Álgebra Lineal

Grado en Ingeniería del Software/Informática

Facultad de Informática, UCM

Álvaro Martínez Pérez

Índice general

Capítulo 1

Conceptos preliminares

1.1. La base de las matemáticas.

Vamos a empezar por introducir algunos de los términos básicos del lenguaje matemático.

Lo primero que necesitamos son las **definiciones**. No se pueden hacer matemáticas si existe ambigüedad en el lenguaje.

Definicion: *Es un enunciado o proposición que expone un concepto o determina un objeto de forma única y precisa.*

Por ejemplo, vamos a suponer que ya conocemos la definición de triángulo y la definición de ángulo recto. Entonces podríamos definir a partir de éstas un objeto nuevo:

Definición 1.1.1. *Se llama* triángulo rectángulo *a aquel en el que uno de sus ángulos es recto.*

Una vez definido (y fijado) este concepto podemos seguir construyendo a partir de ahí: podemos definir lo que son los catetos, la hipotenusa y así construir, sin ninguna ambigüedad, todos los conceptos que vamos a necesitar.

Una vez que tenemos los conceptos podemos empezar a construir matemáticas con ellos. Lo primero que necesitamos es un punto de partida, unas piezas básicas que van a ser nuestras “verdades iniciales” y se llaman **axiomas**. Es el único momento en matemáticas en el que vamos a asumir algo como “cierto” sin más.

Axioma: *Es una proposición verdadera que se admite sin demostración.*

Los axiomas siempre son lo más simples e intuitivamente evidentes que sea posible y siempre se trata de tener el menor número posible de ellos. El objetivo en matemáticas es construir todo el conocimiento sacando conclusiones a partir de los axiomas.

Un ejemplo de axioma sería:

Ejemplo 1.1.2. *Dados dos puntos diferentes se puede trazar una única línea recta que los une.*

Este tipo de afirmaciones son la base sobre la que se sustentan todas las matemáticas. No vamos a entrar en esto por ser un tema algo más complejo pero debemos decir que no son verdades absolutas: si cambiamos los axiomas tenemos otras teorías igual de válidas, sólo son un punto de partida que aceptamos por acuerdo.

Ya tenemos los conceptos y un punto de partida a partir del cual empezar a extraer conclusiones. Veamos cuál es el proceso.

1.2. Implicación, teorema y demostración.

Implicación: *Es una proposición que establece que si la proposición A es verdadera entonces la proposición B también lo es. Se denota $A \Rightarrow B$.* Suele expresarse con una expresión del tipo "Si A, entonces B". En este caso diremos que A es una *condición suficiente* para B (porque es suficiente que se verifique A para saber que se verifica B) y que B es una *condición necesaria* para A (porque no puede verificarse A si no se verifica B porque $A \Rightarrow B$ y habría una contradicción).

Ejemplo 1.2.1. *Si un número n es múltiplo de 4 entonces n es múltiplo de 2. Dicho de otro modo: n es múltiplo de $4 \Rightarrow n$ es múltiplo de 2.*

Si la implicación no es verdadera se denota $A \not\Rightarrow B$.

Ejemplo 1.2.2. *n es múltiplo de $2 \not\Rightarrow n$ es múltiplo de 4.*

Es importante notar que para que una implicación no sea verdadera basta con que exista una única excepción.

Si $A \Rightarrow B$ y $B \Rightarrow A$ decimos que "A es condición necesaria y suficiente para B" o "A si y sólo si B". Se denota $A \Leftrightarrow B$.

Observación 1.2.3. *Dada una implicación $A \Rightarrow B$, si no se cumple B, entonces no se puede cumplir A. Por ejemplo, en la frase: si llueve, entonces la acera se moja, si la acera está seca, significa que no ha llovido. Nótese, sin embargo, que el que la acera está mojada, no implica necesariamente que haya llovido. Puede haberse mojado de cualquier otra manera y la implicación sigue siendo cierta.*

Ejercicio 1.2.4. *En la expresión: Yo iré al cine sólo si tú vas, ¿cuál es la implicación? Nótese que esta expresión no es equivalente a: "si tú vas, yo voy".*

Una **proposición** o un **teorema** son enunciados de la forma $A \Rightarrow B$ donde la proposición A se llama *hipótesis* y la proposición B se llama *tesis.*

Por ejemplo, y ya que conocemos las definiciones de triángulo rectángulo, catetos e hipotenusa, podemos enunciar el Teorema de Pitágoras.

Teorema 1.2.5. *Si un triángulo es rectángulo entonces la suma de los cuadrados de los catetos es igual al cuadrado de la hipotenusa.*

En este caso, la hipótesis es que el triángulo es rectángulo y la tesis que la suma de los cuadrados de los catetos es igual al cuadrado de la hipotenusa.

Por supuesto, en matemáticas cualquier afirmación, cualquier enunciado que no sea un axioma, debe ser *demostrado.*

Una **demostración** es una sucesión de pasos lógicos de forma que, tomando como verdadera una hipótesis, aseguran la veracidad de una tesis.

Por ejemplo:

Proposición 1.2.6. *Dado un número elevado al cuadrado si lo dividimos por 4 el resto sólo puede ser 0 ó 1.*

Demostración. Consideramos un número n. Si n es par entonces $n = 2k$ para algún número k. Si n es impar entonces $n = 2k+1$ para algún k. En el primer caso, $n^2 = (2k)^2 = 2^2k^2 = 4k^2$. Si dividimos $4k^2/4 = k^2$ y el resto es 0. En el segundo caso, $n^2 = (2k+1)^2 = (2k)^2 + 4k + 1 = 4k^2 + 4k + 1$. Al dividir $(4k^2 + 4k + 1)/4$ el cociente es $k^2 + k$ y el resto es 1. □

Siguiendo el razonamiento paso a paso podemos tener la certeza de que el enunciado es cierto. Éste es el modo de proceder en matemáticas. Ahora bien, debemos tener mucho cuidado a la hora de hacer las implicaciones porque a partir de una implicación falsa, podemos demostrar cualquier cosa.

Ejemplo 1.2.7. *La siguiente cadena de implicaciones contiene en realidad dos errores. ¿Cuáles?*

Sean a, b dos números reales tales que $a = b$. Entonces, $a = b \Rightarrow a^2 = ab \Rightarrow a^2 - b^2 = ab - b^2 \Rightarrow (a+b)(a-b) = b(a-b) \Rightarrow a+b = b \Rightarrow b+b = b \Rightarrow 2b = b \Rightarrow 2 = 1$ *¡lo cual es una contradicción!*

1.3. Conjuntos y operaciones con conjuntos.

La definición de conjunto es bastante técnica así que nos conformaremos con una idea intuitiva. Entendemos como **conjunto**, "una colección de elementos".

Veamos algunos ejemplos:

- $A = \{a, b, c\}$ es un conjunto con tres elementos: a, b y c.
- $\mathbb{N}$ es el conjunto de los números naturales: 1, 2, 3...
- $\mathbb{Z}$ es el conjunto de los números enteros: ... -3, -2, -1, 0, 1, 2, 3...
- $\mathbb{R}$ es el conjunto de los números reales.
- $\emptyset$ es el conjunto vacío, es decir, el conjunto que no contiene ningún elemento.
- Podemos definir el conjunto a través de una propiedad como "el conjunto de todos los números pares".

1.4. Lenguaje matemático

. Una de sus características fundamentales es la precisión: no puede haber ambiegüedad alguna en matemáticas. La otra, es que se escribe con la ayuda de una serie de símbolos (cuantificadores, símbolos de relación...) para que las expresiones sean concisas, rápidas de leer y escribir y claras de interpretar. Por supuesto, requiere algún tiempo (y práctica) acostumbrarse a dicho lenguaje. El dominio de este lenguaje es indispensable para leer matemáticas en cualquier libro o tratado y para trabajar con ellas.

Un elemento a que pertenece a un conjunto A se denota $a \in A$. Normalmente se utilizan letras minúsculas para los alementos y mayúsculas para los conjuntos. Por ejemplo, $-3 \in \mathbb{Z}$. Esto, que literalmente dice "-3 pertenece al conjunto de los enteros", habría que leerlo como "-3 es un número entero". Si queremos escribir: "tomamos un número real cualquiera", la forma de escribir esto en lenguaje matemático sería: "sea $a \in \mathbb{R}$" (o cualquier otra letra minúscula en lugar de a).

Si un elemento a no pertenece a un conjunto A se denota $a \notin A$. Por ejemplo, $-3 \notin \mathbb{N}$.

Si hablamos de relaciones entre conjuntos A y B, decimos que A está **incluído** en B si todo elemento de A está en B. Se denota $A \subset B$. Así, por ejemplo, $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{R}$.

La **intersección** de A y B, denotado $A \cap B$ es el conjunto de elementos que pertenecen simultáneamente a A y B.

La **unión** de A y B, denotado $A \cup B$ es el conjunto de elementos que pertenecen a A o a B. Por ejemplo, si $A = \{a, b, c\}$ y $B = \{b, c, d\}$ entonces $A \cap B = \{b, c\}$ y $A \cup B = \{a, b, c, d\}$.

Pregunta 1.4.1. *Si A son los números naturales múltiplos de 2 y B son los números naturales múltiplos de 3, ¿quiénes serán $A \cap B$ y $A \cup B$?*

Símbolos de orden: menor que "$<$", menor o igual que "$\leq$", mayor que "$>$" y mayor o igual que "$\geq$".

Cuantificadores: Se trata de unos símbolos básicos para el razonamiento lógico y, en particular, matemático. Vamos a usar tres:

- Existe. Se denota: $\exists$. Una expresión del tipo "$\exists\, x$..." quiere decir que "existe al menos un elemento x..."

- Para todo. Se denota: $\forall$. Una expresión como "$\forall\, x, y$..." quiere decir que "para todo x, y..."

- Tal que. Se denota $|$. Una expresión como "$\forall x\ \exists y \,|\, y > x$" se leería "para todo x existe y tal que y es mayor que x".

Ejemplo 1.4.2. *Escribir matemáticamente usando cuantificadores la siguiente proposición: Para todo número natural, existe un entero tal que la suma de ambos es 0.*

Solución: $\forall\, n \in \mathbb{N},\ \exists z \in \mathbb{Z} \,|\, n + z = 0$.

Ejercicio 1.4.3. *Sea P el conjunto de números pares e I el de números impares. Traducir de lenguaje matemático a lenguaje humano las siguientes proposiciones y decidir si son ciertas of falsas:*

- $\forall\, a, b \in I,\ a + b \in P$ *y* $ab \in I$.

- $\forall\, a \in P$ *y* $\forall b \in I,\ ab + a \in P$.

- $\forall\, a, b \in I$ *con* $a < b,\ \exists c \in P \,|\, a < c < b$.

- $\forall\, a, b \in I$ *con* $a < b$ *y* $\forall c \in P,\ a < c < b$.

- $\forall\, a, b \in \mathbb{N},\ \exists c \in I \,|\, a^2 < c < b^2$.

- $\forall\, a, b \in \mathbb{N},\ \exists c \in \mathbb{N} \,|\, a^2 + b^2 = c^2$.

Ejemplo 1.4.4. *Traducir de lenguaje matemático a lenguaje humano, o al revés, las siguientes expresiones:*

- *Para todo número natural, hay un número natural mayor.* $\forall n \in \mathbb{N}, \exists m \in \mathbb{N} \,|\, m > n$.
- *Para todo número par, existe un número impar mayor que él.* $\forall\, n = 2k,\ \exists m = 2k' + 1 \,|\, m > n$.
- $\forall x \in \mathbb{R}, x < x^2 \Leftrightarrow x < 0$ *ó* $x > 1$. *Solución: un número real es menor que su cuadrado si y sólo si el número es negativo o mayor que 1.*
- $\forall\, r_1, r_2 \in \mathbb{R} \,|\, r_1 < r_2,\ \exists q \in \mathbb{Q} \,|\, r_1 < q < r_2$. *Solución: dados dos números reales cualesquiera (con uno mayor que otro), existe un número racional entre ambos.*

Ejercicio 1.4.5. *Supongamos que denotamos por L un conjunto de locos, P un conjunto de periódicos y D un conjunto de días. Un loco se denotaría $l \in L$, un periódico, $p \in P$ y un día sería $d \in D$. Para escribir que el loco d lee el periódico p el día d escribiríamos: "l lee p en d". Traducir a lenguaje matemático las siguientes expresiones:*

- *Algún loco habrá que cada día lea todos los periódicos.*
- *Algún loco habrá cada día que lea todos los periódicos.*
- *Cada loco lee algún periódico cada día.*
- *Cada día hay algún periódico que todo los locos leen.*
- *Todos los días habrá algún loco que lea algún periódico.*
- *Todos los locos leen todos los periódicos cada día.*
- *Hubo un día en el que algún loco leyó algún periódico.*

El lenguaje matemático tiene la ventaja de que es preciso y no caben distintas interpretaciones. En cambio, el lenguaje común depende muchas veces de sobreentendidos. Al traducir el lenguaje común a lenguaje matemático hay que tener siempre mucho cuidado para capturar el significado exacto. El siguiente ejemplo muestra la dificultad que puede tener entender una implicación expresada en lenguaje común.

Ejemplo 1.4.6. *La siguiente proposición "Yo voy sólo si tú vas" es equivalente (matemáticamente hablando, es decir, es equivalente como proposición lógica) a una de las siguientes. ¿Sabrías decir a cuál?*

- *si vas tú, voy yo,*

- *si voy yo, es que tú vas,*
- *tú vas y yo no,*
- *no vas y no voy,*
- *voy y tú no vas,*
- *o vamos los dos o ninguno.*

El **producto cartesiano** de dos conjuntos A y B, se denota $A \times B$, y es el conjunto de pares (x, y) con $x \in A$, $y \in B$.

También se puede considerar el producto cartesiano de n conjuntos X_1, X_2, ..., X_n, $X_1 \times X_2 \times \cdots \times X_n$ que será el conjunto de *n-tuplas* $(x_1, x_2, ..., x_n)$ con $x_i \in X_i$ para cada $i = 1, 2, ..., n$. El ejemplo que más vamos a usar va a ser el espacio usual de 3 dimensiones $\mathbb{R}^3$ ó $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ donde los puntos o elementos del conjunto vienen dados por tres números reales: (a, b, c).

Sean A y B conjuntos. Una **aplicación** f de A en B es una correspondencia en la que a cada elemento de A le corresponde un (y sólo un) elemento de B. Se denota

$$\begin{aligned} f\colon A &\to B \\ a &\mapsto f(a) = b \end{aligned}$$

Ejemplo 1.4.7.

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3 \end{aligned}$$

definida en $\mathbb{R}$. *Esta aplicación toma un número* x *y le asocia otro número que es* x^3.

$$\begin{aligned} f\colon \mathbb{R}\backslash\{3\} &\to \mathbb{R} \\ x &\mapsto \frac{3x+1}{x-3} \end{aligned}$$

definida en $\mathbb{R}\backslash\{3\}$, *es decir, todos los números reales menos el 3 (que haría 0 el denominador).*

Definición 1.4.8. *Una aplicación* $f\colon A \to B$ *se dice que es* ***inyectiva*** *si dados dos elementos distintos de* A *sus imágenes por* f *son distintas, es decir, si* $a_1 \neq a_2$ *entonces* $f(a_1) \neq f(a_2)$.

Esto es lo mismo que decir que f es inyectiva si $f(a_1) = f(a_2)$ implica que $a_1 = a_2$.

Definición 1.4.9. *Una aplicación $f\colon A \to B$ se dice que es* ***sobreyectiva*** *o* ***suprayectiva*** *si todo elemento de B es imagen de algún elemento de A, es decir, si $\forall\, b \in B\ \exists\, a \in A$ tal que $f(a) = b$.*

Por último, si tenemos tres conjuntos A, B, C y aplicaciones $f\colon A \to B$ y $g\colon B \to C$, la **composición** $h = g \circ f$ es una aplicación que va de A en C y a cada elemento $x \in A$ le asocia $h(x) = g(f(x))$. Por ejemplo, supongamos $f\colon \mathbb{R} \to \mathbb{R}$ y $g\colon \mathbb{R} \to \mathbb{R}$ tales que $f(x) = \cos x$ y $g(y) = y^2$. Entonces $h = g \circ f$ es la aplicación $h(x) = \cos^2 x$ (es decir, $h(x) = (\cos x)^2$, primero hacemos el coseno, que sería aplicar f y después elevamos lo que haya salido al cuadrado, que sería aplicar g).

Capítulo 2

Variable compleja

2.1. Números complejos y operaciones

2.1.1. Definición

El conjunto más elemental de números es el de los naturales, $\mathbb{N} = \{1, 2, 3, ...\}$. Podemos sumarlos, pero cuando tratamos de restarlos, nos encontramos con que no podemos calcular $n - m$ si $n \leq m$ sin salirnos de $\mathbb{N}$, así que necesitamos definir más números: los enteros $\mathbb{Z}$. Ahora podemos sumar, restar y multiplicar sin problemas, pero cuanto intentamos dividir, podemos calcular algunas divisiones como $6/3$ pero si intentamos dividir $7/3$ nos vuelven a faltar números. Entonces definimos los racionales. Y cuando estos se quedan cortos, añadimos los irracionales, como $\sqrt{2}$ o π para formar el conjunto de los reales. Pero, una vez más, cuando intentamos calcular las raíces de un polinomio, a veces podemos, como en $x^2 - 3x + 2 = 0$, y a veces no, como en $x^2 + 1 = 0$. Para poder encontrar soluciones a cualquier polinomio, definimos un nuevo conjunto de números, los *números complejos*.

Definimios la *unidad imaginaria* i tal que $i^2 = -1$. Esto nos permite resolver cualquier raíz cuadrada negativa ya que $\sqrt{-4} = \sqrt{4 \cdot (-1)} = \sqrt{4}\sqrt{-1} = 2i$.

Un *número complejo* es una expresión de la forma

$$a + bi$$

con $a, b \in \mathbb{R}$. Esta expresión se denomina *forma binómica* del número complejo (veremos otras formas de expresarlos más adelante). El valor a se llama *parte real* y el valor b se llama *parte imaginaria*. Así, si $z = a + bi$ denotamos $\mathrm{Re}(z) = a$ e $\mathrm{Im}(z) = b$.

El conjunto de los números complejos se denota por

$$\mathbb{C} = \{a + bi \, : \, a, b \in \mathbb{R}\}.$$

Este conjunto se puede identificar con $\mathbb{R}^2$ donde al número complejo $a+bi$ le corresponde el punto (a,b). Así, el eje horizontal se denomina eje real y el vertical, eje imaginario.

Observación 2.1.1. *Igual que los naturales son los enteros positivos, los números reales, son los complejos cuya parte imaginaria es 0. Si la parte real es 0, decimos que el número es* imaginario puro.

2.1.2. Suma y resta

La suma y la resta de números complejos se hacen operando la parte real y la parte imaginaria por separado.

- $(3+2i)+(5-4i)=(3+5)+(2i-4i)=8-2i.$
- $(3+2i)-(5-4i)=(3-5)+(2i-(-4i))=-2+6i.$

En general, dados dos números complejos cualesquiera $(a+bi)$ y $(c+di)$, su suma se define como

$$(a+bi)+(c+di)=(a+c)+(b+d)i.$$

Del mismo modo, la resta se define como

$$(a+bi)-(c+di)=(a-c)+(b-d)i.$$

Es inmediato ver que la suma de números complejos cumple las propiedades deseables en cualquier operación. Si

- asociativa: $(z_1+z_2)+z_3=z_1+(z_2+z_3)\ \forall z_1,\ z_2,\ z_3\in\mathbb{C}$,
- conmutativa: $z_1+z_2=z_2+z_1\ \forall z_1,z_2\in\mathbb{C}$,
- tiene elemento neutro: $0+0i$,
- tiene elemento opuesto: $-(a+bi)=-a-bi$.

Ejercicio 2.1.2. *Probar las propiedades anteriores para la suma de complejos.*

2.2. Conjugado, módulo y argumento

Se define el *conjugado* de un número complejo $z = a + bi$ al número:

$$\overline{z} = a - bi,$$

cambiando de signo la parte imaginaria.

Por ejemplo, $\overline{2+3i} = 2 - 3i$, o $\overline{2-3i} = 2 + 3i$.

Ejercicio 2.2.1. *Demostrar las siguientes propiedades del conjugado:*

- $\overline{\overline{z}} = z$
- $\overline{z} = z$ *si y solo si* $z \in \mathbb{R}$
- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$
- $\overline{-z} = -\overline{z}$

Se define el *módulo* de un número complejo $z = a + bi$ al número:

$$|z| = \sqrt{a^2 + b^2}.$$

Ejercicio 2.2.2. *Demostrar las siguientes propiedades del módulo*

- $|z|$ *representa la distancia del punto* z *al 0*
- $|z| = |\overline{z}|$
- $|\operatorname{Re}(z)| \leq |z|$
- $|z_1 + z_2| \leq |z_1| + |z_2|$
- $|z_1 - z_2|$ *representa la distancia entre* z_1 z_2
- $|z - z_0| = R$ *representa la circunferencia de centro* z_0 *y radio* R

Se define el *argumento* de un número complejo $z = a+bi$, $\arg(z)$, al ángulo θ que forma el vector (a, b) con el eje real positivo en sentido antihorario. Nótese que $\theta + 2\pi k$ es igual al ángulo θ para todo $k \in \mathbb{Z}$. Elegiremos representar el argumento por un ángulo entre 0 y 2π radianes (otra posibilidad natural sería considerar el ángulo entre $-\pi$ y π).

Si $a + bi$ tiene argumeto θ y $a \neq 0$ es fácil ver que $tg(\theta) = \frac{b}{a}$. Nótese también que dado un ángulo θ, $tg(\theta) = tg(\theta + \pi)$. Si utilizamos una calculadora, al calcular el arcotangente de un número, nos va a dar un ángulo entre $-\pi/2$ y $\pi/2$. Para determinar el argumento del número complejo, debemos saber primero en qué cuadrante está y ajustar el ángulo que nos devuelve la calculadora. Si el ángulo está en el tercer cuadrante, habrá que sumarle π. Si está en el cuarto cuadrante, la calculadora nos dará un número negativo así que, para obtener un número entre 0 y 2π habrá que sumarle 2π. Por último, si está en el segundo cuadrante, nos saldrá un número negativo en el cuarto cuadrante y, por tanto, habrá que sumarle π. En resumen:

$$\arg(z) = \theta = \begin{cases} \operatorname{arc\,tg}(\frac{b}{a}), & \text{si } z \text{ está en el } 1^{er} \text{ cuadrante,} \\ \frac{\pi}{2} & \text{si } a = 0 \text{ y } b > 0, \\ \operatorname{arc\,tg}(\frac{b}{a}) + \pi, & \text{si } z \text{ está en el } 2^{\underline{o}} \text{ o } 3^{er} \text{ cuadrante,} \\ -\frac{\pi}{2} & \text{si } a = 0 \text{ y } b < 0, \\ \operatorname{arc\,tg}(\frac{b}{a}) + 2\pi, & \text{si } z \text{ está en el 4 cuadrante} \end{cases}$$

Ejemplo 2.2.3.

- *Ver el caso de* $1+i, -1-i, 1-i, -1+i$.
- *Si* $z = -\sqrt{3}+i$, $|z| = 2$ *y* $\theta = \operatorname{arc\,tg}\left(\frac{1}{-\sqrt{3}}\right) + \pi = -\frac{\pi}{6} + \pi = \frac{5\pi}{6}$, *ya que* $-\sqrt{3}+i$ *está en el segundo cuadrante.*

2.2.1. Producto y división

Para multiplicar dos números complejos, simplemente desarrollamos el paréntesis multiplicando término a término:

- $(3+2i)\cdot(5-4i) = 15 - 8i^2 - 12i + 10i = 15 + 8 - 2i = 23 - 2i$.

En general, el *producto* de números complejos se define como

$$(a+bi)(c+di) = (ac-bd) + (ad+bc)i.$$

Es inmediato ver que el producto de números complejos cumple las siguientes propiedades:

- asociativa: $(z_1 \cdot z_2)\cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$ $\forall z_1,\ z_2,\ z_3 \in \mathbb{C}$,
- conmutativa: $z_1 \cdot z_2 = z_2 \cdot z_1$ $\forall z_1, z_2 \in \mathbb{C}$,
- tiene elemento neutro: $(1+0i)$
- es distributivo respecto a la suma: $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$.

Ejercicio 2.2.4. *Probar las propiedades mencionadas para el producto de complejos.*

Podemos definir la división como sigue, multiplicando numerador y denominador por el conjugado del denominador:

- $\frac{2+i}{1-i} = \frac{(2+i)(1+i)}{(1-i)(1+i)} = \frac{2+i^2+2i+i}{1-i^2+i-i} = \frac{1+3i}{2} = \frac{1}{2} + \frac{3}{2}i.$

En general, se define como

$$\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i.$$

Observación 2.2.5. *El producto de números complejos tiene elemento elemento inverso para todo $z = a + bi$ distinto del 0:*

$$(a+bi)^{-1} = \frac{1}{a+bi} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i.$$

Por ejemplo:

$$(3+4i)^{-1} = \frac{3}{3^2+4^2} - \frac{4}{3^2+4^2}i = \frac{3}{25} - \frac{4}{25}i.$$

Ejercicio 2.2.6. *Demostrar las siguientes propiedades del conjugado y el módulo:*

- $\overline{z_1 z_2} = \overline{z_1 z_2}$
- $\overline{z^{-1}} = \overline{z}^{-1}$, *si* $z \neq 0$
- $\overline{z_1 z_2} = \overline{z_1 z_2}$
- $\overline{z^{-1}} = \overline{z}^{-1}$, *si* $z \neq 0$
- $\mathrm{Re}(z) = \frac{z+\overline{z}}{2}$
- $\mathrm{Im}(z) = \frac{z-\overline{z}}{2i}$
- $|z|^2 = z\overline{z}$
- $|z_1 z_2| = |z_1||z_2|$
- *si* $z_2 \neq 0$, $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$

Ejercicio 2.2.7. *Probar que si $z_1 z_2 = 0$ entonces $z_1 = 0$ o $z_2 = 0$.*

Ejercicio 2.2.8. *Demostrar las siguientes propiedades:*

- *si* $z_3 \neq 0$, $\frac{z_1+z_2}{z_3} = \frac{z_1}{z_3} + \frac{z_2}{z_3}$
- *si* $z_2 \neq 0$, $\frac{z_1}{z_2} = z_1 \frac{1}{z_2}$
- *si* $z_1, z_2 \neq 0$, $\frac{1}{z_1 z_2} = (z_1 z_2)^{-1} = z_1^{-1} z_2^{-1} = \frac{1}{z_1}\frac{1}{z_2}$
- $(z_1 + z_2)^2 = z_1^2 + 2z_1 z_2 + z_2^2$
- $(z_1 + z_2)^n = \sum_0^n \binom{n}{k} z_1^{n-k} z_2^k$

2.3. Forma trigonométrica, polar y exponencial

Forma trigonométrica:

Dados el módulo ρ y el argumento θ de un número complejo $z = a + bi$, por trigonometía elemental, tenemos que

$$a = \rho \cos\theta, \qquad b = \rho \operatorname{sen}\theta.$$

Por tanto, podemos expresar el número complejo como

$$z = \rho(\cos\theta + i \operatorname{sen}\theta),$$

lo que se conoce como *forma trigonométrica* de z.

Dos números complejos son iguales si tienen el mismo módulo y sus argumentos se diferencian en $2k\pi$ para algún $k \in \mathbb{Z}$.

Dados dos números complejos en forma trigonométrica, el producto y el cociente se expresan del siguiente modo:

Si $z = \rho\Big(\cos(\theta) + i \operatorname{sen}(\theta)\Big)$, $z' = \rho'\Big(\cos(\theta') + i \operatorname{sen}(\theta')\Big)$

- $$\begin{aligned} zz' =& \rho\Big(\cos(\theta) + i \operatorname{sen}(\theta)\Big)\rho'\Big(\cos(\theta') + i \operatorname{sen}(\theta')\Big) \\ =& \rho\rho'[\cos(\theta+\theta') + i \operatorname{sen}(\theta+\theta')] \end{aligned}$$

  (para demostrarlo, basta hacer los productos uno a uno y usar las fórmulas del coseno y el seno de la suma de ángulos).

- $\frac{z}{z'} = \frac{\rho}{\rho'}(\cos(\theta - \theta') + i \operatorname{sen}(\theta - \theta'))$

  (basta ver que $z' \frac{\rho}{\rho'}\Big(\cos(\theta - \theta') + i \operatorname{sen}(\theta - \theta')\Big) = z$).

Ejercicio 2.3.1. *Comprobar que si tomamos los números complejos* $1 + i = \sqrt{2}(\cos(\pi/4) + i \operatorname{sen}(\pi/4))$ *y* $-2+2i = \sqrt{8}(\cos(3\pi/4) + i \operatorname{sen}(3\pi/4))$ *obtenemos el mismo resultado:*

- *multiplicándolos en forma binómica o en forma trigonométrica, es decir, veamos que*

$$\begin{aligned}(1+i)(-2+2i) &= -4 = 4\cos(\pi)\\ &= \sqrt{2}\Big(\cos(\pi/4)+i\,\mathrm{sen}(\pi/4)\Big)\sqrt{8}\Big(\cos(3\pi/4)+i\,\mathrm{sen}(3\pi/4)\Big).\end{aligned}$$

- *dividiéndolos en forma binómica o en forma trigonométrica, es decir, veamos que*

$$\frac{-2+2i}{1+i} = 2i = 2\,\mathrm{sen}(\pi/2) = \frac{\sqrt{8}\Big(\cos(3\pi/4)+i\,\mathrm{sen}(3\pi/4)\Big)}{\sqrt{2}\Big(\cos(\pi/4)+i\,\mathrm{sen}(\pi/4)\Big)}.$$

Forma polar:

Ya hemos visto cómo un número complejo queda determinado por su módulo y su argumento. Dado un número complejo z con módulo ρ y argumento θ, se denota en *forma polar* como

$$z = \rho_\theta.$$

Por ejemplo, $1+i = \sqrt{2}(\cos(\pi/4)+i\,\mathrm{sen}(\pi/4))$ se podría denotar en forma polar como $\sqrt{2}_{\pi/4}$.

Así, denotado en forma polar, tenemos que

$$zz' = \rho_\theta\rho'_{\theta'} = (\rho\rho')_{\theta+\theta'} \quad \text{y} \quad \frac{z}{z'} = \frac{\rho_\theta}{\rho'_{\theta'}} = \left(\frac{\rho}{\rho'}\right)_{\theta-\theta'}.$$

Forma exponencial:

Por último, usando la fórmula de Euler, $e^{i\theta} = \cos(\theta) + i\,\mathrm{sen}(\theta)$ podemos expresar el número complejo z con módulo ρ y argumento θ en *forma exponencial* simplemente como

$$z = \rho e^{i\theta}$$

Por ejemplo, $1+i = \sqrt{2}(\cos(\pi/4)+i\,\mathrm{sen}(\pi/4))$ se podría denotar en forma exponencial como $\sqrt{2}e^{i\frac{\pi}{4}}$.

Denotado en forma exponencial, tenemos que

$$zz' = \rho e^{i\theta}\rho' e^{i\theta'} = (\rho\rho')e^{i(\theta+\theta')} \quad \text{y} \quad \frac{z}{z'} = \frac{\rho e^{i\theta}}{\rho' e^{i\theta'}} = \frac{\rho}{\rho'}e^{i(\theta-\theta')}.$$

Ejercicio 2.3.2. *Escribir en forma trigonométrica, polar y exponencial el número* $z = -1 - \sqrt{3}i$.

$\rho = 2$, $\theta = arctg\left(\frac{-\sqrt{3}}{-1}\right) + \pi = \frac{\pi}{3} + \pi = \frac{4\pi}{3}$. *Por tanto,*

$$z = 2\left(\cos\left(\frac{4\pi}{3}\right) + i\,\mathrm{sen}\left(\frac{4\pi}{3}\right)\right) = 2_{\frac{4\pi}{3}} = 2e^{i\frac{4\pi}{3}}.$$

Conviene recordar que:

$$1_{\pi/6} = \frac{\sqrt{3}}{2} + \frac{1}{2}i, \quad 1_{\pi/4} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i, \quad 1_{\pi/3} = \frac{1}{2} + \frac{\sqrt{3}}{2}i.$$

A partir de estos, podemos calcular de forma inmediata los puntos $1_{k\pi/6}$ y $1_{j\pi/4}$ para cualquier $j, k \in \mathbb{N}$.

2.4. Potencias de números complejos

Primero notamos que $i^2 = -1$, $i^3 = -i$, $i^4 = 1$ y, a partir de ahí, para todo $4 < k \in \mathbb{N}$, $i^k = i^j$ siendo j el resto de dividir k entre 4.

Teorema 2.4.1 (Fórmula de De Moivre)**.** *Para cualquier número complejo* $z = \rho(\cos\theta + i\,\mathrm{sen}\,\theta)$ *y* $n \in \mathbb{N}$, *se verifica que*

$$z^n = \rho^n(\cos(n\theta) + i\,\mathrm{sen}(n\theta)).$$

Ejercicio 2.4.2. *Demostrar, por inducción, la fórmula de De Moivre.*

En forma polar con $z = \rho_\theta$, tenemos que

$$z^n = \rho^n_{n\theta}.$$

En forma exponencial con $z = \rho e^{i\theta}$, tenemos que

$$z^n = (\rho e^{i\theta})^n = \rho^n e^{in\theta}.$$

Ejercicio 2.4.3. *Ver que* $(1-i)^8 = 16$ *(nótese que* $\arg(1-i) = \frac{7\pi}{4}$*).*

2.5. Raíces de números complejos

Dado un número complejo z se dice que $w \in \mathbb{C}$ es una *raíz n-ésima* de z si $w^n = z$.

Teorema 2.5.1. *Todo número complejo $z = \rho(\cos\theta + i \operatorname{sen}\theta)$ distinto de 0 tiene exactamente n raíces n-ésimas distintas, dadas por*

$$w_k = \sqrt[n]{\rho}\Big(\cos\frac{\theta + 2k\pi}{n} + i \operatorname{sen}\frac{\theta + 2k\pi}{n}\Big).$$

Ejemplo 2.5.2. *Las raíces n-ésimas de la unidad, $1 = 1(\cos 0 + i \operatorname{sen} 0)$ son*

$$w_k = \cos\frac{2k\pi}{n} + i \operatorname{sen}\frac{2k\pi}{n}.$$

En particular, las raíces cúbicas de la unidad serían $w_0 = 1$, $w_1 = \cos\frac{2\pi}{3} + i \operatorname{sen}\frac{2\pi}{3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ y $w_2 = \cos\frac{4\pi}{3} + i \operatorname{sen}\frac{4\pi}{3} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$

Ejercicio 2.5.3. *Calcular las raíces cuartas de $z = 16(\cos\pi + i \operatorname{sen}\pi)$.*

Al igual que con los números reales, dado un polinomio $P(z)$ con coeficientes complejos y un número $\alpha \in \mathbb{C}$ tenemos que

$$P(\alpha) = 0 \text{ si y solo si } z - \alpha \text{ divide a } P(z).$$

Nota: Esto nos permite aplicar el método de Ruffini para calcular raíces de polinomios también en $\mathbb{C}$.

Ejercicio 2.5.4. *Factorizar el polinomio $z^4 - 1$.*

Teorema 2.5.5 (fundamental del álgebra)**.** *Un polinomio de grado $n \geq 1$ con coeficientes en $\mathbb{C}$ tiene exactamente n raíces en $\mathbb{C}$ (contando multiplicidades).*

Proposición 2.5.6. *Dado un polinomio $P(z)$ con coeficientes reales, si $\alpha \in \mathbb{C}$ es una raíz, entonces $\overline{\alpha} \in \mathbb{C}$.*

Ejercicio 2.5.7. *Demostrar la proposición anterior usando las propiedades del conjugado.*

Capítulo 3

Elementos básicos del álgebra lineal

3.1. Matrices

3.1.1. Definición

Una matriz real no es más que un conjunto de números reales dispuestos de cierta manera.

Se llama *matriz real de orden* $m \times n$ (o de m filas y n columnas) al conjunto de $m \cdot n$ números reales dispuestos de la forma

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

Se representan por A, (a_{ij}) o $(a_{ij})_{m \times n}$. El elemento a_{ij} se llama *coeficiente*, el índice i corresponde a la fila y el j a la columna. El conjunto de todas las matrices de orden $m \times n$ se denota $\mathcal{M}_{m \times n}$.

Dos matrices A y B de orden $m \times n$ son iguales si sus coeficientes a_{ij} y b_{ij} son iguales para $i = 1, 2, ..., m$, $j = 1, 2..., n$.

Ejemplo 3.1.1.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \neq \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Hay algunos tipos de matrices que son importantes y les damos un nombre:

- Una matriz de orden $m \times 1$ es una *matriz columna*.
- Una matriz de orden $1 \times n$ es una *matriz fila*.
- Si $m = n$ la matriz es una *matriz cuadrada*.
- Si en una matriz cuadrada todos los elementos fuera de la diagonal son 0 ($a_{ij} = 0 \, \forall \, i \neq j$) decimos que la matriz es diagonal.
- I_n denota la matriz cuadrada de orden n cuyos elementos de la diagonal principal son 1 y los restantes 0 ($a_{ii} = 1 \; \forall \, i$ y $a_{ij} = 0 \; \forall \, i \neq j$). Si no hay ambigüedad en el orden, la denotaremos símplemente como I.
- Si en una matriz cuadrada todos los elementos por encima o por debajo de la diagonal son 0 ($a_{ij} = 0 \, \forall \, i > j$ o $a_{ij} = 0 \, \forall \, i < j$) decimos que la matriz es triangular.
- Una matriz se llama *escalonada por renglones* o simplemente *escalonada* si cumple con las siguientes propiedades:
  1. Todas las filas cero están en la parte inferior de la matriz.
  2. El primer elemento diferente de cero (o pivote) de cada fila está a la derecha del primer elemento diferente de cero (o pivote) de la fila anterior.
- Una matriz escalonada se llama *reducida* si todos los pivotes son 1 y todos los coeficientes encima del pivote son 0.

Ejemplo 3.1.2. *Triangular superior. Si $i > j$ entonces $a_{ij} = 0$:*

$$\begin{pmatrix} 4 & 2 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{pmatrix}$$

Triangular inferior. Si $i < j$ entonces $a_{ij} = 0$:

$$\begin{pmatrix} 4 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -2 & 2 \end{pmatrix}$$

Diagonal. Si $i \neq j$ entonces $a_{ij} = 0$:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

3.1.2. Suma

Dadas dos matrices A y B de orden $m \times n$ definimos la matriz suma $C = A + B$ como la matriz cuyos coeficientes son de la forma: $c_{ij} = a_{ij} + b_{ij}$ para todo i, j.

Nótese que la suma de matrices sólo está definida entre matrices que tengan el mismo orden.

Ejemplo 3.1.3.

$$\begin{pmatrix} -1 & 2 & 3 \\ 1 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 0 & -1 \\ 3 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 4 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix} + \begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix}$$

no está definida.

Propiedades.

Dadas las matrices A, B, C de orden $m \times n$ se verifica:

- Asociativa: $(A + B) + C = A + (B + C)$.
- Existe elemento neutro (representamos como 0 a la matriz tal que $a_{ij} = 0$ para todo i, j): $A + 0 = 0 + A = A$.
- Existe elemento opuesto (representamos como $(-A)$ a la matriz de coeficientes $-a_{ij}$): $A + (-A) = (-A) + A = 0$.
- Conmutativa: $A + B = B + A$.

3.1.3. Multiplicación de una matriz por un número

Dado $\lambda \in \mathbb{R}$, la matriz λA se define como la matriz cuyos coeficientes son λa_{ij}.

Ejemplo 3.1.4. *Si* $\lambda = 3$ *y* $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 0 \end{pmatrix}$ *entonces* $\lambda A = \begin{pmatrix} 3 & 6 & 9 \\ 6 & -3 & 0 \end{pmatrix}$

Propiedades.

Para todo para de números $\lambda, \mu \in \mathbb{R}$ y todo para de matrices A, B de orden $m \times n$ se verifica:

- $(\lambda + \mu)A = \lambda A + \mu A$.
- $\lambda(A + B) = \lambda A + \lambda B$.
- $\lambda(\mu A) = (\lambda\mu)A$.
- $1A = A$.

3.1.4. Multiplicación de matrices

Sea $A = (a_{ij})$ una matriz de orden $m \times k$ y sea $B = (b_{ij})$ una matriz de orden $k \times n$. El producto $AB = C$ se define como la matriz de orden $m \times n$ donde el elemento c_{ij} de la fila i y la columna j de AB es:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{ik}b_{kj}$$

para $i = 1, 2, ..., m$, $j = 1, 2, ..., n$.

Este producto está definido si y solo si el número de columnas de la matriz A es igual al número de filas de la matriz B.

Ejemplo 3.1.5.

- $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 32$
- $\begin{pmatrix} 1 & 0 & -1 & 2 \\ 3 & 2 & 1 & -2 \\ 4 & 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & -1 \\ 2 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 9 \\ 7 & 10 \end{pmatrix}$

Propiedades.

Sean A, BC matrices y $\lambda \in \mathbb{R}$. Suponiendo que las operaciones siguientes están definidas, se verifica:

- Asociativa: $(AB)C = A(BC)$.
- Distributiva: $A(B + C) = AB + AC$ y $(A + B)C = AC + BC$.
- No es, en general, conmutativa: $AB \neq BA$.
- Existe elemento neutro: $AI = IA = A$.
- $\lambda(AB) = (\lambda A)B$

Ejercicio 3.1.6. *Dar un ejemplo de dos matrices A, B tales que $AB \neq BA$.*

3.1.5. Matrices especiales

Traspuesta de una matriz.

La transpuesta de una matriz A, A^t, es la matriz que resulta de cambiar en A las filas por las columnas. Es decir, si los coeficientes de A son a_{ij} los de A^t son a_{ji}.

Ejemplo 3.1.7. $A = \begin{pmatrix} 2 & 1 \\ 0 & 4 \\ 1 & -1 \end{pmatrix}$ $\qquad A^t = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 4 & -1 \end{pmatrix}$

Propiedades

Sean A, B, C matrices y $\lambda \in \mathbb{R}$. Suponiendo que las operaciones siguientes están definidas, se verifica:

- $(A^t)^t = A$.
- $(A + B)^t = A^t + B^t$.
- $(\lambda A)^t = \lambda A^t$.
- $(AB)^t = B^t A^t$.

Matrices simétricas y antisimétricas

Una matriz cuadrada es simétrica si es igual a su traspuesta, $A^t = A$, es decir, $a_{ij} = a_{ji}\ \forall\, i, j$.

Una matriz cuadrada es antisimétrica si es igual a la opuesta de su transpuesta, $A^t = -A$, es decir $a_{ij} = -a_{ji}\ \forall\, i, j$. Si A es antisimétrica, en particular, dado que $a_{ii} = -a_{ii}$ entonces $2a_{ii} = 0$ y $a_{ii} = 0\ \forall\, i$.

Proposición 3.1.8. *Toda matriz cuadrada A puede descomponerse como suma de una matriz simétrica, $(A + A^t)/2$, y una antisimétrica $(A - A^t)/2$.*

Ejemplo 3.1.9.

- $A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 4 & 7 \\ 3 & 7 & -1 \end{pmatrix}$ *es una matriz simétrica.*
- $A = \begin{pmatrix} 0 & 1 & -3 \\ -1 & 0 & 7 \\ 3 & -7 & 0 \end{pmatrix}$ *es una matriz antisimétrica.*
- AA^t *es una matriz simétrica:* $(AA^t)^t = (A^t)^t A^t = AA^t$.

3.1.6. Inversa de una matriz

En el producto entre números reales, hay un *elemento neutro*, es decir, un elemento que multiplicado por cualquier otro número real x nos da x. El elemento neutro de los números reales es el número 1. Si consideramos matrices cuadradas de orden n, como vimos en las propiedades de la multiplicación, el elemento neutro es la matriz I_n.

En la multiplicación de números reales todo elemento posee un *elemento inverso*, esto es, para cualquier número existe otro número tal que el producto de ambos es el elemento neutro. $\forall x \in \mathbb{R}\ \exists x' \in \mathbb{R}$ tal que $x'x = 1 = xx'$. El inverso de x se denota x^{-1} y, en el caso de los números reales $x^{-1} = \frac{1}{x}$.

En el caso de una matriz A, una matriz inversa B debe cumplir que $BA = I = AB$ con el producto de matrices. La primera consecuencia de esto es que la matriz A debe ser cuadrada.

Ejercicio 3.1.10. *Explicar por qué una matriz de orden $m \times n$ con $m \neq n$ no puede tener inversa.*

Además, no toda matriz cuadrada tiene inversa. Si una matriz cuadrada A tiene inversa, ésta se denota A^{-1}.

Propiedades

- Si una matriz admite inversa, ésta es única.
- A^{-1} tiene inversa. De hecho, $(A^{-1})^{-1} = A$.
- Si A y B tienen inversa entonces $(AB)^{-1} = B^{-1}A^{-1} (\neq A^{-1}B^{-1})$.
- Si A tiene inversa entonces $(A^t)^{-1} = (A^{-1})^t$.

3.1.7. Álgebra de matrices

Podemos operar algebraicamente con las matrices igual que hacemos con números en una ecuación. En este caso hay que ser algo más cuidadosos porque, como hemos visto, las propiedades no son las mismas.

No hay propiedad conmutativa en el producto entre matrices, por tanto si tenemos dos expresiones con matrices X, Y donde $X = Y$ y queremos multiplicar ambos lados de la ecuación por la matriz A se cumple que $AX = AY$ o bien $XA = YA$ pero puede ocurrir que $AX \neq YA$.

A su vez, "algo que está multiplicando" no puede "pasar dividiendo". Primero, no existe la división entre matrices. Lo que se puede hacer es multiplicar por el inverso en ambos lados, pero ¡únicamente en el caso de que ya sepamos que existe inverso!

Por ejemplo si A es invertible y sabemos que $XA = Y$ entonces podemos multiplicar por el inverso de A ¡por la derecha! en ambos lados de la ecuación y obtenemos que $X = YA^{-1}$.

Ejemplo 3.1.11. *Si conocemos las matrices A y B (y suponemos que todos los produtos e inversas necesarios se pueden hacer), ¿cómo podemos determinar la matriz X en la siguiente ecuación?*

$$AX = X + B$$

Pensemos primero cómo lo haríamos con números. Si la ecuación fuera $3x = x+4$*, restaríamos* x *en ambos lados de la ecuación para obtener* $3x-x = 4$*, sacaríamos factor común a la* x *para llegar a* $(3-1)x = 4 \Rightarrow 2x = 4$ *(nótese que al sacar factor común a la* x *ha aparecido un 1 que no era explícito antes, como si tuviéramos* $3x - 1x = 4$*) y finalmente dividiríamos por dos ambos lados de la ecuación para llegar a* $x = 2$*. ¿Cómo sería entonces el proceso con matrices?*

En primer lugar, restamos X *en ambos lados obteniendo* $AX - X = B$*. A continuación podemos sacar factor común a la* X*, dado que también se cumple la propiedad distributiva. De nuevo, necesitamos hacer explícito qué es lo que multiplica (por la izquierda) a la matriz* X*. En este caso, sería la matriz unidad,* I*. Llegamos entonces a la expresión* $(A-I)X = B$*. Finalmente, dado que no existe la división en matrices, para despejar la matriz* X *tenemos que multiplicar por el inverso de* $(A - I)$ *por la izquierda en ambos miembros. Nos queda, por tanto,*

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

Ejercicio 3.1.12. *Si conocemos las matrices* A *y* B *(y suponemos que todos los produtos e inversas necesarios se pueden hacer), ¿cómo podemos determinar la matriz* X *en la ecuación*

$$AX = B + AXB$$

?

La resolución podría ser como sigue:

$$AX = B + AXB \Rightarrow AX - AXB = B \Rightarrow A(X - XB) = B \Rightarrow X - XB = A^{-1}B$$
$$\Rightarrow X(I - B) = A^{-1}B \Rightarrow X = A^{-1}B(I - B)^{-1}.$$

Es muy importante darse cuenta de que en el caso de matrices no es lo mismo multiplicar por un lado que por otro (no hay propiedad conmutativa en el producto) así que al sacar factor común o al despejar hay que tener en cuenta siempre el orden de los productos.

Ejercicio 3.1.13. *Asumiendo que conocemos las matrices* A *y* B *(y suponemos que todos los produtos e inversas necesarios se pueden hacer) despejar la matriz* X *en las siguientes ecuaciones:*

- $X^{-1}(A + B)^{-1} = A^{-1}$
- $A^{-1}(A - BX)^t = B$
- $AX - (B^{-1}A^{-1})^{-1} = I$

- $B^{-1}X^t - (AB)^{-1} = I$
- $(X - I)A = X + B$
- $(BX^{-1})^{-1} - B^{-1} = XA$

3.1.8. Cálculo de la inversa de una matriz por el método de Gauss-Jordan

Dada una matriz (cuadrada) invertible A queremos calcular su inversa A^{-1}. Este método consiste en calcularla realizando transformaciones elementales de filas de A. Estas transformaciones son las siguientes:

1º Intercambiar filas.

2º Multiplicar cada fila por un número $\lambda \in \mathbb{R}$, $\lambda \neq 0$.

3º Sumar una fila a otra.

Para calcularla, empezamos escribiendo la matriz A y a su derecha la matriz unidad I obteniendo una matriz de orden $n \times 2n$, $(A : I)$. Entonces se realizan las operaciones elementales por filas en la matriz $(A : I)$ hasta conseguir que A se transforme en I y entonces I se habrá transformando en A^{-1}.

Ejemplo 3.1.14.

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix} \qquad (A : I) = \begin{pmatrix} 1 & -1 & 1 & : & 1 & 0 & 0 \\ 0 & 0 & 1 & : & 0 & 1 & 0 \\ 1 & 2 & 1 & : & 0 & 0 & 1 \end{pmatrix} \rightarrow$$

$$\xrightarrow{F_3:=F_3-F_1} \begin{pmatrix} 1 & -1 & 1 & : & 1 & 0 & 0 \\ 0 & 0 & 1 & : & 0 & 1 & 0 \\ 0 & 3 & 0 & : & -1 & 0 & 1 \end{pmatrix} \xrightarrow{F_2 \leftrightarrow F_3} \begin{pmatrix} 1 & -1 & 1 & : & 1 & 0 & 0 \\ 0 & 3 & 0 & : & -1 & 0 & 1 \\ 0 & 0 & 1 & : & 0 & 1 & 0 \end{pmatrix} \rightarrow$$

$$\xrightarrow{F_2:=\frac{1}{3}F_2} \begin{pmatrix} 1 & -1 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & -\frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & : & 0 & 1 & 0 \end{pmatrix} \xrightarrow{F_1:=F_1+F_2} \begin{pmatrix} 1 & 0 & 1 & : & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & : & -\frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & : & 0 & 1 & 0 \end{pmatrix} \rightarrow$$

$$\xrightarrow{F_1:=F_1-F_3} \begin{pmatrix} 1 & 0 & 0 & : & \frac{2}{3} & -1 & \frac{1}{3} \\ 0 & 1 & 0 & : & -\frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & : & 0 & 1 & 0 \end{pmatrix}$$

Así:

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

3.2. Determinantes

3.2.1. Definición.

A toda matriz cuadrada A se le asocia un número real que se llama **determinante de la matriz**. Se denota $det(A)$ o $|A|$ (con barras verticales en lugar de paréntesis).

La definición general requiere introducir el concepto de *permutación*. Para simplificar la exposición tomaremos como defición su desarrollo. Veamos primero cómo se calculan los determinantes de orden 2 y 3.

- Determinante de orden 2. Si $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ entonces

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

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

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

$-a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$.

Este desarrollo será utilizado con frecuencia. La *regla de Sarrus*, disponible en cualquier referencia, permite memorizarlo cómodamente. Otra forma es hacer el desarrollo, tal y como se describe a continuación, a través de la fila o columna que contenga más ceros.

- Determinante de una matriz A de orden n. Desarrollo de un determinante por los elementos de una línea o columna.

Definición 3.2.1. *Se llama* ***menor complementario*** *del elemento a_{ij} al determinante de orden $n-1$ que se obtiene suprimiendo la fila i y la columna j. El menor complementario de a_{ij} se denota M_{ij}.*

Ejemplo 3.2.2. *Sea* $A = \begin{pmatrix} 4 & 1 & 2 \\ 0 & -1 & 1 \\ 3 & 0 & -2 \end{pmatrix}$. *El menor complementario del elemento a_{22} será por tanto el determinante* $\begin{vmatrix} 4 & 2 \\ 3 & -2 \end{vmatrix} = 4 \cdot (-2) - 3 \cdot 2 = -8 - 6 = -14.$

Definición 3.2.3. *Se llama **adjunto** del elemento a_{ij} al menor complementario del elemento a_{ij} multiplicado por $(-1)^{i+j}$. El adjunto de a_{ij} se denota A_{ij}.*

Nótese que $(-1)^{i+j}$ es igual a 1 si y sólo si i y j son ambos pares o ambos impares y es igual a -1 si uno de ellos es par y el otro impar.

Dada una matriz de orden n, el determinante de A es igual a la suma de los elementos de una fila o columna cualquiera por sus adjuntos correspondientes. Es decir,

$$\text{Por la fila } i\text{: } |A| = \sum_{k=1}^{n} (-1)^{i+k}\, a_{ik}\, M_{ik} = \sum_{k=1}^{n} a_{ik}\, A_{ik}.$$

$$\text{Columna } j\text{: } |A| = \sum_{k=1}^{n} (-1)^{k+j}\, a_{kj}\, M_{kj} = \sum_{k=1}^{n} a_{kj}\, A_{kj}.$$

Ejemplo 3.2.4.

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

Vamos a desarrollarlo por los elementos de la tercera columna:

$$|A| = (-1)^{1+3} 5 A_{13} + (-1)^{3+3}(-8) A_{33} = 5(-149) + (-8)91 = -1473.$$

$$A_{13} = \begin{vmatrix} 4 & 5 & 6 \\ 0 & 9 & 1 \\ 5 & 3 & 3 \end{vmatrix} = -149 \qquad A_{33} = \begin{vmatrix} 3 & 7 & 2 \\ 4 & 5 & 6 \\ 5 & 3 & 3 \end{vmatrix} = 91$$

Ejercicio 3.2.5. *Comprobar que si hacemos el desarrollo de un determinante de orden 3 a través de cualquier fila o columna obtenemos la expresión que vimos anteriormente.*

Definición 3.2.6. *Dada una matriz cuadrada de orden n, se define la **traza** de A como el valor de la suma de todos los elementos de la diagonal principal: $tr(A) := \sum_{i=1}^{n} a_{ii}$.*

Propiedades de los determinantes.

1 El determinante de una matriz y el de su traspuesta es el mismo: $|A| = |A^t|$.

2 Si se intercambian entre sí dos filas o dos columnas el determinante cambia de signo.

3 Si la matriz tiene una fila o columna de ceros, el determinante es 0.

4 Si la matriz tiene dos filas o columnas iguales o proporcionales entonces el determinante es 0.

5 Si se multiplica una sola fila o una sola columna por un número t entonces el determinante es multiplicado por t.

$$\begin{vmatrix} a_{11} & \cdots & ta_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & ta_{nj} & \cdots & a_{nn} \end{vmatrix} = t \cdot \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

6 Si los elementos de la columna j de A se decomponen como una suma, $a_{kj} = b_{kj} + c_{kj}$ para $k = 1, 2, ..., n$, el determinante es la suma de los determinantes, $|A| = |B| + |C|$, donde B y C son iguales a A excepto en la columna j que está formada, respectivamente, por los elementos b_{kj} y c_{kj}.

$$\begin{vmatrix} a_{11} & \cdots & b_{1j} + c_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & b_{nj} + c_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & b_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & b_{nj} & \cdots & a_{nn} \end{vmatrix} +$$

$$+ \begin{vmatrix} a_{11} & \cdots & c_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & c_{nj} & \cdots & a_{nn} \end{vmatrix}$$

7 Si los elementos de la fila i de A se decomponen como una suma, $a_{ik} = b_{ik} + c_{ik}$ para $k = 1, 2, ..., n$, el determinante es la suma de los determinantes, $|A| = |B| + |C|$, donde B y C son iguales a A excepto en la fila i que está formada, respectivamente, por los elementos b_{kj} y c_{kj}.

Ejercicio 3.2.7. *Expresar la propiedad mediante determinantes como en el caso anterior.*

8 Si a una fila (o columna) se le suma un múltiplo cualquiera de otra fila (o columna) el determinante no varía.

Ejemplo 3.2.8.

$$\begin{vmatrix} a_{11} & \cdots & a_{1j}+ta_{1k} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nj}+ta_{nk} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

9 Si A es una matriz triangular el determinante es el producto de los elementos de la diagonal: $|A| = \prod_{i=1}^{n} a_{ii}$.

10 El determinante del producto es el producto de los determinantes: $|AB| = |A| \cdot |B|$.

Podemos usar la propiedad 8 para calcular un determinante de forma sencilla. Operando con las filas o columnas sin variar el valor del determinante podemos hacer que por debajo de la diagonal todo sean ceros. Una vez obtenida una matriz triangular, calcular el determinante es inmediato por la propiedad 9. Las operaciones que podemos hacer con la matriz sin que cambie el determinante se llaman **operaciones elementales**.

Ejemplo 3.2.9.

- *Sea* $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ $\quad |A| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{vmatrix} = 0$

- $\begin{vmatrix} 2 & -3 & 4 \\ 5 & -1 & 6 \\ 8 & 0 & 2 \end{vmatrix} = -86.$

- $\begin{vmatrix} 2a & a & a & a \\ a & 2a & a & a \\ a & a & 2a & a \\ a & a & a & 2a \end{vmatrix} = \begin{vmatrix} 5a & a & a & a \\ 5a & 2a & a & a \\ 5a & a & 2a & a \\ 5a & a & a & 2a \end{vmatrix} = 5a \begin{vmatrix} 1 & a & a & a \\ 1 & 2a & a & a \\ 1 & a & 2a & a \\ 1 & a & a & 2a \end{vmatrix} =$

$$= 5a \begin{vmatrix} 1 & a & a & a \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{vmatrix} = 5a^4.$$

Ejercicio 3.2.10. *Calculamos mediante operaciones elementales*

$$|A| = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 2 & 5 & 0 & 3 \\ -1 & 1 & 2 & 0 \\ 0 & 0 & 5 & -4 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & -3 \\ 0 & 3 & 1 & 3 \\ 0 & 0 & 5 & -4 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & -5 & 12 \\ 0 & 0 & 5 & -4 \end{vmatrix} =$$
$$= \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & -5 & 12 \\ 0 & 0 & 0 & 8 \end{vmatrix} = -40.$$

Ejercicio 3.2.11. *Calcular, desarrollando por una fila o columna y mediante operaciones elementales* $|A| = \begin{vmatrix} 3 & 2 & 1 & -1 \\ 2 & 1 & 0 & -2 \\ 3 & -4 & 1 & 5 \\ 1 & 2 & -3 & 2 \end{vmatrix}$

3.2.2. Aplicaciones de los determinantes

Cálculo de la inversa de una matriz

Llamamos matriz adjunta de A, $Adj(A)$ a la matriz que resulta de sustituir cada elemento de la matriz A por su respectivo adjunto.

Se verifica que:

$$A \cdot (Adj(A))^t = (Adj(A))^t \cdot A = |A| \cdot I \tag{3.1}$$

Proposición 3.2.12. *Una matriz cuadrada es inversible si y sólo si su determinante es distinto de 0. En ese caso,*

$$A^{-1} = \frac{(Adj(A))^t}{|A|}.$$

Demostración. Si A es inversible entonces $|A| \neq 0$:

Si A es inversible, existe una matriz B tal que $AB = BA = I$. Entonces, $|AB| = |BA| = |I|$ y, por las propiedades del determinante, $|A||B| = |B||A| = |I| = 1$, luego $|A| \neq 0$ (de lo contrario $|B| \cdot 0 = 0 \neq 1$).

Si $|A| \neq 0$ entonces A es inversible:

Por la propiedad de la matriz adjunta, ver (3.1), $A \cdot \frac{(Adj(A))^t}{|A|} = \frac{(Adj(A))^t}{|A|} \cdot A = I$[1] luego $A^{-1} = \frac{(Adj(A))^t}{|A|}$. □

[1]Nótese que, para poder dividir por $|A|$ es necesaria la hipótesis de que $|A| \neq 0$

Ejemplo 3.2.13. *Dada la matriz* $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & 1 & 1 \end{pmatrix}$ *vamos a calcular* A^{-1}.

$$\Big(Adj(A)\Big)^t = \begin{pmatrix} -1 & 1 & -3 \\ 8 & -5 & 6 \\ -6 & 3 & -3 \end{pmatrix}$$

$$|A| = -3 \qquad A^{-1} = \frac{(Adj(A))^t}{|A|} = \begin{pmatrix} \frac{1}{3} & \frac{-1}{3} & 1 \\ \frac{-8}{3} & \frac{5}{3} & -2 \\ 2 & -1 & 1 \end{pmatrix}$$

Rango de una matriz

Si en una matriz A se toman k filas y k columnas, los elementos de la intersección de esas filas y columnas forman una submatriz cuyo determinante se llama **menor de orden** k. El menor formado por las k primeras filas y las k primeras columnas de A se llama menor principal de orden k.

Ejemplo 3.2.14. $A = \begin{pmatrix} 1 & 3 & -2 & 1 \\ 4 & -3 & 5 & 3 \\ 2 & -7 & 3 & 5 \end{pmatrix}$

a) $|1|$ *es el menor principal de orden 1.*

b) $\begin{vmatrix} 1 & -2 \\ 2 & 3 \end{vmatrix}$ *es un menor de orden 2.*

c) $\begin{vmatrix} 1 & -2 & 1 \\ 4 & 5 & 3 \\ 2 & 3 & 5 \end{vmatrix}$ *es un menor de orden 3.*

Definición 3.2.15. *Se dice que el* **rango** *de la matriz* A *es* k*, y escribimos* $\mathrm{rg}(A) = k$ *si existe un menor de orden* k *no nulo y todos los menores de orden superior a* k *son nulos.*

Veamos un método para calcular el rango de cualquier matriz A de orden $m \times n$. Para ello necesitamos introducir el concepto de *orlar un menor*. Orlar un menor de orden k consiste en formar un menor añadiendo una fila y una columna para obtener un menor de orden $k+1$.

- Se busca un menor de orden 1 no nulo, es decir, un número distinto de 0 en la matriz.
    - Si no existe, entonces $\mathrm{rg}(A) = 0$ y hemos terminado.
    - Si existe, entonces $\mathrm{rg}(A) \geq 1$ y continuamos el proceso.

- Se calculan los menores de orden 2 que se obtienen orlando el menor de orden 1 no nulo.
  - Si no existe, o son todos nulos entonces $\mathrm{rg}(A) = 1$ y hemos terminado.
  - Si existe, entonces $\mathrm{rg}(A) \geq 2$ y continuamos el proceso.
- Se calculan los menores de orden 3 que se obtienen orlando el menor de orden 2 no nulo.
  - Si no existe, o son todos nulos entonces $\mathrm{rg}(A) = 2$ y hemos terminado.
  - Si existe, entonces $\mathrm{rg}(A) \geq 2$ y continuamos el proceso.
- Repetimos la operación hasta que el proceso termina. Nótese que el rango máximo de una matriz de orden $m \times n$ es menor o igual que m y que n.

Nota 3.2.16. *Este procedimiento permite hallar el rango calculando menos determinantes. Si el rango de una matriz A es k, en lugar de calcular todos los menores de orden $k+1$ bastará con comprobar los que orlan a un determinado menor de orden k.*

Ejemplo 3.2.17. *Calcular el rango de la siguiente matriz:*

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

$$\mathrm{rg}(A) = 3 \textit{ por ser } \begin{vmatrix} 1 & 2 & -2 \\ 4 & -1 & 1 \\ -2 & 1 & 1 \end{vmatrix} = -18 \neq 0.$$

Las operaciones elementales no modifican el valor del rango. Por tanto, otro método para calcularlo consiste en transformar A en una matriz escalonada (en la que al principio de cada fila hay un cero más que en la anterior) mediante operaciones elementales. El rango de A es el número de filas (o columnas) no nulas que tiene la matriz escalonada.

Ejemplo 3.2.18. *Calcular el rango de la matriz A mediante operaciones elementales.*

$$\begin{pmatrix} 1 & 0 & 3 & 1 \\ 1 & -1 & 7 & -1 \\ 2 & 1 & 2 & 4 \\ 5 & 1 & 11 & 7 \end{pmatrix} \sim \begin{matrix} F_2 := F_2 - F_1 \\ F_3 := F_3 - 2F_1 \\ F_4 := F_4 - 5F_1 \end{matrix} \sim \begin{pmatrix} 1 & 0 & 3 & 1 \\ 0 & -1 & 4 & -2 \\ 0 & 1 & -4 & 2 \\ 0 & 1 & -4 & 2 \end{pmatrix} \sim$$

$$\sim \begin{array}{c} F_3 := F_3 - F_2 \\ F_4 := F_4 - F_2 \\ F_2 := -F_2 \end{array} \sim \begin{pmatrix} 1 & 0 & 3 & 1 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Si la matriz contiene algún parámetro, el rango puede depender del valor de éste. Así, deberemos separar los casos posibles, cuando los haya, en función del parámetro.

Ejemplo 3.2.19. *Calcular el rango de la matriz* $A = \begin{pmatrix} 1 & 3 \\ 2 & a \end{pmatrix}$

Es claro que $\mathrm{rg}(A) \geq 1$ *ya que hay coeficientes distintos de 0. También es inmediato ver que* $\mathrm{rg}(A) \leq 2$ *porque el menor más grande posible en la matriz es el de tamaño 2. Así pues, lo único que necesitamos ver es si el determinante de* A *es o no igual a 0.*

Tenemos que $|A| = a - 6$*, luego* $|A| = 0$ *si y sólo si* $a = 6$*. Por tanto, si* $a = 6$*,* $\mathrm{rg}(A) = 1$ *y si* $a \neq 6$*,* $\mathrm{rg}(A) = 2$*.*

Ejemplo 3.2.20. *Calcular el rango de la matriz* $B = \begin{pmatrix} 1 & 3 & 4 \\ 2 & a & 5 \end{pmatrix}$

De nuevo es claro que $\mathrm{rg}(A) \geq 1$ *y* $\mathrm{rg}(A) \leq 2$*. Sin embargo, ahora tenemos un menor de tamaño 2 dado por las columnas 1 y 3,* $M = \begin{pmatrix} 1 & 4 \\ 2 & 5 \end{pmatrix}$ *y tal que* $|M| \neq 0$ *independientemente del parámetro* a*. Por tanto,* $\mathrm{rg}(B) = 2$ *para cualquier valor de* a*.*

Ejemplo 3.2.21. *Calcular el rango de la siguiente matriz:*

$$C = \begin{pmatrix} 3 & -2 & a & 2 \\ 1 & -1 & 1 & 0 \\ a & 2 & -2 & 3 \end{pmatrix}.$$

En este caso, el rango de la matriz puede depender del valor del parámetro a*. Empezamos con el elemento* $a_{11} = 3$ *y vamos orlando:*

$\begin{vmatrix} 3 & -2 \\ 1 & -1 \end{vmatrix} = 1 \neq 0$*, por lo menos el rango es 2 (independientemente del valor de* a*). Orlamos con la tercera fila y la tercera columna:*

$\begin{vmatrix} 3 & -2 & a \\ 1 & -1 & 1 \\ a & 2 & -2 \end{vmatrix} = -4 + a^2$ *Así,* $-4 + a^2 = 0$ *si y sólo si* $a = \pm 2$*.*

Orlamos con la tercera fila y la cuarta columna:

$\begin{vmatrix} 3 & -2 & 2 \\ 1 & -1 & 0 \\ a & 2 & 3 \end{vmatrix} = 2a + 1$ *Así,* $2a + 1 = 0$ *si y sólo si* $a = -\frac{1}{2}$*.*

Por tanto, como a no puede ser al mismo tiempo ± 2 *y* $-\frac{1}{2}$*, uno de los dos determinantes es distinto de 0 y el rango de C es 3.*

3.3. Sistemas de ecuaciones lineales

Una expresión de la forma $a_1x_1 + a_2x_2 + ... + a_nx_n = b$ con $x_1, x_2, ..., x_n$ ariables y $a_1, a_2, ..., a_n, b$ son números dados, constituye una ecuación lineal.

Un conjunto de m ecuaciones lineales en las incognitas $x_1, x_2, ..., x_n$ constituye un **sistema**.

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\} \tag{3.2}$$

Los x_i se denominan *incógnitas*, los a_{ij} *coeficientes* y los b_i *términos independientes.*

Resolver un sistema de ecuaciones es obtener un conjunto de valores de las incognitas que satisfacen simultaneamente a todas las ecuaciones, cada uno de estos conjuntos de valores constituye una solución.

Un sistema que tiene solución es compatible. Si la solución es única se dice que es determinado y si tiene más de una solución indeterminado. Si el sistema no tiene solución es incompatible.

Ejemplo 3.3.1. *1)* $x + y = 0$ *Compatible indeterminado.*

2) $\left.\begin{array}{rcl} x + y &=& 0 \\ x - y &=& 0 \end{array}\right\}$ *Compatible determinado.*

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

Un sistema se puede expresar matricialmente. Sea A la matriz de los coeficientes de (3.2), que es una matriz del tipo (m, n) $B(m, 1)$ la matriz de los coeficientes indeterminados y $X(n, 1)$ es la matriz de las incognitas.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots\cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots\cdots & a_{2n} \\ \cdots & \cdots & \cdots\cdots\cdots & \cdots \\ a_{m1} & a_{m2} & \cdots\cdots\cdots & a_{mn} \end{pmatrix} B = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix} X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$$

El sistema (3.2) lo expresamos por la ecuación matricial:

$$AX = B \tag{3.3}$$

o mediante un sumatorio:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \text{ para } i = 1, 2..., m.$$

La matriz de orden $m \times (n+1)$, $(A : B)$ que se obtiene añadiendo a la matriz A una columna extra a su derecha que es la matriz B, se llama *matriz ampliada.*

Dos sistemas son **equivalentes** si tienen las mismas soluciones.

Proposición 3.3.2. *Un sistema de ecuaciones lineales es equivalente a cualquiera de los sistemas que resultan de realizar operaciones elementales en las filas de su matriz ampliada* $(A : B)$*.*

En los sistemas compatibles es necesario que hallemos métodos de resolución del sistema $AX = B$.

3.3.1. Teorema de Rouché-Frobenius

El sistema de $AX = B$ de m ecuaciones con n incognitas es:

- compatible $\Leftrightarrow$ rg$(A) =$ rg$(A : B)$,
  - compatible y determinado $\Leftrightarrow$ rg$(A) =$ rg$(A : B) = n$,
  - compatible indeterminado $\Leftrightarrow$ rg$(A) =$ rg$(A : B) < n$,
- incompatible si rg$(A) <$ rg$(A : B)$.

Ejemplo 3.3.3. *Estudiar la compatibilidad del sistema:*

$$\left.\begin{array}{rcl} x - 2y + z & = & 6 \\ x - 3z & = & 8 \\ y - 2z & = & 4 \\ 2x - 3y & = & 1 \end{array}\right\}$$

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & -2 \\ 2 & -3 & 0 \end{pmatrix}, \qquad (A : B) = \begin{pmatrix} 1 & -2 & 1 & : & 6 \\ 1 & 0 & -3 & : & 8 \\ 0 & 1 & -2 & : & 4 \\ 2 & -3 & 0 & : & 1 \end{pmatrix}$$

Vemos que rg$(A) = 2$ *ya que* $\begin{vmatrix} 1 & -2 \\ 1 & 0 \end{vmatrix} = 2 \neq 0$ *y si consideramos los orlados, comprobaremos que todos ellos tienen determinante 0:*

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

Para comprobar que el rango es menor que 3 habría que comprobar que todos los menores de orden 3 tienen determinante 0. Gracias al procedimiento de los menores orlados evitamos resolver dos determinantes de orden 3. ¿Puedes ver cuáles serían dichos determinantes?

Sin embargo, $\text{rg}(A:B) = 3$ *ya que* $\begin{vmatrix} 1 & -2 & 6 \\ 1 & 0 & 8 \\ 0 & 1 & 4 \end{vmatrix} = 6 \neq 0.$

Por tanto, $\text{rg}(A) \neq \text{rg}(A:B)$ *y el sistema es* ***no compatible***.

Ejemplo 3.3.4. *Estudiar la compatibilidad del sistema:*

$$\left.\begin{array}{rcl} x - 2y + 3z - 2t & = & -9 \\ y - z + 2t & = & 5 \\ x + 2z + t & = & 0 \end{array}\right\}$$

$$A = \begin{pmatrix} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & 2 \\ 1 & 0 & 2 & 1 \end{pmatrix}, \qquad (A:B) = \begin{pmatrix} 1 & -2 & 3 & -2 & : & -9 \\ 0 & 1 & -1 & 2 & : & 5 \\ 1 & 0 & 2 & 1 & : & 0 \end{pmatrix}$$

Comprobamos que $\text{rg}(A) = 3$ *ya que*

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

Como el número de filas es 3, es inmediato que $\text{rg}(A:B) \leq 3$. *Por tanto,* $\text{rg}(A) = \text{rg}(A:B) = 3$ *y el sistema es* ***compatible***. *Por ser el rango menor que el número de incógnitas, el sistema es* ***indeterminado***.

Ejemplo 3.3.5. *Discutir en función de los parámetros a y b:*

$$\left.\begin{array}{rcl} ax + 4y & = & 1 \\ x + ay & = & b \end{array}\right\}$$

Para discutir el sistema en función de los parámetros usamos el teorema de Rouché-Frobenius igual que si se tratara de números. La diferencia está en que ahora el rango de la matriz del sistema y de la ampliada dependen de los valores de a y b. Lo que hacemos es distinguir los casos (cuándo el sistema es compatible determinado, compatible indeterminado o incompatible) según los valores de a y b.

Primero calculamos los rangos (que dependen de los parámetros).

La matriz del sistema, A, es

$$A = \begin{pmatrix} a & 4 \\ 1 & a \end{pmatrix}$$

y la ampliada, $(A|b)$ es

$(A|b) = \begin{pmatrix} a & 4 & 1 \\ 1 & a & b \end{pmatrix}$.

Está claro que $\mathrm{rg}(A) \geq 1$ *ya que hay al menos un elemento distinto de 0 (el coeficiente* a_{12}*, por ejemplo). Ahora bien,* $|A| = a^2 - 4$ *y por tanto, el rango de la matriz depende de a. Sabemos que si* $|A| = 0$ *el rango es 1 y si* $|A| \neq 0$*, el rango será 2, es decir, si* $a^2 - 4 = 0$ *el rango es 1 y si* $a^2 - 4 \neq 0$ *el rango es 2. Hay que distinguir estos dos casos y seguir razonando a partir de aquí.*

Caso 1: $a^2 - 4 = 0$*, es decir, si* $a = 2$ *o si* $a = -2$ *(ojo con perder soluciones, ¡si* $a = -2$ *el rango es 1!). Entonces, el rango de la ampliada depende únicamente de b.*

Caso 1-*a*). *Si* $a = 2$ *entonces para que el sistema sea compatible tiene que ocurrir que* $\mathrm{rg}(A|b) = 1$*, luego* $\begin{vmatrix} 2 & 1 \\ 1 & b \end{vmatrix} = 0$ *y* $b = \frac{1}{2}$.

Por tanto si $a = 2$ *y* $b = \frac{1}{2}$*, entonces* $\mathrm{rg}(A) = \mathrm{rg}(A|b) = 1(< 2)$ *y el sistema es compatible indeterminado. Si* $a = 2$ *y* $b \neq \frac{1}{2}$*, entonces* $\mathrm{rg}(A) < \mathrm{rg}(A|b) = 2$ *y el sistema es incompatible.*

Caso 1-*b*). *Si* $a = -2$*, para que el sistema sea compatible tiene que ocurrir que* $\begin{vmatrix} -2 & 1 \\ 1 & b \end{vmatrix} = 0$ *y* $b = -\frac{1}{2}$.

Por tanto si $a = -2$ *y* $b = -\frac{1}{2}$*, entonces* $\mathrm{rg}(A) = \mathrm{rg}(A|b) = 1(< 2)$ *y el sistema es compatible indeterminado. Si* $a = -2$ *y* $b \neq -\frac{1}{2}$*, entonces* $\mathrm{rg}(A) < \mathrm{rg}(A|b) = 2$ *y el sistema es incompatible.*

En los casos compatibles, podemos obtener la solución resolviendo un sistema en el que ya no aparecen a y b. Nótese que al ser un sistema compatible indeterminado, la solución no será única.

Caso 2: $a^2 - 4 \neq 0$*, es decir, si* $a \neq 2$ *y* $a \neq -2$*. Entonces,* $2 = \mathrm{rg}(A) \leq \mathrm{rg}(A|b) < 3$ *y, por tanto,* $\mathrm{rg}(A) = \mathrm{rg}(A|b) = 2$ *para cualquier valor de b y el sistema es compatible determinado. La única solución, que depende de los valores concretos que tomen a y b, sería* $x = \frac{a-4b}{a^2-4}$*,* $y = \frac{ab-1}{a^2-4}$*. Nótese que esta solución tiene sentido porque el denominador,* $a^2 - 1$ *ya sabemos que en este caso nunca puede ser 0.*

3.3.2. Método de Gauss y método de Cramer

Método de Gauss. Consiste en determinar un sistema equivalente a $AX = B$ de forma que la matriz ampliada $(A : B)$ sea una matriz escalonada.

Ejemplo 3.3.6.

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

$$A=\begin{pmatrix} 2 & 1 & -1 \\ 1 & -2 & 2 \\ 2 & 1 & 1 \end{pmatrix} \quad B=\begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix} \quad X=\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

La matriz ampliada es:

$$(A:B)=\begin{pmatrix} 2 & 1 & -1 & : & -3 \\ 1 & -2 & 2 & : & 1 \\ 2 & 1 & 1 & : & 5 \end{pmatrix}$$

Se efectúan transformaciones elementales:

$$(A:B) \overset{F_1\leftrightarrow F_2}{\longrightarrow} \begin{pmatrix} 1 & -2 & 2 & : & 1 \\ 2 & 1 & -1 & : & -3 \\ 2 & 1 & 1 & : & 5 \end{pmatrix} \overset{\substack{F_2:=F_2-2F_1 \\ F_3:=F_2-2F_1}}{\longrightarrow} \begin{pmatrix} 1 & -2 & 2 & : & 1 \\ 0 & 5 & -5 & : & -5 \\ 0 & 5 & -3 & : & 3 \end{pmatrix}$$

$$\overset{F_2:=\frac{1}{5}F_2}{\longrightarrow} \begin{pmatrix} 1 & -2 & 2 & : & 1 \\ 0 & 1 & -1 & : & -1 \\ 0 & 5 & -3 & : & 3 \end{pmatrix} \overset{F_3:=F_3-2F_2}{\longrightarrow} \begin{pmatrix} 1 & -2 & 2 & : & 1 \\ 0 & 1 & -1 & : & -1 \\ 0 & 0 & 2 & : & 8 \end{pmatrix}$$

$$\overset{F_3:=\frac{1}{2}F_3}{\longrightarrow} \begin{pmatrix} 1 & -2 & 2 & : & 1 \\ 0 & 1 & -1 & : & -1 \\ 0 & 0 & 1 & : & 4 \end{pmatrix}$$

Se obtiene una matriz escalonada. Tenemos pues el siguiente sistema es equivalente:

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

De aquí, es inmediato que $z=4$*. De ahí,* $y-4=-1$*, es decir,* $y=3$ *y* $x-6+8=1$*, es decir,* $x=-1$*.*

Nota 3.3.7. *Si durante el proceso resulta alguna fila de la forma* $(0,0,...,0:b)$ *con* $b\neq 0$ *el sistema es incompatible.*

Ejercicio 3.3.8. *Estudiar y resolver por Gauss el sistema:*

$$\left.\begin{array}{rcr} x-\ y+2z & = & -3 \\ 2x\quad -\ z & = & 5 \\ 2y+3z & = & 3 \end{array}\right\}$$

La matriz ampliada es:

$$(A:B)=\begin{pmatrix} 1 & -1 & 2 & : & -3 \\ 2 & 0 & -1 & : & 5 \\ 0 & 2 & 3 & : & 3 \end{pmatrix}$$

donde $\mathrm{rg}(A)=3=\mathrm{rg}(A:B)$. *Por tanto, el sistema es compatible determinado.*

Un camino posible para resolver por Gauss sería:

$$\begin{pmatrix} 1 & -1 & 2 & : & -3 \\ 2 & 0 & -1 & : & 5 \\ 0 & 2 & 3 & : & 3 \end{pmatrix} \overset{F_2':=F_2-2F_1}{\longrightarrow} \begin{pmatrix} 1 & -1 & 2 & : & -3 \\ 0 & 2 & -5 & : & 11 \\ 0 & 2 & 3 & : & 3 \end{pmatrix}$$

$$\overset{F_3':=F_3-F_2}{\longrightarrow} \begin{pmatrix} 1 & -1 & 2 & : & -3 \\ 0 & 2 & -5 & : & 11 \\ 0 & 0 & 8 & : & -8 \end{pmatrix}$$

obteniendo el sistema equivalente:

$$\left.\begin{array}{rcr} x-\ y+2z & = & -3 \\ 2y-5z & = & 11 \\ 8z & = & -8 \end{array}\right\}$$

cuyas soluciones sonn $z=-1$, $y=3$, $x=2$.

Proposición 3.3.9. *Dada una matriz, hay una única matriz escalonada reducida equivalente por filas (mediante las operaciones elementales descritas).*

Definición 3.3.10. *Dada una matriz A, a la única matriz escalonada reducida equivalente se la llama* forma normal de Hermite *de la matriz A.*

Proposición 3.3.11. *Dos matrices son equivalentes por filas si y solo si tienen la misma forma normal de Hermite.*

Ejercicio 3.3.12. *Comprobar que son equivalentes por filas las matrices:*

$$A=\begin{pmatrix} 1 & -1 & 2 & -3 \\ 2 & 0 & -1 & 5 \\ 0 & 2 & 3 & 3 \end{pmatrix} \qquad B=\begin{pmatrix} 1 & -3 & -1 & -6 \\ 3 & -1 & 1 & 2 \\ 1 & 1 & 5 & 0 \end{pmatrix}$$

Sol.

$$H=\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Proposición 3.3.13. *Dos sistemas definidos por matrices equivalentes por filas tienen las mismas soluciones.*

Sistemas de Cramer: Un sistema lineal de n ecuaciones con n incognitas en el que la matriz del sistema es regular ($|A| \neq 0$) se denomina de Cramer. Un sistema de Cramer $\sum_{j=1}^{n} a_{ij}x_j = b_i$ es compatible determinado, es decir, tiene solución única.

La solución viene dada por las fórmulas de Cramer:

$$x_i = \frac{A_i}{|A|}$$

donde A_i es el determinante resultante de sustituir la i-ésima columna por la columna de coordenadas $(b_1, b_2, ..., b_n)$ del segundo miembro.

Ejemplo 3.3.14. *Resolvemos usando la fórmula de Cramer el siguiente sistema.*

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

La matriz de coeficientes es $A = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & -1 \end{pmatrix}$,

$B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad |A| = -3.$

$x_1 = \frac{\begin{vmatrix} 1 & -2 & 1 \\ 2 & 3 & -1 \\ 3 & 0 & -1 \end{vmatrix}}{-3} = \frac{-10}{-3} = \frac{10}{3}, \qquad x_2 = \frac{\begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & 3 & -1 \end{vmatrix}}{-3} = \frac{-1}{3},$

$x_3 = \frac{\begin{vmatrix} 1 & -2 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{vmatrix}}{-3} = \frac{9}{-3} = -3.$

El método de Cramer se puede adaptar para resolver también sistemas compatibles indeterminados que no son tipo Cramer. La idea es que las variables que “sobran” pueden situarse como parte de los términos independientes (a la derecha de la igualdad), dejando una matriz tipo Cramer como matriz del sistema. Una vez hecho esto, se resuelve normalmente por Cramer y el resultado es que las soluciones van a depender del valor de estas variables independientes. Veamos un ejemplo.

Ejemplo 3.3.15. *Para la resolución del ejemplo 3.3.4, podemos usar el método de Cramer si pasamos las variables dependientes (las que no están asociadas al menor no nulo de máximo orden) al término independiente. Así, suponiendo* $\mathrm{rg}(A) = \mathrm{rg}(A : B) = k$ *obtenemos* k *ecuaciones con* k *incógnitas:*

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

Ahora la matriz ampliada del sistema es $\begin{pmatrix} 1 & -2 & 3 & : & -9+2t \\ 0 & 1 & -1 & : & 5-2t \\ 1 & 0 & 2 & : & -t \end{pmatrix}$.

Por tanto:

$$x = \frac{\begin{vmatrix} -9+2t & -2 & 3 \\ 5-2t & 1 & -1 \\ -t & 0 & 2 \end{vmatrix}}{1} = 2 - 3t, \qquad y = \frac{\begin{vmatrix} 1 & -9+2t & 3 \\ 0 & 5-2t & -1 \\ 1 & -t & 2 \end{vmatrix}}{1} = 4 - t,$$

$$z = \frac{\begin{vmatrix} 1 & -2 & -9+2t \\ 0 & 1 & 5-2t \\ 1 & 0 & -t \end{vmatrix}}{1} = -1 + t.$$

Ejercicio 3.3.16. *Estudiar y resolver el sistema en función del parámetro* a*:*

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

Solución: si $a = 0$*, incompatible; si* $a \neq 0$*, compatible determinado con solución* $x = -\frac{1}{a}$*,* $y = -\frac{1}{a}$*,* $z = \frac{1}{a}$*.*

Ejercicio 3.3.17. *Estudiar y resolver el sistema en función del parámetro* a*:*

$$\left.\begin{array}{rcl} x + \quad (1-a)z & = & a \\ x + \quad az & = & 1 - a \\ 2x + ay - z & = & 3 \end{array}\right\}$$

Solución: si $a \neq 0, \frac{1}{2}$*, compatible determinaddo con solución* $x = 1$*,* $y = 0$*,* $z = -1$*; si* $a = 0$*, compatible indeterminado con solución* $x = 1$*,* $y = \alpha$ *con* $\alpha \in \mathbb{R}$*,* $z = -1$*; si* $a = \frac{1}{2}$*, compatible indeterminado con solución* $x = \frac{1-\alpha}{2}$*,* $y = 4 + 4\alpha$*,* $z = \alpha$ *con* $\alpha \in \mathbb{R}$*.*

3.3.3. Sistemas homogéneos

Un sistema se dice que es homogéneo si los términos independientes $b_1, ..., b_n$ son todos 0.

Un sistema homogeneo $AX = 0$ es siempre compatible pues el rango de la matriz A coincide con el de la ampliada. De hecho, $x_1 = 0, x_2 = 0, ..., x_n = 0$ es siempre una solución del sistema homogeneo y se denomina *solución trivial.*

Si $\mathrm{rg}(A) = n$ la solución trivial es la única del sistema.

Si $\mathrm{rg}(A) < n$ el sistema tiene infinitas soluciones siendo siempre una de ellas la solución trivial.

Ejemplo 3.3.18. *Resolver los sistemas:*

1)

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

$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & -1 & -2 \\ -1 & 2 & 1 \end{pmatrix}$, $\quad |A| = 12 \neq 0, \quad \mathrm{rg}(A) = 3$. *Por tanto sólo hay una solución que es la trivial:* $x = 0, \quad y = 0 \quad z = 0.$

2)

$$\left.\begin{array}{rcl} 7x-6y+15z & = & 0 \\ 3x+4y+z & = & 0 \end{array}\right\}$$

$A = \begin{pmatrix} 7 & -6 & 15 \\ 3 & 4 & 1 \end{pmatrix}$ $\quad \mathrm{rg}(A) = 2$. *Como el rango es menor que el número de incógnitas, el sistema admite soluciones distintas de la trivial. Podemos resolverlo por Cramer pasando la variable* z *al término independiente. Resolvemos el sistema:*

$$\left.\begin{array}{rcl} 7x-6y & = & -15z \\ 3x+4y & = & -z \end{array}\right\}$$

Resolviendo: $A = \begin{pmatrix} 7 & -6 \\ 3 & 4 \end{pmatrix}$ $\quad |A| = 46.$

$x = \frac{-66}{46}\alpha, \quad y = \frac{38}{46}\alpha, \quad z = \alpha.$

Ponemos presentar las variables x, y *en función de* z *o mejor, como hacemos aquí, presentar las tres variables en función de un parámetro independiente* α.

Capítulo 4

Espacios vectoriales

4.1. Definición y propiedades

Se llama espacio vectorial sobre $\mathbb{R}$ a un conjunto E, no vacío, dotado de dos aplicaciones:

Suma:

$+\colon E \times E \to E$ tal que $\forall\, x, y, z \in E$ se cumplen las siguientes propiedades:

a) Asociativa: $x + (y + z) = (x + y) + z$,

b) Conmutativa: $x + y = y + x$,

c) Existe elemento neutro, $0_E \in E$, que verifica que $x + 0_E = x$,

d) Todo elemento $x \in E$ posee un opuesto, $-x$, tal que $x + (-x) = 0$.

Producto (de un número por un vector):

$*\colon \mathbb{R} \times E \to E$ tal que $\forall x, y \in E$ y $\forall \lambda, \beta \in \mathbb{R}$ se cumple que:

a) $\lambda * (x + y) = \lambda * x + \lambda * y$,

b) $(\lambda + \beta) * x = \lambda * x + \beta * x$,

c) $\lambda(\beta * x) = (\lambda\beta) * x$,

d) $1 * x = x$.

Los elementos de E se denominan **vectores** y los de $\mathbb{R}$ **escalares**.

Ejemplo 4.1.1. *Se puede comprobar que los siguientes espacios con las operaciones indicadas tienen estructura de espacio vectorial.*

1) $\mathbb{R}^n$ *con la suma* $(a_1, ..., a_n) + (b_1, ..., b_n) = (a_1 + b_1, ..., a_n + b_n)$ *y el producto* $\lambda * (a_1, ..., a_n) = (\lambda \cdot a_1, ..., \lambda \cdot a_n)$.

2) El conjunto $\mathcal{M}_{m\times n}$ *con la suma entre matrices y el producto entre un número y una matriz.*

3) El conjunto $\mathbb{R}[x]$ *de polinomios en una variable,* x*, con la suma y el producto por un número:* $\lambda * (a_0x^n + \cdots + a_n) = \lambda \cdot a_0x^n + \cdots + \lambda \cdot a_n$.

4) Las ecuaciones lineales con n *variables* $x_1, ..., x_n$*,* $a_1x_1 + \cdots + a_nx_n = b$*, con la suma* $[a_1x_1 + \cdots + a_nx_n = b] + [a'_1x_1 + \cdots + a'_nx_n = b'] = [(a_1 + a'_1)x_1 + \cdots + (a_n + a'_n)x_n = (b + b')]$ *y el producto por un número* $\lambda * [a_1x_1 + \cdots + a_nx_n = b] = (\lambda \cdot a_1)x_1 + \cdots + (\lambda \cdot a_n)x_n = (\lambda \cdot b)$.

Propiedades.

$\forall x, y \in E$ y $\forall \lambda, \beta \in \mathbb{R}$ las siguientes propiedades se deducen de la definición:

1) $0 * x = 0_E$.

2) $(-\lambda) * x = \lambda * (-x) = -(\lambda * x)$.

3) $\lambda * (x - y) = \lambda * x - \lambda * y$.

4) $(\lambda - \beta) * x = \lambda * x - \beta * x$

5) Si $\lambda * x = 0_E$ entonces $\lambda = 0$ ó $x = 0_E$.[1]

6) Sean $x_1, x_2, ..., x_n$ n vectores del espacio vectorial E. Dados n números cualesquiera $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$, el elemento $x = \lambda_1x_1 + \lambda_2x_2 + \cdots + \lambda_nx_n = \sum_{i=1}^n \lambda_ix_i$, es un elemento de E. Decimos que x es **combinación lineal** de $x_1, x_2, ..., x_n$.[2]

Ejemplo 4.1.2. *En el espacio vectorial* $\mathbb{R}^2$*, el vector* $(-1, 6)$ *es combinación lineal de los vectores* $(3, 4)$ *y* $(-2, 1)$ *ya que* $(-1, 6) = 1 \cdot (3, 4) + 2 \cdot (-2, 1)$.

Ejercicio 4.1.3. *Escribir un ejemplo concreto de combinación lineal con vectores de* $\mathbb{R}^2$ *o* $\mathbb{R}^3$.

[1] Nótese que esto significa que λ es el número 0 ó x es el vector 0, que son cosas diferentes.

[2] Por abreviar, escribimos λ_ix_i en lugar de $\lambda_i * x_i$

4.2. Dependencia e independencia lineal

Se dice que el conjunto de vectores $v_1, ..., v_n \in E$ es **linealmente independiente** si la relación $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_E$ se verifica únicamente cuando $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

En caso contrario, es decir, cuando la relación anterior se cumple con algún α_i no nulo se dice que el conjunto de vectores es linealmente dependiente. También decimos simplemente que los vectores son linealmente independientes o linealmente dependientes.

Ejemplo 4.2.1. *1) Los vectores* $(1,0)$ *y* $(0,1)$ *son linealmente independientes: si* $\alpha_1(1,0) + \alpha_2(0,1) = (0,0)$ *entonces* $(\alpha_1, 0) + (0, \alpha_2) = (\alpha_1, \alpha_2) = (0,0)$ *y, por tanto,* $\alpha_1 = \alpha_2 = 0$.

2) Los vectores $(2,-3)$ *y* $(-4,6)$ *son linealmente dependientes. Planteamos la ecuación* $\alpha_1(2,-3) + \alpha_2(-4,6) = (0,0)$ *y comprobamos que tiene soluciones distintas de* $\alpha_1 = \alpha_2 = 0$.

$$\left.\begin{array}{rcl} 2\alpha_1 - 4\alpha_2 & = & 0 \\ -3\alpha_1 + 6\alpha_2 & = & 0 \end{array}\right\}$$

El conjunto de soluciones es $\alpha_1 = 2\alpha_2$ *luego no son necesariamente 0.*

Ejercicio 4.2.2. *Comprobar si los siguientes conjuntos de vectores son linealmente independientes:*

1) El conjunto de vectores $\{e_1, e_2, ..., e_n\}$ *de* $\mathbb{R}^n$ *con* $e_1 = (1,0,...,0)$, $e_2 = (0,1,0,...,0)$,...,$e_n = (0,...,0,1)$.

2) Los vectores $(1,-2,1)$, $(2,0,1)$, $(0,1,-1)$ *y* $(1,-1,0)$ *en* $\mathbb{R}^3$.

Proposición 4.2.3. *Si un conjunto de vectores es linealmente independiente entonces cualquier subconjunto suyo lo es.*

4.3. Base de un espacio vectorial

Un conjunto de n vectores $v_1, v_2, ..., v_n$ de un espacio vectorial E es un **sistema de generadores** o **sistema generador** de E si todo vector de E es combinación lineal de $v_1, v_2, ..., v_n$, es decir, si $\forall w \in E \;\; \exists \alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R}$ tales que $w = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$.

En este caso, se dice que los vectores $v_1, v_2, ..., v_n$ generan el espacio vectorial E.

Ejemplo 4.3.1. *Los vectores* $\{(1,0),(0,1)\}$ *forman un sistema generador de* $\mathbb{R}^2$. *Comprobamos que, en efecto, para todo vector* $(x,y) \in \mathbb{R}^2$ *existen dos números reales, en este caso* $\alpha_1 = x$, $\alpha_2 = y$ *de forma que* $(x,y) = x(1,0) + y(0,1)$.

Nota 4.3.2. *Nótese que un espacio vectorial no está generado de forma única. Por ejemplo los vectores* $(1,0)$ *y* $(0,1)$ *forman un sistema generador de* $\mathbb{R}^2$ *como también lo forman los vectores* $(1,1)$ *y* $(1,-3)$ *o cualquier otro par de vectores linealmente independientes.*

Pregunta 4.3.3. *¿Es* $\{(2,1),(3,-2),(0,4)\}$ *un sistema generador de* $\mathbb{R}^2$*?*

Pregunta 4.3.4. *¿Es* $\{(3,-1,2),(1,0,-1)\}$ *un sistema generador de* $\mathbb{R}^3$*?*

Definición 4.3.5. *Un conjunto de vectores* $\{v_1, v_2, ..., v_n\}$ *es* ***base*** *de un espacio vectorial* E *si verifica:*

a) $\{v_1, v_2, ..., v_n\}$ *es un sistema generador de* E,

b) $\{v_1, v_2, ..., v_n\}$ *es linealmente independiente.*

Ejemplo 4.3.6. *1)* $\{(1,0),(0,1)\}$ *es una base de* $\mathbb{R}^2$,

2) $\{(1,0),(0,1),(1,1)\}$ *no es una base de* $\mathbb{R}^2$ *ya que es un sistema generador pero no linealmente independiente,*

3) $\{(3,-1,2),(1,0,-1)\}$ *no es una base de* $\mathbb{R}^3$ *ya que es linealmente independiente pero no un sistema generador.*

Pregunta 4.3.7. *¿Es* $\{(1,1,1),(0,1,1),(0,0,1)\}$ *un sistema generador de* $\mathbb{R}^3$*?*

Observación 4.3.8. *En una base, el orden de los vectores importa. Por ejemplo,* $\{(1,0),(0,1)\}$ *y* $\{(0,1),(1,0)\}$ *son dos bases distintas de* $\mathbb{R}^2$.

Teorema 4.3.9 (Existencia de base)**.** *Un espacio vectorial* E *generado por un número finito de elementos siempre admite una base.*

Nótese, sin embargo, que la base no tiene por qué ser única. Por ejemplo, $\{(1,0),(0,1)\}$ y $\{(1,2),(3,-1)\}$ son distintas bases de $\mathbb{R}^2$. De hecho, cualquier familia de n vectores linealmente independiente es base de $\mathbb{R}^n$.

Teorema 4.3.10 (Unicidad de la expresión)**.** *Si* $\{v_1, v_2, ..., v_n\}$ *es una base de* E *entonces cada vector* $w \in E$ *puede expresarse de modo único como combinación lineal de los vectores de la base, es decir, existe una única familia de números* $\alpha_1, \alpha_2, ..., \alpha_n$ *tal que* $w = \sum_{i=1}^{n} \alpha_i v_i$.

Demostración. Supongamos $w = \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \beta_i v_i$. Entonces, $\sum_{i=1}^n \alpha_i v_i - \sum_{i=1}^n \beta_i v_i = 0$. Operando, $\sum_{i=1}^n)(\alpha_i - \beta_i)v_i = 0$ y por ser $\{v_1, v_2, ..., v_n\}$ linealmente independiente, $\alpha_i - \beta_i = 0\ \forall\, i$, es decir, $\alpha_i = \beta_i\ \forall\, i$ y, por tanto, la expresión es única. □

Ejercicio 4.3.11. *Comprobar que, en efecto, hay un único modo de expresar el vector* $(2, 1)$ *como combinación lineal de los vectores* $(1, 3)$ *y* $(2, -1)$*.*

Dada una base de un espacio vectorial $\{v_1, v_2, ..., v_n\}$ y un vector expresado, de forma única, como combinación lineal de dichos vectores $w = \sum_{i=1}^n \alpha_i v_i$, los escalares $\alpha_1, ..., \alpha_n$ se llaman **coordenadas del vector** w **en la base** $\{v_1, v_2, ..., v_n\}$.

Se denota como $w = (\alpha_1, ..., \alpha_n)_{\{v_1, v_2, ..., v_n\}}$ o, si no hay ambigüedad respecto a la base, simplemente $(\alpha_1, ..., \alpha_n)$.

Ejemplo 4.3.12. *Consideramos en* $\mathbb{R}^2$ *un vector* v *que, como combinación lineal de los vectores* $(1, 0), (0, 1)$ *se expresa como* $v = 2\cdot(1, 0)+5\cdot(0, 1)$*. Así, las coordenadas del vector* v *en la base* $\{(1, 0), (0, 1)\}$ *serían* $v = (2, 5)_{\{(1,0),(0,1)\}}$*.*

Si consideramos otra base de $\mathbb{R}^2$*, por ejemplo* $(1, 0), (1, 1)$*, entonces el vector* $v = 2 \cdot (1, 0) + 5 \cdot (0, 1)$ *como combinación lineal de* $(1, 0), (1, 1)$ *sería* $v = -3(1, 0) + 5(1, 1)$ *y, por tanto, las coordenadas son* $(-3, 5)_{\{(1,0),(1,1)\}}$*.*

Nótese que las coordenadas del mismo vector dependen de la base. Si no se indica ninguna base, supondremos que se trata de la base canónica: $e_1, e_2, ..., e_n$, esto es, $(1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1)$.

Ejercicio 4.3.13. *Hallar las coordenadas del vector* $(3, -2, 1)_{\{e_1,e_2,e_3\}}$ *respecto a la base* $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$*.*

Planteamos la combinación lineal

$$(3, -2, 1) = \alpha_1(1, 1, 0) + \alpha_2(1, 0, 1) + \alpha_3(0, 1, 1)$$

de donde obtenemos un sistema:

$$\left.\begin{array}{rcl} 3 & = & \alpha_1 + \alpha_2 \\ -2 & = & \alpha_1 \qquad + \alpha_3 \\ 1 & = & \qquad \alpha_2 + \alpha_3 \end{array}\right\}$$

Resolviendo, obtenemos los valores $\alpha_1 = 0$*,* $\alpha_2 = 3$*,* $\alpha_3 = -2$*. El vector es el* $(0, 3, -2)_{\{(1,1,0),(1,0,1),(0,1,1)\}}$*.*

4.4. Dimensión de un espacio vectorial

Proposición 4.4.1. *Todas las bases de un espacio vectorial tienen el mismo número de elementos.*

Definición 4.4.2. *Se llama **dimensión** de un espacio vectorial E al número de elementos n de una de sus bases. Se denota* $\dim(E) = n$.

Es fácil ver que, según esta definición, la dimensión de $\mathbb{R}^n$ es n.

Pregunta 4.4.3. *¿Cuál es la dimensión del espacio de las matrices cuadradas de orden 2?*

4.5. Subespacios vectoriales

Definición 4.5.1. *Un conjunto F no vacío de un espacio vectorial E es un **subespacio vectorial** si verifica las siguientes propiedades:*

i) $\forall\, v, w \in F,\ v + w \in F.$

ii) $\forall\, v \in F,\ \forall\, \lambda \in \mathbb{R}\ \lambda v \in F.$

ó lo que es equivalente:

i') $\forall\, v, w \in F,\ \forall\, \lambda \in \mathbb{R}\ \lambda v + \mu w \in F.$

- Dado un espacio vectorial E, si definimos $F = \{0_E\}$, F es un subespacio vectorial.
- También es inmediato comprobar que $F = E$ cumple la definición de ser subespacio vectorial.
- Sea $\{v_1, ..., v_p\}$ un conjunto de vectores de un espacio vectorial E. Sea $F = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p \,|\, \alpha_1, \alpha_2, ..., \alpha_p \in \mathbb{R}\}$ el conjunto de todas las **combinaciones lineales de dichos vectores**, entonces F es un subespacio vectorial de E. En este caso, se dice que F es el **subespacio vectorial de E generado por** $\{v_1, ..., v_p\}$. Se denota $\langle v_1, \ldots, v_p\rangle$ o $\mathrm{L}[v_1, \ldots, v_p]$.

Definición 4.5.2. *Se llama **rango** de un conjunto de vectores a la dimensión del subespacio vectorial que genera.*

Proposición 4.5.3. *El rango de un conjunto de vectores es el número máximo de vectores linealmente independientes que posee.*

Ejemplo 4.5.4. *Determinar el subespacio vectorial F de $\mathbb{R}^3$ generado por* $\{(1,0,0),(0,1,0),(1,1,0)\}$*. Podemos aplicar directamente la definición y considerar el conjunto de todas las combinaciones lineales:*

$$F = \{\alpha_1(1,0,0) + \alpha_2(0,1,0) + \alpha_3(1,1,0) \mid \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\}.$$

Por tanto, $F = \{(\alpha_1 + \alpha_3, \alpha_2 + \alpha_3, 0) \mid \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\}$*. A partir de aquí, podemos mejorar la presentación del resultado cambiando los parámetros de referencia. Nótese que al poder asignar cualquier valor a los parámetros, el valor de uno de ellos, digamos* α_3*, es irrelevante. El espacio sería* $F = \{(\alpha_1, \alpha_2, 0) \mid \alpha_1, \alpha_2 \in \mathbb{R}\}$.

Calcular el rango de un conjunto de p vectores de un espacio vectorial E de dimensión n se reduce a calcular el rango de la matriz formada por los p vectores de n componentes:

Ejemplo 4.5.5. *Calculamos el rango del conjunto* $\{(1,0,0),(0,1,0),(1,1,0)\}$. *Formamos una matriz con estos vectores:*
$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$ *Es inmediato comprobar que el rango de esta matriz es 2. Por tanto, la dimensión del espacio vectorial que generan, F, es 2. Como hemos visto, necesitamos únicamente dos parámetros para expresar el conjunto F. El número de parámetros necesarios es también igual a la dimensión.*

Dado que todo subespacio de $\mathbb{R}^n$ debe tener dimensión finita, estará generado por una familia finita de vectores. Así, si $F = \mathrm{L}[v_1, \ldots, v_m]$, con $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})$, tenemos que todos los vectores $(x_1, x_2, \ldots, x_n) \in F$ son de la forma:

$$(x_1, x_2, \ldots, x_n) = \alpha_1(v_{11}, v_{12}, \ldots, v_{1n}) + \alpha_2(v_{21}, v_{22}, \ldots, v_{2n}) + \cdots + \alpha_m(v_{m1}, v_{m2}, \ldots, v_{mn}),$$

con $\alpha_1, \alpha_2, ..., \alpha_m \in \mathbb{R}$. Si expresamos esta igualdad coordenada a coordenada, obtenemos

$$\left.\begin{array}{c} v_{11}\alpha_1 + v_{22}\alpha_2 + \cdots + v_{m1}\alpha_m = x_1 \\ v_{12}\alpha_1 + v_{22}\alpha_2 + \cdots + v_{m2}\alpha_m = x_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ v_{1n}\alpha_1 + v_{2n}\alpha_2 + \cdots + v_{mn}\alpha_m = x_n \end{array}\right\} \tag{4.1}$$

que se denominan *ecuaciones paramétricas* del subespacio. Un subespacio presentado así, decimos que está en forma *paramétrica*.

Proposición 4.5.6. *Sea*

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}\right\}$$

un sistema homogéneo compatible de m *ecuaciones con* n *incógnitas* $x_1, x_2, ..., x_n$.

Sea $F = \{(s_1, s_2, ..., s_n) \mid x_1 = s_1, x_2 = s_2, ..., x_n = s_n$ *es solución al sistema*$\} \subset \mathbb{R}^n$ *y sea* k *el rango del sistema, es decir, el número máximo de ecuaciones linealmente independientes. Entonces:*

- F *es subespacio vectorial,*
- $\dim(F) = n - k$.

El recíproco, también es cierto.

Proposición 4.5.7. *Todo subespacio vectorial* F *de* $\mathbb{R}^n$ *es solución de un sistema lineal homogéneo.*

Decimos que un subespacio vectorial F está definido en forma *implícita* si F es el conjunto de soluciones de un sistema lineal homogéneo. Estas ecuaciones se llaman *ecuaciones implícitas* del subespacio.

Si resovemos un sistema lineal homogéneo que sea compatible indeterminado (en otro caso, la solución única es que todas las variables sean cero y no necesitamos ningún parámetro para el conjunto de soluciones), podemos obtener su forma paramétrica y dar una base del subespacio.

Ejemplo 4.5.8. *Sea* F *el subespacio dado por las soluciones del sistema*

$$\left.\begin{array}{c} x + y + z + t = 0 \\ x + y - z - t = 0 \end{array}\right\} \tag{4.2}$$

Utilizando el método de Gauss, por ejemplo, es inmediato ver que este sistema es equivalente a

$$\left.\begin{array}{c} x = -y \\ z = -t \end{array}\right\} \tag{4.3}$$

Para expresarlo en forma paramétrica necesitamos tantos parámetros como la dimensión del subespacio (nótese que el sistema tiene rango 2 y que necesitamos 4-2 parámetros, ver proposición 4.5.6), en este caso, dos. Una forma sencilla (no la única) de dar una solución paramétrica a este sistema *sería suponer $y = \alpha$, $t = \beta$ ya que las otras variables ya las tenemos despejadas en función de estas. Por tanto, en forma paramétrica nos quedaría:*

$$F = \{(-\alpha, \alpha, -\beta, \beta) \,:\, \alpha,\, \beta \in \mathbb{R}\}.$$

Si ahora quisiéramos dar una base para este espacio, una forma sencilla de hacerlo (por supuesto, hay infinitas bases posibles, esta es solo una de ellas) es tomar un vector por cada parámetro haciendo 1 ese parámetro y 0 el resto. Así,

$$F = \mathrm{L}[(-1,1,0,0),(0,0,-1,1)].$$

Al revés, también es sencillo pasar de la base o de la forma paramétrica a la forma implícita. Para ello, partimos de un sistema en el que expresamos cada variable a partir de su expresión paramétrica, como vimos en (4.1). En este sistema, consideramos que los parámetros (α_i) son las incógnitas y escalonamos por el método de Gauss la matriz correspondiente. Una vez escalonado, las ecuaciones implícitas del subespacio serán aquellas en las que no aparezcan ya parámetros.

Ejemplo 4.5.9. *Tomemos las soluciones del ejemplo 4.5.8 y supongamos que lo que sabemos es que $F = \mathrm{L}[(-1,1,0,0),(0,0,-1,1)]$. Por tanto, si consideramos las combinaciones lineales de estos dos vectores de su base, obtenemos que*

$$F = \{\alpha(-1,1,0,0) + \beta(0,0,-1,1) \,:\, \alpha,\, \beta \in \mathbb{R}\}.$$

Si expresamos esta igualdad coordenada a coordenada, obtenemos el sistema:

$$\left.\begin{array}{r} -\alpha = x \\ \alpha = y \\ -\beta = z \\ \beta = t \end{array}\right\} \tag{4.4}$$

Aplicamos Gauss a este sistema:

$$\begin{pmatrix} -1 & 0 & : & x \\ 1 & 0 & : & y \\ 0 & -1 & : & z \\ 0 & 1 & : & t \end{pmatrix} \xrightarrow{F_1 \leftrightarrow F_2,\ F_3 \leftrightarrow F_4,\ F_2 \leftrightarrow F_3} \begin{pmatrix} 1 & 0 & : & y \\ 0 & 1 & : & t \\ -1 & 0 & : & x \\ 0 & -1 & : & z \end{pmatrix}$$

$$\xrightarrow{F_3' := F_3 + F_1,\ F_4' := F_4 + F_2,} \begin{pmatrix} 1 & 0 & : & y \\ 0 & 1 & : & t \\ 0 & 0 & : & x+y \\ 0 & 0 & : & z+t \end{pmatrix}$$

Observamos que las filas en las que no aparecen los parámetros α y β son las dos últimas. Por tanto, la forma implícita del subespacio sería como solución del sistema:

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

Nótese que este sistema es equivalente al sistema de partida del ejemplo 4.5.8. En ambos casos, es inmediato ver que su forma de Hermite es

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

Por último, caber mencionar que el número de ecuaciones que nos van a salir será n menos el rango de la matriz de este sistema de parámetros. Si dicha matriz tiene rango máximo (esto es, igual al número de parámetros), en particular, si los parámetros vienen de una base del subespacio como en el ejercicio anterior, entonces el número de ecuaciones en la forma implícita será n menos el número de parámetros (ver proposición 4.5.6).

Ejercicio 4.5.10. *Dar las ecuaciones implícitas del subespacio $G = \mathrm{L}[(1,0,2),(3,1,0)]$ de $\mathbb{R}^3$.*

En forma paramétrica obtenemos:

$$\left.\begin{array}{rcrcl} \alpha &+& 3\beta &=& x \\ && \beta &=& y \\ 2\alpha &&&=& z \end{array}\right\}$$

Dado que la matriz del sistema tiene rango 2, podemos adelantar que al escalonar por Gauss nos va a quedar una única ecuación en la que no aparecen parámetros. Escalonamos por Gauss:

$$\begin{pmatrix} 1 & 3 & : & x \\ 0 & 1 & : & y \\ 2 & 0 & : & z \end{pmatrix} \overset{F'3:=F_3-2F_1+6F_2}{\longrightarrow} \begin{pmatrix} 1 & 3 & : & x \\ 0 & 1 & : & y \\ 0 & 0 & : & z-2x+6y \end{pmatrix}$$

Por tanto, la ecuación implícita de G es $z-2x+6y=0$.

4.6. Intersección y suma de subespacios vectoriales

Dados dos subespacios vectoriales F, G de un espacio vectorial E, se llama **intersección** de F y G y se denota $F \cap G$ al conjunto:

$$F \cap G = \{v \in E \mid v \in F \text{ y } v \in G\}.$$

Proposición 4.6.1. *La intersección de dos subespacios vectoriales de E es otro subespacio vectorial de E.*

Observación 4.6.2. *Si tenemos los subespacios F y G expresados en forma implícita por sendos sistemas, los puntos del subespacio $F \cap G$ son exactamente los que satisfacen ambos sistemas. Por tanto, la expresión implícita del subespacio $F \cap G$ viene dada por el sistema con todas las ecuaciones de los dos sistemas.*

En cambio, la unión de subespacios vectoriales en el sentido conjuntista no es un subespacio vectorial.

Ejemplo 4.6.3. *Sean F y G los subespacios de $\mathbb{R}^2$ generados por los vectores $(1,0)$ y $(0,1)$ respectivamente, es decir, $F = \{(x,0) \,|\, x \in \mathbb{R}\}$ y $G = \{(0,y) \,|\, y \in \mathbb{R}\}$.*

Si $F \cup G$ fuera subespacio, el vector $(1,0) + (0,1) = (1,1)$ debería pertenecer a $F \cup G$ pero $(1,1)$ no pertenece ni a F ni a G.

Dados dos subespacios vectoriales F, G de un espacio vectorial E, se llama **suma** de F y G y se denota $F + G$ al conjunto:

$$F + G = \{v + w \,|\, v \in F \text{ y } w \in G\}.$$

Proposición 4.6.4. *El conjunto $F + G$ es un subespacio.*

Proposición 4.6.5. *Si tenemos dos subespacios $F = \mathrm{L}[v_1, \ldots, v_m]$ y $G = \mathrm{L}[w_1, \ldots, w_k]$, entonces*

$$F + G = \mathrm{L}[v_1, \ldots, v_m, w_1, \ldots, w_k].$$

Teorema 4.6.6. *Sean F y G dos subespacios vectoriales de un espacio vectorial E. Entonces:*

$$\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G).$$

Ejemplo 4.6.7. *Consideramos $F = \mathrm{L}[(1,0,0)]$ y $G = \mathrm{L}[(1,1,0),(0,1,0)]$.*

ES fácil comprobar que:

$$F \cap G = \mathrm{L}[(1,0,0)],$$

$$F + G = \mathrm{L}[(1,0,0),(1,1,0),(0,1,0)] = \mathrm{L}[(1,1,0),(0,1,0)] = G.$$

4.7. Matriz del cambio de base de un espacio vectorial

Como hemos visto, las coordenadas de un vector dependen de la base elegida. Fijadas dos bases, B_1, B_2, pretendemos calcular una matriz tal que dado un vector expresado mediante coordenadas respecto a la base B_1 nos permita calcular de forma automática las coordenadas del vector respecto a la base B_2. Esta matriz es la **matriz del cambio de base**.

Sea E un espacio vectorial de dimensión n. Sea $B = \{e_{1,}e_2, \ldots, e_n\}$ una base de E, en la que las coordenadas de un vector $x \in E$ son $(x_{1,}x_2, \ldots, x_n)$.

Sea $B' = \left\{e'_{1,}e'_2, \ldots, e'_n\right\}$ otra base de E y sea $x = (x'_1, x'_2, \ldots, x'_n)_{B'}$.

Los vectores de B puestos en función de los de B' son:.

$$e_1 = a_{11}e'_1 + a_{21}e'_2 + \cdots + a_{n1}e'_n$$

$$e_2 = a_{12}e'_1 + a_{22}e'_2 + \cdots + a_{n2}e'_n$$

....

$$e_n = a_{1n}e'_1 + a_{2n}e'_2 + \cdots + a_{nn}e'_n$$

que también se puede expresar como:

$$e_j = \sum_{i=1}^{n} a_{ij}e'_i \text{ con } j = 1, \ldots, n.$$

Matricialmente, el cambio de base resulta ser:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_i \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ & & \cdots & \cdots & \cdots & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ & & \cdots & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}$$

$A = (a_{ij})$ es la matriz cambio de base de B a B'. Las columnas de A son los vectores $e_1, e_2, ..., e_n$ expresados en la base B'.

Como las columnas de A la forman los vectores de una base, que son linealmente independientes, el rango de A es máximo luego el determinante es no nulo y A es inversible. Así pues,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} = A^{-1} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_i \\ \vdots \\ x'_n \end{pmatrix}$$

Ejemplo 4.7.1. *Consideramos el espacio vectorial* $\mathbb{R}^2$ *y las bases* $B = \{(1,2),(2,1)\}$ *y* $B' = \{(1,0),(0,1)\}$. *Vamos a calcular las coordenadas del vector* $x = (3,4)_B$ *respecto a la base* B'.

Primero, calculamos las coordenadas de los vectores de la base B*:*

$$(1,2) = a_{11}(1,0) + a_{21}(0,1) \text{ de donde } a_{11} = 1 \text{ y } a_{21} = 2.$$

$$(2,1) = a_{12}(1,0) + a_{22}(0,1) \text{ de donde } a_{12} = 2 \text{ y } a_{22} = 1.$$

La matriz cambio de base será: $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$

La relación entre las coordenadas viene dada por la expresión:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$$

resultando que $(3,4)_B = (11,10)_{B'}$.

Si, por el contrario, queremos obtener las coordenadas del vector $x = (3,4)_{B'}$ *respecto a la base* B, *podemos hacer otra vez el procedimiento o calcular la inversa de la matriz obtenida o, simplemente, resolver el sistema:*

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

es decir,

$$\left.\begin{array}{rcl} x_1 + 2x_2 & = & 3 \\ 2x_1 + \ x_2 & = & 4 \end{array}\right\}$$

Resolviendo, $x_1 = \frac{5}{3}$ *y* $x_2 = \frac{2}{3}$. *Por tanto, el vector* $x = (3,4)_{B'}$ *expresado en coordenadas respecto de* B *es* $x = (\frac{5}{3}, \frac{1}{2})_B$.

Ejemplo 4.7.2. *Dar la matriz del cambio de base de* $B = \{(0,1,1),(1,1,1),(3,1,0)\}$ *a* $B' = \{(1,0,1),(-1,0,2),(0,2,5)\}$.

$$(0,1,1) = a_{11}(1,0,1) + a_{21}(-1,0,2) + a_{31}(0,2,5)$$

$$(1,1,1) = a_{12}(1,0,1) + a_{22}(-1,0,2) + a_{32}(0,2,5)$$

$$(3,1,0) = a_{13}(1,0,1) + a_{23}(-1,0,2) + a_{33}(0,2,5)$$

De cada línea obtenemos un sistema de tres ecuaciones con tres incógnitas a_{i1}, a_{i2}, a_{i3}. *Solucionando los sistemas tenemos que la matriz cambio de base es:*

$$A = \begin{pmatrix} -\frac{1}{2} & \frac{1}{6} & \frac{7}{6} \\ -\frac{1}{2} & -\frac{5}{6} & -\frac{11}{6} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Ejercicio 4.7.3. *Determinar la matriz cambio de base de* $B = \{(-1,2),(3,1)\}$ *a* $B' = \{(1,-1),(0,2)\}$.

Solución:

$$\begin{pmatrix} -1 & 3 \\ \frac{1}{2} & 2 \end{pmatrix}$$

4.8. Producto escalar y norma

Definición 4.8.1. *Sean* $x = (x_1, x_2, ..., x_n), y = (y_1, y_2, ..., y_n)$ *dos vectores de* $\mathbb{R}^n$. *Se denomina* ***producto escalar*** *de* x *e* y, $x \cdot y$, *al* número real*:* $x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n$.

Definición 4.8.2. *Se dice que dos vectores* x, y *son* ***ortogonales*** *cuando su producto escalar es 0,* $x \cdot y = 0$.

Definición 4.8.3. *Una base de un espacio vectorial* $\{e_1, e_2, ..., e_n\}$ *es* ortogonal *si* $e_i \cdot e_j = 0 \; \forall \, i \neq j$ *(todos los vectores son ortogonales entre sí).*

Ejemplo 4.8.4. *La base canónica de* $\mathbb{R}^n$ *es ortogonal.*

Definición 4.8.5. *Dado un vector* $x = (x_1, x_2, ..., x_n)$ *de* $\mathbb{R}^n$, *se denomina* ***norma*** *o* ***módulo*** *de* x, $||x||$, *al número real no negativo:*

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

Ejercicio 4.8.6. *Comprobar que dado un vector* x *de* $\mathbb{R}^n$, $||x|| = \sqrt{x \cdot x}$.

Dado un vector x no nulo, llamamos **normalizar** el vector a obtener un vector de la misma dirección y sentido que x pero con norma 1. Para normalizar un vector basta dividirlo entre su norma. Es inmediato comprobar que para cualquier vector no nulo $x = (x_1, x_2, ..., x_n)$, si consideramos $y = \frac{1}{||x||}(x_1, x_2, ..., x_n) = (\frac{x_1}{||x||}, \frac{x_1 2}{||x||}, ..., \frac{x_n}{||x||})$, entonces

$$||y|| = \sqrt{(\frac{x_1}{||x||})^2 + (\frac{x_2}{||x||})^2 + \cdots + (\frac{x_n}{||x||})^2} = \frac{\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}}{||x||} = \frac{||x||}{||x||} = 1.$$

Definición 4.8.7. *Una base de un espacio vectorial* $\{e_1, e_2, ..., e_n\}$ *es* orto-normal *si es ortogonal y además* $||e_i|| = 1\,\forall i$ *(esto es, si es ortogonal y todos los vectores tienen norma 1).*

Ejemplo 4.8.8. *La base canónica de* $\mathbb{R}^n$ *es ortonormal.*

Capítulo 5

Aplicaciones lineales

5.1. Definición y propiedades

Sean E y F dos espacios vectoriales sobre $\mathbb{R}$. Se llama **aplicación lineal** de E en F a toda aplicación f de E en F tal que:

a) $\forall\, x, y \in E \quad f(x+y) = f(x) + f(y)$

b) $\forall\, \lambda \in \mathbb{R} \quad f(\lambda x) = \lambda f(x)$

o, lo que es equivalente:

a') $\forall\, x, y \in E, \forall\, \lambda, \mu \in \mathbb{R} \qquad f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$

Ejercicio 5.1.1. *Comprobar si las siguientes aplicaciones son lineales:*

1) $f\colon \mathbb{R} \to \mathbb{R}$ *tal que* $f(x) = 3x$.

2) $f\colon \mathbb{R} \to \mathbb{R}$ *tal que* $f(x) = 2x + 3$.

3) $f\colon \mathbb{R}^2 \to \mathbb{R}$ *tal que* $f(x, y) = 3x - 2y$.

4) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ *tal que* $f(x, y) = (x^2, x - y)$.

Propiedades.
Dada una aplicación lineal $f\colon E \to F$:

1) $f(0_E) = 0_F$

2) $f(-x) = -f(x)$.

3) $f(x - y) = f(x) - f(y)$.

4) Si el conjunto de vectores de $\{x_1, ..., x_n\} \subset E$ es linealmente dependiente (l. d.), también lo es el conjunto $\{f(x_1), ..., f(x_n)\} \subset F$.

   Demostración. Si $x_1, ..., x_n$ son l. d., existe una combinación lineal $\lambda_1 x_1 + \cdots \lambda_n x_n = 0_E$ con algún $\lambda_j \neq 0$. Entonces, $f(\lambda_1 x_1 + \cdots \lambda_n x_n) = f(0_E)$ y, por las propiedades de la aplicación lineal, $f(\lambda_1 x_1 + \cdots \lambda_n x_n) = \lambda_1 f(x_1) + \cdots \lambda_n f(x_n) = 0_F$ con $\lambda_j \neq 0$. Por tanto, $\{f(x_1), ..., f(x_n)\}$ es l.d. □

5) Si el conjunto de vectores $\{f(x_1), ..., f(x_n)\} \subset F$ es linealmente independiente, también lo es el conjunto $\{x_1, ..., x_n\} \subset E$.

Ejercicio 5.1.2. *Tomemos el conjunto de vectores* $(1,1,0), (2,0,-1), (0,-1,0)$ *en* $\mathbb{R}^3$ *y la aplicación lineal* $f : \mathbb{R}^3 \to \mathbb{R}^2$ *tal que* $f(x,y,z) = (x+y,z)$. *Comprobar si el conjunto* $\{(1,1,0), (2,0,-1), (0,-1,0)\}$ *es linealmente independiente. Estudiar si el conjunto* $\{f((1,1,0)), f((2,0,-1)), f((0,-1,0))\} \subset \mathbb{R}^2$ *es linealmente indpendiente. ¿Se contradice esto con las propiedades de la aplicación lineal?*

Definición 5.1.3. *Sea* $f : E \to F$ *una aplicación lineal. Se define la* ***imagen de*** E ***por*** f *como* $\mathrm{Im}(f) := \{y \in F \mid \exists x \in E, f(x) = y\}$.

Proposición 5.1.4. *La imagen de una aplicación lineal es un subespacio vectorial de* V.

Demostración. Sean $y_1, y_2 \in f(E)$. Entonces, $\exists x_1, x_2 \in E$ tales que $y_1 = f(x_1)$, $y_2 = f(x_2)$. Así, $\forall\, \lambda, \mu \in \mathbb{R}$, $\lambda y_1 + \mu y_2 = \lambda f(x_1) + \mu f(x_2) = f(\lambda x_1 + \mu x_2)$ y, por tanto, $\lambda y_1 + \mu y_2 \in f(E)$ como queríamos demostrar. □

Definición 5.1.5. *La dimensión del subespacio vectorial* $\mathrm{Im}(f)$ *se llama* ***rango de la aplicación lineal*** *y se denota* $\mathrm{rg}(f)$.

Teorema 5.1.6. *Sea* f *una aplicación lineal de* E *en* F *y sea* $\{x_1, ..., x_n\}$ *una base de* E. *Entonces,* $\{f(x_1), ..., f(x_n)\}$ *es un sistema generador de subespacio* $\mathrm{Im}(f) \subset F$.

Definición 5.1.7. *Se llama* ***núcleo*** *de una aplicación lineal* $f : E \to F$, $\mathrm{Ker}(f)$, *al conjunto de elementos de* E *que tienen como imagen el elemento neutro de* F: $\mathrm{Ker}(f) := \{x \in E \mid f(x) = 0_F\}$.

Proposición 5.1.8. *El núcleo de una aplicación lineal* $f : E \to F$ *es un subespacio vectorial de* E.

Demostración. Sean $x, y \in \mathrm{Ker}(f)$. Entonces, $f(x) = 0$, $f(y) = 0$ y $\forall\, \lambda, \mu \in \mathbb{R}$, $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = \lambda 0 + \mu 0 = 0$ como queríamos demostrar. □

Teorema 5.1.9. *Sea $f\colon E \to F$ una aplicación lineal. Entonces:*

$$\dim \mathrm{Im}(f) + \dim \mathrm{Ker}(f) = \dim E.$$

Proposición 5.1.10. *Una aplicación lineal f es inyectiva si y sólo si* $\mathrm{Ker}(f) = 0_E$.

Demostración. Sea f inyectiva. Si $x \in \mathrm{Ker}(f)$, $f(x) = 0_F = f(0_E)$. Como f es inyectiva, $x = 0_E$, luego $\mathrm{Ker}(f) = \{0_E\}$.

Supongamos ahora que $\mathrm{Ker}(f) = \{0_E\}$ y tomemos $x, y \in E$ tales que $f(x) = f(y)$. Entonces, $0_F = f(x) - f(y) = f(x - y)$ y, por tanto, $x - y \in \mathrm{Ker}(f)$. Como $\mathrm{Ker}(f) = \{0_E\}$, $x = y$ y f es inyectiva. □

Por definición, f es sobreyectiva si y sólo si $\mathrm{Im}(f) = F$.

5.1.1. Clasificación de las aplicaciones lineales.

Sea $f\colon E \to F$ una aplicación lineal. Entonces, f es:

- inyectiva $\Leftrightarrow \dim(E) = \mathrm{rg}(f)$,
- sobreyectiva $\Leftrightarrow \dim(F) = \mathrm{rg}(f)$,
- biyectiva $\Leftrightarrow \dim(E) = \mathrm{rg}(f) = \dim(F)$.

Proposición 5.1.11. *Si $f\colon E \to F$ es una aplicación lineal biyectiva entonces existe una aplicación inversa $f^{-1}\colon F \to E$ tal que $f^{-1} \circ f = Id_E$, $f \circ f^{-1} = Id_F$ y f^{-1} es lineal.*

Definición 5.1.12. *A una aplicación lineal biyectiva la llamamos* ***isomorfismo*** *de espacios vectoriales.*

Ejemplo 5.1.13. *Sea $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x, y) = (2x - y, x + y)$. Veamos el núcleo, la imagen y clasifiquemos la aplicación.*

En primer lugar, para calcular el núcleo basta reescribir la definición para el caso concreto de la aplicación f. Así,

$$\ker f = \{(x, y) \,|\, (2x - y, x + y) = (0, 0)\}.$$

Esto se traduce en un sistimea lineal:

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

y la única solución es $(x, y) = (0, 0)$. Por tanto, $\mathrm{Ker}\, f = \{(0, 0)\}$ *y por la proposición 5.1.10, f es inyectiva.*

Dado que la aplicación f va de $\mathbb{R}^2$ en $\mathbb{R}^2$, ambas de dimensión 2, por el Teorema 5.1.9, $dim\,\mathrm{Im}(f) = 2$ que coincide con la dimensión del espacio de llegada ($\mathbb{R}^2$). Por tanto, f es también sobreyectiva y, al ser ambas cosas, biyectiva.

Ejercicio 5.1.14. *Dadas las siguientes aplicaciones, determinar el núcleo y clasificar si son inyectivas, sobreyectivas y/o biyectivas.*

- $f(x, y) = (3x - 2y, 2y, 0)$,
- $g(x, y, z) = (x + y, y - z)$,
- $h(x, y) = (2x - y, ay - x)$ *(en función del parámetro a).*

5.2. Matriz asociada a una aplicación lineal

Por el teorema 5.1.6, sabemos que dada una aplicación lineal $f\colon E \to F$ y una base de E, la imagen de los elementos de la base es un sistema generador del subespacio $\mathrm{Im}(f)$.

Teorema 5.2.1. *Sea $\{x_1, x_2, ..., x_n\}$ una base de E y sea $\{y_1, y_2, ..., y_n\}$ un conjunto cualquiera de n vectores de F. Entonces, existe una única aplicación lineal $f\colon E \to F$ tal que $f(x_i) = y_i\ \forall i = 1, 2, ...n$.*

Asi pues, una aplicación lineal está determinada por la imagen de los elementos de la base.

Sea $B = \{u_1, u_2, ..., u_n\}$ una base de E, $B' = \{v_1, v_2, ..., v_m\}$ una base de F y $f\colon E \to F$ una aplicación lineal.

Sea $x \in E$. Como vimos en el teorema 4.3.10, cualquier vector se puede expresar (de forma única) como combinación lineal de los vectores de la base: $x = x_1u_1 + x_2u_2 + \cdots + x_nu_n$. Sea $y = f(x)$. Si expresamos y como combinación lineal de los vectores de B' tenemos que

$$\begin{aligned} y = y_1v_1 + y_2v_2 + \cdots y_mv_m &= f(x_1u_1 + x_2u_2 + \cdots + x_nu_n) = \\ &= x_1f(u_1) + x_2f(u_2) + \cdots + x_nf(u_n). \end{aligned} \tag{5.1}$$

Si expresamos los vectores $f(u_i)$ en función de la base B':

$$f(u_1) = f_{11}v_1 + f_{21}v_2 + \cdots + f_{m1}v_m$$
$$f(u_2) = f_{12}v_1 + f_{22}v_2 + \cdots + f_{m2}v_m$$

....

$$f(u_n) = f_{1n}v_1 + f_{2n}v_2 + \cdots + f_{mn}v_m$$

En forma matricial, si sustituimos estas expresiones en (5.1) tenemos:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} & \cdots\cdots\cdots & f_{1n} \\ f_{21} & f_{22} & \cdots\cdots\cdots & f_{2n} \\ \cdots & \cdots & \cdots\cdots\cdots & \cdots \\ f_{m1} & f_{m2} & \cdots\cdots\cdots & f_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Es decir, si $(y)_{B'} = (y_1, y_2, ..., y_m)$ denota las coordenadas del vector y en la base B', $(x)_B = (x_1, x_2, ..., x_n)_B$ denota las coordenadas del vector x en la base B y $M(f) = (f_{ij})$ denota la matriz cuyas columnas son las coordenadas de los vectores $f(u_j)$ en la base B' tenemos que:

$$(y)_{B'} = M(f)(x)_B.$$

Decimos que $M(f)$, que depende únicamente de la aplicación f y de las bases B y B' es la **matriz de la aplicación lineal respecto a las bases B y B'**.

Proposición 5.2.2. *Dada una aplicación lineal f, el rango de la aplicación coincide con el rango de la matriz asociada:* $\mathrm{rg}(M(f)) = \mathrm{rg}(f) = \dim(\mathrm{Im}(f))$.

Ejemplo 5.2.3. *Calculamos la matriz de la aplicación $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ tal que $f(x, y, z) = (x + y, y + z)$ respecto a las bases canónicas.*

Calculamos la imagen de los vectores de la base: $f((1, 0, 0)) = (1, 0)$, $f(0, 1, 0) = (1, 1)$, $f(0, 0, 1) = (0, 1)$. Las columnas de la matriz de la aplicación lineal son las coordenadas de estos vectores respecto a la base que tomamos en $\mathbb{R}^2$, en este caso, la canónica. Así pues, la expresión matricial es:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Ejemplo 5.2.4. *Sea $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ la aplicación dada por $f(x, y) = (3y - 2x, 3x + y)$. Determinar la matriz asociada a f respecto a las bases $B_1 = \{(1, 0), (1, 1)\}$ y $B_2 = \{(2, 1), (0, 2)\}$.*

En primer lugar, calculamos las imágenes a través de f de los vectores de la primera base, B_1:

$$f(1, 0) = (-2, 3)$$

$$f(1,1) = (1,4)$$

A continuación, determinamos las coordenadas de los vectores obtenidos en la segunda base, B_2:

$$x(2,1) + y(0,2) = (-2,3)$$

$$x'(2,1) + y'(0,2) = (1,4)$$

Esto produce dos sistemas lineales cuyas soluciones son: $x = -1$, $y = 2$, $x' = \frac{1}{2}$, $y' = \frac{7}{4}$. Por tanto, la matriz asociada a la aplicación f respecto a este par de bases será:

$$\begin{pmatrix} -1 & \frac{1}{2} \\ 2 & \frac{7}{4} \end{pmatrix}$$

A través de esta matriz podemos calcular la imagen de cualquier vector expresado en sus coordenadas en B_1 obteniendo la imagen en las coordenadas en B_2. Nótese que, por ejemplo, que multiplicando por la matriz obtenemos que la imagen del vector $(1,4)_{B_1}$ será el vector $(1,9)_{B_2}$:

$$\begin{pmatrix} -1 & \frac{1}{2} \\ 2 & \frac{7}{4} \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 9 \end{pmatrix}.$$

Es inmediato comprobar que, efectivamente, $(1,4)_{B_1} = (5,4)$ y $f(5,4) = (2,19) = (1,9)_{B_2}$.

Ejercicio 5.2.5. *Sea $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ la aplicación dada por $f(x,y) = (0, 2x + 2y, y - x)$. Determinar la matriz asociada a f respecto a las bases $B_1 = \{(2,1),(1,-1)\}$ y $B_2 = \{(1,0,1),(0,1,-1),(1,1,1)\}$.*

Solución:

$$\begin{pmatrix} -5 & 2 \\ 1 & 2 \\ 5 & -2 \end{pmatrix}$$

5.3. Sobre el conjunto de las aplicaciones lineales

Dados dos espacios vectoriales E y F, el conjunto de todas las aplicaciones lineales de E en F se denota $\pounds\,(E,F)$.

5.3.1. Composición de aplicaciones lineales

Sean E, F, G tres espacios vectoriales sobre $\mathbb{R}$. Sea f una aplicación lineal de E en F y g una aplicación lineal de F en G. Entonces, se verifica que $(g \circ f)$ es una aplicación lineal de E en G.

$$\begin{tikzcd} g \circ f : E \arrow[rr] \arrow[dr, "f"'] & & G \\ & F \arrow[ur, "g"'] & \end{tikzcd}$$

Comprobamos que, efectivamente, cumple la definición:

a) $(g \circ f)(x+y) = g\left[f\ (x+y)\right] = g\left[f(x) + f(y)\right] = g\left[f(x)\right] + g\left[f(y)\right] = (g \circ f)(x) + (g \circ f)(y)\ \forall\, x, y \in E.$

b) $(g \circ f)(\lambda x) = g\ \left[f(\lambda x)\right] = g\ \left[\lambda f\ (x)\right] = \lambda (g \circ f)(x) \quad \forall\, x \in E\ \forall \lambda \in \mathbb{R}.$

Proposición 5.3.1. *$M(g \circ f) = M(g) \cdot M(f)$.*

Ejemplo 5.3.2. *Dadas las aplicaciones lineales $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ y $g \colon \mathbb{R}^2 \to \mathbb{R}^2$ definidas por:*

$$f(x,y) = (3y - 2x, 3x + y) \qquad y \qquad g(s,t) = (2s + t, 4s)$$

obtener la composición $h = g \circ f$.

En primer lugar, determinamos las matrices asociadas respecto a las bases canónicas:

$$M(f) = \begin{pmatrix} -2 & 3 \\ 3 & 1 \end{pmatrix} \quad M(g) = \begin{pmatrix} 2 & 1 \\ 4 & 0 \end{pmatrix}.$$

La matriz asociada a la composición será

$$M(h) = \begin{pmatrix} 2 & 1 \\ 4 & 0 \end{pmatrix} \cdot \begin{pmatrix} -2 & 3 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 7 \\ -8 & 12 \end{pmatrix}.$$

El resultado, por tanto, es $h(x,y) = (7y - x, 12y - 8x)$.

Ejercicio 5.3.3. *Dadas las aplicaciones lineales $f \colon \mathbb{R}^2 \to \mathbb{R}^3$ y $g \colon \mathbb{R}^3 \to \mathbb{R}^2$ definidas por:*

$$f(x,y) = (2x - 3y, 2y, 0) \qquad y \qquad g(s,t,u) = (s + 2t - u, s + u, 2t - u),$$

obtener la composición $h = g \circ f$.

5.3.2. El conjunto de aplicaciones lineales como espacio vectorial

Teorema 5.3.4. *Dados dos espacios vectoriales E y F, el conjunto de las aplicaciones lineales de E en F, $\pounds(E,F) = \{f\colon E \to F \,|\, f \text{ lineal }\}$ con las operaciones:*

- *Suma:* $+ : \pounds(E,F) \times \pounds(E,F) \to \pounds(E,F)$ *tal que* $f+g\colon E \to F$ *es la aplicación que asocia a cada* $x \in E$, $(f+g)(x) := f(x)+g(x)\ \forall\, x \in E$.
- *Producto:* $* : \mathbb{R} \times \pounds(E,F) \to \pounds(E,F)$ *tal que* $\lambda * f\colon E \to F$ *es la aplicación que asocia a cada* $x \in E$, $(\lambda * f)(x) := \lambda * f(x)\ \forall\, x \in E$.

es un espacio vectorial.

Teorema 5.3.5. *Si* $\dim(E) = n$ *y* $\dim(F) = m$ *con bases* B *y* B' *respectivamente, entonces:*

$$M_{B,B'}\colon \pounds(E,F) \to \mathcal{M}_{m\times n}$$

tal que a cada aplicación f *le hace corresponder su matriz asociada respecto a las bases* B *y* B' *es un isomorfismo de espacios vectoriales. En particular,* $\dim(\pounds(E,F)) = m \cdot n$.

5.4. Cambio de base de una aplicación lineal

Sea $f\colon E \to F$ una aplicación lineal, $B_E = \{e_1, e_2, ..., e_n\}$ y $B'_E = \{e'_1, e'_2, ..., e'_n\}$ dos bases de E y $B_F = \{f_1, f_2, ..., f_n\}$ y $B'_F = \{f'_1, f'_2, ..., f'_n\}$ dos bases de F. Sea $w = f(v)$ y sean $v = (x_1, \ldots, x_n)_{B_E} =: x$, $v = (x'_1, \ldots, x'_n)_{B'_E} =: x'$, $w = (y_1, \ldots, y_n)_{B_F} =: y$, $w = (y'_1, \ldots, y'_n)_{B'_F} =: y'$.

Como vimos en el apartado 5.2, tenemos una matriz $A_1 = M_{B_E,B_F}(f)$ que caracteriza la aplicación lineal f respecto a las bases B_E y B_F,

$$y = A_1 x. \tag{5.2}$$

Del mismo modo, tenemos una matriz $A_2 = M_{B'_E,B'_F}(f)$ que caracteriza la aplicación lineal f respecto a las bases B'_E y B'_F,

$$y' = A_2 x'. \tag{5.3}$$

Sea P la matriz de cambio de base de B_E a B'_E. Como vimos en el apartado 4.7:

$$x' = Px. \tag{5.4}$$

Sea Q la matriz de cambio de base de B_F a B'_F. Entonces,

$$y' = Qy. \tag{5.5}$$

Por tanto la relación entre las matrices que caracterizan la aplicación lineal respecto a las distintas bases, A_1 y A_2, es como sigue:

$$y' = Qy \Rightarrow y = Q^{-1}y' \overset{(5,3)}{\Rightarrow} y = Q^{-1}A_2x' \overset{(5,4)}{\Rightarrow} y = Q^{-1}A_2Px \tag{5.6}$$

De (5.2) y (5.6) tenemos que:

$$A_1 = Q^{-1}A_2P.$$

De forma equivalente, o despejando la ecuación, tenemos también que:

$$A_2 = QA_1P^{-1}.$$

Ejemplo 5.4.1. *Sea $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ dada por $f(x, y, z) = (3x + 2y - 4z, x - 5y + 3z)$.*

a) Calcular la matriz asociada a f respecto a las bases canónicas, B_E y B_F, de $\mathbb{R}^3$ y $\mathbb{R}^2$.

b) Calcular la matriz asociada a f cuando las bases son $B'_E = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ y $B'_F = \{(1, 3), (2, 5)\}$.

Solución.

a) La matriz viene determinada por las imágenes de los elementos de la base B_E respecto a la base B_F.

$$\left.\begin{array}{r} f(1,0,0) = (3,1) \\ f(0,1,0) = (2,-5) \\ f(0,0,1) = (-4,3) \end{array}\right\} \text{ de donde } A_1 = M_{B_E,B_F}(f) = \begin{pmatrix} 3 & 2 & -4 \\ 1 & -5 & 3 \end{pmatrix}$$

b) Podemos calcular directamente la matriz de la aplicación lineal respecto a las nuevas bases. Calculamos las imágenes de los elementos de la base B'_E primero respecto a la base canónica B_F.

$$\left.\begin{array}{ll} f(1,1,1) = (1,-1)_{B_F} & \to (1,-1)_{B_F} = a_{11}(1,3) + a_{21}(2,5) \\ f(1,1,0) = (5,-4)_{B_F} & \to (5,-4)_{B_F} = a_{12}(1,3) + a_{22}(2,5) \\ f(1,0,0) = (3,1)_{B_F} & \to (3,1)_{B_F} = a_{13}(1,3) + a_{23}(2,5) \end{array}\right\}$$

Y resolviendo los tres sistemas, tenemos que

$$A_2 = M_{B'_E,B'_F} = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$$

También podemos calcular $A_2 = M_{B'_E,B'_F}(f)$ *usando las matrices del cambio de base:* $A_2 = QA_1P^{-1}$. *Como vimos en el apartado 4.7, las columnas de la matriz cambio de base de* B' *a* B *son los vectores de* B' *en función de* B.

$$M_{B'_E,B_E} = M^{-1}_{B_E,B'_E} = P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

y $M_{B'_F,B_F} = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$

Haciendo los cálculos, tenemos que

$$M^{-1}_{B'_F,B_F} = M_{B_F,B'_F} = Q = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 3 & 2 & -4 \\ 1 & -5 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$$

Definición 5.4.2. *Dos matrices* A_1, A_2 *son* ***equivalentes*** *si existen dos matrices inversibles* P *y* Q *tales que* $A_2 = Q^{-1}A_1P$.

Proposición 5.4.3. *Dos matrices son equivalentes si y solo si son las matrices asociadas a una misma aplicación lineal en distintas bases.*

Definición 5.4.4. *Dos matrices* A_1, A_2 *son* ***semejantes*** *si existe una matriz inversible* P *tal que* $A_2 = P^{-1}A_1P$.

Ejemplo 5.4.5. *Las matrices* $A_1 = \begin{pmatrix} -1 & 2 \\ 3 & 3 \end{pmatrix}$ *y* $A_2 = \begin{pmatrix} 5 & 6 \\ -1 & -3 \end{pmatrix}$ *son semejantes. La matriz de paso es* $P = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$

Comprobar que $A_2 = P^{-1}A_1P$.

Definición 5.4.6. *Una aplicación lineal de un espacio vectorial en sí mismo,* $f : E \to E$, *se denomina* endomorfismo.

Proposición 5.4.7. *Dos matrices son semejantes si y solo si son las matrices asociadas a un mismo endomorfismo en distintas bases.*

5.5. Aplicaciones lineales en $\mathbb{R}^2$ y algunos movimientos del plano

Si $f : \mathbb{R}^2 \to \mathbb{R}^2$ es una apicación lineal, entonces $M(f) \in \mathcal{M}_{2x2}$. Veamos algunas transformaciones básicas del plano y cómo son sus correspondientes matrices.

- Aplicación nula: $f(x,y) = (0,0)$, $M(f) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

- Aplicación identidad: $f(x,y) = (x,y)$, $M(f) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

- Proyección sobre los ejes: eje OX, $f(x,y) = (x,0)$ o eje OY, $g(x,y) = (0,y)$,
$$M(f) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad M(g) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

- Homotecias centradas en (0,0): $f(x,y) = (ax, ay)$, $M(f) = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$

- Giros (en el sentido antihorario) con centro $(0,0)$ y ángulo α: $f(x,y) = (x\cos\alpha - y\operatorname{sen}\alpha, x\operatorname{sen}\alpha + y\cos\alpha)$
$$M(f) = \begin{pmatrix} \cos(\alpha) & -\operatorname{sen}(\alpha) \\ \operatorname{sen}(\alpha) & \cos(\alpha) \end{pmatrix}$$

  Veamos de dónde sale esto. Si consideramos (x,y) como número complejo, $x+iy$, el producto de $1_\alpha \cdot (x+iy)$ nos da un giro de ángulo α en sentido antihorario. Si lo escribimos en forma trigonométrica, tenemos que $1_\alpha = \cos\alpha + i\operatorname{sen}\alpha$. Por tanto, el producto es
$$(\cos\alpha + i\operatorname{sen}\alpha)(x+iy) = x\cos\alpha - y\operatorname{sen}\alpha + (x\operatorname{sen}\alpha + y\cos\alpha)i$$

- Simetría respecto a los ejes: eje OX, $f(x,y) = (x,-y)$ o eje OY, $g(x,y) = (-x,y)$,
$$M(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad M(g) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

- Simetría respecto a la recta $y = mx$ con $m = \tan(\alpha)$: $f(x,y) = (x\cos(2\alpha) + y\operatorname{sen}(2\alpha), x\operatorname{sen}(2\alpha) - y\cos(2\alpha))$
$$M(f) = \begin{pmatrix} \cos(2\alpha) & \operatorname{sen}(2\alpha) \\ \operatorname{sen}(2\alpha) & -\cos(2\alpha) \end{pmatrix}$$

Capítulo 6

Diagonalización de matrices

6.1. Definición y propiedades

Sea E un espacio vectorial sobre $\mathbb{R}$ o $\mathbb{C}$ y $f\colon E \to E$ un *endomorfismo*, esto es, una aplicación lineal de un espacio vectorial en sí mismo.

Pretendemos expresar la aplicación lineal en forma matricial respecto a alguna base B de forma que la matriz sea diagonal. Así pues, la primera pregunta que nos hacemos es: ¿cuándo existe dicha base? Es decir, ¿para qué endomorfismos existe una base B para la que $M_B(f)$ es diagonal?

Propiedad: Si f es un endomorfismo, $B = \{v_1, v_2, ..., v_n\}$ es una base de E y

$$M_B(f) = \begin{pmatrix} a_{11} & 0 & ... & 0 \\ 0 & a_{22} & ... & 0 \\ : & : & ... & 0 \\ 0 & 0 & ... & a_{nn} \end{pmatrix}$$

es diagonal, entonces para cada $v_i \in B$, $f(v_i) = a_{ii}v_i$.

Tenemos que, en la base B, $v_i = (0, ..., 0, 1, 0, ..., 0)_B$. Si representamos la aplicación lineal en forma matricial:

$$\begin{pmatrix} a_{11} & 0 & ... & 0 \\ 0 & a_{22} & ... & 0 \\ : & : & ... & 0 \\ : & : & ... & : \\ 0 & 0 & ... & a_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ : \\ 1 \\ : \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ : \\ a_{ii} \\ : \\ 0 \end{pmatrix}$$

es decir $0v_1 + 0v_2 + \cdots + 0v_{i+i} + a_{ii}v_i + 0v_{i+i} + \cdots + 0v_n = a_{ii}v_i$.

Definición 6.1.1. *Sea f un endomorfismo de un espacio vectorial E sobre $\mathbb{R}$o $\mathbb{C}$. Se llama **vector propio** o **autovector** de f a un elemento $v \in E$ que verifica:*

i) $v \neq 0_E$

ii) Existe un elemento $\lambda \in \mathbb{R}$ *tal que* $f(v) = \lambda v$.

Si denotamos el endomorfismo identidad por Id, la matriz asociada será la matriz unidad, $M(Id) = I$, para cualquier base. Así $f(v) = \lambda v = (\lambda Id)(v)$ y, por las propiedades de la aplicación lineal, $(f - \lambda Id)(v) = 0$. Así, las anteriores propiedades $i), ii)$ son equivalentes a la condición:

i') $v \neq 0$ y $v \in \mathrm{Ker}(f - \lambda Id)$ para algún $\lambda \in \mathbb{R}$.

Definición 6.1.2. *Se dice que* λ *es el* ***valor propio*** *o* ***autovalor*** *asociado al vector propio* v *si* $f(v) = \lambda v$.

Es decir, λ es un **valor propio** si y sólo si $\mathrm{Ker}(f - \lambda Id) \neq 0$, o sea, si y sólo si $(f - \lambda Id)$ no es inyectiva.

Llamamos vectores propios y valores propios de la matriz cuadrada A a los vectores propios y valores propios del endomorfismo asociado f respecto a una base B tal que $M_B(f) = A$.

La ecuación $f(x) = \lambda x$ es equivalente a $Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0$ donde, si f es un endomorfismo de un espacio vectorial E de dimensión n,

$$I = \begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ : & : & ... & : \\ 0 & 0 & ... & 1 \end{pmatrix}_{n \times n}$$

6.2. Cálculo de los valores propios

Sea f un endormorfismo en un espacio vectorial E de dimensión n, B una base de E y $A = M_B(f)$. Como vimos antes, λ es un **valor propio** de f si y sólo si $\mathrm{Ker}(f - \lambda Id) \neq 0 \Leftrightarrow (f - \lambda Id)$ es no inyectivo $\Leftrightarrow M(f - \lambda Id)$ es no inversible $\Leftrightarrow A - \lambda I_n$ es no inversible $\Leftrightarrow det(A - \lambda I_n) = 0$.

Por tanto, podemos encontrar los valores propios resolviendo los posibles valores de λ en la ecuación $det(A - \lambda I_n) = 0$, es decir,

$$\begin{vmatrix} (a_{11} - \lambda) & a_{12} & ... & a_{1n} \\ a_{21} & (a_{22} - \lambda) & ... & a_{2n} \\ : & : & ... & : \\ a_{n1} & a_{n2} & ... & (a_{nn} - \lambda) \end{vmatrix} = 0.$$

Esta ecuación se denomina **ecuación característica**.

Su desarrollo nos da un polinomio $P(\lambda)$ de grado n que se denomina **polinomio característico** de la matriz A. Los valores propios de la matriz A son las raíces de dicho polinomio.

Por el teorema fundamental del álgebra, todo polinomio de grado n admite n raíces reales o complejas (no necesariamente distintas):

$$P(\lambda) = (-1)^n(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdot \;\cdots\; \cdot(\lambda-\lambda_k)^{r_k}$$

donde $r_1 + r_2 + \cdots + r_k = n$.

Definición 6.2.1. *El número de veces que aparece la raíz i, r_i, se llama* ***multiplicidad*** *del valor propio.*

Ejemplo 6.2.2. *1)* $A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}$; $P(\lambda) = \begin{vmatrix} 6-\lambda & -2 \\ -2 & 9-\lambda \end{vmatrix} =$

$$= (6-\lambda)(9-\lambda) - 4 = 0 \to \lambda^2 - 15\lambda + 50 = 0 \to \lambda_1 = 10 \text{ ó } \lambda_2 = 5.$$

2) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

$$P(\lambda) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = \begin{vmatrix} 3-\lambda & 3-\lambda & 3-\lambda \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} =$$

$$= (3-\lambda)\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = (3-\lambda)\begin{vmatrix} 1 & 1 & 1 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = (3-\lambda)\lambda^2 = 0.$$

Por tanto, $(3-\lambda)\lambda^2 = 0 \to \lambda_1 = 0$ *(doble),* $\lambda_2 = 3$.

6.3. Cálculo de los vectores propios

Los vectores propios se obtienen obteniendo soluciones no triviales del sistema homogéneo: $(A - \lambda_i I_n)X = 0$ para cada valor propio λ_i:

$$\begin{pmatrix} (a_{11}-\lambda_i) & a_{12} & \ldots & a_{1n} \\ a_{21} & (a_{22}-\lambda_i) & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & (a_{nn}-\lambda_i) \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

es decir,

$$\left.\begin{array}{rrrr} (a_{11}-\lambda_i)x_1 + & a_{12}x_2 + \cdots + & a_{1n}x_n = 0 \\ a_{21}x_1 + & (a_{22}-\lambda_i)x_2 + \cdots + & a_{2n}x_n = 0 \\ \multicolumn{3}{c}{\dotfill} \\ a_{n1}x_1 + & a_{n2}x_2 + \cdots + & (a_{nn}-\lambda_i)x_n = 0 \end{array}\right\}$$

Nótese que este sistema tiene solución no trivial si y sólo si $det(A-\lambda_i I_n) = 0$, es decir, si y sólo si λ_i es valor propio. Si tiene solución no trivial, quiere decir que el sistema es compatible indeterminado.

Como vimos en la proposición 4.5.6, el conjunto de soluciones del sistema es un espacio vectorial. En particular, si $x = (x_1, x_2, ..., x_n)$ es una solución no trivial, también lo es $\alpha \cdot x = (\alpha x_1, \alpha x_2, ..., \alpha x_n)$, $\forall\, \alpha \neq 0$. Al subespacio vectorial asociado al valor propio λ_i lo denotamos E_{λ_i}. El conjunto de vectores propios asociados al valor propio λ_i es $E_{\lambda_i} - \{0\}$.

Ejemplo 6.3.1. *Veamos los vectores propios de las matrices del ejemplo anterior.*

1) *Los vectores propios asociados a $\lambda_1 = 10$ serán las soluciones no triviales de la ecuación:*

   $(A - \lambda_1 I_2)\, X = 0$ *donde* $A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}$ *y* $(A - 10I_2) = \begin{pmatrix} 6-10 & -2 \\ -2 & 9-10 \end{pmatrix} = \begin{pmatrix} -4 & -2 \\ -2 & -1 \end{pmatrix}$

   de donde obtenemos el sistema: $\left.\begin{array}{r} -4x_1 - 2x_2 = 0 \\ -2x_1 - \ \ x_2 = 0 \end{array}\right\}$

   Observamos que las dos ecuaciones son proporcionales. Resolviendo:

   $2x_1 + x_2 = 0 \to$ *Si* $x_1 = \alpha, x_2 = -2\alpha$ *con* $\alpha \in \mathbb{R}$. *Así, los vectores propios asociados a* $\lambda_1 = 10$ *son de la forma :* $(\alpha, -2\alpha)$ *con* $\alpha \neq 0$. *Dicho de otra manera,* $E_{10} = \{(\alpha, -2\alpha) \mid \alpha \in \mathbb{R}\}$. *Tomando un valor no nulo cualquiera de α podemos expresar el subespacio E_{10} a partir de una base que lo genera:* $E_{10} = \mathrm{L}[(1, -2)]$.

   Los vectores propios asociados a $\lambda_2 = 5$ provienen del sistema:

   $\begin{pmatrix} 6-5 & -2 \\ -2 & 9-5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

   Esto es, $\left.\begin{array}{r} x_1 - 2x_2 = 0 \\ -2x_1 + 4x_2 = 0 \end{array}\right\}$

   Observamos que son proporcionales (el rango es 1) y que, por tanto, una de las ecuaciones es redundante. Resolviendo $x_1 - 2x_2 = 0 \to$ *Si* $x_2 = \beta$, *entonces* $x_1 = 2\beta$. *Los vectores propios asociados a* $\lambda_2 = 5$

son de la forma: $(2\beta, \beta)$ *con* $\beta \neq 0$. *Es decir,* $E_5 = \{(2\beta, \beta) \mid \alpha, \beta \in \mathbb{R}\}$. *Tomando un valor no nulo cualquiera de* β *podemos expresar el subespacio* E_5 *a partir de una base que lo genera:* $E_5 = \mathrm{L}[(2, 1)]$.

2) Vectores propios asociados al valor propio $\lambda_1 = 0$

$(A - 0I_3) = \begin{pmatrix} 1-0 & 1 & 1 \\ 1 & 1-0 & 1 \\ 1 & 1 & 1-0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longleftrightarrow x_1 + x_2 + x_3 = 0.$

Los vectores propios son de la forma $(\alpha, \beta, -\alpha - \beta)$ *con* $\alpha, \beta \neq 0$. $E_0 = \{(\alpha, \beta, -\alpha - \beta) \mid \alpha, \beta \in \mathbb{R}\}$. *Para dar una base de* E_0 *necesitaremos dos vectores que obtenemos dando valores a* α *y* β *de forma que los vectores resultantes sean linealmente independientes. Por ejemplo, tomando* $\alpha = 1, \beta = 0$ *y* $\alpha = 0, \beta = 1$ *resulta:* $E_0 = \mathrm{L}[(1, 0, -1), (0, 1, -1)]$.

Los vectores propios asociados a $\lambda_2 = 3$ *, verifican* $(A - 0I_3) = \begin{pmatrix} 1-3 & 1 & 1 \\ 1 & 1-3 & 1 \\ 1 & 1 & 1-3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longleftrightarrow x_1 = x_2 = x_3$. *Son, por tanto, de la forma* (α, α, α) *con* $\alpha \neq 0$. $E_3 = \{(\alpha, \alpha, \alpha) \mid \alpha \in \mathbb{R}\} = \mathrm{L}[(1, 1, 1)]$.

Nótese que: $\dim(\mathrm{Ker}(A - 0I_3)) = 2$ *siendo* 0 *una raíz doble del polinomio característico y* $\dim(\mathrm{Ker}(A - 3I_3)) = 1$ *siendo* 3 *una raíz simple del polinomio característico.*

Ejercicio 6.3.2. *Dada la matriz*

$$A = \begin{pmatrix} \frac{1}{2} & -3 & 2 & 1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

determinar sus valores propios y hallar un vector propio asociado al valor propio -1.

Solución: *Desarrollando el polinomio característico (lo más fácil es hacerlo por la 3ª fila. Después, desarrollo el determinante 3x3 de nuevo por la 3ª fila) obtenemos* $P(\lambda) = \lambda[(3 - 2 + \lambda) + (-1 - \lambda)(\frac{1}{2} - \lambda)(2 - \lambda)] = -\lambda(1 + \lambda)[1 - (\frac{1}{2} - \lambda)(2 - \lambda)] = \lambda^2(1 + \lambda)(\lambda - \frac{5}{2})$.

Los valores propios son 0 *(doble), -1 y* $\frac{5}{2}$.

El sistema asociado al valor propio -1 *es:*

$\left.\begin{array}{rcl} \frac{3}{2}x - 2y + 2z + t & = & 0 \\ 3y \qquad - t & = & 0 \\ z & = & 0 \\ x & = & 0 \end{array}\right\}$ *de donde* $x = z = 0$ *y* $t = 3y$. *Un vector propio asociado al -1 sería de la forma* $(0, \alpha, 0, 3\alpha)$ *para algún* $\alpha \neq 0$.

6.4. Propiedades de los vectores propios y valores propios

1) Dado un valor propio λ, el conjunto E_λ es un subespacio vectorial de E.

   En efecto, sean v_1 y v_2 dos soluciones al sistema $(A - \lambda I)X = 0$ asociado al valor propio λ.

   $\left.\begin{array}{l} Av_1 = \lambda v_1 \Leftrightarrow f(v_1) = \lambda v_1 \\ Av_2 = \lambda v_2 \Leftrightarrow f(v_2) = \lambda v_2 \end{array}\right\}$

   $\forall\, \alpha, \beta \in \mathbb{R}$, $f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2) = \alpha\lambda v_1 + \beta\lambda v_2 = \lambda(\alpha v_1 + \beta v_2)$.

2) La suma de los n valores propios de una matriz es igual a su traza $(a_{11} + a_{22} + \cdots + a_{nn})$. En particular, dos matrices semejantes tienen el mismo polinomio característico y la misma traza.

3) El producto de los n valores propios de una matriz es igual a su determinante.

Teorema 6.4.1. *Un endomorfismo f de E es diagonalizable si y sólo si existe una base $(v_1, v_2, ..., v_n)$ de E formada por vectores propios de f.*

Demostración. Sea $B = \{v_1, v_2, ..., v_n\}$ una base de vectores propios. Supongamos $\lambda_1, \lambda_2, ..., \lambda_n$ los valores propios asociados: $f(v_i) = \lambda_i v_i$ para $i = 1, \ldots, n$. Como se vió en 5.2, $M_B(f) = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & ... & 0 \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$

Supongamos que f es diagonalizable. Sea $B = \{v_1, v_2, ..., v_n\}$ una base tal que $M_B(f)$ es diagonal: $M_B(f) = \begin{pmatrix} a_{11} & 0 & ... & 0 \\ 0 & a_{22} & ... & 0 \\ \vdots & \vdots & ... & 0 \\ 0 & 0 & ... & a_{nn} \end{pmatrix}$

Es inmediato comprobar que si expresamos los vectores $v_1, v_2, ..., v_n$ en la base B se cumple que $M_B(f)(v_i)_B = a_{ii}v_i$. Por tanto, los v_i son vectores propios. $\square$

Teorema 6.4.2. *Dado un endomorfismo* $f\colon E \to E$ *con* $\dim(E) = n$ *y autovalores* $\lambda_1, \lambda_2, ..., \lambda_k$ *con multiplicidades* $\alpha_1, \alpha_2, ..., \alpha_k$*,* f *es diagonalizable si y sólo si* $\alpha_1 + \cdots + \alpha_k = n$ *y* $\dim(E_{\lambda_i}) = \alpha_i$*.*

Observación 6.4.3. *Nótese que, por el teorema fundamental del álgebra, si* E *es un espacio vectorial sobre* $\mathbb{C}$*, entonces la primera condición del teorema 6.4.2, (*$\alpha_1 + \cdots + \alpha_k = n$*), se da siempre.*

Proposición 6.4.4. *Si un endomorfismo* $f\colon E \to E$ *con* $\dim(E) = n$ *posee* n *autovalores reales diferentes* $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$*, entonces* f *es diagonalizable.*

6.5. Diagonalización de una matriz

Sea f un endormorfismo dado por la matriz $A = M_B(f)$ y supongamos que es diagonalizable. Para calcular la matriz equivalente diagonal primero obtenemos los autovalores, posiblemente repetidos, $\lambda_1, \lambda_2, ..., \lambda_n$. Entonces, la matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ : & : & ... & 0 \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$$

es equivalente a A. Nótese que si tomamos los autovalores en un orden distinto obtenemos otra matriz diagonal equivalente a A.

Veamos ahora cómo obtener una matriz de paso P tal que $D = P^{-1}AP$.

Calculamos para cada valor λ_i el espacio vectorial asociado E_{λ_i}. Si la matriz es diagonalizable, como vimos en el teorema 6.4.2, la dimensión de E_{λ_i} coincide con la multiplicidad de λ_i, α_i. Dicho espacio estará generado por α_i vectores que, como vimos, son vectores propios. Los vectores que forman las bases de los espacios E_{λ_i} serán las columnas de la matriz P en el mismo orden en el que aparecen los autovalores en D.

Ejemplo 6.5.1. *1) Dada la matriz* $A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}$ *vimos que los vectores propios son* $\lambda_1 = 10$ *y* $\lambda_1 = 5$*. Por tanto*

$$D = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}$$

Los subespacios asociados son $E_{10} = \mathrm{L}[(1,-2)]$ *y* $E_5 = \mathrm{L}[(2,1)]$. *Así:*
$P = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$

Efectivamente, $P^{-1} = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$ *y*

$$\begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

2) Dada la matriz $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ *vimos que los vectores propios eran* $\lambda_1 = 0$, *con multiplicidad 2 y* $\lambda_2 = 3$.

La matriz diagonal será: $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

Los subespacios asociados, como vimos en el ejemplo anterior, son $E_0 = \mathrm{L}[(1,0,-1),(0,1,-1)]$, $E_3 = \mathrm{L}[(1,1,1)]$. *La matriz de paso que obtenemos es:*
$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$

Nótese que la matriz diagonal depende del orden en el que tomamos los autovalores. La matriz P *depende de dicho orden y de la elección que hemos hecho de los vectores de las bases de los subespacios* E_{λ_i}.

Ejercicio 6.5.2. *Diagonizar las siguientes matrices calculando la matriz de paso* P*:*

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

6.6. Bases ortonormales y método de Gram-Schmidt

Definición 6.6.1. *Una base de un espacio vectorial* $\{e_1, e_2, ..., e_n\}$ *es ortonormal si:*

i) $e_i \cdot e_j = 0\ \forall\, i \neq j$ *(son ortogonales entre sí).*

ii) $||e_1|| = ||e_2|| = \cdots = ||e_n|| = 1$ *(todos tienen norma 1).*

Ejemplo 6.6.2. *La base canónica de* $\mathbb{R}^n$ *es ortonormal.*

Dada una base cualquiera de E, podemos construir una base ortonormal mediante el **método de Gram-Schmidt**.

Sea $\{e_1, e_2, ..., e_n\}$ una base de E.

Primero, normalizamos $e'_1 = \frac{e_1}{||e_1||}$.

Después, para cada $k = 2, 3, ..., n$, suponiendo $e'_1, ..., e'_{k-1}$ definidos y ortonormales entre sí, hacemos

$$x_k = e_k - \sum_{i=1}^{k-1} \langle e_k, e'_i \rangle e'_i.$$

Nótese que entonces x_k es ortogonal a $e'_1, ..., e'_{k-1}$. Primero, como $e'_1, ..., e'_{k-1}$ son ortonormales entre sí, $\forall i, j \leq k$ $\langle x_i, x_j \rangle = 1$ si $i = j$ y $\langle x_i, x_j \rangle = 0$ si $i \neq j$. Así pues, $\forall j = 1, k-1$ $\langle x_k, e'_j \rangle = \langle e_k - \sum_{i=1}^{k-1} \langle e_k, e'_i \rangle e'_i, e'_j \rangle = \langle e_k, e'_j \rangle - \sum_{i=1}^{k-1} \langle e_k, e'_i \rangle \langle e'_i, e'_j \rangle = \langle e_k, e'_j \rangle - \langle e_k, e'_j \rangle = 0$. Además, por ser los vectores linealmente independientes, x_k siempre es distinto de cero y se puede normalizar.

Finalmente, normalizamos $e'_k = \frac{x_k}{x_k}$.

Así, $\{e'_1, e'_2, ..., e'_n\}$ es una base ortonormal.

Ejemplo 6.6.3. *Sea* $B = \{(1, 0, -1), (0, 1, -1), (1, 1, 1)\}$ *una base de* $\mathbb{R}^3$.

$e'_1 = \frac{e_1}{||e_1||} = \frac{(1,0,-1)}{\sqrt{2}} = (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}})$.

$$x_2 = e_2 - \langle e_2, e'_1 \rangle e'_1 = (0, 1, -1) - \langle (0, 1, -1), (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}) \rangle (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}) =$$

$$= (0, 1, -1) - \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}) = (0, 1, -1) - (\frac{1}{2}, 0, \frac{-1}{2}) = (\frac{-1}{2}, 1, \frac{-1}{2}).$$

$e'_2 = \frac{x_2}{||x_2||} = \frac{(\frac{-1}{2}, 1, \frac{-1}{2})}{\frac{\sqrt{6}}{2}} = \frac{2}{\sqrt{6}}(\frac{-1}{2}, 1, \frac{-1}{2}) = (\frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}})$.

$$x_3 = e_3 - \langle e_3, e'_1 \rangle e'_1 - \langle e_3, e'_2 \rangle e'_2 = (1, 1, 1) - \langle (1, 1, 1), (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}) \rangle$$

$$(\frac{1}{\sqrt{2}},0,\frac{-1}{\sqrt{2}})-\langle(1,1,1),(\frac{-1}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{-1}{\sqrt{6}})\rangle(\frac{-1}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{-1}{\sqrt{6}})=$$

$$=(1,1,1)-0(\frac{1}{\sqrt{2}},0,\frac{-1}{\sqrt{2}})-0(\frac{-1}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{-1}{\sqrt{6}})=(1,1,1).$$

$e'_3=\frac{(1,1,1)}{\sqrt{3}}=(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})$.

6.7. Diagonalización de matrices simétricas

Teorema 6.7.1. *Si una matriz A con coeficientes reales es simétrica, entonces:*

a) Los valores propios son números reales.

b) Si λ_1,λ_2 son dos valores propios distintos de A entonces los vectores propios asociados v_1,v_2 son ortogonales.

c) La matriz es diagonalizable. De hecho, existe una matriz P ortogonal tal que $D=P^{-1}AP=P^tAP$.

Corolario 6.7.2. *Si los valores propios de una matriz simétrica A, $\lambda_1,\lambda_2,\dots,\lambda_n$, son todos distintos y $v_1,v_2,\dots,v_n$ son sus vectores propios asociados, entonces $\{\frac{v_1}{||v_1||},\frac{v_2}{||v_2||},\dots,\frac{v_n}{||v_n||}\}$ es una base ortonormal.*

Proposición 6.7.3. *Si $v_1,v_2,\dots,v_n$ es una base ortonormal de vectores propios de una matriz simétrica A y P es la matriz cuyas columnas son $v_1,v_2,\dots,v_n$, entonces P es ortogonal y $D=P^tAP$.*

Ejemplo 6.7.4. *Sea* $A=\begin{pmatrix}4&0&-1\\0&3&0\\-1&0&4\end{pmatrix}$

El polinomio característico es $P(\lambda)=|A-\lambda I|=\begin{vmatrix}4-\lambda&0&-1\\0&3-\lambda&0\\-1&0&4-\lambda\end{vmatrix}=(4-\lambda)(3-\lambda)(4-\lambda)-(3-\lambda)=-\lambda^3+11\lambda^2-39\lambda+45.$

Los valores propios son las raíces de este polinomio. Para calcularlas podemos factorizar el polinomio usando la regla de Ruffini. Obtenemos así que $P(\lambda)=-(\lambda-3)^2(\lambda-5)$ *y los autovalores son* $\lambda=3$ *y* $\lambda=5$.

Calculamos el espacio vectorial asociado a $\lambda=3$, E_3. *Será el conjunto de soluciones del sistema homogéneo:*

$$\begin{pmatrix} 4-3 & 0 & -1 \\ 0 & 3-3 & 0 \\ -1 & 0 & 4-3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

El sistema resultante es:

$$\left.\begin{array}{rr} x & -z=0 \\ -x & +z=0 \end{array}\right\}$$

de donde $x = z$.

Por tanto, $E_3 = \{(\alpha, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$

Buscamos una base E_3. E_3 *tiene dimensión 2. Asignando valores a los parámetros* α, β *obtenemos dicha base. Haciendo* $\alpha = 1$, $\beta = 0$ *tenemos el vector propio* $(1, 0, 1)$, *Haciendo* $\alpha = 0$, $\beta = 1$ *tenemos el vector propio* $(0, 1, 0)$. *Ambos son linalmente independientes y forman una base:* $E_3 = \mathrm{L}[(1, 0, 1), (0, 1, 0)]$.

Calculamos el espacio vectorial asociado a $\lambda = 5$, E_5. *Será el conjunto de soluciones del sistema homogéneo:*

$$\begin{pmatrix} 4-5 & 0 & -1 \\ 0 & 3-5 & 0 \\ -1 & 0 & 4-5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

El sistema resultante es:

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

de donde $y = 0$, *y* $x = -z$.

Por tanto, $E_5 = \{(\alpha, 0, -\alpha) \mid \alpha \in \mathbb{R}\}$.

Buscamos una base E_5. *Dando un valor cualquiera al único parámetro* α *obtenemos dicha base. Haciendo* $\alpha = 1$, *tenemos que* $(1, 0, -1)$ *es un vector propio tal que* $E_5 = \mathrm{L}[(1, 0, -1)]$.

La matriz de cambio tendría por columnas los vectores propios.

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

Así, la matriz diagonal será $D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ *y* $D = P^{-1}AP$.

Para encontrar la matriz de paso ortogonal debemos usar el método de Gram-Schmidt.

$e_1' = \frac{e_1}{||e_1||} = \frac{(1,0,1)}{\sqrt{2}} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$.

$$x_2 = e_2 - \langle e_2, e_1' \rangle e_1' = (0, 1, 0) - \langle (0, 1, 0), (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) \rangle (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) =$$

$$= (0,1,0) - 0(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}) = (0,1,0).$$

$e'_2 = \frac{x_2}{||x_2||} = (0,1,0).$

$$x_3 = e_3 - \langle e_3, e'_1 \rangle e'_1 - \langle e_3, e'_2 \rangle e'_2 = (1,0,-1) - \langle (1,0,-1), (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) \rangle$$

$$(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) - \langle (1,0,-1), (0,1,0) \rangle (0,1,0) = (1,0,-1) - 0(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) -$$

$$-0(0,1,0) = (1,0,-1).$$

$e'_3 = \frac{(1,0,-1)}{\sqrt{2}} = (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}).$

La matriz de paso sería

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

Así, $D = Q^t A Q$.

6.8. Cálculo de potencias de matrices

Sea A una matriz diagonalizable y sea P la matriz de paso. Así, $D = P^{-1}AP$ y $A = PDP^{-1}$. Entonces

$A^n = A \cdot A \cdot \overset{(n\ veces)}{\cdots} \cdot A = (PDP^{-1}) \cdot (PDP^{-1}) \cdot \overset{(n\ veces)}{\cdots} \cdot (PDP^{-1}) = PD(P^{-1})P)D(P^{-1}P) \cdots (P^{-1}P)DP^{-1} = PD^nP^{-1}.$

Por tanto, para caldular A^n basta calcular P, P^{-1} y D^n.

Si

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

entonces es inmediato comprobar que

$$D^n = \begin{pmatrix} \lambda_1^n & 0 & \ldots & 0 \\ 0 & \lambda_2^n & \ldots & 0 \\ \vdots & \vdots & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_n^n \end{pmatrix}$$

Ejemplo 6.8.1. *Sea* $A = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 4 \end{pmatrix}$

Como vimos en el ejemplo 6.7.4, $D = Q^t A Q$ [1] *con:*

$$D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

Es inmediato ver que $D^n = \begin{pmatrix} 3^n & 0 & 0 \\ 0 & 3^n & 0 \\ 0 & 0 & 5^n \end{pmatrix}$

Por tanto,

$$A^n = QD^nQ^{-1} = QD^nQ^t = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3^n & 0 & 0 \\ 0 & 3^n & 0 \\ 0 & 0 & 5^n \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{pmatrix} =$$

$$\begin{pmatrix} \frac{3^n+5^n}{2} & 0 & \frac{3^n-5^n}{2} \\ 0 & 3^n & 0 \\ \frac{3^n-5^n}{2} & 0 & \frac{3^n+5^n}{2} \end{pmatrix}$$

[1] Podemos hacerlo con cualquier matriz de paso. En este caso, lo hacemos con Q y no con P porque al ser ortogonal el cálculo de la inversa es trivial

Bibliografía

[1] Baro González, E., Tomeo Perucha, V. (2014), Introducción al álgebra lineal

[2] Baro González, E., Tomeo Perucha, V. (2014), Introducción al álgebra lineal Baro González E., Tomeo Perucha V. (2014), Introducción al álgebra lineal [Libro-e]

[3] Merino González, L.M., Santos Aláez, E. (2015), Álgebra lineal con métodos elementales

[4] Ward Brown, J., Churchill, R. V. (2004), Variable compleja y aplicaciones