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

Este sistema es compatible indeterminado (tiene más incógnitas que ecuaciones). Luego el sistema admite soluciones no triviales y por lo tanto X es LD. Las soluciones del sistema son de la forma $a = \frac{5}{2}b$, $c = \frac{3}{2}b$ y b libre. Eligiendo $b = 2$, obtenemos $a = 5$ y $c = 3$. Luego una combinación lineal no trivial que da el vector nulo es

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

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

$$a(1,1,1) + b(1,2,1) + c(2,3,3) = (0,0,0).$$

Esta ecuación equivale al sistema homogéneo

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

Calculando el determinante del sistema, obtenemos $\begin{vmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 3 \end{vmatrix} = 1$. Como el determinante es distinto de cero, la única solución que admite el sistema es la trivial y por lo tanto X es LI.

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

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

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

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

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

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

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

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

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

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

5. Es fácil de probar que si en $\Bbbk^n$ consideramos

$$e_1 = (1,0,\dots,0),\ e_2 = (0,1,0,\dots,0),\ \dots,\ e_n = (0,0,\dots,0,1),$$

entonces $\{e_1, e_2, \dots, e_n\} \subset \Bbbk^n$ es LI. En particular la base canónica $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ de $\mathbb{R}^3$ es un conjunto LI.

Observaciones 4.3.2. 1. Dado $v \in V$, el conjunto $\{v\}$ es LD si y solo si $v = 0$ (esto se debe a la última afirmación de la proposición 4.1.3).

2. Si en $X = \{v_1, \dots, v_n\} \subset V$ se cumple que existe algún i tal que $v_i = 0$, entonces vale

$$0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_n = 0$$

y por lo tanto X es LD.

Proposición 4.3.3. *Un conjunto $X = \{v_1, \ldots, v_n\} \subset V$ (con $n \geq 2$) es LD si y solo si X contiene un vector que es combinación lineal de los restantes.*

Dem. Si v_i es combinación lineal de $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$, entonces v_i se puede escribir de la forma $v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_nv_n$, luego

$$a_1v_1 + \cdots + a_{i-1}v_{i-1} + (-1)v_i + a_{i+1}v_{i+1} + \cdots + a_nv_n = 0$$

es una combinación lineal no trivial de $v_1, \ldots, v_n$. Recíprocamente, si $\{v_1, \ldots, v_n\}$ es LD, entonces existen escalares no todos nulos $b_1, \ldots, b_n$ tales que $b_1v_1+\cdots+b_nv_n = 0$. Luego si $b_i \neq 0$, entonces podemos despejar v_i de la igualdad anterior, obteniendo

$$v_i = \left(\frac{-b_1}{b_i}\right) v_1 + \cdots + \left(\frac{-b_{i-1}}{b_i}\right) v_{i-1} + \left(\frac{-b_{i+1}}{b_i}\right) v_{i+1} + \cdots + \left(\frac{-b_n}{b_i}\right) v_n. \quad \square$$

Corolario 4.3.4. *Un conjunto $\{u, v\}$ es LD si y solo si existe $a \in \Bbbk$ tal que $u = av$ o $v = au$.* $\square$

Ejemplo 4.3.5. En el primer ejemplo 4.3.1.1 vimos que el conjunto $X = \{(-4, 1),\ (1, 1),\ (1, -1)\} \subset \mathbb{R}^2$ es LD y que vale

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

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

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

A continuación veremos que si achicamos un conjunto LI vamos a seguir obteniendo un conjunto LI y que si agrandamos un conjunto LD volvemos a obtener un conjunto LD.

Proposición 4.3.6. *Sean A y B subconjuntos finitos de V, tales que $A \subset B$.*

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

Dem. Supongamos $A = \{v_1, \ldots, v_n\}$ y $B = \{v_1, \ldots, v_n, v_{n+1}, \cdots, v_m\}$. Si A es LD, entonces existen escalares no todos nulos $a_1, \ldots, a_n$ tales que $a_1v_1 + \cdots + a_nv_n = 0$. Luego

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

es una combinación lineal no trivial de los elementos de B que da el vector nulo y por lo tanto B es LD. Supongamos ahora que B es LI. Si A no fuese LI, entonces sería LD y por la primera parte obtendríamos que B es LD, contradiciendo lo supuesto. Luego A es LI[1]. $\square$

Si agrandamos un conjunto LI no tenemos porqué seguir obteniendo un conjunto LI, de la misma forma que si achicamos un LD no tenemos porqué seguir obteniendo un LD. Alcanza con pensar en

$$A = \{(1, 0), (0, 1)\}, \quad B = \{(1, 0), (0, 1), (1, 1)\}, \quad A, B \subset \mathbb{R}^2.$$

Es claro que A es LI, B es LD y es $A \subset B$.

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

Proposición 4.3.7. *Si $\{v_1, \ldots, v_n\} \subset V$ es LI y $v_{n+1} \in V$ no es combinación lineal de $v_1, \ldots, v_n$, entonces $\{v_1, \ldots, v_n, v_{n+1}\}$ es LI.*

[1] En realidad ambas afirmaciones son lógicamente equivalentes, dado que cada una es el contrarrecíproco de la otra.

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

Por razones que veremos más adelante, nos interesa definir si el conjunto vacío $\emptyset$ es LD o LI. Como queremos que se verifiquen todas las propiedades anteriores, entonces la proposición 4.3.6 nos fuerza a definir que el conjunto vacío es LI (si fuese LD, entonces todos los subconjuntos finitos de V serían LD). De esta forma tenemos definido el concepto de LD o LI para todo subconjunto finito de V.

Generadores. Sea $B = \{v_1, \dots, v_n\}$ un subconjunto de V. Decimos que B es un *generador*[2] de V si todo vector de V se puede escribir como combinación lineal de $v_1, \dots, v_n$, es decir, si $[v_1, \dots, v_n] = V$.

Ejemplo 4.3.8. Consideremos en $\mathbb{R}^2$ los conjuntos $A_1 = \{(1,2),(1,1),(-3,5)\}$, $A_2 = \{(1,2),(2,3)\}$ y $A_3 = \{(3,-6),(-5,10)\}$. Vamos a investigar si son generadores de $\mathbb{R}^2$. Empezamos por A_1.

$$(x,y) = a(1,2) + b(1,1) + c(-3,5) \;\Rightarrow\; \left\{ \begin{array}{rcl} a + b - 3c &=& x \\ 2a + b + 5c &=& y \end{array} \right. .$$

Este sistema es compatible indeterminado (para todo valor de x, y). Esto nos dice que A_1 es un generador. Consideremos ahora A_2.

$$(x,y) = a(1,2) + b(2,3) \;\Rightarrow\; \left\{ \begin{array}{rcl} a + 2b &=& x \\ 2a + 3b &=& y \end{array} \right. .$$

Este sistema es compatible determinado (para todo valor de x, y). Luego A_2 es un generador. Consideremos A_3.

$$(x,y) = a(3,-6) + b(-5,10) \;\Rightarrow\; \left\{ \begin{array}{rcl} 3a - 5b &=& x \\ -6a + 10b &=& y \end{array} \right. .$$

Si multiplicamos la ecuación de arriba por 2 y la sumamos a la de abajo, obtenemos $2x + y = 0$. Luego si (x,y) es tal que $2x + y \neq 0$, entonces el sistema es incompatible. Esto nos dice que A_3 no es un generador.

Es claro que si agrandamos un generador, entonces seguimos obteniendo un generador, pero si lo achicamos el conjunto obtenido no tiene por qué seguirlo siendo. El siguiente resultado nos da una condición para poder achicar un generador y que siga siendo generador.

Proposición 4.3.9. *Si $G = \{v_1, \dots, v_n, v_{n+1}\}$ es un generador de V y v_{n+1} es combinación lineal de $v_1, \dots, v_n$, entonces $G' = \{v_1, \dots, v_n\}$ también es un generador de V.*

Dem. Sean $b_1, \dots, b_n \in \mathbb{k}$ tales que $v_{n+1} = b_1 v_1 + \cdots + b_n v_n$. Si $v \in V$ es un vector arbitrario, como G es un generador de V, sabemos que existen $a_1, \dots, a_{n+1} \in \mathbb{k}$ tales que $v = a_1 v_1 + \cdots + a_n v_n + a_{n+1} v_{n+1}$. Luego

$$\begin{aligned} v &= a_1 v_1 + \cdots + a_n v_n + a_{n+1} v_{n+1} = a_1 v_1 + \cdots + a_n v_n + a_{n+1}(b_1 v_1 + \cdots + b_n v_n) \\ &= (a_1 + a_{n+1} b_1) v_1 + \cdots + (a_n + + a_{n+1} b_n) v_n. \end{aligned}$$

Esto prueba que todo vector de V se puede escribir como combinación lineal de G'. □

El siguiente resultado es importante.

Teorema 4.3.10. *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.*

[2]También se usan otros nombres, como *conjunto generador* o *conjunto de generadores* o *sistema generador*, aunque en Uruguay lo normal es llamarle simplemente generador.

Dem. Sea $A = \{v_1, \dots, v_n\}$ un conjunto LI y $G = \{w_1, \dots, w_m\}$ 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 $a_1, \dots, a_m \in \Bbbk$ tales que $v_1 = a_1 w_1 + \cdots + a_m w_m$. Si fuesen $a_1 = \cdots = a_m = 0$, entonces sería $v_1 = 0$, pero esto es imposible porque $v_1 \in A$ y A es LI. Luego alguno de los escalares es no nulo. Eventualmente reordenando los elementos de G, podemos suponer $a_1 \neq 0$. Luego

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

Veamos que $G_1 := \{v_1, w_2, \dots, w_m\}$ es un generador de V. Como G es un generador, entonces $G \cup \{v_1\} = \{v_1, w_1, w_2, \dots, w_m\}$ también es un generador. Pero recién probamos que w_1 es combinación lineal de $v_1, w_2, \dots, w_m$, luego la proposición 4.3.9 implica que $\{v_1, w_2, \dots, w_m\}$ es un generador.

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

$$v_{k+1} = c_1 v_1 + \cdots + c_k v_k + c_{k+1} w_{k+1} + \cdots + c_m w_m. \tag{4.3}$$

Si fuese $c_{k+1} = \cdots = c_m = 0$, entonces sería $v_{k+1} = c_1 v_1 + \cdots + c_k v_k$, contradiciendo que A es LI. Luego existe algún i, con $k+1 \leq i \leq m$, tal que $c_i \neq 0$. Como antes podemos suponer $c_{k+1} \neq 0$, y por lo tanto podemos despejar w_{k+1} de (4.3) obteniendo que w_{k+1} es combinación lineal de $v_1, \dots, v_k, v_{k+1}, w_{k+2}, \dots, w_m$.

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 es la misma que al principio. Como G_k es un generador, entonces

$$G_k \cup \{v_{k+1}\} = \{v_1, \dots, v_k, v_{k+1}, w_{k+1}, w_{k+2}, \dots, w_m\}$$

es también un generador. Pero recién vimos que w_{k+1} es combinación lineal de los otros, luego podemos sacarlo y el conjunto resultante

$$G_{k+1} := \{v_1, \dots, v_k, v_{k+1}, w_{k+2}, \dots, w_m\}$$

sigue siendo un generador.

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 A obteniendo $G_m = \{v_1, \dots, v_m\}$ que es un generador a V. Pero esto es imposible porque en ese caso el elemento $v_{m+1} \in A$ sería combinación lineal de $v_1, \dots, v_m$, contradiciendo que A es LI. Luego necesariamente es $n \leq m$. $\square$

Bases. Sea $B = \{v_1, \dots, v_n\}$ un subconjunto de V. Decimos que B es una *base* de V si B es un generador y además es un conjunto LI. Si $V = \{0\}$ (el espacio trivial), entonces definimos que el conjunto vacío $\emptyset$ es la base de V. Esto es coherente con la definición anterior, ya que $\emptyset$ es LI y $[\emptyset] = \{0\}$.

Proposición 4.3.11. *Un conjunto $B = \{v_1, \dots, v_n\}$ es una base de V si y solo si todo vector de V se puede escribir de forma única como combinación lineal de $v_1, \dots, v_n$.*

Dem. Empezamos probando el directo. Como B es un generador de V, sabemos que todo vector de V se puede escribir como combinación lineal de los elementos de B; lo que tenemos que probar es la unicidad. Sea $v \in V$ y supongamos que podemos escribir v de dos formas

$$v = a_1 v_1 + \cdots + a_n v_n, \qquad v = b_1 v_1 + \cdots + b_n v_n.$$

Restando ambas igualdades obtenemos $(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = 0$. Como B es LI, deducimos $a_i - b_i = 0$, para todo i, es decir $a_i = b_i$, para todo i. Esto prueba la unicidad.

Consideremos ahora el recíproco. Es claro que B es un generador. Para probar que es LI, observemos que si $a_1, \dots, a_n$ son escalares tales que $a_1v_1 + \cdots + a_nv_n = 0$, entonces

$$a_1v_1 + \cdots + a_nv_n = 0v_1 + \cdots + 0v_n.$$

Como hay unicidad en la descomposición, deducimos que es $a_1 = \cdots = a_n = 0$. Luego B es LI. $\square$

Ejemplo 4.3.12. Si consideramos los conjuntos A_1, A_2 y A_3 del ejemplo 4.3.8, entonces A_1 es un generador pero no es base, A_2 es una base y A_3 no es un generador, así que tampoco es una base.

Ejemplo 4.3.13. Consideremos el plano $W = \{(x, y, z) \in \mathbb{R}^3 : \ 2x + y + z = 0\}$. Si $(x, y, z) \in W$, entonces $2x + y + z = 0$ y por lo tanto $z = -2x - y$. Luego podemos escribir

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

Esto nos dice que $B = \{(1, 0, -2), (0, 1, -1)\}$ es un generador de W (notar que $(1, 0, -2)$ y $(0, 1, -1)$ pertenecen a W). Además es fácil de probar que B es LI, luego B es una base del plano W.

En algunos espacios hay ciertas bases estándar a las que se les suele llamar la *base canónica* del espacio.

1. La base canónica de $\mathbb{k}^n$ es el conjunto $\{e_1, e_2, \dots, e_n\}$ definido en el ejemplo 4.3.1.5. Notar que para $n = 2$ y $n = 3$ obtenemos las bases canónicas $\{\mathbf{i}, \mathbf{j}\}$ de $\mathbb{R}^2$ y $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ de $\mathbb{R}^3$.
2. La base canónica de[3] $\mathbb{k}_n[x]$ es el conjunto $\{1, x, \dots, x^n\}$.
3. La base canónica de $M_{m\times n}$ es el conjunto $\{e_{ij} : \ i = 1, \dots, m, \ j = 1, \dots, n\}$, 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 es $B = \{e_{11}, e_{12}, e_{21}, e_{22}\}$, siendo $e_{11} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $e_{12} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $e_{21} = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ y $e_{22} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

Los siguientes resultados son consecuencia del teorema 4.3.10.

Proposición 4.3.14. *Si V es un espacio vectorial, A es un subconjunto LI, B es una base y G es un generador, entonces $\#A \leq \#B \leq \#G$.* $\square$

Dem. Dado que B es base, entonces B es un generador y es LI. Como A es LI y B es un generador, entonces $\#A \leq \#B$, y como B es LI y G es un generador, es $\#B \leq \#G$. $\square$

Proposición 4.3.15. *Si B_1 y B_2 son bases de V, entonces $\#B_1 = \#B_2$.*

Dem. Dado que B_1 y B_2 son bases, entonces son generadores y son LI. Como B_1 es LI y B_2 es un generador, entonces $\#B_1 \leq \#B_2$. Intercambiando lo roles de B_1 y B_2, obtenemos $\#B_2 \leq \#B_1$. Luego $\#B_1 = \#B_2$. $\square$

La proposición anterior implica que podemos dar la siguiente definición.

[3] Acá hay dos posibilidades dependiendo de si escribimos los polinomios en potencias crecientes o decrecientes; en el segundo caso conviene tomar como base canónica a $\{x^n, \dots, x, 1\}$.

Dimensión. Si en un espacio V existe un conjunto finito $B \subset V$ tal que B es una base de V, entonces diremos que V tiene *dimensión finita*, llamamos *dimensión* de V a la cantidad de elementos de B y escribimos $\dim V = \#B$. En caso contrario diremos que V tiene *dimensión infinita*. En particular el espacio trivial $\{0\}$ tiene dimensión finita y $\dim\{0\} = 0$.

Ejemplos 4.3.16. 1. Teniendo en cuenta las bases canónicas de la página 71, deducimos $\dim \mathbb{k}^n = n$, $\dim \mathbb{k}_n[x] = n+1$ y $\dim M_{m\times n} = mn$, para todo m, n.

2. El conjunto $B = \{(1,0,-2),(0,1,-1)\}$ es una base del plano $W = \{(x,y,z) \in \mathbb{R}^3 : \ 2x+y+z=0\}$ (ejemplo 4.3.13), luego $\dim W = 2$.

Notar que ahora la proposición 4.3.14 puede enunciarse de la siguiente manera.

Proposición 4.3.17. *En un espacio de dimensión finita la cantidad de elementos de un conjunto LI es menor o igual que la dimensión del espacio y la cantidad de elementos de un generador es mayor o igual a la dimensión del espacio.* □

El siguiente resultado es útil cuando se conoce la dimensión del espacio.

Proposición 4.3.18. *Sea V un espacio de dimensión finita. Si $A \subset V$ tiene la misma cantidad de elementos que la dimensión de V, entonces para que A sea una base alcanza con que sea un generador o que sea LI.*

Dem. En ambos casos vamos a razonar por el absurdo. Supongamos que A es LI. Si A no es base, entonces no es un generador de V y por lo tanto existe un vector $v \in V$ que no es combinación lineal de A. Esto último implica que $A \cup \{v\}$ es LI (proposición 4.3.7). Entonces $A \cup \{v\}$ es un conjunto LI que tiene más elementos que la dimensión del espacio $\lightning$.

Supongamos ahora que A es un generador. Si A no es base, entonces A no es LI y por lo tanto contiene un vector w que es combinación lineal de los restantes. Luego $A \setminus \{w\}$ sigue siendo un generador de V (proposición 4.3.9) y tiene menos elementos que la dimensión del espacio $\lightning$. □

La siguiente proposición establece un criterio simple para saber si un conjunto es una base de $\mathbb{k}^n$.

Proposición 4.3.19. *Sea $B = \{v_1, \dots, v_n\}$ un subconjunto de $\mathbb{k}^n$ y sea Δ el determinante de la matriz cuyas columnas son $v_1, \dots, v_n$ escritos verticalmente. Entonces B es base de $\mathbb{k}^n$ si y solo si $\Delta \neq 0$.*

Dem. Sean $v_1 = (a_{11}, \dots, a_{n1}), \ \cdots, \ v_n = (a_{1n}, \dots, a_{nn})$. Si $x_1, \dots, x_n \in \mathbb{k}$ verifican $x_1v_1 + \cdots + x_nv_n = 0$, entonces

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

Por la proposición 4.3.18 sabemos que B es base de $\mathbb{k}^n$ si y solo si es LI, y esto ocurre si y solo si la única solución que admite el sistema homogéneo anterior es la trivial, lo cual sucede si y solo si $\Delta \neq 0$. □

Ejemplo 4.3.20. Consideremos los siguientes subconjuntos de $\mathbb{R}^3$: $B_1 = \{(1,2,3), (4,0,1), (3,0,1)\}$ y $B_2 = \{(1,-1,-1), (5,-3,-4), (1,5,2)\}$. Aplicando la proposición anterior obtenemos

$$\begin{vmatrix} 1 & 4 & 3 \\ 2 & 0 & 0 \\ 3 & 1 & 1 \end{vmatrix} = -2 \begin{vmatrix} 4 & 3 \\ 1 & 1 \end{vmatrix} = -2 \neq 0; \qquad \begin{vmatrix} 1 & 5 & 1 \\ -1 & -3 & 5 \\ -1 & -4 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 1 \\ 0 & 2 & 6 \\ 0 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 6 \\ 1 & 3 \end{vmatrix} = 0.$$

Luego B_1 es base y B_2 no lo es.

Los dos resultados siguientes muestran que todo generador contiene una base y que todo conjunto LI está contenido en alguna base.

Proposición 4.3.21. *Si existe un conjunto finito $G \subset V$ tal que G es un generador de V, entonces G contiene una base de V y por lo tanto la dimensión de V es finita.*

Dem. Si G es LI, entonces es base y ya está probado. Si no, entonces G es LD y por lo tanto contiene un vector que es combinación lineal de los restantes. Sea G_1 el conjunto obtenido quitándole a G ese vector. Entonces G_1 es un generador de V (por la proposición 4.3.9) y tiene un elemento menos que G. Si G_1 es LI, entonces es base y ya está probado. Si no, entonces razonamos con G_1 como con G para obtener un generador $G_2 \subset G_1$ con un elemento menos que G_1 y así seguimos. Como la cantidad de elementos de G es finita, en algún momento tenemos que llegar a un subconjunto de G que es generador y LI, y por lo tanto es base. □

Proposición 4.3.22. *Sea V un espacio de dimensión finita. Si $A \subset V$ es LI, entonces A está contenido en alguna base de V.*

Dem. Sea $\dim V = n$ y $A \subset V$ un LI con m elementos. Sabemos que es $m \leq n$. Si $m = n$, entonces la proposición 4.3.18 implica que A es una base de V. Supongamos ahora $m < n$. Entonces A no es una base de V y por lo tanto no es un generador de V. Luego existe un vector $v_{m+1} \in V$ tal que $v_{m+1} \notin [A]$ y por lo tanto $A \cup \{v_{m+1}\}$ es LI. Si $m + 1 = n$, entonces $A \cup \{v_{m+1}\}$ es un conjunto LI con n elementos y por lo tanto es una base de V. Si $m + 1 < n$, entonces repetimos el razonamiento con $A \cup \{v_{m+1}\}$ en lugar de A y así seguimos hasta llegar a un conjunto LI con n elementos, que por lo tanto es base de V. □

Decimos que $A \subset V$ es un conjunto LI *maximal* si A es LI y no está contenido estrictamente en ningún otro conjunto LI. Análogamente, decimos que G es un generador *minimal* si es un generador y no contiene estrictamente a ningún otro generador.

Proposición 4.3.23. *Sea V un espacio de dimensión finita. Si A es un subconjunto de V, entonces las siguientes afirmaciones son equivalentes:*

A es un generador minimal de V, A es un conjunto LI maximal, A es una base de V.

Dem. Sea A una base de V, entonces A es LI y es un generador de V. Si a A le sacamos un elemento, entonces obtenemos un conjunto con menos elementos que la dimensión de V y por lo tanto ese conjunto no pude ser un generador. Esto prueba que una base es un generador minimal. Por otro lado, si a A le agregamos un elemento, entonces obtenemos un conjunto con más elementos que la dimensión de V y por lo tanto ese conjunto no pude ser LI. Esto prueba que una base es un LI maximal.

Sea A un generador minimal de V. Como A es un generador, entonces existe B base de V tal que $B \subset A$. Como B es un generador y $B \subset A$, entonces la minimalidad de A implica $B = A$. Luego A es una base de V. El caso en que A es un LI maximal se prueba en forma análoga, aplicando la proposición 4.3.22. □

Dimensión de subespacios. Definimos la *dimensión* de un subespacio W de un espacio V, como la dimensión de W como espacio vectorial.

Proposición 4.3.24. *Sea V un espacio de dimensión finita. Si W es un subespacio de V, entonces W tiene dimensión finita y* $\dim W \leq \dim V$. *Además, si* $\dim W = \dim V$, *entonces* $W = V$.

Dem. Sea $n = \dim V$. Si $W = \{0\}$, el resultado es obvio. Supongamos ahora $W \neq \{0\}$. Como todo subconjunto de W es también un subconjunto de V, entonces un subconjunto de W con más de n elementos es LD. Esto prueba que todo subconjunto LI de W tiene a lo más n elementos. Sea A un subconjunto LI de

W con la cantidad máxima posible de elementos. Por lo anterior es $\#A \leq n$ y además A es un subconjunto LI maximal de W, luego A es una base de W. Esto a su vez implica $\dim W \leq \dim V$.
Respecto a la segunda afirmación, si $\dim W = \dim V$ y B es una base de W, entonces B es LI y tiene tantos elementos como la dimensión de V, luego B es una base de V y por lo tanto $V = [B] = W$. □

Subespacios de $\mathbb{R}^2$ y de $\mathbb{R}^3$. Si W es un subespacio de $\mathbb{R}^2$, entonces $\dim W \leq \dim \mathbb{R}^2 = 2$. Si $\dim W = 0$, entonces $W = \{0\}$ y si $\dim W = 2$, entonces $W = \mathbb{R}^2$. Si $\dim W = 1$, entonces existe $0 \neq v \in \mathbb{R}^2$ tal que $W = \{tv : \ t \in \mathbb{R}\}$, luego W es una recta que pasa por el origen.

Si W es un subespacio de $\mathbb{R}^3$, entonces $\dim W \leq 3$. Si $\dim W = 0$, entonces $W = \{0\}$, si $\dim W = 3$, entonces $W = \mathbb{R}^3$ y si $\dim W = 1$, W entonces es una recta que pasa por el origen. Si $\dim W = 2$, entonces existe $\{u, v\} \subset \mathbb{R}^3$ conjunto LI tal que $W = \{tu + sv : \ t \in \mathbb{R}\}$ y por lo tanto W es un plano que pasa por el origen.

4.4. Operaciones con subespacios

Sean W_1, W_2 subespacios de V. Como los subespacios son conjuntos, entonces podemos considerar su intersección

$$W_1 \cap W_2 := \{v \in V : \ v \in W_1 \text{ y } v \in W_2\}.$$

Definimos la *suma* de W_1 y W_2, mediante

$$W_1 + W_2 := \{w_1 + w_2 : \ w_1 \in W_1, \ w_2 \in W_2\}.$$

Proposición 4.4.1. *Si W_1 y W_2 son subespacios, entonces $W_1 \cap W_2$ y $W_1 + W_2$ también son subespacios.*

Dem. Probaremos que la suma es un subespacio, dejando la prueba de la intersección como ejercicio. Lo primero es notar que, como el vector nulo 0 está en W_1 y W_2, luego de $0 = 0 + 0$ se deduce que $0 \in W_1 + W_2$. Supongamos ahora que tenemos $u, v \in W_1 + W_2$ y $a \in \Bbbk$. Como $u, v \in W_1 + W_2$, entonces existen $w_1, w_1' \in W_1$ y $w_2, w_2' \in W_2$ tales que $u = w_1 + w_2$ y $v = w_1' + w_2'$. Luego $au + v = aw_1 + w_1' + aw_2 + w_2'$, con $aw_1 + w_1' \in W_1$ y $aw_2 + w_2' \in W_2$. Esto implica $au + v \in W_1 + W_2$. □

Observación 4.4.2. Dados dos subespacios W_1, W_2, su intersección $W_1 \cap W_2$ es el mayor subespacio contenido en W_1 y W_2, mientras que $W_1 + W_2$ es el menor subespacio que contiene a W_1 y W_2. Notar que la unión $W_1 \cup W_2$ en general no es un subespacio.

Ejemplos 4.4.3. 1. Consideremos el espacio $\mathbb{R}^3$.

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

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

Para ver esto último, notar que hay varias formas de escribir un vector de $\mathbb{R}^3$ como suma de uno en W_1 con uno de W_2, por ejemplo

$$(x, y, z) = (x, y, 0) + (0, 0, z), \quad (x, y, z) = (0, y, 0) + (x, 0, z), \quad (x, y, z) = (x/2, y, 0) + (x/2, 0, z).$$

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

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

Esto último sale de escribir $(x, y, 0) = (x, 0, 0) + (0, y, 0)$; esta es la única forma de escribir un vector del plano Oxy como suma de uno en W_1 con uno de W_2.

Ejemplo 4.4.4. Consideremos el espacio $\mathbb{R}^4$ y los subespacios

$$W_1 = \{(x, y, z, t) : \ y = z = t)\}, \quad W_2 = \{(x, y, z, t) : \ x = y, \ z = t)\}.$$

Su intersección es $W_1 \cap W_2 = \{(x, y, z, t) : \ x = y = z = t)\} = \{(x, x, x, x) : \ x \in \mathbb{R})\}$. Consideremos ahora la suma. Primero observar que podemos escribir

$$W_1 = \{(x, y, y, y) : \ x, y \in \mathbb{R})\}, \quad W_2 = \{(x, x, z, z) : \ x, z \in \mathbb{R})\}. \tag{4.4}$$

Luego para saber si un vector (x, y, z, t) está en $W_1 + W_2$ tenemos que ver si la siguiente ecuación en las variables a, b, c, d tiene solución

$$(x, y, z, t) = (a, b, b, b) + (c, c, d, d). \tag{4.5}$$

Operando obtenemos que la ecuación de arriba equivale a

$$(x, y, z, t) = (a + c, b + c, b + d, b + d) \quad \Leftrightarrow \quad \begin{cases} a + c &= x \\ b + c &= y \\ b + d &= z \\ b + d &= t \end{cases}.$$

Las dos últimas ecuaciones implican que, para que haya solución, tiene que ser $z = t$. En ese caso, eliminando la última ecuación obtenemos un sistema compatible (indeterminado). Esto nos dice que, dado $(x, y, z, t) \in \mathbb{R}^4$, si $z = t$, entonces siempre existen $a, b, c, d \in \mathbb{R}$ que verifican (4.5). Luego $W_1 + W_2 = \{(x, y, z, t) : \ z = t\}$.

De ahora en adelante supondremos que V es un espacio de dimensión finita.

Proposición 4.4.5. *Sean W_1 y W_2 dos subespacios de V. Si G_1 un generador de W_1 y G_2 un generador de W_2, entonces $G := G_1 \cup G_2$ es un generador de $W_1 + W_2$.*

Dem. Sean $G_1 = \{u_1, \dots, u_n\}$ y $G_2 = \{v_1, \dots, v_m\}$. Si $v \in W_1 + W_2$, entonces existen $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$. Como G_1 y G_2 son generadores respectivos de W_1 y W_2, entonces existen escalares $a_1, \dots, a_n$ y $b_1, \dots, b_m$ tales que $w_1 = a_1u_1 + \cdots + a_nu_n$ y $w_2 = b_1v_1 + \cdots + b_mv_m$. Luego

$$v = w_1 + w_2 = a_1u_1 + \cdots + a_nu_n + b_1v_1 + \cdots + b_mv_m.$$

Esto prueba que $G = \{u_1, \dots, u_n, v_1, \dots, v_m\}$ es un generador de $W_1 + W_2$. $\square$

Observación 4.4.6. Si W_1 y W_2 dos subespacios de V, B_1 es una base de W_1 y B_2 es una base de W_2, entonces $B := B_1 \cup B_2$ es un generador de $W_1 + W_2$, pero no es necesariamente una base, dado que puede ser LD (ver más adelante el ejemplo 4.4.9).

Teorema 4.4.7. *Si W_1 y W_2 son subespacios de V, entonces*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Dem. Sea $B_0 = \{u_1, \dots, u_n\}$ una base de $W_1 \cap W_2$. Como B_0 es un subconjunto LI de W_1 y de W_2, entonces existen $v_1, \dots, v_p \in W_1$ y $w_1, \dots, w_q \in W_2$, tales que si

$$B_1 = \{u_1, \dots, u_n, v_1, \dots, v_p\} \quad \text{y} \quad B_2 = \{u_1, \dots, u_n, w_1, \dots, w_q\},$$

entonces B_1 es base de W_1 y B_2 es base de W_2. Probaremos que

$$B := B_1 \cup B_2 = \{u_1, \dots, u_n, v_1, \dots, v_p, w_1, \dots, w_q\}$$

es una base de W_1+W_2. Notar que la proposición 4.4.5 implica que B es un generador de W_1+W_2. Veamos ahora que B es LI. Sean $a_i, b_j, c_k \in \Bbbk$ tales que

$$a_1u_1+\cdots+a_nu_n+b_1v_1+\cdots+b_pv_p+c_1w_1\cdots+c_qw_q=0. \tag{4.6}$$

Luego

$$c_1w_1\cdots+c_qw_q=-(a_1u_1+\cdots+a_nu_n+b_1v_1+\cdots+b_pv_p). \tag{4.7}$$

Como el lado izquierdo de (4.7) está en W_2 y el derecho en W_1, deducimos $c_1w_1\cdots+c_qw_q\in W_1\cap W_2$. Como B_0 es una base de $W_1\cap W_2$, entonces existen escalares d_i tales que

$$c_1w_1\cdots+c_qw_q=d_1u_1+\cdots+d_nu_n \quad\Rightarrow\quad c_1w_1\cdots+c_qw_q+(-d_1)u_1+\cdots+(-d_n)u_n=0.$$

Dado que B_2 es LI, deducimos $c_1=\cdots=c_q=d_1=\cdots=d_n=0$. Sustituyendo estos valores en (4.6) obtenemos

$$a_1u_1+\cdots+a_nu_n+b_1v_1+\cdots+b_pv_p=0.$$

Ahora usando que B_1 es LI deducimos $a_1=\cdots=a_n=b_1=\cdots=b_p=0$. Esto prueba que B es LI y por lo tanto es base de W_1+W_2. Luego

$$\dim(W_1+W_2)=\#B=n+p+q=(n+p)+(n+q)-n=\dim W_1+\dim W_2-\dim(W_1\cap W_2). \quad\square$$

Observación 4.4.8. El resultado anterior se suele usar para calcular la dimensión de la suma, ya que en general es más fácil determinar la intersección que la suma.

Ejemplo 4.4.9. Consideremos el ejemplo 4.4.4. Usando la descripción de W_1 y W_2 dada en (4.4) obtenemos que $B_1=\{(1,0,0,0),\,(0,1,1,1)\}$ es una base de W_1 y $B_2=\{(1,1,0,0),\,(0,0,1,1)\}$ es una base de W_2. Además es claro que $\{(1,1,1,1)\}$ es una base de $W_1\cap W_2$. Entonces aplicando el teorema 4.4.7 obtenemos

$$\dim(W_1+W_2)=\dim W_1+\dim W_2-\dim(W_1\cap W_2)=2+2-1=3.$$

Por otro lado aplicando la proposición 4.4.5 obtenemos que

$$B:=B_1\cup B_2=\{(1,0,0,0),\,(0,1,1,1),\,(1,1,0,0),\,(0,0,1,1)\}$$

es un generador de W_1+W_2. Notar que $\#B=4$ y $\dim(W_1+W_2)=3$, Así que B es LD. Operando obtenemos

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

La igualdad anterior nos dice que a B podemos quitarle uno cualquiera de sus vectores y seguiremos obteniendo un generador, así que por ejemplo

$$C=\{(1,0,0,0),\,(0,1,1,1),\,(1,1,0,0)\}$$

es un generador de W_1+W_2. Como además $\#C=\dim(W_1+W_2)$, concluimos que C es una base de W_1+W_2. Notar que de esta forma hallamos una base de W_1+W_2, sin hallar explícitamente W_1+W_2.

Ejemplo 4.4.10. Sea $V=\mathbb{R}^4$ y consideramos los subespacios

$$W_1=\left\{(x,y,z,t)\in\mathbb{R}^4:\ x=0,\ y+z+t=0\right\} \quad \text{y} \quad W_2=\left\{(x,y,z,t)\in\mathbb{R}^4:\ t=0\right\}.$$

Es fácil de probar que es $\dim W_1=2$ y $\dim W_2=3$. Además

$$W_1\cap W_2=\left\{(x,y,z,t)\in\mathbb{R}^4:\ x=t=0,\ y+z=0\right\}=[(0,1,-1,0)].$$

Luego $\dim(W_1\cap W_2)=1$ y por lo tanto

$$\dim(W_1+W_2)=\dim W_1+\dim W_2-\dim(W_1\cap W_2)=2+3-1=4.$$

Como W_1+W_2 es un subespacio de $\mathbb{R}^4$ y vale $\dim(W_1+W_2)=4=\dim\mathbb{R}^4$, concluimos $W_1+W_2=\mathbb{R}^4$.

Suma directa. Sean W_1, W_2 subespacios de V. Si $V = W_1 + W_2$ y $W_1 \cap W_2 = \{0\}$, entonces se dice que V es *suma directa* de W_1 y W_2 y se escribe $V = W_1 \oplus W_2$.

Teorema 4.4.11. *Si W_1 y W_2 subespacios de V, entonces las siguientes condiciones son equivalentes.*

1. *$V = W_1 \oplus W_2$.*
2. *$V = W_1 + W_2$ y si $w_1 \in W_1$ y $w_2 \in W_2$ verifican $w_1 + w_2 = 0$, entonces $w_1 = w_2 = 0$.*
3. *Para todo $v \in V$ existen únicos $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$.*

Dem. $(2 \Rightarrow 1)$. Lo único que hay que probar es que vale $W_1 \cap W_2 = \{0\}$. Si $v \in W_1 \cap W_2$, entonces vale $v + (-v) = 0$, con $v \in W_1$ y $-v \in W_2$; luego es $v = -v = 0$. Esto implica $W_1 \cap W_2 = \{0\}$.

$(1 \Rightarrow 3)$. Sea $v \in V$. Como es $V = W_1 + W_2$, entonces existen $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$. Supongamos que además existen otros elementos $w_1' \in W_1$ y $w_2' \in W_2$ tales que $v = w_1' + w_2'$. Esto implica

$$w_1 + w_2 = w_1' + w_2' \quad \Rightarrow \quad w_1 - w_1' = w_2' - w_2.$$

Como es $w_1 - w_1' \in W_1$ y $w_2 - w_2' \in W_2$, entonces la igualdad de arriba implica $w_1 - w_1' = w_2' - w_2 \in W_1 \cap W_2 = \{0\}$ y por lo tanto $w_1 - w_1' = w_2 - w_2' = 0$, lo cual implica $w_1 = w_1'$ y $w_2 = w_2'$. Esto prueba la unicidad de la descomposición $v = w_1 + w_2$.

$(3 \Rightarrow 2)$. Es claro que vale $V = W_1 + W_2$. Si $w_1 \in W_1$ y $w_2 \in W_2$ verifican $w_1 + w_2 = 0$, como también vale $0 + 0 = 0$ con $0 \in W_1$ y $0 \in W_2$, entonces la unicidad de la descomposición implica $w_1 = w_2 = 0$. $\square$

Ejemplo 4.4.12. Si consideramos los ejemplos 4.4.3, deducimos del primero que el espacio $\mathbb{R}^3$ es suma del plano Oxy con el plano Oxz, pero esta suma no es directa, y del segundo que el plano Oxy es suma directa de la recta Ox y la recta Oy.

Proposición 4.4.13. *Si $V = W_1 \oplus W_2$, entonces $\dim V = \dim W_1 + \dim W_2$. Además, si B_1 es una base de W_1 y B_2 es una base de W_2, entonces $B_1 \cap B_2 = \emptyset$ y $B_1 \cup B_2$ es una base de V.* $\square$

Dem. La primera afirmación es consecuencia del teorema 4.4.7. Consideremos ahora la intersección de B_1 y B_2. Si existiese $v \in B_1 \cup B_2$, entonces $v \in W_1 \cap W_2 = \{0\}$ y por lo tanto $v = 0$; pero eso es imposible porque $v \in B_1$ y B_1 es LI. Luego $B_1 \cap B_2 = \emptyset$. Sabemos que $B_1 \cup B_2$ es un generador de $V = W_1 + W_2$. Además, como es $B_1 \cap B_2 = \emptyset$, obtenemos

$$\#(B_1 \cup B_2) = \#B_1 + \#B_2 = \dim W_1 + \dim W_2 = \dim V,$$

esto implica que $B_1 \cup B_2$ es una base de V. $\square$

Proposición 4.4.14. *Sean W_1, W_2 subespacios de un espacio de dimensión finita V. Si $W_1 \cap W_2 = \{0\}$ y $\dim V = \dim W_1 + \dim W_2$, entonces $V = W_1 \oplus W_2$.*

Dem. Aplicando el teorema 4.4.7 obtenemos $\dim(W_1 + W_2) = \dim W_1 + \dim W_2$. Luego $\dim(W_1 + W_2) = \dim V$ y por lo tanto $V = W_1 + W_2$. Como además es $W_1 \cap W_2 = \{0\}$, concluimos $V = W_1 \oplus W_2$. $\square$

Ejemplo 4.4.15. 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

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

Luego $W_1 \cap W_2 = \{(0, 0, 0)\}$. Por otro lado W_1 es un plano y W_2 una recta (el eje Ox), 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 4.4.16. 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 $W_1 \cap W_2 = \{0\}$ (recordar el ejemplo 4.4.9).

Operaciones con más de dos subespacios. A continuación generalizaremos lo anterior a una cantidad finita arbitraria de subespacios.

Sean $W_1, \ldots, W_m$ subespacios de V. Definimos su *intersección* $\bigcap_{i=1}^m W_i$ y su *suma* $\sum_{i=1}^m W_i$, mediante

$$\bigcap_{i=1}^{m} W_i := \{v \in V : \ v \in W_i, \ \forall i = 1, \ldots, n\}, \qquad \sum_{i=1}^{m} W_i := \left\{ \sum_{i=1}^{m} w_i : \ w_i \in W_i, \ \forall i = 1, \ldots, n \right\}.$$

La prueba de la siguiente proposición es análoga al caso de dos subespacios y queda como ejercicio.

Proposición 4.4.17. *Si $W_1, \ldots, W_m$ son subespacios de V, entonces $\bigcap_{i=1}^m W_i$ y $\sum_{i=1}^m W_i$ son también subespacios.* □

Decimos que una familia $\{W_1, \ldots, W_m\}$ de subespacios de V es *independiente* si verifica la siguiente condición:

$$\text{si } w_1 \in W_1, \ldots, w_m \in W_m \text{ son tales que } \sum_{i=1}^{m} w_i = 0, \ \text{entonces } w_1 = \cdots = w_m = 0.$$

En lo que sigue escribimos

$$\sum_{j \neq i} W_j := W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_m, \quad \forall i = 1, \ldots, m.$$

Proposición 4.4.18. *Sea $\{W_1, \ldots, W_m\}$ una familia de subespacios de V. Entonces las siguientes condiciones son equivalentes:*

1. *La familia $\{W_1, \ldots, W_m\}$ es independiente.*
2. *Para cada $i = 1, \ldots, m$, vale $W_i \cap \sum_{j \neq i} W_j = \{0\}$.*
3. *Para cada $w \in \sum_{i=1}^m W_i$, existen únicos $w_i \in W_i$, $i = 1, \ldots, m$, tales que $w = \sum_{i=1}^m w_i$.*

Dem. $(1 \Rightarrow 2)$. Sea $i \in \{1, \ldots, m\}$. Si $w \in W_i \cap \sum_{j \neq i} W_j$, entonces $w \in W_i$ y existen $w_j \in W_j$, para todo $j \neq i$ tales que

$$w = w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_m \quad \Rightarrow \quad w_1 + \cdots + w_{i-1} + (-w) + w_{i+1} + \cdots + w_m = 0.$$

Como la familia es independiente, entonces todos los sumandos de la suma anterior son nulos, luego $w = 0$.

$(2 \Rightarrow 3)$. Si un elemento $w \in \sum_{i=1}^m W_i$ se escribe de dos formas $w = \sum_{i=1}^m w_i$ y $w = \sum_{i=1}^m w_i'$, con $w_i, w_i' \in W_i$, para todo $i = 1, \ldots, m$, entonces $\sum_{i=1}^m w_i = \sum_{i=1}^m w_i'$ y por lo tanto para cada $i = 1, \ldots, m$ es

$$w_i - w_i' = \sum_{j \neq i}^{m} (w_j' - w_j)$$

Como $w_i - w_i' \in W_i$ y $\sum_{j \neq i}^m (w_j' - w_j) \in \sum_{j \neq i} W_j$, entonces $w_i - w_i' \in W_i \cap \sum_{j \neq i} W_j = \{0\}$. Luego $w_i = w_i'$, para todo $i = 1, \ldots, m$.

$(3 \Rightarrow 1)$. Sean $w_i \in W_i$, $i = 1, \ldots, m$, tales que $\sum_{i=1}^m w_i = 0$. Como $0 \in W_i$ para todo $i = 1, \ldots, m$ y $\sum_{i=1}^m 0 = 0$, entonces la unicidad implica que es $w_i = 0$, para todo $i = 1, \ldots, m$. Luego la familia $\{W_1, \ldots, W_m\}$ es independiente. □

Observación 4.4.19. Una familia de dos subespacios $\{W_1, W_2\}$ es independiente si y solo si $W_1 \cap W_2 = \{0\}$. Una familia de tres subespacios $\{W_1, W_2, W_3\}$ es independiente si y solo si

$$W_1 \cap (W_2 + W_3) = \{0\}, \quad W_2 \cap (W_1 + W_3) = \{0\}, \quad W_3 \cap (W_1 + W_2) = \{0\}.$$

Notar que las condiciones $W_1 \cap W_2 = W_1 \cap W_3 = W_2 \cap W_3 = \{0\}$ no implican las condiciones de arriba. Por ejemplo, si consideramos $V = \mathbb{R}^2$ y

$$W_1 = \{(x, 0) : \ x \in \mathbb{R}\}, \quad W_2 = \{(0, y) : \ y \in \mathbb{R}\}, \quad W_3 = \{(x, x) : \ x \in \mathbb{R}\},$$

entonces es fácil de probar que las intersecciones dos a dos son triviales (son rectas que pasan por el origen), pero $W_1 + W_2 = \mathbb{R}^2$, luego $W_3 \subset W_1 + W_2$ y por lo tanto $W_3 \cap (W_1 + W_2) = W_3 \neq \{(0, 0)\}$.

Decimos que un subespacio W de V es *suma directa* de una familia de subespacios $\{W_1, \ldots, W_m\}$ y escribimos $W = \bigoplus_{i=1}^m W_i$, si $W = \sum_{i=1}^m W_i$ y la familia $\{W_1, \ldots, W_m\}$ es independiente.

La proposición 4.4.18 implica directamente el siguiente resultado.

Corolario 4.4.20. *Si $W_1, \ldots, W_m \subset V$ son subespacios, entonces las siguientes condiciones son equivalentes.*

1. $V = \bigoplus_{i=1}^m W_i$.
2. $V = \sum_{i=1}^m W_i$ *y* $W_i \cap \sum_{j \neq i} W_j = \{0\}$ *para todo* $i = 1, \ldots, m$.
3. *Para todo $v \in V$, existen únicos $w_i \in W_i$, $i = 1, \ldots, m$, tales que $v = \sum_{i=1}^m w_i$.* □

Observaciones 4.4.21.

1. Notar que el corolario anterior implica que vale $V = W_1 \oplus W_2$ si y solo si $V = W_1 + W_2$ y $W_1 \cap W_2 = \{0\}$; esta fue la definición que dimos para la suma directa de dos subespacios.
2. La notación de suma directa puede ser un poco confusa, porque como conjuntos es $\bigoplus_{i=1}^m W_i = \sum_{i=1}^m W_i$. Lo único que estamos diciendo al escribir $W = \bigoplus_{i=1}^m W_i$ es que en $W = \sum_{i=1}^m W_i$ hay una forma única de escribir cada vector $w \in W$ como $w = \sum_{i=1}^m w_i$, con $w_i \in W_i$, para todo i.

De ahora en adelante supondremos que V es un espacio de dimensión finita. La misma idea que usamos en la demostración de la proposición 4.4.5, sirve para probar el siguiente resultado.

Proposición 4.4.22. *Sean $W_1, \ldots, W_m$ subespacios de V y consideremos $W = \sum_{i=1}^m W_i$. Si G_i es un generador de W_i, para todo $i = 1, \ldots, m$, entonces su unión $\bigcup_{i=1}^m G_i$ es un generador de W.* □

A continuación veremos cómo se relaciona la suma directa con la dimensión.

Proposición 4.4.23. *Sea $\{W_1, \ldots, W_m\}$ una familia independiente de subespacios de V. Consideremos $W = \bigoplus_{i=1}^m W_i$ y sea B_i una base de W_i, para todo i. Entonces*

1. $B_i \cap B_j = \emptyset$, *para todo* $i \neq j$.
2. $B = \bigcup_{i=1}^m B_i$ *es una base de* W.
3. $\dim W = \sum_{i=1}^m \dim W_i$.

Dem. Si existiesen $i \neq j$ tales que $B_i \cap B_j \neq \emptyset$, entonces existiría $v \in B_i \cap B_j$. Como B_i es una base, entonces $v \neq 0$; luego

$$v \in W_i \cap W_j \subset W_i \cap \sum_{l \neq i} W_l \quad \Rightarrow \quad W_i \cap \sum_{l \neq i} W_l \neq \{0\} \quad \lightning.$$

Luego $B_i \cap B_j = \emptyset$, para todo $i \neq j$.

Probaremos ahora que B es una base de W. Por la proposición anterior sabemos que B es un generador de W, así que solo resta probar que es LI. Para cada $i = 1, \dots, m$, sea $B_i = \{w_1^i, \dots, w_{l_i}^i\}$. Supongamos que tenemos escalares a_j^i tales que

$$a_1^1 w_1^1 + \cdots + a_{l_1}^1 w_{l_1}^1 + \cdots + a_1^m w_1^m + \cdots + a_{l_m}^m w_{l_m}^m = 0.$$

Como $a_1^i w_1^i + \cdots + a_{l_i}^i w_{l_i}^i \in W_i$ (para todo i) y la familia $\{W_1, \dots, W_m\}$ es independiente, deducimos que es

$$a_1^i w_1^i + \cdots + a_{l_i}^i w_{l_i}^i = 0, \quad \forall i = 1, \dots, m.$$

Luego como cada B_i es LI, deducimos que es $a_j^i = 0$, para todo i, j y por lo tanto B es LI. La tercera afirmación se deduce directamente de las anteriores. $\square$

Corolario 4.4.24. *Sean $W_1, \dots, W_m$ subespacios de V. Si $\{W_1, \dots, W_m\}$ es una familia independiente y $\dim V = \sum_{i=1}^m \dim W_i$, entonces $V = \bigoplus_{i=1}^m W_i$.*

Dem. La proposición anterior implica $\dim(\bigoplus_{i=1}^m W_i) = \sum_{i=1}^m \dim W_i = \dim V$, luego $\bigoplus_{i=1}^m W_i = V$. $\square$

4.5. Dimensión infinita.

Lo que veremos en esta sección es complementario y no se necesita para entender el resto.

Las definiciones que vimos anteriormente de conjuntos LD o LI, generadores y bases, son para espacios de dimensión finita. A continuación veremos las definiciones generales. En esta sección V es un espacio vectorial arbitrario.

Si $X \subset V$ es un subconjunto no vacío (posiblemente infinito), definimos el *subespacio generado* por X como el conjunto de todas las combinaciones lineales (finitas) que podemos lograr con elementos de X, es decir

$$[X] := \{v \in V : \exists n \geq 1,\ v_1, \dots, v_n \in X,\ a_1, \dots, a_n \in \Bbbk,\ \text{tales que } v = a_1 v_1 + \cdots + a_n v_n\}.$$

Notar que si $W \subset V$ es un subespacio tal que $X \subset W$, entonces $[X] \subset W$. Luego el subespacio generado por X es el menor subespacio de V que contiene a X. Como antes definimos $[\emptyset] = \{0\}$.

Decimos que un conjunto $X \subset V$ es *linealmente dependiente* (LD) si existe una cantidad finita de elementos $v_1, \dots, v_n \in X$ y escalares no todos nulos $a_1, \dots, a_n \in \Bbbk$ tales que $a_1 v_1 + \cdots + a_n v_n = 0$. En caso contrario se dice que X es *linealmente independiente* (LI). Notar que X es LD si y solo si X contiene un subconjunto finito que es LD y X es LI si y solo si todos sus subconjuntos finitos son LI. En particular esta definición implica que el conjunto vacío es LI.

Decimos que $A \subset V$ es un *generador* de V si el subespacio generado por A es todo V y decimos que A es una *base* de V si es un generador y es LI. Notar que esto implica que $\emptyset$ es una base del espacio trivial $\{0\}$.

Teorema 4.5.1. *Todo espacio vectorial admite una base.*

Dem. Sea V un espacio no trivial. Consideremos $\mathcal{A} = \{X \subset V : X \text{ es LI}\}$, pensado como conjunto parcialmente ordenado con la inclusión. Sea $\mathcal{C}$ una *cadena* en $\mathcal{A}$ (es decir, $\mathcal{C}$ es un subconjunto no vacío de $\mathcal{A}$ tal que si $X, Y \in \mathcal{C}$, entonces $X \subset Y$ o $Y \subset X$). Consideremos $M = \bigcup_{X \in \mathcal{C}} X$. Si $Z \subset M$ es un subconjunto finito, al ser $\mathcal{C}$ una cadena obtenemos que existe $X \in \mathcal{C}$ tal que $Z \subset X$. Como X es LI, entonces Z también es LI. Luego todo subconjunto finito de M es LI y por lo tanto M es LI. Entonces $M \in \mathcal{A}$ y por lo tanto M es una cota superior de $\mathcal{C}$. Luego el lema de Zorn implica que $\mathcal{A}$ tiene un elemento maximal B. Entonces probamos que en V existe un conjunto B que es un LI maximal, y es fácil de probar que esto implica que B es una base de V. $\square$

Observación 4.5.2. En general todo lo que vimos en dimensión finita para conjuntos LD y LI, generadores y bases, vale también en dimensión infinita (eventualmente acomodando un poco los enunciados). Sin embargo en dimensión infinita las bases tienen menos utilidad que en dimensión finita, ya que en general no es posible determinar una base de un espacio de dimensión infinita, aun cuando en teoría sabemos que existe (ver los ejemplos siguientes). También perdemos los resultados que involucran argumentos de conteo, como por ejemplo la proposición 4.3.18.

El siguiente resultado lo usaremos en los ejemplos siguientes. Recordar que si V es un espacio de dimensión finita n, entonces todo subconjunto de V con más de n elementos es LD (proposición 4.3.17). Luego si en V existe un subconjunto LI infinito, entonces la dimensión de V es infinita.

Ejemplo 4.5.3. Probaremos que el conjunto $B = \{x^n : n = 0, 1, 2, \dots\}$ es una base del espacio de los polinomios $\Bbbk[x]$ y por lo tanto $\Bbbk[x]$ tiene dimensión infinita. Es claro que B es generador, así que solo falta probar que es LI. Para cada $n = 1, 2, \dots$, consideremos el conjunto $B_n = \{1, x, \dots, x^n\}$. Si $a_0, \dots, a_n \in \Bbbk$ son tales que $a_0 1 + a_1 x + \cdots + a_n x^n = 0$ (el polinomio nulo), entonces $a_0 = \cdots = a_n = 0$. Esto prueba que B_n es LI, para todo n. Si C es un subconjunto finito de B, entonces existe algún n tal que $C \subset B_n$ y por lo tanto C es LI (por la proposición 4.3.6). Luego todo subconjunto finito de B es LI y por lo tanto B es LI.

Ejemplos 4.5.4.

1. Pensando al espacio de los polinomios $\mathbb{R}[x]$ como el espacio de las sucesiones reales finitas, entonces el espacio V de las sucesiones reales contiene a $\mathbb{R}[x]$ como subespacio y por lo tanto V tiene dimensión infinita.

2. Si consideramos el espacio $D(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \text{ es derivable}\}$, entonces $\mathbb{R}[x]$ es un subespacio de $D(\mathbb{R})$ (identificando los polinomios con las funciones polinómicas) y por lo tanto $D(\mathbb{R})$ tiene dimensión infinita. Por lo mismo $C(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \text{ es continua}\}$ también tiene dimensión infinita.

Respecto a estos ejemplos, salvo en los polinomios, en los otros casos no se conocen bases de esos espacios.

En espacios de dimensión infnita tiene sentido considerar sumas de familias infinitas de subespacios. Si $\{W_\alpha : \alpha \in \Lambda\}$ es una familia de subespacios de V (posiblemente infinita), entonces su *intersección* es el conjunto

$$\bigcap_{\alpha\in\Lambda} W_\alpha := \{v \in V : v \in W_\alpha, \ \forall \alpha \in \Lambda\}.$$

Es fácil de probar que $\bigcap_{\alpha\in\Lambda} W_\alpha$ es un subespacio de V. En este caso se define $\sum_{\alpha\in\Lambda} W_\alpha$ como el conjunto de todas las sumas (finitas) de vectores que están en $\bigcup_{\alpha\in\Lambda} W_\alpha$. Notar que $\sum_{\alpha\in\Lambda} W_\alpha$ coincide con el subespacio generado por $\bigcup_{\alpha\in\Lambda} W_\alpha$, luego $\sum_{\alpha\in\Lambda} W_\alpha$ también es un subespacio de V. La definición de suma directa también se extiende a familias arbitrarias, en forma análoga a como definimos la suma directa de una familia finita de subespacios. Por ejemplo, con esa definición vale $\mathbb{R}[x] = \bigoplus_{n=0}^{\infty} \mathbb{R}_n[x]$.

Capítulo 5

Transformaciones lineales

En este capítulo estudiaremos las funciones entre espacios vectoriales que preservan la estructura lineal. Dado que los resultados más interesantes se obtienen en espacios de dimensión finita, en este capítulo asumiremos siempre que estamos en esa situación, aunque las definiciones y algunos de los resultados valen en general.

5.1. Definiciones y propiedades básicas

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 $a \in \Bbbk$ y $v \in V$.

La siguiente proposición da una forma más rápida para chequear si una función es una transformación lineal.

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

$$T(au+v) = aT(u) + T(v), \quad \forall u, v \in V,\ a \in \Bbbk. \tag{5.1}$$

Dem. Si T es una transformación lineal, entonces

$$T(au+v) = T(au) + T(v) = aT(u) + T(v), \quad \forall u, v \in V,\ a \in \Bbbk.$$

Recíprocamente, si T verifica (5.1), entonces tomando $a = 1$ en (5.1), obtenemos

$$T(u+v) = T(u) + T(v), \quad \forall u, v \in V.$$

Tomando $u = v = 0$ en esta última igualdad, obtenemos $T(0) = T(0) + T(0)$, lo cual implica $T(0) = 0$. Ahora tomando $v = 0$ en (5.1), obtenemos

$$T(au) = aT(u) + T(0) = aT(u) \quad \Rightarrow \quad T(au) = aT(u), \quad \forall u \in V,\ a \in \Bbbk. \quad \square$$

Ejemplos 5.1.2. Ejemplos de transformaciones lineales.

1. La función *nula* $T : V \to W$ definida por $T(v) = 0$, para todo $v \in V$ (V y W espacios arbitrarios).
2. La función identidad $\mathrm{Id} : V \to V$ definida por $\mathrm{Id}(v) = v$, para todo $v \in V$ (V espacio arbitrario).
3. La función $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x,y) = (x, -y)$, para todo $(x,y) \in \mathbb{R}^2$. Notar que T es la simetría axial respecto al eje Ox.

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

Proposición 5.1.3. *Sea $T : V \to W$ una transformación lineal.*

1. *Si $a_1, \dots, a_n \in \mathbb{k}$ y $v_1, \dots, v_n \in V$ $(n = 1, 2, \dots)$, entonces*
$$T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n).$$
2. *$T(0) = 0$ (la imagen por T del vector nulo de V es el vector nulo de W).*
3. *$T(-v) = -T(v)$, para todo $v \in V$.*
4. *$T(u - v) = T(u) - T(v)$, para todo $u, v \in V$.*

Dem. Ejercicio (la segunda afimación ya fue probada en la demostración de la proposición 5.1.1). □

En lo que sigue veremos cómo son las transformaciones lineales de $\mathbb{k}^n$ en $\mathbb{k}^m$.

Proposición 5.1.4. *Si $T : \mathbb{k}^n \to \mathbb{k}^m$ es una transformación lineal, entonces existen únicos escalares $a_{ij} \in \mathbb{k}$ tales que*
$$T(x_1, \dots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n,\ \dots\ ,\ a_{m1}x_1 + \cdots + a_{mn}x_n), \tag{5.2}$$
para todo $(x_1, \dots, x_n) \in \mathbb{k}^n$.

Dem. Si existen $a_{ij} \in \mathbb{k}$ que verifican (5.2), entonces calculando las imágenes por T de los vectores de la base canónica $B = \{e_1, \dots, e_n\}$, obtenemos
$$T(e_1) = (a_{11}, \dots, a_{m1}),\ \dots\ ,\ T(e_n) = (a_{1n}, \dots, a_{mn}). \tag{5.3}$$
Esto prueba la unicidad de los a_{ij}. Recíprocamente, si definimos los a_{ij} por las fórmulas (5.3), entonces
$$\begin{aligned} T(x_1, \dots, x_n) &= T(x_1e_1 + \cdots + x_ne_n) \\ &= x_1T(e_1) + \cdots + x_nT(e_n) \\ &= x_1(a_{11}, \dots, a_{m1}) + \cdots + x_n(a_{1n}, \dots, a_{mn}) \\ &= (a_{11}x_1 + \cdots + a_{1n}x_n,\ \dots,\ a_{m1}x_1 + \cdots + a_{mn}x_n). \quad \square \end{aligned}$$

Observación 5.1.5. Si escribimos los vectores en forma vertical, entonces la fórmula (5.2) queda
$$T\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} \quad \Rightarrow \quad T\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

En general, dada una matriz $A \in M_{m\times n}$, definimos una función $L_A : \mathbb{k}^n \to \mathbb{k}^m$ mediante $L_A(v) = Av$ para todo $v \in \mathbb{k}^n$. Notar que $L_A : \mathbb{k}^n \to \mathbb{k}^m$ es una transformación lineal:
$$L_A(au + v) = A(au + v) = aAu + Av = aL_A(u) + L_A(v), \quad \forall a \in \mathbb{k},\ uv \in V.$$

Con esta notación, la proposición 5.1.4 nos dice que si $T : \mathbb{k}^n \to \mathbb{k}^m$ es una transformación lineal, entonces existe una única matriz $A \in M_{m\times n}$ tal que $T = L_A$. Además, las columnas de A son los vectores $T(e_1), \dots, T(e_n)$, escritos en forma vertical, siendo $\{e_1, \dots, e_n\}$ la base canónica de $\mathbb{k}^n$.

Ejemplo 5.1.6. Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T(x, y, z) = (x + 2y + 3z,\ x - y)$ para todo $(x, y, z) \in \mathbb{R}^3$. Aplicando la proposición anterior sabemos que T es una transformación lineal. Entonces es $T = L_A$ para cierta matriz $A \in M_{2\times 3}$. Para hallar A podemos usar las fórmulas (5.3) obteniendo $T(1, 0, 0) = (1, 1)$, $T(0, 1, 0) = (2, -1)$ y $T(0, 0, 1) = (3, 0)$, luego $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 0 \end{pmatrix}$.

A continuación veremos que toda transformación lineal queda determinada por sus valores en una base.

Proposición 5.1.7. *Sean V y W dos espacios vectoriales. Si $B = \{v_1, \dots, v_n\}$ es una base de V y $w_1, \dots, w_n$ son vectores arbitrarios en W, entonces existe una única transformación lineal $T : V \to W$ tal que $T(v_i) = w_i$, para todo $i = 1, \dots, n$. La misma está definida por*

$$T(x_1v_1 + \cdots + x_nv_n) = x_1w_1 + \cdots + x_nw_n, \tag{5.4}$$

para todo $x_1, \dots, x_n \in \Bbbk$.

Dem. Si existe una transformación lineal $T : V \to W$ que verifique $T(v_i) = w_i$, para todo i, entonces para todo $x_1, \dots, x_n \in \Bbbk$ 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 que T queda definida por la fórmula (5.4) y por lo tanto es única. Para probar la existencia, como B es base de V, entonces todo vector de V se escribe en forma única como combinación lineal de $v_1, \dots, v_n$. Luego tiene sentido definir una función $T : V \to W$ mediante la fórmula (5.4). Probaremos que T definida de esa manera es una transformación lineal. Sean $v, w \in V$ y $a \in \Bbbk$. Como B es base de V, entonces existen escalares $x_i, y_i \in \Bbbk$, $i = 1, \dots, n$, tales que $v = \sum_{i=1}^n x_iv_i$ y $w = \sum_{i=1}^n y_iv_i$. Entonces

$$\begin{aligned} T(av + w) &= T\left(a\sum_{i=1}^n x_iv_i + \sum_{i=1}^n y_iv_i\right) = T\left(\sum_{i=1}^n (ax_i + y_i)v_i\right) = \sum_{i=1}^n (ax_i + y_i)w_i \\ &= a\sum_{i=1}^n x_iw_i + \sum_{i=1}^n y_iw_i = aT(v) + T(w). \quad \square \end{aligned}$$

Ejemplo 5.1.8. Buscamos encontrar una transformación lineal $T : \mathbb{R}^2 \to M_2(\mathbb{R})$ tal que $T(1, 1) = \left(\begin{smallmatrix} 2 & 3 \\ 3 & 0 \end{smallmatrix}\right)$ y $T(1, -1) = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 2 \end{smallmatrix}\right)$. Como $B = \{(1, 1), (1, -1)\}$ es una base de $\mathbb{R}^2$, entonces podemos aplicar la proposición 5.1.7. Para encontrar T tenemos que ver cómo escribir un vector arbitrario (x, y) de $\mathbb{R}^2$ como combinación lineal de la base B. En este caso es $(x, y) = \frac{x+y}{2}(1, 1) + \frac{x-y}{2}(1, -1)$, luego

$$\begin{aligned} T(x, y) &= T\left(\frac{x + y}{2}(1, 1) + \frac{x - y}{2}(1, -1)\right) = \frac{x + y}{2}T(1, 1) + \frac{x - y}{2}T(1, -1) \\ &= \frac{x + y}{2}\begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix} + \frac{x - y}{2}\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} x + y & x + 2y \\ 2x + y & x - y \end{pmatrix}. \end{aligned}$$

Entonces $T : \mathbb{R}^2 \to M_2(\mathbb{R})$ está definida por $T(x, y) = \left(\begin{smallmatrix} x+y & x+2y \\ 2x+y & x-y \end{smallmatrix}\right)$, para todo $(x, y) \in \mathbb{R}^2$.

Observación 5.1.9. Dados $v_1, \dots, v_n \in V$ y $w_1, \dots, w_n \in W$, no siempre existe una transformación lineal $T : V \to W$ que verifique $T(v_i) = w_i$, para todo $i = 1, \dots, n$. Por ejemplo, si existiese una transformación lineal $T : \mathbb{R}^2 \to M_2(\mathbb{R})$ tal que $T(1, 0) = (2, 3, 3)$ y $T(2, 0) = (0, -1, 1)$, entonces valdría

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

Por otro lado, si buscamos una transformación lineal $T : \mathbb{R}^2 \to M_2(\mathbb{R})$ tal que $T(1, 1) = \left(\begin{smallmatrix} 2 & 3 \\ 3 & 0 \end{smallmatrix}\right)$, entonces existen infinitas transformaciones lineales que lo verifican. Alcanza con completar el vector $(1, 1)$ a una base de $\mathbb{R}^2$ y aplicar la proposición 5.1.7 (el ejemplo 5.1.8 es un caso particular de esta construcción).

Operaciones con transformaciones lineales.

Proposición 5.1.10. *Si $T : V \to W$ y $S : W \to U$ son transformaciones lineales, entonces la función compuesta $S \circ T : V \to U$ es una transformación lineal.*

Dem. Sean $u, v \in V$ y $a \in \Bbbk$. Entonces

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

Recordar que si V es un espacio vectorial y D es un conjunto, entonces el conjunto $\mathcal{F}(D, V)$ formado por las funciones de D en V, es un espacio vectorial con las operaciones punto a punto.

Sean V y W espacios vectoriales y $\mathcal{L}(V, W)$ el conjunto de transformaciones lineales de V en W. El conjunto $\mathcal{L}(V, W)$ está contenido en $\mathcal{F}(V, W)$, el cual por lo que vimos recién es un espacio vectorial. Luego dadas $S, T : V \to W$ transformaciones lineales y $c \in \Bbbk$, entonces tenemos definida su suma $S + T : V \to W$ y el producto por escalar $c \cdot T : V \to W$ mediante

$$(S + T)(v) := S(v) + T(v); \quad (c \cdot T)(v) := c \cdot T(v), \quad \forall v \in V. \tag{5.5}$$

En general escribiremos cT en vez de $c \cdot T$, aunque en ciertas pruebas usaremos la notación con el punto porque es más explícita.

Proposición 5.1.11. *Si V y W son espacios vectoriales, entonces $\mathcal{L}(V, W)$ con las operaciones definidas en* (5.5) *es un subespacio de $\mathcal{F}(V, W)$ y por lo tanto $\mathcal{L}(V, W)$ es un espacio vectorial.*

Dem. Empezamos recordando que la función nula $0 : V \to W$ es una transformación lineal y por lo tanto está en $\mathcal{L}(V, W)$. Probaremos ahora que si $S, T : V \to W$ son transformaciones lineales y $c \in \Bbbk$, entonces $S + T : V \to W$ y $c \cdot T : V \to W$ son también transformaciones lineales. Sean $u, v \in V$ y $a \in \Bbbk$. Entonces

$$\begin{aligned}
(S + T)(a \cdot u + v) &= S(a \cdot u + v) + T(a \cdot u + v) = a \cdot S(u) + S(v) + a \cdot T(u) + T(v) \\
&= a \cdot \big(S(u) + T(u)\big) + S(v) + T(v) = a \cdot (S + T)(u) + (S + T)(v). \\
(c \cdot T)(a \cdot u + v) &= c \cdot T(a \cdot u + v) = c \cdot \big(a \cdot T(u) + T(v)\big) = c \cdot \big(a \cdot T(u)\big) + c \cdot T(v) = ca \cdot T(u) + c \cdot T(v) \\
&= ac \cdot T(u) + c \cdot T(v) = a \cdot \big(c \cdot T(u)\big) + c \cdot T(v) = a \cdot (c \cdot T)(u) + (c \cdot T)(v).
\end{aligned}$$

Luego $S + T \in \mathcal{L}(V, W)$ y $c \cdot T \in \mathcal{L}(V, W)$. $\square$

La siguiente proposición describe cómo se relaciona la composición con la suma y el producto por escalares.

Proposición 5.1.12. *Si $S, S_1, S_2 : U \to V$ y $T, T_1, T_2 : V \to W$ son transformaciones lineales y $a \in \Bbbk$, entonces*

$$T \circ (S_1 + S_2) = T \circ S_1 + T \circ S_2 \quad y \quad (T_1 + T_2) \circ S = T_1 \circ S + T_2 \circ S,$$
$$T \circ (aS) = (aT) \circ S = a(T \circ S).$$

Dem. Sea $v \in V$. Entonces

$$\begin{aligned}
\big(T \circ (S_1 + S_2)\big)(v) &= T\big((S_1 + S_2)(v)\big) = T\big(S_1(v) + S_2(v)\big) = T\big(S_1(v)\big) + T\big(S_2(v)\big) \\
&= (T \circ S_1)(v) + (T \circ S_2)(v) = (T \circ S_1 + T \circ S_2)(v). \\
\big((T_1 + T_2) \circ S\big)(v) &= (T_1 + T_2)\big(S(v)\big) = T_1\big(S(v)\big) + T_2\big(S(v)\big) = (T_1 \circ S)(v) + (T_2 \circ S)(v) \\
&= (T_1 \circ S + T_2 \circ S)(v) \\
\big(T \circ (a \cdot S)\big)(v) &= T\big((a \cdot S)(v)\big) = T\big(a \cdot S(v)\big) = a \cdot T\big(S(v)\big) = a \cdot (T \circ S)(v) = \big(a \cdot (T \circ S)\big)(v). \\
\big((a \cdot T) \circ S\big)(v) &= (a \cdot T)\big(S(v)\big) = a \cdot T\big(S(v)\big) = a \cdot (T \circ S)(v) = \big(a \cdot (T \circ S)\big)(v). \quad \square
\end{aligned}$$

A continuación veremos cómo se comportan las transformaciones lineales de $\Bbbk^n$ en $\Bbbk^m$ respecto a las operaciones anteriores.

Proposición 5.1.13. *1. Si $A, B \in M_{m\times n}$ y $c \in \Bbbk$, entonces $L_{A+B} = L_A + L_B$ y $L_{cA} = cL_A$.*

2. Si $A \in M_{m\times n}$ y $B \in M_{n\times p}$, entonces $L_A \circ L_B = L_{AB}$.

Dem. Sea $v \in \Bbbk^n$. Entonces

$$\begin{aligned} L_{A+B}(v) &= (A+B)v = Av + Bv = L_A(v) + L_B(v) = (L_A + L_A)(v), \\ L_{cA}(v) &= (cA)v = c(Av) = c\big(L_A(v)\big) = (cL_A)(v), \\ (L_A \circ L_B)(v) &= L_A\big(L_B(v)\big) = A(Bv) = (AB)v = L_{AB}(v). \quad \square \end{aligned}$$

5.2. Inyectividad y sobreyectividad.

Sea $T : V \to W$ una transformación lineal. Definimos

$$\operatorname{Ker}(T) := \{v \in V : \ T(v) = 0\}, \qquad \operatorname{Im}(T) := \{w \in W : \ \exists v \in V \text{ tal que } w = T(v)\}.$$

El conjunto $\operatorname{Ker}(T)$ es el *núcleo*[1] de T y el conjunto $\operatorname{Im}(T)$ es la *imagen* o *recorrido* de T.

Proposición 5.2.1. *Si $T : V \to W$ es una transformación lineal, entonces $\operatorname{Ker}(T)$ es un subespacio de V y $\operatorname{Im}(T)$ es un subespacio de W.*

Dem. Para que quede más claro, vamos a escribir 0_V y 0_W a los vectores nulos de V y W, respectivamente. Como T es lineal, entonces verifica $T(0_V) = 0_W$; luego $0_V \in \operatorname{Ker}(T)$ y $0_W \in \operatorname{Im}(T)$.
Sean $v_1, v_2 \in \operatorname{Ker}(T)$ y $a \in \Bbbk$, entonces

$$T(av_1 + v_2) = aT(v_1) + T(v_2) = a0_V + 0_V = 0_V \quad \Rightarrow \quad av_1 + v_2 \in \operatorname{Ker}(T).$$

Sean $w_1, w_2 \in \operatorname{Im}(T)$ y $a \in \Bbbk$, entonces existen $v_1, v_2 \in V$ tales que $w_1 = T(v_1)$ y $w_2 = T(v_2)$. Luego

$$aw_1 + w_2 = aT(v_1) + T(v_2) = T(av_1 + v_2) \quad \Rightarrow \quad aw_1 + w_2 \in \operatorname{Im}(T).$$

Esto termina la prueba de que $\operatorname{Ker}(T) \subset V$ y $\operatorname{Im}(T) \subset W$ son subespacios. $\square$

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

Proposición 5.2.2. *Si $T : V \to W$ es una transformación lineal, entonces T es inyectiva si y solo si $\operatorname{Ker}(T) = \{0\}$.*

Dem. Si T es inyectiva y $v \in V$ es tal que $v \neq 0$, entonces $T(v) \neq T(0) = 0$; luego $\operatorname{Ker}(T) = \{0\}$. Recíprocamente, si $\operatorname{Ker}(T) = \{0\}$ y $u, v \in V$ son tales que $T(u) = T(v)$, entonces $T(u-v) = T(u) - T(v) = 0$. Luego $u - v \in \operatorname{Ker}(T) = \{0\}$ y por lo tanto $u - v = 0$, es decir $u = v$. $\square$

Observación 5.2.3. Dado que una transformación lineal $T : V \to W$ es sobreyectiva si y solo si $\operatorname{Im}(T) = W$, entonces conociendo su núcleo e imagen, podemos saber si es inyectiva, sobreyectiva o biyectiva.

Observación 5.2.4. Sea $T : V \to W$ una transformación lineal. Si $v_1, \dots, v_n \in V$ y $a_1, \dots, a_n \in \Bbbk$, entonces

$$a_1v_1 + \cdots + a_nv_n = 0 \quad \Rightarrow \quad a_1T(v_1) + \cdots + a_nT(v_n) = 0.$$

Esto implica que si A es un subconjunto LD de V, entonces $T(A)$ es un subconjunto LD de W. Por otro lado, si consideramos $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (x - y, y - x)$ y la base canónica $B = \{(1, 0), (0, 1)\}$, entonces $T(B) = \{(1, -1), (-1, 1)\}$. Notar que $T(B)$ no es un conjunto LI ni es un generador de $\mathbb{R}^2$.

Luego las transformaciones lineales llevan conjuntos LD en conjuntos LD, pero en general no llevan conjuntos LI en conjuntos LI, ni generadores en generadores. Por esa razón es que tiene interés la siguiente proposición.

[1] La abreviación Ker viene de *kernel*, que es un término que se suele usar para denominar el núcleo.

Proposición 5.2.5. *Sea $T : V \to W$ una transformación lineal.*

1. *Si T es inyectiva y A es un subconjunto LI de V, entonces $T(A)$ es un subconjunto LI de W.*

2. *Si A es un generador de V, entonces $T(A)$ es un generador de $\operatorname{Im}(T)$. Luego si T es sobreyectiva y A es un generador de V, entonces $T(A)$ es un generador de W.*

3. *Si T es biyectiva y A es una base de V, entonces $T(A)$ es una base de W.*

Dem. La tercera afirmación se deduce de las otras, así que solo hay que probar las dos primeras. Por simplicidad haremos la prueba cuando el conjunto A es finito, pero el resultado vale en general.

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

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

Como A es LI, se concluye que es $a_1 = \dots = a_n = 0$. Luego $T(A)$ es LI.

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

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

Esto muestra que todo elemento de $\operatorname{Im}(T)$ es combinación lineal de $T(A) = \{T(v_1), \dots, T(v_n)\}$. $\square$

Observación 5.2.6. Si $T : V \to W$ es una transformación lineal arbitraria y B es una base de V, entonces $T(B)$ es un generador de $\operatorname{Im}(T)$, pero no es necesariamente una base (ver el ejemplo siguiente).

Ejemplo 5.2.7. Consideremos la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$T(x, y, z) = (x + y + z,\ x + y - z,\ 2x + 2y + 2z,\ -x - y + z). \tag{5.6}$$

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

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

Luego $\operatorname{Ker}(T) = \{(x, -x, 0) :\ x \in \mathbb{R}\} = [(1, -1, 0)]$ y es claro que $\{(1, -1, 0)\}$ es base de $\operatorname{Ker}(T)$.
Para saber si un vector $(a, b, c, d) \in \mathbb{R}^4$ está en la imagen de T, tenemos que averiguar si existe un vector $(x, y, z) \in \mathbb{R}^3$ tal que $(a, b, c, d) = T(x, y, z)$, es decir si el sistema

$$\left\{\begin{array}{rcl} x + y + z & = & a \\ x + y - z & = & b \\ 2x + 2y + 2z & = & c \\ -x - y + z & = & d \end{array}\right. \tag{5.7}$$

es compatible. Si a la tercera ecuación le sumamos la primera multiplicada por -2 y a la cuarta le sumamos la segunda, obtenemos que el sistema anterior es equivalente al siguiente

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

Luego para que el sistema tenga solución, necesariamente debe ser $c = 2a$ y $d = -b$, y es claro que si esto sucede, entonces de las dos primeras ecuaciones siempre podemos despejar x, y, z (no en forma única). Luego el sistema (5.7) es compatible si y solo si $c = 2a$ y $d = -b$, por lo tanto la imagen de T es

$$\operatorname{Im}(T) = \left\{(a, b, c, d) \in \mathbb{R}^4 : \ c = 2a, \ d = -b\right\}. \tag{5.8}$$

Notar que determinar $\operatorname{Im}(T)$ fue bastante más complicado que determinar $\operatorname{Ker}(T)$. Una alternativa es aplicar la proposición 5.2.5. Si consideramos $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ la base canónica de $\mathbb{R}^3$, entonces

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

Luego $T(B) = \{(1, 1, 2, -1), (1, -1, 2, 1)\}$ es un generador de $\operatorname{Im}(T)$ y claramente es LI. Así que $T(B)$ es una base de $\operatorname{Im}(T)$ y por lo tanto

$$\operatorname{Im}(T) = [(1, 1, 2, -1), (1, -1, 2, 1)]. \tag{5.9}$$

Es un ejercicio el verificar que las dos formas de describir $\operatorname{Im}(T)$ dadas por (5.8) y (5.9), coinciden.

El siguiente resultado muestra que el núcleo y la imagen de una transformación lineal están relacionados.

Teorema 5.2.8. *Si $T : V \to W$ es una transformación lineal, entonces*

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

Dem. Sea $B_1 = \{v_1, \dots, v_n\}$ una base de $\operatorname{Ker}(T)$. Como $\{v_1, \dots, v_n\} \subset V$ es LI, entonces existen $w_1, \dots, w_m$ en V tales que $B_2 = \{v_1, \dots, v_n, w_1, \dots, w_m\}$ es una base de V. Luego

$$T(B_2) = \{T(v_1), \dots, T(v_n), T(w_1), \dots, T(w_m)\}$$

es un generador de $\operatorname{Im}(T)$. Pero $T(v_1) = \cdots = T(v_n) = 0$, así que $B_3 = \{T(w_1), \dots, T(w_m)\}$ también es un generador de $\operatorname{Im}(T)$. Veamos que B_3 es LI. Sean $a_1 \dots, a_m \in \mathbb{k}$ tales que $a_1 T(w_1) + \cdots + a_m T(w_m) = 0$. Luego

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

Entonces existen $b_1 \dots, b_n \in \mathbb{k}$ tales que

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

Como B_2 es base de V, deducimos que es $b_1 = \cdots = b_n = a_1 = \cdots = a_m = 0$. Luego B_3 es LI y por lo tanto es base de $\operatorname{Im}(T)$. Entonces

$$\dim \operatorname{Ker}(T) + \dim \operatorname{Im}(T) = \#B_1 + \#B_3 = n + m = \#B_2 = \dim V. \quad \square$$

Ejemplo 5.2.9. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

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

Razonando como en el ejemplo 5.2.7 se obtiene que $\{(-1, -1, 1)\}$ es una base de $\operatorname{Ker}(T)$ y que el conjunto $G = \{(1, 1, 0), (1, 0, 1), (2, 1, 1)\}$ es un generador de $\operatorname{Im}(T)$. Luego $\dim \operatorname{Ker}(T) = 1$ y como $\dim \mathbb{R}^3 = 3$, entonces el teorema anterior implica $\dim \operatorname{Im}(T) = 2$. Como el generador G de $\operatorname{Im}(T)$ tiene tres elementos, para obtener una base de $\operatorname{Im}(T)$ alcanza con tomar dos elementos de G que sean LI; en este caso cualquier par de esos elementos lo verifica, luego

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

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

Isomorfismos. A las transformaciones lineales biyectivas se les llama *isomorfismos*. Si existe un isomorfismo entre dos espacios vectoriales, entonces se dice que los espacios son *isomorfos*. Si dos espacios son isomorfos, entonces existe una correspondencia uno a uno entre sus vectores y esta correspondencia respeta las operaciones. Luego los espacios isomorfos tienen las mismas propiedades como espacios vectoriales.

Proposición 5.2.10. *Si $T : V \to W$ es un isomorfismo, entonces su función inversa $T^{-1} : W \to V$ es también una transformación lineal y por lo tanto 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 $a \in \Bbbk$. Consideremos $v_1 = T^{-1}(w_1)$ y $v_2 = T^{-1}(w_2)$. Como T es lineal, es

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

Dados dos espacios V y W, escribiremos $V \simeq W$ para indicar que V y W son isomorfos. Notar que la relación entre espacios de "ser isomorfos" es de equivalencia, es decir, verifica

$$V \simeq V; \quad V \simeq W \ \Rightarrow \ W \simeq V; \quad V \simeq W \text{ y } W \simeq U \ \Rightarrow \ V \simeq U.$$

La primera propiedad se debe a que la función identidad es un isomorfismo, la segunda a la proposición anterior y la tercera a que claramente la composición de isomorfismos da un isomorfismo.

Proposición 5.2.11. *Si dos espacios son isomorfos, entonces tienen la misma dimensión.*

Dem. Si existe un isomorfismo $T : V \to W$ y B es una base de V, entonces $T(B)$ es una base de W (proposición 5.2.5) y por lo tanto $\dim W = \#T(B) = \#B = \dim V$. $\square$

El siguiente resultado es interesante, porque la inyectividad y la sobreyectividad no suelen estar relacionadas.

Teorema 5.2.12. *Si $T : V \to W$ es 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 la tercera afirmación implica las dos primeras. Así que solo hay que probar que T es inyectiva si y solo si es sobreyectiva.

Recordar que en el teorema 5.2.8 probamos que vale la siguiente fórmula

$$\dim \operatorname{Ker}(T) + \dim \operatorname{Im}(T) = \dim V. \tag{5.10}$$

Si T es inyectiva, entonces $\operatorname{Ker}(T) = \{0\}$ y usando (5.10) deducimos $\dim \operatorname{Im}(T) = \dim V$. Como por hipótesis es $\dim V = \dim W$, entonces $\dim \operatorname{Im}(T) = \dim W$ y por lo tanto $\operatorname{Im}(T) = W$.

Si T es sobreyectiva, es $\operatorname{Im}(T) = W$ y por lo tanto $\dim \operatorname{Im}(T) = \dim W = \dim V$. Entonces usando de nuevo (5.10) deducimos $\dim \operatorname{Ker}(T) = 0$; luego $\operatorname{Ker}(T) = \{0\}$ y por lo tanto T es inyectiva. $\square$

Proposición 5.2.13. *Sean $T : V \to W$ y $S : W \to V$ transformaciones lineales entre espacios de la misma dimensión. Si $S \circ T = \mathrm{Id}_V$, entonces T y S son isomorfismos y $T = S^{-1}$.*

Dem. Es fácil de probar que $S \circ T = \mathrm{Id}_V$ implica que T es inyectivo y S es sobreyectivo (esto vale para funciones, no requiere linealidad). Luego el teorema anterior implica que T y S son isomorfismos. Entonces

$$S \circ T = \mathrm{Id}_V \quad \Rightarrow \quad S^{-1} \circ S \circ T = S^{-1} \circ \mathrm{Id}_V \quad \Rightarrow \quad \mathrm{Id}_W \circ T = S^{-1} \quad \Rightarrow \quad T = S^{-1}. \quad \square$$

Coordenadas. Sea V un espacio un espacio de dimensión n y $B = \{v_1, \dots, v_n\}$ una base *ordenada* de V, es decir una base para la cual fijamos un orden en sus elementos. Si $v \in V$, entonces existen únicos $x_1, \dots, x_n \in \Bbbk$ tales que $v = x_1v_1 + \cdots + x_nv_n$. La n-upla $(x_1, \dots, x_n) \in \Bbbk^n$ son las *coordenadas* del vector v en la base B y escribimos $\operatorname{coord}_B(v) = (x_1, \dots, x_n)$. En lo que sigue vamos a considerar la función $\operatorname{coord}_B : V \to \Bbbk^n$, que asocia a cada vector $v \in V$ su vector de coordenadas $\operatorname{coord}_B(v) \in \Bbbk^n$. Dejamos como ejercicio el probar que $\operatorname{coord}_B : V \to \Bbbk^n$ es una transformación lineal (esto se puede probar directamente o usando la proposición 5.1.7).

Proposición 5.2.14. *Si V es un espacio de dimensión n, entonces $V \simeq \Bbbk^n$.* □

Dem. Sea $B = \{v_1, \dots, v_n\}$ una base ordenada de V. Si $v \in V$ es tal que $\operatorname{coord}_B(v) = (0, \dots, 0)$, entonces $v = 0v_1 + \cdots + 0v_n = 0$. Luego $\operatorname{coord}_B : V \to \Bbbk^n$ es una transformación lineal inyectiva entre dos espacios de la misma dimensión y por lo tanto es un isomorfismo (teorema 5.2.12). □

Ejemplos 5.2.15. En los ejemplos siguientes las bases son ordenadas.

1. Si consideramos la base canónica $B = \{1, x, \dots, x^n\}$ de $\Bbbk_n[x]$, entonces
$$\operatorname{coord}_B(a_0 + a_1x + \cdots + a_nx^n) = (a_0,\, a_1,\, \dots,\, a_n) \in \Bbbk^{n+1}.$$
2. Si $B = \{e_1, \dots, e_n\}$ es la base canónica de $\Bbbk^n$, entonces para cada $(x_1, \dots, x_n) \in \Bbbk^n$ vale $(x_1, \dots, x_n) = x_1e_1 + \cdots + x_ne_n$ y por lo tanto $\operatorname{coord}_B(x_1, \dots, x_n) = (x_1, \dots, x_n)$. Luego $\operatorname{coord}_B = \operatorname{Id} : \Bbbk^n \to \Bbbk^n$.
3. Si $B = \{e_{11}, e_{12}, e_{21}, e_{22}\}$ es la base canónica de $M_{2\times 2}(\Bbbk)$, entonces $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = ae_{11} + be_{12} + ce_{21} + de_{22}$, luego $\operatorname{coord}_B : M_{2\times 2}(\Bbbk) \to \Bbbk^4$ está definida por $\operatorname{coord}_B\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right) = (a, b, c, d)$.
4. El conjunto $B_1 = \{(1, 1),\, (1, -1)\}$ es una base de $\mathbb{R}^2$. Para hallar las coordenadas de un vector genérico $(x, y) \in \mathbb{R}^2$, tenemos que resolver
$$(x, y) = a(1, 1) + b(1, -1) \quad \Rightarrow \quad \left\{ \begin{array}{rcl} a + b & = & x \\ a - b & = & y \end{array} \right. \quad \Rightarrow \quad a = \frac{x + y}{2},\ b = \frac{x - y}{2}.$$
Luego
$$(x, y) = \frac{x + y}{2}(1, 1) + \frac{x - y}{2}(1, -1) \quad \Rightarrow \quad \operatorname{coord}_{B_1}(x, y) = \left(\frac{x + y}{2}, \frac{x - y}{2}\right).$$
5. Consideramos ahora la base $B_2 = \{(1, -1),\, (1, 1)\}$ de $\mathbb{R}^2$. Si $(x, y) \in \mathbb{R}^2$, entonces
$$(x, y) = \frac{x - y}{2}(1, -1) + \frac{x + y}{2}(1, 1) \quad \Rightarrow \quad \operatorname{coord}_{B_2}(x, y) = \left(\frac{x - y}{2}, \frac{x + y}{2}\right).$$
Notar que B_1 y B_2 coinciden como bases, pero no como bases ordenadas. Por eso es que las coordenadas en la base B_1 son distintas que en la base B_2.

Proposición 5.2.16. *Dos espacios son isomorfos si y solo si tienen la misma dimensión.*

Dem. Sean V y W dos espacios. El directo lo vimos en la proposición 5.2.11. Para el recíproco, si $\dim V = \dim W = n$, entonces $V \simeq \Bbbk^n$ y $W \simeq \Bbbk^n$, luego $V \simeq W$. □

Observación 5.2.17. Supongamos que V y W son espacios tales que $\dim V = \dim W = n$. Si elegimos $B = \{v_1, \dots, v_n\}$ y $\mathcal{C} = \{w_1, \dots, w_n\}$ bases ordenadas de V y W, respectivamente, entonces el isomorfismo dado por la proposición anterior es el mapa $T : V \to W$ definido por
$$T(x_1v_1 + \cdots + x_nv_n) = x_1w_1 + \cdots + x_nw_n, \quad \forall x_1, \dots, x_n \in \Bbbk.$$
Este resultado también se puede obtener aplicando la proposición 5.1.7.

5.3. Matriz asociada

En esta sección veremos que dada una transformación lineal $T : V \to W$, si elegimos bases de V y W, entonces a T le podemos hacer corresponder una cierta matriz. Esa matriz contiene toda la información de T, por lo que, conociéndola, conocemos T.

De ahora en adelante vamos a asumir siempre que las bases están ordenadas. Además, usaremos la siguiente notación: si $v_1, \dots, v_n \in \mathbb{k}^m$, entonces $[v_1|\cdots|v_n] \in M_{m\times n}(\mathbb{k})$ es la matriz cuyas columnas son los vectores $v_1, \dots, v_n$ escritos en forma vertical. Por ejemplo, si $v_1 = (1, 2)$, $v_2 = (3, 4)$ y $v_3 = (5, 6)$, entonces

$$[v_1|v_2|v_3] = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

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

$${}_C[T]_B := \left[\operatorname{coord}_C\big(T(v_1)\big)|\cdots|\operatorname{coord}_C\big(T(v_n)\big)\right].$$

Más explícitamente, si escribimos

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

entonces

$${}_C[T]_B = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}. \tag{5.11}$$

Proposición 5.3.1. *Si $T : V \to W$ es una transformación lineal y B y C son bases de V y W respectivamente, entonces*

$$\operatorname{coord}_C\big(T(v)\big) = {}_C[T]_B \cdot \operatorname{coord}_B(v), \quad \forall v \in V. \tag{5.12}$$

En la fórmula anterior, $\operatorname{coord}_B(v)$ *y* $\operatorname{coord}_C\big(T(v)\big)$ *están pensados como vectores columna.*

Dem. Sean $B = \{v_1, \dots, v_n\}$ y $C = \{w_1, \dots, w_m\}$. Definimos ${}_C[T]_B = (a_{ij}) \in M_{m\times n}$ como en (5.11). Si $v \in V$, entonces existen $x_1, \dots, x_n \in \mathbb{k}$ tales que $v = x_1v_1 + \cdots + x_nv_n$. Luego

$$\begin{aligned} T(v) &= T(x_1v_1 + \cdots + x_nv_n) \\ &= x_1T(v_1) + \cdots + x_nT(v_n) \\ &= x_1(a_{11}w_1 + \cdots + a_{m1}w_m) + \cdots + x_n(a_{1n}w_1 + \cdots + a_{mn}w_m) \\ &= (x_1a_{11} + \cdots + x_na_{1n})w_1 + \cdots + (x_1a_{m1} + \cdots + x_na_{mn})w_n \end{aligned}$$

Lo anterior prueba que si $\operatorname{coord}_B(v) = (x_1, \dots, x_n)$, entonces

$$\operatorname{coord}_C\big(T(v)\big) = (x_1a_{11} + \cdots + x_na_{1n},\ \dots,\ x_1a_{m1} + \cdots + x_na_{mn}).$$

Esta última fórmula equivale a (5.12). $\square$

Observación 5.3.2. Si $T : \mathbb{k}^n \to \mathbb{k}^m$ es una transformación lineal, entonces sabemos que existe $A \in M_{m\times n}$ tal que $T = L_A$, es decir, $T(v) = Av$, para todo $v \in \mathbb{k}^n$. Es fácil de probar (usando la proposición anterior o las fórmulas (5.3)) que si $B \subset \mathbb{k}^n$ y $C \subset \mathbb{k}^m$ son las bases canónicas respectivas, entonces $A = {}_C[T]_B$.

Ejemplos 5.3.3.

1. Sea $T : \mathbb{R}_1[x] \to \mathbb{R}^3$ definida por $T(a+bx) = (a+b,\ a+2b,\ 2a+b)$, para todo $(x,y) \in \mathbb{R}^2$. Consideremos la base $B = \{1+x,\ 1-x\}$ y la base canónica $C = \{(1,0,0),\ (0,1,0),\ (0,0,1)\}$. Luego

$$T(1+x) = (2,3,3), \quad T(1-x) = (0,-1,1) \quad \Rightarrow \quad {}_C[T]_B = \begin{pmatrix} 2 & 0 \\ 3 & -1 \\ 3 & 1 \end{pmatrix}.$$

   Observar que en este caso fue fácil obtener ${}_C[T]_B$, porque C es la base canónica de $\mathbb{R}^3$.

2. Consideremos la transformación lineal $T : \mathbb{R}_1[x] \to \mathbb{R}^3$ del ejercicio anterior, pero ahora con las bases $B = \{1+x,\ 1-x\}$ y $C = \{(1,1,1),\ (1,1,0),\ (1,0,0)\}$. Calculando las coordenadas de un vector genérico (x,y,z) en la base C, obtenemos $\operatorname{coord}_C(x,y,z) = (z,\ y-z,\ x-y)$. Luego

$$\operatorname{coord}_C\big(T(1+x)\big) = \operatorname{coord}_C(2,3,3) = (3,0,-1), \quad \operatorname{coord}_C\big(T(1-x)\big) = \operatorname{coord}_C(0,-1,1) = (1,-2,1).$$

   Entonces la matriz asociada es ${}_C[T]_B = \begin{pmatrix} 3 & 1 \\ 0 & -2 \\ -1 & 1 \end{pmatrix}$.

3. Consideremos las bases $B = \{(1,0),\ (1,1)\}$ de $\mathbb{R}^2$ y $C = \{(1,1,1),\ (1,1,0),\ (1,0,0)\}$ de $\mathbb{R}^3$. Supongamos que $T : \mathbb{R}^2 \to \mathbb{R}^3$ es una transformación lineal de la cual su matriz asociada es

$${}_C[T]_B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

   Veamos cómo obtener T. Hallando las coordenadas de un vector genérico (x,y) en la base B obtenemos $\operatorname{coord}_B(x,y) = (x-y,\ y)$. Luego

$$\operatorname{coord}_C\big(T(x,y)\big) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x-y \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \\ x-y \end{pmatrix}$$

   y por lo tanto

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

   Entonces T está definida por $T(x,y) = (2x+y,\ x+2y,\ x+y)$, para todo $(x,y) \in \mathbb{R}^2$.

El siguiente resultado muestra que la fórmula (5.12) caracteriza a la matriz asociada.

Proposición 5.3.4. *Sea $T : V \to W$ una transformación lineal, B una base de V y C una base de V. Si una matriz A verifica*

$$\operatorname{coord}_C\big(T(v)\big) = A \cdot \operatorname{coord}_B(v), \tag{5.13}$$

para todo $v \in V$, entonces $A = {}_C[T]_B$.

Dem. Sea $B = \{v_1, \ldots, v_n\}$. Entonces $\operatorname{coord}_B(v_i) = e_i$, para todo i, siendo $\{e_1, \ldots, e_n\}$ la base canónica de $\Bbbk^n$. Notar que el producto Ae_i nos da la columna i-ésima de A. Por otro lado (5.13) implica

$$\operatorname{coord}_C\big(T(v_i)\big) = Ae_i, \quad \forall i = 1, \ldots, n.$$

Entonces

$$A = [Ae_1|\cdots|Ae_n] = \big[\operatorname{coord}_C\big(T(v_1)\big)|\cdots|\operatorname{coord}_C\big(T(v_n)\big)\big] = {}_C[T]_B\,. \quad \square$$

A continuación veremos cómo se relaciona la matriz aociada con las operaciones en transformaciones lineales.

Proposición 5.3.5. *Sean V, W y U espacios vectoriales. Consideremos bases respectivas B, C y D.*

1. *Sean $T_1, T_2 : V \to W$ dos transformaciones lineales y $a \in \Bbbk$, entonces*

$$ {}_C[T_1 + aT_2]_B = {}_C[T_1]_B + a\,{}_C[T_2]_B\,. \tag{5.14}$$

2. *Sean $T : V \to W$ y $S : W \to U$ dos transformaciones lineales, entonces*

$$ {}_D[S \circ T]_B = {}_D[S]_C \cdot {}_C[T]_B \tag{5.15}$$

Dem. Sabemos que para todo $v \in V$, valen

$$\operatorname{coord}_C\big(T_1(v)\big) = {}_C[T_1]_B \cdot \operatorname{coord}_B(v), \quad \operatorname{coord}_C\big(T_2(v)\big) = {}_C[T_2]_B \cdot \operatorname{coord}_B(v).$$

Entonces usando la linealidad de coord_C, obtenemos

$$\begin{aligned}\operatorname{coord}_C\big((T_1 + aT_2)(v)\big) &= \operatorname{coord}_C\big(T_1(v) + aT_2(v)\big) = \operatorname{coord}_C\big(T_1(v)\big) + a\operatorname{coord}_C\big(T_2(v)\big)\\ &= {}_C[T_1]_B \cdot \operatorname{coord}_B(v) + a\,{}_C[T_2]_B \cdot \operatorname{coord}_B(v) = ({}_C[T_1]_B + a\,{}_C[T_2]_B) \cdot \operatorname{coord}_B(v).\end{aligned}$$

Como esto vale para todo $v \in V$, aplicando la proposición 5.3.4 obtenemos (5.14). Veamos ahora la segunda afirmación. Para todo $v \in V$ y $w \in W$, valen

$$\operatorname{coord}_C\big(T(v)\big) = {}_C[T]_B \cdot \operatorname{coord}_B(v), \quad \operatorname{coord}_D\big(S(w)\big) = {}_D[S]_C \cdot \operatorname{coord}_C(w).$$

Entonces

$$\begin{aligned}\operatorname{coord}_D\big((S \circ T)(v)\big) &= \operatorname{coord}_D\big(S\big(T(v)\big)\big) = {}_D[S]_C \cdot \operatorname{coord}_C\big(T(v)\big) = {}_D[S]_C \cdot \big({}_C[T]_B \cdot \operatorname{coord}_B(v)\big)\\ &= ({}_D[S]_C \cdot {}_C[T]_B) \cdot \operatorname{coord}_B(v), \quad \forall v \in V.\end{aligned}$$

Luego aplicando de nuevo la proposición 5.3.4, obtenemos (5.15). $\square$

Observación 5.3.6. Sea $T : V \to W$ una transformación lineal. Si $B \subset V$ y $C \subset W$ son bases y $A = {}_C[T]_B$, entonces la fórmula (5.12) se puede escribir de la forma

$$\operatorname{coord}_C\big(T(v)\big) = L_A(\operatorname{coord}_B(v))\,, \quad \forall v \in V,$$

lo cual equivale a $\operatorname{coord}_C \circ T = L_A \circ \operatorname{coord}_B$. Esa relación se suele escribir diciendo que el siguiente

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\operatorname{coord}_B"'] & W \arrow[d, "\operatorname{coord}_C"] \\ \Bbbk^n \arrow[r, "L_A"'] & \Bbbk^m \end{tikzcd} \tag{5.16}$$

es un *diagrama conmutativo*. La conmutatividad de (5.16) nos dice que si identificamos un vector con sus coordenadas en la base correspondiente ($v \in V$ con $\operatorname{coord}_B(v) \in \Bbbk^n$ y $w \in W$ con $\operatorname{coord}_C(w) \in \Bbbk^m$), entonces $T : V \to W$ se identifica con la transformación lineal $L_A : \Bbbk^n \to \Bbbk^m$, siendo $A = {}_C[T]_B$.

Recordar que $\mathcal{L}(V,W)$ es el espacio de las transformaciones lineales de V en W.

Proposición 5.3.7. *Sean V y W espacios vectoriales tales que* $\dim V = n$ *y* $\dim W = m$. *Sean B una base de V y C una base de W. Definimos una función $\Phi : \mathcal{L}(V,W) \to M_{m\times n}(\mathbb{k})$ mediante $\Phi(T) = {}_C[T]_B$, para todo $T \in \mathcal{L}(V,W)$. Entonces $\Phi : \mathcal{L}(V,W) \to M_{m\times n}(\mathbb{k})$ es un isomorfismo.*

Dem. La primera parte de la proposición anterior implica que $\Phi : \mathcal{L}(V,W) \to M_{m\times n}(\mathbb{k})$ es una transformación lineal. Además es claro por la fórmula (5.12) que si $T \in \mathcal{L}(V,W)$ verifica ${}_C[T]_B = 0$, entonces $T = 0$. Luego $\operatorname{Ker}(\Phi) = \{0\}$ y por lo tanto Φ es inyectiva.
Probaremos ahora que Φ es sobreyectiva. Dada $A \in M_{m\times n}(\mathbb{k})$, consideremos $L_A \in \mathcal{L}(\mathbb{k}^n, \mathbb{k}^m)$ y definamos $T \in \mathcal{L}(V,W)$ mediante la composición $T := (\operatorname{coord}_C)^{-1} \circ L_A \circ \operatorname{coord}_B$. Luego

$$\operatorname{coord}_C \circ T = L_A \circ \operatorname{coord}_B \quad \Rightarrow \quad \operatorname{coord}_C\big(T(v)\big) = A \cdot \operatorname{coord}_B(v), \quad \forall v \in V.$$

Entonces la proposición 5.3.4 implica ${}_C[T]_B = A$, es decir, $\Phi(T) = A$. Esto prueba la sobreyectividad. $\square$

Corolario 5.3.8. *Si V y W son dos espacios vectoriales, entonces* $\dim \mathcal{L}(V,W) = (\dim V)(\dim W)$.

Esto se debe a que si $\dim V = n$ y $\dim W = m$, entonces $\mathcal{L}(V,W) \simeq M_{m\times n}(\mathbb{k})$ implica

$$\dim \mathcal{L}(V,W) = \dim M_{m\times n}(\mathbb{k}) = mn = (\dim V)(\dim W). \quad \square$$

A continuación veremos cómo se refleja en la matriz asociada el que una transformación lineal sea un isomorfismo. Recordar que si existe un isomorfismo $T : V \to W$, entonces V y W tienen la misma dimensión.

Proposición 5.3.9. *Sea $T : V \to W$ una transformación lineal entre dos espacios de la misma dimensión y sean B y C bases de V y W respectivamente. Entonces T es un isomorfismo si y solo si ${}_C[T]_B$ es invertible. En ese caso vale*

$$ {}_B\left[T^{-1}\right]_C = \left({}_B[T]_C\right)^{-1}. \tag{5.17}$$

Dem. Si T es un isomorfismo, entonces existe su inversa T^{-1} que verifica $T^{-1} \circ T = \operatorname{Id}_V$. Luego

$$ {}_B\left[T^{-1} \circ T\right]_B = {}_B[\operatorname{Id}]_B \quad \Rightarrow \quad {}_B\left[T^{-1}\right]_C \cdot {}_C[T]_B = I$$

Como ambas matrices son cuadradas, esto implica que son invertibles y que vale (5.17).

Recíprocamente, supongamos ahora que ${}_C[T]_B$ es invertible. Si $v \in \operatorname{Ker}(T)$, entonces

$$T(v) = 0 \quad \Rightarrow \quad \operatorname{coord}_C\big(T(v)\big) = 0 \quad \Rightarrow \quad {}_C[T]_B \cdot \operatorname{coord}_B(v) = 0 \quad \Rightarrow \quad \operatorname{coord}_B(v) = 0 \quad \Rightarrow \quad v = 0.$$

Luego $T : V \to W$ es inyectiva y como es $\dim V = \dim W$, concluimos que T es un isomorfismo. $\square$

Corolario 5.3.10. *Si $A \in M_n$, entonces $L_A : \mathbb{k}^n \to \mathbb{k}^n$ es un isomorfismo si y solo si A es invertible. En caso afirmativo es $(L_A)^{-1} = L_{A^{-1}}$.*

Dem. Si C es la base canónica de $\mathbb{k}^n$, entonces ${}_C[L_A]_C = A$. Luego se aplica la proposición 5.3.9. $\square$

A continuación veremos que invertibilidad de una matriz está relacionada con la dependencia lineal de sus filas y de sus columnas. Este resultado será usado en la sección siguiente.

Proposición 5.3.11. *Sea A una matriz cuadrada. Entonces las siguientes afirmaciones son equivalentes.*

1. *La matriz A es invertible.*

2. *Las columnas de A forman un conjunto linealmente independiente.*

3. *Las filas de A forman un conjunto linealmente independiente.*

Dem. Supongamos $A \in M_n(\Bbbk)$ y sean $A_1, \dots, A_n \in \Bbbk^n$ las columnas de A. Recordar que si $\{e_1, \dots, e_n\}$ es la base canónica de $\Bbbk^n$, entonces es $A_i = Ae_i = L_A(e_i)$, para todo $i = 1, \dots, n$; luego $\operatorname{Im}(L_A) = [A_1, \dots, A_n]$. La equivalencia entre las afirmaciones 1 y 2 se debe a lo siguiente.

$$\begin{aligned} A \text{ es invertible} &\Leftrightarrow L_A : \Bbbk^n \to \Bbbk^n \text{ es un isomorfismo} \Leftrightarrow L_A : \Bbbk^n \to \Bbbk^n \text{ es sobreyectiva} \\ &\Leftrightarrow \operatorname{Im}(L_A) = \Bbbk^n \Leftrightarrow [A_1, \dots, A_n] = \Bbbk^n \Leftrightarrow \{A_1, \dots, A_n\} \text{ es un generador de } \Bbbk^n \\ &\Leftrightarrow \{A_1, \dots, A_n\} \text{ es una base de } \Bbbk^n \Leftrightarrow \{A_1, \dots, A_n\} \text{ es LI.} \end{aligned}$$

La afirmación 3 es equivalente a las otras, dado que las filas de A son las columnas de A^t, y A es invertible si y solo si lo es A^t. □

Matriz de cambio de base. Si B y C son dos bases de un mismo espacio V, entonces a la matriz ${}_C[\mathrm{Id}]_B$ se le llama la *matriz de cambio de base* de B a C. En este caso la fórmula (5.12) queda en

$$\operatorname{coord}_C(v) = {}_C[\mathrm{Id}]_B \cdot \operatorname{coord}_B(v), \quad \forall v \in V. \tag{5.18}$$

Esto nos permite conocer las coordenadas de un vector en la base C sabiendo sus coordenadas en B.

Lo que sigue muestra cómo se relacionan ${}_B[\mathrm{Id}]_C$ con ${}_C[\mathrm{Id}]_B$.

Proposición 5.3.12. *Sean B y C dos bases de un mismo espacio V. Entonces la matriz de cambio de base ${}_B[\mathrm{Id}]_C$ es invertible y su inversa es la matriz de cambio de base ${}_C[\mathrm{Id}]_B$.*

Dem. Como la función identidad $\mathrm{Id} : V \to V$ es un isomorfismo, entonces aplicando la proposición 5.3.9 obtenemos que ${}_B[\mathrm{Id}]_C$ es invertible y $({}_B[\mathrm{Id}]_C)^{-1} = {}_C\left[\mathrm{Id}^{-1}\right]_B = {}_C[\mathrm{Id}]_B$. □

Ejemplo 5.3.13. Consideremos en el espacio $\mathbb{R}_1[x]$ los conjuntos $B = \{3x + 2, 2x + 1\}$ y $C = \{x + 1, 1\}$. Dejamos como ejercicio el verifcar que B y C son bases de $\mathbb{R}_1[x]$. Operando, obtenemos que las coordenadas de un polinomio genérico $a+bx$ en la base C son $\operatorname{coord}_C(ax+b) = (a, b-a)$. Luego $\operatorname{coord}_C(3x+2) = (3, -1)$ y $\operatorname{coord}_C(2x + 1) = (2, -1)$ y por lo tanto

$${}_C[\mathrm{Id}]_B = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}.$$

Luego aplicando la proposición 5.3.12 obtenemos

$${}_B[\mathrm{Id}]_C = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}.$$

Luego aplicando la fórmula (5.18) obtenemos

$$\operatorname{coord}_B(ax + b) = {}_B[\mathrm{Id}]_C \cdot \operatorname{coord}_C(ax + b) = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix} \begin{pmatrix} a \\ b-a \end{pmatrix} = \begin{pmatrix} 2b-a \\ 2a-3b \end{pmatrix}.$$

Entonces

$$\operatorname{coord}_B(ax + b) = (2b - a,\ 2a - 3b) \quad \Rightarrow \quad ax + b = (2b - a)(3x + 2) + (2a - 3b)(2x + 1).$$

El siguiente resultado tiene varias aplicaciones, por ejemplo sirve para probar que un conjunto es una base.

Proposición 5.3.14. *Sea $B = \{v_1, \dots, v_n\}$ una base de un espacio V y consideremos $w_1, \dots, w_n \in V$. Sean $a_{ij} \in \Bbbk$ tales que $w_j = \sum_{i=1}^n a_{ij}\, v_i$, para todo $j = 1 \dots, n$. Consideremos $A = (a_{ij}) \in M_n(\Bbbk)$ y $\Delta = \det(A)$. Entonces $C = \{w_1, \dots, w_n\}$ es base de V si y solo si $\Delta \neq 0$. Además, en caso afirmativo es ${}_B[\mathrm{Id}]_C = A$.*

Dem. Si $x_1 \dots, x_n \in \Bbbk$, entonces

$$\sum_{j=1}^n x_j w_j = \sum_{j=1}^n x_j \sum_{i=1}^n a_{ij}\, v_i = \sum_{j=1}^n \sum_{i=1}^n x_j a_{ij}\, v_i = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j \right) v_i.$$

Como $B = \{v_1, \dots, v_n\}$ es LI, entonces vale $\sum_{j=1}^n x_j w_j = 0$ si y solo si $\sum_{j=1}^n a_{ij} x_j = 0$, para todo $i = 1 \dots, n$. Esto último puede escribirse

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ & \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= 0 \end{cases}. \tag{5.19}$$

Como C tiene la misma cantidad de elementos que la dimensión de V, entonces C es una base de V si y solo si es LI, y esto ocurre si y solo si el sistema (5.19) es compatible determinado, lo cual equivale a $\Delta \neq 0$. La última afirmación es consecuencia de cómo están definidos los vectores $w_1, \dots, w_n$. □

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

Proposición 5.3.15. *Sea $T : V \to W$ una transformación lineal. Sean B_1, B_2 bases de V y C_1, C_2 bases de W. Entonces*

$${}_{C_2}[T]_{B_2} = {}_{C_2}[\mathrm{Id}]_{C_1} \cdot {}_{C_1}[T]_{B_1} \cdot {}_{B_1}[\mathrm{Id}]_{B_2},$$

siendo ${}_{C_2}[\mathrm{Id}]_{C_1}$ y ${}_{B_1}[\mathrm{Id}]_{B_2}$ las matrices de cambio de base.

Dem. La prueba consiste en aplicar reiteradamente la fórmula (5.15):

$${}_{C_2}[T]_{B_2} = {}_{C_2}[\mathrm{Id} \circ T \circ \mathrm{Id}]_{B_2} = {}_{C_2}[\mathrm{Id} \circ T]_{B_1} \cdot {}_{B_1}[\mathrm{Id}]_{B_2} = {}_{C_2}[\mathrm{Id}]_{C_1} \cdot {}_{C_1}[T]_{B_1} \cdot {}_{B_1}[\mathrm{Id}]_{B_2}. \quad \square$$

Ejemplo 5.3.16. Queremos saber cómo cambian las coordenadas de un vector en el plano $\mathbb{R}^2$ cuando rotamos nuestro sistema de coordenadas un ángulo θ en sentido positivo (antihorario). Sea $B = \{\mathbf{i}, \mathbf{j}\}$ la base canónica. El nuevo sistema de coordenadas va a tener una base $C = \{\mathbf{i}_\theta, \mathbf{j}_\theta\}$, en que $\mathbf{i}_\theta$ y $\mathbf{j}_\theta$ están definidos por

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

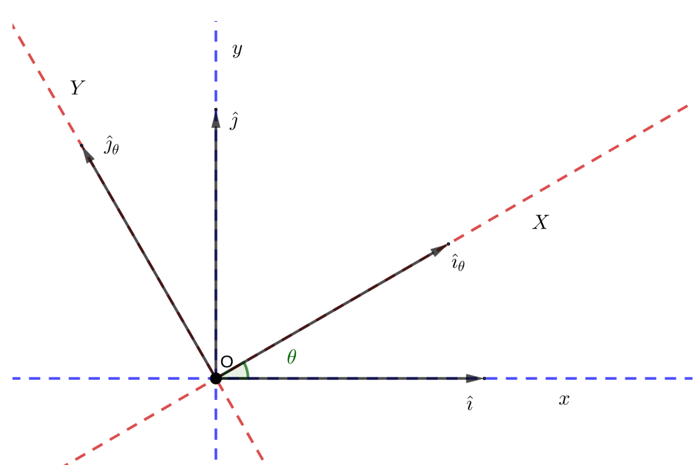

Figura 5.1: Rotación de ejes

Si consideramos un vector $v \in \mathbb{R}^2$, entonces v va a tener ciertas coordenadas (x, y) en la base B y otras coordenadas (X, Y) en la base C, por lo cual

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

Queremos saber cómo se relacionan (x, y) y (X, Y). La matriz de cambio de base de C a B es

$$_B[\mathrm{Id}]_C = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Luego (x, y) se obtiene de (X, Y) usando la fórmula de cambio de base

$$\mathrm{coord}_B(v) = {_B[\mathrm{Id}]_C}\ \mathrm{coord}_C(v) \quad \Rightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \quad \Rightarrow \quad \begin{cases} x = X\cos\theta - Y\operatorname{sen}\theta \\ y = X\operatorname{sen}\theta + Y\cos\theta \end{cases}.$$

Si ahora queremos saber cómo obtener las coordenadas (X, Y) a partir de las coordenadas (x, y), tenemos dos caminos: o las despejamos de la fórmula anterior, o si no usamos la proposición 5.3.12:

$$_B[\mathrm{Id}]_C = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \quad \Rightarrow \quad {_C[\mathrm{Id}]_B} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Luego usando de nuevo la fórmula de cambio de base, obtenemos

$$\mathrm{coord}_C(v) = {_C[\mathrm{Id}]_B}\ \mathrm{coord}_B(v) \quad \Rightarrow \quad \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \Rightarrow \quad \begin{cases} X = x\cos\theta + y\operatorname{sen}\theta \\ Y = -x\operatorname{sen}\theta + y\cos\theta \end{cases}.$$

Resumiendo, las fórmulas para pasar de un sistema de coordenadas al otro son las siguientes.

$$\begin{aligned} v = X\mathbf{i}_\theta + Y\mathbf{j}_\theta \quad &\Rightarrow \quad v = (X\cos\theta - Y\operatorname{sen}\theta)\mathbf{i} + (X\operatorname{sen}\theta + Y\cos\theta)\mathbf{j}, \\ v = x\mathbf{i} + y\mathbf{j} \quad &\Rightarrow \quad v = (x\cos\theta + y\operatorname{sen}\theta)\,\mathbf{i}_\theta + (-x\operatorname{sen}\theta + y\cos\theta)\,\mathbf{j}_\theta. \end{aligned}$$

5.4. Rango

Si A es una matriz, entonces llamamos *rango* o *rango por columnas* de A, a la cantidad máxima de columnas linealmente independientes que tiene A.

Ejemplo 5.4.1. Consideremos las matrices

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 1 & 2 & 0 & 3 \\ 1 & 2 & 0 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Entonces $\mathrm{rango}(A) = 3$, $\mathrm{rango}(B) = 2$, $\mathrm{rango}(C) = 1$ y $\mathrm{rango}(D) = 0$.

Observación 5.4.2. Si $A \in M_n$, entonces el rango de A es n si y solo si las columnas de A forman un conjunto linealmente independiente. Pero en la proposición 5.3.11 vimos que esto equivale a que A sea invertible. Luego $A \in M_n$ es invertible si y solo si el rango de A es n.

En estas notas hemos visto varias condiciones que equivalen a la invertibilidad de una matriz. Las resumimos en el siguiente teorema.

Teorema 5.4.3. *Si $A \in M_n$, entonces las siguientes afirmaciones son equivalentes.*

- *La matriz A es invertible.*
- *Existe una matriz $B \in M_n$ tal que $AB = I$.*
- *Existe una matriz $B \in M_n$ tal que $BA = I$.*
- *El determinante de A es distinto de cero.*
- *Las columnas de A forman un conjunto LI.*
- *Las filas de A forman un conjunto LI.*
- *El rango de A es n.* □

Es claro que el teorema anterior implica lo siguiente.

Corolario 5.4.4. *Si $A \in M_n$, entonces las siguientes afirmaciones son equivalentes.*

- *El determinante de A es cero.*
- *Existe una columna de A que es combinación lineal de las restantes columnas.*
- *Existe una fila de A que es combinación lineal de las restantes filas.*
- *El rango de A es menor que n.* □

A continuación veremos otra forma de pensar el rango.

Proposición 5.4.5. *El rango de $A \in M_{m\times n}(\Bbbk)$ coincide con la dimensión del subespacio de $\Bbbk^m$ generado por las columnas de A.*

Dem. Sean $A_1, \dots, A_n \in \Bbbk^m$ las columnas de A. Consideremos $W = [A_1, \dots, A_n]$ el subespacio de $\Bbbk^m$ generado por $A_1, \dots, A_n$. Si $\operatorname{rango}(A) = r$, entonces existe $B = \{A_{i_1}, \dots, A_{i_r}\}$ conjunto LI de columnas de A. Si existiese otra columna A_k de A que no es combinación lineal de $A_{i_1}, \dots, A_{i_r}$, entonces $\{A_{i_1}, \dots, A_{i_r}, A_k\}$ sería un conjunto LI de columnas de A, contradiciendo la maximalidad de r. Luego todas las otras columnas de A son combinación lineal de $A_{i_1}, \dots, A_{i_r}$ y por lo tanto B es un generador de W. Como además B es LI, concluimos que B es una base de W. Luego $\dim W = \#B = r$. □

Si $T : V \to W$ es una transformación lineal, entonces llamamos *rango* de T a la dimensión de la imagen de T. El siguiente resultado muestra cómo se relacionan las dos definiciones anteriores de rango.

Proposición 5.4.6.

1. *Sea $T \in \mathcal{L}(V, W)$. Si B y C son bases de V y W respectivamente, entonces* $\operatorname{rango}(T) = \operatorname{rango}({}_C[T]_B)$.
2. *Sea $A \in M_{m\times n}$ y consideremos $L_A : \Bbbk^n \to \Bbbk^m$. Entonces* $\operatorname{rango}(A) = \operatorname{rango}(L_A)$.

Dem. Consideremos la primera afirmación. Si $B = \{v_1, \dots, v_n\}$, entonces $T(B) = \{T(v_1), \dots, T(v_n)\}$ es un generador de $\operatorname{Im}(T)$ y por lo tanto el rango de T es la cantidad máxima de vectores LI que hay en $T(B)$. Por otro lado es

$$_C[T]_B = \left[\operatorname{coord}_C\big(T(v_1)\big)|\cdots|\operatorname{coord}_C\big(T(v_n)\big)\right].$$

Como el mapa $\operatorname{coord}_C : W \to \mathbb{k}^m$ ($m = \dim W$) es un isomorfismo (y por lo tanto lleva conjuntos LI en conjuntos LI), deducimos que la cantidad máxima de vectores LI que hay en $T(B)$ coincide con la cantidad máxima de columnas LI que hay en $_C[T]_B$. Luego $\operatorname{rango}(T) = \operatorname{rango}(_C[T]_B)$. La segunda afirmación se deduce de la primera, dado que A es la matriz asociada a L_A en las bases canónicas correspondientes. $\square$

A continuación veremos cómo se relaciona el rango con la equivalencia de matrices. Para eso necesitamos el siguiente resultado.

Proposición 5.4.7. *Sean* $V_1 \xrightarrow{S} V \xrightarrow{T} W \xrightarrow{R} W_1$ *transformaciones lineales.*

1. *Si S es sobreyectiva, entonces* $\operatorname{Im}(T \circ S) = \operatorname{Im}(T)$.

2. *Si R es inyectiva, entonces* $\dim \operatorname{Im}(R \circ T) = \dim \operatorname{Im}(T)$.

3. *Si S y R son isomorfismos, entonces* $\dim \operatorname{Im}(R \circ T \circ S) = \dim \operatorname{Im}(T)$.

Dem. La primera afirmación es válida aun cuando S y T son solo funciones. Para probarlo, notar que siempre vale $\operatorname{Im}(T \circ S) \subset \operatorname{Im}(T)$. Por otro lado, si $w \in \operatorname{Im}(T)$, entonces existe $v \in V$ tal que $w = T(v)$. Como S es sobreyectiva, entonces existe $v_1 \in V_1$ tal que $v = S(v_1)$. Luego $w = T(v) = T\big(S(v_1)\big) = (T \circ S)(v_1)$. Esto implica $\operatorname{Im}(T) \subset \operatorname{Im}(T \circ S)$ y por lo tanto $\operatorname{Im}(T \circ S) = \operatorname{Im}(T)$.

Consideremos la segunda afirmación. Sea $\dim \operatorname{Im}(T) = n$ y $B = \{w_1, \dots, w_n\}$ una base de $\operatorname{Im}(T)$. Probaremos que $C = \{R(w_1), \dots, R(w_n)\}$ es una base de $\operatorname{Im}(R \circ T)$, lo cual implica la segunda afirmación.
Como R es inyectiva y B es LI, se deduce que $C = R(B)$ es LI. Luego solo resta probar que C genera a $\operatorname{Im}(R \circ T)$. Sea $w \in \operatorname{Im}(R \circ T)$. Consideremos $v \in V$ tal que $w = R\big(T(v)\big)$. Como $T(v) \in \operatorname{Im}(T)$ y B es base de $\operatorname{Im}(T)$, entonces existen $a_1, \dots, a_n \in \mathbb{k}$ tales que $T(v) = \sum_{i=1}^n a_i w_i$. Entonces $w = R\big(T(v)\big) = \sum_{i=1}^n a_i R(w_i)$. Esto prueba que C es un generador de $\operatorname{Im}(R \circ T)$.

La tercera afirmación se deduce de las dos anteriores:

$$\dim \operatorname{Im}(R \circ T \circ S) = \dim \operatorname{Im}\big(R \circ (T \circ S)\big) = \dim \operatorname{Im}(T \circ S) = \dim \operatorname{Im}(T). \quad \square$$

Corolario 5.4.8. *Si $A, B \in M_{m \times n}$ son matrices equivalentes, entonces* $\operatorname{rango}(A) = \operatorname{rango}(B)$.

Dem. Como es $A \sim B$, entonces existen matrices invertibles $P \in M_m$ y $Q \in M_n$ tales que $A = PBQ$. Luego $L_A = L_{PBQ} = L_P \circ L_B \circ L_Q$. Como P y Q son invertibles, entonces L_P y L_Q son isomorfismos y por lo tanto $\operatorname{rango}(A) = \dim \operatorname{Im}(L_A) = \dim \operatorname{Im}(L_B) = \operatorname{rango}(B)$. $\square$

Observación 5.4.9. Cuando estudiamos la equivalencia de matrices, vimos que si $A \in M_{m \times n}$ y $A \neq 0$, entonces existe un núnero r, con $1 \leq r \leq \min\{m, n\}$, tal que $A \sim \Phi_r^{m,n} := \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right) \in M_{m \times n}$, siendo $I_r \in M_r$ la matriz identidad. Notar que el rango de $\Phi_r^{m,n}$ es r, dado que sus primeras r columnas son parte de la base canónica de $\mathbb{k}^m$ y las otras columnas son nulas.

Proposición 5.4.10. *Si $A \in M_{m \times n}$ y $A \neq 0$, entonces existe un único r, con $1 \leq r \leq \min\{m, n\}$, tal que $A \sim \Phi_r^{m,n}$. Este número r es el rango de A.*

Dem. Sabemos que existe r, con $1 \leq r \leq \min\{m, n\}$, tal que $A \sim \Phi_r^{m,n}$. Entonces aplicando el corolario anterior obtenemos $\operatorname{rango}(A) = \operatorname{rango}(\Phi_r^{m,n}) = r$. Luego $r = \operatorname{rango}(A)$, lo cual implica la unicidad de r. $\square$

El siguiente teorema muestra que dos matrices son equivalentes si y solo si tienen el mismo rango.

Teorema 5.4.11. *Sean $A, B \in M_{m\times n}$. Entonces $A \sim B$ si y solo si* $\operatorname{rango}(A) = \operatorname{rango}(B)$.

Dem. El directo es el corolario de arriba. Supongamos ahora $\operatorname{rango}(A) = \operatorname{rango}(B) = r$. Si $r = 0$, entonces $A = B = 0$ y por lo tanto $A \sim B$. Si $r > 0$, entonces $A \sim \Phi_r^{m,n}$ y $B \sim \Phi_r^{m,n}$, luego $A \sim B$. □

Observación 5.4.12. La equivalencia de matrices parte al conjunto $M_{m\times n}$ en clases de equivalencia. El teorema anterior nos dice que si $k = \min\{m, n\}$, entonces en $M_{m\times n}$ hay exactamente $k + 1$ clases de equivalencia, que son las correspondientes a la matriz nula y a las matrices $\Phi_1^{m,n}, \dots, \Phi_k^{m,n}$.

Ahora veremos que la trasposición no afecta el rango.

Proposición 5.4.13. *Si $A \in M_{m\times n}$, entonces* $\operatorname{rango}(A) = \operatorname{rango}\left(A^t\right)$.

Dem. Sea $r = \operatorname{rango}(A)$. Entonces $A \sim \Phi_r^{m,n}$ y por lo tanto existen matrices invertibles $P \in M_m$ y $Q \in M_n$ tales que $A = P\Phi_r^{m,n}Q$. Luego

$$A^t = (P\Phi_r^{m,n}Q)^t = Q^t(\Phi_r^{m,n})^t P^t = Q^t\Phi_r^{n,m}P^t.$$

Como P^t y Q^t son matrices invertibles, deducimos $A^t \sim \Phi_r^{n,m}$ y por lo tanto $\operatorname{rango}(A^t) = r = \operatorname{rango}(A)$. □

Si $A \in M_{m\times n}$, entonces llamamos *rango por filas* de A a la cantidad máxima de filas linealmente independientes que hay en A. Como la trasposición de matrices intercambia las filas con las columnas, entonces la proposición anterior prueba que el rango por filas coincide con el rango por columnas.

Observación 5.4.14. Como matrices equivalentes tiene el mismo rango, para calcular el rango de una matriz siempre podemos realizar operaciones elementales en sus filas y columnas, con el objetivo de transformarla en una matriz en la cual sea más simple calcular el rango.

Ejemplo 5.4.15. Sea $A = \begin{pmatrix} -7 & 4 & 1 \\ 4 & -2 & -2 \\ 1 & -2 & 7 \end{pmatrix} \in M_3(\mathbb{R})$. Para calcular el rango de A, observar que si primero le sumamos a la primera columna la segunda multiplicada por 2 y luego a la tercera fila le sumamos la primera multiplicada por 3, obtenemos:

$$\begin{pmatrix} -7 & 4 & 1 \\ 4 & -2 & -2 \\ 1 & -2 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & 1 \\ 0 & -2 & -2 \\ -3 & -2 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & 1 \\ 0 & -2 & -2 \\ 0 & 10 & 10 \end{pmatrix}.$$

Luego el rango de A coincide con el rango de la última matriz, que claramente es 2 (rango por filas).

Ahora veremos otra forma de calcular el rango, usando determinantes.

Sea $A \in M_{m\times n}(\Bbbk)$. Los *menores de orden* h de A ($1 \leq h \leq \min\{m, n\}$) son los determinantes de las submatrices cuadradas de A de orden h, obtenidas suprimiendo $m - h$ filas y $n - h$ columnas de A.

Ejemplo 5.4.16. Si $A = \begin{pmatrix} 1 & 2 & 5 & 0 \\ 3 & 6 & 7 & 0 \\ 4 & 8 & 12 & 0 \end{pmatrix} \in M_{3\times 4}(\mathbb{R})$, entonces A tiene menores de orden 1, 2 y 3. Los menores de orden 1 son simplemente los coeficientes de A. Los menores de orden 3 se obtienen suprimiendo una columna de A:

$$\begin{vmatrix} 2 & 5 & 0 \\ 6 & 7 & 0 \\ 8 & 12 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 5 & 0 \\ 3 & 7 & 0 \\ 4 & 12 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 2 & 0 \\ 3 & 6 & 0 \\ 4 & 8 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 2 & 5 \\ 3 & 6 & 7 \\ 4 & 8 & 12 \end{vmatrix} = 0.$$

Los menores de orden 2 se obtienen suprimiendo una fila y dos columnas de A; por ejemplo los menores de orden 2 que se obtienen eliminando la última fila son:

$$\begin{vmatrix} 1 & 2 \\ 3 & 6 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 5 \\ 3 & 7 \end{vmatrix} = -8, \quad \begin{vmatrix} 2 & 5 \\ 6 & 7 \end{vmatrix} = -16, \quad \begin{vmatrix} 5 & 0 \\ 7 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 2 & 0 \\ 6 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} = 0.$$

Los menores de orden 2 que se obtienen eliminando la fila del medio son:

$$\begin{vmatrix} 1 & 2 \\ 4 & 8 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 5 \\ 4 & 12 \end{vmatrix} = -8, \quad \begin{vmatrix} 2 & 5 \\ 8 & 12 \end{vmatrix} = -16, \quad \begin{vmatrix} 5 & 0 \\ 12 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 2 & 0 \\ 8 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 \\ 4 & 0 \end{vmatrix} = 0.$$

Los menores de orden 2 que se obtienen eliminando la primera fila son:

$$\begin{vmatrix} 3 & 6 \\ 4 & 8 \end{vmatrix} = 0, \quad \begin{vmatrix} 3 & 7 \\ 4 & 12 \end{vmatrix} = 8, \quad \begin{vmatrix} 6 & 7 \\ 8 & 12 \end{vmatrix} = 16, \quad \begin{vmatrix} 7 & 0 \\ 12 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 6 & 0 \\ 8 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 3 & 0 \\ 4 & 0 \end{vmatrix} = 0.$$

Teorema 5.4.17. *Si* $A \in M_{m\times n}(\Bbbk)$*, entonces* $\operatorname{rango}(A) = \max\{\textit{orden de } M : \ M \textit{ menor no nulo de } A\}$.

Dem. Sean $r = \operatorname{rango}(A)$ y $l = \max\{\text{orden de M} : \ \text{M menor no nulo de } A\}$. Queremos probar $r = l$.

Sea B una submatriz cuadrada de A de orden q. Si $q > r$, entonces como todo conjunto de q columnas de A es LD, deducimos que las columnas de B son LD y por lo tanto el determinante de B es cero. Luego si $q > r$, entonces todos los menores de orden q de A son nulos. Esto implica $l \leq r$.
Para probar $l = r$ alcanza con mostrar que existe un menor no nulo de orden r de A. Como el rango de A es r, entonces existen r columnas de A que son LI. Si C es la matriz $m \times r$ formada por esas columnas, entonces el rango de C es r y por lo tanto existen r filas de C que son LI. Sea D la submatriz de C formada por esas filas. Entonces D es una matriz $r \times r$ que tiene r filas LI y por lo tanto su determinante es distinto de cero (proposición 5.3.11). Así que D es un menor de A que es no nulo y tiene orden r. $\square$

Ejemplo 5.4.18. En el ejemplo 5.4.16 obtuvimos que todos los menores de orden 3 son nulos y existe algún menor de orden 2 no nulo, luego el rango de A es 2.

Si $A \in M_{m\times n}(\Bbbk)$, entonces al número $\max\{\text{orden de M} : \ \text{M menor no nulo de } A\}$ se le llama el *rango por determinantes* de A. Lo que hemos probado en esta sección es que, para una matriz dada, es lo mismo el rango por columnas, el rango por filas y el rango por determinantes. Esto nos da tres alternativas para calcular el rango de una matriz.

5.5. Espacio dual

En la proposición 5.1.11 vimos que si V y W son espacios vectoriales, entonces $\mathcal{L}(V, W)$ es también un espacio vectorial. Como el cuerpo $\Bbbk$ es un espacio vectorial, entonces a todo espacio V le podemos asociar el espacio $V^* := \mathcal{L}(V, \Bbbk)$ llamado el *espacio dual* de V. A los elementos de V^* se les suele llamar *funcionales*.

Ejemplos 5.5.1.

1. Aplicando la proposición 5.1.4, obtenemos que una función $\alpha : \Bbbk^n \to \Bbbk$ está en $(\Bbbk^n)^*$ si y solo si existen $a_1, \ldots, a_n \in \Bbbk$ tales que

$$\alpha(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n, \quad \forall (x_1, \ldots, x_n) \in \Bbbk^n.$$

   Así que los elementos de $(\Bbbk^n)^*$ son las funciones de la forma anterior.

2. Recordar que la traza de $A = (a_{ij}) \in M_n(\Bbbk)$ está definida por $\operatorname{tr}(A) = \sum_{i=1}^n a_{ii}$. Haciendo variar A en $M_n(\Bbbk)$ obtenemos la *función traza* $\operatorname{tr} : M_n(\Bbbk) \to \Bbbk$. Sabemos que vale $\operatorname{tr}(A + cB) = \operatorname{tr}(A) + c \operatorname{tr}(B)$, para todo $A, B \in M_n(\Bbbk)$ y $c \in \Bbbk$, luego $\operatorname{tr} \in M_n(\Bbbk)^*$.

Proposición 5.5.2. *Sea $\alpha \in V^*$. Si α no es la función nula, entonces $\alpha : V \to \Bbbk$ es sobreyectiva y* $\dim \operatorname{Ker}(\alpha) = \dim V - 1$.

Dem. Si $\alpha \in V^*$ no es la funcional nula, entonces $\operatorname{Im}(\alpha) \neq \{0\}$. Como $\operatorname{Im}(\alpha) \subset \Bbbk$ es un subespacio y $\dim \Bbbk = 1$, esto implica $\operatorname{Im}(\alpha) = \Bbbk$ y por lo tanto α es sobreyectiva. Luego

$$\dim V = \dim \operatorname{Ker}(\alpha) + \dim \operatorname{Im}(\alpha) = \dim \operatorname{Ker}(\alpha) + 1 \quad \Rightarrow \quad \dim \operatorname{Ker}(\alpha) = \dim V - 1. \quad \square$$

Sea V un espacio vectorial. Si $B = \{e_1, \dots, e_n\}$ es una base de V, entonces para cada $i = 1, \dots, n$ existe una transformación lineal $e_i^* : V \to \Bbbk$ tal que

$$e_i^*(e_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}, \quad \forall j = 1, \dots, n. \tag{5.20}$$

Explícitamente, $e_i^* : V \to \Bbbk$ está definida por $e_i^*(x_1 e_1 + \cdots + x_n e_n) = x_i$, para todo $x_1, \dots, x_n \in \Bbbk$.

Ejemplo 5.5.3. Si $B = \{e_1, \dots, e_n\}$ es la base canónica de $\Bbbk^n$, entonces $e_1^*, \dots, e_n^* \in (\Bbbk^n)^*$ son las funciones definidas por $e_i^*(x_1, \dots, x_n) = x_i$, para todo $(x_1, \dots, x_n) \in \Bbbk^n$.

Proposición 5.5.4. *Sea $B = \{e_1, \dots, e_n\}$ una base de V. Si definimos $e_1^*, \dots, e_n^* \in V^*$ como arriba, entonces:*

1. *Para todo $v \in V$, vale $v = \sum_{i=1}^n e_i^*(v) e_i$.*
2. *Para todo $\alpha \in V^*$, vale $\alpha = \sum_{i=1}^n \alpha(e_i) e_i^*$.*

Dem. Si $v \in V$, entonces existen escalares $x_i \in \Bbbk$ tales que $v = \sum_{i=1}^n x_i e_i$. Luego es $e_i^*(v) = x_i$, para todo i y por lo tanto $v = \sum_{i=1}^n e_i^*(v) e_i$. Sea ahora $\alpha \in V^*$. Si $v \in V$, entonces usando la primera parte obtenemos

$$\alpha(v) = \alpha\left(\sum_{i=1}^n e_i^*(v) e_i\right) = \sum_{i=1}^n \alpha\big(e_i^*(v) e_i\big) = \sum_{i=1}^n e_i^*(v) \alpha(e_i) = \sum_{i=1}^n \alpha(e_i) e_i^*(v) = \left(\sum_{i=1}^n \alpha(e_i) e_i^*\right)(v).$$

Como esto vale para todo $v \in V$, deducimos $\alpha = \sum_{i=1}^n \alpha(e_i) e_i^*$. $\square$

Proposición 5.5.5. *Si $B = \{e_1, \dots, e_n\}$ es una base de V, entonces $B^* := \{e_1^*, \dots, e_n^*\}$ es una base de V^*.*

Dem. Como $\dim \Bbbk = 1$, entonces $\dim V^* = \dim \mathcal{L}(V, \Bbbk) = \dim V$. Además, la segunda parte de la proposición anterior implica que B^* es un generador de V^*. Entonces B^* es un generador de V^* con la misma cantidad de elementos que la dimensión de V^* y por lo tanto B^* es una base de V^*. $\square$

Si $B = \{e_1, \dots, e_n\}$ es una base de V, entonces $B^* = \{e_1^*, \dots, e_n^*\}$ es la *base dual* de B en V^*.

Dual de una transformación lineal. Si $T : V \to W$ es una transformación lineal, entonces definimos $T^* : W^* \to V^*$ mediante

$$T^*(\beta) := \beta \circ T, \quad \forall \beta \in W^*.$$

Notar que $T^*(\beta) : W \to \Bbbk$ está definida de forma tal que el siguiente diagrama conmute

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[dr, dashed, "T^*(\beta)"'] & W \arrow[d, "\beta"] \\ & \Bbbk \end{tikzcd}$$

A la función $T^* : W^* \to V^*$ se le suele llamar la *dual* o la *traspuesta* de T.

Proposición 5.5.6. *Si $T : V \to W$ es una transformación lineal, entonces $T^* : W^* \to V^*$ también es una transformación lineal.*

Dem. Sean $\alpha, \beta \in W^*$ y $a \in \Bbbk$. Entonces $T^*(\alpha + a\beta) = (\alpha + a\beta) \circ T = \alpha \circ T + a(\beta \circ T) = T^*(\alpha) + aT^*(\beta)$. $\square$

Proposición 5.5.7. *Dados V y W, la función $\mathcal{L}(V, W) \to \mathcal{L}(W^*, V^*)$ definida por $T \mapsto T^*$, para toda $T \in \mathcal{L}(V, W)$, es una transformación lineal.*

Dem. Consideremos $T_1, T_2 \in \mathcal{L}(V, W)$ y $a \in \Bbbk$. Sea $\beta \in W^*$, entonces

$$(T_1 + aT_2)^*(\beta) = \beta \circ (T_1 + aT_2) = \beta \circ T_1 + a(\beta \circ T_2) = T_1^*(\beta) + aT_2^*(\beta) = (T_1^* + aT_2^*)(\beta).$$

Como lo anterior vale para todo $\beta \in W^*$, concluimos $(T_1 + aT_2)^* = T_1^* + aT_2^*$. $\square$

La proposición siguiente describe otras propiedades de la correspondencia $T \mapsto T^*$.

Proposición 5.5.8.
1. *Para todo espacio V, vale $(\mathrm{Id}_V)^* = \mathrm{Id}_{V^*} : V^* \to V^*$.*
2. *Si $S : U \to V$ y $T : V \to W$ son transformaciones lineales, entonces $(T \circ S)^* = S^* \circ T^*$.*
3. *Si $T : V \to W$ es un isomorfismo, entonces $T^* : W^* \to V^*$ es un isomorfismo y $(T^*)^{-1} = \left(T^{-1}\right)^*$.*

Dem. La prueba de la primera afirmación es inmediata: $(\mathrm{Id}_V)^*(\alpha) = \alpha \circ \mathrm{Id}_V = \alpha$, para todo $\alpha \in V^*$. Para la segunda, es

$$(T \circ S)^*(\alpha) = \alpha \circ (T \circ S) = (\alpha \circ T) \circ S = S^*(\alpha \circ T) = S^*(\left(T^*(\alpha)\right) = (S^* \circ T^*)(\alpha), \quad \forall \alpha \in W^*.$$

Para la tercera, al ser $T \circ T^{-1} = T^{-1} \circ T = \mathrm{Id}_V$, aplicando las dos partes anteriores obtenemos

$$\left(T \circ T^{-1}\right)^* = \left(T^{-1} \circ T\right)^* = (\mathrm{Id}_V)^* \quad \Rightarrow \quad \left(T^{-1}\right)^* \circ T^* = T^* \circ \left(T^{-1}\right)^* = \mathrm{Id}_{V^*}.$$

Esta última igualdad implica que T^* es invertible y su inversa es $\left(T^{-1}\right)^*$. $\square$

A continuación veremos que, en bases adecuadas, la matriz asociada a T^* es la traspuesta de la matriz asociada a T.

Proposición 5.5.9. *Sea $T : V \to W$ una transformación lineal. Si B y C son bases respectivas de V y W, entonces*

$${}_{B^*}[T^*]_{C^*} = \left({}_C[T]_B\right)^t.$$

Dem. Sean $B = \{e_1, \dots, e_n\}$ y $C = \{f_1, \dots, f_m\}$. Supongamos ${}_C[T]_B = (a_{ij})$ y ${}_{B^*}[T^*]_{C^*} = (b_{ij})$, entonces

$$T(e_j) = \sum_{i=1}^{m} a_{ij} f_i, , \quad \forall j = 1, \dots, n, \tag{5.21}$$

$$T^*(f_l^*) = \sum_{k=1}^{n} b_{kl} e_k^*, \quad \forall l = 1, \dots, m. \tag{5.22}$$

Aplicando la segunda parte de la proposición 5.5.4 y la fórmula (5.21), obtenemos

$$\begin{aligned} T^*(f_l^*) &= \sum_{k=1}^{n} \left(T^*(f_l^*)\right)(e_k)\, e_k^* = \sum_{k=1}^{n} f_l^*(T(e_k))\ e_k^* = \sum_{k=1}^{n} f_l^* \left(\sum_{i=1}^{m} a_{ik} f_i \right)\ e_k^* = \sum_{k=1}^{n} \sum_{i=1}^{m} a_{ik} f_l^*(f_i)\, e_k^* \\ &= \sum_{k=1}^{n} \sum_{i=1}^{m} a_{ik} \delta_{li}\, e_k^* = \sum_{k=1}^{n} a_{lk}\, e_k^*. \end{aligned}$$

Comparando esta expresión de $T^*(f_l^*)$ con la de (5.22), deducimos $a_{lk} = b_{kl}$, para todo l, k. $\square$

Apéndice A

Conjuntos, relaciones y funciones

Si bien en este texto asumimos que el lector tiene los conocimientos básicos sobre teoría de conjuntos que se suelen ver en los cursos de enseñanza secundaria, vamos a hacer un breve repaso de los mismos con énfasis en las relaciones y funciones, lo que de paso nos permite fijar ciertas ideas y notaciones que se usan en este texto. En lo que sigue no vamos a ser muy formales y a menudo usaremos conceptos que no definiremos pero que deben ser bien conocidos por el lector.

En teoría de conjuntos partimos de tres conceptos primitivos, que son *conjunto*, *elemento* y *pertenencia*. En general usaremos letras mayúsculas para los conjuntos, minúsculas para los elementos y el signo $\in$ para la pertenencia. El símbolo $\emptyset$ denota al conjunto *vacío*, que es el conjunto que no tiene elementos. Decimos que dos conjuntos son *iguales* si tienen los mismos elementos. En particular estamos asumiendo que los conjuntos no tienen elementos repetidos y que no están ordenados. Por ejemplo, vale

$$\{a, b\} = \{b, a\} = \{a, a, b\}.$$

Decimos que el conjunto A está *contenido* en el conjunto B o que A es un *subconjunto* de B y escribimos $A \subset B$ si todos los elementos de A están en B. Si además existen elementos de B que no están en A, entonces decimos que A está contenido *estrictamente* en B y escribimos[1] $A \subsetneq B$. Observar que A y B son iguales si y solo si A está contenido en B y B está contenido en A. En símbolos,

$$A = B \quad \Leftrightarrow \quad A \subset B \text{ y } B \subset A.$$

Dados dos subconjuntos A y B de un conjunto X, su *unión* $A \cup B$ y su *intersección* $A \cap B$ se definen mediante

$$A \cup B := \{x \in X : \ x \in A \text{ o } x \in B\}, \qquad A \cap B := \{x \in X : \ x \in A \text{ y } x \in B\}.$$

También definimos su *resta* $A \setminus B$ mediante

$$A \setminus B := \{x \in A : \ x \notin B\}.$$

La unión e intersección de dos conjuntos se generaliza naturalmente a una familia arbitraria $\{A_\alpha\}_{\alpha \in \Lambda}$ de subconjuntos de X

$$\bigcup_{\alpha \in \Lambda} A_\alpha := \{x \in X : \ \exists \alpha \in \Lambda \text{ tal que } x \in A_\alpha\}, \qquad \bigcap_{\alpha \in \Lambda} A_\alpha := \{x \in X : \ x \in A_\alpha, \ \forall \alpha \in \Lambda\}.$$

Cuando consideremos un conjunto cuyos elementos son a su vez conjuntos, a veces le llamaremos *familia* para evitar confusiones. Dado un conjunto A, su conjunto *potencia* o *conjunto de las partes* de A es la familia de todos los subconjuntos de A; la potencia de A la escribiremos $\mathcal{P}(A)$ (también se escribe 2^A).

[1] Conviene tener cuidado con estas notaciones, porque hay quien escribe $A \subseteq B$ en vez de $A \subset B$ y $A \subset B$ en vez de $A \subsetneq B$.

Para simplificar, de ahora en adelante asumiremos que los conjuntos con los que trabajamos son no vacíos. Eso nos evita entrar en discusiones técnicas, que no ayudan en una primera aproximación a estos temas.

Un *par ordenado* es un conjunto de dos elementos en el cual fijamos un orden para los mismos. Formalmente, dados dos elementos a, b el par ordenado (a, b) se define por $(a, b) := \big\{\{a\},\ \{a, b\}\big\}$. De esta forma vale

$$(a, b) = (c, d) \quad \Leftrightarrow \quad a = c \text{ y } b = d.$$

Dados dos conjuntos A y B, su *producto cartesiano* es el conjunto $A \times B := \{(a, b) : \ a \in A \text{ y } b \in B\}$. Muchos conceptos matemáticos se definen por medio de relaciones entre conjuntos. Una *relación* R entre dos conjuntos A y B es, por definición, un subconjunto de $A \times B$. Las relaciones que más aparecen, son las relaciones de orden, las relaciones de equivalencia y las funciones.

Relaciones de orden. Un *orden parcial* en un conjunto A es una relación $R \subset A \times A$ que verifica las siguiente propiedades.

- *Reflexiva*: $(a, a) \in R$, para todo $a \in A$.
- *Antisimétrica*: si $(a, b) \in R$ y $(b, a) \in R$, entonces $a = b$.
- *Transitiva*: si $(a, b) \in R$ y $(b, c) \in R$, entonces $(a, c) \in R$.

En general se suele usar el símbolo $\leq$ para el orden parcial y escribir $a \leq b$ para $(a, b) \in R$. Con esta notación las propiedades anteriores se escriben:

$$a \leq a,\ \forall a \in A; \quad a \leq b \text{ y } b \leq a \ \Rightarrow\ a = b; \quad a \leq b \text{ y } b \leq c \ \Rightarrow\ a \leq c.$$

Un *conjunto parcialmente ordenado* es un par $(A, \leq)$, en el cual A es un conjunto y $\leq$ es un orden parcial en A. Si un orden parcial verifica que todos los elementos de A son *comparables*, es decir que dados $a, b \in A$ vale siempre $a \leq b$ o $b \leq a$, entonces se dice que el orden $\leq$ es *total* y que $(A, \leq)$ es un *conjunto totalmente ordenado*. Por ejemplo el conjunto de los números reales con el orden usual es un conjunto totalmente ordenado. Por otro lado, dado un conjunto arbitrario A, su potencia $\mathcal{P}(A)$ es un conjunto parcialmente ordenado con el orden dado por la inclusión, es decir, definiendo $X \leq Y$ cuando $X \subset Y$. Este último orden en general no es un orden total (depende del conjunto A).

Supongamos que $(A, \leq)$ es un conjunto parcialmente ordenado. Veremos algunas definiciones.

- Si $a, b \in A$ verifican $a \leq b$ y $a \neq b$, entonces escribimos $a < b$.
- Un elemento $a \in A$ es *maximal* si no existe $b \in A$ tal que $a < b$.
- Si B es un subconjunto de A, entonces un elemento $a \in A$ es una *cota superior* de B si a verifica $b \leq a$, para todo $b \in B$.
- Una *cadena* en A es un subconjunto B de A tal que B es un conjunto totalmente ordenado, con el orden de A a restringido a B .

Con las definiciones anteriores estamos en condiciones de enunciar el siguiente.

Lema de Zorn: si en un conjunto parcialmente ordenado se cumple que toda cadena tiene una cota superior, entonces el conjunto tiene algún elemento maximal.

El lema de Zorn no se puede deducir de los otros axiomas de la teoría de conjuntos, pero se puede probar introduciendo el *axioma de elección*, aunque en realidad son equivalentes. Se utiliza para probar propiedades en conjuntos infinitos, por ejemplo, para demostrar que todo espacio vectorial admite una base.

Relaciones de equivalencia. Sea A un conjunto. Una relación $R \subset A \times A$ es *de equivalencia* si verifica las siguiente propiedades.

- *Reflexiva*: $(a, a) \in R$, para todo $a \in A$.
- *Simétrica*: si $(a, b) \in R$, entonces $(b, a) \in R$.
- *Transitiva*: si $(a, b) \in R$ y $(b, c) \in R$, entonces $(a, c) \in R$.

En este caso se suele usar el símbolo $\sim$ para la relación R y escribir $a \sim b$ para $(a, b) \in R$. Con esta notación las propiedades anteriores se escriben:

$$a \sim a,\ \forall a \in A; \quad a \sim b \Rightarrow b \sim a; \quad a \sim b \text{ y } b \sim c \ \Rightarrow \ a \sim c.$$

Supongamos que $\sim$ es una relación de equivalencia en A. Si $a \in A$, entonces la *clase de equivalencia* de a es el conjunto $\overline{a} := \{x \in A : \ x \sim a\}$. Al conjunto de las clases de equivalencia $A/{\sim} := \{\overline{a} : \ a \in A\}$ se le llama el *conjunto cociente* de A por la relación $\sim$. El conjunto cociente $A/{\sim}$ es un subconjunto de $\mathcal{P}(A)$.
Las relaciones de equivalencia en A están directamente relacionadas con las particiones de A. Una *partición* de un conjunto A es un subconjunto $\mathcal{X}$ de $\mathcal{P}(A)$ que verifica las siguientes condiciones:

$$\text{si } X \in \mathcal{X} \Rightarrow X \neq \emptyset; \quad \text{si } X, Y \in \mathcal{X} \text{ y } X \cap Y \neq \emptyset \Rightarrow X = Y; \quad \bigcup_{X \in \mathcal{X}} X = A.$$

Observar que si $a, b \in A$ y $\overline{a} \cap \overline{b} \neq \emptyset$, entonces es $a \sim b$ y por lo tanto $\overline{a} = \overline{b}$. Además si $a \in A$, entonces es $a \in \overline{a}$, luego $\overline{a} \neq \emptyset$, para todo $a \in A$, y también es claro que vale $\bigcup_{a \in A} \overline{a} = A$. Esto prueba que el conjunto cociente $A/{\sim}$ es una partición de A. Recíprocamente, si $\mathcal{X} \subset \mathcal{P}(A)$ es una partición de A y definimos una relación $\sim$ en A mediante $a \sim b$ si y solo si existe $X \in \mathcal{X}$ tal que $a, b \in X$, entonces es fácil de probar que $\sim$ es una relación de equivalencia en A. Luego existe una correspondencia uno a uno entre relaciones de equivalencia en A y particiones de A.

Funciones. La idea de función entre dos conjuntos, es la de una correspondencia que asocia a cada elemento de uno de ellos un único elemento del otro. Lo que sigue es una manera de formalizar esa idea. Una *función* es una terna (X, Y, f) en la cual X e Y son dos conjuntos y $f \subset X \times Y$ es una relación que verifica que para todo $x \in X$ existe un único $y \in Y$ tal que $(x, y) \in f$.

La notación $f : X \to Y$ quiere decir que (X, Y, f) es una función; en ese caso decimos que f es una función *de X en Y* y escribimos $y = f(x)$ en lugar de $(x, y) \in f$. Si $f : X \to Y$ es una función, entonces X es su *dominio* e Y es su *codominio*. La *imagen* o *recorrido* de f es el conjunto $\operatorname{Im}(f) := \{y \in Y : \ \exists\, x \in X \text{ tal que } y = f(x)\}$.

En todo conjunto X tenemos la función *identidad* $\mathrm{id}_X : X \to X$, definida por $\mathrm{id}_X(x) = x$, para todo $x \in X$.

Si $f : X \to Y$ y $g : Y \to Z$ son dos funciones, entonces su *composición* es la función $g \circ f : X \to Z$ definida por $(g \circ f)(x) = g\big(f(x)\big)$, para todo $x \in X$. La composición verifica las siguientes propiedades.

- Si $f : X \to Y$ es una función, entonces $\mathrm{id}_Y \circ f = f \circ \mathrm{id}_X = f$.
- La composición es *asociativa*, es decir si tenemos funciones $f : X \to Y$, $g : Y \to Z$ y $h : Z \to T$, entonces $(h \circ g) \circ f = h \circ (g \circ f)$.

A continuación veremos algunas definiciones. Sea $f : X \to Y$ una función.

- f es *inyectiva*, si $x_1 \neq x_2$ implica $f(x_1) \neq f(x_2)$; esto equivale a $f(x_1) = f(x_2)$ implica $x_1 = x_2$.
- f es *sobreyectiva*, si para todo $y \in Y$ existe $x \in X$ tal que $f(x) = y$; esto equivale a $\operatorname{Im}(f) = Y$.

- f es *biyectiva*, si es inyectiva y sobreyectiva.

A continuación veremos cómo se relacionan estos conceptos con la composición. Sean $f : X \to Y$ y $g : Y \to Z$ funciones y consideremos su composición $g \circ f : X \to Z$. Entonces

- Si f y g son inyectivas o sobreyectivas, entonces $g \circ f$ es inyectiva o sobreyectiva, respectivamente.
- Si $g \circ f$ es inyectiva, entonces f es inyectiva y si $g \circ f$ es sobreyectiva, entonces g es sobreyectiva.

  Estas propiedades implican las siguientes.

- Si f y g son biyectivas, entonces $g \circ f$ es biyectiva.
- Si $g \circ f$ es biyectiva, entonces g es sobreyectiva y f es inyectiva.

Invertibilidad. Decimos que una función $f : X \to Y$ es *invertible* si existe una función $g : Y \to X$ tal que

$$g\big(f(x)\big) = x,\ f\big(g(y)\big) = y, \quad \forall x \in X,\ y \in Y \quad \Leftrightarrow \quad g \circ f = \mathrm{id}_X,\ f \circ g = \mathrm{id}_Y. \tag{A.1}$$

Notar que si otra función $h : Y \to X$ verifica lo mismo que g, entonces

$$h = h \circ \mathrm{id}_Y = h \circ (f \circ g) = (h \circ f) \circ g = \mathrm{id}_X \circ g = g \quad \Rightarrow \quad h = g.$$

Luego si existe una función g que verifique (A.1), entonces g es única. En ese caso a g se le llama la *inversa* de f y se escribe $f^{-1} = g$. Notar que $f^{-1} : Y \to X$ queda caracterizada por verificar

$$f^{-1} \circ f = \mathrm{id}_X, \quad f \circ f^{-1} = \mathrm{id}_Y. \tag{A.2}$$

Si f es invertible, como id_X e id_Y son biyectivas, entonces de (A.2) se deduce que f es biyectiva. Recíprocamente, si f es biyectiva, entonces para cada $y \in Y$ existe un único $x \in X$ tal que $f(x) = y$, eso nos permite definir una función $g : Y \to X$ mediante $g(y) = x$. Notar que de acuerdo a esta definición es

$$g\big(f(x)\big) = x,\ f\big(g(y)\big) = y, \quad \forall x \in X,\ y \in Y.$$

Eso nos dice que f es invertible y $f^{-1} = g$. En resumen, probamos que una función $f : X \to Y$ es invertible si y solo si es biyectiva, y en ese caso su inversa se obtiene como acabamos de ver.

Notar que (A.2) implica que si f es invertible, entonces f^{-1} es invertible y su inversa es f, es decir $\left(f^{-1}\right)^{-1} = f$. A su vez esto implica que si f es biyectiva, entonces f^{-1} también es biyectiva. Respecto a la composición, si $f : X \to Y$ y $g : Y \to Z$ son funciones biyectivas, entonces es fácil de probar que $g \circ f$ es biyectiva y vale $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Decimos que dos conjuntos X e Y son *equipotentes* si existe una función biyectiva $f : X \to Y$. Notar que la relación entre conjuntos de "ser equipotentes" es de equivalencia. Si X e Y son conjuntos finitos, entonces son equipotentes si y solo si tienen la misma cantidad de elementos.

Finalmente notar que si $\sim$ es una relación de equivalencia en un conjunto A y $A/\sim$ es el conjunto cociente, entonces existe una función sobreyectiva $\pi : A \to A/\sim$ definida por $\pi(a) = \overline{a}$, para todo $a \in A$. Esta función π se llama la *proyección canónica* de A en $A/\sim$.