ÁLGEBRA II / ÁLGEBRA - NOTAS DEL TEÓRICO

SILVINA RIVEROS, ALEJANDRO TIRABOSCHI Y AGUSTÍN GARCÍA IGLESIAS.

Año 2021
FAMAF - UNC

ÍNDICE GENERAL

ÍNDICE DE FIGURAS

PREFACIO

Las siguientes notas se han utilizado para el dictado del curso "Álgebra II / Álgebra" del primer año de las licenciaturas y profesorados de FAMAF. Han sido las notas principales en el dictado del año 2018 y 2020, y se limitan casi exclusivamente al contenido dictado en el curso. Las partes señaladas con (*) y los apéndices son optativos.

Estas notas están basadas principalmente en *Apuntes de Álgebra II - Año 2005* de Silvina Riveros y han sido revisadas, modificadas y ampliadas por Alejandro Tiraboschi y Agustín García Iglesias.

También hemos utilizado como bibliografía de apoyo los siguientes:

- *Álgebra Lineal.* Autores: Gabriela Jerónimo, Juan Sabia y Susana Tesauri. Año 2008. Puede descargarse del Departamento de Matemática de la UBA, en la dirección

  http://mate.dm.uba.ar/~jeronimo/algebra_lineal/AlgebraLineal.pdf

- *Linear Algebra.* Autores: Jim Hefferon. Se descarga en

  http://joshua.smcvt.edu/linearalgebra/

Otra bibliografía de interés:

- *Álgebra Lineal.* Autor: Serge Lang. Año: 1976. Editorial: Fondo Educativo Interamericano. Ficha en biblioteca de FAMAF: http://bit.ly/2stBTzs

- *Álgebra Lineal.* Autores: Kenneth Hoffman y Ray Kunze. Año: 1973. Editorial: Prentice Hall. Ficha en biblioteca de FAMAF: http://bit.ly/2tn3eRc

Contenidos mínimos

Resolución de ecuaciones lineales. Matrices. Operaciones elementales. Matriz inversa. Espacios vectoriales sobre $\mathbb{R}$ y $\mathbb{C}$. Subespacios. Independencia lineal. Bases y dimensión Rectas y planos en $\mathbb{R}^n$. Transformaciones lineales y matrices. Isomorfismos. Cambio de bases. Núcleo e imagen de transformaciones lineales. Rango fila y columna. Determinante de una matriz. Cálculo y propiedades básicas. Espacios con producto interno. Desigualdad de Cauchy-Schwartz. Desigualdad triangular. Teorema de Pitágoras. Ortonormalización de Gram-Schmidt. Ecuaciones de rectas y planos en $\mathbb{R}^n$. Distancias. Introducción a vectores y valores propios. Aplicaciones. Diagonalización de matrices simétricas.

Parte I

VECTORES Y SISTEMAS LINEALES EN $\mathbb{R}^n$

1 VECTORES

El concepto de vector es básico para el estudio de funciones de varias variables y proporciona la motivación geométrica para todo el curso. Por lo tanto, las propiedades de los vectores, tanto algebraicas como geométricas, serán discutidas en forma resumida en este capítulo.

1.1 Álgebra lineal en $\mathbb{R}^2$ y $\mathbb{R}^3$

Sabemos que se puede usar un número para representar un punto en una línea, una vez que se selecciona la longitud de una unidad.

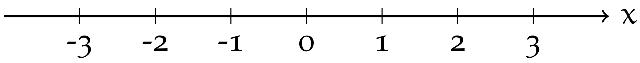


Figura 1: La recta real y algunos números enteros.

Se puede usar un par de números (x, y) para representar un punto en el plano. Estos pueden ser representados como en la figura 2.

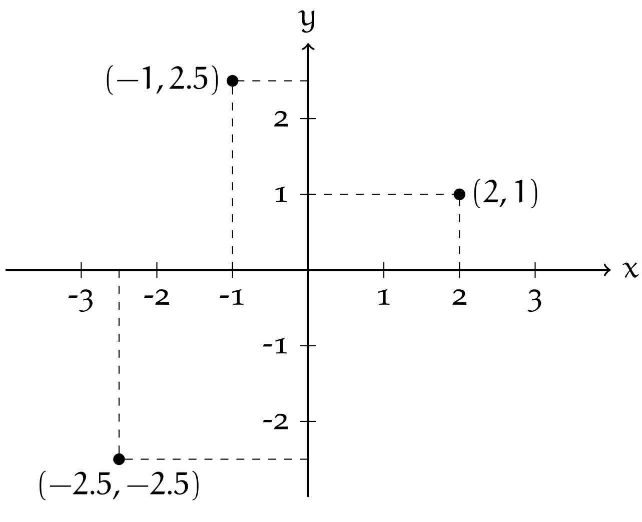


Figura 2: Representación gráfica de los puntos $(2, 1)$, $(-1, 2.5)$ y $(-2.5, -2.5)$ en $\mathbb{R}^2$.

Ahora observamos que un triple de números (x, y, z) se puede usar para representar un punto en el espacio, es decir, espacio tridimensional, o 3-espacio. Simplemente, introducimos un eje más. La figura 3 ilustra esto.

En lugar de usar (x, y, z), también suele usarse la notación (x_1, x_2, x_3). La línea podría llamarse el 1-espacio, y el plano podría llamarse el 2-espacio.

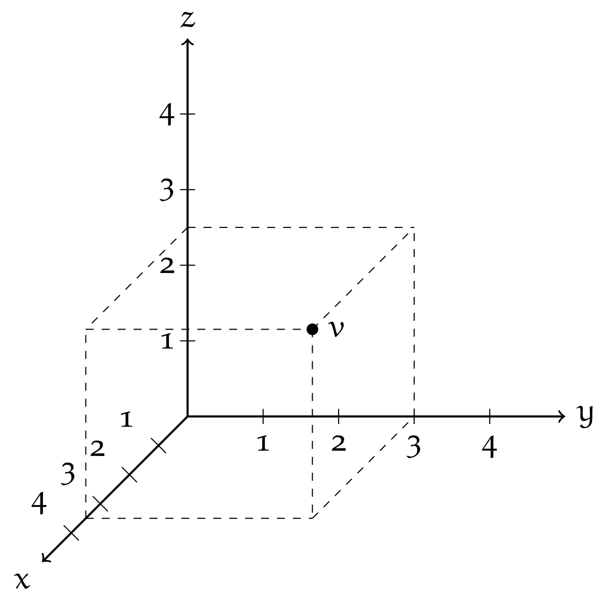


Figura 3: Representación gráfica del punto $v = (3.5, 3, 2.5)$ en $\mathbb{R}^3$.

Por lo tanto, podemos decir que un solo número representa un punto en el 1-espacio. Un par representa un punto en el 2-espacio. Un triple representa un punto en el 3-espacio.

Aunque no podemos hacer un dibujo para generalizar lo anterior a 4-espacios, no hay nada que nos impida considerar un cuádruple de números y decretar que este es un punto en el 4-espacio. Un quíntuple sería un punto en el 5-espacio, luego vendría un séxtuple, séptuple, óctuple, etc.

Podemos generalizar y definir un punto en el n-espacio, para n un entero positivo, como una n-tupla de números. Vamos a denotar tal n-tupla con letras v, w, u, ... y usaremos otras letras minúsculas para los números. Si $v = (x_1, x_2, \dots, x_n)$, llamamos a los números $x_1, x_2, \dots, x_n$ las *coordenadas* del punto v. Más precisamente, x_i será la *coordenada* i*-ésima* de v. Por ejemplo, en el 3-espacio, 2 es la primera coordenada del punto $(2, 3, -4)$, y -4 es la tercera coordenada. Denotamos a los n-espacios por $\mathbb{R}^n$. Para formalizar:

Definición 1.1.1. Sea $\mathbb{R}$ el cuerpo de los números reales, entonces

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

Todo v en $\mathbb{R}^n$ será llamado *punto*. Alternativamente, también podemos decir que v es un *vector en el origen* o simplemente un *vector*.

Observación. Debido a que separamos las coordenadas de un vector con comas, no es conveniente utilizar la notación española que inicia con coma la parte decimal de un número. Por ejemplo, en este apunte a "dos coma cuatro" lo escribiremos "2.4" y así para cualquier número real.

La mayoría de nuestros ejemplos tendrán lugar cuando $n = 2$ o $n = 3$. Por lo tanto, el lector puede visualizar cualquiera de estos dos casos a lo largo del apunte. Para ello usaremos el *sistema de coordenadas cartesianas* para representar los elementos de $\mathbb{R}^2$ y $\mathbb{R}^3$, tal como se ha hecho en las figuras 2 y 3.

Ejemplo. Un ejemplo clásico de 3-espacio es, por supuesto, el espacio en el que vivimos. Después de seleccionar un origen y un sistema de coordenadas, podemos describir la posición de un punto (cuerpo, partícula, etc.) mediante 3 coordenadas. Además, como se sabía hace mucho tiempo, es conveniente extender este espacio a un espacio de 4 dimensiones, donde la cuarta coordenada es el tiempo, seleccionándose el origen del tiempo, por ejemplo, como el nacimiento de Cristo, aunque esto es puramente arbitrario. Entonces, un punto con coordenada de tiempo negativo es un punto antes de Cristo, y un punto con coordenada de tiempo positiva es un punto después de Cristo.

Sin embargo, no es que obligatoriamente "el tiempo es la cuarta dimensión". El espacio 4-dimensional anterior es solo un ejemplo posible. Hagamos un ejemplo relacionado a la economía: tomamos como coordenadas la cantidad de dinero gastado por una industria a lo largo de un año. Por ejemplo, podríamos tener un espacio de 6 dimensiones con coordenadas correspondientes a las siguientes industrias: 1. acero, 2. automotriz, 3. productos agrícolas, 4. productos químicos, 5. indumentaria y 6. transporte. Las coordenadas de las 6-tuplas representarían el gasto anual de las industrias correspondientes. Por ejemplo,

$$(1000, 800, 550, 300, 700, 200)$$

significaría que la industria del acero gastó 1000 en un año determinado, la automotriz 800, etc.

También podemos visualizar los 3-espacios como "productos de espacios de dimensiones inferiores". Por ejemplo, podemos ver las coordenadas de los 3-espacios como dos coordenadas en un 2-espacio acompañada por una coordenada en el 1-espacio. Esto es, (x, y, z) indica el mismo punto que $((x, y), z)$. Esto se escribe como

$$\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}^1.$$

Utilizamos el signo del producto, que no debe confundirse con otros "productos", como el producto de los números. Del mismo modo, podemos escribir

$$\mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}^1.$$

Hay otras formas de expresar $\mathbb{R}^4$ como un producto, a saber:

$$\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2.$$

Esto significa que al punto $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ lo podemos describir por el par ordenado $((x_1, x_2), (x_3, x_4)) \in \mathbb{R}^2 \times \mathbb{R}^2$.

En general, dado $n > 1$, y n_1, n_2 tal que $n_1 + n_2 = n$, tenemos

$$\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}.$$

De forma más general aún, dado $n > 1$, y $n_1, \ldots, n_k$ tal que $n_1 + \cdots + n_k = n$, tenemos

$$\mathbb{R}^n = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k}.$$

Ahora vamos a definir cómo sumar los puntos de $\mathbb{R}^n$. Si v, w son dos puntos, digamos en el 2-espacio, definimos $v + w$ como el punto cuyas coordenadas son la suma de cada coordenada. Es decir, si, por ejemplo, $v = (1, 2)$ y $w = (-3, 5)$, entonces $v + w = (-2, 7)$. En 3-espacios la definición es análoga. Por ejemplo, si $v = (-1, y, 3)$ y $w = (x, 7, -2)$, entonces $v + w = (x - 1, y + 7, 1)$, con $x, y \in \mathbb{R}$.

En dos y tres dimensiones podemos definir Dados $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$ o $(x_1, x_2, x_3), (y_1, y_2, y_3) \in \mathbb{R}^3$, definimos

- $(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2)$,
- $(x_1, x_2, x_3) + (y_1, y_2, y_3) := (x_1 + y_1, x_2 + y_2, x_3 + y_3)$.

Generalizando,

Definición 1.1.2. Si $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$, definimos la suma de los dos vectores como:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n),$$

Observemos que se satisfacen las siguientes propiedades: sean v, w, u en $\mathbb{R}^n$, entonces

S1. $v + w = w + v$ (*conmutatividad de la suma*),

S2. $(v + w) + u = v + (w + u)$ (*asociatividad de la suma*),

S3. si definimos

$$0 = (0, \ldots, 0),$$

el punto cuyas coordenadas son todas 0, el *vector cero*, entonces

$$v + 0 = 0 + v = v,$$

(*existencia de elemento neutro de la suma*).

S4. si $v = (x_1, \ldots, x_n)$, definimos $-v = (-x_1, \ldots, -x_n)$. Entonces

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

(*existencia de opuesto* o *inverso aditivo*).

Estas propiedades se deducen casi trivialmente de la definición de suma, coordenada a coordenada, y de la validez de las propiedades en el caso de la recta real. Como es usual en otros contextos ya conocidos, si $v, w \in \mathbb{R}^n$, entonces denotamos $v - w := v + (-w)$.

Ejemplo. Vimos al final del ejemplo de la p. 7 que una n-tupla puede representar cuestiones relacionadas con las finanzas. En nuestro ejemplo una 6-tupla representaba el gasto anual de determinadas actividades económicas, por ejemplo los gastos en los años 2000 y 2001 son

$$\begin{array}{lcl} 2000 & \to & (1000, 800, 550, 300, 700, 200) \\ 2001 & \to & (1200, 700, 600, 300, 900, 250) \end{array}$$

Luego los costos totales en los dos años son

$$\begin{aligned} &(1000, 800, 550, 300, 700, 200) + (1200, 700, 600, 300, 900, 250) = \\ &= (1000 + 1200, 800 + 700, 550 + 600, 300 + 300, 700 + 900, 200 + 250) \\ &= (2200, 1500, 1350, 600, 1600, 450). \end{aligned}$$

En el ejemplo anterior es claro que la suma de puntos se corresponde con lo que nosotros esperamos que ocurra.

En el plano y en el espacio la suma se puede hacer en forma "geométrica". Veamos ahora hagamos una interpretación geométrica de la suma en el plano.

En álgebra lineal a veces resultará conveniente pensar a cada punto como un *vector* que comienza en el origen. Los vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$ se pueden graficar como "flechas" que parten del origen y llegan a las coordenadas del punto. Veamos en los siguientes ejemplos que está interpretación es útil.

Ejemplo. Sea $v = (2, 3)$ y $w = (-1, 1)$. Entonces $v + w = (1, 4)$. En el dibujo de los puntos involucrados aparece un *paralelogramo* (fig. 4)

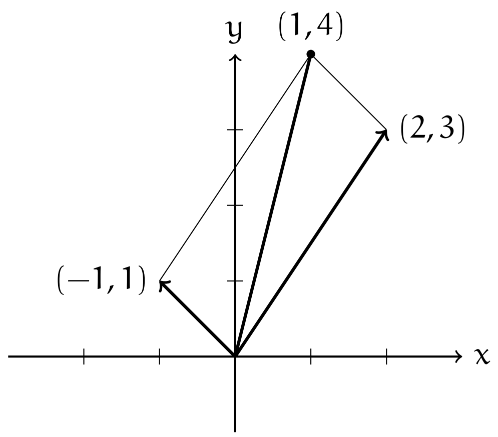


Figura 4: Ejemplo de la ley del paralelogramo.

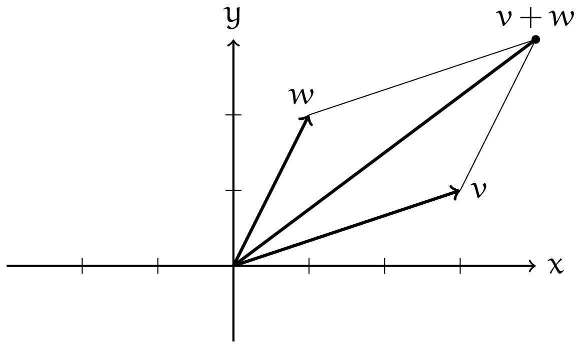


Figura 5: La ley del paralelogramo.

Ejemplo. Sea $v = (3, 1)$ y $w = (1, 2)$. Entonces

$$v + w = (4, 3).$$

Esta suma la representamos en la fig. 5.

Vemos de nuevo que en la representación geométrica aparece un paralelogramo. La razón por la cual la figura que aparece es un paralelogramo se puede dar en términos de la geometría plana de la siguiente manera. Obtenemos $v = (1, 2)$ comenzando desde el origen $0 = (0, 0)$, y moviéndonos 1 unidad hacia la derecha y 2 hacia arriba. Para obtener $v + w$, comenzamos desde v, y de nuevo nos movemos 1 unidad a la derecha y 2 hacia arriba. Así, el segmento entre o y w, y entre v y $v + w$ son las hipotenusas de los triángulos rectángulos cuyos catetos correspondientes son de la misma longitud y paralelos. Los segmentos anteriores son por lo tanto paralelos y de la misma longitud, como se ilustra en la fig. 6. Esta forma geométrica de visualizar la suma de dos vectores en $\mathbb{R}^2$ es conocida como *ley del parelogramo.*

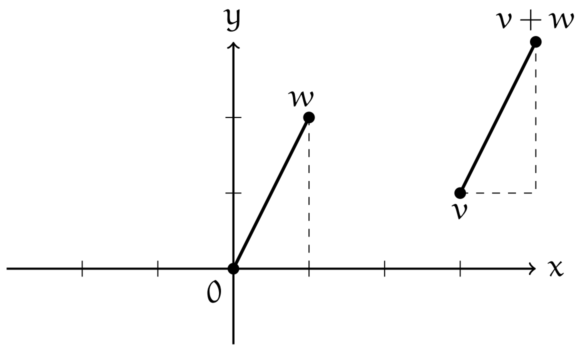


Figura 6: La ley del paralelogramo.

Ejemplo. Sea el punto $v = (3, 1)$, entonces $-v = (-3, -1)$. Si dibujamos v y $-v$ vemos que $-v$ es un vector del mismo "tamaño" que v pero con la dirección opuesta. Podemos ver a $-v$ como la reflexión de v a través del origen (fig. 7).

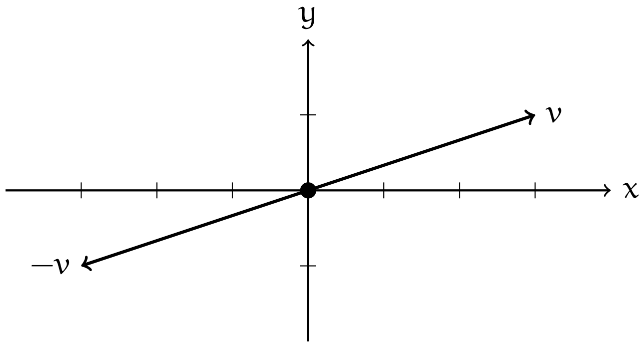


Figura 7: El opuesto de un vector.

La resta de dos vectores también se puede representar geométricamente: restemos al vector v el vector w. Como primera opción podemos encontrar el vector $-w$ y sumarlo a v aplicando la ley del paralelogramo. Esto es equivalente a lo siguiente: los vectores v y w determinan el triángulo determinado por los puntos 0, v y w. Entonces, el lado determinado por w y v, en ese sentido, trasladado al origen es el vector $v - w$ (fig. 8). Claramente, esta forma geométrica de hacer la resta es de nuevo una aplicación de la ley del paralelogramo, pues $(v - w) + w = v$.

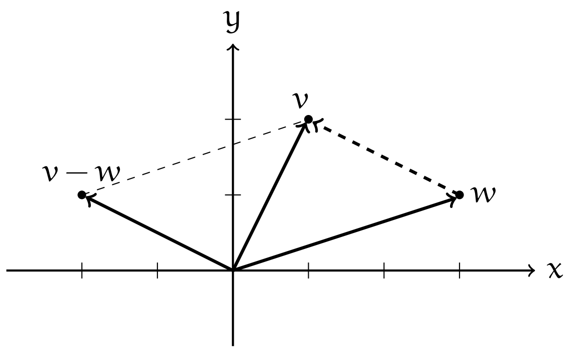


Figura 8: Resta de vectores.

Ahora consideraremos la multiplicación de un vector v por un número.

Definición 1.1.3. Sea $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$, entonces

$$\lambda.v = (\lambda x_1, \ldots, \lambda x_n).$$

También denotamos a esta multiplicación por λv.

Ejemplo. Si $v = (2, -1, 5)$ y $\lambda = 7$, entonces $\lambda v = (14, -7.35)$.

Es fácil verificar las siguientes reglas: dados $v, w \in \mathbb{R}^n$,

P1. $1.v = v$.

P2. $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v$, para todo $\lambda_1, \lambda_2 \in \mathbb{R}$.

D1. $\lambda(v+w) = \lambda v + \lambda w$, para todo $\lambda \in \mathbb{R}$ (*propiedad distributiva*).

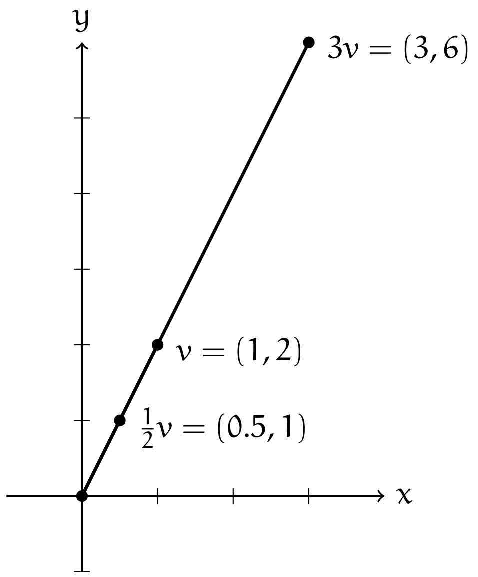
D2. $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ para todo $\lambda_1, \lambda_2 \in \mathbb{R}$ (*propiedad distributiva*).

También tengamos en cuenta que

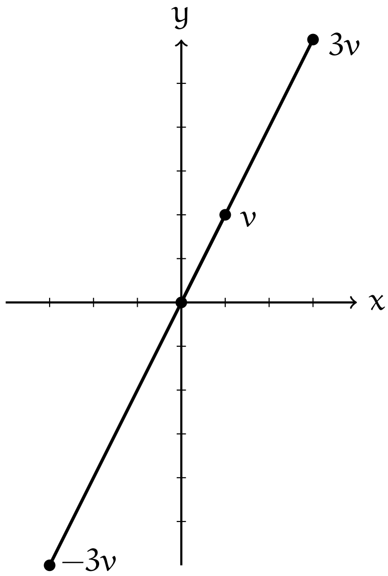
$$(-1)v = -v.$$

¿Cuál es la representación geométrica de la multiplicación de un vector por un número?

Ejemplo. Sea $v = (1,2)$ y $\lambda = 3$. Luego $\lambda v = (3,6)$ como en la siguiente figura:



(a)



(b)

La multiplicación por 3 equivale a "estirar" v por 3. Del mismo modo, $\frac{1}{2}v$ equivale a estirar v en $\frac{1}{2}$, es decir, reducir v a la mitad de su tamaño. En general, si t es un número con $t > 0$, interpretamos tv como un punto en la misma dirección que v con tamaño t-veces el tamaño de v. De hecho, decimos que v y w tienen la misma dirección si existe un número $\lambda > 0$ tal que $v = \lambda w$. La multiplicación por un número negativo invierte la dirección. Así, $-3v$ se representa como en la figura anterior, en la parte (b). Decimos que v y w (ninguno de los cuales es cero) tienen direcciones opuestas si existe un número $\lambda < 0$ tal que $v = \lambda w$. Por lo tanto, $-v$ tiene dirección opuesta a v.

Más allá de las interpretaciones geométricas, hemos definido en forma algebraica la suma de vectores en $\mathbb{R}^n$ y la multiplicación de un vector por un escalar, y estas operaciones tienen ciertas propiedades de interés.

Concluyendo, las definiciones y resultados más importantes de esta sección son:

Sean $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$, definimos

- $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$,

- $\lambda.v := (\lambda x_1, \ldots, \lambda x_n)$.

Dados v, w, u en $\mathbb{R}^n$, se verifican

S1. $v + w = w + v$ (*conmutatividad de la suma*),

S2. $(v + w) + u = v + (w + u)$ (*asociatividad de la suma*),

S3. sea $0 := (0, \ldots, 0)$, el *vector cero*, entonces $0 + v = v + 0 = v$ (*existencia de elemento neutro de la suma*).

S4. Si $v = (x_1, \ldots, x_n)$, entonces $-v := (-x_1, \ldots, -x_n)$ y se satisface $v + (-v) = (-v) + v = 0$ (*existencia de opuesto* o *inverso aditivo*).

P1. $1.v = v$.

P2. $\lambda_1(\lambda_2 v) = (\lambda_1\lambda_2)v$, para todo $\lambda_1, \lambda_2 \in \mathbb{R}$.

D1. $\lambda(v + w) = \lambda v + \lambda w$, para todo $\lambda \in \mathbb{R}$ (*propiedad distributiva*).

D2. $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ para todo $\lambda_1, \lambda_2 \in \mathbb{R}$ (*propiedad distributiva*).

Verán más adelante que las propiedades anteriores son muy parecidas a los "axiomas" que se utilizan en el capítulo 3 para definir espacios vectoriales abstractos (ver definición 3.1.1).

Definición 1.1.4. Dado, $n \in \mathbb{N}$, para cada $i \in \{1, ..., n\}$, se denota $e_i \in \mathbb{R}^n$ al vector cuyas coordenadas son todas 0 excepto la coordenada i que es un 1.

$$e_i := (0, ..., 1, ..., 0)$$

El conjunto $\{e_1, ..., e_n\}$ se llama *base canónica* de $\mathbb{R}^n$.

Ejemplo. En $\mathbb{R}^3$ los vectores son $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$

Estos vectores jugarán un rol central en la materia, principalmente, por la siguiente propiedad.

Proposición 1.1.5. *Todo vector de $\mathbb{R}^n$ se escribe como combinación lineal de la base canónica. Explícitamente, si $(x_1, ..., x_n) \in \mathbb{R}^n$ entonces*

$$(x_1, ..., x_n) = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

La demostración es trivial pero por ahora no la haremos.

Ejemplo.

$$\begin{aligned}(1, 2, 3) &= (1, 0, 0) + (0, 2, 0) + (0, 0, 3)\\ &= 1(1, 0, 0) + 2(0, 1, 0) + 3(0, 0, 1)\\ &= 1e_1 + 2e_2 + 3e_3\end{aligned}$$

§ *Ejercicios*

(1) Dados $v = (-1, 2 - 0)$, $w = (2, -3, -1)$ y $u = (1, -1, 1)$, calcular:

   a) $2v + 3w - 5u$,

   b) $5(v + w)$,

   c) $5v + 5w$ (y verificar que es igual al vector de arriba).

1.2 EL PRODUCTO ESCALAR

En 2-espacios, dados dos vectores $v = (x_1, x_2)$ y $w = (y_l, y_2)$, definimos su *producto escalar* como

$$\langle v, w\rangle := x_1 y_1 + x_2 y_2.$$

Para el caso de 3-espacios, sean $v = (x_1, x_2, x_3)$ y $w = (y_l, y_2, y_3)$, entonces el *producto escalar de* v *y* w es

$$\langle v, w\rangle := x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Finalmente, en los n-espacios, generalizamos la definición de la manera obvia:

Definición 1.2.1. Sean $v = (x_1, \dots, x_n)$ y $w = (y_l, \dots, y_n)$ vectores de $\mathbb{R}^n$, el *producto escalar de* v *y* w se define como

$$\langle v, w\rangle := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Es importante notar que este producto es un número real. Por ejemplo, si

$$v = (1, 3, -2) \quad \text{y} \quad w = (-1, 4, -3),$$

entonces

$$\langle v, w\rangle = -1 + 12 + 6 = 17.$$

Por el momento, no le damos una interpretación geométrica a este producto escalar y veremos esto en la sección 1.3. Ahora derivaremos algunas propiedades importantes.

Proposición 1.2.2. *Sean* v, w, u *tres vectores en* $\mathbb{R}^n$, *entonces*

P1. $\langle v, w\rangle = \langle w, v\rangle$.

P2.

$$\langle v, w + u\rangle = \langle v, w\rangle + \langle v, u\rangle = \langle w + u, v\rangle.$$

P3. *Si* λ *es un número, entonces*

$$\langle \lambda v, w\rangle = \lambda\langle v, w\rangle \quad y \quad \langle v, \lambda w\rangle = \lambda\langle v, w\rangle.$$

P4. *Si $v = 0$ es el vector cero, entonces $\langle v, v\rangle = 0$, de lo contrario*

$$\langle v, v\rangle > 0$$

Demostración. Expresemos los tres vectores en coordenadas: $v = (x_1, \dots, x_n)$, $w = (y_1, \dots, y_n)$, $u = (z_1, \dots, z_n)$.

P1.

$$x_1y_1 + x_2y_2 + \cdots + x_ny_n = y_1x_1 + y_2x_2 + \cdots + y_nx_n$$

porque para cualquiera de los dos números x, y, tenemos que $xy = yx$. Esto prueba la propiedad .

Para P2, sea $u = (z_1, \dots, z_n)$. Entonces

$$w + u = (y_1 + z_1, \dots, y_n + z_n)$$

y

$$\begin{aligned}\langle v, w+u\rangle &= \langle (x_1, \dots, x_n), (y_1 + z_1, \dots, y_n + z_n)\rangle \\ &= x_1(y_1 + z_1) + \cdots x_n(y_n + z_n) \\ &= x_1y_1 + x_1z_1 + \cdots x_ny_n + x_nz_n\end{aligned}$$

Reordenando los términos obtenemos

$$\langle v, w+u\rangle = x_1y_1 + \cdots + x_ny_n + x_1z_1 + \cdots + x_nz_n,$$

que no es otra cosa que $\langle v, w\rangle + \langle v, u\rangle$.

Dejamos la propiedad P3 como ejercicio.

Finalmente probemos P4. Observemos que

$$\langle v, v\rangle = x_1^2 + x_2^2 + \cdots + x_n^2. \tag{1.2.1}$$

Como $x_i^2 \geqslant 0$ para todo i, entonces $\langle v, v\rangle \geqslant 0$. Además, es claro que si v tiene todas las coordenadas iguales a 0, entonces $\langle v, v\rangle = 0$. En el caso que $v \neq 0$, entonces, existe algún i tal que $x_i \neq 0$, por lo tanto $x_i^2 > 0$ y por la ecuación (1.2.1), tenemos que $\langle v, v\rangle > 0$. □

Por la propiedad P1 diremos que el producto escalar es *simétrico*, por las propiedades P2 y P3 diremos que es una *forma bilineal* y, finalmente, por la propiedad P4 diremos que es *definido positivo.*

El producto escalar $\langle v, w\rangle$ puede ser igual a 0 para determinados vectores, incluso ambos distintos de 0. Por ejemplo, si $v = (1, 2, 3)$ y $w = (2, 1, -\frac{4}{3})$, entonces

$$\langle v, w\rangle = 2 + 2 - 4 = 0.$$

Definición 1.2.3. Decimos que dos vectores v y w en $\mathbb{R}^n$ son *perpendiculares* u *ortogonales* si $\langle v, w\rangle = 0$. Cuando v y w son ortogonales denotamos $v \perp w$.

Por el momento, no es claro que en el plano la definición anterior coincida con nuestra noción geométrica e intuitiva de perpendicularidad. Esto lo veremos en la siguiente sección. Aquí nos limitaremos a observar un ejemplo.

Ejemplo. En $\mathbb{R}^3$ consideremos los vectores

$$e_1 = (1, 0, 0), \quad e_2 = (0, 1, 0), \quad e_3 = (0, 0, 1),$$

representados en la fig. 9

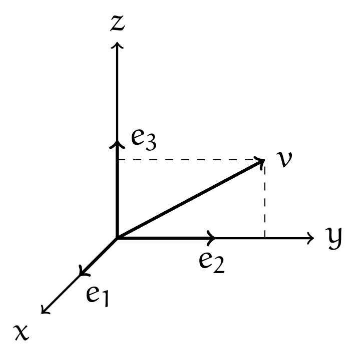


Figura 9: Vectores canónicos en $\mathbb{R}^3$.

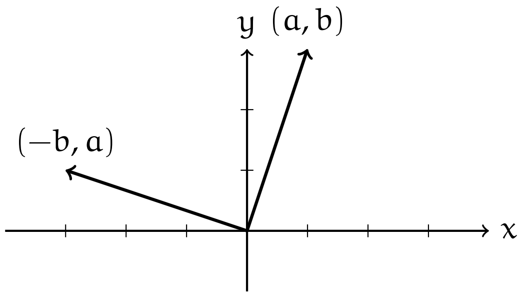
Luego, vemos que $\langle e_i, e_j \rangle = 0$, si $i \neq j$ y por lo tanto e_i es perpendicular a e_j si $i \neq j$, lo cual concuerda con nuestra intuición.

Observemos que si $v = (x_1, x_2, x_3)$, entonces $\langle v, e_i \rangle = x_i$. Por lo tanto, si la coordenada i-ésima de v es cero, v es ortogonal a e_i. Esto nos dice, por ejemplo, que si v es un vector contenido en el plano que incluye e_2 y e_3, es decir si la primera coordenada es cero, entonces v es ortogonal a e_1.

Ejemplo. Sea (a, b) un vector en $\mathbb{R}^2$, entonces $(-b, a)$ es un vector ortogonal a (a, b) debido a que

$$\langle (a, b), (-b, a) \rangle = a \cdot b + (-b) \cdot a = 0.$$

Si graficamos con un ejemplo, $a = 1$, $b = 3$; vemos que esto se corresponde con nuestra intuición de perpendicularidad.



§ *Ejercicios*

(1) Calcular los siguientes productos escalares.

a) $\langle(-1, 2-0), (2, -3, -1)\rangle$,

b) $\langle(4, -1), (-1, 2)\rangle$.

(2) Dados $v = (-1, 2-0)$, $w = (2, -3, -1)$ y $u = (1, -1, 1)$, verificar que:

$$\langle 2v + 3w, -u\rangle = -2\langle v, u\rangle - 3\langle w, u\rangle$$

(3) Sea $v = (x_1, x_2, x_3) \in \mathbb{R}^3$ y sea e_1, e_2 y e_3 la base canónica de $\mathbb{R}^3$ (ver definición 1.1.4). Verificar que

$$v = x_1e_1 + x_2e_2 + x_3e_3 = \langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2 + \langle v, e_3\rangle e_3.$$

(4) Probar, usando sólo las propiedades **P1**, **P2**, y **P3** del producto escalar, que dados $v, w, u \in \mathbb{R}^n$ y $\lambda_1, \lambda_2 \in \mathbb{R}$,

a) se cumple:

$$\langle \lambda_1 v + \lambda_2 w, u\rangle = \lambda_1\langle v, u\rangle + \lambda_2\langle w, u\rangle.$$

b) Si $\langle v, w\rangle = 0$, es decir si v y w son ortogonales, entonces

$$\langle \lambda_1 v + \lambda_2 w, \lambda_1 v + \lambda_2 w\rangle = \lambda_1^2\langle v, v\rangle + \lambda_2^2\langle w, w\rangle.$$

(5) Probar que

a) $(2, 3, -1)$ y $(1, -2, -4)$ son ortogonales.

b) $(2, -1)$ y $(1, 2)$ son ortogonales. Dibujar en el plano.

(6) Encontrar

a) un vector no nulo ortogonal a $(3, -4)$,

b) un vector no nulo ortogonal a $(2, -1, 4)$,

c) un vector no nulo ortogonal a $(2, -1, 4)$ y $(0, 1, -1)$,

1.3 LA NORMA DE UN VECTOR

Si v es vector, entonces $\langle v, v\rangle \geqslant 0$ y definimos como la *norma de* v o *longitud de* v al número

$$||v|| = \sqrt{\langle v, v\rangle}.$$

Cuando v pertenece al plano y $v = (x, y)$, entonces $||v|| = \sqrt{x^2 + y^2}$ y si graficamos el vector en la fig. 10, vemos que la noción de norma o longitud en $\mathbb{R}^2$ se deduce del teorema de Pitágoras.

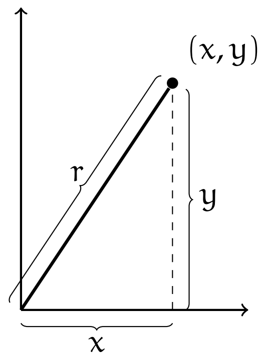


Figura 10: $r = \sqrt{x^2 + y^2}$.

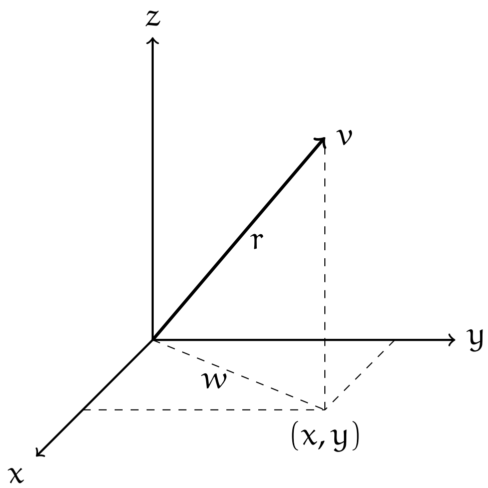


Figura 11: $w = \sqrt{x^2 + y^2}$, $r = \sqrt{w^2 + z^2} = \sqrt{x^2 + y^2 + z^2}$.

Si $n = 3$, el dibujo es como en la fig. 11, para $v = (x, y, z)$. Es decir, por la aplicación reiterada del teorema de Pitágoras obtenemos que la longitud de v es $\sqrt{x^2 + y^2 + z^2}$.

En general, si $v = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, entonces

$$||v|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

y la aplicación reiterada del teorema de Pitágoras nos dice que esta es la definición correcta de longitud o norma de un vector.

Proposición 1.3.1. *Sea $v \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$, entonces*

$$||\lambda v|| = |\lambda| ||v||.$$

Demostración. $||\lambda v||^2 = \langle \lambda v, \lambda v \rangle$, por la propiedad P3 del producto escalar,

$$\langle \lambda v, \lambda v \rangle = \lambda \langle v, \lambda v \rangle = \lambda^2 \langle v, v \rangle.$$

Es decir $||\lambda v||^2 = \lambda^2 ||v||^2$, por lo tanto (sacando raíz cuadrada), $||\lambda v|| = |\lambda| ||v||$. □

El producto escalar no sólo es útil para definir la longitud de un vector, sino que también nos dice cual es el ángulo entre dos vectores: sean $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ en $\mathbb{R}^2$; veremos a continuación que

$$\langle v_1, v_2 \rangle = ||v_1|| \, ||v_2|| \cos(\theta), \tag{1.3.1}$$

donde θ es el ángulo comprendido entre v_1 y v_2.

Sea α_1 el ángulo comprendido entre v_1 y el eje horizontal y α_2 el ángulo comprendido entre v_2 y el eje horizontal. Entonces,

$$v_1 = ||v_1||(\cos(\alpha_1), \operatorname{sen}(\alpha_1)), \quad v_2 = ||v_2||(\cos(\alpha_2), \operatorname{sen}(\alpha_2)),$$

por lo tanto

$$\langle v_1, v_2 \rangle = ||v_1|| \, ||v_2||(\cos(\alpha_1)\cos(\alpha_2) + \operatorname{sen}(\alpha_1)\operatorname{sen}\alpha_2)).$$

Por otro lado, por la propiedad de la suma de los cosenos tenemos que

$$\cos(\alpha_1)\cos(\alpha_2) + \operatorname{sen}(\alpha_1)\operatorname{sen}(\alpha_2) = \cos(\alpha_1 - \alpha_2). \tag{1.3.2}$$

Es decir,

$$\langle v_1, v_2 \rangle = ||v_1|| \, ||v_2|| \cos(\alpha_1 - \alpha_2),$$

y precisamente, $\theta = \alpha_1 - \alpha_2$ es el ángulo comprendido entre v_1 y v_2.

Esto se puede generalizar a $\mathbb{R}^3$ y ahí en vez de la fórmula (1.3.2) se debe usar la ley esférica de los cosenos. Los resultados se puede generalizar a $\mathbb{R}^n$ y en general vale que si $v_1, v_2 \in \mathbb{R}^n$, entonces el ángulo comprendido entre v_1 y v_2 es

$$\theta = \arccos\left(\frac{\langle v_1, v_2 \rangle}{||v_1|| \, ||v_2||}\right). \tag{1.3.3}$$

Terminaremos esta sección dando la noción de distancia entre dos vectores o dos puntos.

Definición 1.3.2. Sea $v, w \in \mathbb{R}^n$, entonce las *distancia* entre v y w es $||v - w||$.

Vemos en la fig. 12 que la norma del vector $v - w$ es la longitud del segmento que une w con v.

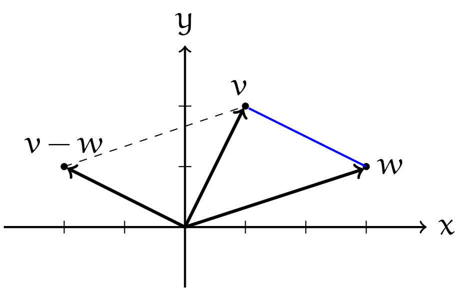


Figura 12: Distancia de v a w.

Una de las desigualdades más notables referentes a la norma de un vector es la *desigualdad triangular*:

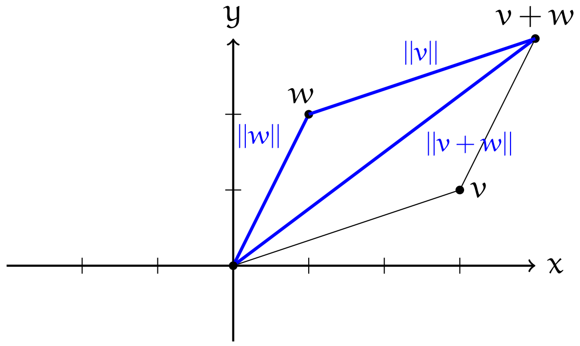
Proposición 1.3.3. *Sean $v, w \in \mathbb{R}^n$, entonces*

$$||v + w|| \leqslant ||v|| + ||w||,$$

y la igualdad se cumple sólo cuando w es múltiplo de v.

Demostración. Ejercicio. □

La desigualdad triangular expresa en forma algebraica el resultado, más conocido, "en todo triángulo, un lado es menor que la suma de los otros dos", que graficamos a continuación.



§ *Ejercicios*

(1) Encontrar la longitud de los vectores.

$$(a)\ (2, 3), \qquad (b)\ (t, t^2), \qquad (c)\ (\cos\phi, \operatorname{sen}\phi).$$

(2) Calcular $\langle v, w\rangle$ y el ángulo entre v y w para los siguientes vectores.

$$(a)\ v = (2, 2), w = (1, 0), \qquad (b)\ v = (-5, 3, 1), w = (2, -4, -7).$$

(3) Dados $v, w, \in \mathbb{R}^n$, probar que si $\langle v, w\rangle = 0$, es decir si v y w son ortogonales, entonces

$$||v + w||^2 = ||v||^2 + ||w||^2.$$

¿Cuál es el nombre con que se conoce este resultado en $\mathbb{R}^2$?

(4) Sean $v, w \in \mathbb{R}^2$, probar usando solo la definición explícita del producto escalar en $\mathbb{R}^2$ que

$$|\langle v, w\rangle| \leqslant ||v||\,||w|| \qquad \text{(Desigualdad de Schwarz).}$$

[Ayuda: elevar al cuadrado y aplicar la definición.]

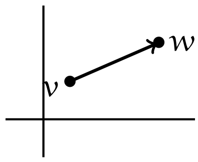


Figura 13: Un vector afín.

1.4 VECTORES AFINES

En esta sección veremos el concepto de vector afín, que nos servirá para entender más geométricamente los conceptos de rectas y planos en $\mathbb{R}^2$ y $\mathbb{R}^3$, respectivamente (secciones 1.5 y 1.6). Definimos un *vector afín* como un par ordenado de puntos v y w, que escribimos $\overrightarrow{vw}$ y lo visualizamos como una flecha entre v y w. Llamamos a v el *punto inicial* y w el *punto final* del vector afín (fig. 13).

Sean $\overrightarrow{vw}$ y $\overrightarrow{pq}$ dos vectores afines. Diremos que son *equivalentes* si $w - v = q - p$.

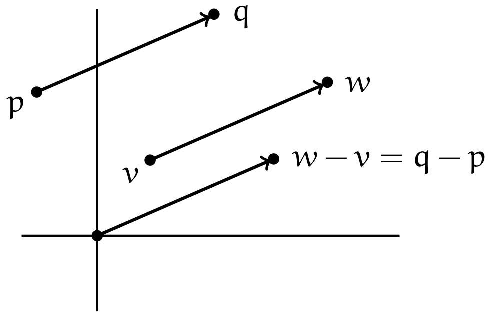


Figura 14: Dos vectores equivalentes.

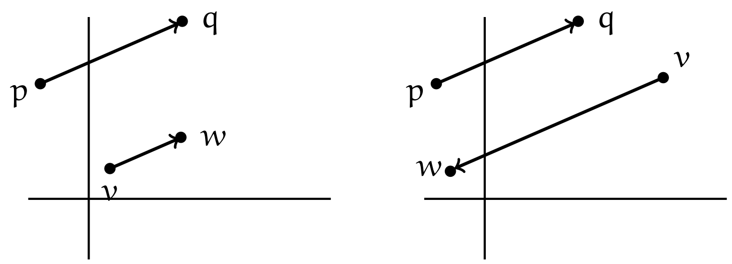
Cada vector afín $\overrightarrow{vw}$ es equivalente a uno cuyo punto de inicial es el origen, pues $\overrightarrow{vw}$ es equivalente a $\overrightarrow{0(w-v)}$ (ver fig. 14).

Claramente este es el único vector cuyo punto inicial es el origen y que es equivalente a $\overrightarrow{vw}$. Si visualizamos la ley del paralelogramo en el plano, entonces está claro que la equivalencia de dos vectores afines se puede interpretar geométricamente diciendo que las longitudes de los segmentos de línea determinadas por el par de puntos son iguales, y que las "direcciones" de los dos vectores son las mismos.

A una n-tupla la podemos interpretar como un vector cuyo punto inicial es el origen. En vista de esto, llamaremos, como lo venimos haciendo, a una n-tupla punto o vector, dependiendo de la interpretación que tenemos en mente.

Se dice que dos vectores afines $\overrightarrow{vw}$ y $\overrightarrow{pq}$ son *paralelos* si hay un número $\lambda \neq 0$ tal que $w - v = \lambda(q - p)$. Se dice que tienen la *misma dirección* si hay un número $\lambda > 0$ tal que $w - v = \lambda(q - p)$, y que tienen *direcciones opuestas* si hay un número $\lambda < 0$ tal que $w - v = \lambda(q - p)$.

En los siguientes dibujos, ilustramos vectores afines paralelos. En el primer dibujo con la misma dirección, en el segundo, con direcciones opuestas.



§ *Ejercicios*

(1) En cada uno de los siguientes casos determinar si los vectores $\overrightarrow{vw}$ y $\overrightarrow{xy}$ son equivalentes y/o paralelos.

a) $v = (1, -1)$, $w = (4, 3)$, $x = (-1, 5)$, $y = (5, 2)$.

b) $v = (1, -1, 5)$, $w = (-2, 3, -4)$, $x = (3, 1, 1)$, $y = (-3, 9, -17)$.

1.5 RECTAS EN $\mathbb{R}^2$

Conocemos de la secundaria y de cursos anteriores el concepto de recta, por ejemplo en el sitio online EcuRed dice:

"Una recta puede ser expresada mediante una ecuación del tipo $y = mx + b$*, donde* x*,* y *son variables en un plano. En dicha expresión* m *es denominada pendiente de la recta y está relacionada con la inclinación que toma la recta respecto a un par de ejes que definen el Plano. Mientras que* b *es el término independiente y es el valor del punto en el cual la recta corta al eje vertical en el plano."*

Dicho en otros términos una recta, según esta definición, es el conjunto de puntos $(x, y) \in \mathbb{R}^2$ que satisfacen la ecuación $y = mx + b$ y puede verse como el gráfico de la función $f(x) = mx + b$. Si, por ejemplo, $m = \frac{1}{2}$ y $b = 1$, podemos dibujar la recta en el plano cerca del origen, como en fig.15.

Sin embargo, con la definición anterior no es posible considerar las rectas verticales. Las rectas verticales están dadas por una ecuación del tipo $x = b$, es decir son todos los puntos (x, y) tal que $x = b$ e y puede tomar cualquier valor. Por ejemplo, la recta $x = 2.5$ se grafica como en la fig. 16.

No es difícil dar una definición que englobe todas las rectas posibles del plano:

Definición 1.5.1 (Definición general de la recta)**.** Sean $a, b, c \in \mathbb{R}$ y tal que a, b no son simultáneamente 0. La *recta con ecuación implícita*

$$ax + by = c, \tag{1.5.1}$$

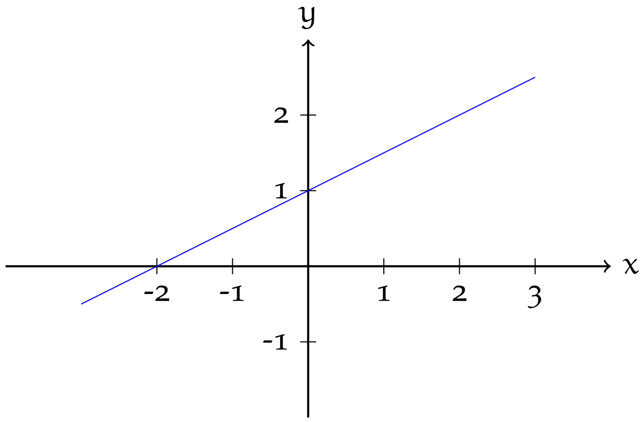


Figura 15: La recta $y = \frac{1}{2}x + 1$.

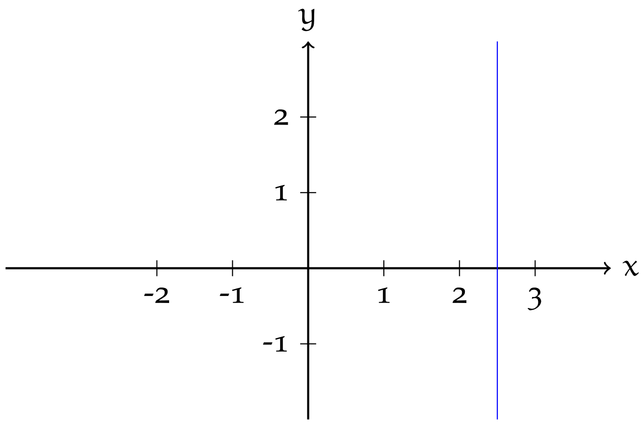


Figura 16: La recta $x = 2.5$.

es el conjunto de puntos (x, y) en $\mathbb{R}^2$ que satisfacen la ecuación (1.5.1). Es decir, si denotamos L a la recta,

$$L = \{(x, y) \in \mathbb{R}^2 : ax + by = c\}.$$

Observar que si $b \neq 0$, entonces la recta es $\; y = -\dfrac{a}{b}x + \dfrac{c}{b}\;$ y que si $b = 0$, entonces $a \neq 0$ y la recta es $\; x = \dfrac{c}{a}$.

Observación. Si consideramos el vector (a, b) en $\mathbb{R}^2$, $c \in \mathbb{R}$ y L la recta definida por los puntos (x, y) tal que $ax + by = c$, entonces L es la recta formada por el conjunto de puntos (x, y) en $\mathbb{R}^2$ que satisfacen

$$\langle (x, y), (a, b) \rangle = c.$$

Ahora bien, consideremos (x_0, y_0) un punto de la recta, entonces, obviamente tenemos que $\langle (x_0, y_0), (a, b) \rangle = c$, por lo tanto la recta se puede describir como los puntos (x, y) que satisfacen la ecuación

$$\langle (x, y), (a, b) \rangle = \langle (x_0, y_0), (a, b) \rangle.$$

Por la propiedad P2 del producto escalar, llegamos a la conclusión que

$$L = \{(x, y) \in \mathbb{R}^2 : \langle (x, y) - (x_0, y_0)\, ,\, (a, b) \rangle = 0\}.$$

Sea $v_0 = (x_0, y_0)$ y $v = (x, y)$, representemos gráficamente la situación:

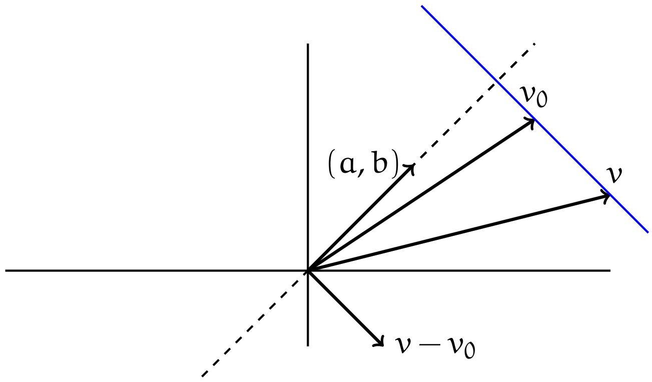


Figura 17: Una recta en el plano.

La recta L es, entonces, la recta perpendicular a (a, b) y que pasa por v_0.

El razonamiento también es posible hacerlo en el otro sentido:

Resultado 1.5.2. La ecuación implícita de la recta L perpendicular a (a, b) y que pasa por (x_0, y_0) es

$$ax + by = \langle (x_0, y_0), (a, b) \rangle.$$

Ejemplo. Encontrar la ecuación implícita de la recta que pasa por $(2, -1)$ y es perpendicular a $(-2, 3)$.

Solución. Por lo visto anteriormente la recta esta formada por los puntos (x, y) tales que

$$-2x + 3y = c$$

y debemos determinar el valor de c. Como $(2, -1)$ pertenece a la recta

$$c = -2 \cdot 2 + 3 \cdot (-1) = -7.$$

Luego, la ecuación implícita de la recta es

$$-2x + 3y = -7.$$

□

Una definición equivalente de recta es la siguiente:

Definición 1.5.3. Sean $v, w \in \mathbb{R}^2$ tal que $w \neq 0$. Sea

$$L = \{v + tw : t \in \mathbb{R}\}.$$

Diremos entonces que L *es la recta que pasa por* v *paralela a* w.

Observemos que la recta L está dada por todos los puntos que se obtienen de la función

$$X(t) = v + tw, \quad \text{para } t \in \mathbb{R}. \tag{1.5.2}$$

En el espacio $\mathbb{R}^2$, diremos que (1.5.2) es la *ecuación paramétrica* o la *representación paramétrica* de la recta L que pasa por el punto v y es paralela a $w \neq 0$.

Podemos representar una recta dada en forma paramétrica como en la figura 18. Cuando damos tal representación paramétrica, podemos pensar

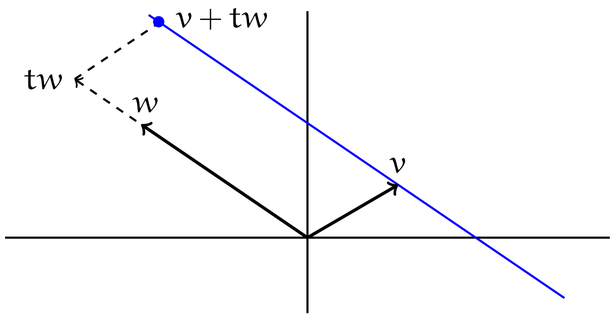


Figura 18: Una recta en el plano.

en un móvil que comienza en el punto v en el tiempo $t = 0$, y moviéndose en la dirección de w. En el momento t, el móvil está en la posición $v + tw$. Por lo tanto, podemos interpretar físicamente la representación paramétrica como una descripción del movimiento, en que w se interpreta como la velocidad del móvil. En un momento dado t, el móvil está en el punto $X(t) = v + tw$ que es llamada la *posición* del móvil en el tiempo t.

Esta representación paramétrica también es útil para describir el conjunto de los puntos que se encuentran en el segmento de línea entre dos puntos dados. Sean v, u dos puntos, entonces el segmento entre v y u consiste en todos los puntos

$$S(t) = v + t(u - v) \quad \text{con} \quad 0 \leqslant t \leqslant 1. \tag{1.5.3}$$

Observar que en tiempo 0, $S(0) = v$ y en tiempo 1, $S(1) = v + (u - v) = u$. Como t "va" de 0 a 1, el móvil va de v a u, en linea recta.

Extendiendo a ambos lados el segmento, podemos describir la recta que pasa por v y u por la ecuación paramétrica (fig. 19)

$$S(t) = v + t(u - v) \quad \text{con} \quad t \in \mathbb{R}.$$

Ejemplo. Encontrar una representación paramétrica para la recta que contiene los puntos $(1, -3, 1)$ y $(-2, 4, 5)$.

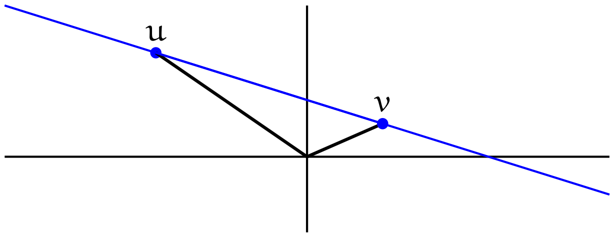


Figura 19: La recta que pasa por v y u.

Solución. Llamemos $v = (1, -3, 1)$ y $u = (-2, 4, 5)$. Entonces $u - v = (-2, 4, 5) - (1, -3, 1) = (-3, 7, 4)$ y la representación paramétrica de la recta que pasa por u y v es

$$X(t) = v + t(u - v) = (1, -3, 1) + t(-3, 7, 4), \quad t \in \mathbb{R}.$$

□

Ahora discutiremos la relación entre una representación paramétrica y la ecuación implícita de una recta en el plano.

Supongamos que trabajamos en el plano y tenemos $v, w \in \mathbb{R}^2$ con $w \neq 0$ y la recta descrita en forma paramétrica:

$$X(t) = v + tw.$$

Sea $v = (x_1, y_1)$, $w = (x_2, y_2)$, entonces, todo punto de la recta es de la forma

$$(x, y) = (x_1, y_1) + t(x_2, y_2) = (x_1 + tx_2, y_1 + ty_2),$$

es decir, los puntos de la recta X son los (x, y) tal que

$$x = x_1 + tx_2, \qquad y = y_1 + ty_2,$$

para $t \in \mathbb{R}$. Dado que $(x_2, y_2) \neq 0$, podemos despejar t de alguna de las ecuaciones y usando la otra ecuación eliminamos t y obtenemos una ecuación implícita. Veremos esto en un ejemplo.

Ejemplo. Sean $v = (2, 1)$ y $w = (-1, 5)$ y sea X la recta que pasa por v en la dirección w. Encontrar la ecuación implícita de L.

Solución. La representación paramétrica de la recta que pasa por v en la dirección de w es

$$X(t) = (2, 1) + t(-1, 5) = (2 - t, 1 + 5t).$$

Es decir, si miramos cada coordenada,

$$x = 2 - t, \qquad y = 1 + 5t. \tag{*}$$

Despejando t de la primera ecuación obtenemos $t = 2 - x$. Reemplazando este valor de t en la segunda ecuación obtenemos $y = 1 + 5t = 1 + 5(2 - x)t = y = 11 - 5x$, luego

$$5x + y = 11, \tag{**}$$

que es la ecuación implícita de la recta.

□

Esta eliminación de t muestra que cada par (x, y) que satisface la representación paramétrica (*) para algún valor de t también satisface la ecuación (**).

Recíprocamente, de la ecuación implícita podemos obtener la representación paramétrica.

Ejemplo. Encontrar la representación paramétrica de la recta definida por

$$5x + y = 11.$$

Solución. Supongamos que tenemos un par de números (x, y) que satisfacen la ecuación implícita, $5x + y = 11$, luego $y = (-5)x + 11$, remplazando x por t (sólo por notación) obtenemos que

$$Y(t) = (t, -5t + 11)$$

es la representación paramétrica de la recta. □

De los ejemplos anteriores se deduce que la recta

$$X(t) = (2 - t, 1 + 5t)$$

es la misma que la recta

$$Y(t) = (t, -5t + 11).$$

Observar que, pese a que hablamos de "la representación paramétrica de la recta", una recta tiene muchas formas de ser representada paramétricamente.

Los procedimientos de los ejemplos anteriores se pueden generalizar a cualquier recta y de esa forma se puede demostrar que la definición paramétrica y la definición implícita de la recta son equivalentes.

Finalmente, podemos obtener la representación paramétrica de la recta a partir de un vector ortogonal a ella y otro vector perteneciente a ella.

Proposición 1.5.4. *Sean* $(a, b), (x_0, y_0) \in \mathbb{R}^2$ *con* $(a, b) \neq 0$. *La recta perpendicular a* (a, b) *que pasa por* (x_0, y_0) *es*

$$L = \{(x_0, y_0) + t(b, -a) \mid t \in \mathbb{R}\}$$

Demostración. El vector $(b, -a)$ es perpendicular a a (a, b) y por lo tanto tiene la dirección de la recta. Luego la ecuación paramétrica de la recta es $v_0 + t(b, -a)$ para algún v_0 en la recta. Como (x_0, y_0) pertenece a la recta, obtenemos el resultado que queríamos probar. □

Ejemplo. Encontrar una representación paramétrica para la recta que contiene los puntos $(2, 2)$ y y es perpendicular a $(2, 1)$.

Solución. El vector ortogonal a $(2, 1)$ es $(1, -2)$. Luego:

$$\begin{aligned} L &= \{(2,2) + t(1,-2) \mid t \in \mathbb{R}\} \\ &= \{(2+t, 2-2t) \mid t \in \mathbb{R}\} \end{aligned}$$

□

Debemos observar que en $\mathbb{R}^3$ no alcanza una sola ecuación lineal del tipo $ax + by + cz = d$ para definir una recta. Veremos en la sección siguiente que una ecuación lineal define un plano en $\mathbb{R}^3$. Genéricamente hablando, con las soluciones de una ecuación en $\mathbb{R}^n$ se obtiene un objeto "con una dimensión menos". Todo esto quedará claro al final de la materia cuando estudiemos subespacios vectoriales de un espacio vectorial.

§ *Ejercicios*

(1) Sea R_1 la recta que pasa por $p_1 = (2, 0)$ y es ortogonal a $(1, 3)$.
   - *a)* Dar la descripción paramétrica e implícita de R_1.
   - *b)* Graficar en el plano a R_1.
   - *c)* Dar un punto p por el que pase R_1 distinto a p_1.
   - *d)* Verificar si $p + p_i$ y $-p$ pertenece a R_1

(2) Repetir el ejercicio anterior con las siguientes rectas.
   - *a)* R_2: recta que pasa por $p_2 = (0, 0)$ y es ortogonal a $(1, 3)$.
   - *b)* R_3: recta que pasa por $p_3 = (1, 0)$ y es paralela a R_1.

(3) Calcular, numérica y graficamente, las intersecciones $R_1 \cap R_2$ y $R_1 \cap R_3$.

(4) Sea $L = \{(x, y) \in \mathbb{R}^2 : ax + by = c\}$ una recta en $\mathbb{R}^2$. Sean p y q dos puntos por los que pasa L.
   - *a)* ¿Para qué valores de c puede asegurar que $(0, 0) \in L$?
   - *b)* ¿Para qué valores de c puede asegurar que $\lambda q \in L$?, donde $\lambda \in \mathbb{R}$.
   - *c)* ¿Para qué valores de c puede asegurar que $p + q \in L$?

(5) Sea L una recta en $\mathbb{R}^2$. Probar que L pasa por $(0,0)$ si y sólo si pasa por $p+\lambda q$ para todo par de puntos distintos p y q de L y para todo $\lambda \in \mathbb{R}$.

1.6 planos en $\mathbb{R}^3$

En la sección anterior vimos (aunque no lo demostramos) que existe una equivalencia entre la definición implícita y la definición paramétrica de la recta. En esta sección definiremos un plano en $\mathbb{R}^3$ utilizando la forma implícita, que es la forma más usual y además es geométricamente intuitiva. Luego veremos la definición del plano en su versión paramétrica .

Comenzaremos, debido a que es más simple, con planos que pasan por el origen, como el de la fig. 20.

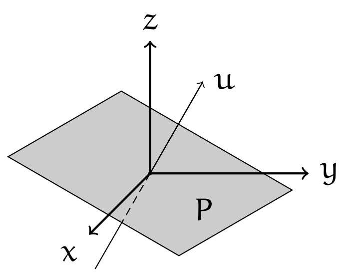


Figura 20: El plano P y u, un vector perpendicular al plano.

En este caso, es claro que el plano está determinado por un vector perpendicular al mismo, es decir si P es un plano que pasa por el origen y u es un punto de $\mathbb{R}^3$, no nulo, tal que $u \perp P$, entonces

$$P = \{v \in \mathbb{R}^3 : \langle v, u\rangle = 0\}.$$

Sea ahora un plano P que no pasa por el origen. Tomo $v_0 \in P$ y entonces observamos que

$$P_0 = \{v - v_0 : v \in P\} \tag{1.6.1}$$

es un plano que pasa por el origen (pues $v_0 - v_0 \in P_0$). Luego, si u perpendicular a P_0 tenemos que

$$P_0 = \{w : \langle w, u\rangle = 0\}. \tag{1.6.2}$$

De las ecuaciones (1.6.1) y (1.6.2) deducimos que

$$v \in P \Leftrightarrow v - v_0 \in P_0 \Leftrightarrow \langle v - v_0, u\rangle = 0,$$

es decir

$$P = \{v \in \mathbb{R}^3 : \langle v - v_0, u\rangle = 0\}.$$

Observemos que $\langle v - v_0, u\rangle = 0$ sii $\langle v, u\rangle - \langle v_0, u\rangle = 0$ sii $\langle v, u\rangle = \langle v_0, u\rangle$. Es decir, si $d = \langle v_0, u\rangle$, tenemos

$$P = \{v \in \mathbb{R}^3 : \langle v, u\rangle = d\}.$$

Esta interpretación geométrica del plano se puede formalizar en la siguiente definición.

Definición 1.6.1. Sean $a, b, c, d \in \mathbb{R}$ tal que $(a, b, c) \neq (0, 0, 0)$ y sea

$$P = \{(x, y, z) : ax + by + cz = d\}.$$

Entonces diremos que P es un *plano con ecuación implícita* $ax + by + cz = d$ y que (a, b, c) es un *vector normal al plano* P. A esta forma de describir el plano también suele llamársela la *ecuación normal del plano.*

Observar que la ecuación $ax + by + cz = d$ no es más que la ecuación $\langle (x, y, z), (a, b, c) \rangle = d$.

Ejemplo. El plano determinado por la ecuación

$$2x - y + 3z = 5$$

es perpendicular al vector $(2, -1, 3)$. Si queremos encontrar un punto en ese plano, por supuesto que tenemos muchas opciones. Podemos dar un valor arbitrario a x e y, y luego despejamos z. Para obtener un punto concreto, sea $x = 1$, $y = 1$. Luego resolvemos para z, a saber

$$3z = 5 - 2 + 1 = 4,$$

luego $z = \dfrac{4}{3}$ y entonces

$$(1, 1, \frac{4}{3})$$

es un punto en el plano.

Se dice que dos planos son *paralelos* (en el 3-espacio) si sus vectores normales son paralelos, es decir son proporcionales. Se dice que son *perpendiculares* si sus vectores normales son perpendiculares. El *ángulo entre dos planos* se define como el ángulo entre sus vectores normales.

Como $(a, b, c) \neq (0, 0, 0)$, entonces una de las tres componentes del vector normal al plano no es cero. Supongamos que $a \neq 0$, luego es fácil despejar x en función de las constantes a, b, c y d; y las variables y y z, por lo tanto cada coordenada del plano depende paramétricamente de y y z y así obtenemos una ecuación paramétrica de P (que depende de 2 parámetros). Se puede hacer de forma análoga cuando $b \neq 0$ o $c \neq 0$.

Ejemplo. Dado el plano $P = \{(x, y, z) : x - 2y + z = 1\}$, hallaremos una ecuación paramétrica de P. Como $x - 2y + z = 1$ sii $x = 2y - z + 1$, tenemos que

$$P = \{(2y - z + 1, y, z) : y, z \in \mathbb{R}\},$$

o, escrito de una forma más estándar,

$$P = \{(2s - t + 1, s, t) : s, t \in \mathbb{R}\}.$$

Observemos que $(2s-t+1, s, t) = (1,0,0)+(2s-t, s, t) = (1,0,0)+(2s, s, 0)+(-t, 0, t) = (1,0,0)+s(2,1,0)+t(-1,0,1)$, por lo tanto, podemos también escribir

$$P = \{(1,0,0)+s(2,1,0)+t(-1,0,1) : s,t \in \mathbb{R}\}.$$

Cualquiera de las formas paramétricas de describir P es correcta, pero la última es la que se utiliza para definir formalmente el plano en forma paramétrica.

Definición 1.6.2. Sean $v, w_1, w_2 \in \mathbb{R}^3$ tal que w_1,w_2 no nulos y tal que w_2 no sea un múltiplo de w_1. Sea

$$P = \{v + sw_1 + tw_2 : s,t \in \mathbb{R}\}.$$

Diremos entonces que P *es el plano a través de* v *paralelo a los vectores* w_1 *y* w_2.

Claramente, en la definición de arriba, el vector v pertenece al plano y el plano

$$P_0 = \{sw_1 + tw_2 : s,t \in \mathbb{R}\}$$

es el plano que pasa por el origen y paralelo a P.

Ya hemos visto que de la ecuación implícita del plano podemos pasar a al ecuación paramétrica fácilmente. Es un poco menos directo pasar de la ecuación paramétrica a la ecuación implícita, pero podemos describir un procedimiento general: sea $P = \{v + sw_1 + tw_2 : s,t \in \mathbb{R}\}$, entonces $v \in P$ y $P_0 = \{sw_1 + tw_2 : s,t \in \mathbb{R}\}$ es el plano paralelo a P que pasa por el origen. Si encontramos $u \neq 0$ tal que $\langle u, w_1\rangle = 0$ y $\langle u, w_2\rangle = 0$, entonces $\langle sw_1 + tw_2, u\rangle = 0$ para s, t arbitrarios y

$$P_0 = \{(x,y,z) : \langle (x,y,z), u\rangle = 0\}.$$

Sea $d = \langle v, u\rangle$, entonces $\langle v + sw_1 + tw_2, u\rangle = \langle v, u\rangle = d$, para s, t arbitrarios. Es decir

$$P = \{(x,y,z) : \langle (x,y,z), u\rangle = d\}.$$

Ejemplo. Sea P el plano definido en forma paramétrica por

$$P = \{(1,1,0)+s(-1,0,-1)+t(0,1,-2) : s,t \in \mathbb{R}\}.$$

Encontremos la ecuación implícita de P.

Sea $u = (a,b,c)$, entonces

$$\begin{aligned} \langle u, (-1,0,-1)\rangle = 0 &\quad\Leftrightarrow\quad -a-c = 0, \\ \langle u, (0,1,-2)\rangle = 0 &\quad\Leftrightarrow\quad b-2c = 0. \end{aligned}$$

Estas dos ecuaciones se cumplen sii $a = -c$ y $b = 2c$, es decir si $u = (-c, 2c, c)$. Si, por ejemplo, $c = 1$, tenemos $u = (-1,2,1)$, luego el plano paralelo a P que pasa por el origen es

$$P_0 = \{(x,y,z) : -x+2y+z = 0\}.$$

Como $\langle (1,1,0),(-1,2,1)\rangle = 1$, obtenemos

$$\begin{aligned} P &= \{(x,y,z) : \langle (x,y,z),(-1,2,1)\rangle = 1\} \\ &= \{(x,y,z) : -x+2y+z = 1\}. \end{aligned}$$

§ *Ejercicios*

(1) Sea $v_0 = (2,-1,1)$.

a) Describir paramétricamente el conjunto $P_1 = \{w \in \mathbb{R}^3 : \langle v_0, w\rangle = 0\}$.

b) Describir paramétricamente el conjunto $P_2 = \{w \in \mathbb{R}^3 : \langle v_0, w\rangle = 1\}$.

c) ¿Qué relación hay entre P_1 y P_2?

(2) Escribir la ecuación paramétrica y la ecuación normal de los siguientes planos.

a) π_1: el plano que pasa por $(0,0,0)$, $(1,1,0)$, $(1,-2,0)$.

b) π_2: el plano que pasa por $(1,2,-2)$ y es perpendicular a la recta que pasa por $(2,1,-1)$, $(3,-2,1)$.

c) $\pi_3 = \{w \in \mathbb{R}^3 : w = s(1,2,0) + t(2,0,1) + (1,0,0);\ s,t \in \mathbb{R}\}$.

(3) ¿Cuáles de las siguientes rectas cortan al plano π_3 del ejercicio *c)*? Describir la intersección en cada caso.

(a) $\{w : w = (3,2,1) + t(1,1,1)\}$, (b) $\{w : w = (1,-1,1) + t(1,2,-1)\}$,
(c) $\{w : w = (-1,0,-1) + t(1,2,-1)\}$, (d) $\{w : w = (1,-2,1) + t(2,-1,1)\}$.

1.7 BASES ORTONORMALES EN $\mathbb{R}^n$ (*)

Definición 1.7.1. Diremos que $\mathcal{B} = \{u_1,\dots,u_n\} \subset \mathbb{R}^n$ es una *base ortogonal* o *BO* de $\mathbb{R}^n$ si $u_i \perp u_j$ cuando $i \neq j$. Diremos $\mathcal{B}$ que es una *base ortonormal* o *BON* si es una base ortogonal y $\|u_i\| = 1$ para todo i.

Ejemplo.

(1) La base canónica $\mathcal{C} = \{e_1,\dots,e_n\}$ es una base ortonormal de $\mathbb{R}^n$.

(2) Cualesquiera dos vectores ortogonales en $\mathbb{R}^2$ forman una base ortogonal, por ejemplo, $(1,-1),(1,1)$.

(3) $\mathcal{B}_1 = \{(\cos\theta, \operatorname{sen}\theta, 0), (-\operatorname{sen}\theta, \cos\theta, 0), (0,0,1)\}$ es una BON de $\mathbb{R}^3$

Demostración.

$$\begin{aligned}
\langle(\cos\theta, \operatorname{sen}\theta, 0), (\cos\theta, \operatorname{sen}\theta, 0)\rangle &= \cos^2\theta + \operatorname{sen}^2\theta = 1\\
\langle(-\operatorname{sen}\theta, \cos\theta, 0), (-\operatorname{sen}\theta, \cos\theta, 0)\rangle &= \operatorname{sen}^2\theta + \cos^2\theta = 1\\
\langle(0,0,1), (0,0,1)\rangle &= 1.
\end{aligned}$$

Es decir todos los vectores tiene norma 1. Ahora bien,

$$\begin{aligned}
\langle(\cos\theta, \operatorname{sen}\theta, 0), (-\operatorname{sen}\theta, \cos\theta, 0)\rangle &= -\operatorname{sen}\theta\cos\theta + \operatorname{sen}\theta\cos\theta = 0\\
\langle(\cos\theta, \operatorname{sen}\theta, 0), (0,0,1)\rangle &= 0\\
\langle(-\operatorname{sen}\theta, \cos\theta, 0), (0,0,1)\rangle &= 0.
\end{aligned}$$

Es decir, todos los vectores son ortogonales entre sí. Luego $\mathcal{B}_1$ es una BON.

□

(4) $\mathcal{B}_2 = \{(3/5, 4/5, 0), (-4/5, 3/5, 0), (0,0,1)\}$ es una BON de $\mathbb{R}^3$.

Esto se prueba en fórma análoga al ítem anterior: primero se verifica que todos los vectores tengan norma 1 y luego que dos vectores distintos en $\mathcal{B}_2$ sean ortogonales.

Observación. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortogonal de $\mathbb{R}^n$, entonces si $u_i = \dfrac{v_i}{||v_i||}$ el conjunto $\mathcal{B}' = \{u_1, \ldots, u_n\}$ es una base ortonormal de $\mathbb{R}^n$.

Recordemos que si $\mathcal{C} = \{e_1, \ldots, e_n\}$ es la base canónica de $\mathbb{R}^n$ y $v = (x_1, \ldots, x_n)$, entonces $v = x_1e_1 + x_2e_2 + \cdots + x_ne_n$. Como $\langle v, e_i\rangle = x_i$, podemos reescribir

$$v = \langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2 + \cdots + \langle v, e_n\rangle e_n.$$

El siguiente teorema generaliza la fórmula anterior a cualquier base ortonormal. La prueba la podremos hacer recién en el capítulo 5.

Teorema 1.7.2. *Sea $\mathcal{B} = \{u_1, \ldots, u_n\}$ una BON de $\mathbb{R}^n$, y $v \in \mathbb{R}^n$, entonces*

$$v = \langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2 + \cdots + \langle v, u_n\rangle u_n. \qquad \square$$

La facilidad de escribir cualquier vector como combinación lineal de los vectores de una base ortonormal es una propiedad de suma importancia con aplicaciones en la física y la matemática.

Corolario 1.7.3. *Sea $\mathcal{B} = \{w_1, \ldots, w_n\}$ una BO de $\mathbb{R}^n$, y $v \in \mathbb{R}^n$, entonces*

$$v = \frac{\langle v, w_1\rangle}{\langle w_1, w_1\rangle}w_1 + \frac{\langle v, w_2\rangle}{\langle w_2, w_2\rangle}w_2 + \cdots + \frac{\langle v, w_n\rangle}{\langle w_n, w_n\rangle}w_n.$$

Demostración. Como $\{w_1,\ldots,w_n\}$ es una BO, entonces $\left\{\frac{w_1}{||w_1||},\ldots,\frac{w_n}{||w_n||}\right\}$ es una BON. Por lo tanto,

$$\begin{aligned}
v &= \langle v, \frac{w_1}{||w_1||}\rangle\frac{w_1}{||w_1||} + \langle v, \frac{w_2}{||w_2||}\rangle\frac{w_2}{||w_2||} + \cdots + \langle v, \frac{w_n}{||w_n||}\rangle\frac{w_n}{||w_n||} \\
&= \frac{\langle v, w_1\rangle}{||w_1||^2}w_1 + \frac{\langle v, w_2\rangle}{||w_2||^2}w_2 + \cdots + \frac{\langle v, w_n\rangle}{||w_n||^2}w_n \\
&= \frac{\langle v, w_1\rangle}{\langle w_1, w_1\rangle}w_1 + \frac{\langle v, w_2\rangle}{\langle w_2, w_2\rangle}w_2 + \cdots + \frac{\langle v, w_n\rangle}{\langle w_n, w_n\rangle}w_n.
\end{aligned}$$

□

Corolario 1.7.4. *Sea* $\mathcal{B} = \{u_1,\ldots,u_n\}$ *una BON de* $\mathbb{R}^n$*, y* $v \in \mathbb{R}^n$*, entonces*

$$||v||^2 = \langle v, u_1\rangle^2 + \langle v, u_2\rangle^2 + \cdots + \langle v, u_n\rangle^2.$$

Demostración. Por el teorema, $v = \sum_{i=1}^n \langle v, u_i\rangle u_i$, luego

$$\begin{aligned}
||v||^2 &= \langle \sum_{i=1}^n \langle v, u_i\rangle u_i, \sum_{j=1}^n \langle v, u_j\rangle u_j\rangle = \sum_{i,j=1}^n \langle \langle v, u_i\rangle u_i, \langle v, u_j\rangle u_j\rangle \\
&= \sum_{i,j=1}^n \langle v, u_i\rangle\langle v, u_j\rangle\langle u_i, u_j\rangle = \sum_{i1}^n \langle v, u_i\rangle\langle v, u_i\rangle \\
&= \sum_{i=1}^n \langle v, u_i\rangle^2.
\end{aligned}$$

□

Proposición 1.7.5 (Proceso de ortogonalización de Gram-Schmidt)**.** *Sean* $\{w_1,\ldots,w_k\}$ *tales que* w_i *es no nulo y* $w_i \perp w_j$ *si* $i \neq j$*. Sea* $v \in \mathbb{R}^n$*. Entonces*

$$w = v - \frac{\langle v, w_1\rangle}{\langle w_1, w_1\rangle}w_1 - \frac{\langle v, w_2\rangle}{\langle w_2, w_2\rangle}w_2 - \cdots - \frac{\langle v, w_k\rangle}{\langle w_k, w_k\rangle}w_k,$$

satisface que $\langle w, w_i\rangle = 0$ *para* $0 \leqslant i \leqslant k$*.*

Demostración. Simplemente debemos calcular $\langle w, w_i\rangle$, utilizando el hecho de que el producto escalar es una forma bilineal.

$$\begin{aligned}
\langle w, w_i\rangle &= \langle v - \sum_{j=1}^k \frac{\langle v, w_j\rangle}{\langle w_j, w_j\rangle}w_j, w_i\rangle = \langle v, w_i\rangle - \sum_{j=1}^k \frac{\langle v, w_j\rangle}{\langle w_j, w_j\rangle}\langle w_j, w_i\rangle \\
&= \langle v, w_i\rangle - \langle v, w_i\rangle = 0.
\end{aligned}$$

□

Observación. Si el w resultante de la proposición anterior es no nulo, entonces el conjunto $\{w_1, \ldots, w_k, w\}$ es un conjunto de vectores ortogonales entre si. Por lo tanto, podemos partir de un vector no nulo e inductivamente ir encontrando conjuntos cada vez más grandes de vectores ortogonales entre si. De esta forma, al final del proceso, podremos obtener una base ortogonal.

Ejemplo. Sea $w_1 = (2, -1, 1)$.

(1) Encontrar $w_2, w_3 \in \mathbb{R}^3$ de tal forma que $\{w_1, w_2, w_3\}$ sea una BO.

(2) Dado $u = (3, 1, 5)$ escribir V como combinación lineal de w_1, w_2, w_3

Solución. (1) Consideremos $v = e_1$ y usemos Gram-Schmidt para encontrar, a partir de v y w_1 un vector w_2 ortogonal a w_1. Obtengamos w ortogonal a w_1 por Gram-Schmidt:

$$\begin{aligned} w &= v - \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = (1, 0, 0) - \frac{\langle (1,0,0), (2,-1,1) \rangle}{||(2,-1,1)||^2}(2,-1,1) \\ &= (1,0,0) - \frac{2}{6}(2,-1,1) = (1,0,0) - (\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}) \\ &= (\frac{1}{3}, \frac{1}{3}, -\frac{1}{3}) \end{aligned}$$

Obtuvimos $w = (\frac{1}{3}, \frac{1}{3}, -\frac{1}{3})$ ortogonal a w_1. Por comodidad, multiplicamos el vector por 3 y así obtenemos $w_2 = (1, 1, -1)$ ortogonal a w_1.

Ahora tenemos w_1, w_2 tales que $w_1 \perp w_2$. Encontremos w_3 ortogonal a ambos. Sea $v = e_2$ y por Gram-Schmidt obtengamos w ortogonal a w_1, w_2:

$$\begin{aligned} w &= v - \frac{\langle v, w_1 \rangle}{||w_1||^2} w_1 - \frac{\langle v, w_2 \rangle}{||w_2||^2} w_2 \\ &= (0,1,0) - \frac{\langle (0,1,0), w_1 \rangle}{||w_1||^2} w_1 - \frac{\langle (0,1,0), w_2 \rangle}{||w_2||^2} w_2 \\ &= (0,1,0) + \frac{1}{6}(2,-1,1) - \frac{1}{3}(1,1,-1) \\ &= (0, -\frac{1}{2}, -\frac{1}{2}). \end{aligned}$$

Luego si, por comodidad, consideramos $w_3 = -2w = (0, 1, 1)$ obtenemos que

$$\mathcal{B} = \{(2,-1,1), (1,1,-1), (0,1,1)\}$$

es una BO.

(2) Por el corolario 1.7.3:

$$\begin{aligned} (3,1,5) &= \frac{\langle (3,1,5), w_1 \rangle}{||w_1||^2} w_1 + \frac{\langle (3,1,5), w_2 \rangle}{||w_1||^2} w_2 + \frac{\langle (3,1,5), w_3 \rangle}{||w_3||^2} w_3 \\ &= \frac{5}{3}(2,-1,1) - \frac{1}{3}(1,1,-1) + 3(0,1,1). \end{aligned}$$

□

Ejemplo. Sea P el plano definido en forma paramétrica por

$$P = \{\lambda(1,1,0) + \mu(1,1,1) : \lambda, \mu \in \mathbb{R}\}.$$

Dar la ecuación implícita de P.

Solución. En este tipo de problemas también es útil el proceso de ortogonalización de Gram-Schmidt. Consideremos $w_1 = (1,1,0)$ y $v = (1,1,1)$, entonces por Gram-Schmidt obtenemos w_2 ortogonal a w_1:

$$\begin{aligned} w_2 &= v - \frac{\langle v, w_1\rangle}{\langle w_1, w_1\rangle} w_1 = (1,1,1) - \frac{\langle (1,1,1),(1,1,0)\rangle}{||(1,1,0)||^2}(1,1,0) \\ &= (1,1,1) - (1,1,0) \\ &= (0,0,1). \end{aligned}$$

Observar que el vector w_2 es ortogonal a w_1 y es igual a $-(1,1,0) + (1,1,1)$, luego $w_2 \in P$. En definitiva, $w_1 = (1,1,0)$ y $w_2 = (0,0,1)$ son dos vectores ortogonales que pertenecen a P, por lo tanto

$$P = \{\lambda w_1 + \mu w_2 : \lambda, \mu \in \mathbb{R}\}.$$

Debemos ahora encontrar un vector u ortogonal al plano, es decir ortogonal a w_1 y w_2. Sea $v = e_1$, por Gram-Schmidt podemos hacer

$$\begin{aligned} u &= v - \frac{\langle v, w_1\rangle}{||w_1||^2} w_1 - \frac{\langle v, w_2\rangle}{||w_2||^2} w_2 \\ &= (1,0,0) - \frac{\langle (1,0,0), w_1\rangle}{||w_1||^2} w_1 - \frac{\langle (1,0,0), w_2\rangle}{||w_2||^2} w_2 \\ &= (1,0,0) - \frac{1}{2}(1,1,0) \\ &= (\frac{1}{2}, -\frac{1}{2}, 0). \end{aligned}$$

Por lo tanto la ecuación implícita del plano es

$$P = \{(x,y,z) \in \mathbb{R}^3 : \frac{1}{2}x - \frac{1}{2}y = 0\}.$$

□

2 SISTEMAS LINEALES

En este capítulo estudiaremos en forma sistemática los sistemas de ecuaciones lineales, es decir las soluciones de un conjunto finito de ecuaciones donde la relación entre las incógnitas se expresa en forma lineal.

2.1 SISTEMAS DE ECUACIONES LINEALES

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

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array} \tag{2.1.1}$$

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

Llamaremos a (2.1.1) un *sistema de m ecuaciones lineales con n incógnitas.* A una n-tupla $(x_1, \ldots, x_n)$ de elementos de $\mathbb{K}^n$ que satisface cada una de las ecuaciones de (2.1.1) la llamaremos una *solución del sistema*. Si $y_1 = \cdots = y_m = 0$, el sistema se llamará *homogéneo.* En caso contrario el sistema se denominará *no homogéneo.*

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

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

Ejemplo 2.1.1. Resolvamos ahora un sistema de ecuaciones homogéneo sencillo:

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

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

$$\begin{array}{lllcrclcrcl} \text{Si hacemos} & -2(2)+(1) & \text{obtenemos:} & (1') & -7x_2 - 7x_3 & = & 0 & \Rightarrow & x_2 & = & -x_3 \\ \text{Si hacemos} & 3(1)+(2) & \text{obtenemos:} & (2') & 7x_1 + 7x_3 & = & 0 & \Rightarrow & x_1 & = & -x_3, \end{array}$$

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

En el ejemplo anterior hemos encontrado soluciones por *eliminación de incógnitas*, es decir multiplicando por constantes adecuadas ciertas ecuaciones y sumándolas hemos eliminado en (1) a x_1 y en (2) a x_2, con lo cual la solución del sistema se deduce inmediatamente por pasaje de término.

Ejemplo 2.1.2. Encontrar las soluciones (x, y, z) del sistema de ecuaciones:

$$\begin{array}{cccccc} (1) & x & & +2z & = & 1 \\ (2) & x & -3y & +3z & = & 2 \\ (3) & 2x & -y & +5z & = & 3 \end{array} \tag{S1}$$

Es decir, queremos encontrar los números reales x, y y z que satisfagan las ecuaciones anteriores.

Solución. Veremos que la única solución es $(x, y, z) = (-1, 0, 1)$. El método que usaremos, similar al del ejemplo anterior, será el de eliminación de variables o incógnitas: vemos en el sistema que queremos resolver 8 variables, algunas repetidas. Trataremos de eliminar en cada ecuación la mayor cantidad de variables posibles de tal forma de llegar a una formulación equivalente del sistema que nos de inmediatamente la solución.

Supongamos que (x, y, z) es una solución de nuestro sistema. Entonces también vale que:

$$\begin{array}{rcrcccccc} & (2) & & x & -3y & +3z & = & 2 \\ (-1)\cdot & (1) & (-1)\cdot & (x & & +2z) & = & (-1)\cdot 1 \\ \hline & (2') & & & -3y & +z & = & 1 \end{array}$$

Por lo tanto (x, y, z) también es solución del sistema

$$\begin{array}{cccccc} (1) & x & & +2z & = & 1 \\ (2') & & -3y & +z & = & 1 \\ (3) & 2x & -y & +5z & = & 3 \end{array} \tag{S2}$$

Dado que (x, y, z) es solución del sistema (S2), entonces también vale que:

$$\begin{array}{rcrcccccc} & (3) & & 2x & -y & +5z & = & 3 \\ (-2)\cdot & (1) & (-2)\cdot & (x & & +2z) & = & (-2)\cdot 1 \\ \hline & (3') & & & -y & +z & = & 1 \end{array}$$

Por lo tanto (x, y, z) también es solución del sistema

$$\begin{array}{cccccc} (1) & x & & +2z & = & 1 \\ (2') & & -3y & +z & = & 1 \\ (3') & & -y & +z & = & 1 \end{array} \tag{S3}$$

Dado que (x, y, z) es solución del sistema (S3), entonces también vale que:

$$\begin{array}{rcrccccc} & (2') & & -3y & +z & = & 1 \\ (-3)\cdot & (3') & (-3)\cdot & (-y & +z) & = & (-3)\cdot 1 \\ \hline & (2'') & & & -2z & = & -2 \end{array}$$

Por lo tanto (x, y, z) también es solución del sistema

$$\begin{array}{rrrl} x & & +2z & = 1 \\ & & -2z & = -2 \\ & -y & +z & = 1 \end{array} \quad \text{equivalentemente} \quad \begin{array}{rrrl} x & & +2z & = 1 \\ & & z & = 1 \\ & -y & +z & = 1 \end{array}$$

Dado (x, y, z) es solución del sistema

$$\begin{array}{crrrl} 1 & x & & +2z & = 1 \\ 2'' & & & z & = 1 \\ 3' & & -y & +z & = 1 \end{array} \tag{S4}$$

o equivalentemente

$$\begin{array}{crrrl} 1 & x & & +2z & = 1 \\ 3' & & -y & +z & = 1 \\ 2'' & & & z & = 1 \end{array} \tag{S5}$$

Entonces también vale que:

$$\begin{array}{rrcr} & 1 & x + 2z = & 1 \\ (-2)\cdot & 2'' & (-2)\cdot(\quad z\ ) = & (-2)\cdot 1 \\ \hline & 1' & x \quad = & -1 \end{array} \quad \text{y} \quad \begin{array}{rrcr} & 3' & -y + z = & 1 \\ (-1)\cdot & 2'' & (-1)\cdot(\quad z\ ) = & (-1)\cdot 1 \\ \hline & 3'' & -y \quad = & 0 \end{array}$$

Por lo tanto (x, y, z) también es solución del sistema

$$\begin{array}{rrrcr} x & & & = & -1 \\ & -y & & = & 0 \\ & & z & = & 1 \end{array} \tag{S6}$$

o equivalentemente

$$\begin{array}{rcr} x & = & -1 \\ y & = & 0 \\ z & = & 1 \end{array} \tag{S7}$$

En resumen, supusimos que (x, y, z) es una solución del sistema

$$\begin{array}{rrrl} x & & +2z & = 1 \\ x & -3y & +3z & = 2 \\ 2x & -y & +3z & = 1 \end{array}$$

y probamos que

$$x = -1 \quad y = 0, \quad z = 1.$$

□

Tanto en el ejemplo 2.1.1 como en el ejemplo 2.1.2 eliminamos variables usando alguna de las siguientes operaciones entre ecuaciones:

E1. multiplicar una ecuación por una constante no nula,

E2. sumar a una ecuación una constante por otra, y

E3. permutar ecuaciones.

Veremos en las secciones siguientes que estas operaciones son “reversibles” y que con ellas podemos reducir todo sistema de ecuaciones a uno cuyas soluciones son obvias.

Ejemplo. Así como haciendo operaciones del tipo E1, E2 y E3 en la ecuaciones del ejemplo 2.1.2 llegamos de

$$\begin{array}{rrrcr} x & & +2z & = & 1 \\ x & -3y & +3z & = & 2 \\ 2x & -y & +5z & = & 3 \end{array} \qquad a \qquad \begin{array}{rcr} x & = & -1 \\ y & = & 0 \\ z & = & 1, \end{array}$$

haciendo las “operaciones inversas” (que son del mismo tipo) podemos llegar de

$$\begin{array}{rcr} x & = & -1 \\ y & = & 0 \\ z & = & 1. \end{array} \qquad a \qquad \begin{array}{rrrcr} x & & +2z & = & 1 \\ x & -3y & +3z & = & 2. \\ 2x & -y & +5z & = & 3 \end{array}$$

Luego, ambos sistemas son tienen las mismas soluciones.

2.2 EQUIVALENCIA DE SISTEMAS DE ECUACIONES LINEALES

Dado el sistema

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array} \tag{2.2.1}$$

donde $y_1, \ldots, y_m$ y $a_{i,j}$ $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$ son números en $\mathbb{K}$, si multiplicamos cada ecuación por c_i $(1 \leqslant i \leqslant m)$ y sumamos miembro a miembro obtenemos

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

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

$$(c_1a_{11} + c_2a_{21} + \cdots + c_ma_{m1})x_1 + \cdots + (c_1a_{1n} + c_2a_{2n} + \cdots + c_ma_{mn})x_n = \\ = c_1y_1 + c_2y_2 + \cdots + c_my_m,$$

o, escrito de otra forma,

$$\left(\sum_{i=1}^{m} c_i a_{i1}\right) x_1 + \cdots + \left(\sum_{i=1}^{m} c_i a_{in}\right) x_n = \sum_{i=1}^{m} c_i y_i, \tag{2.2.2}$$

la cual es una *combinación lineal* de las ecuaciones dadas en (2.3.4). Observar que la ecuación (2.2.2), es una ecuación lineal con n incógnitas, es decir es del mismo tipo que cada una de las ecuaciones que componen el sistema de ecuaciones original.

Proposición 2.2.1. *Sean* $c_1, \ldots, c_m$ *en* $\mathbb{K}$. *Si* $(x_1, \ldots, x_n) \in \mathbb{K}^n$ *es solución del sistema de ecuaciones*

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

entonces $(x_1, \ldots, x_n)$ *también es solución de la ecuación*

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

Demostración. Por hipótesis

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

Luego,

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

y esta, como vimos, es otra escritura de la ecuación (2.2.2). □

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

Definición 2.2.2. Decimos que dos sistemas de ecuaciones lineales son *equivalentes* si cada ecuación de un sistema es combinación lineal del otro.

Teorema 2.2.3. *Dos sistemas de ecuaciones lineales equivalentes tienen las mismas soluciones.*

Demostración. Sea

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

equivalente a

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

En particular, las ecuaciones de (**) se obtienen a partir de combinaciones lineales de las ecuaciones del sistema (*). Luego, por proposición 2.2.1, si $(x_1, \ldots, x_n)$ es solución de (*), también será solución de cada una de las ecuaciones de (**) y por lo tanto solución del sistema.

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

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

Ejemplo. Encontrar las soluciones del siguiente sistema de ecuaciones

$$\begin{array}{cccccccc} 1 & 2x_1 & + & 4x_2 & - & 6x_3 & = & 0 \\ 2 & 3x_1 & - & x_2 & + & 5x_3 & = & 0. \end{array} \tag{S0}$$

Solución. Si reemplazamos la ecuación 1 por 1/2, obtenemos el sistema

$$\begin{array}{cccccccc} 1' & x_1 & + & 2x_2 & - & 3x_3 & = & 0 \\ 2' & 3x_1 & - & x_2 & + & 5x_3 & = & 0. \end{array} \tag{S1}$$

Reemplazando $2'$ por $2' - 3\,1'$, obtenemos

$$\begin{array}{cccccccc} 1'' & x_1 & + & 2x_2 & - & 3x_3 & = & 0 \\ 2'' & & - & 7x_2 & + & 14x_3 & = & 0. \end{array} \tag{S2}$$

Reemplazando $2''$ por $2''/(-7)$, obtenemos

$$\begin{array}{cccccccc} 1''' & x_1 & + & 2x_2 & - & 3x_3 & = & 0 \\ 2''' & & & x_2 & - & 2x_3 & = & 0. \end{array} \tag{S3}$$

Reemplazando $1'''$ por $1''' - 2\,2'''$, obtenemos

$$\begin{array}{cccccccc} 1'''' & x_1 & & & + & x_3 & = & 0 \\ 2'''' & & & x_2 & - & 2x_3 & = & 0. \end{array} \tag{S4}$$

Luego $x_1 = -x_3$ y $x_2 = 2x_3$, y esto nos dice que las soluciones son de la forma $\{(-x_3, 2x_3, x_3) : x_3 \in \mathbb{R}\}$.

Por otro lado, observar que

- a partir de (S4) podemos obtener (S3) reemplazando $1''''$ por $1'''' + 2\,2''''$;
- a partir de (S3) podemos obtener (S2) reemplazando $2'''$ por $-7\,2'''$;
- a partir de (S2) podemos obtener (S1) reemplazando $2''$ por $2'' + 3\,1''$;
- a partir de (S1) podemos obtener (S0) reemplazando $1'$ por $2\,1'$.

Es decir los sistemas (S0) y (S4) son equivalentes y por lo tanto tienen las mismas soluciones. Como el conjunto de soluciones de (S4) es $\{(-x_3, 2x_3, x_3) : x_3 \in \mathbb{R}\}$, éste también es el conjunto de soluciones del sistema original. □

Ejemplo. Encontrar las soluciones del siguiente sistema de ecuaciones

$$\begin{array}{cccccccc} 1 & 2x_1 & - & x_2 & + & x_3 & = & 1 \\ 2 & x_1 & + & 3x_2 & + & 3x_3 & = & 2 \\ 3 & x_1 & + & & + & 2x_3 & = & 1. \end{array}$$

(observar que en 3 el coeficiente de x_2 es cero.)

Solución. Si

reemplazamos 1 por	$1 - 2\,2 +$	obtenemos:	$1'$	$-7x_2 - 5x_3 = -3,$
reemplazamos 2 por	$2 + 3\,1$	obtenemos:	$2'$	$7x_1 + 6x_3 = 5,$
no cambiamos	3	y obtenemos:	$3'$	$x_1 + 2x_3 = 1.$

Ahora reemplazando $2'$ por $2' - 7\,1'$, obtenemos $-8x_3 = -2$, que es equivalente a que $x_3 = \frac{1}{4}$. Por lo tanto, obtenemos el sistema

$$\begin{array}{cccc} 1'' & -7x_2 - 5x_3 & = & -3 \\ 2'' & x_3 & = & \frac{1}{4} \\ 3'' & x_1 + 2x_3 & = & 1, \end{array}$$

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

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

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

2.3 MATRICES

En esta sección introduciremos el concepto de matriz y veremos un sistema de ecuaciones se puede describir en el lenguaje de las matrices. También veremos que sistemas de ecuaciones lineales equivalentes se corresponden con matrices equivalentes por filas. Esto nos permitirá, en la próxima sección, explicitar en forma clara y concisa el método de Gauss.

Debemos tener claro que el lenguaje de las matrices en el contexto de esta sección, no es más que una notación más cómoda para el problema de resolver sistemas de ecuaciones lineales.

Estudiaremos la solución de un sistema de ecuaciones lineales

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m. \end{array} \tag{2.3.1}$$

Observemos que podemos escribir los coeficientes de las fórmulas de la izquierda en un arreglo rectangular de m filas y n columnas:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2.3.2}$$

También podemos escribir los $x_1, \ldots, x_n$ e $y_1, \ldots, y_n$ como *matriz columna*

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \tag{2.3.3}$$

Definición 2.3.1. Sea $\mathbb{K}$ cuerpo. Una *matriz* $m \times n$ o de *orden* $m \times n$ es un arreglo rectangular de elementos de $\mathbb{K}$ con m filas y n columnas. A cada elemento de la matriz la llamamos *entrada* o *coeficiente*. Si A es una matriz $m \times n$, denotamos $[A]_{ij}$ la entrada que se ubica en la fila i y la columna j. Al conjunto de matrices de orden $m \times n$ con entradas en $\mathbb{K}$ lo denotamos $M_{m \times n}(\mathbb{K})$, o simplemente $M_{m \times n}$ si $\mathbb{K}$ está sobreentendido.

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

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

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

.

Usualmente escribiremos a una matriz $m \times n$ con entradas $[A]_{ij} = a_{ij}$ como en (2.3.2). A esta matriz también la podemos denotar como $A = [a_{ij}]$. Dos matrices $A = [a_{ij}]$ y $B = [b_{ij}]$, de orden $m \times n$, son iguales si $a_{ij} = b_{ij}$ para todo $i = 1, \ldots, m$ y $j = 1, \ldots, n$. Es decir, dos matrices son iguales si los elementos que ocupan la misma posición en ambas matrices coinciden.

Como hicimos al comienzo de la sección, a un sistema de m ecuaciones con n incógnitas le asignaremos una matriz $m \times n$ lo cual nos permitirá trabajar en forma más cómoda y, como veremos en la próxima sección, podremos resolver los sistemas de ecuaciones lineales en forma algorítmica, realizando operaciones elementales por fila en las matrices correspondientes.

Sean

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

Entonces, podemos escribir el sistema de ecuaciones (2.3.1) como

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}. \tag{2.3.4}$$

En forma resumida:

$$AX = Y. \tag{2.3.5}$$

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

2.3.1 *Operaciones elementales por fila*

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

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

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

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

y

$$F_r + F_s = \begin{bmatrix} a_{r1} + a_{s1} & a_{i2} + a_{s2} & \cdots & a_{in} + a_{sn} \end{bmatrix}.$$

Diremos que la fila i es nula si

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

Definición 2.3.2. Sea $A = [a_{ij}]$ una matriz $m \times n$, diremos que e es una *operación elemental por fila* si aplicada a la matriz A se obtiene $e(A)$ de la siguiente manera:

E1. multiplicando la fila r por una constante $c \neq 0$, o

E2. cambiando la fila F_r por $F_r + tF_s$ con $r \neq s$, para algún $t \in \mathbb{K}$, o

E3. permutando la fila r por la fila s.

E1, E2 y E3 son tres tipos de operaciones elementales, más precisamente sea

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

entonces

E1. si multiplicamos la fila r por $c \neq 0$,

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

con $c \neq 0$, o

E2. si $r \neq s$, multiplicamos la fila s por $t \in \mathbb{K}$ y la sumamos a la fila r,

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

E3. La última operación elemental es permutar la fila r por la fila s:

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

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

E1. Multiplicar la fila r por $c \neq 0$

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

E2. Si $r \neq s$, multiplicar la fila s por $t \in \mathbb{K}$ y sumarla a la fila r

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

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

E3. Permutar la fila r por la fila s

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

Ejemplo. Sea

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

Ejemplificaremos las operaciones elementales
E1. Multipliquemos la fila 2 por -2, obtenemos

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

E2. Sumemos a la fila 3 dos veces la fila 1,

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

E3. Permutemos la fila 2 con la fila 3.

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

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

Teorema 2.3.3. *A cada operación elemental por fila e le corresponde otra operación elemental e' (del mismo tipo que e) tal que $e'(e(A)) = A$ y $e(e'(A)) = A$. En otras palabras, la operación inversa de una operación elemental es otra operación elemental del mismo tipo.*

Demostración.

E1. La operación inversa de multiplicar la fila r por $c \neq 0$ es multiplicar la misma fila por $1/r$.

E2. La operación inversa de multiplicar la fila s por $t \in \mathbb{K}$ y sumarla a la fila r es multiplicar la fila s por $-t \in \mathbb{K}$ y sumarla a la fila r.

E3. La operación inversa de permutar la fila r por la fila s es la misma operación.

□

Definición 2.3.4. Sean A y B dos matrices $m \times n$. Diremos que B es *equivalente por filas* a A, si B se puede obtener de A por un número finito de operaciones elementales por fila.

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

(1) $A \sim A$ (reflexiva),

(2) $A \sim B$, entonces $B \sim A$ (simétrica), y

(3) si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

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

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

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

con $e_1, \ldots, e_k$ operaciones elementales por fila. Por el teorema 2.3.3, tenemos $e'_1, \ldots, e'_{k-1}, e'_k$ operaciones elementales inversas de $e_1, \ldots, e_{k-1}, e_k$, respectivamente. Luego,

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

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

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

Ejemplo. Veamos que la matriz

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

es equivalente por fila a la matriz

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

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

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

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

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

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & -1 \end{bmatrix} \xrightarrow{F_3 \leftrightarrow F_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{F_3-2F_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{F_1-3F_2}$$

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

□

Definición 2.3.5. Consideremos un sistema como en (2.3.1) y sea A la matriz correspondiente al sistema. La *matriz ampliada* del sistema es

$$A' = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & y_1 \\ a_{21} & \cdots & a_{2n} & y_2 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & y_m \end{array}\right] \tag{2.3.6}$$

que también podemos denotar

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

Teorema 2.3.6. *Sea* $[A|Y]$ *la matriz ampliada de un sistema no homogéneo y sea* $[B|Z]$ *una matriz que se obtiene a partir de* $[A|Y]$ *por medio de operaciones elementales. Entonces, los sistemas correspondientes a* $[A|Y]$ *y* $[B|Z]$ *tienen las mismas soluciones.*

Demostración. Supongamos que $[B|Z]$ se obtiene por una operación elemental por fila a partir de $[A|Y]$, entonces las ecuaciones de $[B|Z]$ son combinaciones lineales de las ecuaciones de $[A|Y]$. Como toda operación elemental por fila tiene inversa, podemos obtener $[A|Y]$ a partir de $[B|Z]$ y por lo tanto las ecuaciones de $[A|Y]$ son combinaciones lineales de las ecuaciones de $[B|Z]$. Es decir $[A|Y]$ y $[B|Z]$ determinan sistemas de ecuaciones lineales equivalentes y por lo tanto tiene las mismas soluciones (teorema 2.2.3).

En el caso que $[B|Z]$ se obtenga a partir $[A|Y]$ haciendo varias operaciones elementales, se aplica el razonamiento de arriba las veces que sea necesario. □

Ejemplo. Resolvamos el siguiente sistema:

$$\begin{aligned} 2x_1 - x_2 + x_3 + 2x_4 &= 2 \\ x_1 - 4x_2 - x_4 &= 1 \\ 2x_1 + 6x_2 - x_3 + 3x_4 &= 0, \end{aligned} \tag{2.3.7}$$

para $x_i \in \mathbb{R}$ $(1 \leqslant i \leqslant 4)$.

La matriz ampliada correspondiente a este sistema de ecuaciones es

$$\left[\begin{array}{cccc|c} 2 & -1 & 1 & 2 & 2 \\ 1 & -4 & 0 & -1 & 1 \\ 2 & 6 & -1 & 3 & 0 \end{array}\right].$$

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

$$\left[\begin{array}{cccc|c} 2 & -1 & 1 & 2 & 2 \\ 1 & -4 & 0 & -1 & 1 \\ 2 & 6 & -1 & 3 & 0 \end{array}\right] \xrightarrow{F_1 \leftrightarrow F_2} \left[\begin{array}{cccc|c} 1 & -4 & 0 & -1 & 1 \\ 2 & -1 & 1 & 2 & 2 \\ 2 & 6 & -1 & 3 & 0 \end{array}\right] \xrightarrow{F_2 - 2F_1} \left[\begin{array}{cccc|c} 1 & -4 & 0 & -1 & 1 \\ 0 & 7 & 1 & 4 & 0 \\ 2 & 6 & -1 & 3 & 0 \end{array}\right]$$

$$\xrightarrow{F_3 - 2F_1} \left[\begin{array}{cccc|c} 1 & -4 & 0 & -1 & 1 \\ 0 & 7 & 1 & 4 & 0 \\ 0 & 14 & -1 & 5 & -2 \end{array}\right] \xrightarrow{F_3 - 2F_2} \left[\begin{array}{cccc|c} 1 & -4 & 0 & -1 & 1 \\ 0 & 7 & 1 & 4 & 0 \\ 0 & 0 & -3 & -3 & -2 \end{array}\right] \xrightarrow{F_3/(-3)} \left[\begin{array}{cccc|c} 1 & -4 & 0 & -1 & 1 \\ 0 & 7 & 1 & 4 & 0 \\ 0 & 0 & 1 & 1 & \frac{2}{3} \end{array}\right]$$

$$\xrightarrow{F_2 - F_3} \left[\begin{array}{cccc|c} 1 & -4 & 0 & -1 & 1 \\ 0 & 7 & 0 & 3 & -\frac{2}{3} \\ 0 & 0 & 1 & 1 & \frac{2}{3} \end{array}\right] \xrightarrow{F_2/7} \left[\begin{array}{cccc|c} 1 & -4 & 0 & -1 & 1 \\ 0 & 1 & 0 & \frac{3}{7} & -\frac{2}{21} \\ 0 & 0 & 1 & 1 & \frac{2}{3} \end{array}\right] \xrightarrow{F_1 + 4F_2} \left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{5}{7} & \frac{13}{21} \\ 0 & 1 & 0 & \frac{3}{7} & -\frac{2}{21} \\ 0 & 0 & 1 & 1 & \frac{2}{3} \end{array}\right].$$

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

$$\begin{aligned} x_1 + \frac{5}{7}x_4 &= \frac{13}{21} \\ x_2 + \frac{3}{7}x_4 &= -\frac{2}{21} \\ x_3 + x_4 &= \frac{2}{3}, \end{aligned}$$

luego

$$\begin{aligned} x_1 &= -\frac{5}{7}x_4 + \frac{13}{21} \\ x_2 &= -\frac{3}{7}x_4 - \frac{2}{21} \\ x_3 &= -x_4 + \frac{2}{3}. \end{aligned}$$

Por lo tanto, el conjunto de soluciones del sistema de ecuaciones (2.3.7) es

$$\left\{ (-\frac{5}{7}t + \frac{13}{21}, -\frac{3}{7}t - \frac{2}{21}, -t + \frac{2}{3}, t) : t \in \mathbb{R} \right\}.$$

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

Ejemplo. Consideremos ahora el siguiente sistema sobre los números complejos:

$$\begin{aligned} 2x_1 + ix_2 &= 0 \\ -ix_1 + 3x_2 &= 0 \\ x_1 + 2x_2 &= 0. \end{aligned} \tag{2.3.8}$$

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

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

Luego el sistema (2.3.8) es equivalente al sistema $x_1 = x_2 = 0$, que resulta ser la única solución.

2.4 MÉTODO DE GAUSS

Ahora avanzaremos en una forma sistemática para hallar todas las soluciones de un sistema de ecuaciones.

2.4.1 *Matrices reducidas por filas*

Definición 2.4.1. Una matriz A de $m \times n$ se llama *reducida por filas* o *MRF* si

(a) la primera entrada no nula de una fila de A es 1. Este 1 es llamado *1 principal*.

(b) Cada columna de A que contiene un 1 principal tiene todos los otros elementos iguales a 0.

Una matriz A de $m \times n$ es *escalón reducida por fila* o *MERF* si, es MRF y

c) todas las filas cuyas entradas son todas iguales a cero están al final de la matriz, y

d) en dos filas consecutivas no nulas el 1 principal de la fila inferior está más a la derecha que el 1 principal de la fila superior.

Ejemplo. Las siguientes matrices son MRF, pero no MERF:

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & 3 \end{bmatrix} \text{ no cumple (c)}, \qquad \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \text{ no cumple (d)}.$$

Las siguientes matrices, no son MRF:

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

Las siguientes son MERF:

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

En general una matriz MERF tiene la forma

$$\begin{bmatrix} 0 & \cdots & 1 & * & 0 & * & * & 0 & * & * \\ 0 & \cdots & 0 & \cdots & 1 & * & * & 0 & * & * \\ \vdots & & \vdots & & \vdots & & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 1 & * & * \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 \end{bmatrix} \tag{2.4.1}$$

Definición 2.4.2. Sea Id_n la matriz $n \times n$ definida

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

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

Observar que Id_n es una matriz escalón reducida por fila.

Teorema 2.4.3. *Toda matriz* $m \times n$ *sobre* $\mathbb{K}$ *es equivalente por fila a una matriz escalón reducida por fila.*

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

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

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

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

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

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

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

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

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

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

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

Finalmente, intercambiando filas podemos entonces obtener una matriz escalón reducida por fila.

□

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

Observación 2.4.4 (MÉTODO PARA REDUCIR A MERF).

(1) Nos ubicamos en la primera fila.

(2) Si la fila es 0 y no es la última, pasar a la fila siguiente.

(3) Si la fila no es 0,

a) si el primera entrada no nula está en la columna k y su valor es c, dividir la fila por c (ahora la primera entrada no nula vale 1),

b) con operaciones elementales del tipo $F_r + tF_s$ hacer 0 todas las entradas en la columna k (menos la de la columna actual).

(4) Si la fila es la última, pasar al ((5)). Si la fila no es la última, pasar a la fila siguiente e ir al paso ((2)).

(5) Permutar las filas hasta obtener una MERF.

Ejemplo. Ejemplifiquemos con la matriz que aparece en el ejemplo de la p. 50, es decir con la matriz

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

Siguiendo estrictamente el algoritmo:

$$\begin{bmatrix} 2 & -1 & 1 & 2 \\ 1 & -4 & 0 & -1 \\ 2 & 6 & -1 & 3 \end{bmatrix} \xrightarrow{F_1/2} \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 1 & -4 & 0 & -1 \\ 2 & 6 & -1 & 3 \end{bmatrix} \xrightarrow[F_3-2F_1]{F_2-F_1} \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 0 & -\frac{7}{2} & -\frac{1}{2} & -2 \\ 0 & 7 & -2 & 1 \end{bmatrix}$$

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

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

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

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

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

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

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

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

Por lo tanto el conjunto de soluciones del sistema es

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

2.4.2 *Método de Gauss*

Consideremos el siguiente sistema no homogéneo de m ecuaciones lineales con n incógnitas:

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

con $a_{ij} \in \mathbb{K}$. Sea $[A|Y]$ la matriz asociada a este sistema. Mediante operaciones elementales de fila obtenemos una matriz $[B|Z]$, donde B es matriz escalón reducida por fila (teorema 2.4.3). Por el teorema 2.3.6 los sistemas de ecuaciones $AX = Y$ y $BX = Z$ tiene las mismas soluciones.

Veamos ahora formalmente cuales son las soluciones del sistema $BX = Z$. Sea r el número de filas no nulas de B y $k_1, \ldots, k_r$ las columnas donde aparecen los primeros 1's en las primeras r filas. Entonces, $k_1 < k_2 < \cdots < k_r$ y el sistema de ecuaciones asociado a B es:

$$\begin{array}{rcl} x_{k_1} + \sum_{j \neq k_1,\ldots,k_r} b_{1j}\, x_j &=& z_1 \\ x_{k_2} + \sum_{j \neq k_1,\ldots,k_r} b_{2j}\, x_j &=& z_2 \\ \vdots & & \vdots \\ x_{k_r} + \sum_{j \neq k_1,\ldots,k_r} b_{rj}\, x_j &=& z_r \\ 0 &=& z_{r+1} \\ \vdots & & \\ 0 &=& z_m. \end{array} \tag{2.4.2}$$

y, por lo tanto, el sistema tiene solución si y solo si $z_{r+1} = \cdots = z_m = 0$ y en ese caso las soluciones son:

$$\begin{array}{rl} x_{k_1} &= z_1 - \sum_{j \neq k_1,\ldots,k_r} b_{1j}\, x_j \\ x_{k_2} &= z_2 - \sum_{j \neq k_1,\ldots,k_r} b_{2j}\, x_j \\ \vdots & \qquad \vdots \\ x_{k_r} &= z_r - \sum_{j \neq k_1,\ldots,k_r} b_{rj}\, x_j \end{array}. \tag{2.4.3}$$

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

Las soluciones del sistema son, entonces, los $(x_1, \ldots, x_n) \in \mathbb{K}^n$ tal que $x_{k_1}, \ldots, x_{k_r}$ satisfacen las ecuaciones (2.4.3) y $x_i \in \mathbb{K}$ para x_i libre.

El método anterior es denominado *método de Gauss.*

En particular, hemos probado:

Teorema 2.4.5. *Sea A matriz $n \times n$ con coeficientes en $\mathbb{K}$. Si A es equivalente por filas a B una MERF y B tiene filas nulas, entonces si el sistema $AX = Y$ tiene solución, tiene infinitas soluciones.*

Demostración. Sea r el número de filas no nulas de B, como B tiene filas nulas, entonces $r < n$. Como el sistema tiene solución y $r < n$ hay al menos una variable libre. Esto implica que hay infinitas soluciones.

Por ser A y B equivalentes por fila, $AX = Y$ y $BX = Z$ tienen las mismas soluciones y, por consiguiente, $AX = Y$ tiene infinitas soluciones. □

Teorema 2.4.6. *Sea A una matriz $m \times n$ con $m < n$ e Y matriz $m \times 1$. Entonces, si el sistema de ecuaciones lineales $AX = Y$ tiene solución, tiene infinitas soluciones.*

Demostración. La matriz A es equivalente por fila a una matriz B escalón reducida por fila y B tiene r filas no nulas con $r \leqslant m < n$. Como el sistema tiene solución y $r < n$ hay al menos una variable libre. Por lo tanto, $BX = Z$ (y $AX = Y$) tienen infinitas soluciones. □

Lema 2.4.7. *Sea R una matriz $n \times n$ escalón reducida por fila tal que no tiene filas nulas. Entonces $R = \mathrm{Id}_n$.*

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

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

Teorema 2.4.8. *Sea A una matriz $n \times n$. Entonces, A es equivalente por filas a la matriz Id_n si y sólo si el sistema $AX = Y$ tiene un única solución.*

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

$$\begin{aligned} x_1 &= z_1 \\ x_2 &= z_2 \\ &\vdots \\ x_n &= z_n \end{aligned}$$

cuya única solución es la solución trivial.

($\Leftarrow$) Sea R la matriz escalón reducida por filas asociada a A. Por hipótesis, $AX = Y$ tiene una sola solución y por lo tanto $RX = Z$, para algún Z, tiene una sola solución. Luego, no hay variables libres, es decir hay n filas no nulas en R, como R tiene n filas, lo anterior implica que R no tiene filas nulas. Entonces, por el lema anterior, $R = \mathrm{Id}_n$. □

Ejemplo. Resolvamos el sistema

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 2 \\ 5x_2 - x_3 &= 0. \end{aligned}$$

La matriz aumentada correspondiente a este sistema es

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

apliquemos el método de Gauss:

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

Luego, el sistema se reduce a

$$\begin{aligned} x_1 + 3/5x_3 &= 1 \\ x_2 - 1/5x_3 &= 0. \end{aligned}$$

Es decir,

$$\begin{aligned} x_1 &= -3/5x_3 + 1 \\ x_2 &= 1/5x_3. \end{aligned}$$

En consecuencia, las soluciones de esta ecuación son

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

2.5 ÁLGEBRA DE MATRICES

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

2.5.1 *Algunos tipos de matrices*

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

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

es cuadrada de orden 3.

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

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

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

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}. \tag{2.5.1}$$

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

$$\begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \end{bmatrix}, \tag{2.5.2}$$

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

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

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

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

Por ejemplo, la matriz nula 2×3 es

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

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

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

$$\begin{bmatrix} 2 & -1 & 3 & 1 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}. \tag{2.5.3}$$

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

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

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

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

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

2.5.2 *Suma de matrices*

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

Veamos un ejemplo, consideremos las siguientes matrices:

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

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

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

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

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

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

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

2.5.3 *Multiplicación de matrices*

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

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}. \tag{2.5.4}$$

Es decir, los elementos que ocupan la posición ij en la matriz producto, se obtienen sumando los productos que resultan de multiplicar los elementos de la fila i en la primera matriz por los elementos de la columna j de la segunda matriz. Al producto de A por B lo denotamos AB.

Es muy importante recalcar que por la definición, se puede multiplicar una matriz $m \times n$ por una matriz $r \times p$, sólo si $n = r$ y en ese caso, la multiplicación resulta ser una matriz $m \times p$.

Podemos visualizar la multiplicación así:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} \cdots & b_{1j} & \cdots \\ \cdots & b_{2j} & \cdots \\ & \vdots & \\ \cdots & b_{nj} & \cdots \end{bmatrix} = \begin{bmatrix} & \vdots & \\ \cdots & \sum_{k=1}^{n} a_{ik} \cdot b_{kj} & \cdots \\ & \vdots & \end{bmatrix}$$

Observación 2.5.1. Sean $A = [a_{ij}]$ matriz $m \times n$ y $B = [b_{ij}]$ matriz $n \times p$, entonces si multiplicamos la matriz que se forma con la fila i de A por la matriz que determina la columna j de B, obtenemos el coeficiente ij de AB. Esquemáticamente

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

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

El lector recordará el producto escalar definido en el capítulo 1 y notará que el coeficiente ij de AB es el producto escalar de la fila i de A por la columna j de B, ambos pensados como vectores.

Ejemplo. Si

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

como A es 2×2 y B es 2×3, la matriz AB será 2×3 y aplicando la regla (2.5.4), obtenemos:

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

Observemos que, debido a nuestra definición, no es posible multiplicar B por A, pues no está definido multiplicar una matriz 2×3 por una 2×2.

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

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

Entonces, por un lado,

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

y por otro lado,

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

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

- Asociativa:

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

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

$$A\, \mathrm{Id}_n = A = \mathrm{Id}_m\, A,$$

- Distributiva:

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

y

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

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

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

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

$$A^0 = \mathrm{Id}_n, \quad A^m = A^{m-1}A,$$

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

Observación 2.5.2. Un caso especial de multiplicación es la multiplicación por matrices diagonales. Sea $n \in \mathbb{N}$ y

$$\mathrm{diag}(d_1, d_2, \ldots, d_n) := \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

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

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

Esto es claro, pues si denotamos $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$, el coeficiente ij de DA es la fila i de D por la columna j de A, es decir

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

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

Si B es $m \times n$, el lector podrá comprobar que

$$B\ \mathrm{diag}(d_1, d_2, \ldots, d_n) = \begin{bmatrix} d_1 C_1 & d_2 C_2 & \cdots & d_n C_n \end{bmatrix},$$

donde $C_1, C_2, \ldots, C_n$ son las columnas de B.

Finalmente, de lo visto más arriba respecto a la multiplicación por una matriz diagonal obtenemos:

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

para $k \in \mathbb{N}$.

Otras observaciones importantes:

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

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

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

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

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

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

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

2.5.4 *Multiplicación de una matriz por un escalar*

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

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

Por ejemplo,

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

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

$$cA = \begin{bmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix} \tag{2.5.5}$$

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

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

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

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

2.6 MATRICES ELEMENTALES

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

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

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array} \tag{2.6.1}$$

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

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

entonces

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

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

Definición 2.6.1. Una matriz $m \times m$ se dice *elemental* si fue obtenida por medio de una única operación elemental a partir de la matriz identidad Id_m.

Ejemplo. Veamos cuales son las matrices elementales 2×2:

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

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

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

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

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

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

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

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

$$a_{ij} = \begin{cases} 1 & \text{si } i = j \text{ e } i \neq k, \\ c & \text{si } i = j = k, \\ 0 & \text{si } i \neq j. \end{cases} \tag{2.6.2}$$

Gráficamente,

$$\begin{array}{cc} & \begin{array}{ccccc} & & \overset{k}{\downarrow} & & \end{array} \\ \overset{k}{\rightarrow} & \begin{bmatrix} 1 & 0 & & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & c & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & & \cdots & 1 \end{bmatrix} \end{array}$$

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

$$a_{ij} = \begin{cases} 1 & \text{si } i = j \\ c & \text{si } i = r,\ j = s, \\ 0 & \text{otro caso.} \end{cases} \tag{2.6.3}$$

Gráficamente,

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

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

$$a_{ij} = \begin{cases} 1 & \text{si } (i = j,\ i \neq r,\ i \neq s) \text{ o } (i = r,\ j = s) \text{ o } (i = s,\ j = r) \\ 0 & \text{otro caso.} \end{cases} \tag{2.6.4}$$

Gráficamente,

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

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

Teorema 2.6.2. *Sea* e *una operación elemental por fila y sea* E *la matriz elemental* $E = e(\mathrm{Id})$. *Entonces* $e(A) = EA$.

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

E1. Sea $c \in \mathbb{K}$, y sea e la operación elemental de a la fila 2 le sumarle la fila 1 multiplicada por c. Entonces. $E := e(\mathrm{Id}_2)$ resulta en la matriz elemental:

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

Ahora bien,

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

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

E2. Sea $c \in \mathbb{K}$, y sea e la operación elemental de a la fila 2 le sumarle la fila 1 multiplicada por c. Entonces. $E := e(\mathrm{Id}_2)$ resulta en la matriz elemental:

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

Luego

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

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

E3. Finalmente, sea e la operación elemental que intercambia la fila 1 por la fila 2. Entonces, $E := e(\mathrm{Id}_2)$ es la matriz

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

Luego

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

□

Corolario 2.6.3. *Sean* A *y* B *matrices* $m \times n$. *Entonces* B *equivalente por filas a* A *si y sólo si* $B = PA$ *donde* P *es producto de matrices elementales. Más aún, si* $B = e_k(e_{k-1}(\cdots(e_1(A))\cdots))$ *con* $e_1, e_2, \ldots, e_k$ *operaciones elementales de fila y* $E_i = e_i(\mathrm{Id})$ *para* $i = 1, \ldots, k$, *entonces* $B = E_k E_{k-1} \cdots E_1 A$.

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

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

Sea $E_i = e_i(\mathrm{Id}_m)$, entonces, por el teorema anterior $A_1 = E_1 A$ y $A_i = E_i A_{i-1}$ $(i = 2, \ldots, k)$. Por lo tanto $B = E_k A_{k-1}$, en otras palabras $B = E_k E_{k-1} \cdots E_1 A$, luego $P = E_k E_{k-1} \cdots E_1$.
($\Leftarrow$) Si $B = PA$, con $P = E_k E_{k-1} \cdots E_1$ donde $E_i = e_i(\mathrm{Id}_m)$ es una matriz elemental, entonces

$$B = PA = E_k E_{k-1} \cdots E_1 A \overset{\text{Teor. 2.6.2}}{=\!=} e_k(e_{k-1}(\cdots(e_1(A))\cdots)).$$

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

2.7 MATRICES INVERTIBLES

Definición 2.7.1. Sea A una matriz $n \times n$ con coeficientes en $\mathbb{K}$. Una matriz $B \in M_{n \times n}(\mathbb{K})$ es *inversa de* A si $BA = AB = \mathrm{Id}_n$. En ese caso, diremos que A es *invertible*.

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

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

Proposición 2.7.2. *Sea* $A \in M_{n \times n}(\mathbb{K})$,

(1) sean $B, C \in M_{n \times n}(\mathbb{K})$ *tales que* $BA = \mathrm{Id}_n$ *y* $AC = \mathrm{Id}_n$, *entonces* $B = C$;

(2) si A *invertible la inversa es única.*

Demostración. (1)

$$B = B\,\mathrm{Id}_n = B(AC) = (BA)C = \mathrm{Id}_n\,C = C.$$

(2) Sean B y C inversas de A, es decir $BA = AB = \mathrm{Id}_n$ y $CA = AC = \mathrm{Id}_n$. En particular, $BA = \mathrm{Id}_n$ y $AC = \mathrm{Id}_n$, luego, por (1), $B = C$. □

Definición 2.7.3. Sea $A \in M_{n \times n}(\mathbb{K})$ invertible. A la única matriz inversa de A la llamamos *la matriz inversa de* A y la denotamos A^{-1}.

Veremos más adelante que si una matriz $n \times n$ admite una inversa a izquierda, es decir si existe B tal que $BA = \mathrm{Id}_n$, entonces la matriz es invertible. Lo mismo vale si A admite inversa a derecha.

Ejemplo. Sea A la matriz

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

Entonces, A es invertible y su inversa es

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

Esto se resuelve comprobando que $AA^{-1} = \mathrm{Id}_3$ (por lo dicho más arriba es innecesario comprobar que $A^{-1}A = \mathrm{Id}_3$).

Observación. No toda matriz tiene inversa, por ejemplo la matriz nula (cuyos coeficientes son todos iguales a 0) no tiene inversa pues $0 \,.\, A = 0 \neq \mathrm{Id}$. También existen matrices no nulas no invertibles, por ejemplo la matriz

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

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

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

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

Teorema 2.7.4. *Sean A y B matrices $n \times n$ con coeficientes en $\mathbb{K}$. Entonces*

(1) si A invertible, entonces A^{-1} es invertible y su inversa es A, es decir $(A^{-1})^{-1} = A$;

(2) si A y B son invertibles, entonces AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.

Demostración. (1) La inversa a izquierda de A^{-1} es A, pues $AA^{-1} = \mathrm{Id}_n$. Análogamente, la inversa a derecha de A^{-1} es A, pues $A^{-1}A = \mathrm{Id}_n$. Concluyendo: A es la inversa de A^{-1}.

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

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}\,\mathrm{Id}_n\,B = B^{-1}B = \mathrm{Id}_n,$$

y, análogamente,

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\,\mathrm{Id}_n\,A^{-1} = AA^{-1} = \mathrm{Id}_n\,.$$

□

Observación. Si $A_1, \ldots, A_k$ son invertibles, entonces $A_1 \ldots A_k$ es invertible y su inversa es
$$(A_1 \ldots A_k)^{-1} = A_k^{-1} \ldots A_1^{-1}.$$
El resultado es una generalización del punto (2) del teorema anterior y su demostración se hace por inducción en k (usando (2) del teorema anterior). Se deja como ejercicio al lector.

Observación. La suma de matrices invertibles no necesariamente es invertible, por ejemplo $A + (-A) = 0$ que no es invertible.

Teorema 2.7.5. *Una matriz elemental es invertible.*

Demostración. Sea E la matriz elemental que se obtiene a partir de Id_n por la operación elemental e. Se e_1 la operación elemental inversa (teorema 2.3.3) y $E_1 = e_1(\mathrm{Id}_n)$. Entonces
$$\begin{aligned} EE_1 &= e(e_1(\mathrm{Id}_n)) = \mathrm{Id}_n \\ E_1E &= e_1(e(\mathrm{Id}_n)) = \mathrm{Id}_n\,. \end{aligned}$$
Luego $E_1 = E^{-1}$. □

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

(1) Si $c \neq 0$,
$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1/c & 0 \\ 0 & 1 \end{bmatrix} \text{ y } \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1/c \end{bmatrix},$$

(2) si $c \in \mathbb{K}$, ,
$$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix} \text{ y } \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -c \\ 0 & 1 \end{bmatrix}.$$

(3) Finalmente,
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

En el caso general tenemos:

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

(2)

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

(3)

$$\text{La inversa de}\quad \begin{array}{rc} & \begin{array}{ccccccc} & & r & & s & & \\ & & \downarrow & & \downarrow & & \end{array} \\ \begin{array}{r} \\ \\ \xrightarrow{r} \\ \\ \xrightarrow{s} \\ \\ \\ \end{array} & \begin{bmatrix} 1 & \cdots & & \cdots & & \cdots & 0 \\ \vdots & \ddots & & & & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & & \cdots & & \cdots & 1 \end{bmatrix} \end{array} \quad \text{es la misma matriz.}$$

Teorema 2.7.6. *Sea* A *matriz* $n \times n$ *con coeficientes en* $\mathbb{K}$. *Las siguientes afirmaciones son equivalentes*

i) A *es invertible,*

ii) A *es equivalente por filas a* Id_n,

iii) A *es producto de matrices elementales.*

Demostración.

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

Recordemos que las matrices escalón reducidas por fila si tienen filas nulas, ellas se encuentran al final. Ahora bien, si la última fila de R es nula entonces, RB tiene la última fila nula también y por lo tanto no puede ser igual a la identidad, es decir, en ese caso R no es invertible, lo cual produce un absurdo. Concluyendo: la última fila (la fila n) de R no es nula y como es MERF, R no tiene filas nulas. Por lo tanto $R = \mathrm{Id}_n$ (lema 2.4.7) y, entonces, A es equivalente por filas a Id_n.

ii) $\Rightarrow$ *iii)* Como A es equivalente por filas a Id_n, al ser la equivalencia por filas una relación de equivalencia, tenemos que Id_n es equivalente por filas a A, es decir existen $E_1, \ldots, E_k$ matrices elementales, tales que $E_1 E_2, \ldots, E_k \mathrm{Id}_n = A$. Por lo tanto, $A = E_1 E_2, \ldots, E_k$ producto de matrices elementales.

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

Corolario 2.7.7. *Sean A y B matrices $m \times n$. Entonces, B es equivalente por filas a A si y sólo si existe matriz invertible P de orden $m \times m$ tal que $B = PA$.*

Demostración.

($\Rightarrow$) B es equivalente por filas a A, luego existe P matriz producto de matrices elementales tal que $B = PA$. Como cada matriz elemental es invertible (teorema 2.7.5) y el producto de matrices invertibles es invertible (teorema 2.7.4(2)), se deduce que P es invertible.

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

Corolario 2.7.8. *Sea A matriz $n \times n$. Sean $e_1, \ldots, e_k$ las operaciones elementales por filas que reducen a A a una MERF y esta MERF es la identidad, es decir $e_1(e_2(\cdots(e_k(A))\cdots)) = \mathrm{Id}_n$. Entonces, A invertible y las mismas operaciones elementales aplicadas a Id_n nos llevan a A^{-1}, es decir $e_1(e_2(\cdots(e_k(\mathrm{Id}_n))\cdots)) = A^{-1}$.*

Demostración. Por el teorema anterior, al ser A equivalente por filas a la identidad, A es invertible. Sean las matrices elementales $E_i = e_i(\mathrm{Id}_n)$ para $i = 1, \ldots, k$, entonces (ver corolario 2.6.3) $E_1 E_2 \ldots E_k A = \mathrm{Id}_n$, por lo tanto, multiplicando por A^{-1} a derecha en ambos miembros,

$$\begin{aligned} E_1 E_2 \ldots E_k A A^{-1} = \mathrm{Id}_n A^{-1} \quad &\Leftrightarrow \\ E_1 E_2 \ldots E_k \mathrm{Id}_n = A^{-1} \quad &\Leftrightarrow \\ e_1(e_2(\cdots(e_k(\mathrm{Id}_n))\cdots)) = A^{-1}. & \end{aligned}$$

□

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

Ejemplo. Calculemos la inversa (si tiene) de

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

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

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

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

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

El lector desconfiado podrá comprobar, haciendo el producto de matrices, que $AA^{-1} = A^{-1}A = \mathrm{Id}_2$. □

Teorema 2.7.9. *Sea A matriz $n \times n$ con coeficientes en $\mathbb{K}$. Entonces, las siguientes afirmaciones son equivalentes.*

i) *A es invertible.*

ii) *El sistema $AX = Y$ tiene una única solución para toda matriz Y de orden $n \times 1$.*

iii) *El sistema homogéneo $AX = 0$ tiene una única solución trivial.*

Demostración.

i) ⇒ *ii)* Sea X_0 solución del sistema $AX = Y$, luego

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

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

ii) ⇒ *iii)* Es trivial, tomando $Y = 0$.

iii) ⇒ *i)* Sea R la matriz escalón reducida por filas equivalente a A, es decir $R = PA$ con P invertible y R es MERF. Si R tiene una fila nula, entonces por teorema 2.4.5, el sistema $AX = 0$ tiene más de una solución, lo cual es absurdo. Por lo tanto, R no tiene filas nulas. Como es una matriz cuadrada y es MERF, tenemos que $R = \mathrm{Id}_n$. Luego A es equivalente por filas a Id_n y por teorema 2.7.6 se deduce que A es invertible.

□

Corolario 2.7.10. *Sea A una matriz $n \times n$ con coeficientes en $\mathbb{K}$. Si A tiene inversa a izquierda, es decir si existe B matriz $n \times n$ tal que $BA = \mathrm{Id}_n$, entonces A es invertible. Lo mismo vale si A tiene inversa a derecha.*

Demostración. Supongamos que A tiene inversa a izquierda y que B sea la inversa a izquierda, es decir $BA = \mathrm{Id}_n$. El sistema $AX = 0$ tiene una única solución, pues $AX_0 = 0 \Rightarrow BAX_0 = B0 \Rightarrow X_0 = 0$. Luego, A es invertible (y su inversa es B).

Supongamos que A tiene inversa a derecha y que C sea la inversa a derecha, es decir $AC = \mathrm{Id}$. Por lo demostrado más arriba, C es invertible y su inversa es A, es decir $AC = \mathrm{Id}$ y $CA = \mathrm{Id}$, luego A es invertible. □

Terminaremos la sección calculando algunas matrices inversas usando el corolario 2.7.8.

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

Solución.

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 3 & 2 & 4 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right] \xrightarrow{F_2-3F_1} \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 5 & -2 & -3 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right] \xrightarrow{F_2\leftrightarrow F_3}$$

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

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 8 & -3 & 1 & -5 \end{array}\right] \xrightarrow{F_3/8} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{array}\right] \xrightarrow{F_2+2F_3}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{4} & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{3}{8} & \frac{1}{8} & -\frac{5}{8} \end{array}\right].$$

Por lo tanto

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

□

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

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

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

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

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

$$\begin{bmatrix} 1 & \frac{b}{a} \\ 0 & \frac{ad-bc}{a} \end{bmatrix} \overset{a/(ad-bc)\,F_2}{\longrightarrow} \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{bmatrix} \overset{F1-b/a\,F_2}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

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

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \overset{F_1/a}{\longrightarrow} \begin{bmatrix} \frac{1}{a} & 0 \\ 0 & 1 \end{bmatrix} \overset{F_2-cF_1}{\longrightarrow} \begin{bmatrix} \frac{1}{a} & 0 \\ -\frac{c}{a} & 1 \end{bmatrix} \overset{a/(ad-bc)\,F_2}{\longrightarrow} \begin{bmatrix} \frac{1}{a} & 0 \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \overset{F1-b/a\,F_2}{\longrightarrow}$$

$$\longrightarrow \begin{bmatrix} \frac{1}{a} + \frac{bc}{a(ad-bc)} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}.$$

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

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{2.7.1}$$

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

$$\begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \overset{F_1 \leftrightarrow F_2}{\longrightarrow} \begin{bmatrix} c & d \\ 0 & b \end{bmatrix} \overset{F_1/c}{\longrightarrow} \begin{bmatrix} 1 & \frac{d}{c} \\ 0 & b \end{bmatrix} \overset{F_2/b}{\longrightarrow} \begin{bmatrix} 1 & \frac{d}{c} \\ 0 & 1 \end{bmatrix} \overset{F_1-d/cF_2}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

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

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \overset{F_1 \leftrightarrow F_2}{\longrightarrow} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \overset{F_1/c}{\longrightarrow} \begin{bmatrix} 0 & \frac{1}{c} \\ 1 & 0 \end{bmatrix} \overset{F_2/b}{\longrightarrow} \begin{bmatrix} 0 & \frac{1}{c} \\ \frac{1}{b} & 0 \end{bmatrix} \overset{F_1-d/cF_2}{\longrightarrow} \begin{bmatrix} -\frac{d}{bc} & \frac{1}{c} \\ \frac{1}{b} & 0 \end{bmatrix}.$$

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

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

Es decir, la expresión de la inversa es igual a (2.7.1) (considerando que $a = 0$).

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

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

(ejercicio).

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

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{2.7.2}$$

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

2.8 DETERMINANTE

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

El determinante, permite, entre otras cosas,

- determinar si una matriz cuadrada es invertible,
- dar una fórmula cerrada para la inversa de una matriz invertible.

Como consecuencia de lo anterior, el determinante permite determinar si un sistema de n ecuaciones lineales con n incógnitas admite una única solución o no, y en el caso de que exista una única solución, dar una fórmula cerrada de esa solución.

Una forma de definir determinante es con una fórmula cerrada que usa el *grupo de permutaciones*. Esta forma de definir determinante está fuera del alcance de este curso. La forma que usaremos nosotros para definir determinante es mediante una definición recursiva: para calcular el determinante de una matriz $n \times n$, usaremos el cálculo del determinante para matrices $n - 1 \times n - 1$, que a su vez se calcula usando el determinante de matrices $n - 2 \times n - 2$ y así sucesivamente hasta llegar al caso base, que es el caso de matrices 1×1.

Definición 2.8.1. Sea $A \in M_n(\mathbb{K})$. Sean i, j tal que $1 \leqslant i, j \leqslant n$. Entonces $A(i|j)$ es la matriz $n-1 \times n-1$ que se obtiene eliminando la fila i y la columna j de A.

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

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

Definición 2.8.2. Sea $n \in \mathbb{N}$ y $A = [a_{ij}] \in M_n(\mathbb{K})$, entonces el *determinante de* A, denotado $\det(A)$ se define como:

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

(n) si $n > 1$,

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

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

Observemos, que con las definiciones introducidas tenemos

$$\det(A) = \sum_{i=1}^{n} a_{i1} C^A_{i1}. \tag{2.8.1}$$

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

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

Observación (Determinantes 2×2). Calculemos el determinante de las matrices 2×2. Sea

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

entonces

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

Cuando estudiamos la matrices invertibles 2×2 (ejemplo de p. 75), vimos que A es invertible si y solo si $ad - bc \neq 0$, es decir

$$A \textit{ es invertible si y solo si } \det A \neq 0. \tag{2.8.2}$$

Este resultado se generaliza para matrices $n \times n$. Más aún, la fórmula (2.7.1), que aquí reescribimos como

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

se generaliza también para matrices cuadradas de cualquier dimensión (ver el corolario C.2.4).

Observación (DETERMINANTES 3×3). Calculemos el determinante de las matrices 3×3. Sea

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

entonces

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

Observar que el determinante de una matriz 3×3 es una sumatoria de seis términos cada uno de los cuales es de la forma $\pm a_{1i_1}a_{2i_2}a_{3i_3}$ e $i_1i_2i_3$ puede ser cualquier permutación de 123. La fórmula

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}, \tag{2.8.3}$$

no es fácil de recordar, pero existe un procedimiento sencillo que nos permite obtenerla y es el siguiente:

(1) a la matriz original le agregamos las dos primeras filas al final,

(2) "sumamos" cada producto de las diagonales descendentes y "restamos" cada producto de las diagonales ascendentes.

$$\begin{array}{cccc} + & & & \\ & a_{11} & a_{12} & a_{13} \\ + & & & \\ & a_{21} & a_{22} & a_{23} \\ + & & & \\ & a_{31} & a_{32} & a_{33} \\ - & & & \\ & a_{11} & a_{12} & a_{13} \\ - & & & \\ & a_{21} & a_{22} & a_{13} \\ - & & & \end{array} \tag{2.8.4}$$

Es decir,

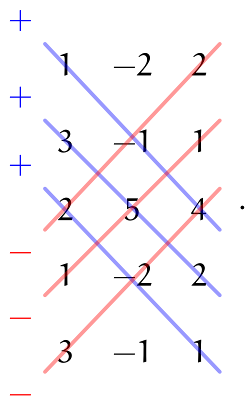
(a) se suman $a_{11}a_{22}a_{33}$, $a_{21}a_{32}a_{13}$, $a_{31}a_{12}a_{23}$, y

(b) se restan $a_{31}a_{22}a_{13}$, $a_{11}a_{32}a_{23}$, $a_{21}a_{12}a_{33}$.

Ejemplo. Calcular el determinante de

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

La forma más sencilla es ampliando la matriz y calculando:



Luego

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

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

Proposición 2.8.3. *Sea* $A\in M_n(\mathbb{K})$ *matriz triangular superior cuyos elementos en la diagonal son* $d_1,\dots,d_n$. *Entonces* $\det A = d_1.d_2.\dots.d_n$.

Demostración. Podemos demostrar el resultado por inducción sobre n: es claro que si $n=1$, es decir si $A=[d_1]$, el determinante vale d_1. Por otro lado, si $n>1$, observemos que $A(1|1)$ es también triangular superior con valores $d_2,\dots,d_n$ en la diagonal principal. Entonces, usamos la definición de la fórmula (2.8.1) y observamos que el desarrollo por la primera columna solo tiene un término, pues esta columna solo tiene un coeficiente no nulo, el d_1 en la primera posición. Por lo tanto,

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

□

Corolario 2.8.4. $\det \mathrm{Id}_n = 1$.

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

Corolario 2.8.5. *Si* R *es una MERF, entonces*

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

Demostración. Si R no tiene filas nulas es igual a Id_n (lema 2.4.7), luego $\det R = 1$. En general, R es una matriz triangular superior y si tiene alguna fila nula r, entonces el coeficiente en la diagonal de la fila r es igual a 0 y por lo tanto $\det R = 0$. □

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

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

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

luego

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

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

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

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

Por lo tanto,

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

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

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

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

por lo tanto

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

Todos los resultado del ejemplo anterior se pueden generalizar.

Teorema 2.8.6. *Sea* $A \in M_n(\mathbb{K})$ *y sean* $1 \leqslant r, s \leqslant n$.

(1) Sea $c \in \mathbb{K}$ *y* B *la matriz que se obtiene de* A *multiplicando la fila* r *por* c*, es decir* $A \xrightarrow{cF_r} B$*, entonces* $\det B = c \det A$.

(2) Sea $c \in \mathbb{K}$, $r \neq s$ *y* B *la matriz que se obtiene de* A *sumando a la fila* r *la fila* s *multiplicada por* c*, es decir* $A \xrightarrow{F_r + cF_s} B$*, entonces* $\det B = \det A$.

(3) Sea $r \neq s$ *y sea* B *la matriz que se obtiene de* A *permutando la fila* r *con la fila* s*, es decir* $A \xrightarrow{F_r \leftrightarrow F_s} B$*, entonces* $\det B = -\det A$.

Demostración. Ver los teoremas C.1.1, C.1.4, C.1.3 y sus demostraciones. □

Este resultado nos permite calcular el determinante de matrices elementales.

Corolario 2.8.7. *Sea* $n \in \mathbb{N}$ *y* $c \in \mathbb{K}$*. Sean* $1 \leqslant r, s \leqslant n$*, con* $r \neq s$.

(1) Si $c \neq 0$*, la matriz elemental que se obtiene de multiplicar por* c *la fila* r *de* Id_n*, tiene determinante igual a* c.

(2) Sea $r \neq s$*. La matriz elemental que se obtiene de sumar a la fila* r *de* Id_n *la fila* s *multiplicada por* c*, tiene determinante* 1.

(3) Finalmente, si $r \neq s$*, la matriz elemental que se obtiene de intercambiar la fila* r *por la fila* s *de* Id_n *tiene determinante* -1.

Demostración. Se deduce fácilmente del teorema anterior y del hecho de que $\det \mathrm{Id}_n = 1$. □

Corolario 2.8.8. *Sea* $A \in M_n(\mathbb{K})$.

(1) Si A *tiene dos filas iguales, entonces* $\det A = 0$.

(2) Si A *tiene una fila nula, entonces* $\det A = 0$.

Demostración. (1) Sea A matriz donde $F_r = F_s$ con $r \neq s$. Luego, intercambiando la fila r por la fila s obtenemos la misma matriz. Es decir $A \xrightarrow{F_r \leftrightarrow F_s} A$. Por el teorema 2.8.6 (3), tenemos entonces que $\det A = -\det A$, por lo tanto $\det A = 0$.

(2) Sea F_r una fila nula de A, por lo tanto multiplicar por 2 esa fila no cambia la matriz. Es decir $A \xrightarrow{2F_r} A$. Por el teorema 2.8.6 (1), tenemos entonces que $\det A = 2 \det A$, por lo tanto $\det A = 0$. □

Teorema 2.8.9. *Sean* $A, B \in M_n(\mathbb{K})$, *entonces*

(1) A *invertible si y solo si* $\det(A) \neq 0$.

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

Demostración. Ver el teoremas C.1.8 y C.1.9 y . □

Corolario 2.8.10. *Sean* $A, B \in M_n(\mathbb{K})$, *entonces*

(1) si A *invertible* $\det(A^{-1}) = \det(A)^{-1}$,

(2) $\det(AB) = \det(BA)$.

Demostración. 1. Por teorema 2.8.9, $\det(AA^{-1}) = \det(A)\det(A^{-1})$. Como $AA^{-1} = \mathrm{Id}_n$, entonces $1 = \det(\mathrm{Id}_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$. Por lo tanto $\det(A^{-1}) = 1/\det(A)$.

2. $\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA)$. □

Definición 2.8.11. Sea A una matriz $m \times n$ con coeficientes en $\mathbb{K}$. La *transpuesta* de A, denotada A^t, es la matriz $n \times m$ que en la fila i y columna j tiene el coeficiente $[A]_{ji}$. Es decir

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

Si A es una matriz $n \times n$, diremos que es *simétrica* si $A^t = A$.

Ejemplo. Si

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

entonces

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

Ejemplo. Si

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

entonces

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

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

Ejemplo. Si
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 4 & 7 \end{bmatrix},$$
entonces $A^t = A$, es decir A es simétrica.

Proposición 2.8.12. *Sea* A *matriz* $m \times n$.

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

(2) Si B *matriz* $n \times k$, *entonces*
$$(AB)^t = B^t A^t.$$

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

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

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

(3)
$$\begin{aligned} A \text{ invertible } &\Leftrightarrow \text{existe } B \text{ matriz } n \times n \text{ tal que } AB = \mathrm{Id}_n = BA \\ &\Leftrightarrow (AB)^t = \mathrm{Id}_n^t = (BA)^t \\ &\Leftrightarrow B^t A^t = \mathrm{Id}_n = A^t B^t \\ &\Leftrightarrow B^t \text{ es la inversa de } A^t. \end{aligned}$$

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

Observar que por inducción no es complicado probar que si $A_1, \ldots, A_k$ son matrices, entonces
$$(A_1 \ldots A_k)^t = A_k^t \ldots A_1^t.$$

Ejemplo. Veamos las transpuesta de las matrices elementales 2×2.

(1) Si $c \neq 0$, multiplicar por c la primera fila y multiplicar c por la segunda fila son, respectivamente,
$$E = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \text{ y } E = \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix},$$
por lo tanto E^t es la misma matriz en ambos casos.

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

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

por lo tanto

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

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

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

por lo tanto $E^t = E$

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

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

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

Teorema 2.8.13. *Sea* $A \in M_n(\mathbb{K})$*, entonces* $\det(A) = \det(A^t)$

Demostración. Ver el teorema C.1.11. □

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

Proposición 2.8.14. *Sea* $A \in M_n(\mathbb{K})$ *matriz triangular inferior cuyos elementos en la diagonal son* $d_1, \ldots, d_n$*. Entonces* $\det A = d_1.d_2.\ldots.d_n$.

Demostración. Si A es triangular inferior con elementos en la diagonal $d_1, \ldots, d_n$, entonces A^t es triangular superior con elementos en la diagonal $d_1, \ldots, d_n$. Por la proposición 2.8.3, $\det A^t = d_1 \ldots d_n$. Por el teorema 2.8.13 obtenemos el resultado. □

Teorema 2.8.15. *Sea* $A \in M_n(\mathbb{K})$ *y sean* $1 \leqslant r, s \leqslant n$.

(1) Sea $c \in \mathbb{K}$ *y* B *la matriz que se obtiene de* A *multiplicando la columna* r *por* c*, entonces* $\det B = c \det A$.

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

(3) Sea B *la matriz que se obtiene de* A *permutando la columna* r *con la fila* s*, entonces* $\det B = -\det A$.

Demostración. Las operaciones por columna del enunciado se traducen a operaciones por fila de la matriz A^t. Luego, aplicando los resultados del teorema 2.8.6 y usando el hecho de que $\det(A) = \det(A^t)$ y $\det(B) = \det(B^t)$ en cada caso, se deduce el corolario. □

Corolario 2.8.16. *Sea* $A \in M_n(\mathbb{K})$.

(1) *Si* A *tiene dos columnas iguales, entonces* $\det A = 0$.

(2) *Si* A *tiene una columna nula, entonces* $\det A = 0$.

Demostración. (1) Si A tiene dos columnas iguales, entonces A^t tiene dos filas iguales, luego, por corolario 2.8.8 (1), $\det A^t = 0$ y por lo tanto $\det A = 0$.

(2) Si A tiene una columna nula, entonces A^t tiene una fila nula, luego, 2.8.8 (2), $\det A^t = 0$ y por lo tanto $\det A = 0$. □

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

Teorema 2.8.17. *El determinante de una matriz* A *de orden* $n \times n$ *puede ser calculado por la expansión de los cofactores en cualquier columna o cualquier fila. Más específicamente,*

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

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

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

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

Demostración. Ver la demostración de el teorema C.1.12. □

2.9 AUTOVALORES Y AUTOVECTORES

La multiplicación de una matriz por un vector es una transformación lineal, las cuales veremos en el capítulo 4, y estas juegan un rol muy importante en toda la matemática, pasando por el álgebra, el análisis, la geometría, etc. Como ya vimos, es sencillo multiplicar a izquierda por matrices diagonales:

$$\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1x_1 \\ d_2x_2 \\ \vdots \\ d_nx_n \end{bmatrix}$$

(ver observación 2.5.2). En particular, si e_i es el vector columna que tiene un 1 en la fila i y todas las demás filas iguales a 0, entonces

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

y esta propiedad caracteríza las matrices diagonales, es decir una matriz D es diagonal si y solo si $De_i = d_i e_i$ para $1 \leqslant i \leqslant n$. Dicho de otra forma una matriz es diagonal si y solo si al aplicarla sobre algún e_i obtenemos un múltiplo de e_i.

Aunque no siempre es posible caracterizar una matriz A por los vectores v que al aplicarle A obtenemos un múltiplo de v, el estudio de este tipo de vectores resulta importante para obtener información de la matriz y propiedades de la misma.

Definición 2.9.1. Sea $A \in \mathbb{K}^{n\times n}$. Se dice que $\lambda \in \mathbb{K}$ es un *autovalor* de A y si existe $v \in \mathbb{K}^n$ no nulo tal que

$$Av = \lambda v.$$

En ese caso decimos que v es un *autovector* asociado a λ

Observación. Nada impide, por definición, que un autovalor pueda valer 0, pero un autovector *nunca* puede ser 0.

Ejemplo. 1 es un autovalor de Id_n y todo $v \in \mathbb{K}^n$ es un autovector asociado a 1 pues

$$\mathrm{Id}_n\, v = v$$

Ejemplo. 0 es un autovalor de $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ es un autovector asociado a 0 pues

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

Observación. La existencia de autovalores dependen del cuerpo donde estamos trabajando. Por ejemplo sea $A \in \mathbb{R}^{n\times n}$, definida por

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

Entonces, A no tiene autovalores reales. Veamos por qué.

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \tag{2.9.1}$$

Si λ fuera un autovalor y $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ fuera autovector, tendríamos

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix}. \tag{2.9.2}$$

Luego, por (2.9.1) y (2.9.2), $-x_2 = \lambda x_1$ y $x_1 = \lambda x_2$, entonces $-x_2 = \lambda x_1 = \lambda^2 x_2$. Si $\lambda \neq 0$, entonces $\lambda^2 > 0$, y eso implica que $x_2 = 0$ y en consecuencia $x_1 = 0$. Si $\lambda = 0$, también $x_1 = x_2 = 0$. Es decir, en ambos casos $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ y no es autovector.

Veremos más adelante que si permitimos autovalores complejos entonces esta matriz tiene autovalores.

Definición 2.9.2. Dado $i \in \{1, \ldots, n\}$, como ya vimos se denota e_i al vector columna de $\mathbb{K}^n$ cuyas coordenadas son todas ceros excepto la coordenada i que es un 1

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

El conjunto $\{e_1, \ldots, e_n\}$ se llama *base canónica* de $\mathbb{K}^n$.

Ejemplo. En $\mathbb{K}^3$ la base canónica es $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Ejemplo. Sea $D \in \mathbb{K}^{n \times n}$ una matriz diagonal con entradas $\lambda_1, \lambda_2, \ldots, \lambda_n$. Entonces e_i es un autovector con autovalor λ_i $\forall\, i \in \{1, \ldots, n\}$

El conjunto de todos los autovectores con un mismo autovalor es *invariante por la suma y la multiplicación por escalares.* En particular los múltiplos de un autovector son autovectores con el mismo autovalor.

Definición 2.9.3. Sea $A \in \mathbb{K}^{n \times n}$ y $\lambda \in \mathbb{K}$ un autovalor de A. El *autoespacio* asociado a λ es

$$V_\lambda = \{v \in \mathbb{K}^n \mid Av = \lambda v\}.$$

Es decir, V_λ es el conjunto formado por todos los autovectores asociados a λ y el vector nulo.

Teorema 2.9.4. *Sea A matriz $n \times n$ y $\lambda \in \mathbb{K}$. Si v, w pertenecen a V_λ, el autoespacio de A asociado a λ, entonces $v + tw \in V_\lambda$ para cualquier $t \in \mathbb{K}$.*

Demostración.

$$A(v + tw) = Av + tAw = \lambda v + t\lambda w = \lambda(v + tw).$$

□

Proposición 2.9.5. *Sea* A *matriz* $n \times n$ *y* $v, w \in \mathbb{K}^n$ *autovectores con autovalores* $\lambda, \mu \in \mathbb{K}$*, respectivamente. Entonces.* $\lambda \neq \mu$ *implica que* $v \neq w$*. Es decir, autovectores con autovalores distintos son distintos.*

Demostración. Supongamos que $v = w$, entonces $Av = \lambda v$ y $Av = \mu v$. Luego, $\lambda v = \mu v$ y por lo tanto

$$(\lambda - \mu)v = \begin{bmatrix} (\lambda - \mu)v_1 \\ \vdots \\ (\lambda - \mu)v_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Como $v \neq 0$ por ser autovector, alguna de sus coordenadas es no nula. Entonces $\lambda - \mu$ tiene que ser 0 o dicho de otro modo $\lambda = \mu$, lo cual es un absurdo. □

Problema. Hallar los autovalores de $A \in \mathbb{K}^{n \times n}$ y para cada autovalor, describir explícitamente el autoespacio asociado.

- En otras palabras nos preguntamos que $\lambda \in \mathbb{K}$ y que $v \in \mathbb{K}^n$ satisfacen

$$Av = \lambda v \Longleftrightarrow Av - \lambda v = 0 \Longleftrightarrow (A - \lambda \operatorname{Id})v = 0.$$

- La última igualdad es un sistema de ecuaciones lineales. Queremos ver entonces si existe un $v \in \mathbb{K}^n$ no nulo que sea solución del sistema homogéneo

$$(A - \lambda \operatorname{Id})X = 0. \tag{*}$$

- Un sistema $BX = 0$ tiene solución no trivial sii $\det(B) = 0$. Por lo tanto (*) tiene solución no trivial si y sólo si

$$\det(A - \lambda \operatorname{Id}) = 0.$$

Estos sencillos pasos demuestran lo siguiente.

Proposición 2.9.6. $\lambda \in \mathbb{K}$ *es un autovalor de* A *y* $v \in \mathbb{K}^n$ *es un autovector asociado a* λ *si y sólo si*

- $\det(A - \lambda \operatorname{Id}) = 0$
- v *es solución del sistema homogéneo* $(A - \lambda \operatorname{Id})X = 0$

Esta es casi la respuesta a nuestro problema. Para dar una respuesta más operativa introducimos el siguiente polinomio.

Definición 2.9.7. Sea $A \in \mathbb{K}^{n \times n}$. El *polinomio característico* de A es

$$\chi_A(x) = \det(A - x \operatorname{Id})$$

Ejemplo. El polinomio característico de Id_n es

$$\chi_{\mathrm{Id}_n}(x) = (1-x)^n$$

Demostración. $\mathrm{Id} - x\,\mathrm{Id} = (1-x)\,\mathrm{Id}$ es una matriz diagonal con $(1-x)$ en todas las entradas de la diagonal. Entonces el determinante es el producto de la diagonal. □

En general, si $A = [a_{ij}]$ matriz $n \times n$, tenemos que

$$\chi_A(x) = \det(A - x\ \mathrm{Id}) = \det \begin{bmatrix} a_{11}-x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-x \end{bmatrix}$$

y el polinomio característico de A es un polinomio de grado n, más precisamente

$$\chi_A(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Esto se puede demostrar por inducción.

Ejemplo. El polinomio característico de $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ es $\chi_A(x) = x^2$

Ejemplo. Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, entonces $\chi_A(x) = (a-x)(d-x) - bc$.

Demostración. $A - x\,\mathrm{Id} = \begin{bmatrix} a-x & b \\ c & d-x \end{bmatrix}$ y usamos la fórmula del determinante de una 2×2. □

Proposición 2.9.8. *Sea* $A \in \mathbb{K}^{n\times n}$. *Entonces* $\lambda \in \mathbb{K}$ *es autovalor si y sólo si* λ *es raíz del polinomio característico de* A.

Demostración.

$$\begin{aligned} \lambda \text{ es autovalor } &\Leftrightarrow \text{existe } v \neq 0 \text{ tal que } Av = \lambda v \\ &\Leftrightarrow 0 = Av - \lambda v = Av - \lambda\,\mathrm{Id}\,v = (A - \lambda\,\mathrm{Id})v \\ &\Leftrightarrow (A - \lambda\,\mathrm{Id})X = 0 \text{ tiene solución no trivial} \\ &\Leftrightarrow \chi_A(\lambda) = \det(A - \lambda\,\mathrm{Id}) = 0 \\ &\Leftrightarrow \lambda \text{ es raíz del polinomio característico.} \end{aligned}$$

□

Observación. Sea $A \in \mathbb{K}^{n\times n}$, entonces podemos aplicar el siguiente método para encontrar autovalores y autovectores de A.

(1) Calcular $\chi_A(x) = \det(A - x\,\mathrm{Id}\,)$,

(2) Encontrar las raíces $\lambda_1, \dots, \lambda_k$ de $\chi_A(x)$.

No siempre es posible hacerlo, pues no hay una fórmula o método general para encontrar las raíces de polinomios de grado 5 o superior.

(3) Para cada i con $1 \leqslant i \leqslant k$ resolver el sistema de ecuaciones lineales:

$$(A - \lambda_i\,\mathrm{Id})X = 0.$$

Las soluciones no triviales de este sistema son los autovectores con autovalor λ_i.

Ejemplo. Encontrar autovalores y autovectores de la matriz

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

Solución.

(1) $\chi_A(x) = \det \begin{bmatrix} 3-x & -2 \\ 1 & -x \end{bmatrix} = x^2 - 3x + 2 = (x-1)(x-2).$

(2) Los autovalores de A son las raíces de $\chi_A(x)$: 1 y 2.

(3) Debemos resolver los sistemas de ecuaciones:

$$(A - \mathrm{Id})X = 0, \qquad (A - 2\,\mathrm{Id})X = 0.$$

Es decir, debemos resolver los sistemas

$$\begin{bmatrix} 3-1 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{o, equivalentemente,} \quad \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (\text{S1})$$

$$\begin{bmatrix} 3-2 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{o, equivalentemente,} \quad \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (\text{S2})$$

(S1) $\begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} \stackrel{F_1-2F_2}{\longrightarrow} \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \Rightarrow x_1 - x_2 = 0 \Rightarrow (t,t)$ es solución.

(S2) $\begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \stackrel{F_2-F_1}{\longrightarrow} \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \Rightarrow x_1 - 2x_2 = 0 \Rightarrow (2t,t)$ es solución.

De lo anterior concluimos:

- Los autovalores de A son 1 y 2.
- El auto espacio correspondiente al autovalor 1 es

$$V_1 = \{t(1,1) : t \in \mathbb{R}\}.$$

- El auto espacio correspondiente al autovalor 2 es

$$V_2 = \{t(2,1) : t \in \mathbb{R}\}.$$

□

Ejemplo. Sea $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^2$. Encontrar los autovalores *reales* de A.

Solución. $A - x\,\mathrm{Id} = \begin{bmatrix} -x & -1 \\ 1 & -x \end{bmatrix}$, luego

$$\chi_A(x) = x^2 + 1.$$

El polinomio no tiene raíces reales, por lo tanto no existen autovalores reales (obviamente no hay autovectores). □

Sin embargo si nos dicen

Encontrar autovalores y autovectores complejos de la matriz $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$,

la respuesta va a ser diferente.

Lo que ocurre es que

$$\chi_A(x) = x^2 + 1 = (x+i)(x-i),$$

y este polinomio *sí* tiene raíces complejas: i y $-i$, por lo tanto i y $-i$ son los autovalores de A.

Averigüemos los autovalores planteando los sistemas de ecuaciones correspondientes, es decir $A - \lambda\,\mathrm{Id}\,x = 0$ para $\lambda = i, -i$:

$$\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{S1}$$

$$\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{S2}$$

Resolvamos los sistemas:

(S1) $\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \xrightarrow{F_1+iF_2} \begin{bmatrix} 0 & 0 \\ 1 & -i \end{bmatrix} \Rightarrow x_1 - ix_2 = 0 \Rightarrow (i\omega, \omega)$ es solución ($\omega \in \mathbb{C}$).

(S2) $\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \xrightarrow{F_1-iF_2} \begin{bmatrix} 0 & 0 \\ 1 & i \end{bmatrix} \Rightarrow x_1 + ix_2 = 0 \Rightarrow (-i\omega, \omega)$ es solución ($\omega \in \mathbb{C}$).

Luego A tiene dos autovalores, i y $-i$, y

$$V_i = \{\omega(i,1) : \omega \in \mathbb{C}\}, \qquad V_{-i} = \{\omega(-i,1) : \omega \in \mathbb{C}\}.$$

Nunca está de más comprobar los resultados:

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

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -i \end{bmatrix} = (-i) \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

Parte II

ÁLGEBRA LINEAL

3 ESPACIOS VECTORIALES

En este capítulo estudiaremos en forma general las combinaciones lineales sobre conjuntos abstractos. En el primer capítulo desarrollamos el método de Gauss para la resolución de sistemas de ecuaciones lineales. En el método de Gauss se usan sistemáticamente las combinaciones lineales de las filas de una matriz. Podemos ver estas filas como elementos de $\mathbb{K}^n$ y nuestro primer impulso para el estudio de las combinaciones lineales sería trabajar en este contexto, es decir en $\mathbb{K}^n$. Sin embargo, muchos de los resultados sobre combinaciones lineales en $\mathbb{K}^n$ son aplicables también a conjuntos más generales y de gran utilidad en la matemática. Por lo tanto, en este capítulo nuestros "espacios vectoriales" (espacios donde pueden hacerse combinaciones lineales de vectores) serán espacios abstractos, pero usualmente haremos referencia a los espacios vectoriales "concretos" (los $\mathbb{K}^n$) que ya conocemos.

3.1 DEFINICIÓN Y EJEMPLOS DE ESPACIOS VECTORIALES

Definición 3.1.1. Sea $\mathbb{K}$ cuerpo. Un *espacio vectorial* sobre $\mathbb{K}$ o un $\mathbb{K}$-*espacio vectorial* , consiste de un conjunto V no vacío, cuyos elementos son llamados *vectores*, junto a '+' y '.' tal que

(*a*) '$+: V \times V \to V$' es una operación, llamada *adición* o *suma de vectores*, tal que a dos vectores $v, w \in V$ les asigna otro vector $v + w \in V$,

(*b*) '$.: \mathbb{K} \times V \to V$' es una operación tal que a $\lambda \in \mathbb{K}$ y $v \in V$ le asigna el vector $\lambda.v$ (o simplemente λv). '.' es llamada el *producto por escalares*.

Además, estas operaciones deben satisfacer

S1. $v + w = w + v$, para $v, w \in V$ (*conmutatividad de la suma*),

S2. $(v + w) + u = v + (w + u)$, para $v, w, u \in V$ (*asociatividad de la suma*),

S3. existe un único vector 0, llamado *vector cero*, tal que $0 + v = v + 0 = v$, para todo $v \in V$ (*existencia de elemento neutro de la suma*).

S4. Para cada $v \in V$, existe un único vector $-v$ tal que $v + (-v) = (-v) + v = 0$ (*existencia de opuesto* o *inverso aditivo*).

P1. $1.v = v$ para todo $v \in V$.

P2. $\lambda_1(\lambda_2 v) = (\lambda_1\lambda_2)v$, para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $v \in V$.

D1. $\lambda(v+w) = \lambda v + \lambda w$, para todo $\lambda \in \mathbb{K}$ y todo $v, w \in V$ (*propiedad distributiva*).

D2. $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ para todo $\lambda_1, \lambda_2 \in \mathbb{K}$ y todo $v \in V$ (*propiedad distributiva*).

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

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

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

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

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

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

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

Ejemplo. $\mathbb{K}^n$. Sea $\mathbb{K}$ cuerpo, y sea

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

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

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

(*b*) $\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$.

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

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

S1. $x + y = y + x$, pues $x_i + y_i = y_i + x_i$, $1 \leqslant i \leqslant n$.

S2. $(x+y)+z = x+(y+z)$, pues $(x_i+y_i)+z_i = x_i+(y_i+z_i)$, $1 \leqslant i \leqslant n$.

S3. Sea $0 = (0, \ldots, 0)$, entonces $0 + x = (0 + x_1, \ldots, 0 + x_n) = (x_1, \ldots, x_n) = x$.

S4. Sea $-x = (-x_1, \ldots, -x_n)$, entonces $x + (-x) = (x_1 - x_1, \ldots, x_n - x_n) = (0, \ldots, 0)$.

P1. $1.x = (1.x_1, \ldots, 1.x_n) = (x_1, \ldots, x_n) = x$.

P2. $\lambda_1(\lambda_2 x) = (\lambda_1\lambda_2)x$ pues $\lambda_1(\lambda_2 x_i) = (\lambda_1\lambda_2)x_i$, $1 \leqslant i \leqslant n$.

D1. $\lambda(x+y) = \lambda x + \lambda y$, pues $\lambda(x_i + y_i) = \lambda x_i + \lambda y_i$, $1 \leqslant i \leqslant n$.

D2. $(\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$, pues $(\lambda_1 + \lambda_2)x_i = \lambda_1 x_i + \lambda_2 x_i$, $1 \leqslant i \leqslant n$.

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

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

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

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

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

Ejemplo. POLINOMIOS. Sea

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

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

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

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

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

Por ejemplo,

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

Ejemplo. ESPACIOS DE FUNCIONES. Sean

$$\begin{aligned} F(\mathbb{R}) &= \{f : \mathbb{R} \to \mathbb{R} : \text{ tal que } f \text{ es una función}\}, \\ C(\mathbb{R}) &= \{f : \mathbb{R} \to \mathbb{R} : \text{ tal que } f \text{ es una función continua}\}. \end{aligned}$$

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

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

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

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

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

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

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

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

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

Ejemplo. REALES POSITIVOS. Consideremos el conjunto de los números reales positivos:

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

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

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

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

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

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

Proposición 3.1.2. *Sea* V *un espacio vectorial sobre el cuerpo* $\mathbb{K}$. *Entonces,*

(1) $\lambda.0 = 0$, *para todo* $\lambda \in \mathbb{K}$;

(2) $0.v = 0$, *para todo* $v \in V$;

(3) si $\lambda \in \mathbb{K}, v \in V, v \neq 0$ *y* $\lambda.v = 0$, *entonces* $\lambda = 0$;

(4) $(-1).v = -v$, *para todo* $v \in V$.

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

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

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

(2) Análoga a (1).

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

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

□

3.2 SUBESPACIOS VECTORIALES

Definición 3.2.1. Sea V un espacio vectorial sobre $\mathbb{K}$. diremos que $W \subset V$ es *subespacio de* V si $W \neq \emptyset$ y

(a) si para cualesquiera $w_1, w_2 \in W$, se cumple que $w_1 + w_2 \in W$ y

(b) si $\lambda \in \mathbb{K}$ y $w \in W$, entonces $\lambda w \in W$.

Observación 3.2.2. Si W subespacio de V, entonces $0 \in W$: como $W \neq \emptyset$, tomo cualquier $w \in W$ y por (a) tenemos que $0.w \in W$. Ya vimos en la proposición 3.1.2(2) que $0.w = 0$ y por lo tanto $0 \in W$.

Observación 3.2.3. Si W subespacio de V y $w \in W$, entonces $-w \in W$: hemos visto (proposición 3.1.2(4)) que $(-1)w = -w$, luego por (b) de la definición de subespacio $-w \in W$.

Teorema 3.2.4. *Sea* V *un espacio vectorial sobre* $\mathbb{K}$ *y* W *subespacio de* V. *Entonces* W *con las operaciones suma y producto por escalares de* V *es un espacio vectorial.*

Demostración. Para que W sea espacio vectorial sus operaciones deben satisfacer los axiomas de la definición de espacio vectorial (definición 3.1.1).

Por la observación 3.2.2, el 0 del espacio vectorial pertenece al subespacio.

Por la observación 3.2.3 concluimos que $-w \in W$. Es decir el opuesto de un vector en W también pertenece a W.

Teniendo en cuenta estos dos hechos y que las operaciones en V satisfacen los axiomas de la definición 3.1.1 (y por lo tanto en W también), queda demostrado que W, con las operaciones heredadas de V, es espacio vectorial. $\square$

Ejemplo. Veremos ahora una serie de ejemplos de subespacios vectoriales.

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

(2) Sea V un $\mathbb{K}$-espacio vectorial y sea $v \in V$, entonces
$$W = \{\lambda v : \lambda \in \mathbb{K}\}$$
es un subespacio vectorial. En efecto

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

b) $\lambda_1 v \in W$, con $\lambda_1 \in \mathbb{K}$ y $\lambda \in \mathbb{K}$, entonces $\lambda(\lambda_1 v) = (\lambda\lambda_1)v \in W$.

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

(3) Sean $V = \mathbb{K}^n$ y $1 \leqslant j \leqslant n$. Definimos
$$W = \left\{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{K}\ (1 \leqslant i \leqslant n), x_j = 0\right\}.$$
Es decir W es el subconjunto de V de todas las n-tuplas con la coordenada j igual a 0. Por ejemplo si $j = 1$
$$W = \{(0, x_2, \ldots, x_n) : x_i \in \mathbb{K}\ (2 \leqslant i \leqslant n)\}.$$
Veamos que este último es un subespacio:

a) si $(0, x_2, \ldots, x_n), (0, y_2, \ldots, y_n) \in W$, entonces $(0, x_2, \ldots, x_n) + (0, y_2, \ldots, y_n) = (0, x_2 + y_2, \ldots, x_n + y_n)$, el cual pertenece a W.

b) Por otro lado, si $\lambda \in \mathbb{K}$, $\lambda(0, x_2, \ldots, x_n) = (0, \lambda x_2, \ldots, \lambda x_n) \in W$.

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

(4) El conjunto $\mathbb{R}[x] = \{P(x) : P(x) \text{ es polinomio en } \mathbb{R}\}$, es subespacio de $F(\mathbb{R})$, pues $\mathbb{R}[x] \subset F(\mathbb{R})$ y las operaciones de suma y producto por un escalar son cerradas en $\mathbb{R}[x]$.

(5) De forma análoga, el conjunto $\mathbb{R}[x]$ es subespacio de $C(\mathbb{R})$, el espacio de funciones continuas de $\mathbb{R}$.

(6) Sea $W = \left\{A \in M_n(\mathbb{K}) : A^t = A\right\}$. Es claro que $A \in W$ si y sólo si $[A]_{ij} = [A]_{ji}$. Veamos que W es subespacio de $M_n(\mathbb{K})$:

a) sean $A = [a_{ij}]$, $B = [b_{ij}]$ tales que $A = A^t$ y $B = B^t$, entonces debemos verificar que $A + B \in W$, es decir que la transpuesta de $A + B$ es la misma matriz: ahora bien, $[A + B]_{ij} = a_{ij} + b_{ij}$, luego

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

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

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

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

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

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

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

Sea

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

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

a) si $x, y \in W$, es decir si $Ax = 0$ y $Ay = 0$, entonces $A(x + y) = Ax + Ay = 0 + 0 = 0$, luego , $x + y \in W$;

b) si $\lambda \in \mathbb{K}$ y $x \in W$, entonces $A(\lambda x) = \lambda Ax = \lambda.0 = 0$, luego $\lambda x \in W$.

Definición 3.2.5. Sea V espacio vectorial sobre $\mathbb{K}$ y $v_1, \ldots, v_n$ vectores en V. Dado $v \in V$, diremos que v *es combinación lineal de los* $v_1, \ldots, v_n$ si existen escalares $\lambda_1, \ldots, \lambda_n$ en $\mathbb{K}$, tal que

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

Ejemplo.

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

$$v = iv_1 + 2v_2.$$

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

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

entonces

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

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

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

$$\begin{aligned} v &= iv_1+2v_2+0v_3, \qquad \text{y también} \\ v &= (i-1)v_1+v_2+v_3. \end{aligned}$$

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

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

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

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

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

entonces $v=\lambda_1v_1+\cdots+\lambda_nv_n$ se traduce, en coordenadas, a

$$\begin{aligned}(b_1,\ldots,b_m)&=\lambda_1(a_{11},\ldots,a_{m1})+\cdots+\lambda_n(a_{1n},\ldots,a_{mn})\\ &=(\lambda_1a_{11}+\cdots+\lambda_na_{1n},\ldots,\lambda_1a_{m1}+\cdots+\lambda_na_{mn}).\end{aligned}$$

Luego, v es combinación lineal de los vectores $v_1,\ldots,v_n\in\mathbb{K}^m$ si y sólo si el sistema de ecuaciones:

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

con incógnitas $\lambda_1,\ldots,\lambda_n$ tiene solución.

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

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

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

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

Ahora bien, usando el método de Gauss

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

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

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

Teorema 3.2.6. *Sea V un espacio vectorial sobre $\mathbb{K}$ y sean $v_1, \ldots, v_k \in V$. Entonces*

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

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

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

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

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

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

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

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

□

Definición 3.2.7. Sea V un espacio vectorial sobre $\mathbb{K}$ y sean $v_1, \ldots, v_k \in V$. Al subespacio vectorial $W = \{\lambda_1 v_1 + \cdots + \lambda_k v_k : \lambda_1, \ldots, \lambda_k \in \mathbb{K}\}$ de las combinaciones lineales de $v_1, \ldots, v_k$ se lo denomina *subespacio generado por* $v_1, \ldots, v_k$ y se lo denota

$$\begin{array}{rcll} W & = & \langle v_1, \ldots, v_k \rangle, & \text{o} \\ W & = & \operatorname{gen}\{v_1, \ldots, v_k\} & \text{o} \\ W & = & \operatorname{span}\{v_1, \ldots, v_k\}. & \end{array}$$

Además, en este caso, diremos que el conjunto $S = \{v_1, \ldots, v_k\}$ *genera* al subespacio W o que los vectores $v_1, \ldots, v_k$ *generan* W.

Teorema 3.2.8. *Sea V un espacio vectorial sobre $\mathbb{K}$. Entonces la intersección de subespacios vectoriales es un subespacio vectorial.*

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

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

entonces

(a) si $w_1, w_2 \in W$, tenemos que $w_1, w_2 \in W_i$ para todo $i \in I$, luego, como W_i es subespacio vectorial, $w_1 + w_2 \in W_i$ para todo $i \in I$, por lo tanto $w_1 + w_2 \in W$;

(b) si $\lambda \in \mathbb{K}$ y $w \in W$, $w \in W_i$ para todo $i \in I$ y, por lo tanto, $cw \in W_i$ para todo $i \in I$. En consecuencia $w \in W$.

$\square$

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

Teorema 3.2.9. *Sea V un espacio vectorial sobre $\mathbb{K}$ y sean $v_1, \ldots, v_k \in V$. Entonces, la intersección de todos los subespacios vectoriales que contienen a $v_1, \ldots, v_k$ es igual a $\langle v_1, \ldots, v_k \rangle$.*

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

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

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

Definición 3.2.10. Sea V un espacio vectorial sobre $\mathbb{K}$ y sean $S_1, \ldots, S_k$ subconjuntos de V. definimos

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

el conjunto *suma de los* $S_1, \ldots, S_k$.

Teorema 3.2.11. *Sea V un espacio vectorial sobre $\mathbb{K}$ y sean $W_1, \ldots, W_k$ subespacios de V. Entonces $W = W_1 + \cdots + W_k$ es un subespacio de V.*

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

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

(b) $\lambda v = \lambda(v_1 + \cdots + v_k) = \lambda v_1 + \cdots + \lambda v_k \in W_1 + \cdots + W_k$, pues como W_i es subespacio de V, tenemos que $\lambda v_i \in W_i$.

$\square$

Proposición 3.2.12. *Sea V un espacio vectorial sobre $\mathbb{K}$ y sean $v_1, \ldots, v_r$ elementos de de V. Entonces*

$$\langle v_1, \ldots, v_r \rangle = \langle v_1 \rangle + \cdots + \langle v_r \rangle.$$

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

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

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

Ejemplo. Veremos una serie de ejemplos de subespacios, suma e intersección de subespacios.

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

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

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

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

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

Luego, también podríamos escribir

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

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

$$W_1 = \left\{\begin{bmatrix} x_1 & x_2 \\ x_3 & 0 \end{bmatrix} : x_1, x_2, x_3 \in \mathbb{C}\right\}, \qquad W_2 = \left\{\begin{bmatrix} y_1 & 0 \\ 0 & y_2 \end{bmatrix} : y_1, y_2 \in \mathbb{C}\right\}.$$

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

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

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

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ x_3 & 0 \end{bmatrix} + \begin{bmatrix} y_1 & 0 \\ 0 & y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 & x_2 \\ x_3 & y_2 \end{bmatrix},$$

y esto se cumple tomando $x_1 = a, y_1 = 0, x_2 = b, x_3 = c, y_2 = d$.

Por otro lado

$$\begin{aligned} W_1 \cap W_2 &= \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in W_1, \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in W_2\right\} \\ &= \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} : \begin{matrix} (a = x_1, b = x_2, c = x_3, d = 0) \wedge \\ (a = y_1, b = c = 0, d = y_2) \end{matrix}\right\} \\ &= \left\{\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} : a \in \mathbb{C}\right\}. \end{aligned}$$

Finalmente, daremos la definición de espacio fila y espacio columna de una matriz, definiciones que veremos posteriormente son de importancia en la teoría.

Definición 3.2.13. Sea $A = [a_{ij}] \in M_{m\times n}(\mathbb{K})$. El *vector fila* i es el vector $(a_{i1}, \ldots, a_{in}) \in \mathbb{K}^n$. El *espacio fila* de A es el subespacio de $\mathbb{K}^n$ generado por los m vectores fila de A. De forma análoga, se define el vector columna j al vector $(a_{1j}, \ldots, a_{mj}) \in \mathbb{K}^m$ y el *espacio columna* de A es el subespacio de $\mathbb{K}^m$ generado por los n vectores columna de A.

Ejemplo. Sea
$$A = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{C}^{3\times 5},$$
entonces, por definición, el espacio fila es el subespacio generados por las filas de la matriz:
$$W = \langle (1,2,0,3,0), (0,0,1,4,0), (0,0,0,0,1) \rangle_{\mathbb{C}}.$$
También, como vimos en (1) del ejemplo de la p. 107, el espacio fila puede ser caracterizado de forma implícita:
$$W = \left\{ (x_1,x_2,x_3,x_4,x_5) \in \mathbb{C}^5 : x_2 = 2x_1, x_4 = 3x_1 + 4x_3 \right\}.$$

3.3 BASES Y DIMENSIÓN

Definición 3.3.1. Sea V un espacio vectorial sobre $\mathbb{K}$. Un subconjunto S de V se dice *linealmente dependiente* (o simplemente, *LD* o *dependiente*) si existen vectores distintos $v_1,\ldots,v_n \in S$ y escalares $\lambda_1,\ldots,\lambda_n$ de $\mathbb{K}$, no todos nulos, tales que
$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0.$$
Un conjunto que no es linealmente dependiente se dice *linealmente independiente* (o simplemente, *LI* o *independiente*).

Si el conjunto S tiene solo un número finito de vectores $v_1,\ldots,v_n$, se dice, a veces, que los $v_1,\ldots,v_n$ son LD (o LI), en vez de decir que S es LD (o LI, respectivamente).

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

(LI-1) $\forall\, \lambda_1,\ldots,\lambda_n$ en $\mathbb{K}$ tal que $\lambda_i \neq 0$ para algún i, entonces $\lambda_1 v_1 + \cdots + \lambda_n v_n \neq 0$, o ,

(LI-2) si $\lambda_1,\ldots,\lambda_n$ en $\mathbb{K}$ tales que $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$, entonces $0 = \lambda_1 = \cdots = \lambda_n$.

El enunciado (LI-1) se deduce intuitivamente negando la definición de linealmente dependiente y el resultado (LI-2) es el contrarrecíproco de (LI-1).

Para los interesados, lo anterior es un ejercicio de lógica: ser LD se puede enunciar
$$(\exists \lambda_1,\ldots,\lambda_n : (\exists i : \lambda_i \neq 0) \wedge (\lambda_1 v_1 + \cdots + \lambda_n v_n = 0)). \tag{LD}$$

Recordar que $\neg(\exists \lambda : P \wedge Q) \equiv (\forall \lambda : \neg P \vee \neg Q)$ y que $\neg P \vee \neg Q \equiv P \Rightarrow \neg Q$. Luego la negación de (LD), es decir ser LI, es
$$(\forall \lambda_1,\ldots,\lambda_n : (\exists i : \lambda_i \neq 0) \Rightarrow \lambda_1 v_1 + \cdots + \lambda_n v_n \neq 0). \tag{LI-1}$$

Como $(P \Rightarrow Q) \equiv (\neg Q \Rightarrow \neg P)$, el contrarrecíproco, la propiedad (LI-1) es equivalente a

$$(\forall \lambda_1, \ldots, \lambda_n : (\lambda_1 v_1 + \cdots + \lambda_n v_n = 0) \Rightarrow (\forall i : \lambda_i = 0)). \qquad \text{(LI-2)}$$

Las siguientes afirmaciones son consecuencias fácilmente deducibles de la definición.

(1) Todo conjunto que contiene un conjunto linealmente dependiente es linealmente dependiente.

(2) Todo subconjunto de un conjunto linealmente independiente es linealmente independiente.

(3) Todo conjunto que contiene el vector 0 es linealmente dependiente; en efecto, $1.0 = 0$.

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

$$\begin{array}{rcl} \lambda_1 - \lambda_2 & = & 0 \\ -\lambda_1 + \lambda_2 & = & 0 \\ \lambda_1 + \lambda_2 & = & 0 \end{array}.$$

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

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

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

Ejemplo. Sea $\mathbb{K}$ cuerpo. En $\mathbb{K}^3$ los vectores

$$\begin{aligned} v_1 &= (\ \ 3,\ 0, -3) \\ v_2 &= (-1,\ 1,\ \ 2) \\ v_3 &= (\ \ 4,\ 2, -2) \\ v_4 &= (\ \ 2,\ 1,\ \ 1) \end{aligned}$$

son linealmente dependientes, pues

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

Por otro lado, los vectores

$$\begin{aligned} e_1 &= (1, 0, 0) \\ e_2 &= (0, 1, 0) \\ e_3 &= (0, 0, 1) \end{aligned}$$

son linealmente independientes.

Observación. En general, en $\mathbb{K}^m$, si queremos determinar si $v_1, \ldots, v_n$ es LI, planteamos la ecuación

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = (0, \ldots, 0),$$

que, viéndola coordenada a coordenada, es equivalente a un sistema de m ecuaciones lineales con n incógnitas (que son $\lambda_1, \ldots, \lambda_n$). Si la única solución es la trivial entonces $v_1, \ldots, v_n$ es LI. Si hay alguna solución no trivial, entonces $v_1, \ldots, v_n$ es LD.

Definición 3.3.2. Sea V un espacio vectorial. Una *base* de V es un conjunto $\mathcal{B} \subseteq V$ tal que

(1) $\mathcal{B}$ genera a V, y

(2) $\mathcal{B}$ es LI.

El espacio V es de *dimensión finita* si tiene una base finita, es decir con un número finito de elementos.

Ejemplo (Base canónica de $\mathbb{K}^n$). Sea el espacio vectorial $\mathbb{K}^n$ y sean

$$\begin{array}{rcl} e_1 & = & (1, 0, 0, \ldots, 0) \\ e_2 & = & (0, 1, 0, \ldots, 0) \\ & & \ldots\ldots \\ e_n & = & (0, 0, 0, \ldots, 1) \end{array}$$

(e_i es el vector con todas sus coordenadas iguales a cero, excepto la coordenada i que vale 1). Entonces veamos que $e_1, \ldots, e_n$ es una base de $\mathbb{K}^n$.

(1) Si $(x_1, \ldots, x_n) \in \mathbb{K}^n$, entonces

$$(x_1, \ldots, x_n) = x_1 e_1 + \cdots + x_n e_n.$$

Por lo tanto, $e_1, \ldots, e_n$ genera a $\mathbb{K}^n$.

(2) Si

$$x_1 e_1 + \cdots + x_n e_n = 0,$$

entonces

$$\begin{aligned} (0, \ldots, 0) &= x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + \cdots + x_n(0, 0, \ldots, 1) \\ &= (x_1, 0, \ldots, 0) + (0, x_2, \ldots, 0) + \cdots + (0, 0, \ldots, x_n) \\ &= (x_1, x_2, \ldots, x_n). \end{aligned}$$

Luego, $x_1 = x_2 = \cdots = x_n = 0$ y por lo tanto $e_1, \ldots, e_n$ es LI.

Para $1 \leqslant i \leqslant n$, al vector e_i se lo denomina el i-*ésimo vector canónico* y a la base $\mathcal{B}_n = \{e_1, \ldots, e_n\}$ se la denomina la *base canónica* de $\mathbb{K}^n$.

Ejemplo. Sea P una matriz $n \times n$ invertible con elementos en el cuerpo $\mathbb{K}$. Entonces si $C_1, \ldots, C_n$ son los vectores columna de P (ver definición 3.2.13), estos forman una base de $\mathbb{K}^n$. Eso se verá como sigue. Si $X = (x_1, \ldots, x_n) \in \mathbb{K}^n$, lo podemos ver como columna y

$$PX = x_1 C_1 + \cdots + x_n C_n.$$

Como $PX = 0$ tiene solo la solución trivial $X = 0$, se sigue que $\{C_1, \ldots, C_n\}$ es un conjunto linealmente independiente. ¿Por qué generan $\mathbb{K}^n$? Sea $Y \in \mathbb{K}^n$, si $X = P^{-1}Y$, entonces $Y = PX$, esto es

$$Y = x_1 C_1 + \cdots + x_n C_n.$$

Así, $\{C_1, \ldots, C_n\}$ es una base de $\mathbb{K}^n$.

Ejemplo. Sea $\mathbb{K}_n[x]$ el conjunto de polinomios de grado menor que n con coeficientes en $\mathbb{K}$:

$$\mathbb{K}_n[x] = \left\{ a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1} : a_0, \ldots, a_{n-1} \in \mathbb{K} \right\}.$$

Entonces $1, x, x^2, \ldots, x^n$ es una base de $\mathbb{K}_n[x]$. Es claro que los $1, x, x^2, \ldots, x^n$ generan $\mathbb{K}_n[x]$. Por otro lado, si $\lambda_0 + \lambda_1 x + \lambda_2 x^2 + \cdots + \lambda_{n-1} x^{n-1} = 0$, tenemos que $\lambda_0 = \lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = 0$.

Ejemplo (Base canónica de $M_{m\times n}(\mathbb{K})$). Sean $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant m$ y $E_{ij} \in M_{m\times n}(\mathbb{K})$ definida por

$$[E_{ij}]_{kl} = \begin{cases} 1 & \text{si } i = k \text{ y } j = l, \\ 0 & \text{otro caso.} \end{cases}$$

Es decir E_{ij} es la matriz cuyas entradas son todas iguales a 0, excepto la entrada ij que vale 1. En el caso 2×2 tenemos la matrices

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

Volviendo al caso general, es claro que si $A = [a_{ij}] \in M_{m\times n}(\mathbb{K})$, entonces

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}, \tag{3.3.1}$$

luego $\{E_{ij}\}_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ genera $M_{m\times n}(\mathbb{K})$. También, por la ecuación (3.3.1), es claro que si $\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij} = 0$, entonces $a_{ij} = 0$ para todo i y j. Luego, $\{E_{ij}\}_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ es LI.

Concluyendo, $\{E_{ij}\}_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ es una base de $M_{m\times n}(\mathbb{K})$ y se la denomina la *base canónica* de $M_{m\times n}(\mathbb{K})$.

Observación. ¿Todo espacio vectorial tiene una base? La respuesta es *sí*. Sin embargo, la demostración de este hecho no es sencilla y requiere de herramientas de la teoría de conjuntos, en particular del Lema de Zorn. El lector interesado podrá ver el artículos sobre bases de un espacio vectorial en la Wikipedia: https://es.wikipedia.org/wiki/Base_(álgebra) y una demostración de la existencia de bases para cualquer espacio vectorial en http://fernandorevilla.es/blog/2014/06/22/existencia-de-base-en-todo-espacio-vectorial.

Más allá, de la dificultad en la demostración, supondremos siempre que todo espacio vectorial tiene una base.

Si S es un conjunto finito denotemos $|S|$ al *cardinal* de S es decir, la cantidad de elementos de S.

Teorema 3.3.3. *Sea V un espacio vectorial generado por un conjunto finito de vectores $w_1, \ldots, w_m$. Entonces todo conjunto independiente de vectores de V es finito y contiene a lo más m elementos.*

Demostración. Sea $V = \langle w_1, \ldots, w_m \rangle$ y $S \subset V$. El enunciado del teorema es equivalente a decir:

$$\text{si } S \text{ es LI } \Rightarrow |S| \leqslant m.$$

Para demostrar este teorema es suficiente probar el contrarrecíproco del enunciado, es decir:

$$\text{si } |S| > m \Rightarrow S \text{ es LD},$$

o, dicho de otra forma, todo subconjunto S de V que contiene más de m vectores es linealmente dependiente. Sea S un tal conjunto, entonces $S = \{v_1, \ldots, v_n\}$ con $n > m$. Como $w_1, \ldots, w_m$ generan V, existen escalares a_{ij} en $\mathbb{K}$ tales que

$$v_j = \sum_{i=1}^{m} a_{ij} w_i, \qquad (1 \leqslant j \leqslant n).$$

Probaremos ahora que existen $x_1, \ldots, x_n \in \mathbb{K}$ no todos nulos, tal que $x_1 v_1 + \cdots + x_n v_n = 0$. Ahora bien, para cualesquiera $x_1, \ldots, x_n \in \mathbb{K}$ tenemos

$$\begin{aligned}
x_1 v_1 + \cdots + x_n v_n &= \sum_{j=1}^{n} x_j v_j \\
&= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} w_i \\
&= \sum_{j=1}^{n} \sum_{i=1}^{m} (x_j a_{ij}) w_i \\
&= \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} x_j a_{ij} \Big) w_i. \qquad\qquad (*)
\end{aligned}$$

Si cada coeficiente que multiplica a cada w_i es nulo, entonces $x_1v_1 + \cdots + x_nv_n = 0$. Vamos a ver ahora que existen $x_, \ldots, x_n$ no todos nulos tal que los coeficientes que multiplica a w_i en $(*)$ sean todos nulos. Esto se debe a que el sistema de ecuaciones

$$\sum_{j=1}^{n} x_j a_{ij} = 0, \qquad (1 \leqslant i \leqslant m)$$

tiene m ecuaciones y $n > m$ incógnitas, luego, por el teorema 2.4.6, existen escalares $x_1, \ldots, x_n \in \mathbb{K}$ no todos nulos, tal que $\sum_{j=1}^{n} x_j a_{ij} = 0$, $(1 \leqslant i \leqslant m)$ y, por $(*)$

$$x_1v_1 + \cdots + x_nv_n = \sum_{i=1}^{m} (\sum_{j=1}^{n} x_j a_{ij}) w_i = \sum_{i=1}^{m} 0 \cdot w_i = 0,$$

con algún $x_i \neq 0$. Esto quiere decir que los $v_1, \ldots, v_n$ son LD. □

Corolario 3.3.4. *Si V es un espacio vectorial de dimensión finita, entonces dos bases cualesquiera de V tienen el mismo número de elementos.*

Demostración. Como V es de dimensión finita, tiene una base finita $\mathcal{B}$ de m vectores, es decir, $\mathcal{B}$ es base de V y $|\mathcal{B}| = m$. Sea $\mathcal{B}'$ otra base de V, como $\mathcal{B}$ genera V y $\mathcal{B}'$ es un conjunto LI, entonces, por el teorema anterior, $|\mathcal{B}'| \leqslant m$. Sea $n = |\mathcal{B}'|$, entonces $n \leqslant m$. Por otro lado $\mathcal{B}'$ es base y, por lo tanto, genera V y $\mathcal{B}$ es LI, luego, por el teorema anterior nuevamente, $m \leqslant n$, y en consecuencia $m = n$. □

Hemos demostrado, si V es un espacio vectorial de dimensión finita y $\mathcal{B}, \mathcal{B}'$ dos bases de V, entonces $|\mathcal{B}| = |\mathcal{B}'|$. Esto nos permite hacer la siguiente definición.

Definición 3.3.5. Sea V espacio vectorial de dimensión finita. Diremos que n es *la dimensión de* V y denotaremos $\dim V = n$, si existe una base de V de n vectores. Si $V = \{0\}$, entonces definimos $\dim V = 0$.

Ejemplo. Sean $m, n \in \mathbb{N}$.

(1) $\dim \mathbb{K}^n = n$, pues la base canónica tiene n elementos.

(2) $\dim M_{m\times n}(\mathbb{K}) = mn$, pues la base canónica de $M_{m\times n}(\mathbb{K})$ tiene mn elementos.

(3) $\dim \mathbb{K}_n[x] = n$, pues $1, x, x^2, \ldots, x^{n-1}$ es una base.

Corolario 3.3.6. *Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$. Entonces*

(1) cualquier subconjunto de V con más de n vectores es linealmente dependiente;

(2) *ningún subconjunto de* V *con menos de* n *vectores puede generar* V.

Demostración. (1) Sea $\{v_1, \ldots, v_n\}$ una base de V, entonces $v_1, \ldots, v_n$ generan V, luego, por el teorema 3.3.3, cualquier subconjunto de V que contenga más de n vectores es LD.

(2) Sea S subconjunto de V con $m < n$ vectores. Si S genera V, entonces todo subconjunto de más de m vectores es LD (teorema 3.3.3), por lo tanto, un subconjunto de n vectores es LD. En consecuencia, no puede haber una base de n elementos, lo cual contradice la hipótesis.

□

Lema 3.3.7. *Sea* S *un subconjunto LI de un espacio vectorial* V. *Suponga que* w *es un vector de* V *que no pertenece al subespacio generado por* S. *Entonces* $S \cup \{w\}$ *es LI.*

Demostración. Suponga que $v_1, \ldots, v_n$ son vectores distintos de S y sean $\lambda_i, \lambda \in \mathbb{K}$ tales que

$$\lambda_1 v_1 + \cdots + \lambda_n v_n + \lambda w = 0. \tag{3.3.2}$$

Debemos probar que $\lambda_i = 0$, $1 \leqslant i \leqslant n$, y $\lambda = 0$. Supongamos que $\lambda \neq 0$, entonces podemos dividir la ecuación por λ y haciendo pasaje de término obtenemos

$$w = \left(-\frac{\lambda_1}{\lambda}\right) v_1 + \cdots \left(-\frac{\lambda_n}{\lambda}\right) v_n.$$

Luego w estaría en el subespacio generado por S, lo cual contradice la hipótesis.

Por lo tanto $\lambda = 0$ y, en consecuencia

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

Como S es un conjunto linealmente independiente, todo $\lambda_i = 0$. □

Teorema 3.3.8. *Sea* V *espacio vectorial de dimensión finita* n *y* S_0 *un subconjunto LI de* V. *Entonces* S_0 *es finito y existen* $w_1, \ldots, w_m$ *vectores en* V *tal que* $S_0 \cup \{w_1, \ldots, w_m\}$ *es una base de* V.

Demostración. Se extiende S_0 a una base de V, como sigue. Si S_0 genera V, entonces S_0 es una base de V y está demostrado. Si S_0 no genera V, por el lema anterior se halla un vector w_1 en V tal que el conjunto $S_1 = S_0 \cup \{v_1\}$ es independiente. Si S_1 genera V, está demostrado. Si no, se aplica el lema para obtener un vector w_2 en V tal que el conjunto $S_2 = S_1 \cup \{w_2\} = S_0 \cup \{w_1, w_2\}$ es independiente. Si se continúa de este modo, entonces (y en no más de $\dim V$ de etapas) se llega a un conjunto

$$S_m = S_0 \cup \{w_1, \ldots, w_m\}$$

que es independiente y que genera V (si no, continuamos), por lo tanto S_m es base de V. □

Es decir, todo subconjunto LI de un espacio vectorial de dimensión finita se puede completar a una base.

Corolario 3.3.9. *Sea W es un subespacio de un espacio vectorial de dimensión finita n y S_0 un subconjunto LI de W. Entonces, S_0 se puede completar a una base de W.*

Demostración. Como S_0 es un conjunto linealmente independiente de W, entonces S_0 es también un subconjunto linealmente independiente de V; como V es de dimensión finita, S_0 no tiene más de n elementos y por lo tanto es finito.

Como W es un espacio vectorial, aplicando el teorema anterior completamos a una base de W. □

Corolario 3.3.10. *Sea V espacio vectorial de dimensión finita y $V \neq \{0\}$, entonces* $\dim V > 0$.

Demostración. Como $V \neq \{0\}$, existe $v \in V$ con $v \neq 0$. Entonces, $S_0 = \{v\}$ es LI, pues $\lambda v = 0 \Rightarrow \lambda = 0$. Por el teorema anterior, S_0 se extiende a una base $\mathcal{B}$. Como $|\mathcal{B}| \geqslant |S_0| = 1$, tenemos que $\dim V > 0$. □

Corolario 3.3.11. *Si W es un subespacio propio de un espacio vectorial de dimensión finita V, entonces W es de dimensión finita y* $\dim W < \dim V$.

Demostración. Si $W = \{0\}$, entonces $\dim W = 0$, como $W \subsetneq V$, tenemos que V es no nulo y por lo tanto $\dim W = 0 < \dim V$.

Si $W \neq \{0\}$, sea S subconjunto LI de W. Claramente S también es LI en V y por lo tanto $|S| < \dim(V)$. El axioma de buena ordenación nos garantiza que existe S subconjunto LI de W con $|S|$ máximo.

Veamos que S genera W. Si S no generara a W, entonces existiría $w \in W$ y $w \notin \langle S \rangle$. Como S es LI, por lema 3.3.7, $S \cup \{w\}$ es LI, está incluido en W y tiene cardinal mayor a S. Esto es un absurdo por la maximalidad de S.

Por lo tanto S es un conjunto LI que genera W, es decir, S es una base de W.

Como W es un subespacio propio de V existe un vector v en V que no está en W. Agregando v a la base S de W se obtiene un subconjunto LI de V (lema 3.3.7). Así, $\dim W < \dim V$. □

Hemos visto que si V es un espacio de dimensión finita, entonces todo conjunto LI se puede extender a una base. Veremos ahora que dado un conjunto finito de generadores, existe un subconjunto que es una base.

Teorema 3.3.12. *Sea $V \neq 0$ espacio vectorial y S un conjunto finito de generadores de V, entonces existe un subconjunto $\mathcal{B}$ de S que es una base.*

Demostración. Sea
$$C = \{|R| : R \subseteq S \ \wedge \ R \text{ es LI}\}.$$

Como V no es nulo y S genera V, $C \neq \emptyset$. C es un subconjunto no vacío de $\mathbb{N}$, acotado superiormente por $|S|$ y por lo tanto tiene máximo. Sea n el máximo de C entonces existe $\mathcal{B} \subseteq S$ tal que $|\mathcal{B}| = n$ y $\mathcal{B}$ es LI. Veremos que $\mathcal{B}$ es una base. Para ello, como $\mathcal{B}$ es LI, sólo falta ver que $\mathcal{B}$ genera a V.

Supongamos que existe $v \in S$ tal que $v \notin \langle \mathcal{B} \rangle$. Por el lema 3.3.7, entonces $\mathcal{B} \cup \{v\}$ es LI y este subconjunto LI de S tiene $n+1$ elementos, lo cual contradice la maximalidad de n. Es claro entonces, que $v \in S \Rightarrow v \in \mathcal{B}$, es decir $S \subset \langle B \rangle$. Como $S \subset \langle B \rangle$, entonces $V = \langle S \rangle \subset \langle B \rangle$, es decir $V = \langle B \rangle$. $\square$

Teorema 3.3.13. *Si W_1, y W_2 son subespacios de dimensión finita de un espacio vectorial, entonces $W_1 + W_2$ es de dimensión finita y*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Demostración. El conjunto $W_1 \cap W_2$ es un subespacio de W_1 y W_2 y por lo tanto un espacio vectorial de dimensión finita. Sea $u_1, \ldots, u_k$ una base de $W_1 \cap W_2$, por el teorema 3.3.8, existen $v_1, \ldots, v_n$ vectores en W_1 y $w_1, \ldots, w_m$ vectores en W_2 tal que

$$\{u_1, \ldots, u_k, v_1, \ldots, v_n\} \quad \text{es una base de } W_1,$$

y

$$\{u_1, \ldots, u_k, w_1, \ldots, w_m\} \quad \text{es una base de } W_2.$$

Es claro que, el subespacio $W_1 + W_2$ es generado por los vectores

$$u_1, \ldots, u_k, v_1, \ldots, v_n, w_1, \ldots, w_m.$$

Veamos que estos vectores forman un conjunto independiente. En efecto, suponga que

$$\sum \lambda_i u_i + \sum \gamma_i v_i + \sum \mu_i w_i = 0, \tag{3.3.3}$$

luego

$$\sum \mu_i w_i = -\sum \lambda_i u_i - \sum \gamma_i v_i.$$

Por lo tanto, $\sum \mu_i w_i \in (W_1 \cap W_2) + W_1 = W_1$. Es decir, $\sum \mu_i w_i \in W_2$ y $\sum \mu_i w_i \in W_1$, por lo tanto $\sum \mu_i w_i \in (W_1 \cap W_2)$, y entonces

$$\sum \mu_i w_i = \sum \alpha_i u_i \Rightarrow 0 = \sum \alpha_i u_i - \sum \mu_i w_i.$$

Como $\{u_1, \ldots, u_k, w_1, \ldots, w_m\}$ es una base y por lo tanto LI, tenemos que $0 = \alpha_i = \mu_j$, para todo i, j. Por lo tanto, por (3.3.3),

$$\sum \lambda_i u_i + \sum \gamma_i v_i = 0. \tag{3.3.4}$$

Como $\{u_1, \ldots, u_k, v_1, \ldots, v_n\}$ es una base de W_1, tenemos que también $0 = \lambda_i = \gamma_j$ para todo i, j. Luego $0 = \lambda_i = \gamma_j = \mu_r$, para cualesquiera i, j, r y por lo tanto $u_1, \ldots, u_k, v_1, \ldots, v_n, w_1, \ldots, w_m$ es LI y como generaban a $W_1 + W_2$ resultan ser una base de $W_1 + W_2$, por lo tanto $\dim(W_1 + W_2) = k + n + m$.

Finalmente,

$$\begin{aligned}\dim W_1 + \dim W_2 &= (k+n)+(k+m)\\ &= k+(k+n+m)\\ &= \dim(W_1\cap W_2)+\dim(W_1+W_2).\end{aligned}$$

□

3.4 DIMENSIONES DE SUBESPACIOS

Dada $A \in M_{m\times n}(\mathbb{K})$, ya hemos visto que las soluciones del sistema $AX = 0$ forman un subespacio vectorial. Sea R la MERF equivalente por filas a A y r la cantidad de filas no nulas de R. Ahora bien, cada fila no nula está asociada a una variable principal y las $n-r$ variables restantes son variables libres que generan todas las soluciones. El hecho de que tenemos $n-r$ variables libres no dice que hay $n-r$ vectores LI que generan W, y por lo tanto, $\dim W = n-r$. Esto lo veremos en el ejemplo que sigue. La demostración de hecho mencionado más arriba se verá en el capítulo correspondiente a transformaciones lineales (capítulo 4).

Ejemplo. Encontrar una base del subespacio

$$W = \left\{(x,y,z,w)\in\mathbb{R} : \begin{array}{rcl} x-y-3z+\ \ w &=& 0\\ y+5z+3w &=& 0\end{array}\right\}.$$

Solución. W está definido implícitamente y usando el método de Gauss podemos describirlo paramétricamente, pues:

$$\begin{bmatrix}1&-1&-3&1\\0&1&5&3\end{bmatrix}\xrightarrow{F_1+F_2}\begin{bmatrix}1&0&2&4\\0&1&5&3\end{bmatrix}.$$

Por lo tanto, el sistema de ecuaciones que define W es equivalente a

$$\begin{array}{rcl} x+2z+4w &=& 0\\ y+5z+3w &=& 0,\end{array}$$

es decir

$$\begin{array}{rcl} x &=& -2z-4w\\ y &=& -5z-3w,\end{array}$$

y entonces

$$\begin{aligned}W &= \{(-2z-4w,-5z-3w,z,w) : z,w\in\mathbb{R}\}\\ &= \{(-2,-5,1,0)z+(-4,-3,0,1)w : z,w\in\mathbb{R}\}\\ &= \langle(-2,-5,1,0),(-4,-3,0,1)\rangle.\end{aligned}$$

Concluimos entonces que $(-2,-5,1,0),(-4,-3,0,1)$ es una base de W y, por lo tanto, su dimensión es 2. □

Teorema 3.4.1. *Sean* A *matriz* $m \times n$ *con coeficientes en* $\mathbb{K}$, P *matriz* $m \times m$ *invertible y* $B = PA$. *Entonces el el espacio fila de* A *es igual al espacio fila de* B.

Demostración. Sea $A = [a_{ij}]$, $P = [p_{ij}]$ y $B = [b_{ij}]$. Como $B = PA$, tenemos que la fila i de B es

$$\begin{aligned}(b_{i1}, \ldots, b_{in}) &= (F_i(P).C_1(A), \ldots, F_i(P).C_n(A)) \\ &= (\sum_{j=1}^{m} p_{ij}a_{j1}, \ldots, \sum_{j=1}^{m} p_{ij}a_{jn}) \\ &= \sum_{j=1}^{m} p_{ij}(a_{j1}, \ldots, a_{jn}).\end{aligned}$$

Luego, cada vector fila de B se puede obtener como combinación lineal de los vectores fila de A, y por lo tanto el espacio fila de B está incluido en el espacio fila de A.

Ahora bien, como P invertible, podemos multiplicar por P^{-1} a izquierda la fórmula $B = PA$, y obtenemos $P^{-1}B = P^{-1}PA = A$. Haciendo el mismo razonamiento que arriba concluimos que también el espacio fila de A está incluido en el espacio fila de B y por lo tanto son iguales. □

Corolario 3.4.2. *Sean* A *matriz* $m \times n$ *y* R *la MRF equivalente por filas a* A. *Entonces, el espacio fila de* A *es igual al espacio fila de* R *y las filas no nulas de* R *forman una base del espacio fila de* A.

Demostración. $R = PA$, donde P es una matriz $m \times m$ invertible, luego, por el teorema anterior, el espacio fila de A es igual al espacio fila de R. Calculemos ahora cual es la dimensión del espacio fila de R. Veamos que filas no nulas de R son LI.

Recordemos que por definición de MRF cada fila no nula comienza con un 1 y en esa coordenada todas las demás filas tienen un 0, por lo tanto una combinación lineal no trivial resulta en un vector no nulo: si v es una fila no nula de R, con el 1 principal en la coordinada i y $\lambda \neq 0$, entonces λv vale λ en la posición i y esta coordenada no puede ser anulada por la combinación de otras filas. □

Corolario 3.4.3. *Sean* A *matriz* $n \times n$. *Entonces,* A *es invertible si y sólo si las filas de* A *son una base de* $\mathbb{K}^n$.

Demostración. Si A es invertible entonces la MERF de A es la identidad, por lo tanto el espacio fila de A genera $\mathbb{K}^n$.

Por otro lado, si el espacio fila de A genera $\mathbb{K}^n$, el espacio fila de la MERF es $\mathbb{K}^n$ y por lo tanto la MERF de A es la identidad y en consecuencia A es invertible.

Hemos probado que A es invertible si y sólo si las n filas de A generan $\mathbb{K}^n$. Como $\dim \mathbb{K}^n = n$, todo conjunto de n generadores es una base. □

El corolario 3.4.2 nos provee un método para encontrar una base de un subespacio de $\mathbb{K}^n$ generado por m vectores: si $v_1, \ldots, v_m \in \mathbb{K}^n$ y $W = \langle v_1, \ldots, v_m \rangle$, consideramos la matriz

$$A = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

donde las filas son los vectores $v_1, \ldots, v_m$. Luego calculamos R, una MRF equivalente por filas a A, y si R tiene r filas no nulas, las r filas no nulas son una base de W y, por consiguiente, $\dim W = r$.

Ejemplo. Encontrar una base de $W = \langle (1, 0, 1), (1, -1, 0), (5, -3, 2) \rangle$.

Solución. Formemos la matriz cuyas filas son los vectores que generan W, es decir

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

Entonces

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 5 & -3 & 2 \end{bmatrix} \xrightarrow[F_3-5F_1]{F_2-F_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & -3 & -3 \end{bmatrix} \xrightarrow{-F_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & -3 & -3 \end{bmatrix} \xrightarrow{F_3-3F_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Por lo tanto, $\dim W = 2$ y $(1, 0, 1), (0, 1, 1)$ es una base de W. □

El método que nos provee el corolario 3.4.2 nos permite encontrar una base de un subespacio vectorial de $\mathbb{K}^n$ a partir de un conjunto de generadores del subespacio. Como vimos en el teorema 3.3.12, en todo conjunto finito de generadores existe un subconjunto que es una base. El siguiente teorema nos permite encontrar uno de tales subconjuntos.

Teorema 3.4.4. *Sea $v_1, \ldots, v_r$ vectores en $\mathbb{K}^n$ y $W = \langle v_1, \ldots, v_r \rangle$. Sea A la matriz formada por las filas $v_1, \ldots, v_r$ y R una MRF equivalente por filas a A que se obtiene sin el uso de permutaciones de filas (ver observación 2.4.4). Si $i_1, i_2, \ldots, i_s$ son las filas no nulas de R, entonces $v_{i_1}, v_{i_2}, \ldots, v_{i_s}$ es una base de W.*

Demostración. Se hará por inducción sobre r.

Si $r = 1$ es trivial ver que vale la afirmación.

Supongamos que tenemos el resultado probado para $r - 1$ (hipótesis inductiva).

Sea $W' = \langle v_1, \ldots, v_{r-1} \rangle$ y sea A' la matriz formada por las $r - 1$ filas $v_1, \ldots, v_{r-1}$. Sea R' la MRF equivalente por filas a A' que se obtiene sin usar permutaciones de filas. Por hipótesis inductiva, si $i_1, i_2, \ldots, i_s$ son las filas no nulas de R', entonces $v_{i_1}, v_{i_2}, \ldots, v_{i_s}$ es una base de W'.

Sea
$$R_0 = \begin{bmatrix} R' \\ v_r \end{bmatrix}.$$
Si $v_r \in W'$, entonces $v_{i_1}, v_{i_2}, \ldots, v_{i_s}$ es una base de W y
$$R = \begin{bmatrix} R' \\ 0 \end{bmatrix}$$
es la MRF de A.

Si $v_r \notin W'$, entonces $v_{i_1}, v_{i_2}, \ldots, v_{i_s}, v_r$ es una base de W (lema 3.3.7) y la MRF de A tiene la última fila no nula. □

Ejemplo. Sea $S = \{(1,0,1),(1,-1,0),(5,-3,2)\}$ y $W = \langle S \rangle$. Encontrar una base de W que sea un subconjunto de S.

Solución. Hemos visto en el ejemplo de la p. 120 que una MRF de A es
$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$
y que la misma se obtiene sin usar permutaciones. Esta matriz tiene las dos primeras filas no nulas, por lo tanto, $\{(1,0,1),(1,-1,0)\}$ es una base de W. □

Finalmente, terminaremos esta sección con un teorema que resume algunas equivalencias respecto a matrices invertibles.

Teorema 3.4.5. *Sea A matriz $n \times n$ con coeficientes en $\mathbb{K}$. Entonces son equivalentes*

(1) A es invertible.

(2) A es equivalente por filas a Id_n.

(3) A es producto de matrices elementales.

(4) El sistema $AX = Y$ tiene una única solución para toda matriz Y de orden $n \times 1$.

(5) El sistema homogéneo $AX = 0$ tiene una única solución trivial.

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

(7) Las filas de A son LI.

(8) Las columnas de A son LI.

Demostración. Por teoremas 2.7.6 y 2.7.9, tenemos que (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4) $\Leftrightarrow$ (5).

(1) $\Leftrightarrow$ (6). Por teorema 2.8.9.

(1) $\Leftrightarrow$ (7). Por corolario 3.4.3.

(1) $\Leftrightarrow$ (8). A invertible $\Leftrightarrow$ A^t invertible $\Leftrightarrow$ las filas de A^t son LI $\Leftrightarrow$ las columnas de A son LI. $\square$

4 TRANSFORMACIONES LINEALES

Las transformaciones lineales son las funciones con las que trabajaremos en álgebra lineal. Se trata de funciones entre espacios vectoriales que son compatibles con la estructura, es decir con la suma y el producto por escalares.

4.1 TRANSFORMACIONES LINEALES

Definición 4.1.1. Sean V y W dos espacios vectoriales sobre el cuerpo $\mathbb{K}$. Una *transformación lineal* de V en W es una función $T : V \to W$ tal que

(1) $T(v + v') = T(v) + T(v')$, para $v, v' \in V$,

(2) $T(\lambda v) = \lambda T(v)$, para $v \in V$, $\lambda \in \mathbb{K}$.

Observación. $T : V \to W$ es transformación lineal si y sólo si

(a) $T(\lambda v + v') = \lambda T(v) + T(v')$, para $v, v' \in V$, $\lambda \in \mathbb{K}$.

Algunas veces usaremos esto último para comprobar si una aplicación de V en W es una transformación lineal.

Ejemplo. Si V es cualquier espacio vectorial, la transformación identidad Id, definida por $\operatorname{Id} v = v$ ($v \in V$), es una transformación lineal de V en V. La transformación cero 0, definida por $0v = 0$, es una transformación lineal de V en V.

Ejemplo. Sea $T : \mathbb{K}^3 \to \mathbb{K}^2$ definida por

$$T(x_1, x_2, x_3) = (2x_1 - x_3, -x_1 + 3x_2 + x_3).$$

Entonces, T es una transformación lineal. La demostración la veremos en la observación que sigue a este ejemplo.

Observar que si

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

entonces

$$\begin{bmatrix} 2 & 0 & -1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_3 \\ -x_1 + 3x_2 + x_3 \end{bmatrix}.$$

Es decir, si $\mathcal{C}_n$ es la báse canónica de $\mathbb{K}^n$ y $[x]_{\mathcal{C}_3}$ es la matriz de x en la base canónica, entonces

$$A.[x]_{\mathcal{C}_3} = [T(x)]_{\mathcal{C}_2}.$$

Observación 4.1.2. Sea $T : \mathbb{K}^n \to \mathbb{K}^m$. En general si $T(x_1, \ldots, x_n)$ en cada coordenada tiene una combinación lineal de los $x_1, \ldots, x_n$, entonces T es una transformación lineal. Mas precisamente, si T está definida por

$$\begin{aligned} T(x_1, \ldots, x_n) &= (a_{11}x_1 + \cdots + a_{1n}x_n, \ \ldots \ , a_{m1}x_1 + \cdots + a_{mn}x_n) \\ &= (\sum_{j=1}^{n} a_{1j}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j), \end{aligned}$$

con $a_{ij} \in \mathbb{K}$, entonces T es lineal.

Demostración. Se puede hacer directamente como ejercicio. También se demuestra más adelante en la observación 4.2.6. □

Ejemplo 4.1.3 (Transformaciones de $\mathbb{R}^2$ en $\mathbb{R}^2$). Las rotaciones y reflexiones en $\mathbb{R}^2$ son transformaciones lineales. Sea $\theta \in \mathbb{R}$ tal que $0 \leqslant \theta \leqslant 2\pi$, definimos

$$\begin{array}{rccc} R_\theta : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x, y) & \mapsto & (x \cos\theta - y \operatorname{sen}\theta, y\cos\theta + x \operatorname{sen}\theta) \end{array}$$

Observemos que si escribimos el vector (x, y) en coordenadas polares, es decir si

$$(x, y) = r(\cos\alpha, \operatorname{sen}\alpha), \quad r > 0, \ 0 \leqslant \alpha < 2\pi,$$

entonces

$$\begin{aligned} R_\theta(x, y) &= R_\theta(r\cos\alpha, r\operatorname{sen}\alpha) \\ &= (r\cos\alpha\cos\theta - r\operatorname{sen}\alpha\operatorname{sen}\theta, r\operatorname{sen}\alpha\cos\theta + r\cos\alpha\operatorname{sen}\theta) \\ &= (r\cos(\alpha+\theta), r\operatorname{sen}(\alpha+\theta)) \\ &= r(\cos(\alpha+\theta), \operatorname{sen}(\alpha+\theta)). \end{aligned}$$

Por lo tanto $R_\theta(x, y)$ es el vector (x, y) rotado θ grados en sentido antihorario y en consecuencia R_θ es denominada la *rotación antihoraria en* θ *radianes*. No es difícil verificar R_θ es una transformación lineal.

Otras transformaciones lineales importantes de $\mathbb{R}^2$ en $\mathbb{R}^2$ son

$$S_h(x, y) = (x, -y) \quad \text{y} \quad S_v(x, y) = (-x, y).$$

La primera es la reflexión en el eje x y la segunda la reflexión en el eje y. Las siguientes afirmaciones se comprueban algebraicamente en forma sencilla, pero nos podemos convencer de ellas por su interpretación geométrica:

$$\begin{array}{ll} R_\theta \circ R_\varphi = R_{\theta+\varphi}, & \qquad (\text{rotar } \varphi \text{ y } \theta = \text{rotar } \theta + \varphi) \\ R_{\pi/2} \circ S_h \circ R_{-\pi/2} = S_v & \end{array}$$

Ejemplo. Sea $V = \mathbb{R}[x]$ el espacio vectorial de los polinomios con coeficientes reales. Definimos $D : V \to V$, por

$$D(P)(x) = P'(x), \quad x \in \mathbb{R}.$$

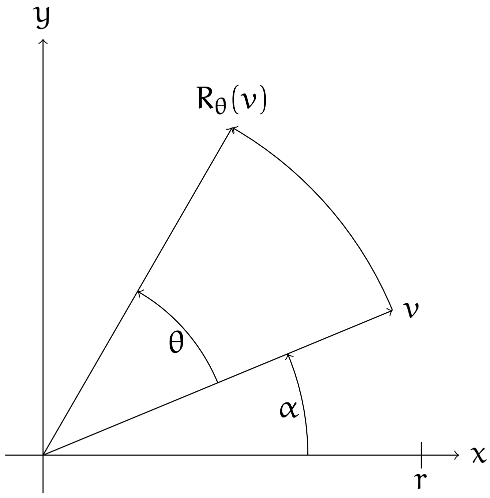


Figura 21: Rotación θ grados.

Observemos primero que la derivada de un polinomio es un polinomio, pues

$$(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0)' = n a_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

Además D es lineal, pues $(f+g)' = f' + g'$ y $(\lambda f)' = \lambda f'$, para f, g funciones derivables y $\lambda \in \mathbb{R}$.

Observación. Sean V y W dos espacios vectoriales sobre el cuerpo $\mathbb{K}$ y $T : V \to W$ un transformación lineal. Entonces $T(0) = 0$

Demostración. $T(0) = T(0+0) = T(0) + T(0)$, por lo tanto

$$-T(0) + T(0) = -T(0) + T(0) + T(0) \;\Rightarrow\; 0 = 0 + T(0) \;\Rightarrow\; 0 = T(0).$$

□

Observación. Las transformaciones lineales preservan combinaciones lineales, es decir si $T : V \to W$ es una transformación lineal, $v_1, \ldots, v_k \in V$ y $\lambda_1, \ldots + \lambda_k \in \mathbb{K}$, entonces

$$T(\lambda_1 v_1 + \cdots + \lambda_k v_k) = \lambda_1 T(v_1) + \cdots + \lambda_k T(v_k).$$

Observar que el caso $k = 2$ se demuestra de la siguiente manera

$$T(\lambda_1 v_1 + \lambda_2 v_2) = T(\lambda_1 v_1) + T(\lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2).$$

El caso general se demuestra por inducción.

Teorema 4.1.4. *Sean V un espacio vectorial de dimensión finita sobre el cuerpo $\mathbb{K}$ y $\{v_1, \ldots, v_n\}$ una base ordenada de V. Sean W un espacio vectorial sobre el mismo cuerpo y $\{w_1, \ldots, w_n\}$, vectores cualesquiera de W. Entonces existe una única transformación lineal T de V en W tal que*

$$T(v_j) = w_j, \quad j = 1, \ldots, n.$$

Demostración. Recordemos que si $v \in V$, existen únicos $a_1, \ldots, a_n \in \mathbb{K}$ (las coordenadas de v) tal que

$$v = a_1 v_1 + \cdots + a_n v_n.$$

Luego para este vector v definimos

$$T(v) = a_1 w_1 + \cdots + a_n w_n.$$

Entonces, T es una correspondencia bien definida que asocia a cada vector v de V un vector $T(v)$ de W. De la definición queda claro que $T(v_j) = w_j$ para cada j. Para ver que T es lineal, sea

$$w = b_1 v_1 + \cdots + b_n v_n,$$

y sea $\lambda \in \mathbb{K}$. Ahora

$$\begin{aligned} \lambda v + w &= \lambda(a_1 v_1 + \cdots + a_n v_n) + b_1 v_1 + \cdots + b_n v_n \\ &= (\lambda a_1 + b_1) v_1 + \cdots + (\lambda a_n + b_n) v_n \end{aligned}$$

con lo que, por definición

$$T(\lambda v + w) = (\lambda a_1 + b_1) w_1 + \cdots + (\lambda a_n + b_n) w_n.$$

Por otra parte

$$\begin{aligned} \lambda T(v) + T(w) &= \lambda(a_1 w_1 + \cdots + a_n w_n) + b_1 w_1 + \cdots + b_n w_n \\ &= (\lambda a_1 + b_1) w_1 + \cdots + (\lambda a_n + b_n) w_n, \end{aligned}$$

y así

$$T(\lambda v + w) = \lambda T(v) + T(w).$$

Finalmente, debemos probar la unicidad de T. Sea $S : V \to W$ transformación lineal tal que $S(v_j) = w_j$ para $1 \leqslant j \leqslant n$. Entonces, si $v \in V$ un vector arbitrario, $v = \sum_i a_i v_i$ y

$$S(v) = S(\sum_i a_i v_i) \sum_i a_i S(v_i) = \sum_i a_i w_i = \sum_i a_i T(v_i) = T(\sum_i a_i v_i) = T(v)$$

□

El teorema 4.1.4 es muy elemental, pero por su importancia ha sido presentado detalladamente.

Ejemplo. Usando el teorema 4.1.4, podemos demostrar la observación 4.1.2 de la siguiente manera: sea $\mathcal{C}_n = \{e_1, \ldots, e_n\}$ es la base canónica de $\mathbb{K}^n$ y sea $T : \mathbb{K}^n \to \mathbb{K}^m$ la única transformación lineal tal que

$$T(e_j) = (a_{1j}, \ldots, a_{mj}), \quad j = 1, \ldots, n$$

Entonces,

$$T(x_1, \ldots, x_n) = (a_{11} x_1 + \cdots + a_{1n} x_n, \ldots, a_{m1} x_1 + \cdots + a_{mn} x_n).$$

es la transformación lineal resultante.

Ejemplo. Los vectores

$$\begin{aligned} v_1 &= (1,2) \\ v_2 &= (3,4) \end{aligned}$$

son linealmente independientes y, por tanto, forman una base de $\mathbb{R}^2$. De acuerdo con el teorema 4.1.4, existe una única transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ tal que

$$\begin{aligned} T(v_1) &= (3,2,1) \\ T(v_2) &= (6,5,4). \end{aligned}$$

Para poder describir T respecto a las coordenadas canónicas debemos calcular $T(e_1)$ y $T(e_2)$, ahora bien,

$$\begin{aligned} (1,0) &= c_1(1,2) + c_2(3,4) \\ (0,1) &= c_3(1,2) + c_4(3,4) \end{aligned}$$

y resolviendo este sistema de cuatro ecuaciones con cuatro incógnitas obtenemos

$$\begin{array}{ccccc} (1,0) & = & -2(1,2) & + & (3,4) \\ (0,1) & = & \frac{3}{2}(1,2) & - & \frac{1}{2}(3,4) \end{array}$$

Luego,

$$\begin{aligned} T(1,0) &= -2T(1,2) + T(3,4) = -2(3,2,1) + (6,5,4) = (0,1,2) \\ T(0,1) &= \frac{3}{2}T(1,2) - \frac{1}{2}T(3,4) = \frac{3}{2}(3,2,1) - \frac{1}{2}(6,5,4) = (\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}) \end{aligned}$$

Entonces

$$T(x_1,x_2) = x_1(0,1,2) + x_2(\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}) = (\frac{3}{2}x_2, x_1 + \frac{1}{2}x_2, 2x_1 - \frac{1}{2}x_2)$$

4.2 NÚCLEO E IMAGEN DE UNA TRANSFORMACIÓN LINEAL

Definición 4.2.1. Sean V, W espacios vectoriales sobre un cuerpo $\mathbb{K}$ y sea $T : V \to W$ una transformación lineal. Definimos

$$\begin{aligned} \mathrm{Im}(T) &:= \{w \in W : \text{existe } v \in V, \text{ tal que } T(v) = w\} = \{T(v) : v \in V\}, \\ \mathrm{Nu}(T) &:= \{v \in V : T(v) = 0\}. \end{aligned}$$

A $\mathrm{Im}(T)$ lo llamamos la *imagen* de T y a $\mathrm{Nu}(T)$ el *núcleo* de T.

Teorema 4.2.2. *Sean V, W espacios vectoriales sobre un cuerpo $\mathbb{K}$ y sea $T : V \to W$ una transformación lineal; entonces $\mathrm{Im}(T) \subset W$ y $\mathrm{Nu}(T) \subset V$ son subespacios vectoriales.*

Demostración. $\mathrm{Im}(T) \neq \emptyset$, pues $0 = T(0) \in \mathrm{Im}(T)$.

Si $T(v_1), T(v_2) \in \mathrm{Im}(T)$ y $\lambda \in \mathbb{K}$, entonces $T(v_1) + T(v_2) = T(v_1 + v_2) \in \mathrm{Im}(T)$ y $\lambda T(v_1) = T(\lambda v_1) \in \mathrm{Im}(T)$.

$\mathrm{Nu}(T) \neq \emptyset$ pues $T(0) = 0$ y por lo tanto $0 \in \mathrm{Nu}(T)$.

Si $v, w \in V$ tales que $T(v) = 0$ y $T(w) = 0$, entonces, $T(v + w) = T(v) + T(w) = 0$. por lo tanto $v + w \in \mathrm{Nu}(T)$. Si $\lambda \in \mathbb{K}$, entonces $T(\lambda v) = \lambda T(v) = \lambda.0 = 0$, luego $\lambda v \in \mathrm{Nu}(T)$. □

Definición 4.2.3. Sean V, W espacios vectoriales sobre un cuerpo $\mathbb{K}$ y sea $T : V \to W$ una transformación lineal. Supongamos que V es de dimensión finita.

(1) El *rango* de T es la dimensión de la imagen de T.

(2) La *nulidad* de T es la dimensión del núcleo de T.

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

$$T(x, y, z) = x + 2y + 3z.$$

Encontrar una base del núcleo y de la imagen.

Solución. Es claro que como T no es 0, la imagen es todo $\mathbb{R}$ (y por lo tanto cualquier $r \in \mathbb{R}, r \neq 0$ es base de la imagen).

Con respecto al núcleo, debemos encontrar una base del subespacio

$$\mathrm{Nu}(T) = \{(x, y, z) : x + 2y + 3z = 0\}.$$

Como $x + 2y + 3z = 0 \Leftrightarrow x = -2y - 3z$, luego,

$$\mathrm{Nu}(T) = \{(-2s - 3t, s, t) : s, t \in \mathbb{R}\}. \tag{4.2.1}$$

Ahora bien, $(-2s - 3t, s, t) = s(-2, 1, 0) + t(-3, 0, 1)$, por lo tanto

$$\mathrm{Nu}(T) =< (-2, 1, 0), (-3, 0, 1) >,$$

y como $(-2, 1, 0), (-3, 0, 1)$ son LI, tenemos que forman una base del núcleo. □

La expresión (4.2.1), que depende de dos parámetros (s y t) que son independientes entre ellos, es llamada la *descripción paramétrica* del núcleo

Todas las transformaciones lineales entre $\mathbb{R}^n$ y $\mathbb{R}^m$ son de la forma "multiplicar por una matriz". Más aún, toda transformación lineal entre espacios vectoriales de dimensión finita se puede expresar de esta forma. Así que analizaremos un poco más en detalle este tipo de transformaciones.

Observación 4.2.4. Sea $A \in \mathbb{R}^{m \times n}$ y consideramos la función T:

$$\begin{array}{rccc} T: & \mathbb{R}^n & \to & \mathbb{R}^m \\ & v & \mapsto & Av. \end{array}$$

Entonces T es una transformación lineal.

Demostración. Debemos ver que T respeta suma y producto por escalares. Sean $v_1, v_2 \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$ entonces

$$T(v_1 + \lambda v_2) = A(v_1 + \lambda v_2) = Av_1 + \lambda A v_2 = T(v_1) + \lambda T(v_2)$$

□

Definición 4.2.5. Sea $A \in \mathbb{R}^{m\times n}$ y sea T la transformación lineal

$$\begin{array}{rccc} T: & \mathbb{R}^n & \to & \mathbb{R}^m \\ & v & \mapsto & Av. \end{array}$$

Diremos que T es *la transformación lineal asociada a* A o *la transformación lineal inducida por* A. Muchas veces denotaremos a esta transformación lineal con el mismo símbolo que la matriz, es decir, en este caso con A.

Ejemplo. Consideremos la matriz $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$.

Entonces si $v = (x, y, z)$,

$$A(v) = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y + z \\ 2x + 2y + 2z \end{bmatrix}$$

En particular, $(1, -1, 0) \in \mathrm{Nu}(A)$ pues $A(1, -1, 0) = 0$ y

$$\begin{aligned} A(1, 0, 0) &= (1, 2) \in \mathrm{Im}(A) \\ A(0, 1, \pi) &= (1 + \pi, 2 + 2\pi) \in \mathrm{Im}(A) \end{aligned}$$

Observación 4.2.6. Sea $T : \mathbb{K}^n \to \mathbb{K}^m$ definida por

$$T(x_1, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, \ \ldots \ , a_{m1}x_1 + \cdots + a_{mn}x_n)$$

con $a_{ij} \in \mathbb{K}$, entonces

$$T(x) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Es decir, T es la transformación lineal inducida por la matriz $A = [a_{ij}]$. Esto, en particular, demuestra la observación 4.1.2.

Proposición 4.2.7. *Sea* $A \in \mathbb{R}^{m\times n}$ *y* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *la transformación lineal asociada. Entonces*

- *El núcleo de* T *es el conjunto de soluciones del sistema homogéneo* $AX = 0$
- *La imagen de* T *es el conjunto de los* $b \in \mathbb{R}^m$ *para los cuales el sistema* $AX = b$ *tiene solución*

Demostración. Se demuestra fácilmente escribiendo las definiciones de los respectivos subconjuntos.

$$v \in \mathrm{Nu}(T) \Leftrightarrow Av = 0 \Leftrightarrow v \text{ es solución de } AX = 0.$$

$$b \in \mathrm{Im}(T) \Leftrightarrow \exists v \in \mathbb{R}^n \text{ tal que } Av = b \Leftrightarrow AX = b \text{ tienen solución.}$$

□

Ejemplo. Sea $T : \mathbb{R}^3 \to \mathbb{R}^4$, definida

$$T(x, y, z) = (x + y,\ x + 2y + z,\ 3y + 3z,\ 2x + 4y + 2z).$$

(1) Describir $\mathrm{Nu}(T)$ en forma paramétrica y dar una base.

(2) Describir $\mathrm{Im}(T)$ en forma paramétrica y dar una base.

Solución. La matriz asociada a esta transformación lineal es

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

Debemos encontrar la descripción paramétrica de

$$\begin{aligned} \mathrm{Nu}(T) &= \{v = (x, y, z) : A.v = 0\} \\ \mathrm{Im}(T) &= \{y = (y_1, y_2, y_3, y_4) : \text{ tal que } \exists v \in \mathbb{R}^3, A.v = y\} \end{aligned}$$

En ambos casos, la solución depende de resolver el sistema de ecuaciones cuya matriz asociada es A:

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & y_1 \\ 1 & 2 & 1 & y_2 \\ 0 & 3 & 3 & y_3 \\ 2 & 4 & 2 & y_4 \end{array}\right] \xrightarrow[F_4 - 2F_1]{F_2 - F_1} \left[\begin{array}{ccc|c} 1 & 1 & 0 & y_1 \\ 0 & 1 & 1 & -y_1 + y_2 \\ 0 & 3 & 3 & y_3 \\ 0 & 2 & 2 & -2y_1 + y_4 \end{array}\right]$$

$$\xrightarrow[\substack{F_3 - 3F_2 \\ F_4 - 2F_2}]{F_1 - F_2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 2y_1 - y_2 \\ 0 & 1 & 1 & -y_1 + y_2 \\ 0 & 0 & 0 & 3y_1 - 3y_2 + y_3 \\ 0 & 0 & 0 & -2y_2 + y_4 \end{array}\right].$$

Luego,

$$T(x, y, z) = (y_1, y_2, y_3, y_4) \quad \Leftrightarrow \quad \left\{ \begin{aligned} x - z &= 2y_1 - y_2 \\ y + z &= -y_1 + y_2 \\ 0 &= 3y_1 - 3y_2 + y_3 \\ 0 &= -2y_2 + y_4 \end{aligned} \right. \qquad (*)$$

Si hacemos $y_1 = y_2 = y_3 = y_4 = 0$, entonces las soluciones del sistema describen el núcleo de T, es decir

$$\begin{aligned} \mathrm{Nu}(T) &= \{(x,y,z) : x - z = 0, y + z = 0\} = \{(s,-s,s) : s \in \mathbb{R}\} \\ &= \{s(1,-1,1) : s \in \mathbb{R}\} \end{aligned}$$

que es la forma paramétrica del $\mathrm{Nu}(T)$. Una base del núcleo de T es $\{(1,-1,1)\}$.

En el sistema (*) las dos primeras ecuaciones no imponen ninguna restricción sobre los y_i (por ejemplo si hacemos $z = 0$ resulta $x = 2y_1 - y_2$, $y = -y_1 + y_2$). Claramente, las últimas dos ecuaciones sí establecen condiciones sobre los y_i y resulta entonces que

$$\mathrm{Im}(T) = \{(y_1,y_2,y_3,y_4) : \text{ tal que } 0 = 3y_1 - 3y_2 + y_3 \text{ y } 0 = -2y_2 + y_4\}$$

Resolviendo este sistema, obtenemos

$$\begin{aligned} \mathrm{Im}(T) &= \{(-\frac{1}{3}s + \frac{1}{2}t, \frac{1}{2}t, s, t) : s,t \in \mathbb{R}\} \\ &= \{s(-\frac{1}{3},0,1,0) + t(\frac{1}{2},\frac{1}{2},0,1) : s,t \in \mathbb{R}\} \end{aligned}$$

que es la descripción paramétrica $\mathrm{Im}(T)$. Es claro que $\{(-\frac{1}{3},0,1,0),(\frac{1}{2},\frac{1}{2},0,1)\}$ es una base de $\mathrm{Im}(T)$. □

He aquí uno de los resultados más importantes del álgebra lineal.

Teorema 4.2.8. *Sean V, W espacios vectoriales sobre un cuerpo $\mathbb{K}$ y sea $T : V \to W$ una transformación lineal. Suponga que V es de dimensión finita. Entonces*

$$\mathit{rango}(T) + \mathit{nulidad}(T) = \dim V.$$

Demostración. Sean

$$\begin{aligned} n &= \dim V \\ k &= \dim(\mathrm{Nu}\, T) = \mathit{nulidad}(T). \end{aligned}$$

Entonces debemos probar que

$$n - k = \dim(\mathrm{Im}\, T) = \mathit{rango}(T).$$

Sea $\{v_1,\ldots,v_k\}$ una base de $\mathrm{Nu}\, T$. Existen vectores $\{v_{k+1},\ldots,v_n\}$, en V tales que $\{v_1,\ldots,v_n\}$ es una base de V. Para probar el teorema, demostraremos que $\{Tv_{k+1},\ldots,Tv_n\}$ es una base para la imagen de T.

(1) $\{Tv_{k+1},\ldots,Tv_n\}$ *genera la imagen de* T.

Si $w \in \mathrm{Im}(T)$, entonces existe $v \in V$ tal que $T(v) = w$, como $\{v_1,\ldots,v_n\}$ es base de V, existen $\lambda_1,\ldots,\lambda_n \in \mathbb{K}$, tal que $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, por lo tanto

$$\begin{aligned} w &= T(v) \\ &= \lambda_1 T(v_1) + \cdots + \lambda_k T(v_k) + \lambda_{k+1} T(v_{k+1}) + \cdots + \lambda_n T(v_n) \\ &= 0 + \cdots + 0 + \lambda_{k+1} T(v_{k+1}) + \cdots + \lambda_n T(v_n) \\ &= \lambda_{k+1} T(v_{k+1}) + \cdots + \lambda_n T(v_n). \end{aligned}$$

Por lo tanto, $\{Tv_{k+1}, \ldots, Tv_n\}$ genera la imagen de T.

(2) $\{Tv_{k+1}, \ldots, Tv_n\}$ *es un conjunto linealmente independiente.*

Para ver que $\{Tv_{k+1}, \ldots, Tv_n\}$ es linealmente independiente, suponga que se tienen escalares μ_i tales que

$$\sum_{i=k+1}^{n} \mu_i T v_i = 0,$$

luego

$$0 = \sum_{i=k+1}^{n} \mu_i T v_i = T(\sum_{i=k+1}^{n} \mu_i v_i).$$

Por lo tanto $v = \sum_{i=k+1}^{n} \mu_i v_i \in \operatorname{Nu}(T)$. Como $\{v_1, \ldots, v_k\}$ es una base de $\operatorname{Nu} T$, existen escalares λ_i tales que

$$v = \sum_{i=1}^{k} \lambda_i v_i,$$

es decir

$$\sum_{j=k+1}^{n} \mu_j v_j = \sum_{i=1}^{k} \lambda_i v_i.$$

Luego

$$\begin{aligned} 0 &= \sum_{i=1}^{k} \lambda_i v_i - (\sum_{j=k+1}^{n} \mu_j v_j) \\ &= \lambda_1 v_1 + \cdots + \lambda_k v_k - \mu_{k+1} v_{k+1} - \cdots - \mu_n v_n. \end{aligned}$$

Como $\{v_1, \ldots, v_n\}$ es una base, y por lo tanto un conjunto LI, tenemos que $0 = \lambda_1 = \cdots = \lambda_k = \mu_{k+1} = \cdots = \mu_n$, y en particular $0 = \mu_{k+1} = \cdots = \mu_n$. Por lo tanto $\{Tv_{k+1}, \ldots, Tv_n\}$ es un conjunto linealmente independiente. □

Sea A una matriz $m \times n$ con coeficientes en $\mathbb{K}$. El *rango fila* de A es la dimensión del subespacio de $\mathbb{K}^n$ generado por las filas de A, es decir la dimensión del espacio fila de A. El *rango columna* de A es es la dimensión del subespacio de $\mathbb{K}^m$ generado por las columna de A. Un consecuencia importante del teorema 4.2.8 es le siguiente resultado.

Teorema 4.2.9. *Si A es una matriz $m \times n$ con coeficientes en $\mathbb{K}$, entonces*

$$\textit{rango fila}\ (A) = \textit{rango columna}\ (A).$$

Demostración. Sea T la transformación lineal

$$\begin{aligned} T : \quad \mathbb{K}^{n\times 1} &\rightarrow \mathbb{K}^{m\times 1} \\ X &\mapsto AX. \end{aligned}$$

Observar que
$$\mathrm{Nu}(T) = \{X \in \mathbb{K}^{n\times 1} : AX = 0\}.$$
Es decir $\mathrm{Nu}(T)$ es el subespacio de soluciones del sistema homogéneo $AX = 0$. Ahora bien, si $k = $ *rango fila* (A), ya hemos dicho (capítulo 3, sección 3.4) que la dimensión del subespacio de soluciones del sistema homogéneo $AX = 0$ es $n - k$. Luego
$$\textit{rango fila } (A) = \dim V - \textit{nulidad}(T). \tag{4.2.2}$$

Por otro lado
$$\mathrm{Im}(T) = \{AX : X \in \mathbb{K}^{n\times 1}\}.$$
Ahora bien,
$$AX = \begin{bmatrix} a_{11}x_1 + \cdots a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots a_{mn}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$
Es decir, que la imagen de T es el espacio generado por las columnas de A. Por tanto,
$$\textit{rango}(T) = \textit{ rango columna } (A).$$
Por el teorema 4.2.8
$$\textit{rango}(T) = \dim V - \textit{nulidad}(T),$$
y por lo tanto
$$\textit{rango columna } (A) = \dim V - \textit{nulidad}(T). \tag{4.2.3}$$

Obviamente, las igualdades (4.2.2) y (4.2.3) implican
$$\textit{rango fila } (A) = \textit{rango columna } (A).$$
□

Definición 4.2.10. Si A es una matriz $m \times n$ con coeficientes en $\mathbb{K}$, entonces el *rango* de A es el rango fila de A (que es igual al rango columna).

4.3 ISOMORFISMOS DE ESPACIOS VECTORIALES

Definición 4.3.1. Sean V, W espacios vectoriales sobre un cuerpo $\mathbb{K}$ y sea $T : V \to W$ una transformación lineal.

(1) T es *epimorfismo* si T es suryectiva, es decir si $\mathrm{Im}(T) = W$.

(2) T es *monomorfismo* si T es inyectiva (o 1-1), es decir si dados $v_1, v_2 \in V$ tales que $T(v_1) = T(v_2)$, entonces $v_1 = v_2$.

Observación. T es epimorfismo si y sólo si

$$T \text{ es lineal y } \forall w \in W,\ \exists v \in V \text{ tal que } T(v) = w.$$

Esto se deduce inmediatamente de la definiciones de función suryectiva y de $\operatorname{Im}(T)$

T es monomorfismo si y sólo si

$$T \text{ es lineal y } \forall v_1, v_2 \in V:\ v_1 \neq v_2 \Rightarrow T(v_1) \neq T(v_2).$$

Esto se obtiene aplicando el contrarrecíproco a la definición de función inyectiva.

Proposición 4.3.2. *Sea $T : V \to W$ una transformación lineal. Entonces T es monomorfismo si y sólo si $\operatorname{Nu}(T) = 0$.*

Demostración. ($\Rightarrow$) Debemos ver que $\operatorname{Nu}(T) = 0$, es decir que si $T(v) = 0$, entonces $v = 0$. Ahora bien, si $T(v) = 0$, como $T(0) = 0$, tenemos que $T(v) = T(0)$, y como T es inyectiva, implica que $v = 0$.

($\Leftarrow$) Sean $v_1, v_2 \in V$ tal que $T(v_1) = T(v_2)$. Entonces

$$0 = T(v_1) - T(v_2) = T(v_1 - v_2).$$

Por lo tanto, $v_1 - v_2 \in \operatorname{Nu}(T)$. Por hipótesis, tenemos que $v_1 - v_2 = 0$, es decir $v_1 = v_2$. $\square$

Observación. Sea $T : V \to W$ transformación lineal,

(1) T es epimorfismo si y sólo si $\operatorname{Im}(T) = W$ si y solo si $\mathrm{rango}(T) = \dim W$.

(2) T es monomorfismo si y sólo si $\operatorname{Nu}(T) = 0$ si y sólo si $\mathrm{nulidad}(T) = 0$.

Proposición 4.3.3. *Sea $T : V \to W$ transformación lineal. Entonces,*

(1) T es monomorfismo si y sólo si T de un conjunto LI es LI.

(2) T es epimorfismo si y sólo si T de un conjunto de generadores de V es un conjunto de generadores de W.

Demostración. Haremos la demostración para el caso de dimensión finita, pero en el caso general la demostración es similar.

(1) ($\Rightarrow$) Sea $\{v_1, \ldots, v_n\}$ un conjunto LI en V y sean $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tales que

$$\lambda_1 T(v_1) + \cdots + \lambda_n T(v_n) = 0,$$

entonces

$$0 = T(\lambda_1 v_1 + \cdots + \lambda_n v_n).$$

Como T es inyectiva, por proposición 4.3.2,

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0,$$

lo cual implica que $\lambda_1, \ldots, \lambda_n$ son todos nulos. Por lo tanto, $T(v_1), \ldots, T(v_n)$ son LI.

(1) ($\Leftarrow$) Sea $v \in V$ tal que $T(v) = 0$. Veremos que eso implica que $v = 0$. Ahora bien, sea $\{v_1, \ldots, v_n\}$ una base de V, entonces existen $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tales que

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n,$$

por lo tanto

$$0 = T(v) = T(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 T(v_1) + \cdots + \lambda_n T(v_n).$$

Como $\{v_1, \ldots, v_n\}$ es LI, por hipótesis, $\{T(v_1), \ldots, T(v_n)\}$ es LI y, por lo tanto, $\lambda_1, \ldots, \lambda_n$ son todos nulos. Luego $v = 0$. Es decir probamos que el núcleo de T es 0, luego por proposición 4.3.2, T es monomorfismo.

(1) ($\Leftarrow$ *alternativa*) Sea $v \in V$ tal que $T(v) = 0$. Si $v \neq 0$, entonces $\{v\}$ es un conjunto LI en V. Luego, $\{T(v)\}$ es un conjunto LI en W y por lo tanto $T(v) \neq 0$. Así, si $T(v) = 0$ entonces $v = 0$ y por lo tanto T es un monomorfismo.

(2) ($\Rightarrow$) Sea $\{v_1, \ldots, v_n\}$ un conjunto de generadores de V y sea $w \in W$. Como T es epimorfismo, existe $v \in V$ tal que $T(v) = w$. Ahora bien,

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n, \;\text{ para algún } \lambda_1, \ldots, \lambda_n \in \mathbb{K},$$

por lo tanto,

$$w = T(v) = T(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 T(v_1) + \cdots + \lambda_n T(v_n).$$

Es decir, cualquier $w \in W$ se puede escribir como combinación lineal de los $T(v_1), \ldots, T(v_n)$ y, por lo tanto, generan W.

(2) ($\Leftarrow$) Sea $\{v_1, \ldots, v_n\}$ una base de V, por hipótesis $T(v_1), \ldots, T(v_n)$ generan W, es decir dado cualquier $w \in W$, existen $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tales que

$$w = \lambda_1 T(v_1) + \cdots + \lambda_n T(v_n),$$

y por lo tanto $w = T(v)$, con

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

□

Definición 4.3.4. Si V y W son espacios vectoriales sobre el cuerpo $\mathbb{K}$, toda transformación lineal $T : V \to W$ suryectiva e inyectiva, se dice *isomorfismo* de V sobre W. Si existe un isomorfismo de V sobre W, se dice que V es *isomorfo* a W y se denota $V \cong W$.

Observe que V es trivialmente isomorfo a V, ya que el operador identidad es un isomorfismo de V sobre V.

Proposición 4.3.5. *Sea* $T : V \to W$ *transformación lineal. Entonces* T *es un isomorfismo si y solo si* T *de una base de* V *es una base de* W.

Demostración. ($\Rightarrow$) Sea $\mathcal{B}$ base de V. Como T es isomorfismo, T es mono y epi, luego por proposición 4.3.3, $T(\mathcal{B})$ es LI y genera W, es decir, es base de W.

($\Leftarrow$) Sea $\mathcal{B}$ base de V y $T : V \to W$ transformación lineal tal que $T(\mathcal{B})$ es base. Por lo tanto, manda un conjunto LI a un conjunto LI y un conjunto de generadores de V a un conjunto de generadores de W. Por proposición 4.3.3, T es mono y epi, por lo tanto T es un isomorfismo. □

Corolario 4.3.6. *Sean* V *y* W *dos* $\mathbb{K}$*-espacios vectoriales de dimensión finita tal que* V *es isomorfo a* W. *Entonces* $\dim(V) = \dim(W)$.

Demostración. Como V es isomorfo a W, existe un isomorfismo $T : V \to W$. Por la proposición anterior si $v_1, \ldots, v_n$ es base de V, entonces $T(v_1), \ldots, T(v_n)$ es base de W. Por lo tanto, $\dim(V) = n = \dim(W)$. □

Recordemos que si una función $f : X \to Y$ es suryectiva e inyectiva, es decir biyectiva, existe su inversa, la cual también es biyectiva. La inversa se denota $f^{-1} : Y \to X$ y viene definida por

$$f^{-1}(y) = x \Leftrightarrow f(x) = y.$$

Teorema 4.3.7. *Sea* $T : V \to W$ *un isomorfismo. Entonces* $T^{-1} : W \to V$ *es lineal y, por lo tanto, también es un isomorfismo.*

Demostración.

Sean $w_1, w_2 \in W$, probemos que $T^{-1}(w_1 + w_2) = T^{-1}(w_1) + T^{-1}(w_2)$.

Sean $v_1 = T^{-1}(w_1)$, $v_2 = T^{-1}(w_2)$. Por lo tanto $T(v_1) = w_1$ y $T(v_2) = w_2$. Ahora bien,

$$\begin{aligned} T^{-1}(w_1 + w_2) = T^{-1}(T(v_1) + T(v_2)) = T^{-1}(T(v_1 + v_2)) = \\ = (T^{-1} \circ T)(v_1 + v_2) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2). \end{aligned}$$

Sean $w \in W$ y $\lambda \in \mathbb{K}$, probemos que $T^{-1}(\lambda w) = \lambda T^{-1}(w)$.

Sea $v = T^{-1}(w)$, entonces

$$T^{-1}(\lambda w) = T^{-1}(\lambda T(v)) = T^{-1}(T(\lambda v)) = (T^{-1} \circ T)(\lambda v) = \lambda v = \lambda T^{-1}(w).$$

□

Ejercicio. Sean V, W y Z espacios vectoriales sobre el cuerpo $\mathbb{K}$ y sean $T : V \to W$, $S : W \to Z$ isomorfismos. Entonces,

(1) $S \circ T : V \to Z$ también es un isomorfismo y

(2) $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.

Como ya se ha dicho, V es isomorfo a V vía la identidad. Por el teorema anterior, si V es isomorfo a W, entonces W es isomorfo a V. Por el ejercicio anterior, si V es isomorfo a W y W es isomorfo a Z, entonces V es isomorfo a Z. En resumen, el isomorfismo es una relación de equivalencia sobre la clase de espacios vectoriales. Si existe un isomorfismo de V sobre W, se dirá a veces que *V y W son isomorfos*, en vez de que V es isomorfo a W. Ello no será motivo de confusión porque V es isomorfo a W, si, y solo si, W es isomorfo a V.

Teorema 4.3.8. *Sean V, W espacios vectoriales de dimensión finita sobre $\mathbb{K}$ tal que $\dim V = \dim W$. Sea $T : V \to W$ transformación lineal. Entonces, son equivalentes:*

(a) *T es un isomorfismo.*

(b) *T es monomorfismo.*

(c) *T es epimorfismo.*

(d) *Si $\{v_1, \ldots, v_n\}$ es una base de V, entonces $\{T(v_1), \ldots, T(v_n)\}$ es una base de W.*

Demostración (*). Sea $n = \dim V = \dim W$.

(a) $\Rightarrow$ (b). Como T es isomorfismo, es biyectiva y por lo tanto inyectiva.

(b) $\Rightarrow$ (c). T monomorfismo, entonces $\mathrm{nulidad}(T) = 0$ (proposición 4.3.2). Luego, como $\mathrm{rango}(T) + \mathrm{nulidad}(T) = \dim V$, tenemos que $\mathrm{rango}(T) = \dim V$. Como $\dim V = \dim W$, tenemos que $\dim \mathrm{Im}(T) = \dim W$ y por lo tanto $\mathrm{Im}(T) = \dim W$. En consecuencia, T es suryectiva.

(c) $\Rightarrow$ (a). T es suryectiva, entonces $\mathrm{rango}(T) = n$, luego $\mathrm{nulidad}(T) = 0$, por lo tanto $\mathrm{Nu}(T) = 0$ y en consecuencia T es inyectiva. Como T es suryectiva e inyectiva es un isomorfismo.

Hasta aquí probamos que (a), (b) y (c) son equivalentes, luego si probamos que (a), (b) o (c) $\Rightarrow$ (d) y que (d) $\Rightarrow$ (a), (b) o (c), estaría probado el teorema.

(a) $\Rightarrow$ (d). Sea $\{v_1, \ldots, v_n\}$ una base de V, entonces $\{v_1, \ldots, v_n\}$ es LI y genera V. Por proposición 4.3.3, tenemos que $\{T(v_1), \ldots, T(v_n)\}$ es LI y genera W, por lo tanto $\{T(v_1), \ldots, T(v_n)\}$ es una base de W.

(d) $\Rightarrow$ (a). Como T de una base es una base, entonces T de un conjunto LI es un conjunto LI y T de un conjunto de generadores de V es un conjunto de generadores de W. Por lo tanto, por proposición 4.3.3, T es monomorfismo y epimorfismo, luego T es un isomorfismo. $\square$

Corolario 4.3.9. *Sean V, W espacios vectoriales de dimensión finita sobre $\mathbb{K}$ tal que $\dim V = \dim W$. Entonces V y W son isomorfos.*

Demostración. Sea $\{v_1, \ldots, v_n\}$ es una base de V y $\{w_1, \ldots, w_n\}$ es una base de W. Poe teorema 4.1.4 existe una única transformación lineal $T : V \to W$ tal que

$$T(v_i) = w_i, \qquad i = 1, \ldots, n.$$

Por el teorema anterior, T es un isomorfismo. $\square$

Ejemplo. $\mathbb{K}_n[x] = \{a_0 + a_1x + \cdots + a_{n-1}x^{n-1} : a_0, a_1, \ldots, a_{n-1} \in \mathbb{K}\}$ es isomorfo a $\mathbb{K}^n$, esto es consecuencia inmediata del corolario anterior, pues ambos tienen dimensión n. Explícitamente, $1, x, \ldots, x^{n-1}$ es base de $\mathbb{K}_n[x]$ y sea $e_1, \ldots, e_n$ la base canónica de $\mathbb{K}^n$, entonces un isomorfismo de $\mathbb{K}_n[x]$ a $\mathbb{K}^n$ viene dado por la única transformación lineal $T : \mathbb{K}_n[x] \to \mathbb{K}^n$ tal que

$$T(x^i) = e_{i+1}, \qquad i = 0, \ldots, n-1.$$

Ejemplo. $M_{m\times n}(\mathbb{K})$ es isomorfo a $\mathbb{K}^{mn}$. El isomorfismo viene dado por $T : M_{m\times n}(\mathbb{K}) \to \mathbb{K}^{mn}$ tal que

$$T(E_{ij}) = e_{(i-1)n+j}, \qquad i = 1, \ldots, m,\ j = 1, \ldots, n.$$

Por ejemplo, en el caso 2×2,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mapsto (1, 0, 0, 0) \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mapsto (0, 1, 0, 0)$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \mapsto (0, 0, 1, 0) \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \mapsto (0, 0, 0, 1).$$

4.4 ÁLGEBRA DE LAS TRANSFORMACIONES LINEALES (*)

En el estudio de las transformaciones lineales de V en W es de fundamental importancia que el conjunto de estas transformaciones hereda una estructura natural de espacio vectorial. El conjunto de las transformaciones lineales de un espacio V en sí mismo tiene incluso una estructura algebraica mayor, pues la composición ordinaria de funciones da una “multiplicación” de tales transformaciones.

Observemos primero que si X conjunto y W espacio vectorial sobre el cuerpo $\mathbb{K}$, entonces

$$F(X, W) := \{f : X \to W\},$$

es decir el conjunto de funciones de X en W es un espacio vectorial sobre $\mathbb{K}$ con la suma y el producto por escalares definido:

$$\begin{aligned} (f+g)(x) &= f(x) + g(x), & f, g &\in F(X, W),\ x \in X \\ (\lambda f)(x) &= \lambda f(x), & f &\in F(X, W),\ x \in X,\ \lambda \in \mathbb{K}. \end{aligned}$$

La demostración de esto es sencilla y se basa en el hecho que W es un espacio vectorial.

Teorema 4.4.1. *Sean V y W espacios vectoriales sobre el cuerpo $\mathbb{K}$ Sean $T, S : V \to W$ transformaciones y $\mu \in \mathbb{K}$. Entonces, $T + S$ y μT son transformaciones lineales de V en W.*

Demostración. Sean $v, v' \in V$ y $\lambda \in \mathbb{K}$, entonces

$$\begin{aligned}
(T+S)(\lambda v + v') &= T(\lambda v + v') + S(\lambda v + v') && \text{(definición de } T+S\text{)}\\
&= \lambda T(v) + T(v') + \lambda S(v) + S(v') && \text{(T y S lineales)}\\
&= \lambda(T(v) + S(v)) + T(v') + S(v') && \\
&= \lambda((T+S)(v)) + (T+S)(v') && \text{(definición de } T+S\text{)}\\
&= \lambda(T+S)(v) + (T+S)(v') && \text{(definición de } \lambda(T+S)\text{).}
\end{aligned}$$

que dice que $T+U$ es una transformación lineal. En forma análoga, si $\mu \in \mathbb{K}$,

$$\begin{aligned}
(\mu T)(\lambda v + v') &= \mu T(\lambda v + v') && \text{(definición de } \mu T\text{)}\\
&= \mu\lambda T(v) + \mu T(v') && \text{(T lineal)}\\
&= \lambda\mu T(v) + \mu T(v') && \\
&= \lambda(\mu T)(v) + (\mu T)(v') && \text{(definición de } \mu T\text{).}
\end{aligned}$$

que dice que μT es una transformación lineal. □

Corolario 4.4.2. *Sean V y W espacios vectoriales sobre el cuerpo $\mathbb{K}$. Entonces, el conjunto de transformaciones lineales de V en W es un subespacio vectorial de $F(V, W)$.*

Se denotará $L(V, W)$ al espacio vectorial de las transformaciones lineales de V en W.

Teorema 4.4.3. *Sean V, W y Z espacios vectoriales sobre el cuerpo $\mathbb{K}$. Sean $T : V \to W$ y $U : W \to Z$ transformaciones lineales. Entonces la función compuesta $U \circ T$ definida por $(U \circ T)(v) = U(T(v))$ es una transformación lineal de V en Z.*

Demostración. Sean $v, v' \in V$ y $\lambda \in \mathbb{K}$, entonces

$$\begin{aligned}
(U \circ T)(\lambda v + v') &= U(T(\lambda v + v')) && \text{(definición de composición)}\\
&= U(\lambda T(v) + T(v')) && \text{(T lineal)}\\
&= \lambda U(T(v)) + U(T(v')) && \text{(U lineal)}\\
&= \lambda(U \circ T)(v) + (U \circ T)(v') && \text{(definición de composición).}
\end{aligned}$$

□

Para simplificar, a veces denotaremos la composición por yuxtaposición, es decir
$$U \circ T = UT.$$

En lo que sigue debemos interesarnos principalmente en transformaciones lineales de un espacio vectorial en sí mismo. Como se tendrá a menudo que escribir "T es una transformación lineal de V en V", se dirá más bien: "T es un operador lineal sobre V".

Definición 4.4.4. Si V es un espacio vectorial sobre el cuerpo $\mathbb{K}$, un *operador lineal* sobre V es una transformación lineal de V en V.

Cuando en el teorema 4.4.3, consideramos $V = W = Z$, tenemos que U y T son opera- dores lineales en el espacio V, y por lo tanto la composición UT es también un operador lineal sobre V. Así, el espacio $L(V, V)$ tiene una "multiplicación" definida por composición. En este caso el operador TU también está definido, y debe observarse que en general $UT \neq TU$, es decir, $UT - TU \neq 0$. Se ha de advertir de manera especial que si T es un operador lineal sobre V, entonces se puede componer T con T. Se usará para ello la notación $T^2 = TT$, y en general $T^n = T \cdots T$ (n veces) para $n = 1, 2, 3, \ldots$ Si $T \neq 0$, se define $T^0 = \mathrm{Id}_V$, el operador identidad.

Lema 4.4.5. *Sea V un espacio vectorial sobre el cuerpo $\mathbb{K}$; sean U, T y S operadores lineales sobre V y sea λ un elemento de $\mathbb{K}$. Denotemos Id_V el operador identidad. Entonces*

(1) $U = \mathrm{Id}_V U = U \mathrm{Id}_V$,

(2) $U(T + S) = UT + US$, $(T + S)U = TU + SU$,

(3) $\lambda(UT) = (\lambda U)T = U(\lambda T)$.

Demostración. (1) es trivial.

Demostraremos $U(T + S) = UT + US$ de (2) y todo lo demás se dejará como ejercicio. Sea $v \in V$, entonces

$$\begin{aligned}
U(T+S)(v) &= U((T+S)(v)) && \text{(definición de composición)}\\
&= U(T(v)+S(v)) && \text{(definición de } T+S)\\
&= U(T(v)) + U(S(v)) && (U \text{ lineal})\\
&= UT(v) + US(v) && \text{(definición de composición).}
\end{aligned}$$

□

El contenido de este lema, y algunos otros resultados sobre composición de funciones de un conjunto en si mismo (como ser la asociatividad), dicen que el espacio vectorial $L(V, V)$, junto con la operación de composición, es lo que se conoce tomo una álgebra asociativa sobre $\mathbb{K}$, con identidad (ver https://es.wikipedia.org/wiki/Álgebra_asociativa).

4.5 COORDENADAS

Una de las características útiles de una base $\mathcal{B}$ en un espacio vectorial V de dimensión n es que permite introducir coordenadas en V en forma análoga a las "coordenadas naturales", x_i, de un vector $v = (x_1, \ldots, x_n)$ en el espacio $\mathbb{K}^n$. En este esquema, las coordenadas de un vector v en V, respecto de la base $\mathcal{B}$, serán los escalares que sirven para expresar v como combinación lineal de los vectores de la base. En el caso de la base canónica $e_1, \ldots, e_n$ de $\mathbb{K}^n$ tenemos

$$v = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i e_i.$$

por lo tanto x_i es la coordenada i-ésima de v respecto a la base canónica.

En forma análoga veremos que si $v_1, \ldots, v_n$ es una base de V, entonces existe una única forma de escribir

$$v = \sum_{i=1}^{n} x_i v_i,$$

y los valores x_i serán las *coordenadas de* v en la base dada.

Definición 4.5.1. Si V es un espacio vectorial de dimensión finita, una *base ordenada* de V es una sucesión finita de vectores linealmente independiente y que genera V.

La diferencia entre la definición de "base" y la de "base ordenada", es que en la última es importante el orden de los vectores de la base. Si la sucesión $v_1, \ldots, v_n$ es una base ordenada de V, entonces el conjunto $\{v_1, \ldots, v_n\}$ es una base de V. La base ordenada es el conjunto, juntamente con el orden dado. Se incurrirá en un pequeño abuso de notación y se escribirá

$$\mathcal{B} = \{v_1, \ldots, v_n\}$$

diciendo que $\mathcal{B}$ es una base ordenada de V.

Proposición 4.5.2. *Sea V espacio vectorial de dimensión finita y sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ordenada de V. Entonces, para cada $v \in V$, existen únicos $x_1, \ldots, x_n \in \mathbb{K}$ tales que*

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Demostración. Como $v_1, \ldots, v_n$ generan V, es claro que existen $x_1, \ldots, x_n \in \mathbb{K}$ tales que $v = x_1 v_1 + \cdots + x_n v_n$. Sean $y_1, \ldots, y_n \in \mathbb{K}$ tales que $v = y_1 v_1 + \cdots + y_n v_n$. Veremos que $x_i = y_i$ para $1 \leqslant i \leqslant n$.

Como $v = \sum_{i=1}^{n} x_i v_i$ y $v = \sum_{i=1}^{n} y_i v_i$, restando miembro a miembro obtenemos

$$0 = \sum_{i=1}^{n} (x_i - y_i) v_i.$$

Ahora bien, $v_1, \ldots, v_n$ son LI, por lo tanto todos los coeficientes de la ecuación anterior son nulos, es decir $x_i - y_i = 0$ para $1 \leqslant i \leqslant n$ y entonces $x_i = y_i$ para $1 \leqslant i \leqslant n$. □

La proposición anterior permite, dada una base ordenada, asociar a cada vector una n-tupla que serán la coordenadas del vector en esa base.

Definición 4.5.3. sea V espacio vectorial de dimensión finita y sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ordenada de V, si $v \in V$ y

$$v = x_1 v_1 + \cdots + x_n v_n,$$

entonces x_i *es la coordenada* i*-ésima de* v y denotamos

$$[v]_{\mathcal{B}} = (x_1, \ldots, x_n).$$

También nos será útil describir a v como una matriz $n \times 1$ y en ese caso hablaremos de *la matriz de* v *en la base* $\mathcal{B}$:

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

(Usamos la misma notación).

Ejemplo. Sea $\mathcal{B} = \{(1,-1), (2,3)\}$ base ordenada de $\mathbb{R}^2$. Encontrar las coordenadas de $(1,0)$ y $(0,1)$ en la base $\mathcal{B}$.

Solución. Debemos encontrar $x_1, x_2 \in \mathbb{R}$ tal que

$$(1,0) = x_1(1,-1) + x_2(2,3).$$

Es decir

$$\begin{aligned} x_1 + 2x_2 &= 1 \\ -x_1 + 3x_2 &= 0. \end{aligned}$$

Resolviendo el sistema de ecuaciones obtenemos $x_1 = \frac{3}{5}$ y $x_2 = \frac{1}{5}$, es decir

$$(1,0) = \frac{3}{5}(1,-1) + \frac{1}{5}(2,3) \quad \text{o equivalentemente} \quad (1,0) = (\frac{3}{5}, \frac{1}{5})_{\mathcal{B}}.$$

De forma análoga podemos ver que

$$(0,1) = -\frac{2}{5}(1,-1) + \frac{1}{5}(2,3) \quad \text{o equivalentemente} \quad (0,1) = (-\frac{2}{5}, \frac{1}{5})_{\mathcal{B}}.$$

□

Proposición 4.5.4. *Sea* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *una base ordenada de* V *un* $\mathbb{K}$*-espacio vectorial. Entonces*

(1) $[v + w]_{\mathcal{B}} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}$, *para* $v, w \in V$,

(2) $[\lambda v]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}}$, *para* $\lambda \in \mathbb{K}$ *y* $v \in V$.

Demostración. (1) Si $v = x_1v_1 + \cdots + x_nv_n$ y $w = y_1v_1 + \cdots + y_nv_n$, entonces

$$v + w = (x_1 + y_1)v_1 + \cdots + (x_n + y_n)v_n,$$

luego

$$[v + w]_{\mathcal{B}} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}.$$

(2) Si $v = x_1v_1 + \cdots + x_nv_n$ y $\in \mathbb{K}$, entonces

$$\lambda v = (\lambda x_1)v_1 + \cdots + (\lambda x_n)v_n,$$

luego

$$[\lambda v]_{\mathcal{B}} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \lambda [v]_{\mathcal{B}}.$$

□

4.6 MATRIZ DE UNA TRANSFORMACIÓN LINEAL

Sea V un espacio vectorial de dimensión n sobre el cuerpo $\mathbb{K}$, y sea W un espacio vectorial de dimensión m sobre $\mathbb{K}$. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ordenada de V, y $\mathcal{B}' = \{w_1, \ldots, w_m\}$ una base ordenada de W. Si T es cualquier transformación lineal de V en W, entonces T está determinada por su efecto sobre los vectores v_j, puesto que todo vector de V es combinación lineal de ellos. Cada uno de los n vectores Tv_j se expresa de manera única como combinación lineal

$$Tv_j = \sum_{i=1}^{m} a_{ij}w_i \qquad (4.6.1)$$

de los w_i. Los escalares $a_{1j}, \ldots, a_{mj}$ son las coordenadas de Tv_j en la base ordenada $\mathcal{B}'$. Por consiguiente, la transformación T está determinada por los mn escalares a_{ij} mediante la expresión (4.6.1).

Definición 4.6.1. Sean V y W espacios vectoriales de dimensión finita con bases ordenadas $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{B}' = \{w_1, \ldots, w_m\}$, respectivamente. Sea $T : V \to W$ una transformación lineal tal que

$$Tv_j = \sum_{i=1}^{m} a_{ij}w_i.$$

A A La matriz $m \times n$ definida por $[A]_{ij} = a_{ij}$ se la denomina *la matriz de* T *respecto a las bases ordenadas* $\mathcal{B}$ *y* $\mathcal{B}'$; y se la denota

$$[T]_{\mathcal{B}\mathcal{B}'} = A.$$

.

Ejemplo. Sea $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida

$$T(x, y, z) = (2x + y, 3y, x + 4z, z).$$

Sean $B = \{e_1, e_2, e_3\}$ la base canónica de $\mathbb{R}^3$ y $B' = \{e_1, e_2, e_3, e_4\}$ la base canónica de $\mathbb{R}^4$. Entonces

$$\begin{array}{rcrcrcrcrcrcr} T(e_1) & = & (2,0,1,0) & = & 2e_1 & + & 0.e_2 & + & e_3 & + & 0.e_4 \\ T(e_2) & = & (1,3,0,0) & = & e_1 & + & 3e_2 & + & 0.e_3 & + & 0.e_4 \\ T(e_3) & = & (0,0,4,1) & = & 0.e_1 & + & 0.e_2 & + & 4e_3 & + & e_4 \end{array}$$

Por lo tanto

$$[T]_{\mathcal{B}\mathcal{B}'} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

Observar que si escribimos los vectores en coordenadas con respecto a las bases canónicas, tenemos que

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + y \\ 3y \\ x + 4z \\ z \end{bmatrix}$$

o más formalmente

$$[T]_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{B}'}.$$

Observación. Recordemos que si $A = [a_{ij}]$ matriz $m \times n$, el operador lineal asociado a A si se define por

$$\begin{array}{rccc} T : & \mathbb{R}^n & \to & \mathbb{R}^m \\ & v & \mapsto & Av. \end{array}.$$

Es decir

$$T(x_1, \ldots, x_n) := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}. \tag{4.6.2}$$

Sean $\mathcal{C}_n$ y $\mathcal{C}_m$ las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente, entonces

$$[T]_{\mathcal{C}_n\mathcal{C}_m} = A.$$

Proposición 4.6.2. *Sea V y W un espacios vectoriales de dimensión n y m respectivamente y sea $T : V \to W$ una transformación lineal. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ordenada de V, y $\mathcal{B}' = \{w_1, \ldots, w_n\}$ una base ordenada de W. Entonces*

$$[T]_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{B}'}, \quad \forall v \in V. \tag{4.6.3}$$

Demostración. Si

$$Tv_j = \sum_{i=1}^{m} a_{ij} w_i$$

entonces $[T]_{ij} = a_{ij}$. Sea $v \in$, entonces $v = x_1 v_1 + \cdots + x_n v_n$ con $x_i \in \mathbb{K}$, por lo tanto

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Ahora bien,

$$\begin{aligned} T(v) = T(\sum_{j=1}^{n} x_j v_j) = \sum_{j=1}^{n} x_j T(v_j) = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij}) w_i = \\ = (\sum_{j=1}^{n} a_{1j}) w_1 + (\sum_{j=1}^{n} a_{2j}) w_2 + \cdots + (\sum_{j=1}^{n} a_{mj}) w_m \end{aligned}$$

y, por lo tanto,

$$[T(v)]_{\mathcal{B}'} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} \\ \sum_{j=1}^{n} a_{2j} \\ \vdots \\ \sum_{j=1}^{n} a_{mj} \end{bmatrix}. \tag{4.6.4}$$

Por otro lado,

$$[T]_{\mathcal{B}\mathcal{B}'} [v]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} \\ \sum_{j=1}^{n} a_{2j} \\ \vdots \\ \sum_{j=1}^{n} a_{mj} \end{bmatrix}. \tag{4.6.5}$$

De las ecuaciones (4.6.4) y (4.6.5) se deduce la formula (4.6.3). □

Teorema 4.6.3 (*). *Sea V y W un espacios vectoriales de dimensión n y m respectivamente y $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{B}' = \{w_1, \ldots, w_m\}$ dos bases ordenadas de V y W respectivamente. Entonces*

$$\kappa : L(V, W) \to M_{m \times n}(\mathbb{K})$$

definida

$$T \mapsto [T]_{\mathcal{B}\mathcal{B}'},$$

es un isomorfismos de espacios vectoriales.

Demostracion. Primero probaremos que κ es lineal y luego que tiene inversa.

Sean $T, T' \in L(V, W)$ y $\lambda \in \mathbb{K}$, veamos que $\kappa(\lambda T + T') = \lambda\kappa(T) + \kappa(T')$, es decir

$$[\lambda T + T']_{\mathcal{B}\mathcal{B}'} = \lambda [T]_{\mathcal{B}\mathcal{B}'} + [T']_{\mathcal{B}\mathcal{B}'}. \tag{4.6.6}$$

Para $1 \leqslant j \leqslant n$, sean

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \qquad \text{y} \qquad T'(v_j) = \sum_{i=1}^{m} a'_{ij} w_i,$$

es decir

$$[T]_{\mathcal{B}\mathcal{B}'} = [a_{ij}] \qquad \text{y} \qquad [T']_{\mathcal{B}\mathcal{B}'} = [a'_{ij}],$$

entonces

$$\begin{aligned}
(\lambda T + T')(v_j) &= \lambda T(v_j) + T'(v_j) \\
&= \lambda \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} a'_{ij} w_i \\
&= \sum_{i=1}^{m} (\lambda a_{ij} + a'_{ij}) w_i,
\end{aligned}$$

por lo tanto

$$[\lambda T + T']_{\mathcal{B}\mathcal{B}'} = [\lambda a_{ij} + a'_{ij}] = \lambda [T]_{\mathcal{B}\mathcal{B}'} + [T']_{\mathcal{B}\mathcal{B}'}$$

y hemos probado (4.6.6) y, en consecuencia, κ es lineal.

Definamos ahora la inversa de κ: sea $A = [a_{ij}]$ matriz $m \times n$ y sea $T : V \to W$ la única transformación lineal que satisface, para $1 \leqslant j \leqslant n$, que

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Es claro que esta aplicación tiene dominio en $M_{m \times n}(\mathbb{K})$ y su imagen está contenida en $L(V, W)$. Más aún, es muy sencillo comprobar que es la aplicación inversa a κ.

□

Teorema 4.6.4. *Sean V, W y Z espacios vectoriales de dimensión finita sobre el cuerpo $\mathbb{K}$; sean $T : V \to W$ y $U : W \to Z$ transformaciones lineales. Si $\mathcal{B}$, $\mathcal{B}'$ y $\mathcal{B}''$ son bases ordenadas de los espacios V, W y Z, respectivamente, entonces*

$$[UT]_{\mathcal{B}\mathcal{B}''} = [U]_{\mathcal{B}'\mathcal{B}''} [T]_{\mathcal{B}\mathcal{B}'}. \tag{4.6.7}$$

Demostración. Sean

$$\mathcal{B} = \{v_1, \ldots, v_n\}, \quad \mathcal{B}' = \{w_1, \ldots, w_m\}, \quad \mathcal{B}'' = \{z_1, \ldots, z_l\}$$

y

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \ 1 \leqslant j \leqslant n; \qquad U(w_i) = \sum_{k=1}^{l} b_{ki} z_k, \ 1 \leqslant i \leqslant m.$$

Es decir

$$[T]_{\mathcal{B}\mathcal{B}'} = [a_{ij}] \qquad \text{y} \qquad [U]_{\mathcal{B}'\mathcal{B}''} = [b_{ij}].$$

Entonces

$$\begin{aligned}(UT)(v_j) &= U(\sum_{i=1}^{m} a_{ij} w_i) \\ &= \sum_{i=1}^{m} a_{ij} U(w_i) \\ &= \sum_{i=1}^{m} a_{ij} \sum_{k=1}^{l} b_{ki} z_k \\ &= \sum_{k=1}^{l} (\sum_{i=1}^{m} b_{ki} a_{ij}) z_k.\end{aligned}$$

Luego el coeficiente kj de la matriz $[UT]_{\mathcal{B}\mathcal{B}''}$ es $\sum_{i=1}^{m} b_{ki} a_{ij}$ que es igual a la fila k de $[U]_{\mathcal{B}'\mathcal{B}''}$ por la columna j de $[T]_{\mathcal{B}\mathcal{B}'}$, en símbolos, si $A = [T]_{\mathcal{B}\mathcal{B}'}$, $B = [U]_{\mathcal{B}'\mathcal{B}''}$ y $C = [UT]_{\mathcal{B}\mathcal{B}''}$, entonces

$$[C]_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij} = F_k(B) C_j(A) = [BA]_{kj}.$$

□

Corolario 4.6.5. *Sean V espacio vectorial de dimensión finita, $\mathcal{B} = \{v_1, \ldots, v_n\}$ base ordenada de V y $T, U : V \to V$ operadores lineales. Entonces*

(1) $[UT]_{\mathcal{B}} = [U]_{\mathcal{B}} [T]_{\mathcal{B}}$.

(2) Si $\mathrm{Id} : V \to V$ es el operador identidad, entonces $[\mathrm{Id}]_{\mathcal{B}} = \mathrm{Id}$, donde Id es la matriz identidad $n \times n$.

(3) Si T es invertible, entonces $[T]_{\mathcal{B}}$ es una matriz invertible y

$$[T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}.$$

Demostración. (1). Es inmediato del teorema anterior tomado $\mathcal{B}' = \mathcal{B}'' = \mathcal{B}$.
(2). $\mathrm{Id}(v_i) = v_i$ y por lo tanto

$$[\mathrm{Id}]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathrm{Id}\,.$$

(3). $\mathrm{Id} = TT^{-1}$, luego

$$\mathrm{Id} = [\mathrm{Id}]_{\mathcal{B}} = [TT^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}} [T^{-1}]_{\mathcal{B}}.$$

Análogamente, $\mathrm{Id} = T^{-1}T$, luego

$$\mathrm{Id} = [\mathrm{Id}]_{\mathcal{B}} = [T^{-1}T]_{\mathcal{B}} = [T^{-1}]_{\mathcal{B}}[T]_{\mathcal{B}}.$$

Por lo tanto $[T]_{\mathcal{B}}^{-1} = [T^{-1}]_{\mathcal{B}}$. □

Teorema 4.6.6 (Matriz de cambio de base). *Sea* V *un espacio vectorial de dimensión finita sobre el cuerpo* $\mathbb{K}$ *y sean*

$$\mathcal{B} = \{v_1, \ldots, v_n\}, \qquad \mathcal{B}' = \{w_1, \ldots, w_n\}$$

bases ordenadas de V. *Sea* T *es un operador lineal sobre* V*. Entonces existe* P *una matriz invertible* $n \times n$ *tal que*

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P.$$

Más aún, $P = [\mathrm{Id}]_{\mathcal{B}'\mathcal{B}}$ *y* $P^{-1} = [\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}$ *, donde* Id *al operador identidad de* V. *Es decir*

$$[T]_{\mathcal{B}'} = [\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}[T]_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}'\mathcal{B}}. \tag{4.6.8}$$

Demostración. Tenemos que $T = \mathrm{Id}\, T$ y $T = T\, \mathrm{Id}$, luego

$$\begin{aligned}
[T]_{\mathcal{B}'\mathcal{B}'} &= [\mathrm{Id}\, T]_{\mathcal{B}'\mathcal{B}'} \\
&= [\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}[T]_{\mathcal{B}'\mathcal{B}} && \text{(teorema 4.6.4)} \\
&= [\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}[T\, \mathrm{Id}]_{\mathcal{B}'\mathcal{B}} \\
&= [\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}[T]_{\mathcal{B}\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}'\mathcal{B}} && \text{(teorema 4.6.4).}
\end{aligned}$$

Por lo tanto $[T]_{\mathcal{B}'} = [\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}[T]_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}'\mathcal{B}}$.
Falta probar que $[\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}$ es el inverso de $[\mathrm{Id}]_{\mathcal{B}'\mathcal{B}}$. Ahora bien,

$$\begin{aligned}
[\mathrm{Id}]_{\mathcal{B}\mathcal{B}'}[\mathrm{Id}]_{\mathcal{B}'\mathcal{B}} &= [\mathrm{Id}]_{\mathcal{B}\mathcal{B}} = \mathrm{Id}, \\
[\mathrm{Id}]_{\mathcal{B}'\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}\mathcal{B}'} &= [\mathrm{Id}]_{\mathcal{B}'\mathcal{B}'} = \mathrm{Id}\,.
\end{aligned}$$

□

La fórmula (4.6.8) es importante por si misma y debemos recordarla.

Definición 4.6.7. Sea V un espacio vectorial de dimensión finita sobre el cuerpo $\mathbb{K}$ y sean $\mathcal{B}$ y $\mathcal{B}'$ bases ordenadas de V. La matriz $P = [\mathrm{Id}]_{\mathcal{B}'\mathcal{B}}$ es llamada la *matriz de cambio de base* de la base $\mathcal{B}'$ a la base $\mathcal{B}$.

La matriz de cambio de base nos permite calcular los cambios de coordenadas: dadas dos bases ordenadas $\mathcal{B}$ y $\mathcal{B}'$, y dadas las coordenadas de v en la base $\mathcal{B}'$ es decir la matriz columna $[v]_{\mathcal{B}'}$ ¿Cómo averiguamos las coordenadas de v en la base $\mathcal{B}$? La respuesta la da la siguiente observación.

Proposición 4.6.8. *Sea* V *un espacio vectorial de dimensión finita sobre el cuerpo* $\mathbb{K}$*, sean* $\mathcal{B}$*,* $\mathcal{B}'$ *bases ordenadas de* V *y* P *la matriz de cambio de base de* $\mathcal{B}'$ *a* $\mathcal{B}$*. Entonces*

$$[v]_{\mathcal{B}} = P\,[v]_{\mathcal{B}'}, \quad \forall v \in V.$$

Demostración. Como $P = [\mathrm{Id}]_{\mathcal{B}'\mathcal{B}}$, por la proposición 4.6.2 tenemos que

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

□

El teorema 4.6.6 nos permite definir el determinante de un operador lineal. Sea V un espacio vectorial de dimensión finita sobre el cuerpo $\mathbb{K}$ y T un operador lineal sobre V. Sean $\mathcal{B}$, $\mathcal{B}'$ bases ordenadas de V, entonces $[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P$, para P una matriz invertible. Por lo tanto,

$$\det([T]_{\mathcal{B}'}) = \det(P^{-1}[T]_{\mathcal{B}}P) = \det([T]_{\mathcal{B}}PP^{-1}) = \det([T]_{\mathcal{B}}).$$

Es decir, el determinante de la matriz de T en cualquier base siempre es igual.

Definición 4.6.9. Sea V un espacio vectorial de dimensión finita sobre el cuerpo $\mathbb{K}$ y T un operador lineal sobre V. El *determinante de* T es el determinante de la matriz de T en alguna base de V.

4.7 OPERADORES DIAGONALIZABLES

Vimos en la sección 2.9 la definición de autovalores y autovectores de una matriz. Por otro lado, en la sección 4.6 vimos que dada una base podemos asignarle a cada transformación lineal una matriz. En esta sección veremos, entro otros temas, los autovalores y autovectores desde una perspectiva de las transformaciones lineales. Por lo dicho anteriormente verán que muchos conceptos y demostraciones se repiten o son similares al caso de la matrices.

Sea V espacio vectorial de dimensión finita. Un operador lineal en V es *diagonalizable* si existe una base ordenada $\mathcal{B} = \{v_1, \ldots, v_n\}$ de V y $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tal que

$$T(v_i) = \lambda_i v_i, \qquad 1 \leqslant i \leqslant n. \tag{4.7.1}$$

En general, los operadores diagonalizables permiten hacer cálculos sobre ellos en forma sencilla, por ejemplo el núcleo del operador definido por (4.7.1) es $\mathrm{Nu}(T) = \langle v_i : \lambda_i = 0 \rangle$ y su imagen es $\mathrm{Im}(T) = \langle v_i : \lambda_i \neq 0 \rangle$ (vermos la demostración de estos resultado más adelante). Otra propiedad importante de los operadores diagonalizables es que la matriz de la transformación lineal en una base adecuada es diagonal (de allí viene el nombre de diagonalizable). En el caso del operador definido por (4.7.1) tenemos que

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

No todo operador lineal es diagonalizable y no es inmediato, ni sencillo, de la definición de un operador lineal decidir si es diagonalizable o no. En esta sección veremos herramientas para estudiar un operador lineal T y su posible diagonalización. La ecuación (4.7.1) sugiere se estudien los vectores que son transformados por T en múltiplos de sí mismos.

Definición 4.7.1. Sea V un espacio vectorial sobre el cuerpo $\mathbb{K}$ y sea T un operador lineal sobre V. Un *valor propio* o *autovalor* de T es un escalar λ de $\mathbb{K}$ tal que existe un vector no nulo $v \in V$ con $T(v) = \lambda v$. Si λ es un autovalor de T, entonces

(1) cualquier $v \in V$ tal que $T(v) = \lambda v$ se llama un *vector propio* o *autovector* de T asociado al valor propio λ;

(2) la colección de todos los $v \in V$ tal que $T(v) = \lambda v$ se llama *espacio propio* o *autoespacio* asociado a λ.

Los valores propios se llaman también a menudo raíces características, eigenvalores. valores característicos o valores espectrales. Nosotros usaremos, preferentemente, "autovalores".

Sea ahora $\lambda \in \mathbb{K}$, definimos

$$V_\lambda := \{v \in V : Tv = \lambda v\}.$$

Observar que $V_\lambda \neq 0$ si y sólo si λ es autovalor y en ese caso V_λ es el autoespacio asociado a λ.

Teorema 4.7.2. *Sea V un espacio vectorial y sea $T : V \to V$ una aplicación lineal. Sea $\lambda \in \mathbb{K}$ entonces, V_λ es subespacio de V.*

Demostración. Sean $v_1, v_2 \in V$ tales que $Tv_1 = \lambda v_1$ y $Tv_2 = \lambda v_2$. Entonces

$$T(v_1 + v_2) = T(v_1) + T(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2),$$

es decir si $v_1, v_2 \in V_\lambda$, probamos que $v_1 + v_2 \in V_\lambda$.

Sea ahora $c \in F$, entonces $T(cv_i) = cT(v_1) = c\lambda v_1 = \lambda(cv_1)$. Por lo tanto, si $v_1 \in V_\lambda$ y $c \in F$, probamos que $cv_1 \in V_\lambda$.

Esto termina de probar el teorema. □

Teorema 4.7.3. *Sea V espacio vectorial y sea $T : V \to V$ una aplicación lineal. Sean $v_1, \ldots, v_m$ autovectores de T, con autovalores $\lambda_1, \ldots, \lambda_m$ respectivamente. Suponga que estos autovalores son distintos entre si, esto es, $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces $v_1, \ldots, v_m$ son linealmente independientes.*

Demostración. Hagamos la demostración por inducción sobre m.

Caso base. Si $m = 1$, no hay nada que demostrar puesto que un vector no nulo el LI.

Paso inductivo. Supongamos que el enunciado es verdadero para el caso $m-1$ con $m>1$, (hipótesis inductiva o HI), y probemos entonces que esto implica que es cierto para m. Debemos ver que si

$$c_1v_1 + c_2v_2 + \cdots c_mv_m = 0 \qquad (*)$$

entonces $c_1 = \cdots c_m = 0$. Multipliquemos $(*)$ por λ_1, obtenemos:

$$c_1\lambda_1v_1 + c_2\lambda_1v_2 + \cdots c_m\lambda_1v_m = 0. \qquad (**)$$

También apliquemos T a $(*)$ y obtenemos

$$c_1\lambda_1v_1 + c_2\lambda_2v_2 + \cdots c_m\lambda_mv_m = 0. \qquad (***)$$

Ahora a $(**)$ le restamos $(***)$ y obtenemos:

$$c_2(\lambda_1-\lambda_2)v_2 + \cdots c_m(\lambda_1-\lambda_m)v_m = 0. \qquad (4.7.2)$$

Como, por hipótesis inductiva, $v_2, \ldots, v_m$ son LI, tenemos que $c_i(\lambda_1-\lambda_i)=0$ para $i \geqslant 2$. Como $\lambda_1-\lambda_i \neq 0$ para $i \geqslant 2$, obtenemos que $c_i = 0$ para $i \geqslant 2$. Por $(*)$ eso implica que $c_1 = 0$ y por lo tanto $c_i = 0$ para todo i. □

Corolario 4.7.4. *Sea V espacio vectorial de dimensión n y sea $T : V \to V$ una aplicación lineal que tiene n autovectores $v_1, \ldots, v_n$ cuyos autovalores $\lambda_1, \ldots, \lambda_n$ son distintos entre si. Entonces $\{v_1, \ldots, v_n\}$ es una base de V.*

Recordemos que si T es una transformación lineal, el determinante de T se define como el determinante de la matriz de la transformación lineal en una base dada y que este determinante no depende de la base.

Definición 4.7.5. Sea $A \in M_n(\mathbb{K})$ el *polinomio característico* de A es

$$\chi_A(x) = \det(A - x\,\mathrm{Id}),$$

donde x es una indeterminada.

Sea V espacio vectorial de dimensión finita y sea $T : V \to V$ lineal, el *polinomio característico de* T es $\chi_T(x) = \det(T - x\,\mathrm{Id})$.

En general, si $A = [a_{ij}]$ matriz $n \times n$, tenemos que

$$\chi_A(x) = \det(A - x\,\mathrm{Id}) = \det \begin{bmatrix} a_{11}-x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-x \end{bmatrix} \qquad (4.7.3)$$

y el polinomio característico de A es un polinomio de grado n, más precisamente

$$\chi_A(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

Esto se puede demostrar fácilmente por inducción.

Ejemplo. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ y su matriz en la base canónica es

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

entonces

$$\det \begin{bmatrix} a-x & b \\ c & d-x \end{bmatrix} = (a-x)(d-x) - bc = x^2 - (a+d)x + (ad-bc).$$

Es decir,

$$\chi_T(x) = x^2 - (a+d)x + (ad-bc).$$

Ejemplo 4.7.6. Consideremos la transformación lineal de $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por (con abuso de notación incluido)

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 10x - 10y + 6z \\ 8x - 8y + 6z \\ -5x + 5y - 3z \end{bmatrix}.$$

Es decir, si $\mathcal{C}$ es la base canónica de $\mathbb{R}^3$,

$$[T]_{\mathcal{C}} = \begin{bmatrix} 10 & -10 & 6 \\ 8 & -8 & 6 \\ -5 & 5 & -3 \end{bmatrix}.$$

Entonces el polinomio característico de T es

$$\det \begin{bmatrix} 10-x & -10 & 6 \\ 8 & -8-x & 6 \\ -5 & 5 & -3-x \end{bmatrix} = -x^3 - x^2 + 6x.$$

Es posible factorizar esta expresión y obtenemos

$$\chi_A(x) = -x(x-2)(x+3).$$

Proposición 4.7.7. *Sea V espacio vectorial de dimensión n y sea $T : V \to V$ lineal. Entonces $\lambda \in \mathbb{K}$ es autovalor si y sólo si λ es raíz del polinomio característico.*

Demostración.
($\Rightarrow$) Si λ es autovalor, entonces existe $v \in V$, no nulo, tal que $Tv = \lambda v$, luego

$$0 = \lambda Tv - v = Tv - \lambda \operatorname{Id} v = (T - \lambda \operatorname{Id})v.$$

Por lo tanto, $T - \lambda \operatorname{Id}$ no es invertible, lo cual implica que $0 = \det(T - \lambda \operatorname{Id}) = \chi_T(\lambda)$. Es decir, λ es raíz del polinomio característico.

($\Leftarrow$) Si λ es raíz del polinomio característico, es decir si $0 = \chi_T(\lambda) = \det(T - \lambda \operatorname{Id})$, entonces $T - \lambda \operatorname{Id}$ no es una transformación lineal invertible, por lo tanto su núcleo es no trivial. Es decir existe $v \in V$ tal que $(T - \lambda \operatorname{Id})v = 0$, luego $Tv = \lambda v$, por lo tanto v es autovector con autovalor λ. □

Repetimos ahora algunos conceptos ya expresados al comienzo de la sección.

Definición 4.7.8. Sea V espacio vectorial de dimensión finita y sea $T : V \to V$ lineal. Diremos que T *es diagonalizable* si existe una base de V de autovectores de T.

En el caso que T sea una transformación lineal diagonalizable y $\mathcal{B} = \{v_1, \ldots, v_n\}$ sea una base de autovectores con autovalores $\lambda_1, \ldots, \lambda_n$, entonces

$$T(v_i) = \lambda_i v_i, \qquad 1 \leqslant i \leqslant n,$$

y, por lo tanto, la matriz de T en la base $\mathcal{B}$ es diagonal, más precisamente

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

Ejemplo. Consideremos la transformación lineal de $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida en el ejemplo 4.7.6.

Ya vimos que el polinomio característico de esta aplicación es

$$\chi_A(t) = t(t-2)(t+3).$$

Luego, por proposición 4.7.7, los autovalores de A son 0, 2 y -3. Debido al corolario 4.7.4 existe una base de autovectores de T. Veamos cuales son. Si λ autovalor de T, para encontrar los autovectores con autovalor λ debemos resolver la ecuación $Tv - \lambda v = 0$, en este caso sería

$$\begin{bmatrix} 10-\lambda & -10 & 6 \\ 8 & -8-\lambda & 6 \\ -5 & 5 & -3-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

para $\lambda = 0, 2, -3$. Resolviendo estos tres sistemas, obtenemos que

$$V_0 = \{(y, y, 0) : y \in \mathbb{R}\}, \quad V_2 = \{(-2z, -z, z) : z \in \mathbb{R}\}, \quad V_{-3} = \{(-2z, -2z, z) : z \in \mathbb{R}\}.$$

Por lo tanto, $\{(1, 1, 0), (-2, -1, 1), (-2, -2, 1)\}$ es una base de autovectores de la transformación lineal.

Proposición 4.7.9. *Sea V espacio vectorial de dimensión n y sea $T : V \to V$ lineal tal que tiene una base de autovectores $\mathcal{B} = \{v_1, \ldots, v_n\}$ con autovalores $\lambda_1, \ldots, \lambda_n$. Entonces* $\operatorname{Nu}(T) = \langle v_i : \lambda_i = 0 \rangle$ *e* $\operatorname{Im}(T) = \langle v_i : \lambda_i \neq 0 \rangle$.

Demostración. Reordenemos la base de tal forma que $\lambda_i = 0$ para $1 \leqslant i \leqslant k$ y $\lambda_i \neq 0$ para $k < i \leqslant n$. Todo $v \in V$ se escribe en términos de la base como

$$v = x_1 v_1 + \cdots + x_k v_k + x_{k+1} v_{k+1} + \cdots + x_n v_n, \quad (x_i \in \mathbb{K}),$$

y entonces
$$T(v) = \lambda_{k+1}x_{k+1}v_{k+1} + \cdots + \lambda_n x_n v_n. \tag{4.7.4}$$
Luego, $T(v) = 0$ si y sólo si $x_{k+1} = \cdots = x_n = 0$, y esto se cumple si y solo si $v = x_1v_1 + \cdots + x_kv_k$, es decir $v \in \langle v_i : \lambda_i = 0\rangle$. También es claro por la ecuación (4.7.4) que

$$\begin{aligned}\operatorname{Im}(T) &= \{\lambda_{k+1}x_{k+1}v_{k+1} + \cdots + \lambda_n x_n v_n : x_i \in \mathbb{K}\}\\ &= \{\mu_{k+1}v_{k+1} + \cdots + \mu_n v_n : \mu_i \in \mathbb{K}\}\\ &= \langle v_i : \lambda_i \neq 0\rangle.\end{aligned}$$

□

Ejemplo. Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el operador definido por $T(x, y) = (y, x)$. Probar que T es diagonalizable y encontrar una base de autovectores.

Demostración. Por la proposición 4.7.7, los autovalores de T son las raíces del polinomio característico, es decir las raíces de

$$\chi_T(\lambda) = \det\begin{bmatrix}-\lambda & 1\\ 1 & -\lambda\end{bmatrix} = \lambda^2 - 1 = (\lambda-1)(\lambda+1).$$

Luego los autovalores son 1 y -1. Para hallar un autovector con autovalor 1 debemos resolver la ecuación $T(x, y) = (x, y)$. Ahora bien,

$$(x, y) = T(x, y) = (y, x),$$

luego $x = y$ y claramente $(1, 1)$ es autovector con autovalor 1.

Por otro lado $T(x, y) = -(x, y)$, implica que $(y, x) = -(x, y)$, es decir $y = -x$ y claramente podemos elegir $(1, -1)$ como autovector con autovalor -1.

Luego $\mathcal{B} = \{(1, 1), (1, -1)\}$ es una base de $\mathbb{R}^2$ de autovectores de T. □

No todas la matrices son diagonalizables, como veremos en el ejemplo a continuación.

Ejemplo. Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el operador definido por $T(x, y) = (2x - y, x + 4y)$. Probar que T tiene un único autovalor λ cuyo autoespacio $V_\lambda = \{v \in \mathbb{R}^2 : Tv = \lambda v\}$ es de dimensión 1.

Demostración. La matriz de T en la base canónica es

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

Por la proposición 4.7.7, los autovalores de T son las raíces del polinomio característico, es decir las raíces de

$$\det\begin{bmatrix}2-x & -1\\ 1 & 4-x\end{bmatrix} = (2-x)(4-x) + 1 = x^2 - 6x + 9 = (x-3)^2.$$

Es decir el único autovalor posible es 3.

Debemos ver para que valores $(x, y) \in \mathbb{R}^2$ se satisface la ecuación

$$T(x, y) = 3(x, y).$$

tiene solución. Esta ecuación es equivalente a

$$\begin{array}{rcll} (2x - y, x + 4y) & = & (3x, 3y) & \Rightarrow \\ 2x - y = 3x & , & x + 4y = 3y & \Rightarrow \\ -y = x & , & x = -y & \Rightarrow \\ y & = & -x & \end{array}$$

Luego $V_3 = \{(x, -x) : x \in \mathbb{R}\}$ que es de dimensión 1.

□

Proposición 4.7.10. *Sea* T *un operador lineal diagonalizable sobre un espacio vectorial* V *de dimensión finita. Sean* $\lambda_1, \ldots, \lambda_k$ *los autovalores distintos de* T*. Entonces, el polinomio característico de* T *es*

$$\chi_T(x) = (\lambda_1 - x)^{d_1} \ldots (\lambda_k - x)^{d_k}$$

con

$$d_i = \dim V_{\lambda_i},$$

para $i = 1, \ldots, k$.

Demostración (∗). T es un operador lineal diagonalizable y $\lambda_1, \ldots, \lambda_k$ los valores propios distintos de T. Entonces existe una base ordenada $\mathcal{B}$ con respecto a la cual T está representado por una matriz diagonal; es decir, los elementos de la diagonal son los escalares λ_j cada uno de los cuales se repite un cierto número de veces. Más específicamente, si $v_{j1}, \ldots, v_{jd_j}$ son los vectores en $\mathcal{B}$ con autovalor λ_j ($1 \leqslant j \leqslant k$), reordenamos la base de tal forma que primero estén los autovectores con autovalor λ_1, a continuación los de autovalor λ_2, etc.:

$$\mathcal{B} = \{v_{11}, \ldots, v_{1d_1}, \ldots, v_{k1}, \ldots, v_{kd_k}\}.$$

Ahora bien, si $v \in V$, entonces

$$\begin{aligned} v &= x_1 v_{11} + \cdots + x_{d_1} v_{1d_1} + \cdots + x_n v_{n1} + \cdots + x_{d_n} v_{nd_n} \\ &= v_1 + v_2 + \cdots + v_k \end{aligned}$$

con $v_i = x_i v_{i1} + \cdots + x_{d_i} v_{id_i} \in V_{\lambda_i}$. Luego

$$T(v) = \lambda_1 v_1 + \lambda_2 v_2 \cdots + \lambda_k v_k \tag{4.7.5}$$

Veamos que $V_{\lambda_i} =< v_{i1}, \ldots, v_{id_i} >$ para $1 \leqslant i \leqslant k$. Es claro que $< v_{i1}, \ldots, v_{id_i} > \subset V_{\lambda_i}$. Probemos ahora que, $V_{\lambda_i} \subset < v_{i1}, \ldots, v_{id_i} >$: si $v \in V_{\lambda_i}$, entonces $T(v)$ es como en (4.7.5) y, por lo tanto, si $v_j \neq 0$ para $j \neq i$ entonces

$T(v) \neq \lambda_j v$, lo que contradice la hipótesis. Es decir $v = v_i \in < v_{i1}, \ldots, v_{id_i} >$. Hemos probado que $V_{\lambda_i} =< v_{i1}, \ldots, v_{id_i} >$ y como $v_{i1}, \ldots, v_{id_i}$ son LI, entonces $\dim V_{\lambda_i} = d_i$.

Por otro lado, la matriz de T en la base $\mathcal{B}$ tiene la forma

$$\begin{bmatrix} \lambda_1 \operatorname{Id}_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 \operatorname{Id}_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \operatorname{Id}_n \end{bmatrix}$$

donde Id_j es la matriz identidad $d_j \times d_j$. Luego, el polinomio característico de T es el producto

$$(\lambda_1 - x)^{d_1} \ldots (\lambda_k - x)^{d_k}.$$

□

Ejemplo. Sea T un operador lineal sobre $\mathbb{R}^3$ representado en la base ordenada canónica por la matriz

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

El polinomio característico de A es

$$\chi_A(x) = \det \begin{bmatrix} 5-x & -6 & -6 \\ -1 & 4-x & 2 \\ 3 & -6 & -4-x \end{bmatrix} = -x^3 + 5x^2 - 8x + 4 = -(x-2)^2(x-1).$$

¿Cuáles son las dimensiones de los espacios de los vectores propios asociados con los dos valores propios? Se deben resolver las ecuaciones asociadas a las matrices

$$A - 2\operatorname{Id} = \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix}$$

y

$$A - \operatorname{Id} = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}.$$

Las soluciones de estos sistemas son los autoespacios de autovalor 2 y 1 respectivamente. En el primer caso,

$$\begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} \xrightarrow[F_3+3F_2]{F_1+3F_2} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Luego, la solución del sistema asociado a $A - 2\,\mathrm{Id}$ es

$$V_2 = \{(2y + 2z, y, z) : y, z \in \mathbb{R}\} =< (2,1,0), (2,0,1) >$$

cuya dimensión es 2.

Por otro lado,

$$\begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix} \underset{F_3+3F_2}{\overset{F_1+4F_2}{\longrightarrow}} \begin{bmatrix} 0 & 6 & 2 \\ -1 & 3 & 2 \\ 0 & 3 & 1 \end{bmatrix} \underset{F_2-F_3}{\overset{F_1-2F_3}{\longrightarrow}} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 3 & 1 \end{bmatrix}.$$

Luego, la solución del sistema asociado a $A - \mathrm{Id}$ es

$$V_1 = \{(z, -\frac{1}{3}z, z) : z \in \mathbb{R}\} =< (1, -\frac{1}{3}, 1) > .$$

Entonces, una base de autovectores de T podría ser

$$\mathcal{B} = \{(2,1,0), (2,0,1), (1, -\frac{1}{3}, 1)\}$$

y en esa base la matriz de la transformación lineal es

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

4.8 OPERADORES SIMÉTRICOS EN $\mathbb{R}^n$

Definición 4.8.1. Sea T un operador lineal en $\mathbb{R}^n$, diremos que T es un *operador simétrico* si la matriz de T en la base canónica es simétrica, es decir si $[T]_{\mathcal{C}}^t = [T]_{\mathcal{C}}$

Observar, como ya hemos visto anteriormente, que en $\mathbb{R}^n$ el producto escalar es

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_i x_i y_i = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Es decir, si usamos la convención que un vector en $\mathbb{R}^n$ se escribe como una matriz columna (de n filas y una columna), tenemos que dados $x, y \in \mathbb{R}^n$,

$$\langle x, y \rangle = x^t y.$$

Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ un operador simétrico y A la matriz asociada a T, es decir $A = [T]_{\mathcal{C}}$, donde $\mathcal{C}$ es la base canónica. Si trabajamos en las coordenadas canónicas es claro que $T(x) = Ax$ y debido a esto a menudo intercambiaremos T por A y viceversa.

Definición 4.8.2. Sea V espacio vectorial y $T : V \to V$ operador lineal. Sea $U \subset V$ subespacio de V. Diremos que U es *invariante por* T si $T(U) := \{T(v) : v \in U\} \subset U$.

Proposición 4.8.3. *Si* T *es un operador simétrico, entonces*

a*)*
$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

b*) Si* W *un subespacio de* $\mathbb{R}^n$ *invariante por* T*, entonces* $W^\perp$ *es invariante por* T*.*

Demostración. Sea $A = [T]_{\mathcal{C}}$.
(a) Notar que

$$\langle Ax, y\rangle = (Ax)^t y = x^t A^t y = x^t A y = \langle x, Ay\rangle$$

(b) Sea $u \in W^\perp$, debemos ver que $Au \in W^\perp$, es decir que $\langle Au, w\rangle = 0$ para todo $w \in W$. Ahora bien,

$$\langle Au, w\rangle = (Au)^t w = u^t A^t w = u^t A w = u^t (Aw) = \langle u, Aw\rangle = 0.$$

La última igualdad es válida pues $u \in W^\perp$ y como $w \in W$, por hipótesis $Aw \in W$. □

Veremos ahora que un operador simétrico, o equivalentemente, una matriz simétrica, es diagonalizable, es decir que hay una base de autovectores del operador o, equivalentemente, que existe una matriz P invertible tal que $P^{-1}AP$ es diagonal.

Usaremos el siguiente resultado sin demostración.

Teorema 4.8.4 (Teorema fundamental del álgebra)**.** *Todo polinomio no constante con coeficientes complejos tiene al menos una raíz compleja. Es decir si*

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0, \textit{ con } a_i \in \mathbb{C},\ a_n \neq 0 \textit{ y } n \geqslant 1,$$

entonces existe $\alpha \in \mathbb{C}$ *tal que* $p(\alpha) = 0$*.*

Pese a llamarse "Teorema fundamental del álgebra", este resultado no suele demostrarse en los cursos de álgebra, pues su demostración requiere del uso de análisis matemático.

Si α es raíz de p, un polinomio de grado n, por el teorema del resto, $p(x) = (x-\alpha)p_1(x)$, con p_1 un polinomio de grado $n-1$. Aplicando inductivamente este procedimiento, podemos deducir:

Corolario 4.8.5. *Si* p *es un polinomio de de grado* $n \geqslant 1$ *con coeficientes en* $\mathbb{C}$*, entonces*

$$p(x) = c(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

con $c, \alpha_i \in \mathbb{C}$*.*

Observación 4.8.6. Recordemos que si $a + bi \in \mathbb{C}$, a es la *parte real* y b es la *parte imaginaria*. El conjugado $a + b_i$ es $\overline{a+bi} = a - bi$. La conjugación cumple que $\overline{\overline{z}} = z$, $\overline{z+w} = \overline{z} + \overline{w}$ y $\overline{zw} = \overline{z}\,\overline{w}$ ($z, w \in \mathbb{C}$). Recordemos también que $z\overline{z} = |z|^2$.

Si $x \in \mathbb{C}^n$, entonces cada coordenada de x es un número complejo, es decir $x_i = a_i + ib_i$, con $a_i, b_i \in \mathbb{R}$. Luego si $v = (a_1, \ldots, a_n)$ y $w = (b_1, \ldots, b_n)$, tenemos que $x = v + wi$ con $v, w \in \mathbb{R}^n$. En este caso, diremos que v es la parte real de x y w la parte imaginaria. También podemos extender la conjugación a $\mathbb{C}^n$ y $\mathbb{C}^{n\times m}$ coordenada a coordenada y entonces no es difícil verificar que si $A, B \in \mathbb{C}^{n\times m}$

$$\overline{\overline{A}} = A, \qquad \overline{A+B} = \overline{A} + \overline{B},$$

y que si $A \in \mathbb{C}^{n\times m}$, $B \in \mathbb{C}^{m\times k}$, $\alpha \in \mathbb{C}$, entonces $\overline{\alpha A} = \overline{\alpha}\,\overline{A}$ y además

$$\overline{\alpha AB} = \overline{\alpha}\,\overline{A}\,\overline{B} = \overline{A}\,\overline{\alpha B}.$$

Notar también que si $z = (z_1, \ldots, z_n)$,

$$z^t\,\overline{z} = [z_1 \ \ldots \ z_n] \begin{bmatrix} \overline{z_1} \\ \vdots \\ \overline{z_n} \end{bmatrix} = |z_1|^2 + \cdots + |z_n|^2,$$

que es > 0 si el vector no es nulo. Denotaremos la expresión de arriba como $||z||^2$.

Teorema 4.8.7. *Sea* T *un operador simétrico de* $\mathbb{R}^n$*. Entonces existe* $\lambda \in \mathbb{R}$ *autovalor real de* T.

Demostración. Sea $A = [T]_{\mathcal{C}}$. Extendamos T a una transformación lineal de $T : \mathbb{C}^n \to \mathbb{C}^n$ de manera natural, con el producto de matrices $T(x) = Ax$ con $x \in \mathbb{C}^n$. Sea χ_A el polinomio característico de A. Por el teorema fundamental del álgebra, existe $\lambda \in \mathbb{C}$ tal que $\chi_A(\lambda) = 0$. Luego existe $x \in \mathbb{C}^n$, no nulo, tal que $Ax = \lambda x$. Veremos que λ es un número real. Por un lado, como A tiene coeficientes reales, tenemos que $A = \overline{A}$ y entonces:

$$x^t A\overline{x} = x^t\overline{A}\,\overline{x} = x^t\overline{Ax} = x^t\overline{\lambda x} = \overline{\lambda}x^t\overline{x} = \overline{\lambda}||x||^2.$$

Por otro lado, como A es simétrica,

$$x^t A\overline{x} = x^t A^t\overline{x} = (Ax)^t\overline{x} = (\lambda x)^t\overline{x} = \lambda x^t\overline{x} = \lambda||x||^2.$$

Por lo tanto, $\lambda = \overline{\lambda}$, lo cual nos dice que $\lambda \in \mathbb{R}$. Es decir, existe un vector $x \in \mathbb{C}^n$ no nulo y $\lambda \in \mathbb{R}$, tal que $Ax = \lambda x$. Si $x = v + iw$ con $v, w \in \mathbb{R}^n$, entonces

$$\lambda v + i\lambda w = \lambda x = Ax = Av + iAw.$$

Como A es una matriz real $Av, Aw \in \mathbb{R}^n$ y como $\lambda \in \mathbb{R}$, tenemos que $Av = \lambda v$ y $Aw = \lambda w$ Como $x = v + iw$ es no nulo, entonces o v o w son no nulos y por lo tanto hay al menos un autovector en $\mathbb{R}^n$ con autovalor $\lambda \in \mathbb{R}$. □

El siguiente resultado, el *teorema espectral*, requiere para su demostración una generalización del resultado anterior para espacios de producto interno de dimensión finita y matrices (transformaciones lineales) simétricas respecto a este producto interno. Todos estos conceptos y resultados son generalizaciones sencillas, pero llevan algún tiempo desarrollarlas y el lector interesado las puede ver en la sección 5.3. La demostración que daremos del teorema espectral se hará por inducción en la dimensión del espacio y utiliza en el paso inductivo la generalización del teorema 4.8.7.

Teorema 4.8.8 (Teorema espectral)**.** *Sea A matriz simétrica $n \times n$. Entonces existe $\mathcal{U} = \{u_1, \ldots, u_n\}$ una BON de $\mathbb{R}^n$ de autovectores de A.*

Idea de la demostración (*). Se hará por inducción en $n = \dim(V)$.

Si $n = 1$ es trivial.

Supongamos que vale para $n-1$ con $n > 1$, y probaremos el resultado para n. Por el teorema 4.8.7 existe $\lambda \in \mathbb{R}$ y $x \in \mathbb{R}^n$ tal que $Ax = \lambda x$. Si $u_n = x/||x||$, u_n tiene norma 1 y cumple también que $Au_n = \lambda u_n$. Sea $W = \langle u_n \rangle$. Entonces W es invariante y por proposición 4.8.3 tenemos que $W^\perp$ es invariante por A. De esta forma, podemos considerar entonces a A como una transformación lineal de $W^\perp$ a $W^\perp$.

No es difícil verificar que el producto interno canónico es un producto interno en $W^\perp$ y que $A : W^\perp \to W^\perp$ es una matriz simétrica con este producto interno restringido.

Como $\dim(W^\perp) = n-1$, por hipótesis inductiva existe $\{u_1, \ldots, u_{n-1}\}$ una BON de $W^\perp$ de autovectores de $A : W^\perp \to W^\perp$. Es claro entonces que $\mathcal{U} = \{u_1, \ldots, u_n\}$ una BON de $\mathbb{R}^n$ de autovectores de A. $\square$

Corolario 4.8.9. *Sea A matriz simétrica $n \times n$, entonces A es diagonalizable.*

Ejemplo. Encontremos autovalores y autovectores de la matriz

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

Como es una matriz simétrica sabemos que es diagonalizable, es decir tiene una base de autovectores. El polinomio característicos es

$$\chi_A(x) = \det \begin{bmatrix} 2-x & -1 & 0 \\ -1 & 2-x & -1 \\ 0 & -1 & 2-x \end{bmatrix} = -x^3 + 6x^2 - 10x + 4.$$

Ahora bien, las raíces de $-x^3 + 6x^2 - 10x + 4$ son

$$\begin{aligned} \lambda_1 &= 2 + \sqrt{2} \\ \lambda_2 &= 2 \\ \lambda_3 &= 2 - \sqrt{2} \end{aligned}$$

Para averiguar los autovectores debemos plantear las ecuaciones $Ax = \lambda_i x$, que resultan en los siguiente sistema de ecuaciones

$$\begin{aligned} 2x_1 - x_2 &= \lambda_i x_1 \\ -x_1 + 2x_2 - x_3 &= \lambda_i x_2 \\ -x_2 + 2x_3 &= \lambda_i x_3, \end{aligned}$$

$(i = 1, 2, 3)$, o equivalentemente,

$$\begin{aligned} (2-\lambda_i)x_1 + x_2 &= 0 \\ -x_1 + (2-\lambda_i)x_2 - x_3 &= 0 \\ -x_2 + (2-\lambda_i)x_3 &= 0, \end{aligned}$$

$(i = 1, 2, 3)$. En el caso de $\lambda_1 = 2 + \sqrt{2}$, resulta

$$\begin{aligned} -\sqrt{2}x_1 + x_2 &= 0 \\ -x_1 - \sqrt{2}x_2 - x_3 &= 0 \\ -x_2 - \sqrt{2}x_3 &= 0, \end{aligned}$$

cuya solución es $\lambda(1, -\sqrt{2}, 1)$. Si continuamos resolviendo los sistemas de ecuaciones, podemos encontrar la siguiente base de autovectores:

$$\begin{aligned} v_1 &= (1, -\sqrt{2}, 1) \\ v_2 &= (-1, 0, 1) \\ v_3 &= (1, \sqrt{2}, 1). \end{aligned}$$

5 PRODUCTO INTERNO

Las propiedades algebraicas de $\mathbb{R}^n$ no son suficientes para hacer frente a ciertas nociones geométricas como ángulos, perpendicularidad y longitud. Hemos visto en el capítulo 1 que con la introducción del producto escalar pudimos definir y trabajar los conceptos previamente mencionados. En este capítulo daremos algunas propiedades adicionales del producto escalar, veremos que las matrices simétricas pueden ser interpretadas a partir del producto escalar y, finalmente, probaremos que las matrices simétricas son diagonalizables.

5.1 PRODUCTO INTERNO

Recordemos que en el capítulo 1 hemos visto que el producto escalar entre dos vectores $x, y \in \mathbb{R}^n$ se define como

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i.$$

También recordemos que si $x \in \mathbb{R}^n$, entonces la norma de x es $||x|| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$.

Como hemos visto en el capítulo 1 el producto escalar cumple cuatro propiedades básicas, que hemos llamado P1 (simetría), P2 y P3 (bilinealidad o linealidad en cada variable), y P4 (positividad). Estas son las únicas propiedades que usaremos, y no la definición explícita de producto escalar, para deducir los resultados de esta sección.

Definición 5.1.1. Sea V un espacio vectorial y una función

$$\langle\, ,\, \rangle : V \times V \to \mathbb{R}.$$

Diremos que $\langle\, ,\, \rangle$ es un *producto interno* si para todo $v, w, u \in \mathbb{R}^n$, se satisface:

P1.

$$\langle v, w \rangle = \langle w, v \rangle.$$

P2.

$$\langle v, w + u \rangle = \langle v, w \rangle + \langle v, u \rangle = \langle w + u, v \rangle.$$

P3. Si $\lambda \in \mathbb{R}$, entonces

$$\langle \lambda v, w \rangle = \lambda \langle v, w \rangle \quad \text{y} \quad \langle v, \lambda w \rangle = \lambda \langle v, w \rangle.$$

P4. Si $v = 0$ es el vector cero, entonces $\langle v, v\rangle = 0$, de lo contrario

$$\langle v, v\rangle > 0$$

Es decir $\langle\ ,\ \rangle$ es una forma bilineal (P2 y P3), simétrica (P1) y positiva (P4)

Obviamente el producto escalar en $\mathbb{R}^n$ es un producto interno, que llamaremos el *producto interno canónico* de $\mathbb{R}^n$. Los resultados de esta sección valen en general para un producto interno en un espacio vectorial de dimensión finita, pero tendremos siempre en mente el producto escalar en $\mathbb{R}^n$.

Ejemplo. El producto escalar es uno entre muchos de los productos internos que podemos tener en $\mathbb{R}^n$, por ejemplo, en $\mathbb{R}^3$, la función definida:

$$\langle(x_1, x_2, x_3), (y_1, y_2, y_3)\rangle = 2x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2 - x_2y_3 - x_3y_2 + 2x_3y_3$$

Es un producto interno (ejercicio).

Ejemplo. También se puede definir un producto interno en un espacio de dimensión infinita, como veremos a continuación.

Sea $E = C^0([a, b])$ el espacio vectorial cuyos elementos son las funciones continuas $f : [a, b] \to \mathbb{R}$. Se puede definir un producto interno en E de la siguiente manera: sean $f, g \in C^0([a, b])$, entonces

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Usando las propiedades de la integral es sencillo ver que $\langle,\rangle$ es una 2-forma, bilineal y simétrica. Por propiedades de las funciones continuas se demuestra que además la 2-forma es positiva.

Este producto interno se utiliza en el estudio de series de Fourier.

Proposición 5.1.2. *Sean* $x, y \in \mathbb{R}^n$. *Entonces,*

(1) *Si* $c \in \mathbb{R}$, *tenemos* $||cx|| = |c|||x||$.

(2) $||x + y||^2 = ||x||^2 + ||y||^2 + 2\langle x, y\rangle$.

Demostración.

Demostración de (1). Es exactamente, proposición 1.3.1 (que se demuestra usando P3).

Demostración de (2).

$$\begin{aligned}
||u + v||^2 &= \langle u + v, u + v\rangle \\
&= \langle u, u + v\rangle + \langle v, u + v\rangle \\
&\overset{(P2)}{=} \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&\overset{(P1)}{=} \langle u, u\rangle + 2\langle u, v\rangle + \langle v, v\rangle \\
&= ||u||^2 + ||v||^2 + 2\langle u, v\rangle.
\end{aligned}$$

□

Recordemos que dos vectores x, y de $\mathbb{R}^n$ son perpendiculares u ortogonales si $\langle x, y\rangle = 0$, lo cual era denotado $x \perp y$.

Definición 5.1.3. Sea $X \subset \mathbb{R}^n$, diremos que X es un *conjunto ortogonal* si $v \perp w$ para $v, w \in X$, $v \neq w$. Diremos que X es un *conjunto ortonormal* si X es ortogonal y todos los vectores de X son *unitarios* (es decir $||v|| = 1$ para $v \in X$).

Proposición 5.1.4. *Sea*

$$X = \{v_1, \ldots, v_r\} \subset \mathbb{R}^n$$

un conjunto ortogonal. Sea

$$X' = \left\{ \frac{v_1}{||v_1||}, \ldots, \frac{v_r}{||v_r||} \right\}.$$

Entonces X' es un conjunto ortonormal.

Demostración. Para demostrar esto debemos ver que dos vectores distintos de X' son ortogonales y que cada vector de X' es de norma 1.

Sea $i \neq j$, entonces

$$\langle \frac{v_i}{||v_i||} | \frac{v_j}{||v_j||} \rangle = \frac{1}{||v_i|| ||v_j||} \langle v_i | v_j \rangle = 0.$$

Por otro lado,

$$\langle \frac{v_i}{||v_i||} | \frac{v_i}{||v_i||} \rangle = \frac{1}{||v_i||^2} \langle v_i | v_i \rangle = \frac{1}{||v_i||^2} ||v_i||^2 = 1.$$

□

Teorema 5.1.5. *Sea $X \subset \mathbb{R}^n$ un conjunto ortogonal. Entonces X es LI.*

Demostración. Sea $X = \{v_1, \ldots, v_r\}$ y sea $a_1, \ldots, a_r$ en F tales que $\sum_{i=1}^r a_i v_i = 0$. Entonces, dado j con $1 \leqslant j \leqslant r$, tenemos

$$0 = \langle \sum_{i=1}^r a_i v_i, v_j \rangle = \sum_{i=1}^r a_i \langle v_i, v_j \rangle = a_j \langle v_j, v_j \rangle = a_j ||v_j||^2.$$

Como X es un conjunto ortogonal, $||v_j|| > 0$, luego $a_j = 0$ para cualquier j. Es decir hemos probado que todos los coeficientes de la suma son cero y por lo tanto X es LI. □

Proposición 5.1.6.

(1) *Teorema de Pitágoras: si $u \perp v$, entonces $||u + v||^2 = ||u||^2 + ||v||^2$.*

(2) *Ley del Paralelogramo: $||u + v||^2 + ||u - v||^2 = 2||u||^2 + 2||v||^2$.*

Demostración. Ambas demostraciones se hacen desarrollando las fórmulas y usando las propiedades del producto escalar.

Demostración de 1.

$$||u+v||^2 = \langle u+v, u+v\rangle = \langle u, u+v\rangle + \langle v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle.$$

Las igualdades de arriba se deben a la bilinealidad del producto interno. Ahora bien, como $u \perp v$, tenemos que $0 = \langle u,v\rangle = \langle v,u\rangle$, luego

$$||u+v||^2 = \langle u,u\rangle + \langle v,v\rangle = ||u||^2 + ||v||^2.$$

Demostración de 2.

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

□

Definición 5.1.7. Si $X \subset \mathbb{R}^n$ es ortogonal (ortonormal) y es base, diremos que X es una *base ortogonal* (resp. *base ortonormal*) o diremos que X es *BO* (resp. *BON*).

Ejemplo.

(1) La base canónica de $\mathbb{R}^n$ es ortonormal.

(2) Si $u = (1,1)$, $v = (1,-1)$, entonces u, v es una base ortogonal.

Proposición 5.1.8. *Sea* $X = \{v_1, \ldots, v_n\}$ *una base ortogonal, entonces*

$$X' = \left\{\frac{v_1}{||v_1||}, \ldots, \frac{v_n}{||v_n||}\right\}$$

es una base ortonormal.

Demostración. Hemos probado en la proposición 5.1.4 que X' es un conjunto ortonormal. Por teorema 5.1.5. X' es un conjunto LI. Veamos ahora que es X' genera a V.

Sea $v \in V$, como X es base de V, en particular genera a V, luego existen $a_i \in R$, tal que $v = \sum_i a_i v_i$. Luego

$$v = \sum_i a_i v_i = \sum_i a_i \frac{||v_i||}{||v_i||} v_i = \sum_i (a_i ||v_i||) \frac{v_i}{||v_i||}.$$

Luego X' es un conjunto de generadores de V. □

Ejemplo. (1) Si $u = (1,1)$, $v = (1,-1)$, entonces $||u|| = ||v|| = \sqrt{2}$ y $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$ es una base ortonormal.

Observación 5.1.9. No es difícil ver en un dibujo que

$$\operatorname{pr}_u(v) := \frac{\langle u|v\rangle}{\langle u|u\rangle}u$$

es la proyección de v en u y que $(v - \operatorname{pr}(v)) \perp u$. Es decir, los vectores

$$u, \quad v - \frac{\langle u|v\rangle}{\langle u|u\rangle}u,$$

son ortogonales.

falta

Figura 22: Proyección de v en u cuando $||v|| = 1$.

Esto, además de la interpretación geométrica, lo podemos demostrar algebraicamente:

$$\langle v - \frac{\langle u,v\rangle}{\langle u,u\rangle}u, u\rangle = \langle v,u\rangle - \frac{\langle u,v\rangle}{\langle u,u\rangle}\langle u,u\rangle = \langle v,u\rangle - \langle u,v\rangle = 0.$$

Proposición 5.1.10 (Desigualdad de Schwarz). *Sean* $u, v \in \mathbb{R}^n$. *Entonces*

$$|\langle u,v\rangle| \leqslant ||u||||v||.$$

Demostración. Primero daremos una demostración basada en los resultados geométricos vistos en el capítulo 1, luego, el lector interesado, podrá leer una demostración algebraica del resultado.

Demostración 1. Esta demostración es válida para el producto escalar en $\mathbb{R}^n$. La fórmula (1.3.1) nos dice que

$$\langle u,v\rangle = ||u||||v||\cos(\theta)$$

donde θ es el ángulo comprendido entre u y v. Como $|\cos(\theta)| \leqslant 1$, tenemos

$$|\langle u,v\rangle| = ||u||\,||v||\,|\cos(\theta)| \leqslant ||u||\,||v||.$$

Demostración 2 (∗). Sea $c = \dfrac{\langle u,v\rangle}{\langle u,u\rangle} = \dfrac{\langle u,v\rangle}{||u||^2}$, entonces, por la observación 5.1.9, tenemos que $v - cu$ es ortogonal a u. Ahora bien,

$$v = (v - cu) + cu$$

y $(v - cu) \perp cu$. Por Pitágoras

$$\begin{aligned} ||v||^2 &= ||v - cu||^2 + ||cu||^2 \\ &= ||v - cu||^2 + |c|^2||u||^2. \end{aligned}$$

Como $||v - cu||^2 \geqslant 0$, tenemos que $|c|^2||u||^2 \leqslant ||v||^2$ y sacando raíces cuadradas obtenemos

$$|c|||u|| \leqslant ||v|| \Rightarrow \frac{|\langle u,v\rangle|}{||u||^2}||u|| \leqslant ||v|| \Rightarrow \frac{|\langle u,v\rangle|}{||u||} \leqslant ||v|| \Rightarrow |\langle u,v\rangle| \leqslant ||v||||u||.$$

□

Teorema 5.1.11 (Desigualdad triangular)**.** *Sean* $u, v \in \mathbb{R}^n$*, entonces*

$$||u+v|| \leqslant ||u|| + ||v||$$

Demostración. Probar

$$||u+v||^2 \leqslant (||u|| + ||v||)^2$$

desarrollando el lado izquierdo de la desigualdad como $\langle u+v, u+v\rangle$ y el lado derecho por el cálculo del binomio al cuadrado. Luego usar la desigualdad de Schwarz. □

Ahora vamos a demostrar el proceso de ortonormalización de Gram-Schmidt, que consta de un algoritmo que permite pasar de una base cualquiera $\{v_1, \ldots, v_n\}$ de $\mathbb{R}^n$ a una base ortonormal $\{u_1, \ldots, u_n\}$, con la importante propiedad de que, para m con $1 \leqslant m \leqslant n$, el subespacio generado por los vectores $\{u_1, \ldots, u_m\}$ es el mismo que el subespacio generado por los vectores $\{v_1, \ldots, v_m\}$.

La idea del proceso es sencillo para dos vectores: sean $v_1, v_2 \in \mathbb{R}^n$ no nulos y no proporcionales, vimos en la observación 5.1.9 que los vectores

$$w_1 = v_1, \quad w_2 = v_2 - \operatorname{pr}_{v_1}(v_2) = v_2 - \frac{\langle v_1, v_2\rangle}{\langle v_1, v_1\rangle} v_1$$

son ortogonales. Ahora bien, $v_1 = w_1$ y $v_2 = \frac{\langle v_1, v_2\rangle}{\langle v_1, v_1\rangle} w_1 + w_2$, luego w_1, w_2 generan el mismo subespacio que v_1, v_2. Concluyendo, dados v_1, v_2 dos vectores LI, w_1, w_2 son dos vectores ortogonales que generan el mismo subespacio. Para $n > 2$ la idea es similar.

Proposición 5.1.12 (Proceso de ortogonalización de Gram-Schmidt)**.** *Sea* $\{v_1, \ldots, v_n\}$ *una base de* $\mathbb{R}^n$*. Entonces existe una base ortogonal* $\{w_1, \ldots, w_n\}$ *tal que el subespacio generado por los vectores* $\{w_1, \ldots, w_m\}$ *es el mismo que el subespacio generado por* $\{v_1, \ldots, v_m\}$ *(*$1 \leqslant m \leqslant n$*). Explícitamente, la base es*

$$
\begin{aligned}
w_1 &= v_1, && (1)\\
w_2 &= v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1, && (2)\\
w_3 &= v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2, && (3)\\
\vdots\ & \ \vdots \\
w_n &= v_n - \frac{\langle v_n, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_n, w_2\rangle}{\langle w_2, w_2\rangle} w_2 - \cdots - \frac{\langle v_n, w_{n-1}\rangle}{\langle w_{n-1}, w_{n-1}\rangle} w_{n-1}. && (n)
\end{aligned}
$$

En forma más breve, para $1 \leqslant i \leqslant n$*,*

$$w_i = v_i - \sum_{j=1}^{i-1} \frac{\langle v_i, w_j\rangle}{\langle w_j, w_j\rangle} w_j \qquad (i)$$

Demostración (∗). Haremos la demostración por inducción sobre n.

Para $n = 1$ el resultado es trivial.

Supongamos que el resultado valga para $k-1 > 0$, es decir $\{w_1, \ldots, w_{k-1}\}$ es ortogonal y $\operatorname{span}(w_1, \ldots, w_{k-1}) = \operatorname{span}(v_1, \ldots, v_{k-1})$. Probemos el resultado para k. Si $i < k$,

$$\begin{aligned}\langle w_k, w_i\rangle &= \langle v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j\rangle}{\langle w_j, w_j\rangle} w_j, w_i\rangle = \langle v_k, w_i\rangle - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j\rangle}{\langle w_j, w_j\rangle}\langle w_j, w_i\rangle \\ &= \langle v_k, w_i\rangle - \langle v_k, w_i\rangle = 0.\end{aligned}$$

Es decir $\langle w_k, w_i\rangle = 0$ para todo $i < k$. Por consiguiente, $\{w_1, \ldots, w_k\}$ es ortogonal.

Demostremos ahora que $\operatorname{span}\{w_1, \ldots, w_m\} = \operatorname{span}\{v_1, \ldots, v_m\}$ para $1 \leqslant m \leqslant n$.

$\operatorname{span}\{w_1, \ldots, w_m\} \subset \operatorname{span}\{v_1, \ldots, v_m\}$: por la fórmula (i) es claro que w_m es combinación lineal de v_m y $w_1, \ldots, w_{m-1}$. Por hipótesis inductiva, los $w_1, \ldots, w_{m-1}$ son combinación lineal de los $v_1, \ldots, v_{m-1}$, luego los $w_1, \ldots, w_m$ son combinación lineal de los $v_1, \ldots, v_m$.

$\operatorname{span}\{v_1, \ldots, v_m\} \subset \operatorname{span}\{w_1, \ldots, w_m\}$: Como

$$v_k = w_k + \sum_{j=1}^{k-1} \frac{\langle v_k, w_j\rangle}{\langle w_j, w_j\rangle} w_j,$$

tenemos que $\operatorname{span}\{v_1, \ldots, v_m\} \subset \operatorname{span}\{w_1, \ldots, w_m\}$. □

Observación. Sea W subespacio de $\mathbb{R}^n$, entonces existe una base ortogonal de W. Esto se deduce del proceso de ortogonalización de Gram-Schmidt: sea $v_1, \ldots, v_k$ una base de W y completamos a $v_1, \ldots, v_n$ una base de $\mathbb{R}^n$. Por Gram-Schmidt obtenemos una BO $w_1, \ldots, w_n$ tal que el subespacio generado por $w_1, \ldots, w_i$ es igual al subespacio generado por $v_1, \ldots, v_i$ para $1 \leqslant i \leqslant n$. En particular $W = \langle v_1, \ldots, v_k\rangle = \langle w_1, \ldots, w_k\rangle$ y por lo tanto $w_1, \ldots, w_k$ es una BON de W.

En la práctica, dada una base $v_1, \ldots, v_k$ de W, con los primeros k pasos del proceso de ortogonalización de Gram-Schmidt obtenemos $w_1, \ldots, w_k$ una base ortogonal de W.

Ejemplo. Encontrar una base ortogonal del subespacio de $\mathbb{R}^3$ generado por los vectores $(1, 2, -1)$ y $(-2, -1, 0)$

Solución. Por Gram-Schmidt:

$$\begin{aligned}w_1 &= (1, 2, -1), \\ w_2 &= (-2, -1, 0) - \frac{\langle (-2, -1, 0), w_1\rangle}{\langle w_1, w_1\rangle} w_1,\end{aligned}$$

es una base ortogonal de W. Calculemos

$$\begin{aligned}
w_2 &= (-2,-1,0) - \frac{\langle(-2,-1,0),(1,2,-1)\rangle}{\langle(1,2,-1),(1,2,-1)\rangle}(1,2,-1) \\
&= (-2,-1,0) - \frac{-4}{6}(1,2,-1) \\
&= (-2,-1,0) - (\frac{-2}{3},\frac{-4}{3},\frac{2}{3}) \\
&= (\frac{-4}{3},\frac{1}{3},\frac{-2}{3}).
\end{aligned}$$

Para simplificar, multiplicamos a w_2 por 3 y obtenemos que

$$(1,2,-1),(-4,1,-2)$$

es una BO de W.

□

Definición 5.1.13. Sean U, W subespacios de $\mathbb{R}^n$. Diremos que U *es ortogonal a* W y denotaremos $U \perp W$ si para todo $u \in U$ y para todo $w \in W$ tenemos que $\langle u, w\rangle = 0$.

Si X es subconjunto de $\mathbb{R}^n$, definimos

$$X^\perp := \{u \in \mathbb{R}^n : \langle u, x\rangle = 0, \forall\, x \in X\} = \{u \in \mathbb{R}^n : \langle u, X\rangle = 0\}.$$

Proposición 5.1.14. *Sea* $X \subset \mathbb{R}^n$*, entonces* $X^\perp$ *es un subespacio de* $\mathbb{R}^n$*.*

Demostración. Debemos probar que si $u, v \in X^\perp$ y $c \in \mathbb{R}$, entonces $cu + v \in X^\perp$, es decir que para todo $x \in X$, se cumple que $\langle cu + v, x\rangle = 0$. Ahora bien,

$$\langle cu + v, x\rangle = c\langle u, x\rangle + \langle v, x\rangle = 0.$$

□

Definición 5.1.15. Sea $\mathbb{R}^n$ espacio vectorial con producto interno $\langle\ ,\ \rangle$ y sea X subconjunto de $\mathbb{R}^n$. Diremos que $X^\perp$ es el *subespacio ortogonal a* X *en* $\mathbb{R}^n$.

5.2 LA ADJUNTA DE UNA TRANSFORMACIÓN LINEAL (*)

Mostraremos en esta sección como el producto interno nos permite asociar a cada transformación lineal $T : V \to W$ una nueva transformación lineal $T^* : W \to V$ llamada la adjunta de T.

Teorema 5.2.1. *Sean* V, W *espacios vectoriales de dimensión finita y con producto interno* $\langle\ ,\ \rangle$ *y* $\langle\ ,\ \rangle$ *respectivamente (se denotan igual). Sea* $T : V \to W$ *lineal, entonces existe una única* $T^* : W \to V$ *que cumple*

$$\langle Tv, w\rangle = \langle v, T^*w\rangle, \tag{5.2.1}$$

para $v \in V$, $w \in W$ *(el producto de la izquierda es en* W *y el de la derecha en* V*).*

Demostración. Sea $\{v_1, \ldots, v_n\}$ una BON de V y $\{w_1, \ldots, w_m\}$ una BON de W, observemos que la coordenada j (en V) de T^*w_i debe cumplir

$$\langle T^* w_i, v_j \rangle = \langle w_i, Tv_j \rangle. \tag{5.2.2}$$

Por lo tanto, definimos

$$T^*(w_i) = \sum_{j=1}^{n} \langle w_i, Tv_j \rangle v_j,$$

y extendemos linealmente a una transformación lineal $T^* : W \to V$. Claramente T^* está bien definida y es lineal (por definición). La unicidad está garantizada por la ecuación (5.2.2).

Finalmente, debemos comprobar que se verifica la ecuación (5.2.3): sean $w \in W$ y $v \in V$, entonces $w = \sum_{i=1}^{m} \langle w_i, w \rangle w_i$ y $v = \sum_{j=1}^{n} \langle v_j, v \rangle v_j$. Reemplazando en la ecuación (5.2.3) w y v por su desarrollo en las bases se obtiene la igualdad. Para el lector curioso, a continuación desarrollamos la demostración:

$$\begin{aligned}
\langle v, T^* w \rangle &= \langle v, T^*(\sum_{i=1}^{m} \langle w_i, w \rangle w_i) \rangle \\
&= \langle v, \sum_{i=1}^{m} \langle w_i, w \rangle T^*(w_i) \rangle \\
&= \sum_{i=1}^{m} \langle w_i, w \rangle \langle v, T^*(w_i) \rangle \\
&= \sum_{i=1}^{m} \langle w_i, w \rangle \langle v, \sum_{j=1}^{n} \langle w_i, Tv_j \rangle v_j \rangle \\
&= \sum_{i=1}^{m} \langle w_i, w \rangle \sum_{j=1}^{n} \langle w_i, Tv_j \rangle \langle v, v_j \rangle.
\end{aligned}$$

Por otro lado, como $v = \sum_{j=1}^{n} \langle v_j, v \rangle v_j$, entonces

$$\sum_{j=1}^{n} \langle w_i, Tv_j \rangle \langle v, v_j \rangle = \sum_{j=1}^{n} \langle w_i, \langle v, v_j \rangle Tv_j \rangle = \langle w_i, T(\sum_{j=1}^{n} \langle v, v_j \rangle v_j) \rangle = \langle w_i, Tv \rangle,$$

por lo tanto

$$\langle v, T^* w \rangle = \sum_{i=1}^{m} \langle w_i, w \rangle \langle w_i, Tv \rangle = \langle \sum_{i=1}^{m} \langle w_i, w \rangle w_i, Tv \rangle = \langle w, Tv \rangle$$

$\square$

Definición 5.2.2. Sean V, W espacios vectoriales de dimensión finita y con producto interno $\langle\ ,\ \rangle$ y $\langle\ ,\ \rangle$ respectivamente. Sea $T : V \to W$ lineal, entonces a la única $T^* : W \to V$ que cumple

$$\langle Tv, w\rangle = \langle v, T^* w\rangle, \tag{5.2.3}$$

para $v \in V$, $w \in W$ se la denomina la *adjunta de* T.

Observación. El caso más interesante, y que pasaremos a estudiar ahora, es cuando $T : V \to V$, es decir cuando el espacio de llegada y de partida es el mismo, y por lo tanto también $T^* : V \to V$.

Ejemplo. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal $T(x, y, z) = (3x + y, 2x - y + 3z, x)$. Calcular T^* y la matriz de T y T^* en la base canónica.

Solución 1. La observación principal para hacer el cálculo de T^* es que dada cualquier transformación lineal S, tenemos que

$$\langle e_i, S(v)\rangle = t_i \Leftrightarrow S(v) = (t_1, \ldots, t_n).$$

Aplicado a este caso,

$$\begin{aligned}
\langle e_1, T^*(x, y, z)\rangle &= \langle T(e_1), (x, y, z)\rangle = \langle (3, 2, 1), (x, y, z)\rangle = 3x + 2y + z\\
\langle e_2, T^*(x, y, z)\rangle &= \langle T(e_2), (x, y, z)\rangle = \langle (1, -1, 0), (x, y, z)\rangle = x - y\\
\langle e_3, T^*(x, y, z)\rangle &= \langle T(e_3), (x, y, z)\rangle = \langle (0, 3, 0), (x, y, z)\rangle = 3y.
\end{aligned}$$

Por lo tanto

$$T^*(x, y, z) = (3x + 2y + z, x - y, 3y).$$

La matriz de T en la base canónica es

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

y la matriz de T^* en la base canónica es

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

□

Observemos que en el ejemplo anterior la matriz de la adjunta es la transpuesta de la matriz de la transformación original. Veremos ahora, que este es un resultado general.

Teorema 5.2.3. *Sea* V *espacio vectorial de dimensión finita con producto interno* $\langle\ ,\ \rangle$ *y sea* $\mathcal{U} = \{u_1, \ldots, u_n\}$ *una BON de* V*. Sea* $T : V \to V$ *una transformación lineal y* A *la matriz de* T *en la base* $\mathcal{U}$*, es decir* $[T]_{\mathcal{U}} = A$*.*

Entonces, $[T^*]_{\mathcal{U}} = A^t$*, es decir, la matriz de* T^* *en la base* $\mathcal{U}$ *es la transpuesta de* A*.*

Demostración. Observemos que como $T(u_j) = \sum_i a_{ij} u_i$, entonces

$$\langle Tu_j, u_i \rangle = a_{ij}.$$

Luego

$$a_{ij} = \langle Tu_j, u_i \rangle = \langle u_j, T^* u_i \rangle.$$

Es decir que $T^*(u_i) = \sum_j a_{ij}$. Es decir $[T^*]_{\mathcal{U}} = A^t$. □

Ejemplo. Resolveremos nuevamente, en forma más sencilla, el ejemplo anterior.

Solución 2. Como $T(x, y, z) = (3x + y, 2x - y + 3z, x)$, la matriz de T en la base canónica es

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

Por lo tanto, por teorema 5.2.3, la matriz de T^* en la base canónica es

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

Luego $T^*(x, y, z) = (3x + 2y + z, x - y, 3y)$. □

Proposición 5.2.4. *Sean* V *espacio vectorial de dimensión finita con producto interno* $\langle\ ,\ \rangle$ *y* $T, S : V \to V$ *transformaciones lineales. Entonces*

(1) $\mathrm{Id}^* = \mathrm{Id}$.

(2) *Si* $c \in \mathbb{R}$*, entonces* $(cR)^* = cR^*$.

(3) $(R + S)^* = R^* + S^*$.

(4) $(RS)^* = S^* R^*$.

(5) $R^{**} = R$.

Demostración. (1). Es trivial.

(2). Por definición de adjunta $(cR)^*$ es la única transformación lineal tal que

$$\langle cR(v), w \rangle = \langle v, (cR)^*(w) \rangle, \quad \forall v, w \in V.$$

Ahora bien

$$\langle v, cR^*(w)\rangle = c\langle v, R^*(w)\rangle = c\langle R(v), w\rangle = \langle cR(v), w\rangle.$$

Es decir $(cR)^* = cR^*$.

(3). Como en el caso anterior, debemos demostrar que $R^* + S^*$ es la única transformación lineal tal que

$$\langle (R+S)(v), w\rangle = \langle v, (R^* + S^*)(w)\rangle, \quad \forall v, w \in V.$$

Ahora bien,

$$\begin{aligned}\langle v, (R^* + S^*)(w)\rangle &= \langle v, R^*(w) + S^*(w)\rangle = \langle v, R^*(w)\rangle + \langle v, S^*(w)\rangle \\ &= \langle R(v), w\rangle + \langle S(v), w\rangle = \langle R(v) + S(v), w\rangle = \langle (R+S)(v), w\rangle.\end{aligned}$$

(4).

$$\begin{aligned}\langle v, (S^*R^*)(w)\rangle &= \langle v, S^*(R^*(w))\rangle = \langle S(v), R^*(w)\rangle \\ &= \langle R(S(v)), w\rangle = \langle (RS)(v), w\rangle.\end{aligned}$$

Por lo tanto $(RS)^* = S^*R^*$.

(5). Por definición de adjunta de R^*, tenemos que $(R^*)^* = R^{**}$ es la única transformación lineal tal que

$$\langle R^*(v), w\rangle = \langle v, R^{**}(w)\rangle, \ \forall\, v, w \in V.$$

Ahora bien, por la definición de adjunta de R sabemos que

$$\langle R^*(v), w\rangle = \langle v, R(w)\rangle, \ \forall\, v, w \in V.$$

Luego $R = R^{**}$.

□

Teorema 5.2.5. *Sea* $T : V \to W$ *una transformación lineal entre espacios vectoriales de dimensión finita con producto interno. Entonces,*

(1) $\operatorname{Nu}(T^*) = \operatorname{Im}(T)^\perp$,

(2) $\operatorname{Im}(T^*) = \operatorname{Nu}(T)^\perp$,

(3) $\operatorname{Nu}(T) = \operatorname{Im}(T^*)^\perp$,

(4) $\operatorname{Im}(T) = \operatorname{Nu}(T^*)^\perp$.

Demostración. La primera afirmación es la que requiera más trabajo, pues las otras se deducen fácilmente de la primera y del hecho que $T^{**} = T$ y $U^{\perp\perp} = U$.

(1)

$$\begin{aligned} w \in \mathrm{Nu}(T^*) &\Leftrightarrow T^*(w) = 0 \\ &\Leftrightarrow \langle v, T^*(w) \rangle = 0, \ \forall v \in V \\ &\Leftrightarrow \langle T(v), w \rangle = 0, \ \forall v \in V \\ &\Leftrightarrow w \in \mathrm{Im}(T)^{\perp}. \end{aligned}$$

(2)

$$\mathrm{Im}(T^*) = (\mathrm{Im}(T^*)^{\perp})^{\perp} \stackrel{((1))}{=} \mathrm{Nu}(T^{**})^{\perp} = \mathrm{Nu}(T)^{\perp}.$$

(3)

$$\mathrm{Nu}(T) = \mathrm{Nu}(T^{**}) \stackrel{((1))}{=} \mathrm{Im}(T^*)^{\perp}.$$

(4)

$$\mathrm{Im}(T) = \mathrm{Im}(T)^{\perp\perp} \stackrel{((1))}{=} \mathrm{Nu}(T^*)^{\perp}.$$

□

5.3 OPERADORES AUTOADJUNTOS (*)

En esta sección todos los espacios vectoriales serán sobre $\mathbb{R}$ y de dimensión finita.

Generalizaremos ahora el concepto de matriz simétrica.

Definición 5.3.1. Sea V un espacio vectorial con producto interno y $T : V \to V$ una transformación lineal. Diremos que T es una *transformación lineal autoadjunta* si $T^* = T$. En ese caso, también suele decirse que T es un *operador lineal autoadjunto*.

Claramente, en $\mathbb{R}^n$ con el producto interno canónico, la multiplicación a izquierda de un vector columna por una matriz simétrica es un operador autoadjunto.

Del teorema 5.2.3 (y un poco más) se deduce el siguiente resultado.

Proposición 5.3.2. *Sea V un espacio vectorial con producto interno y $T : V \to V$ una transformación lineal. Entonces T es un operador lineal autoadjunto si y sólo si para cualquier $\mathcal{U}$ BON de V, la matriz de T en la base $\mathcal{U}$ es simétrica.*

Demostración. ($\Rightarrow$) Por teorema 5.2.3, si A es la matriz de T, entonces la matriz de T^* es A^t. Como $T = T^*$, entonces $A = A^t$.

($\Leftarrow$) Por hipótesis, $[T]_{\mathcal{U}} = [T]_{\mathcal{U}}^t$. Pero por el teorema 5.2.3, tenemos que $[T^*]_{\mathcal{U}} = [T]_{\mathcal{U}}^t$. Por lo tanto $[T^*]_{\mathcal{U}} = [T]_{\mathcal{U}}$, lo cual implica que $T = T^*$. □

Ejemplo 5.3.3. Sea $P : V \to V$ una proyección ortogonal, entonces P es un operador autoadjunto.

Veamos que es así: sea $W \subset V$, tal que P proyecta ortogonalmente a W, es decir $V = W \oplus W^{\perp}$, con $P(w) = w$, $w \in W$ y $P(w') = 0$, $w' \in W^{\perp}$.

Entonces si $v_1, v_2 \in V$, tenemos que $v_1 = w_1 + w_1', v_2 = w_2 + w_2'$ con $w_1, w_2 \in W$, $w_1', w_2' \in W^{\perp}$. Luego

$$\langle P(v_1), v_2 \rangle = \langle w_1, w_2 + w_2' \rangle = \langle w_1, w_2 \rangle = \langle w_1, P(v_2) \rangle = \langle v_1, P(v_2) \rangle.$$

Proposición 5.3.4. *Sea V un espacio vectorial con producto interno. Entonces, el conjunto de operadores lineales autoadjuntos es un espacio vectorial.*

Demostración. El resultado se deduce fácilmente de la proposición 5.2.4 (2) y (3). □

Proposición 5.3.5. *Sean S y T dos operadores lineales autoadjuntos. Entonces, ST es autoadjunto si y sólo si S y T conmutan.*

Demostración. ($\Rightarrow$) Como ST es autoadjunto, tenemos que $ST = (ST)^*$. Por proposición 5.2.4 (4) tenemos que $(ST)^* = T^*S^*$, y como S, T son autoadjuntos $T^*S^* = TS$. Reconstruyendo las igualdades tenemos

$$ST = (ST)^* = T^*S^* = TS,$$

es decir, S y T conmutan.

($\Leftarrow$)

$$(ST)^* = T^*S^* = TS = ST.$$

□

Ejemplo. Sean $T, S : \mathbb{R}^2 \to R^2$ operadores lineales definidos

$$T(x, y) = (x, 2y), \quad S(x, y) = (y, x).$$

Calculemos T^* y S^*. T^* debe satisfacer que

$$\begin{aligned} \langle e_1, T^*(x, y) \rangle &= \langle T(e_1), (x, y) \rangle = \langle e_1, (x, y) \rangle = x, \\ \langle e_2, T^*(x, y) \rangle &= \langle T(e_2), (x, y) \rangle = \langle 2e_2, (x, y) \rangle = 2y. \end{aligned}$$

Es decir, $T^* = T$. Análogamente, se muestra que $S^* = S$. Ahora bien

$$\begin{aligned} \langle e_1, (TS)^*(x, y) \rangle &= \langle TS(e_1), (x, y) \rangle = \langle T(e_2), (x, y) \rangle = \langle 2e_2, (x, y) \rangle = 2y, \\ \langle e_2, (TS)^*(x, y) \rangle &= \langle TS(e_2), (x, y) \rangle = \langle T(e_1), (x, y) \rangle = \langle e_1, (x, y) \rangle = x. \end{aligned}$$

Es decir $(TS)^*(x, y) = (2y, x)$. Por otro lado, $TS(x, y) = T(y, x) = (y, 2x)$. Luego $(TS)^* \neq TS$, es decir TS no es autoadjunto.

Esto ocurre pues, $TS(x, y) = (y, 2x)$ es distinto a $ST(x, y) = S(x, 2y) = (2y, x)$. Es decir, S y T no conmutan.

Ejemplo. En el ejemplo 5.3.3 vimos que si P es una proyección ortogonal, entonces es operador autoadjunto. Veamos ahora que es diagonalizable: sea P la proyección ortogonal a W, y tomemos $\mathcal{U}_0 = \{u_1, \ldots, u_k\}$ una base de W y $\mathcal{U}_1 = \{u_{k+1}, \ldots, u_n\}$ una base de $W^\perp$, luego $\mathcal{U} = \{u_1, \ldots, u_n\}$ es una base de V con la siguiente particularidad

$$P(u_i) = u_i, \quad 1 \leqslant i \leqslant k, \qquad P(u_i) = 0, \quad k+1 \leqslant i \leqslant n.$$

Luego, la base $\mathcal{U}$ consta de autovectores, de los cuales los primeros k tienen autovalor 1 y los siguientes tienen autovalor 0.

Veremos ahora la demostración completa de que un operador autoadjunto es diagonalizable, es decir que hay una base de autovectores del operador.

Proposición 5.3.6. *Sea V un espacio vectorial con producto interno y $T : V \to V$ una transformación lineal. Sea W un subespacio de V invariante por T. Entonces $W^\perp$ es invariante por T^*.*

Demostración. Debemos ver que $T^*(W^\perp) \subset W^\perp$, es decir que $\langle T^*(W^\perp), W\rangle = 0$. Pero,

$$\langle T^*(W^\perp), W\rangle = \langle W^\perp, T(W)\rangle \subseteq \langle W^\perp, W\rangle = 0.$$

□

De lo cual se deduce inmediatamente:

Corolario 5.3.7. *Si $T : V \to V$ una transformación lineal autoadjunta y W un subespacio de V invariante por T, entonces $W^\perp$ es invariante por T.*

Observación 5.3.8. Sea V espacio vectorial sobre $\mathbb{R}$, $\langle\ ,\ \rangle$ un producto interno y $\mathcal{U} = \{u_1, \ldots, u_n\}$ una BON de V. Si $x, y \in V$ con $x = \sum x_i u_i$ y $y = \sum y_i u_i$, entonces

$$\langle x, y\rangle = \sum_i \sum_j \langle x_i, y_j\rangle = \sum_i x_i y_i.$$

Por otro lado

$$[x]_{\mathcal{U}}^t\, [y]_{\mathcal{U}} = \sum_i x_i y_i.$$

Es decir, si damos por sobrentendida la base y denotamos $x = [x]_{\mathcal{U}}$, $y = [y]_{\mathcal{U}}$, tenemos que

$$\langle x, y\rangle = x^t y.$$

En este contexto, si $T : V \to V$ lineal y $A = [T]_{\mathcal{U}}$, entonces podemos pensar a la transformación lineal como $A : \mathbb{R}^n \to \mathbb{R}^n$ y al producto interno como el producto interno canónico de $\mathbb{R}^n$.

Teorema 5.3.9. *Sea $T : V \to V$ operador autoadjunto. Entonces existe $\lambda \in \mathbb{R}$ autovalor de T.*

Demostración. Si $\mathcal{U}$ una BON de V y $A = [T]_{\mathcal{U}}$. Como vimos en la observación 5.3.8, podemos pensar a la transformación lineal como $A : \mathbb{R}^n \to \mathbb{R}^n$ y al producto escalar como el canónico en $\mathbb{R}^n$. Observar que como T es autoadjunta, entonces A es simétrica y la demostración se obtiene directamente del teorema 4.8.7. □

Teorema 5.3.10 (Teorema espectral)**.** *Sea* $T : V \to V$ *un operador autoadjunto. Entonces existe* $\mathcal{U} = \{u_1, \ldots, u_n\}$ *una BON de* V *de autovectores de* T*.*

Demostración. Se hará por inducción en $n = \dim(V)$.

Si $n = 1$ es trivial.

Supongamos que vale para $n-1$ con $n > 1$, y probaremos el resultado para n. Por le teorema 5.3.9 existe $\lambda \in \mathbb{R}$ y $v \in V$ tal que $Av = \lambda v$. Si $u_n = v/||v||$, u_n tiene norma 1 y cumple también que $Au_n = \lambda u_n$. Sea $W = \langle u \rangle$. Entonces por corolario 5.3.7, $W^\perp$ es invariante. Podemos considerar entonces a T como una transformación lineal de $W^\perp$ a $W^\perp$. Como $\dim(W^\perp) = n-1$, por hipótesis inductiva existe $\{u_1, \ldots, u_{n-1}\}$ una BON de $W^\perp$ de autovectores de $T : W^\perp \to W^\perp$. Es claro entonces que $\mathcal{U} = \{u_1, \ldots, u_n\}$ una BON de V de autovectores de $T : V \to V$. □

Observación. La recíproca del teorema anterior también es válida: si $\mathcal{U} = \{u_1, \ldots, u_n\}$ es una BON de V de autovectores de T, entonces T es autoadjunto. Esto se debe a que la matriz de T en la base $\mathcal{U}$ es diagonal y por lo tanto simétrica (ver proposición 5.3.2). También se puede demostrar directamente dando $v, w \in V$, escribiendo cada uno en términos de la base y viendo que $\langle T(v), w\rangle = \langle v, T(w)\rangle$.

Proposición 5.3.11. *Sea* $T : V \to V$ *autoadjunto. Si* $\lambda_1, \ldots, \lambda_k$ *son los autovalores de* T*, entonces*

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k},$$

y esta suma directa es ortogonal, es decir $V_{\lambda_i} \perp V_{\lambda_j}$ *si* $i \neq j$*. (Recordemos que* $V_\lambda = \{v \in V : T(v) = \lambda v\}$*, es el autoespacio con autovalor* λ*).*

Demostración. Como existe $\mathcal{U} = \{u_1, \ldots, u_n\}$ una BON de V de autovectores de T, es claro que los vectores de la base con autovalor λ_i generan V_{λ_i} y que la suma de los V_{λ_i} genera todo. Debemos ver ahora que

$$V_{\lambda_i} \cap \sum_{j \neq i} V_{\lambda_j} = 0$$

. Pero esto es claro porque (reordenando $\mathcal{U}$)

$$V_{\lambda_i} = \operatorname{span}\{u_1, \ldots, u_k\} \quad \text{y} \quad \sum_{j \neq i} V_{\lambda_j} = \operatorname{span}\{u_{k+1}, \ldots, u_n\}.$$

□

Definición 5.3.12. Diremos que un operador lineal $T : V \to V$ es *no negativo*, y escribiremos $T \geqslant 0$, cuando T es autoadjunto y además $\langle T(v), v\rangle \geqslant 0$ para todo $v \in V$. En el caso que $\langle T(v), v\rangle > 0$ para todo $v \in V$, diremos que T es un operador *positivo* y escribiremos $T > 0$.

Teorema 5.3.13. *Un operador autoadjunto $T : V \to V$ es no negativo si y sólo si sus autovalores son todos $\geqslant 0$. Por otro lado, T es positivo si y solo si sus autovalores son todos > 0.*

Demostración. Demostremos la primera afirmación.

($\Rightarrow$) Sea $v \in V$ con autovalor λ, entonces

$$0 \leqslant \langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda ||v||^2.$$

Como $||v||^2 > 0$, tenemos que $\lambda \geqslant 0$.

($\Leftarrow$) Sea $v \in V$. Debemos ver que $\langle T(v), v\rangle \geqslant 0$. Sea $\mathcal{U} = \{u_1, \ldots, u_n\}$ una BON de V de autovectores de T, entonces $v = \sum_i a_i u_i$, luego

$$\langle T(v), v\rangle = \langle T(\sum_i a_i u_i), \sum_j a_j u_j\rangle = \langle \sum_i \lambda_i a_i u_i, \sum_j a_j u_j\rangle = \sum_i \sum_j \lambda_i a_i a_j \langle u_i, u_j\rangle = \sum_i \lambda_i a_i^2.$$

Como por hipótesis $\lambda_i \geqslant 0$, tenemos que $\sum_i \lambda_i a_i^2 \geqslant 0$, por lo tanto $\langle T(v), v\rangle \geqslant 0$.

La segunda afirmación se prueba de manera totalmente análoga (cambiando $\geqslant$ por $>$). □

Observación. En la demostración del teorema anterior hemos demostrado que si T tiene una BON $\mathcal{U} = \{u_1, \ldots, u_n\}$ de autovectores con autovalores $\lambda_1, \ldots, \lambda_n$, entonces si $v = \sum_i a_i u_i$,

$$\langle T(v), v\rangle = \sum_{i=1}^{n} \lambda_i a_i^2 \tag{5.3.1}$$

Corolario 5.3.14. *Sea $T \geqslant 0$. Si $\langle T(v), v\rangle = 0$, entonces $T(v) = 0$.*

Demostración. Sea $\mathcal{U} = \{u_1, \ldots, u_n\}$ una BON de autovectores de T con autovalores $\lambda_1, \ldots, \lambda_n$.

Reordenemos la BON de tal forma que $\lambda_1, \ldots, \lambda_k$ sean no nulos y $\lambda_{k+1}, \ldots, \lambda_n$ sean cero. Sea $v \in V$, entonces $v = \sum_i a_i u_i$ y $T(v) = \sum_{i=1}^{k} \lambda_i a_i u_i$. Por la ecuación (5.3.1) tenemos que

$$\langle T(v), v\rangle = \sum_{i=1}^{n} \lambda_i a_i^2 = \sum_{i=1}^{k} \lambda_i a_i^2.$$

Si $\langle T(v), v\rangle = 0$, tenemos entonces que $a_1 = \cdots = a_k = 0$ y por lo tanto $T(v) = \sum_{i=1}^{k} \lambda_i a_i u_i = 0$.

□

Corolario 5.3.15. *Sea* T *operador lineal, entonces* $T > 0$ *si y sólo si* $T \geqslant 0$ *y* T *inversible.*

Demostración. ($\Rightarrow$) Como $T > 0$, claramente $T \geqslant 0$. Por otro lado, si $v \neq 0$, entonces $\langle T(v), v\rangle > 0$, luego $T(v) \neq 0$, por lo tanto T es inyectiva, luego es biyectiva.

($\Leftarrow$) Sea $v \in V$, $v \neq 0$. Si $\langle T(v), v\rangle = 0$, entonces por el corolario 5.3.14, tenemos que $T(v) = 0$, lo cual no puede ser pues T es inversible. Por lo tanto, $\langle T(v), v\rangle \neq 0$ y como $T \geqslant 0$, $\langle T(v), v\rangle > 0$. □

Definición 5.3.16. Una matriz $A \in M_n(\mathbb{R})$ se dice *no negativa* (resp. *positiva*) si el operador lineal asociado es no negativo (resp. positivo).

Si $A \in M(n \times n)$ s y $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$, tenemos que

$$Av = (\sum_i a_{i1}x_i, \ldots, \sum_i a_{in}x_i),$$

luego

$$\langle A(v), v\rangle = \langle (\sum_i a_{i1}x_i, \ldots, \sum_i a_{in}x_i), (x_1, \ldots, x_n)\rangle = \sum_{i,j=1}^{n} a_{ij}x_i x_j.$$

Es decir, una matriz A es no negativa si para todo $(x_1, \ldots, x_n) \in \mathbb{R}^n$, se cumple que

$$\sum_{i,j=1}^{n} a_{ij}x_i x_j \geqslant 0.$$

Análogamente, si A es positiva entonces

$$\sum_{i,j=1}^{n} a_{ij}x_i x_j > 0$$

para todo $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Ejemplo. Sea A una matriz 2×2 simétrica, es decir

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

con $a, b, c \in \mathbb{R}$. Veamos cuando A es definida positiva.

El polinomio característico de A es

$$\det \begin{bmatrix} a-\lambda & b \\ b & c-\lambda \end{bmatrix} = (a-\lambda)(c-\lambda) - b^2 = \lambda^2 + (-a-c)\lambda + ac - b^2$$

Averiguando las raíces de este polinomio y exigiendo que ambas sean mayores que cero (hacer la cuenta), llegamos a que $A > 0$ si y solo si $ac - b^2 > 0$ y $a, c > 0$. Análogamente, $A \geqslant 0$ si y solo si $ac - b^2 \geqslant 0$ y $a, c \geqslant 0$.

Así, por ejemplo $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} > 0$ y $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \geqslant 0$.

Definición 5.3.17. Sea $T : V \to V$ una transformación lineal. Una transformación lineal $S : V \to V$ se llama *raíz cuadrada de* T si $S^2 = T$.

Teorema 5.3.18. *Sea* $T \geqslant 0$*, entonces existe una única* $S \geqslant 0$ *tal que* $S^2 = T$*. En el caso de ser la* T *positiva, la* S *también resulta positiva.*

Demostración. Sean $\lambda_1, \dots, \lambda_n$ los autovalores de T. Entonces, por proposición 5.3.11,

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k},$$

con $V_{\lambda_i} \perp V_{\lambda_j}$ si $i \neq j$. Es decir, si $v \in V$, entonces existen únicos $v_i \in V_{\lambda_i}$ tal que $v = v_1 + \cdots + v_k$ y $\langle v_i, v_j \rangle = 0$ si $i \neq j$.

Observar que $T(v) = \sum_{i=1}^k \lambda_i v_i$. Definimos entonces, $S(v) = \sum_{i=1}^k \sqrt{\lambda_i} v_i$. Claramente, S es lineal. S es no negativa pues

$$\langle S(v), v \rangle = \langle \sum_{i=1}^k \sqrt{\lambda_i} v_i, \sum_{j=1}^k v_j \rangle = \sum_{i=1}^k \sum_{j=1}^k \sqrt{\lambda_i} \langle v_i, v_j \rangle = \sum_{i=1}^k \sqrt{\lambda_i}.$$

Además,

$$S^2(v) = S(\sum_{i=1}^k \sqrt{\lambda_i} v_i) = \sum_{i=1}^k \sqrt{\lambda_i} S(v_i) = \sum_{i=1}^k \lambda_i v_i = T(v).$$

Es claro que si T es positiva, S también resulta positiva.

Es un poco más complicado demostrar que es única. Sea $R \geqslant 0$ tal que $R^2 = T$. Demostraremos que $R = S$ en varios pasos.

(1) Como $R^2 = T$, entonces R conmuta con T, pues $RT = RR^2 = R^3 = R^2 R = TR$.

(2) $R(V_{\lambda_i}) \subseteq V_{\lambda_i}$. Esto se debe a que R y T conmutan: dado $v_i \in V_{\lambda_i}$

$$T(R(v_i)) = R(T(v_i)) = R(\lambda_i v_i) = \lambda_i R(v_i).$$

Es decir $R(v_i)$ es un autovector de T con autovalor λ_i y por lo tanto pertenece a V_{λ_i}.

(3) Como $R(V_{\lambda_i}) \subseteq V_{\lambda_i}$, podemos restringir y correstringir R a V_{λ_i} y obtenemos $R_i = R : V_{\lambda_i} \to V_{\lambda_i}$. Es fácil comprobar que R_i es autoadjunta y por lo tanto existe una BON de V_{λ_i} de autovectores de R_i. Sea w un vector de la base con autovalor μ, entonces $R_i(w) = \mu w$ y

$$\lambda_i w = T(w) = R^2(w) = R_i^2(w) = R_i(\mu w) = \mu^2 w.$$

Luego $\mu = \sqrt{\lambda_i}$. Es decir todo vector de la base de V_{λ_i} es autovector de R con autovalor $\sqrt{\lambda_i}$, por lo tanto $R(v_i) = \sqrt{\lambda_i} v_i$ para todo $v_i \in V_{\lambda_i}$. Luego si $v = v_1 + \cdots + v_k$ con $v_i \in V_{\lambda_i}$:

$$R(v) = R(v_1) + \cdots + R(v_k) = \lambda_1 v_1 + \cdots + \lambda_k v_k = S(v).$$

□

Observación. Es claro que si T es autoadjunto, entonces T^2 es no negativo. Si además T es inversible, entonces T^2 es positivo.

En el caso de operadores que no son autoadjuntos esto no es cierto. Por ejemplo, si

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

entonces A es el operador que rota los vectores 90ž y aplicado dos veces es una rotación de 180ž, es decir

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

que no es una matriz positiva.

Ejemplos de operadores no negativos se obtiene a partir de la siguiente proposición.

Proposición 5.3.19. *Sean V, W espacios vectoriales de dimensión finita con producto interno dado y sea $T : V \to W$ lineal. Los operadores $T^*T : V \to V$ y $TT^* : W \to W$ son no negativos y tiene el mismo rango que T.*

Demostración.

$$(T^*T)^* = T^*T^{**} = T^*T,$$

y por lo tanto T^*T es autoadjunto. De forma análoga se prueba que TT^* es autoadjunto.

$$\langle (T^*T)(v), v\rangle = \langle T^*(T(v)), v\rangle = \langle T(v), T(v)\rangle = ||T(v)||^2 \geqslant 0,$$

luego T^*T es no negativo. De forma análoga se prueba que TT^* es no negativo.

Para ver que el rango de T^*T es igual al rango de T, demostraremos a continuación que $\mathrm{Nu}(T^*T) = \mathrm{Nu}(T)$ (y por el teorema de la dimensión se deduce que $\dim(\mathrm{Im}(T^*T)) = \dim(\mathrm{Im}(T))$).

La inclusión $\mathrm{Nu}(T) \subseteq \mathrm{Nu}(T^*T)$ es obvia: $v \in \mathrm{Nu}(T) \Rightarrow T(v) = 0 \Rightarrow T^*(T(v)) = 0 \Rightarrow v \in \mathrm{Nu}(T^*T)$. Por otro lado

$$\begin{aligned} v \in \mathrm{Nu}(T^*T) &\Rightarrow T^*T(v) = 0 \\ &\Rightarrow T(v) \in \mathrm{Nu}(T^*) \\ &\Rightarrow T(v) \in \mathrm{Im}(T)^\perp \qquad \text{(pues } \mathrm{Nu}(T^*) = \mathrm{Im}(T)^\perp \text{ por teorema 5.2.5)} \\ &\Rightarrow T(v) \in \mathrm{Im}(T) \cap \mathrm{Im}(T)^\perp \\ &\Rightarrow T(v) = 0 \\ &\Rightarrow v \in \mathrm{Nu}(T). \end{aligned}$$

□

Corolario 5.3.20. *Una transformación lineal $T : V \to W$ es inyectiva si y sólo si T^*T es inversible. T es suryectiva si y sólo si TT^* es inversible.*

Demostración. En efecto, T es inyectiva $\Leftrightarrow$ rango$(T) = \dim(V) \Leftrightarrow$ rango$(T^*T) = \dim(V) \Leftrightarrow T^*T$ es inversible.

Análogamente, T es suryectiva $\Leftrightarrow$ rango$(T) = \dim(W) \Leftrightarrow$ rango$(TT^*) = \dim(W) \Leftrightarrow TT^*$ es inversible. □

Observación. Obviamente, si trabajamos con matrices todos los teoremas anteriores se pueden reformular fácilmente. Por ejemplo, si A matriz $n \times m$, entonces las matrices $A^tA \in M(m \times m)$ y $AA^t \in M(n \times n)$ son no negativas y tiene el mismo rango que A.

Ejemplo. Sea $A = \begin{bmatrix} 1 & 2 & -1 \\ -2 & 4 & 0 \end{bmatrix}$, (que es una matriz de rango 2), entonces

$$AA^t = \begin{bmatrix} 6 & 6 \\ 6 & 18 \end{bmatrix}, \qquad A^tA = \begin{bmatrix} 5 & -6 & -1 \\ -6 & 18 & -2 \\ -1 & -2 & 1 \end{bmatrix}.$$

Por el teorema anterior ambas matrices son no negativas y tienen rango 2. En consecuencia, AA^t is inversible y por lo tanto positiva.

A continuación una extensión del Teorema Espectral, válida para transformaciones lineales cualesquiera.

Teorema 5.3.21 (Teorema de los valores singulares)**.** *Sea $T : V \to W$ una transformación lineal de rango r entre espacios de dimensión finita con producto interno. Entonces, existen bases ortonormales $\{v_1, \ldots, v_n\}$ de V y $\{w_1, \ldots, w_m\}$ de W tales que $T(v_i) = \lambda_i w_i$, $T^*(w_i) = \lambda_i v_i$, con $\lambda_i > 0$ para $i = 1, \ldots, r$. Además $T(v_i) = 0$, $T^*(w_i) = 0$ para $i > r$.*

Demostración. Por la proposición 5.3.19, el operador T^*T es no negativo y tiene rango r. Como T^*T es autoadjunto, no negativo y de rango r; por el Teorema Espectral existe una base ortonormal $\{v_1, \ldots, v_n\}$ de V tal que $T^*T(v_i) = \mu_i v_i$ con $\mu_i > 0$ para $1 \leqslant i \leqslant r$ y $\mu_i = 0$ para $i > r$. Definimos $\lambda_i = \sqrt{\mu_i}$.

Ahora bien,

$$\langle T(v_i), T(v_j)\rangle = \langle v_i, T^*T(v_j)\rangle = \langle v_i, \lambda_i^2 v_j\rangle = \lambda_i^2 \langle v_i, v_j\rangle \qquad (1 \leqslant i, j \leqslant n). \tag{5.3.2}$$

Como $\lambda_i \neq 0$ para $1 \leqslant i \leqslant r$, tenemos que $\{T(v_1), \ldots, T(v_r)\}$ es un conjunto ortogonal y si definimos $w_i = T(v_i)/\lambda_i$ $(1 \leqslant i \leqslant r)$, entonces $\{w_1, \ldots, w_r\}$ es un conjunto ortonormal de W. Podemos completar este conjunto a una base y por Gram-Schmidt y normalización obtenemos $\{w_1, \ldots, w_m\}$ un conjunto ortonormal de W que además cumple que $T(v_i) = \lambda_i w_i$ para $1 \leqslant i \leqslant r$.

Por otro lado, de la ecuación (5.3.2) obtenemos que

$$||T(v_i)||^2 = \langle T(v_i), T(v_i)\rangle = \lambda_i^2 ||v_i||^2 = 0, \qquad i > r$$

y por lo tanto $T(v_i) = 0$ para $i > r$. Es decir, hemos probado las afirmaciones sobre T. Veamos ahora que $T^*(w_i) = \lambda_i v_i$ para $i = 1, \ldots, r$ y $T^*(w_i) = 0$ para $i > r$. Si $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$,

$$\langle T^*(w_i), v_j\rangle = \langle w_i, T(v_j)\rangle = \begin{cases} \langle w_i, \lambda_j w_j\rangle = \lambda_j \langle w_i, w_j\rangle & 1 \leqslant j \leqslant r \\ \langle w_i, 0\rangle = 0 & r+1 \leqslant j \leqslant n \end{cases}$$

□

Los número positivos $\lambda_1, \ldots, \lambda_r$ son llamados los *valores singulares* de la transformación $T : V \to W$ de rango r.

Observación. Observar que, por la demostración del teorema, los valores singulares son las raíces cuadradas de los autovalores del operador T^*T.

Observación. En las hipótesis del teorema anterior es claro que

(1) $\{w_1, \ldots, w_r\}$ es una BON de $\mathrm{Im}(T)$.

(2) $\{v_{r+1}, \ldots, v_n\}$ es una BON de $\mathrm{Nu}(T)$.

(3) $\{v_1, \ldots, v_r\}$ es una BON de $\mathrm{Im}(T^*)$.

(4) $\{w_{r+1}, \ldots, w_n\}$ es una BON de $\mathrm{Nu}(T^*)$.

Observación 5.3.22. Por el teorema anterior, si $T : V \to W$ operador de rango r y $\lambda_1, \ldots, \lambda_r$ son los valores singulares de T, entonces hay una BON $\{v_1, \ldots, v_n\}$ de V y una BON $\{w_1, \ldots, w_m\}$ de W tal que

$$\begin{array}{lll} (T^*T)v_i = \lambda_i^2 v_i, & (TT^*)w_i = \lambda_i^2 w_i, & 1 \leqslant i \leqslant r, \\ (T^*T)v_j = 0, & (TT^*)w_k = 0, & r+1 \leqslant j \leqslant n,\ r+1 \leqslant k \leqslant m. \end{array}$$

Observación. El teorema de los valores singulares se traduce fácilmente al lenguaje de las matrices.

Sea A *matriz* $m \times n$ *de rango* r*, entonces, existen bases ortonormales* $\{v_1, \ldots, v_n\}$ *de* $\mathbb{R}^n$ *y* $\{w_1, \ldots, w_m\}$ *de* $\mathbb{R}^m$ *tales que* $Av_i = \lambda_i w_i$, $A^t w_i = \lambda_i v_i$, *con* $\lambda_i > 0$ *para* $i = 1, \ldots, r$. *Además* $Av_i = 0$, $A^t w_i = 0$ *para* $i > r$.

Observar que los valores singulares son las raíces cuadradas de los autovalores de la matriz A^tA.

5.4 OPERADORES ANTISIMÉTRICOS Y OPERADORES ORTOGONALES (*)

Definición 5.4.1. Sea $T : V \to V$ un operador lineal de V espacio vectorial de dimensión finita con producto interno. diremos que T es *antisimétrico* si $\langle Tv, w\rangle = -\langle v, Tw\rangle$ para $v, w \in V$, es decir si $T^* = -T$. Por otro lado diremos que T es *ortogonal* si T es inversible y $T^* = T^{-1}$.

Ejemplo. Sea $T(x, y, z) = (-y + 2z, x + 3z, -2x - 3y)$. Veamos que T es antisimétrico.

$$\begin{aligned}
\langle T^*(x, y, z), e_1\rangle &= \langle (x, y, z), Te_1\rangle = \langle (x, y, z), (0, 1, -2)\rangle = y - 2z.\\
\langle T^*(x, y, z), e_2\rangle &= \langle (x, y, z), Te_2\rangle = \langle (x, y, z), (-1, 0, -3)\rangle = -x - 3z.\\
\langle T^*(x, y, z), e_3\rangle &= \langle (x, y, z), Te_3\rangle = \langle (x, y, z), (2, 3, 0)\rangle = 2x + 3y.
\end{aligned}$$

Luego, $T^*(x, y, z) = (y - 2z, -x - 3z, 2x + 3y) = -T(x, y, z)$.

Definición 5.4.2. Una matriz $A \in M_n(\mathbb{R})$ es *antisimétrica* si $A^t = -A$. Diremos que A es *ortogonal* si A tiene inversa y $A^t = A^{-1}$.

Ejercicio. T es un operador antisimétrico si y solo si la matriz de T en cualquier base ortonormal es antisimétrica.

Observación. Sea $T : V \to V$ operador lineal ($\dim(V) \leqslant \infty$). Entonces

$$T = \frac{T + T^*}{2} + \frac{T - T^*}{2}$$

Ahora bien,

$$\begin{aligned}
(T + T^*)^* &= T^* + T^{**} = T^* + T = T + T^*,\\
(T - T^*)^* &= T^* - T^{**} = T^* - T = -(T - T^*).
\end{aligned}$$

Es decir, $\dfrac{T + T^*}{2}$ es un operador simétrico y $\dfrac{T - T^*}{2}$ es un operador antisimétrico. Concluyendo: todo operador lineal es suma de un operador simétrico y un operador antisimétrico.

Proposición 5.4.3. *Sea $A \in M_n(\mathbb{R})$. Entonces A ortogonal si y solo si las n columnas de A forman um conjunto ortonormal en $\mathbb{R}^n$.*

Demostración. Recordemos que los vectores de $\mathbb{R}^n$ se escriben como columnas cuando los vemos como matrices, es decir son matrices $n \times 1$

($\Leftarrow$) Sea $\{u_1, \ldots, u_n\}$ una BON de $\mathbb{R}^n$ y

$$A = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}.$$

Entonces

$$A^t = \begin{bmatrix} u_1^t \\ \vdots \\ u_n^t \end{bmatrix}$$

que es la matriz cuyas filas son los vectores de la base.

Si $v \in \mathbb{R}^n$, tenemos que

$$A^t v = \begin{bmatrix} u_1^t v \\ \vdots \\ u_n^t v \end{bmatrix} = \begin{bmatrix} \langle u_1, v\rangle \\ \vdots \\ \langle u_n, v\rangle. \end{bmatrix}$$

Como $v = \sum_i \langle u_i, v\rangle u_i$, es claro que si $v \neq 0$, entonces existe i tal que $\langle u_i, v\rangle u_i \neq 0$. Por lo tanto $A^t v = 0 \Rightarrow v = 0$, luego A^t es inyectiva y por lo tanto biyectiva.

Ahora bien, el coeficiente ij de la matriz A^tA es $u_i^t u_j = \langle u_i, u_j\rangle$ que es 1 si $i = j$ y 0 si $i \neq j$. Es decir los únicos coeficientes no nulos son los de la diagonal y valen 1 y por lo tanto $A^tA = I$. Además,

$$A^tA = I \Rightarrow A^tAA^{-1} = A^{-1} \Rightarrow A^t = A^{-1}.$$

($\Rightarrow$) Es fácil comprobar (con el mismo tipo de cuentas) que si $A^tA = I$, entonces los vectores columna de A forman una BON.

□

Para remarcar lo demostrado, decimos A es una matriz ortogonal si y sólo si

$$A = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix},$$

donde $\{u_1, \ldots, u_n\}$ una BON de $\mathbb{R}^n$.

Ejemplo. Las matrices ortogonales 2×2 son de la forma

$$A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \qquad \text{o} \qquad A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

para $\theta \in [0, 2\pi]$.

Lema 5.4.4. *Sean* $\mathcal{U} = \{u_1, \ldots, u_n\}$ *y* $\mathcal{V} = \{v_1, \ldots, v_n\}$ *dos BON de* V. *Entonces, si* $A = [T]_{\mathcal{U}\mathcal{V}}$, *implica que* $[T^*]_{\mathcal{V}\mathcal{U}} = A^t$.

Demostración. Sea $A = [a_{ij}]$. Tenemos

$$T(u_j) = \sum_i a_{ij} w_i \;\Rightarrow \langle T(u_j), w_i\rangle = a_{ij}.$$

Ahora bien,

$$a_{ij} = \langle T(u_j), w_i\rangle = \langle u_j, T^*(w_i)\rangle \Rightarrow T^*(w_i) = \sum_j a_{ij} u_j$$

y por lo tanto $[T^*]_{\mathcal{V}\mathcal{U}} = A^t$.

□

Teorema 5.4.5. *Sean* $\mathcal{U} = \{u_1, \ldots, u_n\}$ *y* $\mathcal{V} = \{v_1, \ldots, v_n\}$ *dos BON de* V. *Entonces,* $T : V \to V$ *es ortogonal si y sólo si* $[T]_{\mathcal{U}\mathcal{V}}$ *es ortogonal.*

Demostración. Por el lema anterior,

$$A^tA = [T^*]_{\mathcal{V}\mathcal{U}}[T]_{\mathcal{U}\mathcal{V}} = [T^*T]_{\mathcal{V}},$$
$$AA^t = [T]_{\mathcal{U}\mathcal{V}}[T^*]_{\mathcal{V}\mathcal{U}} = [TT^*]_{\mathcal{U}}$$

Luego $T^*T = \mathrm{Id}$ si y solo si $A^tA = I$ y $TT^* = \mathrm{Id}$ si y solo si $AA^t = I$. □

Teorema 5.4.6. *Sea $T : V \to V$ un operador en un espacio vectorial de dimensión finita con producto interno. Entonces son equivalentes,*

(1) *T es ortogonal.*

(2) *T preserva producto interno, es decir $\langle Tv, Tw\rangle = \langle v, w\rangle$, para cualesquiera $v, w \in V$.*

(3) *Si $\mathcal{U} = \{u_1, \ldots, u_n\}$ es una BON de V, entonces $\{Tu_1, \ldots, Tu_n\}$ es también una BON de V.*

Demostración. (1) ⇒ (2) $\langle Tv, Tw\rangle = \langle v, T^*Tw\rangle = \langle v, w\rangle$.

(2) ⇒ (3) $\langle Tu_i, Tu_j\rangle = \langle u_i, u_j\rangle = \delta_{ij}$ (donde δ_{ij} es el símbolo de Kroneker, $\delta_{ij} = 0$ si $i \neq j$ y $\delta_{ii} = 1$), luego $\{Tu_1, \ldots, Tu_n\}$ es un conjunto ortonormal de n elementos, luego son LI y por consiguiente una BON.

(3) ⇒ (1) Sea $\mathcal{V} = \{Tu_1, \ldots, Tu_n\}$. Entonces, $[T]_{\mathcal{U}\mathcal{V}} = I$ y por lo tanto es ortogonal. Por el teorema 5.4.5, T es ortogonal. □

Veremos a continuación la descomposición polar y la descomposición en valores singulares (DVS) de un operador lineal o una matriz.

Teorema 5.4.7 (Descomposición polar)**.** *Sea $T : V \to V$ operador lineal en un espacio vectorial de dimensión finita con producto interno. Entonces T admite una descomposición*

$$T = PU$$

donde $P : V \to V$ es no negativa y $U : V \to V$ es ortogonal.

Demostración. De acuerdo al teorema de los valores singulares,existen bases ortonormales $\{v_1, \ldots, v_n\}$ y $\{w_1, \ldots, w_n\}$ de V tales que $Tv_i = \lambda_i w_i$, $T^*w_i = \lambda_i v_i$, con $\lambda_i > 0$ para $i = 1, \ldots, r$. Además $Tv_i = 0$, $T^*w_i = 0$ para $r < i \leqslant n$. Definiendo $\lambda_i = 0$ para $r < i \leqslant n$, tenemos que $Tv_i = \lambda_i w_i$, $T^*w_i = \lambda_i v_i$, para $i = 1, \ldots, n$.

Definimos P y U como los operadores que satisfacen $P(w_i) = \lambda_i w_i$ y $Uv_i = w_i$, para $i = 1, \ldots, n$. Es claro que P es autoadjunto y semidefinido y como U lleva una base ortonormal en otra base ortonormal, entonces U es ortonormal (teorema 5.4.6).

Verifiquemos ahora que $T = PU$:

$$PU(v_i) = P(w_i) = \lambda_i w_i = T(v_i), \qquad (1 \leqslant i \leqslant n).$$

Por lo tanto PU y T coinciden en una base y esto implica que son iguales. □

Sea A matriz real $m \times n$. Una *descomposición en valores singulares* o *DVS* de A es una factorización del tipo $A = U\Sigma V^t$ con $U \in M_m(\mathbb{R})$, $V \in M_n(\mathbb{R})$ ortogonales y $\Sigma \in M_{m\times n}(\mathbb{R})$ una matriz formada con los valores singulares de A en su diagonal principal ordenados de mayor a menor.

De acuerdo al teorema de los valores singulares,existen bases ortonormales $\{v_1, \ldots, v_n\}$ y $\{w_1, \ldots, w_n\}$ de V tales que $Tv_i = \lambda_i w_i$, $T^* w_i = \lambda_i v_i$, con $\lambda_i > 0$ para $i = 1, \ldots, r$. Además $Tv_i = 0$, $T^* w_i = 0$ para $r < i \leqslant n$. Reordenando las bases, podemos suponer que $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r > 0$.

Sean

$$U = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}, \qquad V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}. \tag{5.4.1}$$

Teorema 5.4.8.

$$A = U\Sigma V^t$$

Demostración.

$$AV = A \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} Av_1 & \cdots & Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 w_1 & \cdots & \lambda_r w_r & 0 & \cdots & 0 \end{bmatrix}.$$

Como Σ es diagonal, no es difícil verificar que

$$U\Sigma = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \Sigma = \begin{bmatrix} \lambda_1 w_1 & \cdots & \lambda_n w_r & 0 & \cdots & 0 \end{bmatrix}.$$

Por lo tanto,

$$AV = U\Sigma.$$

Como V es una matriz ortogonal $V^{-1} = V^t$, luego multiplicando a derecha por V^t la ecuación anterior, obtenemos

$$A = U\Sigma V^t.$$

□

Parte III

APÉNDICES

A

NÚMEROS COMPLEJOS

A.1 CUERPOS

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

Definición A.1.1. Un conjunto $\mathbb{K}$ es un *cuerpo* si es un anillo de división conmutativo, es decir, un anillo conmutativo con unidad en el que todo elemento distinto de cero es invertible respecto del producto. Por tanto, un cuerpo es un conjunto $\mathbb{K}$ en el que se han definido dos operaciones, '+' y '$\cdot$', llamadas *adición* y *multiplicación* respectivamente, que cumplen las siguientes propiedades. En la siguiente lista de axiomas a, b, c denotan elementos arbitrarios de $\mathbb{K}$, y 0 y 1 denotan elementos especiales de $\mathbb{K}$ que cumplen las propiedades especificadas más abajo.

I1. $a + b$ y $a \cdot b$ pertenecen a $\mathbb{K}$.

I2. *Conmutatividad.* $a + b = b + a$; $ab = ba$.

I3. *Asociatividad.* $(a + b) + c = a + (b + c)$; $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

I4. *Existencia de elemento neutro.* Existen números $0, 1 \in \mathbb{K}$ con $0 \neq 1$ tal que $a + 0 = a$; $a \cdot 1 = a$.

I5. *Distributividad.* $a \cdot (b + c) = a \cdot b + a \cdot c$.

I6. *Existencia del inverso aditivo.* Por cada a en $\mathbb{K}$ existe un único $-a$ en $\mathbb{K}$ tal que $a + (-a) = 0$.

I7. *Existencia de inverso multiplicativo.* Si a es distinto de 0, existe un único elemento $a^{-1} \in \mathbb{K}$ tal que $a \cdot a^{-1} = 1$.

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

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

De forma análoga, usaremos la notación

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

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

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

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

También podemos deducir

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

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

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

y más generalmente

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

y también

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

así como

$$a \cdot 0 = 0,$$

todas reglas familiares de la aritmética elemental.

A.1.1 *Un cuerpo finito*

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

$$0 + 0 = 0, \qquad 0 + 1 = 1, \qquad 1 + 0 = 1, \qquad 1 + 1 = 0$$

y el producto $\cdot \colon \mathbb{F}_2 \times \mathbb{F}_2 \to \mathbb{F}_2$ como

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

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

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

A.2 NÚMEROS COMPLEJOS

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

Definición A.2.1. Los *números complejos* es el conjunto $\mathbb{C}$ de los pares ordenados (a, b), denotados $a + ib$, con a, b en $\mathbb{R}$, con las operaciones '+' y '·', definidas

$$(a + ib) + (c + id) := (a + c) + i(c + d), \tag{A.2.1}$$

$$(a + ib) \cdot (c + id) := (ac - bd) + i(ad + bc). \tag{A.2.2}$$

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

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

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

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

La definición de la suma de dos números complejos no debería sorprendernos, pues es la suma "coordenada a coordenada". La definición del producto se basa en que deseamos que $i^2 = -1$, es decir que i sea la solución de la ecuación polinómica $x^2 + 1 = 0$, y que el producto sea distributivo.

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

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

y por lo tanto $i^2 + 1 = -1 + 1 = 0$.

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

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

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

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

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

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

Proposición A.2.2. *Sean* $0 = 0 + i \cdot 0, 1 = 1 + i \cdot 0 \in \mathbb{C}$. *Entonces,* $\mathbb{C}$ *con las operaciones '+' y '·', definidas en (A.2.1) y (A.2.2), respectivamente, es un cuerpo con elementos neutros* 0 *y* 1, *y*

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

Demostración. Ejercicio. □

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

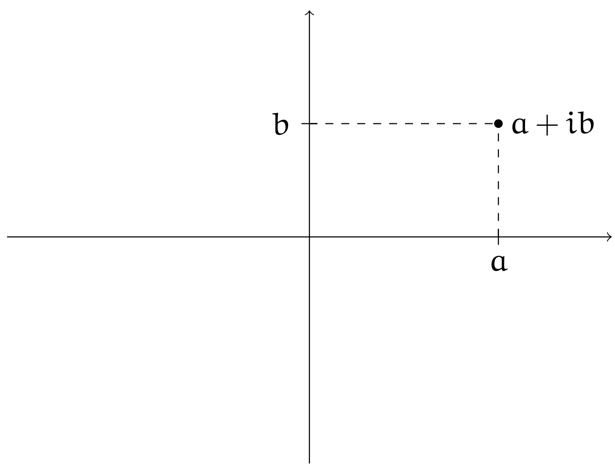


Figura 23: Representación gráfica de los números complejos.

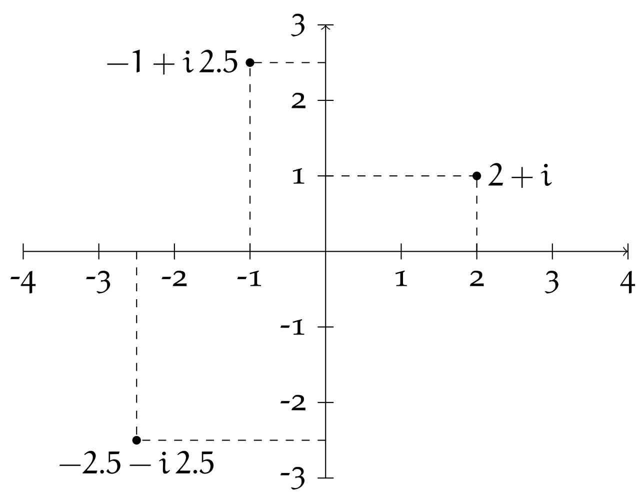


Figura 24: Ejemplos de la representación gráfica de los números complejos.

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

Definición A.2.3. Sea $z = a + ib \in \mathbb{C}$. El *módulo* de z es

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

El *conjugado* de z es

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

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

Proposición A.2.4. *Sean z y w números complejos.*

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

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

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

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

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

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

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

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

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

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

Ejercicio. Determinar el número complejo $2 - 3i + \dfrac{i}{1 - i}$.

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

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

□

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

$$x = r\,\mathrm{sen}(\theta), \qquad y = r\cos(\theta)$$

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

$$a = |z|\,\mathrm{sen}(\theta), \qquad b = |z|\cos(\theta),$$

es decir

$$z = |z|(\cos(\theta) + i\,\mathrm{sen}(\theta)). \tag{A.2.3}$$

Si $z \in \mathbb{C}$, la fórmula (A.2.3) e llamada la *forma polar* de z y θ es llamado el *argumento de* z.

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

$$e^{i\theta} = \cos(\theta) + i\,\mathrm{sen}(\theta). \tag{A.2.4}$$

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

$$z = re^{i\theta}$$

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

Proposición A.2.5. *Sean* $z_1 = r_1e^{i\theta_1}$, $z_2 = r_2e^{i\theta_2}$, *entonces*

$$z_1z_2 = r_1r_2\,e^{i(\theta_1+\theta_2)}.$$

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

$$\begin{aligned} z_1z_2 &= r_1r_2(\cos(\theta_1) + i\,\mathrm{sen}(\theta_1))(\cos(\theta_2) + i\,\mathrm{sen}(\theta_2)) \\ &= r_1r_2(\cos(\theta_1)\cos(\theta_2) + i\cos(\theta_1)\,\mathrm{sen}(\theta_2) + i\,\mathrm{sen}(\theta_1)\cos(\theta_2) + i^2\,\mathrm{sen}(\theta_1)\,\mathrm{sen}(\theta_2)) \\ &= r_1r_2((\cos(\theta_1)\cos(\theta_2) - \mathrm{sen}(\theta_1)\,\mathrm{sen}(\theta_2)) + i(\,\mathrm{sen}(\theta_1)\cos(\theta_2) + \cos(\theta_1)\,\mathrm{sen}(\theta_2))) \\ &\overset{(*)}{=} r_1r_2(\cos(\theta_1 + \theta_2) + i\,\mathrm{sen}(\theta_1 + \theta_2)) = r_1r_2e^{i(\theta_1+\theta_2)}. \end{aligned}$$

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

Observación (Identidad de Euler). Los alumnos que conozcan las series de Taylor reconocerán inmediatamente la fórmula

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!}x^n,$$

donde x es un número real. Ahora bien, remplacemos x por $i\theta$ y obtenemos

$$\begin{aligned} e^{i\theta} &= \sum_{n=0}^{\infty} \frac{1}{n!}(i\theta)^n \\ &= \sum_{k=0}^{\infty} \frac{1}{(2k)!}(i\theta)^{2k} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}(i\theta)^{2k+1}. \end{aligned} \tag{*}$$

No es difícil ver que $i^{2k} = (-1)^k$ y por lo tanto $i^{2k+1} = i^{2k} \cdot i = (-1)^k i$. Por lo tanto, por (*),

$$\begin{aligned} e^{i\theta} &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\theta^{2k} + i\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\theta^{2k+1} \\ &= \cos(\theta) + i\,\mathrm{sen}(\theta), \end{aligned}$$

recuperando así la fórmula (A.2.4), llamada *fórmula de Euler.* Observemos que especializando la fórmula en π obtenemos

$$e^{i\pi} = \cos(\pi) + i\,\mathrm{sen}(\pi) = -1.$$

Escrito de otra forma

$$e^{i\pi} - 1 = 0. \tag{A.2.5}$$

Esta última expresión es denominada la *identidad de Euler* y es considerada una de las fórmulas más relevantes de la matemática, pues comprende las cinco constantes matemáticas más importantes.

Las cinco constantes son:

(1) El número **0**.

(2) El numero **1**.

(3) El número **1**, número irracional que es la relación entre la circunferencia de un círculo y su diámetro. Es aproximadamente $3.14159\ldots$.

(4) El número **e**, también un número irracional. Es la base de los logaritmos naturales y surge naturalmente a través del estudio del interés compuesto y el cálculo. El número e está presente en una gran cantidad de ecuaciones importantes. Es aproximadamente $2.71828\ldots$.

(5) El número **i**, el más fundamental de los números imaginarios.

B

FUNCIONES POLINÓMICAS

En este apéndice se definirán las funciones polinómicas y se mostrarán algunas de sus propiedades fundamentales. Trabajaremos sobre $\mathbb{K}$ cuerpo con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$.

B.1 DEFINICIÓN DE FUNCIONES POLINÓMICAS

Definición B.1.1. Una función $f : \mathbb{K} \to \mathbb{K}$ es *polinomial* o *polinómica* o directamente decimos que f es un *polinomio*, si existen $a_0, a_1, \ldots, a_n \in \mathbb{K}$ tal que

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \tag{B.1.1}$$

para todo $x \in \mathbb{K}$. En este caso diremos que f *tiene grado* $\leqslant n$.

Estaríamos tentados en decir que una función polinómica como en (B.1.1) con $a_n \neq 0$ tiene grado n, pero debemos ser cuidadosos pues todavía no sabemos si la escritura de una función polinómica es única. Es decir, existe la posibilidad de f se escriba de otra forma y por lo tanto el coeficiente más significativo sea diferente. Veremos más adelante que esto no puede ocurrir.

Sea f un polinomio. Si c es un número tal que $f(c) = 0$, entonces llamamos a c *una raíz de* f. Veremos en un momento que un polinomio distinto de cero puede tener solo un número finito de raíces, y daremos un límite para la cantidad de estas raíces.

Ejemplo. Sea $f(x) = x^2 - 3x + 2$. Entonces $f(1) = 0$ y por lo tanto, 1 es una raíz de f. Además, $f(2) = 0$. Por lo tanto, 2 es también una raíz de f.

Ejemplo. Sean $a, b, c \in \mathbb{R}$ y $f(x) = ax^2 + bx + c$, un polinomio en $\mathbb{R}$. Si $b^2 - 4ac = 0$, entonces el polinomio tiene una raíz real, que es

$$-\frac{b}{2a}.$$

Si $b^2 - 4ac > 0$, entonces el polinomio tiene dos raíces reales distintas que son

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{y} \quad \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

En el caso que $b^2 - 4ac < 0$ el polinomio no tiene raíces reales.

Teorema B.1.2. *Sea f un polinomio de grado $\leqslant n$ y sea c una raíz. Entonces existe un polinomio g de grado $\leqslant n-1$ tal que para todo x se cumple*

$$f(x) = (x - c)g(x).$$

Demostración. Escribamos $f(x)$ en función de las potencias de x:

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Veremos a continuación que f puede también escribirse en potencias de $x - c$: escribamos

$$x = (x - c) + c,$$

luego

$$f(x) = a_n((x-c)+c)^n + a_{n-1}((x-c)+c)^{n-1} + \cdots + a_1((x-c)+c) + a_0.$$

Expandiendo las potencias de los binomios $((x-c)+c)^k$ $(1 \leqslant k \leqslant n)$, obtenemos

$$f(x) = b_n(x-c)^n + b_{n-1}(x-c)^{n-1} + \cdots + b_1(x-c) + b_0,$$

para ciertos $b_0, b_1, \ldots, b_n \in \mathbb{K}$. Como $f(c) = 0$, entonces $0 = f(c) = b_0$, luego

$$\begin{aligned} f(x) &= b_n(x-c)^n + b_{n-1}(x-c)^{n-1} + \cdots + b_1(x-c) \\ &= (x-c)(b_n(x-c)^{n-1} + b_{n-1}(x-c)^{n-2} + \cdots + b_1) \\ &= (x-c)g(x), \end{aligned}$$

con $g(x) = b_n(x-c)^{n-1} + b_{n-1}(x-c)^{n-2} + \cdots + b_1$, que es una función polinómica de grado $\leqslant n-1$, y vemos que nuestro teorema está probado. $\square$

El polinomio f es el *polinomio nulo* si $f(x) = 0$ para toda $x \in \mathbb{K}$. Si f es el polinomio nulo, denotamos $f = 0$.

Teorema B.1.3. *Sea f un polinomio de grado $\leqslant n$ tal que*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

y $a_n \neq 0$. Entonces f tiene a lo más n raíces.

Demostración. Lo probaremos haciendo inducción sobre n.

Si $n = 0$, $a_0 \neq 0$, es decir $f(x) = a_0 \neq 0$, que es lo que teníamos que probar (f no tiene raíces).

Sea $n > 0$. Sea c raíz de f. Por el teorema B.1.2,

$$f(x) = (x-c)g(x),$$

con

$$g(x) = b_{n-1}x^{n-1} + \cdots + b_1 x + b_0.$$

Es claro que $b_{n-1} = a_n \neq 0$ y por lo tanto, por hipótesis inductiva, $g(x)$ tiene a lo más $n-1$ raíces. Ahora bien

$$0 = f(x) = (x-c)g(x) \quad \Leftrightarrow \quad x - c = 0 \text{ o } g(x) = 0.$$

Es decir x es raíz de f si y solo si $x = c$ o x es raíz de g. Como g tiene a lo más $n-1$ raíces, f tiene a lo más n raíces. $\square$

Corolario B.1.4. *Sea* f *un polinomio y*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

con $a_n \neq 0$*. Entonces* n *y* $a_0, \ldots, a_n$ *están determinados de forma única.*

Demostración. Debemos ver que si

$$f(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0,$$

con $b_m \neq 0$, entonces $m = n$ y $a_i = b_i$ para $1 \leqslant i \leqslant n$.

Supongamos que $m \leqslant n$ (el caso $n \leqslant m$ es similar), entonces definiendo $b_k = 0$ para $m < k \leqslant n$, tenemos que

$$f(x) = b_n x^n + b_{n-1} x^{n-1} + \cdots + b_1 x + b_0.$$

Por lo tanto si

$$h(x) = (a_n - b_n) x^n + (a_{n-1} - b_{n-1}) x^{n-1} + \cdots + (a_1 - b_1) x + (a_0 - b_0), \tag{B.1.2}$$

tenemos que $h(x) = f(x) - f(x) = 0$, es un polinomio que admite infinitas raíces. Supongamos que algún coeficiente de la expresión (B.1.2) de h sea no nulo. Sea k el máximo coeficiente de la expresión (B.1.2) no nulo. Por el teorema B.1.3, h no puede tener más de k raíces, absurdo pues hemos dicho que tenía infinitas. El absurdo vino de suponer que algún $a_i - b_i$ era no nulo. Por lo tanto $a_i - b_i = 0$ para $0 \leqslant i \leqslant n$, es decir $a_i = b_i$ para $1 \leqslant i \leqslant n$.

□

Este corolario nos dice que la escritura de un polinomio f como

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

con $a_n \neq 0$, es única. Diremos entonces que n es el *grado* de f y lo denotaremos $\operatorname{gr}(f) = n$. En el caso del polinomio 0, el grado no está definido y se usa la convención $\operatorname{gr}(0) = -\infty$.

Diremos también que $a_0, \ldots, a_n$ son los *coeficientes* de f, a_0 es el *término constante* de f y a_n el *coeficiente principal.*

Observemos que si f y g son polinomios con

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \quad \text{y} \quad g(x) = b_n x^n + \cdots + b_1 x + b_0,$$

entonces como $ax^i + bx^i = (a+b)x^i$, tenemos que $f + g$ es un polinomio definido por

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

Por otro lado, debido a que $(ax^i)(bx^j) = abx^{i+j}$, el producto de dos polinomios también es un polinomio. Más precisamente,

Proposición B.1.5. *Sean* f *y* g *polinomios de grado* n *y* m, *respectivamente. Entonces* fg *es un polinomio de grado* $n+m$

Demostración. Sean

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \quad \text{y} \quad g(x) = b_m x^m + \cdots + b_1 x + b_0,$$

con $a_n, b_m \neq 0$. Entonces, debido a que $x^i x^j = x^{i+j}$,

$$(fg)(x) = a_n b_m x^{n+m} + h(x), \tag{B.1.3}$$

con $h(x)$ un polinomio de grado menor a $n+m$. Por lo tanto, el coeficiente principal de fg es $a_n b_m \neq 0$ y, entonces, fg tiene grado $n+m$. □

Ejemplo. Sean $f(x) = 4x^3 - 3x^2 + x + 2$ y $g(x) = x^2 + 1$. Entonces,

$$\begin{aligned}(f+g)(x) &= (4+0)x^3 + (-3+1)x^2 + (1+0)x + (2+1)\\ &= 4x^3 - 2x^2 + x + 3,\end{aligned}$$

y

$$\begin{aligned}(fg)(x) &= (4x^3 - 3x^2 + x + 2)(x^2+1)\\ &= (4x^3 - 3x^2 + x + 2)x^2 + (4x^3 - 3x^2 + x + 2)1\\ &= 4x^5 - 3x^4 + x^3 + 2x^2 + 4x^3 - 3x^2 + x + 2\\ &= 4x^5 - 3x^4 + 5x^3 - x^2 + x + 2\end{aligned}$$

B.2 DIVISIÓN DE POLINOMIOS

Si f y g son polinomios, entonces no necesariamente la función f/g está bien definida en todo punto y puede que tampoco sea un polinomio. Cuando trabajamos con enteros, en cursos anteriores, probamos la existencia del algoritmo de división, más precisamente.

Sean n, d *enteros positivos. Entonces existe un entero* r *tal que* $0 \leqslant r < d$ *un entero* $q \geqslant 0$ *tal que*

$$n = qd + r.$$

Ahora describiremos un procedimiento análogo para polinomios.

Algoritmo de División. *Sean* f *y* g *polinomios distintos de cero. Entonces existen polinomios* q, r *tales que* $\mathrm{gr}(r) < \mathrm{gr}(g)$ *y tales que*

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

A $q(x)$ lo llamamos el *cociente* de la *división polinomial* y a $r(x)$ lo llamamos el *resto* de la división polinomial.

No veremos aquí la demostración del algoritmo de división, basta decir que es muy similar a la demostración del algoritmo de división para números enteros. En los siguientes ejemplos se verá como se calculan el cociente y resto de la división polinomial.

Ejemplo. Sean $f(x) = 4x^3 - 3x^2 + x + 2$ y $g(x) = x^2 + 1$. Para encontrar la división polinomial, debemos multiplicar por un monomio ax^k a $g(x)$ de tal forma que el coeficiente principal de $ax^k g(x)$ sea igual al coeficiente principal de $f(x)$. En este caso, multiplicamos a $g(x)$ por $4x$ y nos queda

$$f(x) = 4xg(x) + r_1(x) = (4x^3 + 4x) + (-3x^2 - 3x + 2)$$

Ahora, con $r_1(x) = -3x^2 - 3x + 2$ hacemos el mismo procedimiento, es decir multiplicamos por -3 a $g(x)$ y vemos que es lo que "falta":

$$r_1(x) = (-3)g(x) + r(x) = (-3x^2 - 3) + (-3x + 5).$$

Como $r(x) = -3x + 5$ tiene grado menor que 2, tenemos que

$$\begin{aligned} f(x) &= 4xg(x) + r_1(x) \\ &= 4xg(x) + (-3)g(x) + r(x) \\ &= (4x - 3)g(x) + r(x). \end{aligned}$$

Es decir,

$$f(x) = q(x)g(x) + r(x),$$

con $q(x) = 4x - 3$ y $r(x) = -3x + 5$.

Observemos que se puede hacer un esquema parecido a la división de números enteros, el cual nos facilita el cálculo:

$$\begin{array}{rrrrl} 4x^3 & -3x^2 & +x & +2 & = (x^2+1)(4x-3) - 3x + 5 \\ -4x^3 & & -4x & & \\ \hline & -3x^2 & -3x & +2 & \\ & 3x^2 & & +3 & \\ \hline & & -3x & +5 & \end{array}$$

Ejemplo. Sean

$$f(x) = 2x^4 - 3x^2 + 1 \quad \text{y} \quad g(x) = x^2 - x + 3.$$

Deseamos encontrar $q(x)$ y $r(x)$ como en el algoritmo de Euclides. Haciendo la división como en el ejercicio anterior:

$$\begin{array}{rrrrrl} 2x^4 & & -3x^2 & & +1 & = (x^2 - x + 3)(2x^2 + 2x - 7) - 13x + 22 \\ -2x^4 & +2x^3 & -6x^2 & & & \\ \hline & 2x^3 & -9x^2 & & & \\ & -2x^3 & +2x^2 & -6x & & \\ \hline & & -7x^2 & -6x & +1 & \\ & & 7x^2 & -7x & +21 & \\ \hline & & & -13x & +22 & \end{array}$$

Es decir $q(x) = 2x^2 + 2x - 7$ y $r(x) = -13x + 22$.

Observemos que el algoritmo de división nos dice que si dividimos un polinomio por uno de grado 1, entonces el resto es una constante (que puede ser 0). Más aún:

Teorema B.2.1 (Teorema del resto)**.** *Sea* f *polinomio y* $c \in \mathbb{K}$*. Entonces, el resto de dividir* f *por* $x - c$ *es* $f(c)$*.*

Demostración. Por el algoritmo de Euclides

$$f(x) = q(x)(x - c) + r,$$

con r de grado < 1, es decir $r \in \mathbb{K}$. Ahora bien

$$f(c) = q(c)(c - c) + r = r,$$

luego $f(c)$ es el resto de dividir f por $x - c$. □

Observar que esto nos da otra prueba del teorema B.1.3: $f(c) = 0$, luego por teorema del resto $f(x) = q(x)(x - c)$.

C

DETERMINANTE

En el apéndice se harán las demostraciones de los resultados correspondientes a la sección de determinantes (sección 2.8).

C.1 DETERMINANTES

Lo primero que veremos será la demostración del teorema 2.8.6. Los tres resultados de ese teorema los demostraremos en forma separada: serán los teoremas C.1.1, C.1.3 y C.1.4.

Teorema C.1.1. *Sea* $A \in M_n(\mathbb{K})$ *y sea* $c \in \mathbb{K}$ *y* B *la matriz que se obtiene de* A *multiplicando la fila* r *por* c*, es decir* $A \xrightarrow{cF_r} B$*, entonces* $\det B = c \det A$.

Demostración. Si multiplicamos la fila r por c obtenemos

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \vdots \\ ca_{r1} & ca_{r2} & \cdots & ca_{rn} \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \cdots & a_{nn} \end{bmatrix}.$$

Observemos que al hacer el desarrollo por la primera columna obtenemos

$$|B| = \sum_{i=1}^{r-1} a_{i1} C^B_{1i} + ca_{r1} C^B_{r1} + \sum_{i=r+1}^{n} a_{i1} C^B_{1i}.$$

Ahora bien, si $i \neq r$, la matriz $B(i|1)$ es la matriz $A(i|1)$ con una fila multiplicada por c, luego $|B(i|1)| = c|A(i|1)|$ y, en consecuencia $C^B_{i1} = c\, C^A_{i1}$. Además, $B(r|1) = A(r|1)$, luego $C^B_{r1} = C^A_{r1}$. Por lo tanto, reemplazando en la ecuación anterior C^B_{i1} por $c\, C^A_{i1}$ si $i \neq r$ y C^B_{r1} por C^A_{r1}, obtenemos

$$|B| = \sum_{i=1}^{r-1} a_{i1} c\, C^A_{1i} + ca_{r1} C^A_{r1} + \sum_{i=r+1}^{n} a_{i1} c C^A_{1i} = c|A|.$$

□

Lema C.1.2. *Sean* A, B, C *matrices* $n \times n$ *tal que*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

y

$$C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{r1} + b_{r1} & a_{r2} + b_{r2} & \cdots & a_{rn} + b_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Es decir B *es igual a* A *pero con la fila* r *cambiada y* C *es como* A *y* B *excepto en la fila* r *donde cada coeficiente el la suma del de* A *y* B *correspondiente. Entonces* $\det(C) = \det(A) + \det(B)$.

Demostración. Se hará por inducción en n. Para $n = 1$, del resultado se reduce a probar que $\det[a + b] = \det[a] + \det[b]$, lo cual es trivial, pues el determinante en matrices 1×1 es la identidad.

Primero consideremos el caso $r = 1$. En este caso tenemos que $A(1|1) = B(1|1) = C(1|1)$, pues en la única fila que difieren las matrices es en la primera. Además, si $i > 1$, $A(i|1)$, $B(i|1)$ y $C(i|1)$ son iguales, excepto que difieren en la primera fila donde los coeficientes de $C(i|1)$ son la suma de los de $A(i|1)$ y $B(i|1)$, entonces, por hipótesis inductiva, $\det C(i|1) = \det A(i|1) + \det B(i|1)$. Concluyendo, tenemos que

$$\begin{aligned} &\det A(1|1) = \det B(1|1) = \det C(1|1), \\ &\det C(i|1) = \det A(i|1) + \det B(i|1), \qquad i > 1 \end{aligned},$$

lo cual implica que

$$\begin{aligned} &C^C_{11} = C^A_{11} = C^B_{11}, \\ &C^C_{i1} = C^A_{i1} + C^B_{i1}, \qquad i > 1 \end{aligned}.$$

Luego

$$\begin{aligned} \det C &= (a_{11} + b_{11})C^C_{11} + \sum_{i=2}^{n} a_{i1}C^C_{i1} \\ &= a_{11}C^C_{11} + b_{11}C^C_{11} + \sum_{i=2}^{n} a_{i1}(C^A_{i1} + C^B_{i1}) \\ &= a_{11}C^A_{11} + b_{11}C^B_{11} + \sum_{i=2}^{n} a_{i1}(C^A_{i1} + C^B_{i1}) \\ &= a_{11}C^A_{11} + \sum_{i=2}^{n} a_{i1}C^A_{i1} + b_{11}C^B_{11} + \sum_{i=2}^{n} a_{i1}C^B_{i1} \\ &= \det A + \det B. \end{aligned}$$

El caso $r > 1$ se demuestra de manera similar o, si se prefiere, puede usarse el teorema C.1.4, observando que la permutación entre la fila 1 y la fila r cambia el signo del determinante. □

Teorema C.1.3. *Sea* $A \in M_n(\mathbb{K})$*. Sea* $c \in \mathbb{K}$ *y* B *la matriz que se obtiene de* A *sumando a la fila* r *la fila* s *multiplicada por* c*, es decir* $A \overset{F_r + cF_s}{\longrightarrow} B$*, entonces* $\det B = \det A$*.*

Demostración. A y B difieren solo en la fila r, donde los coeficientes de B son los los de A más c por los de la fila s. Luego si

$$A = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_r \\ \vdots \\ F_n \end{bmatrix}, \qquad B = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_r + cF_s \\ \vdots \\ F_n \end{bmatrix}, \qquad A' = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ cF_s \\ \vdots \\ F_n \end{bmatrix},$$

el lema anterior nos dice que

$$\det B = \det A + \det A'. \tag{C.1.1}$$

Ahora bien, por teorema C.1.1,

$$\det A' = c \begin{vmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_s \\ \vdots \\ F_n \end{vmatrix},$$

y este último determinante es cero, debido a que la matriz tiene dos filas iguales. Luego, $\det B = \det A$. □

Teorema C.1.4. *Sea* $A \in M_n(\mathbb{K})$ *y sean* $1 \leqslant r, s \leqslant n$*. Sea* B *la matriz que se obtiene de* A *permutando la fila* r *con la fila* s*, es decir* $A \overset{F_r \leftrightarrow F_s}{\longrightarrow} B$*, entonces* $\det B = -\det A$*.*

Demostración. Primero probaremos el teorema bajo el supuesto de que la fila 1 es permutada con la fila k, para $k > 1$. Esto será suficiente para probar el teorema, puesto que intercambiar las filas k y k_0 es equivalente a realizar tres permutaciones de filas: primero intercambiamos las filas 1 y k, luego las filas 1 y k_0, y finalmente intercambiando las filas 1 y k. Cada permutación cambia el signo del determinante y al ser tres permutaciones, el intercambio de la fila k con la fila k_0 cambia el signo.

La prueba es por inducción en n. El caso base $n = 1$ es completamente trivial. (O, si lo prefiere, puede tomar $n = 2$ como el caso base, y el teorema es fácilmente probado usando la fórmula para el determinante de una matriz 2×2). Las definiciones de los determinantes de A y B son:

$$\det(A) = \sum_{i=1}^{n} a_{i1} C_{i1}^{A} \quad \text{y} \quad \det(B) = \sum_{i=1}^{n} b_{i1} C_{i1}^{B}.$$

Supongamos primero que $i \neq 1, k$. En este caso, está claro que $A(i|1)$ y $B(i|1)$ son iguales, excepto que dos filas se intercambian. Por lo tanto, por hipótesis inductiva $C_{i1}^{A} = -C_{i1}^{B}$. Ya que también $a_{i1} = b_{i1}$, tenemos entonces que

$$a_{i1} C_{i1}^{A} = -b_{i1} C_{i1}^{B}, \quad \text{para } i \neq 1, k. \tag{C.1.2}$$

Queda por considerar los términos $i = 1$ y $i = k$. Nosotros afirmamos que

$$-a_{k1} C_{k1}^{A} = b_{11} C_{11}^{B} \quad \text{y} \quad -a_{11} C_{11}^{A} = b_{k1} C_{k1}^{B}. \tag{C.1.3}$$

Si probamos esto, entonces

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{n} a_{i1} C_{i1}^{A} \\
&= a_{11} C_{11}^{A} + \sum_{i=2}^{k-1} a_{i1} C_{i1}^{A} + a_{k1} C_{k1}^{A} + \sum_{i=k+1}^{n} a_{i1} C_{i1}^{A} \qquad \text{por (C.1.2) y (C.1.3)} \\
&= -b_{k1} C_{k1}^{B} - \sum_{i=2}^{k-1} b_{i1} C_{i1}^{B} - b_{11} C_{11}^{B} - \sum_{i=k+1}^{n} b_{i1} C_{i1}^{B} \\
&= -\sum_{i=1}^{n} b_{i1} C_{i1}^{B} = -\det(B).
\end{aligned}$$

Luego el teorema está probado. Por lo tanto debemos probar (C.1.3). Por simetría, basta probar la primera identidad de (C.1.3), es decir que $a_{k1} C_{k1}^{A} = -b_{11} C_{11}^{B}$.

Para esto, primero debemos observar que $a_{k1} = b_{11}$, por lo tanto sólo hace falta probar que $-C_{k1}^{A} = C_{11}^{B}$. En segundo lugar, debemos tener en cuenta que $B(1|1)$ se obtiene de $A(k|1)$ reordenando las filas $1, 2, \ldots, k-1$ de $A(k|1)$ en el orden $2, 3, \ldots, k-1, 1$. Este reordenamiento puede hacerse permutando la fila 1 con la fila 2, luego permutando esa fila con la fila 3, etc., terminando con una permutación con la fila $k-1$. Esto es un total de $k-2$ permutaciones de fila. Asi que, por hipótesis inductiva,

$$\det(B(1|1)) = (-1)^{k-2} \det(A(k|1)) = (-1)^{k} \det(A(k|1)) = -(-1)^{k+1} \det(A(k|1)),$$

es decir $C_{11}^{B} = -C_{k1}^{A}$. Esto completa la demostración del teorema. □

Observación. Del resultado anterior se deduce fácilmente que si una matriz tiene dos filas iguales entonces su determinante es 0. Esto se debe a que, intercambiando las dos filas iguales obtenemos la misma matriz, pero calculando el determinante con el teorema anterior vemos que cambia de signo y el único número en $\mathbb{K}$ que es igual a su opuesto es el 0.

Corolario C.1.5. *Consideremos matrices elementales en* $\mathbb{K}^{n \times n}$.

(1) *Sea* E *la matriz elemental que se obtiene multiplicando por* $c \neq 0$ *la matriz* Id_n. *Entonces* $\det(E) = c$.

(2) *Sea* E *la matriz elemental que se obtiene a partir de* Id_n *sumando* c *veces* F_r *a* F_s *(*$r \neq s$*). Entonces* $\det(E) = 1$.

(3) *Sea* E *la matriz elemental que se obtiene a partir de* Id_n *de permutando la* F_r *con* F_s *(*$r \neq s$*). Entonces* $\det(E) = -1$.

Demostración. Se demuestra trivialmente considerando que en todos los casos $E = e(\mathrm{Id}_n)$ donde e es una operación elemental por fila, considerando que $\det(\mathrm{Id}_n) = 1$ y aplicando los teoremas C.1.1, C.1.3 y C.1.4, según corresponda. □

A continuación veremos que el determinante del producto de matrices es el producto de los determinantes de las matrices.

Teorema C.1.6. *Sea* $A \in M_n(\mathbb{K})$ *y* E *una matriz elemental* $n \times n$. *Entonces*

$$\det(EA) = \det E \det A. \tag{C.1.4}$$

Demostración. En todos los casos $EA = e(A)$ donde e es una operación elemental por fila (teorema 2.6.2).

(1) Si $c \neq 0$, y E es la matriz elemental que se obtiene de multiplicar por c la fila r de Id_n, luego

$$\det(EA) = \det(e(A)) \overset{\text{Teor. C.1.1}}{=} c \cdot \det(A) \overset{\text{Cor. C.1.5.(1)}}{=} \det(E)\det(A).$$

(2) Si E es la matriz elemental que se obtiene de sumar a la fila r de Id_n la fila s multiplicada por c, entonces $\det E = 1$. Por otro lado $\det(EA) = \det(A)$, por lo tanto $\det(EA) = \det(E)\det(A)$.

(3) Finalmente, si E es la matriz elemental que se obtiene de intercambiar la fila r por la fila s de Id_n, entonces $\det E = -1$. Por otro lado $\det(EA) = -\det(A)$, por lo tanto $\det(EA) = \det(E)\det(A)$. □

Corolario C.1.7. *Sea* $A \in M_n(\mathbb{K})$ *y* $E_1, \ldots, E_k$ *matrices elementales* $n \times n$. *Entonces*

$$\det(E_k E_{k-1} \ldots E_1 A) = \det(E_k)\det(E_{k-1}) \ldots \det(E_1)\det(A).$$

Demostración. Por la aplicación reiterada del teorema C.1.6 tenemos,

$$\begin{aligned} \det(E_k E_{k-1} \ldots E_1 A) &= \det(E_k)\det(E_{k-1} \ldots E_1 A) \\ &= \det(E_k)\det(E_{k-1})\det(E_{k-2} \ldots E_1 A) \\ &\;\;\vdots \\ &= \det(E_k)\det(E_{k-1})\det(E_{k-2}) \ldots \det(E_1)\det(A). \end{aligned}$$

□

Teorema C.1.8. *$A \in \mathbb{K}^{n\times n}$ es invertible si y solo si $\det(A) \neq 0$.*

Demostración. ($\Rightarrow$) A invertible, luego por el teorema 2.7.6, A es producto de matrices elementales, es decir $A = E_1E_2\cdots E_k$ donde $E_1, E_2, \ldots, E_k$ son matrices elementales.

Por el corolario anterior, $\det(A) = \det(E_1)\det(E_2)\ldots\det(E_k)$. Como el determinante de matrices elementales es distinto de cero,

$$\det(A) = \det(E_1)\det(E_2)\ldots\det(E_k) \neq 0.$$

($\Leftarrow$) Sean $E_1, E_2, \ldots, E_k$ matrices elementales tales que $R = E_1E_2\cdots E_kA$ y R es MERF. Luego,

$$\det(R) = \det(E_1)\det(E_2)\cdots\det(E_k)\det(A).$$

Como los determinantes de matrices elementales son no nulos

$$\frac{\det(R)}{\det(E_1)\det(E_2)\cdots\det(E_k)} = \det(A). \tag{*}$$

Supongamos que R no es la identidad. Entonces, por el corolario 2.8.5, $\det(R) = 0$, por lo tanto, $\det(A) = 0$, lo cual contradice la hipótesis y llegamos a un absurdo.

Esto implica que $R = \mathrm{Id}_n$ y en consecuencia A es equivalente por filas a Id_n y por lo tanto invertible. □

Teorema C.1.9. *Sean $A, B \in M_n(\mathbb{K})$, entonces*

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

Demostración. Separemos la prueba en dos casos (a) A es invertible y (b) A no es invertible.

(a) Si A invertible, entonces $A = E_1\cdots E_k$ producto de matrices elementales. Por lo tanto $AB = E_1\cdots E_kB$, luego por el corolario C.1.7 $\det(AB) = \det(E_1)\cdots\det(E_k)\det(B) = \det(A)\det(B)$.

(b) Si A no invertible, entonces A es equivalente por filas a una MERF R con la última fila nula. Es decir $R = E_1\cdots E_kA$ y R tiene la última fila nula, por lo tanto $A = E_k^{-1}E_{k-1}^{-1}\ldots E_1^{-1}R$.

Como R tiene la última fila nula, no es difícil ver que RB tiene tiene también la última fila nula y por lo tanto $\det(RB) = 0$. Luego

$$\det(AB) = \det(E_k^{-1})\ldots\det(E_1^{-1})\det(RB) = 0.$$

Como $\det(A) = 0$, tenemos también qure

$$\det(A)\det(B) = 0.$$

□

Haremos ahora la demostración del teorema 2.8.13.

Teorema C.1.10. *Sea* E *matriz elemental, entonces* E^t *es matriz elemental del mismo tipo y* $\det(E) = \det(E^t)$.

Demostración. Si $c \neq 0$ y E es la matriz elemental que se obtiene de multiplicar por c la fila r de Id_n, es claro que $E^t = E$ y por lo tanto $\det(E) = \det(E^t)$.

Si E es la matriz elemental que se obtiene de sumar a la fila r de Id_n la fila s multiplicada por $c \in \mathbb{K}$, entonces E^t es la matriz elemental que se obtiene de sumar a la fila s de Id_n la fila r multiplicada por c. Luego, $\det(E) = \det(E^t) = 1$.

Finalmente, si E es la matriz elemental que se obtiene de intercambiar la fila r por la fila s de Id_n,entonces $E^t = E$ y por lo tanto $\det(E) = \det(E^t)$. □

Teorema C.1.11. *Sea* $A \in M_n(\mathbb{K})$, *entonces* $\det(A) = \det(A^t)$

Demostración. Si A es invertible, entonces $A = E_k E_{k-1} \ldots E_1$ con E_i elemental, por lo tanto $\det(A) = \det(E_k)\det(E_{k-1}) \ldots \det(E_1)$. Luego,

$$\det(A^t) = \det(E_1^t \ldots E_k^t) = \det(E_1^t) \ldots \det(E_k^t) = \det(E_1) \ldots \det(E_k) = \det(A).$$

Si A no es invertible, entonces A^t no es invertible y en ese caso $\det(A) = \det(A^t) = 0$. □

Finalmente, demostremos el teorema 2.8.17.

Teorema C.1.12. *El determinante de una matriz* A *de orden* $n \times n$ *puede ser calculado por la expansión de los cofactores en cualquier columna o cualquier fila. Más específicamente,*

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

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

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

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

Demostración. (1) Primero hagamos la demostración para $j = 2$, es decir para el desarrollo por la segunda columna. Escribamos A en función de sus columnas, es decir

$$A = \begin{bmatrix} C_1 & C_2 & C_3 & \cdots & C_n \end{bmatrix},$$

donde C_k es la columna k de A. Sea $B = [b_{ij}]$ la matriz definida por

$$B = \begin{bmatrix} C_2 & C_1 & C_3 & \cdots & C_n \end{bmatrix}.$$

Entonces, $\det(B) = -\det(A)$. Por otro lado, por la definición de determinante,

$$\begin{aligned}
\det(B) &= \sum_{i=1}^{n} b_{i1} C^B_{i1} \\
&= \sum_{i=1}^{n} b_{i1} (-1)^{i+1} B(i|1) \\
&= \sum_{i=1}^{n} a_{i2} (-1)^{i+1} B(i|1).
\end{aligned}$$

Ahora bien, es claro que $B(i|1) = A(i|2)$, por lo tanto

$$\det(B) = \sum_{i=1}^{n} a_{i2}(-1)^{i+1} A(i|2) = -\sum_{i=1}^{n} a_{i2} C_{i2}.$$

Es decir, $\det(A) = -\det(B) = \sum_{i=1}^{n} a_{i2} C_{i2}$.

El caso $j > 2$ se demuestra de forma similar: si B es la matriz

$$B = \begin{bmatrix} C_j & C_1 & C_2 & \cdots & C_{j-1} & C_{j+1} & \cdots & C_n \end{bmatrix}.$$

entonces $\det(B) = (-1)^{j-1} \det(A)$, pues son necesarios $j-1$ permutaciones para recuperar la matriz A (es decir, llevar la columna j a su lugar). Como $B(i|1) = A(i|j)$, desarrollando por la primera columna el determinante de B obtenemos el resultado.

(2) Observemos primero que $A^t(j|i) = A(i|j)^t$, por lo tanto, si calculamos $\det(A^t)$ por desarrollo por columna i, obtenemos

$$\begin{aligned}
\det A = \det(A^t) &= \sum_{j=1}^{n} [A^t]_{ji} (-1)^{i+j} \det(A^t(j|i)) \\
&= \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det(A(i|j)^t) \\
&= \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det(A(i|j)).
\end{aligned}$$

$\square$

C.2 REGLA DE CRAMER

Veremos ahora que la inversa de una matriz invertible se puede escribir en términos de determinantes de algunas matrices relacionadas y esto, junto a otros resultados, nos permitirá resolver ecuaciones lineales con n-variables y n-incógnitas cuya matriz asociada es invertible.

Teorema C.2.1. *Sea* A *matriz* $n \times n$, *entonces*

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} A = \begin{array}{c} \begin{bmatrix} 0 & \cdots & 0 & \det A & 0 & \cdots & 0 \end{bmatrix} . \\ \uparrow i \end{array}$$

Es decir, la matriz fila formada por los cofactores correspondientes a la columna i *multiplicada por la matriz* A *es igual a la matriz fila con valor* $\det A$ *en la posición* i *y 0 en las otras posiciones.*

Demostración. Si C_j denota la matriz formada por la columna j de A debemos probar que

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = \sum_{k=1}^{n} a_{kj} C_{ki} = \begin{cases} \det(A) & \text{si } j = i \\ 0 & \text{si } j \neq i. \end{cases}$$

Ahora bien,

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_i = \sum_{j=1}^{n} C_{ji} a_{ji},$$

y esto último no es más que el cálculo del determinante por desarrollo de la columna i, es decir, es igual a $\det(A)$.

Para ver el caso $i \neq j$, primero observemos que si

$$B = \begin{bmatrix} C_1 & C_2 & \cdots & C_j & \cdots & C_j & \cdots & C_{n-1} & C_n \end{bmatrix},$$
$$\qquad\qquad\uparrow i \qquad\quad \uparrow j$$

es decir, B es la matriz A donde reemplazamos la columna i por la columna j, entonces como B tiene dos columnas iguales, $\det(B) = 0$. Por lo tanto, si calculamos el determinante de B por el desarrollo en la columna i, obtenemos

$$0 = \det(B) = \sum_{k=1}^{n} a_{kj} C_{ki}. \tag{C.2.1}$$

Por otro lado,

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = \sum_{k=1}^{n} C_{ki} a_{kj},$$

luego, por la ecuación (C.2.1) tenemos que

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = 0$$

si $i \neq j$.

□

Definición C.2.2. Sea A matriz $n \times n$, la *matriz de cofactores* es la matriz cuyo coeficiente ij vale C_{ij}. La matriz de cofactores de A se denota $\operatorname{cof}(A)$. La *matriz adjunta de A* es $\operatorname{adj}(A) = \operatorname{cof}(A)^t$.

Teorema C.2.3. *Sea* A *matriz* $n \times n$, *entonces*

$$\operatorname{adj}(A).A = \det(A)\,\mathrm{Id}_n\,.$$

Demostración. Observar que la fila i de $\operatorname{adj}(A)$ es $\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni}\end{bmatrix}$. Por lo tanto, la fila i de $\operatorname{adj}(A).A$ es

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni}\end{bmatrix} A,$$

que por el teorema C.2.1 es una matriz fila con el valor $\det A$ en la posición i y todos los demás coeficientes iguales a 0. Luego

$$\operatorname{adj}(A).A = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn}\end{bmatrix}.A = \begin{bmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det A\end{bmatrix} = \det(A)\,\mathrm{Id}_n$$

□

Corolario C.2.4. *Si* A *es invertible, entonces*

$$A^{-1} = \frac{1}{\det A}\operatorname{adj} A.$$

Demostración.

$$\frac{1}{\det A}\operatorname{adj} A.A = \frac{1}{\det A}\det A\,\mathrm{Id}_n = \mathrm{Id}_n\,.$$

□

Teorema C.2.5 (Regla de Cramer). *Sea* $AX = Y$ *un sistema de ecuaciones tal que* $A \in M_n(\mathbb{K})$ *es invertible. Entonces, el sistema tiene una única solución* $(x_1, \ldots, , x_n)$ *con*

$$x_j = \frac{\det A_j}{\det A}, \qquad j = 1, \ldots, n,$$

donde A_j *es la matriz* $n \times n$ *que se obtiene de* A *remplazando la columna* j *de* A *por* Y.

Demostración. Haremos la demostración para matrices 3×3. La demostración en el caso general es completamente análoga.

Como A es invertible, existe A^{-1} y multiplicamos la ecuación a izquierda por A^{-1} y obtenemos que $A^{-1}AX = A^{-1}Y$, es decir $X = A^{-1}Y$ y esta es la única solución. Luego

$$\begin{aligned} A^{-1}Y &= \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33}\end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3\end{bmatrix} \\ &= \frac{1}{\det A}\begin{bmatrix} y_1C_{11} + y_2C_{21} + y_3C_{31} \\ y_1C_{12} + y_2C_{22} + y_3C_{32} \\ y_1C_{13} + y_2C_{23} + y_3C_{33}\end{bmatrix} \qquad (*) \end{aligned}$$

Ahora bien, $y_1C_{11} + y_2C_{21} + y_3C_{31}$ es el cálculo de determinante por desarrollo de la primera columna de la matriz

$$\begin{bmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_1 & a_{32} & a_{33} \end{bmatrix},$$

y, de forma análoga, el segundo y tercer coeficiente de la matriz $(*)$ son el determinante de las matrices 3×3 que se obtienen de A remplazando la columna 2 y 3, respectivamente, de A por Y. Es decir

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A^{-1}Y = \frac{1}{\det A}\begin{bmatrix} \det A_1 \\ \det A_2 \\ \det A_3 \end{bmatrix} = \begin{bmatrix} \frac{\det A_1}{\det A} \\ \frac{\det A_2}{\det A} \\ \frac{\det A_3}{\det A} \end{bmatrix},$$

luego $x_j = \dfrac{\det A_j}{\det A}$ para $j = 1, 2, 3$. □

Ejemplo. Resolvamos usando la regla de Cramer el siguiente sistema:

$$\begin{aligned} x_1 + x_2 - x_3 &= 6 \\ 3x_1 - 2x_2 + x_3 &= -5 \\ x_1 + 3x_2 - 2x_3 &= 14. \end{aligned}$$

La matriz asociada al sistema es

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

Luego

$$A_1 = \begin{bmatrix} 6 & 1 & -1 \\ -5 & -2 & 1 \\ 14 & 3 & -2 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 6 & -1 \\ 3 & -5 & 1 \\ 1 & 14 & -2 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 1 & 6 \\ 3 & -2 & -5 \\ 1 & 3 & 14 \end{bmatrix},$$

y

$$\det A = -3, \qquad \det A_1 = -3, \qquad \det A_2 = -9, \qquad \det A_3 = 6.$$

Por lo tanto,

$$\begin{aligned} x_1 &= \frac{\det A_1}{\det A} = \frac{-3}{-3} = 1 \\ x_2 &= \frac{\det A_2}{\det A} = \frac{-9}{-3} = 3 \\ x_3 &= \frac{\det A_3}{\det A} = \frac{6}{-3} = -2. \end{aligned}$$

Observación. En general, la regla de Cramer no es utilizada para resolver sistemas de ecuaciones. El método de Gauss, además de ser mucho más rápido, tiene el beneficio de ser un método que abarca más casos y nos da más información que la regla de Cramer. Por ejemplo, el método de Gauss también funciona para matrices no invertibles, mientras que la regla de Cramer no. Más aún, el método de Gauss se puede utilizar para resolver sistemas con matrices no cuadradas (es decir, sistemas de ecuaciones con números no iguales de variables y ecuaciones), mientras que la regla de Cramer no.

Con respecto a la velocidad de resolución, para matrices $n \times n$ el método de Gauss requiere "alrededor de n^3 operaciones", mientras que la regla de Cramer requiere "alrededor de n^4 operaciones", haciendo el primer método n veces más rápido que el segundo. En lenguaje técnico, el método de Gauss tiene complejidad $O(n^3)$ y la regla de Cramer complejidad $O(n^4)$ (ver https://es.wikipedia.org/wiki/Eficiencia_Algorítmica)