Observación. El conjunto vacío está contenido en cualquier conjunto, es decir $\emptyset \subseteq A$, para todo conjunto A. En efecto, la implicación "$x \in \emptyset \Rightarrow x \in A$" es verdadera, pues su antecedente es falso.

Definición 2.2 *Se dice que el conjunto A es igual al conjunto B, si $A \subseteq B$ y $B \subseteq A$. Lo indicamos* $A = B$

La relación de inclusión no excluye la igualdad de los conjuntos. Si $A \subseteq B$ y además $A \neq B$, se dice que A es un subconjunto *propio* o una parte *propia* de B, o que A está contenido estrictamente en B. Lo notaremos $A \subset B$.

Conjunto de partes de un conjunto

Dado un conjunto A, se puede considerar siempre el conjunto formado por los subconjuntos de A, el cual recibe el nombre de conjunto de las partes de A, y se indica $\mathcal{P}(A)$.

$$\mathcal{P}(A) = \{X : X \subseteq A\}.$$

Observemos que $\mathcal{P}(A)$ nunca es vacío, pues como $\emptyset \subseteq A$ y $A \subseteq A$, entonces $\emptyset$ y A son elementos de $\mathcal{P}(A)$, es decir, $\emptyset \in \mathcal{P}(A)$ y $A \in \mathcal{P}(A)$.

2.2 Unión de conjuntos

Definición 2.3 *Dados dos conjuntos A y B, se llama unión de A y B, y se indica $A \cup B$, al conjunto formado por todos los elementos que pertenecen a A o a B. En notación*

2.3 Intersección de conjuntos

Definición 2.4 *Dados dos conjuntos A y B, se llama intersección de A y B, y se indica $A \cap B$, al conjunto cuyos elementos son los elementos comunes a A y a B, es decir los elementos que pertenecen simultáneamente a los dos conjuntos.*

2.4 Complemento

Definición 2.7 *Dados dos conjuntos A y B, tales que $A \subseteq B$, se llama complemento de A relativo a B, y lo notamos $\mathcal{C}_B A$, al conjunto formado por todos los elementos de B que no pertenecen a A.*

2.5 Diferencia

Definición 2.8 *Dados dos conjuntos A y B, se llama diferencia entre A y B, en ese orden, al conjunto formado por los elementos que pertenecen a A y no pertenecen a B. Se nota $A - B$.*

Observación. La siguiente propiedad es muy útil y resulta en forma inmediata de la definición de diferencia: $A - B = A \cap B'$.

Teorema 2.9 $A \subseteq B \Leftrightarrow B' \subseteq A'$.

Leyes de De Morgan.

1. $(A \cap B)' = A' \cup B'$.
2. $(A \cup B)' = A' \cap B'$.

Antes de demostrarlas, conviene tener presente que:

- $x \notin A \cup B$ significa que $x \notin A$ **y** $x \notin B$.
- $x \notin A \cap B$ significa que $x \notin A$ **o** $x \notin B$.

2.6 Producto cartesiano

Dados dos conjuntos A y B, y objetos cualesquiera $a \in A$ y $b \in B$, consideraremos *pares ordenados* de primera coordenada a y segunda coordenada b, que notaremos (a, b). Dos pares ordenados (a, b) y (a', b') son iguales si y sólo si $a = a'$ y $b = b'$.

Definición 2.10 *Dados dos conjuntos A y B, llamamos producto cartesiano de A y B, y lo notamos $A \times B$, al conjunto formado por todos los pares ordenados (a, b), con $a \in A$ y $b \in B$. Es decir,*

$$A \times B = \{(a, b) : a \in A \ y \ b \in B\}.$$

3.1 Definiciones

Definición 3.1 *Dados dos conjuntos A y B, una relación binaria R entre los elementos de A y los elementos de B (o en $A \times B$), es un subconjunto del producto cartesiano $A \times B$:*

$$R \subseteq A \times B.$$

Se escribe aRb si $(a, b) \in R$. En particular, si $A = B$, entonces R se llama una relación binaria **en** A.

Definición 3.2 *Dada una relación $R \subseteq A \times A$, diremos que R es reflexiva si aRa para todo $a \in A$. ($(a, a) \in R$ para todo $a \in A$).*

Definición 3.3 *Diremos que una relación $R \subseteq A \times A$ es simétrica si se verifica que: Si aRb entonces bRa. ($(a, b) \in R \Rightarrow (b, a) \in R$).*

Definición 3.4 *Diremos que una relación $R \subseteq A \times A$ es antisimétrica si se verifica que:*
Si aRb y bRa entonces $a = b$. ($(a, b) \in R$ y $(b, a) \in R \Rightarrow a = b$).
Equivalentemente: Si $a \neq b$ entonces $a \not R b$ o bien $b \not R a$. ($a \neq b \Rightarrow (a, b) \notin R$ o bien $(b, a) \notin R$).

Nota: Otra expresión equivalente de la propiedad antisimétrica es la siguiente: Si $a \neq b$ y aRb entonces $b \not R a$. ($a \neq b$ y $(a, b) \in R \Rightarrow (b, a) \notin R$).

Definición 3.5 *Diremos que una relación $R \subseteq A \times A$ es transitiva, si se verifica que: Si aRb y bRc entonces aRc. ($(a, b) \in R$ y $(b, c) \in R \Rightarrow (a, c) \in R$).*

Relación	Reflexiva	Simétrica	Antisimétrica	Transitiva
$\emptyset$	no	sí	sí	sí
$\{(a,a)\}$	no	sí	sí	sí
$\{(a,b)\}$	no	no	sí	sí
$\{(b,a)\}$	no	no	sí	sí
$\{(b,b)\}$	no	sí	sí	sí
$\{(a,a),(a,b)\}$	no	no	sí	sí
$\{(a,a),(b,a)\}$	no	no	sí	sí
$\{(a,a),(b,b)\}$	sí	sí	sí	sí
$\{(a,b),(b,a)\}$	no	sí	no	no
$\{(a,b),(b,b)\}$	no	no	sí	sí
$\{(b,a),(b,b)\}$	no	no	sí	sí
$\{(a,a),(a,b),(b,a)\}$	no	sí	no	no
$\{(a,a),(a,b),(b,b)\}$	sí	no	sí	sí
$\{(a,a),(b,a),(b,b)\}$	sí	no	sí	sí
$\{(a,b),(b,a),(b,b)\}$	no	sí	no	no
$A \times A$	sí	sí	no	sí

3.2 Relaciones de equivalencia

Definición 3.6 *Diremos que una relación R en A es de equivalencia si es reflexiva, simétrica y transitiva.*

Partición de un conjunto

La noción de partición de un conjunto que definiremos a continuación, está ligada a la propiedad fundamental que tienen las relaciones de equivalencia.

Dado un conjunto A, diremos que una familia $\{A_i\}_{i\in I}$ de subconjuntos de A constituye una partición de A si:

1. $A_i \neq \emptyset$, para todo $i \in I$,
2. Si $i \neq j$ entonces $A_i \cap A_j = \emptyset$,
3. $\bigcup_{i\in I} A_i = A$,

donde $\bigcup_{i\in I} A_i = \{x : x \in A_i, \text{ para algún } i \in I\ \}$.

Es decir, todos los subconjuntos de la familia son no vacíos, son disjuntos dos a dos y su unión es A.

Partición inducida por una relación de equivalencia. Clases de equivalencia

Definición 3.7 *Dada una relación de equivalencia R definida en un conjunto A, llamaremos "clase de equivalencia de un elemento $x \in A$", al conjunto de todos los elementos de A que están en relación con x. Notaremos C_x.*

$$C_x = \{z \in A : zRx\}.$$

Observar que las clases de equivalencia de elementos de A son subconjuntos de A.

Propiedades de las clases de equivalencia

1. $C_x \neq \emptyset$, para todo $x \in A$. (Ninguna clase es vacía).
2. $C_x = C_y \Leftrightarrow xRy$.
3. $C_x \neq C_y \Rightarrow C_x \cap C_y = \emptyset$.

Teorema 3.8 *Toda relación de equivalencia definida en un conjunto A, determina una partición de A formada por las clases de equivalencia distintas.*

Conjunto cociente

Definición 3.9 *Se llama conjunto cociente al conjunto cuyos elementos son las clases de equivalencia distintas determinadas por una relación de equivalencia R definida en un conjunto A. Se nota A/R.*

Relación de equivalencia determinada por una partición

Vamos a ver ahora la recíproca del teorema anterior, es decir, que dada una partición de A se puede definir una relación de equivalencia de modo tal que las clases de equivalencia correspondientes coincidan con los subconjuntos de la partición dada.

Teorema 3.10 *Sea $\{A_i\}_{i\in I}$ una partición de un conjunto A. Sea R la relación definida en A por: $(x, y) \in R \Leftrightarrow$ existe $i \in I$ tal que $x \in A_i$ e $y \in A_i$. Entonces R es una relación de equivalencia y las clases de equivalencia correspondientes son los subconjuntos de la partición dada.*

Congruencia módulo m

Un ejemplo muy importante de relación de equivalencia es la congruencia módulo m.

Definición 3.11 *Sea $\mathbb{Z}$ el conjunto de los números enteros y sea $m \in \mathbb{Z}$ fijo. Dos números enteros a y b se dicen congruentes módulo m, si y sólo si $a - b$ es un múltiplo de m. O sea,*

$$a \equiv b \pmod{m} \Leftrightarrow a - b = k \cdot m, \text{ con } k \in \mathbb{Z}.$$

Por ejemplo, $10 \equiv 2 \pmod 4$, $11 \equiv 25 \pmod 2$, $10 \equiv 13 \pmod 3$, $a \equiv b \pmod 0$ si y sólo si $a = b$, $a \equiv b \pmod 1$ para todo a, $b \in \mathbb{Z}$.

Teorema 3.12 *La relación de congruencia para un número entero fijo m, definida en el conjunto de los números enteros, es una relación de equivalencia.*

3.3 Relaciones de orden

Definición 3.14 *Una relación binaria R definida en un conjunto A es de orden si es reflexiva, antisimétrica y transitiva.*

Definición 4.1 *Dados dos conjuntos no vacíos A y B, una función o aplicación de A en B es una relación $f \subseteq A \times B$ tal que para todo $a \in A$, existe uno y sólo un elemento $b \in B$ tal que $(a, b) \in f$.*

Dada una función $f : A \to B$, el conjunto A se llama el *dominio* de f y B el *codominio*.

Nota. De la definición de función resulta que dadas dos funciones $f : A \to B,\ \ g : A \to B$, entonces $f = g$ si y sólo si $f(x) = g(x)$, para todo $x \in A$.

Imagen e imagen completa inversa

Para cada subconjunto $X \subseteq A$, se llama *imagen* de X por f al siguiente subconjunto de B: $f(X) = \{y \in B: \text{ existe } x \in X : y = f(x)\}$.

En particular, si $X = A$, la *imagen* de f (o rango de f) es $f(A) = Im(f) = \{y \in B: \text{ existe } x \in A : y = f(x)\}$.

4.1 Funciones inyectivas, epiyectivas y biyectivas

Definición 4.2 *Diremos que una función $f : A \to B$ es inyectiva, o que f es una función biunívoca, si elementos distintos en A tienen imágenes diferentes en B, esto es, si se verifica:*

$$x \neq x' \ \Rightarrow\ f(x) \neq f(x'),$$

o equivalentemente,

$$f(x) = f(x') \ \Rightarrow\ x = x'.$$

Definición 4.3 *Diremos que una función $f : A \to B$ es epiyectiva, suryectiva, sobreyectiva, o que f es de A sobre B, si $Imf = B$, esto es, cualquiera que sea $y \in B$, existe un $x \in A$ tal que $f(x) = y$.*

Definición 4.4 *Diremos que una aplicación $f : A \to B$ es biyectiva o que f es una correspondencia biunívoca de A sobre B si f es inyectiva y epiyectiva.*

4.2 Composición de funciones

Definición 4.5 *Sean A, B, C conjuntos y sean $f : A \to B$ y $g : B \to C$ dos funciones. Se llama composición de f y g a la función, indicada con $g \circ f : A \to C$, definida por:*

$$(g \circ f)(x) = g(f(x)),\ \textit{para todo } x \in A.$$

Proposición 4.6 *Sean A, B, C, D conjuntos y $f : A \to B$, $g : B \to C$, $h : C \to D$. Entonces vale la ley asociativa siguiente:*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Proposición 4.7 *Sea $f : A \to B$ una función. Entonces f es biyectiva si y sólo si existe una función $g : B \to A$ tal que $g \circ f = I_A$ y $f \circ g = I_B$.*

Definición 4.8 *Dada una función $f : A \to B$, se llama inversa de f a toda función $g : B \to A$ con las propiedades siguientes: $g \circ f = I_A$ y $f \circ g = I_B$.*

Según el teorema anterior, una función posee una inversa si y sólo si es biyectiva.

Proposición 4.9 *Si $f : A \to B$ posee una inversa, ésta es única.*

Proposición 4.10 *Sean las funciones $f : A \to B$ y $g : B \to C$.*

Si f y g son inyectivas	*entonces*	*$g \circ f$ es inyectiva.*
Si f y g son epiyectivas	*entonces*	*$g \circ f$ es epiyectiva.*
Si f y g son biyectivas	*entonces*	*$g \circ f$ es biyectiva.*

4.3 Relación de equivalencia asociada a una función

Una de las formas más frecuentes de definir una relación de equivalencia en un conjunto A que sea el dominio de una cierta función f, es considerar como equivalentes los puntos de A en los que f toma el mismo valor.

En efecto, sea dada una función $f : A \to B$. Si definimos $x \sim y \Leftrightarrow f(x) = f(y)$, para $x, y \in A$, entonces $\sim$ es una relación de equivalencia. La verificación es sencilla y se deja como ejercicio.

Así pues, toda función determina una relación de equivalencia en su dominio, de la que diremos que es *asociada* a la función.

Proposición 4.12 *Sea A un conjunto y sea R una relación de equivalencia definida sobre A. Entonces la aplicación natural $\pi : A \to A/R$ es una función epiyectiva cuya relación de equivalencia asociada es R.*

Teorema 5.5 (Criterio de demostración.)
Sea $P(n)$ una proposición relativa al número natural n y supongamos que:

1. *$P(1)$ es verdadera.*
2. *Para todo $k \in \mathbb{N}$, si $P(k)$ es verdadera, entonces $P(k+1)$ es verdadera.*

Entonces $P(n)$ es verdadera para todo $n \in \mathbb{N}$.

Teorema 5.8 (Segunda forma del Principio de Inducción).
Sea $P(n)$ una proposición sobre el número natural n y supongamos que:

a) *$P(1)$ es verdadera.*

b) *Para todo $n \in \mathbb{N}$, $n > 1$, si $P(k)$ es verdadera para todo $k < n$, entonces $P(n)$ es verdadera.*

Entonces $P(n)$ es verdadera para todo $n \in \mathbb{N}$.

Teorema 5.6 *$\mathbb{N}$ es cerrado con respecto a la suma y a la multiplicación.*

(a) *Si $a, b \in \mathbb{N}$, entonces $a + b \in \mathbb{N}$.*

(b) *Si $a, b \in \mathbb{N}$, entonces $a \cdot b \in \mathbb{N}$.*

Proposición 5.9 *Si $x, y \in \mathbb{Z}$ entonces $x + y \in \mathbb{Z}$, $x \cdot y \in \mathbb{Z}$ y $x - y \in \mathbb{Z}$.*

Proposición 5.10 *Si x, $y \in \mathbb{Q}$ entonces $x + y \in \mathbb{Q}$, $x \cdot y \in \mathbb{Q}$, $x - y \in \mathbb{Q}$ y $\frac{x}{y} \in \mathbb{Q}$ $(y \neq 0)$.*

Teorema 5.15 *Sea a un número real positivo y sea n un número natural. Entonces existe un único número positivo b tal que $b^n = a$. Se nota $b = \sqrt[n]{a}$, y decimos que b es la raíz n-ésima de a.*

En base al teorema anterior, podemos asegurar, por ejemplo, la existencia de $\sqrt{2}$. Vamos a probar que $\sqrt{2}$ no es racional.

Supongamos que $\sqrt{2}$ es racional, esto es, existen números enteros a y b, $b \neq 0$, tales que:

$$\sqrt{2} = \frac{a}{b},$$

es decir, existen dos enteros a y b tales que $\left(\frac{a}{b}\right)^2 = 2$. Podemos suponer que $\frac{a}{b}$ es una fracción irreducible, esto es, a y b no poseen factores comunes propios.

De lo anterior resulta $a^2 = 2\,b^2$, lo que implica que a^2 es par, esto es, a es par. $(*)$
Tenemos que $a = 2\,c$ y entonces $4\,c^2 = 2\,b^2$, es decir, $2\,c^2 = b^2$. Razonando como antes obtenemos que b^2 es par concluyendo que b es par. $(**)$

$(*)$ y $(**)$ contradicen la hipótesis sobre la irreducibilidad de la fracción $\frac{a}{b}$.

Definición 6.1 *Sean $a, b \in \mathbb{Z}$. Se dice que a divide a b, y escribimos $a \mid b$ (con una barra vertical), si existe $k \in \mathbb{Z}$ tal que $b = a \cdot k$.*

En este caso diremos también que a *es un factor* de b, que b *es divisible* por a ó que b *es múltiplo* de a .

Si a no divide a b escribimos $a \not| \; b$.

(2) Si $a \mid b$ y $b \mid c$ entonces $a \mid c$.

(4) Si $c \mid a$ y $c \mid b$ entonces $c \mid ax + by$, $x, y \in \mathbb{Z}$.

(1) $c \mid a + b$ no implica que $c \mid a$ ó $c \mid b$. Por ejemplo, $6 \mid 4 + 8$, pero $6 \not| \; 4$ y $6 \not| \; 8$.

(2) Sin embargo, si $c \mid a + b$ y se sabe que $c \mid a$ entonces $c \mid b$ (pues $c \mid (a + b) - a$).

6.1 Algoritmo de la división entera

Teorema 6.2 (Algoritmo de la división entera). *Dados dos enteros a y b, $b \neq 0$ existen enteros q y r, llamados respectivamente el cociente y el resto de dividir a por b, unívocamente determinados tales que:*

$$a = b \cdot q + r \quad \textit{con} \quad 0 \leq r < |b|.$$

Sean a, b dos enteros, y supongamos que al menos uno de ellos es distinto de cero. Indiquemos $D(a, b)$ el conjunto de todos los divisores comunes de a y b, es decir,

$$D(a, b) = \{c \in \mathbb{Z} : c \mid a \text{ y } c \mid b\}.$$

Es claro que $D(a, b) \neq \emptyset$, y como $D(a, b) = D(a) \cap D(b)$, se tiene que $D(a, b)$ es finito.

El mayor de todos los elementos de $D(a, b)$ se llama el *máximo común divisor* de a y b, y se nota (a, b).

Proposición 6.3 *Si a y b son enteros, $b \neq 0$ y r es el resto de dividir a por b entonces $D(a, b) = D(b, r)$.*

Algoritmo de Euclides (Existencia del máximo común divisor)

	q_1	q_2	q_3			q_{n-1}	q_n	q_{n+1}
a	b	r_1	r_2	$\cdots\cdots\cdots$	r_{n-3}	r_{n-2}	r_{n-1}	r_n
r_1	r_2	r_3			r_{n-1}	r_n	0	

Teorema 6.5 *Dados* $a, b \in \mathbb{Z}, b \neq 0$, *existen enteros* x *e* y *tales que* $(a, b) = xa + yb$.

Definición 6.7 *Dos enteros* a *y* b *se dicen relativamente primos o coprimos si* $(a, b) = 1$.

Proposición 6.8 (Euclides). *Sean* $a, b, c \in \mathbb{Z}$. *Si* $c \mid a \cdot b$ *y* $(a, c) = 1$ *entonces* $c \mid b$.

Proposición 6.9 *Si* $a \mid n$ *y* $b \mid n$ *y* $(a, b) = 1$, *entonces* $a \cdot b \mid n$.

Observación. Si la ecuación diofántica $ax + by = c$ tiene una solución (x_0, y_0), entonces todas las soluciones de esta ecuación están dadas por los pares de enteros (x, y) tales que

$$\begin{cases} x = x_0 + t \cdot b \\ y = y_0 - t \cdot a \end{cases}, \quad t \in \mathbb{Z} \qquad (*)$$

En efecto, es fácil ver que cualquier par de enteros de la forma $(*)$ satisface la ecuación, ya que $a(x_0 + t \cdot b) + b(y_0 - t \cdot a) = ax_0 + by_0 = c$.

Sean a y b dos enteros y notemos $M(a, b)$ el conjunto de todos los múltiplos no negativos comunes de a y b. Es claro que $M(a, b) = M(a) \cap M(b)$. Vamos a probar el siguiente

Teorema 6.12 *Si* a *y* b *son dos enteros, entonces existe* $m \in M(a, b)$ *tal que si* $c \in M(a, b)$, *entonces* $m \mid c$.

Teorema 6.16 (Teorema Fundamental de la Aritmética). *Todo número entero* a *distinto de* $0, 1, -1$, *o bien es un número primo, o bien se puede escribir como* ± 1 *por un producto de números primos positivos. Esta representación de un entero como producto de primos es única, salvo el orden de los factores.*

Teorema 6.17 *Sea* $n \in \mathbb{Z}$, $n > 1$. *Si* n *no es primo entonces existe un primo* p *tal que* $p \mid n$ *y* $p \leq \sqrt{n}$.

Ejemplo. Sea $a = 75 = 3^1 \cdot 5^2$. Los divisores positivos de a son: $3^0 \cdot 5^0 = 1$, $3^0 \cdot 5^1 = 5$, $3^0 \cdot 5^2 = 25$, $3^1 \cdot 5^0 = 3$, $3^1 \cdot 5^1 = 15$, $3^1 \cdot 5^2 = 75$.

Si $a > 1$, $a = p_1^{e_1} \cdot p_2^{e_2} \cdot \dots \cdot p_s^{e_s}$ y con $d(a)$ notamos al número de los divisores positivos de a, resulta en forma inmediata del teorema anterior que: $d(a) = (e_1 + 1) \cdot (e_2 + 1) \cdot \dots \cdot (e_s + 1)$.

Corolario 6.19 *Si* $a = p_1^{e_1} \cdot p_2^{e_2} \cdot \cdots \cdot p_s^{e_s}$ *y* $b = p_1^{t_1} \cdot p_2^{t_2} \cdot \cdots \cdot p_s^{t_s}$, $e_i \geq 0$, $t_i \geq 0$, *entonces*

$$(a, b) = p_1^{m_1} \cdot p_2^{m_2} \cdot \cdots \cdot p_s^{m_s}$$

y

$$[a, b] = p_1^{M_1} \cdot p_2^{M_2} \cdot \cdots \cdot p_s^{M_s},$$

donde m_i *y* M_i *representan, respectivamente, el menor y el mayor de los números* e_i *y* t_i.

Demostración. Es inmediata. □

Ejemplo. De $280 = 2^3 \cdot 5 \cdot 7$ y $693 = 3^2 \cdot 7 \cdot 11$ podemos escribir

$$\begin{aligned} 280 &= 2^3 \cdot 3^0 \cdot 5 \cdot 7 \cdot 11^0 \\ 693 &= 2^0 \cdot 3^2 \cdot 5^0 \cdot 7 \cdot 11. \end{aligned}$$

Luego

$$(280, 693) = 2^0 \cdot 3^0 \cdot 5^0 \cdot 7 \cdot 11^0 = 7$$
$$[280, 693] = 2^3 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 = 27720.$$

Teorema 6.20 *Sea* $b \in \mathbb{N}$, $b > 1$. *Para todo entero* $a > 0$, *existen únicos enteros* $a_0, a_1, \ldots, a_n$, *con* $0 \leq a_i < b$, $i = 0, 1, \ldots, n$ *y* $a_n > 0$ *tales que*

$$a = a_n b^n + \cdots + a_2 b^2 + a_1 b + a_0.$$

Escribimos $a = a_n \ldots a_2 a_1 a_{0(b)}$, *ó* $a = a_n \ldots a_2 a_1 a_0$, *si no hay riesgo de confusión. Los coeficientes* a_i *se llaman las cifras y* b *se llama la base de la representación.*

7.1 Definición y propiedades

Definición 7.1 *Sea* $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ *el conjunto de todos los pares ordenados* (a, b) *de números reales sobre el cual definimos las operaciones siguientes:*

1. **Suma:** *Dados* (a, b) *y* $(c, d) \in \mathbb{C}$, $(a, b) + (c, d) = (a + c, b + d)$.
2. **Producto:** *Dados* (a, b) *y* $(c, d) \in \mathbb{C}$, $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$.

El conjunto $\mathbb{C} = \{(a, b) : a,\ b \in \mathbb{R}\}$, *junto con las operaciones de suma y producto definidas, recibe el nombre de conjunto de los números complejos.*

(a) Suma: $z + z' = (a + b\,i) + (c + d\,i) = (a + c) + (b + d)\,i$

(b) Diferencia: $z - z' = (a + b\,i) - (c + d\,i) = (a - c) + (b - d)\,i$

(c) Producto: $z \cdot z' = (a + b\,i) \cdot (c + d\,i) = (ac - bd) + (ad + bc)\,i$

(d) Cociente: $\dfrac{z}{z'} = \dfrac{a + b\,i}{c + d\,i}, \quad z' \neq 0.$

Definición 7.3 *Dado el número complejo $z = (a, b)$ (ó $z = a+b\,i$), llamaremos* **conjugado** *de z al número complejo $\overline{z} = (a, -b)$ (ó $\overline{z} = a - b\,i$).*

Definición 7.4 *Dado un número complejo $z = a + b\,i$, se llama* módulo *de z al número real, positivo o nulo, $\rho = \sqrt{a^2 + b^2}$. Lo notaremos $\rho = |z|$.*

Definición 7.5 *Dado un número complejo $z = a + bi$, $z \neq 0$, se llama* **argumento** *de z, al ángulo formado por el semieje real positivo y el segmento $\overline{OP}$, y a menos de un múltiplo entero de 2π, tomándose como sentido positivo el sentido antihorario.*

Si α es el argumento de z, también serán argumento de z los ángulos de la forma: $\alpha + 2k\pi$, con $k \in \mathbb{Z}$, o sea, que los valores que toma "argumento de z" son:

$$\ldots,\ \alpha - 3\cdot 2\pi,\ \alpha - 2\cdot 2\pi,\ \alpha - 2\pi,\ \alpha,\ \alpha + 2\pi,\ \alpha + 2\cdot 2\pi,\ \ldots$$

Llamaremos **argumento principal** de z, al argumento de z comprendido entre 0 y 2π, es decir, que α es el argumento principal de z si $0 \leq \alpha < 2\pi$. Para indicar que α es el argumento de z, escribiremos $\alpha = \arg z$.

Esta forma de escribir los números complejos se llama **forma polar o trigonométrica** y suele abreviarse también: $z = \rho_\alpha$.

Observando la figura resulta que:

$$\rho = \sqrt{a^2 + b^2} = |z| \quad \text{y} \quad \tan\alpha = \frac{b}{a}\,, \quad \text{esto es,} \quad \alpha = \arctan\frac{b}{a}.$$

$$\begin{aligned} z = a + bi &= \rho\cdot\cos\alpha + \rho\cdot\operatorname{sen}\alpha\cdot i \\ z = a + bi &= \rho(\cos\alpha + i\operatorname{sen}\alpha) \end{aligned}$$

Observaciones

1. El argumento de $z = (0, 0)$ no está definido.
2. El número complejo z es real positivo $\Leftrightarrow$ su argumento principal es $\alpha = 0$. Por ejemplo, $3 = 3\cdot\text{cis}\,0$.
3. El número complejo z es real negativo $\Leftrightarrow$ su argumento principal es $\alpha = \pi$. Así, $-2 = 2\cdot\text{cis}\,\pi$.
4. El número complejo z es imaginario puro $\Leftrightarrow$ su argumento principal es $\alpha = \dfrac{\pi}{2}$ ó $\alpha = 3\,\dfrac{\pi}{2}$. Por ejemplo, $i = \text{cis}\,\dfrac{\pi}{2}$, $-i = \text{cis}\,\dfrac{3\pi}{2}$.

7.2 Operaciones en forma polar

Producto de números complejos en forma polar

Teorema 7.6 *Si* $z = \rho \text{ cis } \alpha$ *y* $z' = \rho' \text{ cis } \alpha'$, *entonces* $z \cdot z' = \rho\rho' \text{cis } (\alpha + \alpha')$

Cociente de números complejos en forma polar

$$\frac{z}{z'} = \frac{\rho}{\rho'} \text{ cis } (\alpha - \alpha').$$

Potenciación de números complejos

Teorema 7.7 *Sea* $z \in \mathbb{C}$, $z = \rho(\cos \alpha + i \text{ sen } \alpha)$. *Entonces para todo* $n \in \mathbb{Z}$, *es:*

$$z^n = [\rho(\cos \alpha + i \text{ sen } \alpha)]^n = \rho^n(\cos n\alpha + i \text{ sen } n\alpha).$$

Radicación de números complejos

Dado un número complejo z y $n \in \mathbb{N}$, llamaremos raíz n-ésima de z a todo número complejo w tal que $w^n = z$. Notaremos con $\sqrt[n]{z}$ al conjunto de todas las raíces n-ésimas de z.

Se plantea entonces el problema de averiguar si dado un número complejo z, existe siempre alguna raíz n-ésima de z, y se demuestra el siguiente teorema.

Teorema 7.8 *Dado un número complejo* $z = \rho \text{ cis } \alpha$, *y* $n \in \mathbb{N}$, *existen exactamente* n *raíces* n*-ésimas de* z, *que son de la forma*

$$r_k = \sqrt[n]{\rho} \text{ cis } \frac{\alpha + 2k\pi}{n}, \quad \textit{con } k = 0, 1, 2, 3, \ldots, n-1.$$

7.3 Raíces de la unidad

Sea el número complejo $z = 1$. Entonces $\rho = 1$ y $\alpha = 0$, y por lo tanto, las raíces n-ésimas de 1 están dadas por:

$$\sqrt[n]{1 \text{ cis } 0°} = \sqrt[n]{1} \text{ cis } \frac{2k\pi}{n}, \text{ con } k = 0, 1, \cdots, n-1.$$

Una de las raíces n-ésimas de 1 es 1, que se obtiene para $k = 0$. Se puede observar que, si n es un número par, existen dos raíces reales que son $+1$ y -1, y si n es impar existe una sola raíz real, que es 1. Geométricamente, los afijos de las n raíces n-ésimas de la unidad, son los vértices de un polígono regular de n lados inscripto en una circunferencia de radio 1.

Raíces primitivas de la unidad

Consideremos el siguiente ejemplo: las raíces cuartas de la unidad son 1, i, -1 y $-i$. De estas cuatro raíces se observa que dos de ellas, 1 y -1, son tales que elevándolas a exponentes naturales **menores** que 4 también se obtiene la unidad. Por ejemplo: $1^1 = 1$, $1^2 = 1$, $1^3 = 1$, $(-1)^2 = 1$.

Sin embargo, las otras dos raíces, i y $-i$, son tales que el menor exponente natural al cual hay que elevarlas para obtener la unidad es 4, ya que: $i^1 = i$, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$ y $(-i)^1 = -i$, $(-i)^2 = -1$, $(-i)^3 = i$, $(-i)^4 = 1$.

Esta clasificación de las raíces cuartas de 1 nos lleva a la siguiente definición.

Definición 7.10 *Se llaman raíces primitivas de n-ésimo orden de la unidad, a aquellas raíces n-ésimas de la unidad que no son raíces de la unidad de un orden menor que n. Es decir, si ϵ es una raíz primitiva de n-ésimo orden de la unidad, entonces $\epsilon^n = 1$ y no existe $k \in \mathbb{N}$ tal que $k < n$ y $\epsilon^k = 1$.*

Teorema 7.11 *Las raíces n-ésimas primitivas de la unidad se obtienen dando a k en la fórmula*

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

valores relativamente primos con n y menores que n.

Teorema 7.12 *Sea u una raíz n-ésima de 1. u es una raíz n-ésima primitiva si y sólo si u^0, u, u^2, ..., u^{n-1} son todas las raíces n-ésimas de 1.*

Definición 8.4 *Se llama grado de un polinomio no nulo $f(X)$, y se nota gr $f(X)$, al mayor índice i tal que $a_i \neq 0$.*

Al polinomio nulo no se le atribuye grado.

Si $gr\ f(X) = n$, diremos que $a_n X^n$ es el *término principal* de $f(X)$, y que a_n es el *coeficiente principal*. $f(X)$ se dice *mónico* si $a_n = 1$.

Para **el grado del producto** de dos polinomios, vale la siguiente propiedad, para $f(X) \neq 0$ y $g(X) \neq 0$:

$$gr\ (f(X) \cdot g(X)) = gr\ f(X) + gr\ g(X).$$

Observación. Los únicos polinomios inversibles con respecto a la multiplicación son las constantes no nulas. En efecto, si un polinomio $f(X)$ tiene inverso $g(X)$, es decir, si $f(X) \cdot g(X) = 1$, entonces $gr\ f(X) + gr\ g(X) = 0$, de donde $gr\ f(X) = gr\ g(X) = 0$, y por lo tanto, $f(X)$ y $g(X)$ son constantes no nulas.

Teorema 8.5 (Algoritmo de la división). *Dados dos polinomios $f(X)$ y $g(X) \in K[X]$, con $g(X)$ no nulo, existen dos únicos polinomios $q(X)$ y $r(X) \in K[X]$, llamados cociente y resto respectivamente de dividir $f(X)$ por $g(X)$, tal que $f(X) = q(X)g(X) + r(X)$, donde $gr\ r(X) < gr\ g(X)$ ó $r(X) = 0$.*

Definición 8.6 *Dados dos polinomios $f(X), g(X) \in K[X]$, se dice que $f(X)$ divide a $g(X)$ si existe un polinomio $h(X) \in K[X]$ tal que $g(X) = h(X) \cdot f(X)$. Se escribe $f|g$.*

Definición 8.7 *Dados dos polinomios $f = f(X)$ y $g = g(X)$, un polinomio $d = d(X)$ se llama un máximo común divisor de $f(X)$ y $g(X)$ si se verifica que:*

1. *$d|f$ y $d|g$.*
2. *Si $d'(X)$ es un polinomio tal que $d'|f$ y $d'|g$, entonces $d'|d$.*

Notaremos $d = (f, g)$.

Si $d(X)$ y $d'(X)$ son dos m.c.d. de $f(X)$ y $g(X)$, por la propiedad 8, $d(X)$ y $d'(X)$ difieren en una constante. Si se considera el m.c.d. mónico, entonces el m.c.d. es *único*.

Análogamente a lo que sucede en $\mathbb{Z}$, es posible aplicar el algoritmo de Euclides para calcular el máximo común divisor de dos polinomios $f(X)$ y $g(X)$.

Definición 8.8 *Un polinomio no constante $f(X) \in K[X]$ se dice irreducible o primo en $K[X]$ si no se puede expresar como producto de dos polinomios no constantes de $K[X]$. (Equivalentemente, si los únicos divisores de $f(X)$ en $K[X]$ son las constantes no nulas y los polinomios $k \cdot f(X)$, $k \in K$, $k \neq 0$.)*

Definición 8.9 *Dos polinomios $f(X)$ y $g(X)$ se dicen relativamente primos si $(f, g) = 1$.*

Mencionemos los siguientes resultados fundamentales cuya demostración omitimos, pues es igual a la vista en $\mathbb{Z}$:

Sean $f(X)$, $g(X)$ y $h(X) \in K[X]$. Entonces:

1. Si $f|g \cdot h$ y $(f, g) = 1$, entonces $f|h$.
   En particular, si $f|g \cdot h$ y f es irreducible, entonces $f|g$ ó $f|h$.

2. Si $f|h$, $g|h$ y $(f, g) = 1$, entonces $f \cdot g|h$.

Teorema 8.10 (El Teorema Fundamental de la Aritmética en K[X]). *Todo elemento $f(X) \in K[X]$ no constante, puede escribirse en la forma:*

$$f = k \cdot p_1^{e_1} \cdot p_2^{e_2} \cdot \dots \cdot p_s^{e_s},$$

donde $k \in K$, p_i son polinomios irreducibles mónicos distintos en $K[X]$ y $e_i \in \mathbb{N}$. Esta factorización es única salvo el orden de los factores.

Teorema 8.11 (Teorema del resto, o Teorema de Bézout). *El resto de dividir un polinomio $f(X)$ por otro de la forma $X - a$ es $f(a)$, es decir, es el valor del polinomio en a.*

Teorema 8.18 (Teorema Fundamental del Algebra). *Todo polinomio no constante con coeficientes en $\mathbb{C}$, tiene por lo menos una raíz en $\mathbb{C}$.*

Corolario 8.19 *Los únicos polinomios irreducibles en $\mathbb{C}[X]$ son los de primer grado.*

Teorema 8.20 *Si z es una raíz compleja de un polinomio $f(X)$* **con coeficientes reales**, *entonces $\overline{z}$ también es raíz de $f(X)$. Además, z y $\overline{z}$ tienen el mismo orden de multiplicidad.*

Corolario 8.21 *En $\mathbb{R}[X]$, los únicos polinomios irreducibles son los de grado 1 y los de segundo grado de la forma $a(X^2 + bX + c)$, con $b^2 - 4c < 0$.*

Demostración. Es consecuencia inmediata de la observación anterior. □

Corolario 8.22 *Todo polinomio con coeficientes reales de grado impar tiene por lo menos una raíz real.*

Teorema 8.23 (Regla de Laguerre-Thibault.) *Sea $f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$, $a_i \in \mathbb{R}$, con $a_n > 0$. Si al dividir $f(X)$ por $X - a$, con $a \geq 0$, todos los coeficientes del cociente, y el resto, son no negativos, entonces a es una cota superior de las raíces reales de $f(X)$.*

2. Número de las raíces reales positivas y negativas de un polinomio con coeficientes reales

Daremos sin demostración la siguiente:

Regla de Descartes. El número de raíces reales positivas de un polinomio $f(X)$ con coeficientes reales, contadas tantas veces como su orden de multiplicidad, es *menor o igual* que el número de variaciones de signo de los coeficientes de $f(X)$, y difiere del mismo en un número par.

Lo mismo vale para el número de raíces reales negativas: es *menor o igual* que el número de variaciones de signo de los coeficientes de $f(-X)$ y difiere del mismo en un número par.

Ejemplo. Hallar el número de raíces reales positivas y negativas de

$$f(X) = X^4 - 3X^3 - 5X^2 + 7X - 3.$$

Como en los coeficientes de $f(X)$ hay tres variaciones de signo, el número de raíces reales positivas de $f(X)$ es 3 ó 1.

Por otro lado, $f(-X) = X^4 + 3X^3 - 5X^2 - 7X - 3$, y este polinomio tiene una sola variación de signo en sus coeficientes. Entonces el número de raíces reales negativas es 1.

3. Raíces racionales de un polinomio con coeficientes racionales.

Consideremos la siguiente ecuación: $2X^3 + \frac{1}{2}X^2 + \frac{1}{3}X + 2 = 0$.

Si la multiplicamos por el m.c.m. de los denominadores, o sea, por 6, es: $12X^3+3X^2+2X+12 = 0$, que es una ecuación con coeficientes enteros que tiene las mismas raíces que la anterior. Es decir, cuando tenemos un polinomio con coeficientes racionales, lo podemos transformar en un polinomio con coeficientes enteros que tiene las mismas raíces, multiplicándolo por el m.c.m. de los denominadores.

Teorema 8.24 (de Gauss). *Si un número racional* $\frac{p}{q}$, *con* p *y* q *relativamente primos, es raíz de un polinomio* $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ *con coeficientes enteros, entonces* $p|a_0$ *y* $q|a_n$.

En resumen: Dado un polinomio no constante $f(X)$, el procedimiento para hallar las raíces de $f(X)$ es el siguiente:

1. Se acotan las raíces reales.
2. Se determina el número de raíces reales positivas y negativas.
3. Se determinan las raíces racionales.
4. Una vez determinada una raíz de $f(X)$, se analiza su orden de multiplicidad y se obtiene un polinomio de menor grado que $f(X)$ cuyas raíces son también raíces de $f(X)$.

4. El Teorema de Bolzano-Weierstrass.

El siguiente teorema juega un papel importante en la localización de las raíces de un polinomio con coeficientes reales.

Teorema 8.25 (Bolzano-Weierstrass). *Sea* $f(X) \in \mathbb{R}[X]$ *y sean* $a, b \in \mathbb{R}$, $a < b$. *Si* $f(b)$ *y* $f(a)$ *son no nulos y tienen diferente signo, entonces existe un número* $c \in (a, b)$ *tal que* $f(c) = 0$.

4. **Criterio de divisibilidad por 11.** Un número entero es múltiplo de 11 si la suma algebraica alternada de sus cifras es múltiplo de 11.
Escribamos
$$a = a_na_{n-1}\ldots a_1a_0 = a_n10^n + a_{n-1}10^{n-1} + \ldots + a_110 + a_0$$

5. **Criterio de divisibilidad por 3.** Un número entero es múltiplo de 3 si la suma de sus cifras es múltiplo de 3.
Escribamos
$$a = a_na_{n-1}\ldots a_1a_0 = a_n10^n + a_{n-1}10^{n-1} + \ldots + a_110 + a_0$$
$$= a_n(9+1)^n + a_{n-1}(9+1)^{n-1} + \ldots + a_1(9+1) + a_0$$
$$= a_n\sum_{k=0}^{n}\binom{n}{k}9^{n-k}\cdot 1^k + a_{n-1}\sum_{k=0}^{n-1}\binom{n-1}{k}9^{n-1-k}\cdot 1^k + \ldots + a_1(9+1) + a_0.$$

10 Sistemas de ecuaciones lineales, matrices y determinantes

10.1 Matrices

Una **matriz** $m \times n$ (o de orden $m \times n$) es un cuadro de números con m filas (horizontales) y n columnas (verticales):

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

Definición 10.1 *La suma de dos matrices A y B de orden $m \times n$ es otra matriz $m \times n$ que se obtiene sumando los elementos correspondientes de A y B. En símbolos, si $A = (a_{ij})$ y $B = (b_{ij})$, entonces $A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$.*

Propiedades

Sean A, B, C matrices $m \times n$. Entonces :

(S_1) $(A + B) + C = A + (B + C)$.

(S_2) $A + B = B + A$.

(S_3) La matriz de orden $m \times n$, $0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$ es tal que $0 + A = A + 0 = A$ para toda matriz de orden $m \times n$ A.

Definición 10.2 *El producto de un número real k y la matriz A es la matriz $k \cdot A$ que se obtiene multiplicando por el número k cada uno de los elementos de A. En símbolos,*

$$k \cdot A = k \cdot (a_{ij}) = (ka_{ij}).$$

Propiedades

Sean A, B matrices de orden $m \times n$ y λ, λ' números complejos. Entonces:

(1) $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$.

(2) $(\lambda + \lambda') \cdot A = \lambda \cdot A + \lambda' \cdot A$.

(3) $(\lambda \lambda') \cdot A = \lambda \cdot (\lambda' \cdot A)$.

(4) $1 \cdot A = A$.

Definición 10.3 *Sea A una matriz $m \times n$, y sea B una matriz $n \times p$. El producto de A y B es la matriz $A \cdot B$ de orden $m \times p$ definida por:*

$$A \cdot B = (c_{ij}),$$

Propiedades Siempre que el producto pueda efectuarse, se verifican

(M_1) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

(M_2) $A \cdot (B + C) = A \cdot B + A \cdot C$, $(B + C) \cdot A = B \cdot A + C \cdot A$.

(M_3) Se llama matriz identidad de orden n a una matriz $n \times n$ en la que $a_{ii} = 1$ para todo i, $1 \leq i \leq n$, y $a_{ij} = 0$ si $i \neq j$. Se nota I ó I_n.

Propiedades

(1) $(A^T)^T = A$.

(2) $(A + B)^T = A^T + B^T$.

(3) $(\lambda \cdot A)^T = \lambda \cdot A^T;\ \ \lambda \in \mathbb{C}$.

(4) $(A \cdot B)^T = B^T \cdot A^T$, si el producto $A \cdot B$ está definido.

10.2 Matrices y sistemas de ecuaciones

Consideremos un sistema de dos ecuaciones lineales con dos incógnitas:

$$\left\{ \begin{array}{ccccc} ax & + & by & = & e \\ cx & + & dy & = & f \end{array} \right.$$

(El nombre de lineales, que también se utiliza para mayor número de incógnitas, se debe a que las ecuaciones del tipo $ax + by = e$ se representan en el plano por una recta.)

Un par ordenado de números (x_0, y_0) se llama *solución* del sistema si al reemplazar x por x_0 e y por y_0 se satisfacen ambas ecuaciones. Geométricamente, esto sucede cuando el punto (x_0, y_0) es el punto de intersección de las rectas r_1 y r_2, de ecuación $ax + by = e$ y $cx + dy = f$, respectivamente.

Hay tres posibilidades para el conjunto de soluciones:

1. Puede ser vacío, es decir, las rectas r_1 y r_2 no se cortan. Decimos entonces que el sistema es *incompatible.*

2. Tiene exactamente un punto. Geométricamente significa que las rectas se cortan. Decimos que el sistema es *compatible determinado.*

3. Contiene infinitos puntos, es decir, r_1 y r_2 son coincidentes. En este caso decimos que el sistema es *compatible indeterminado.*

1. Intercambiar dos filas de la matriz.

2. Reemplazar una fila por la que se obtiene al multiplicarla por un número distinto de cero.

3. Reemplazar una fila por la que se obtiene al sumarle a esa fila otra fila previamente multiplicada por un número.

Estas operaciones elementales transforman la matriz del sistema en otra matriz que corresponde a un sistema de ecuaciones equivalente al dado.

La aplicación de las operaciones elementales tiene por objetivo obtener una matriz que tenga:

- Ceros debajo de la diagonal principal.

- Unos en la diagonal principal (eventualmente puede aparecer algún cero).

10.3 Determinantes

Determinantes de segundo y tercer orden

Dada una matriz cuadrada de 2×2, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, el valor de su determinante se define como

$$\det\, A = |A| = a_{11}a_{22} - a_{12}a_{21}.$$

Consideremos ahora una matriz cuadrada de tercer orden

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

Entonces el determinante de A se define como

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

donde cada término consta de un producto de tres factores, uno de cada fila y uno de cada columna. Si a cada término le asociamos la permutación σ de $\{1,2,3\}$ determinada por los subíndices (por ejemplo, al término $a_{11}a_{23}a_{32}$ le asociamos $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ entonces se asigna a ese término el signo $(-1)^s$, donde s es el número de inversiones de la permutación σ.

Definición 10.4 *Se llama determinante de la matriz cuadrada A a la suma de los $n!$ productos de la forma $(-1)^s a_{1\alpha_1}a_{2\alpha_2}\cdots a_{n\alpha_n}$. Se nota* $\det\ A$ *ó* $|A|$.

$$\det\ A = |A| = \sum (-1)^s a_{1\alpha_1}a_{2\alpha_2}\cdots a_{n\alpha_n}.$$

$\det\ A$ *se dice un determinante de orden n.*

Propiedad 10.5 *El determinante de una matriz coincide con el determinante de su traspuesta. Es decir,* $\det\ A = \det\ A^T$.

Propiedad 10.6 *Si una de las columnas de la matriz A está constituida por ceros, entonces* $\det\ A = 0$.

Propiedad 10.7 *Si A' es la matriz que se obtiene de la matriz A intercambiando dos columnas, entonces* $\det\ A' = -\det\ A$.

Propiedad 10.8 *Si A es una matriz con dos columnas iguales, entonces* $\det\ A = 0$.

Propiedad 10.9 *Si A' es la matriz que se obtiene multiplicando todos los elementos de una columna de la matriz A por un número k, entonces* $\det\ A' = k \cdot \det\ A$.

Propiedad 10.10 *Si A es una matriz que tiene dos columnas proporcionales, entonces* $\det\ A = 0$.

Propiedad 10.11 *Si $A = (a_{ij})$ y la columna r-ésima se puede expresar $a_{ir} = b_{ir} + c_{ir}$, $1 \leq i \leq n$, entonces* $\det\ A = \det\ B + \det\ C$ *donde B y C coinciden con A salvo en la columna r-ésima, en la que B tiene los elementos b_{ir} y C tiene los elementos c_{ir}, $1 \leq i \leq n$.*

Propiedad 10.12 *Si en una matriz A una columna es combinación lineal de otras, entonces* $\det\ A = 0$.

Propiedad 10.13 *Si A' es la matriz que se obtiene al sumar a una columna de una matriz A una combinación lineal de otras columnas de A, entonces* $\det\ A' = \det\ A$.

Cálculo de determinantes

De las propiedades anteriores resulta que si se pasa de una matriz A a otra matriz por medio de una operación elemental, entonces

(I) El determinante cambia de signo si se intercambian dos columnas (filas).

(II) El determinante queda multiplicado por un escalar no nulo si se reemplaza una columna (fila) por un múltiplo no nulo de esa columna (fila).

(III) El determinante no varía si a una columna (fila) se le suma un múltiplo de otra.

Por consiguiente, un método para calcular el determinante de una matriz A consiste en transformar A en una matriz A' triangular, cuyo determinante es de cálculo inmediato. Según lo anterior, $\det A$ y $\det A'$ diferirán a lo sumo en un escalar.

Desarrollo de un determinante por los elementos de una línea (fila o columna).

Dada una matriz $A = (a_{ij})$ se llama **menor complementario** del elemento a_{ij} y se nota M_{ij} al determinante de la matriz que se obtiene a partir de A suprimiendo la i-ésima fila y la j-ésima columna.
Se llama **complemento algebraico o cofactor** del elemento a_{ij} de A al número $(-1)^{i+j} \cdot M_{ij}$.

Lema 10.14 *Sea a_{ij} un elemento cualquiera de una matriz cuadrada A. Si en $D = \det A$ agrupamos todos los términos que contienen a a_{ij} y escribimos $D = a_{ij} \cdot A_{ij} + ($términos que no contienen a $a_{ij})$, entonces A_{ij} es el complemento algebraico del elemento a_{ij}, esto es $A_{ij} = (-1)^{i+j} \cdot M_{ij}$.*

Teorema 10.15 *El determinante de una matriz $A = (a_{ij})$ es igual a la suma de los elementos de una línea (fila o columna) multiplicados cada uno de ellos por sus respectivos complementos algebraicos.*

Corolario 10.16 *Si $A = (a_{ij})$, la suma de los elementos de una fila (columna) multiplicados cada uno de ellos por los complementos algebraicos de los elementos correspondientes de otra fila (columna) distinta es cero.*

Es decir, $a_{s1} \cdot A_{t1} + a_{s2} \cdot A_{t2} + \cdots + a_{sn} \cdot A_{tn} = 0$ y $a_{1s} \cdot A_{1t} + a_{2s} \cdot A_{2t} + \cdots + a_{ns} \cdot A_{nt} = 0$, si $s \neq t$.

Propiedad 10.17 *Si A es una matriz de orden n y α es un número entonces* $\det(\alpha \cdot A) = \alpha^n \cdot \det A$.

Propiedad 10.18 *El determinante de un producto de matrices es el producto de los determinantes de cada una de ellas, es decir, si A y B son matrices cuadradas del mismo orden, entonces* $\det(A \cdot B) = \det A \cdot \det B$.

Matriz Inversa

Si A es una matriz de orden n y si existe una matriz B tal que $A \cdot B = B \cdot A = I_n$, donde I_n es la matriz unidad de orden n, entonces A se dice **inversible** y B se llama la **inversa** de A.

Es fácil ver que si existe una matriz B en esas condiciones, es única. Se escribe $B = A^{-1}$.

No toda matriz tiene inversa. Por ejemplo, si $A = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}$ no existe ninguna matriz B tal que $A \cdot B = B \cdot A = I_n$.

Definición 10.19 *Si $A = (a_{ij})$ es una matriz de orden n, definimos la* **matriz adjunta** *de A, y la notamos $Adj A$, de la siguiente manera: $Adj A = (A_{ij})$ donde A_{ij} es el complemento algebraico del elemento a_{ij}.*

$$Adj A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n1} & \cdots & A_{nn} \end{pmatrix}$$

Teorema 10.20 *Sea A una matriz de orden n. Entonces A es inversible si y sólo si* $\det A \neq 0$. *En ese caso,*

$$A^{-1} = \frac{(Adj\ A)^T}{\det\ A}.$$

10.4 Característica de una matriz

Sea $A = (a_{ij})$ una matriz $n \times m$. Una *submatriz* de A es una matriz construída tomando la intersección de ciertas filas y columnas de A. Por ejemplo,

$$\begin{pmatrix} 1 & -1 & 3 \\ 4 & -1 & 2 \end{pmatrix} \quad \text{es una submatriz de} \quad \begin{pmatrix} 1 & -1 & 2 & 3 \\ 3 & 0 & 1 & 4 \\ 4 & -1 & 5 & 2 \end{pmatrix}.$$

Se llama *menor de A de orden k* al determinante de cualquier submatriz B de A de orden $k \times k$.

Definición 10.21 *La* CARACTERÍSTICA *(o* RANGO*) de una matriz A $n \times m$ es el número entero k = Car A definido por la siguiente condición: existe un menor de A de orden k no nulo, y no existe ningún menor de A de orden mayor que k no nulo. La matriz de un tal menor de orden k se llama* submatriz principal. *También el* menor *se llama* principal.

Observación. Si en la matriz A hay un menor de orden k no nulo y todos los menores de orden $k+1$ son nulos, entonces Car $A = k$. En efecto, cualquier menor de orden $k+2$ será también cero (desarrollarlo por los elementos de una fila o columna). De la misma manera, los menores de orden $k+3$, $k+4$, ... son nulos.

Sea A una matriz de orden $n \times m$. Si indicamos con $C_1, C_2, \ldots, C_m$ las columnas de A, diremos que la columna C_i es combinación lineal de las columnas C_j y C_k, con $1 \leq j, k \leq m$, si existen números α, β tales que $C_i = \alpha \cdot C_j + \beta \cdot C_k$. En forma análoga se define una combinación lineal de un número cualquiera de columnas, y una combinación lineal de dos o más filas.

Lema 10.22 *Sea $A = (a_{ij})$ una matriz de orden n. Si Car $A = n - 1$, entonces una columna (fila) de A es combinación lineal de las columnas (filas) de una submatriz principal cualquiera de A.*

Proposición 10.23 *Todas las columnas (filas) de una matriz $n \times m$ son combinación lineal de las columnas (filas) de una submatriz principal.*

Proposición 10.24 *Sea A una matriz y sea B la matriz obtenida de A al sumar a una fila (columna) de A una combinación lineal de las demás. Sean $i_1, i_2, \ldots, i_k$ y $j_1, j_2, \ldots, j_k$ filas y columnas respectivamente de una matriz principal de A. Entonces $i_1, i_2, \ldots, i_k$ y $j_1, j_2, \ldots, j_k$ son filas y columnas de una matriz principal de B. En particular, A y B tienen la misma característica.*

2. Primeras operaciones algebraicas

La suma y el producto por escalares en $\mathbb{R}^2$ y $\mathbb{R}^3$ se definen en forma similar a lo que hicimos con matrices, es decir, componente a componente. Por ejemplo, con esas definiciones tendríamos

- $(1, 3) + (-1, 7) = (0, 10)$
- $\alpha(1 + p, q, r^2) = (\alpha + \alpha p, \alpha q, \alpha r^2)$ si $\alpha \in \mathbb{R}$
- ...etc

Tienen las mismas propiedades que las respectivas operaciones de matrices, por ejemplo, las que nos permiten escribir $(\alpha + \beta)(x, y, z) = \alpha(x, y, z) + \beta(x, y, z)$.

Definimos la **suma de vectores** (pensados como flechas) con la llamada *regla del paralelogramo*, como sigue. Para sumar los vectores $\mathbf{a} = (a_1, a_2, a_3)$ y $\mathbf{b} = (b_1, b_2, b_3)$ consideramos el paralelogramo que tiene $\mathbf{a}$ y $\mathbf{b}$ como lados adyacentes (ver figura 1). La suma $\mathbf{a} + \mathbf{b}$ se define como el vector que va a lo largo de la diagonal.

El **producto de un escalar por un vector** también tiene una descripción geométrica. Si λ (lambda) es un escalar (es decir, un número real) y $\mathbf{a}$ es un vector, se define $\lambda\mathbf{a}$ como sigue:

- Si $\lambda > 0$, entonces $\lambda\mathbf{a}$ tiene la misma dirección y sentido que $\mathbf{a}$, pero su longitud se multiplica por λ.
- Si $\lambda < 0$, entonces $\lambda\mathbf{a}$ tiene sentido opuesto al de $\mathbf{a}$, y su longitud se multiplica por $|\lambda|$ (por ejemplo, $-2\mathbf{a}$ medirá el doble que $\mathbf{a}$).
- Si $\lambda = 0$, entonces $\lambda\mathbf{a} = \mathbf{0}$.

A decir verdad, si $\mathbf{a} = \mathbf{0}$ la descripción anterior no se puede aplicar literalmente (porque no se le asigna dirección ni sentido), pero simplemente es $\lambda\mathbf{0} = \mathbf{0}$ para todo $\lambda \in \mathbb{R}$.

Definimos a los **vectores**, también llamados *vectores libres*, como segmentos dirigidos en el plano o en el espacio (a diferencia de antes, no pedimos que partan del origen). Aquellos segmentos que se pueden obtener uno de otro mediante traslaciones paralelas (sin rotaciones) se consideran como el mismo vector.

5. Módulo

El módulo de un vector $\mathbf{a}$ es su longitud. Se denota $\|\mathbf{a}\|$, y es un número real no negativo. Puede dar 0, pero sólo en el caso de que $\mathbf{a}$ sea el vector nulo.

En el plano, si $\mathbf{a} = (a_1, a_2)$, usando el teorema de Pitágoras se tiene $\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2}$. En el espacio, si $\mathbf{a} = (a_1, a_2, a_3)$ entonces $\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$. En realidad estas fórmulas son válidas si elegimos ejes que sean perpendiculares entre sí y con la misma unidad de medida, quedando el módulo expresado en esa misma unidad de medida. Existen situaciones donde esto no es así.

Observemos, de la definición de producto por un escalar, que $\|\lambda\mathbf{a}\| = |\lambda|\,\|\mathbf{a}\|$. Es decir, los escalares "salen" del módulo, pero en valor absoluto.

Un vector de módulo 1 recibe el nombre de *vector unitario*. Notemos que podemos "transformar" cualquier vector no nulo en uno unitario dividiéndolo por su módulo. Esto es, si $\mathbf{a} \neq \mathbf{0}$, entonces $\mathbf{a}/\|\mathbf{a}\|$ tiene la misma dirección y sentido que $\mathbf{a}$, y módulo 1. Este proceso se llama *normalizar* el vector $\mathbf{a}$.

6. Base standard de vectores

Una vez elegido un sistema de coordenadas, es conveniente definir tres vectores particulares a lo largo de los ejes x, y, z:

$$\begin{aligned} \mathbf{i} &= (1,0,0) \\ \mathbf{j} &= (0,1,0) \\ \mathbf{k} &= (0,0,1) \end{aligned}$$

Estos forman la llamada *base standard de vectores* (figura 3). También se los denota a veces $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. En el plano, tendremos una componente menos y un vector menos: $\mathbf{i} = (1,0)$, $\mathbf{j} = (0,1)$.

7. Producto escalar

Dados dos vectores $\mathbf{a} = (a_1, a_2, a_3)$ y $\mathbf{b} = (b_1, b_2, b_3)$ se define una nueva operación llamada *producto escalar* mediante

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

El resultado es un número real (un *escalar*). A veces a esta operación se la llama producto interno o producto punto. Una notación alternativa es $\langle \mathbf{a}, \mathbf{b} \rangle$ en vez de $\mathbf{a} \cdot \mathbf{b}$. Observemos que hay una similitud con las operaciones que realizamos entre filas y columnas al calcular producto de matrices.

Para vectores en el plano, la operación es similar solo que sin el término final.

8. Producto vectorial

Se define únicamente para vectores en el espacio (no en el plano), y se calcula así. Si $\mathbf{a} = (a_1, a_2, a_3)$ y $\mathbf{b} = (b_1, b_2, b_3)$ su producto vectorial es

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1).$$

El resultado es un vector. A esta operación se la llama también producto cruz. A veces van a ver la notación $\mathbf{a} \wedge \mathbf{b}$ en vez de $\mathbf{a} \times \mathbf{b}$. Sin embargo, no la usaremos ya que la notación $\mathbf{a} \wedge \mathbf{b}$ se usa habitualmente en física y matemática para denotar otro tipo de objeto llamado *bivector*, que si bien tiene similitudes con el producto vectorial, no es lo mismo.

Hay una forma más fácil de recordar las operaciones, que es calcular la siguiente expresión como si fuera un determinante, desarrollando por la primera fila:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2)\mathbf{i} + (a_3 b_1 - a_1 b_3)\mathbf{j} + (a_1 b_2 - a_2 b_1)\mathbf{k}.$$

Interpretación geométrica Si alguno de los vectores es $\mathbf{0}$, el producto vectorial da $\mathbf{0}$. Si ninguno es $\mathbf{0}$, el resultado de calcular el producto vectorial $\mathbf{a} \times \mathbf{b}$ se puede describir geométricamente como un vector que cumple las siguientes condiciones, que enunciaremos sin demostración:

1. Es perpendicular a los dos vectores $\mathbf{a}$ y $\mathbf{b}$.
2. Su módulo es $\|\mathbf{a}\| \, \|\mathbf{b}\| \, |\operatorname{sen} \theta|$, donde θ es el ángulo que forman $\mathbf{a}$ y $\mathbf{b}$.

9. Proyección ortogonal

Es importante en muchas aplicaciones el cálculo de la proyección ortogonal de un vector $\mathbf{a}$ sobre otro $\mathbf{b} \neq \mathbf{0}$. Geométricamente lo que hacemos es ubicar los dos vectores con un origen común y trazar un eje conteniendo al vector $\mathbf{b}$ (figura 6). El sentido positivo del eje es el dado por el propio vector $\mathbf{b}$. Trazamos una línea perpendicular a ese eje, que pase por el extremo de $\mathbf{a}$. Dicha línea corta al eje en un punto, que nos proporciona una "lectura" de un valor numérico (positivo, negativo o 0) sobre el eje. Ese valor numérico es la proyección ortogonal de $\mathbf{a}$ sobre $\mathbf{b}$, y se denota $\operatorname{proy}_{\mathbf{b}} \mathbf{a}$. Se tienen las siguientes fórmulas:

$$\operatorname{proy}_{\mathbf{b}} \mathbf{a} = \|\mathbf{a}\| \cos \theta$$

$$\operatorname{proy}_{\mathbf{b}} \mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|}$$

10. Producto mixto o producto triple

Hay una operación entre tres vectores que combina un producto escalar con un producto vectorial. Por su importancia, recibe una denominación propia, que es la de producto mixto o triple. El producto mixto de tres vectores en el espacio $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ (en ese orden) es $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$. Una forma fácil de hacer esta operación, sin necesidad de hacer las dos operaciones en forma sucesiva, es poner $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ como filas en un determinante de 3×3:

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

11. Bases

En el plano, una *base ordenada* (o simplemente *base*) está formada por dos vectores no nulos y no paralelos, dados en un orden concreto. La escribimos con notación de conjuntos, por ejemplo,

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

pero entendiendo que interesa el orden en que están dados: hay un primer y un segundo vector. En la figura 8, los vectores $\mathbf{a}$ y $\mathbf{b}$ forman una base en los casos que corresponden a $d > 0$ y $d < 0$.

En el espacio, una base ordenada estará formada por tres vectores no nulos y no coplanares. En la figura 7, los vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ forman una base en el primer y tercer caso, pero no en el segundo.

Tanto en el plano como en el espacio, se puede saber si ciertos vectores dados forman una base si:

1. La cantidad de vectores es correcta (2 en el plano, 3 en el espacio)

2. El determinante formado poniendo los vectores como filas de una matriz (de 2×2 o 3×3 respectivamente) es distinto de 0.

11.1. Combinaciones lineales

Una característica fundamental de las bases es que *todo vector se puede escribir como combinación lineal de los vectores de una base, y además en forma única.* Es decir, si $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, entonces todo vector $\mathbf{v}$ del espacio puede escribirse como

$$\mathbf{v} = \lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \lambda_3 \mathbf{b}_3$$

12. Ecuación paramétrica de la recta

En el espacio, sea P_0 un punto y $\mathbf{d}_L$ (o $\vec{d}_L$) un vector no nulo (en breve veremos por qué lo designamos así). Consideremos la recta que pasa por P_0 y tiene la misma dirección que $\mathbf{d}_L$ (figura 10).

Llamemos L a dicha recta[1] (observemos que L es un conjunto de puntos). Llamemos (x_0, y_0, z_0) a las coordenadas de P_0 y (u_1, u_2, u_3) a las componentes de $\mathbf{d}_L$. Entonces la recta L tiene la siguiente ecuación:

$$L : \begin{cases} x = x_0 + \lambda u_1 \\ y = y_0 + \lambda u_2 \\ z = z_0 + \lambda u_3 \end{cases} \qquad \lambda \in \mathbb{R}.$$

13. Ecuación del plano en el espacio

En el espacio, podemos describir un plano (es usual llamarlo π por ejemplo[2]) a partir de un punto del mismo (es decir, de $P_0 \in \pi$) y un vector no nulo $\mathbf{n}_\pi$ (o $\vec{n}_\pi$) , que llamamos *vector normal* a π. Este vector será perpendicular al plano π. En este contexto, la palabra "normal" significa "perpendicular". Si $\mathbf{n}_\pi = (a, b, c)$ entonces la ecuación del plano (también llamada ecuación implícita del plano) es

$$\pi : ax + by + cz + d = 0,$$

10 - Transformaciones lineales

Una transformacion lineal de V en W es una funci´on T : V → W (en este contexto llamaremos habitualmente T a dicha funci´on) que verifica las siguientes propiedades:

- T(u + v) = T(u) + T(v) para todo u ∈ V , v ∈ V ;
- T(λu) = λT(u) para todo u ∈ V , λ ∈ R

Una *combinación lineal* de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ es una expresión de la forma

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n$$

$$T(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n) = \lambda_1 T(\mathbf{v}_1) + \lambda_2 T(\mathbf{v}_2) + \cdots + \lambda_n T(\mathbf{v}_n)$$

A veces abreviamos TL (transformación lineal) y CL (combinación lineal).

11 - Autovalores y autovectores

Un autovector de T es un vector (no nulo!) tal que al aplicarle T se ve multiplicado por algun numero λ.

Un autovector de T es un vector v ∈ V , no nulo, tal que T(v) = λv para algun λ ∈ R.

Un autovalor de un operador lineal son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección.

1. Definición e interpretación geométrica

En estas notas, nos centraremos en estudiar el caso en que V es alguno de los espacios vectoriales $\mathbb{R}^2$, $\mathbb{R}^3$, (o $\mathbb{R}^n$ en general), E_2, E_3. La teoría se podría desarrollar sin cambios para cualquier espacio vectorial, aunque no lo haremos (ver el principio de las notas de transformaciones lineales).

Sea $T\colon V \to V$ una transformación lineal (observemos que en este contexto dominio y codominio deben ser el mismo conjunto). Un *autovector* de T es un vector $\mathbf{v} \in V$, no nulo, tal que

$$T(\mathbf{v}) = \lambda \mathbf{v}$$

para algún $\lambda \in \mathbb{R}$. Dicho número λ se llama un *autovalor* de T, asociado al autovector $\mathbf{v}$. En otras palabras, un autovector de T es un vector (no nulo!) tal que al aplicarle T se ve multiplicado por algún número λ.

Proposición 1. *Sea $T\colon V \to V$ una TL, y sea $\mathbf{v} \in V$ un autovector de T asociado al autovalor λ. Entonces, para cualquier $\alpha \in \mathbb{R}$, $\alpha \neq 0$, el vector $\alpha\mathbf{v}$ también será un autovector de T asociado al mismo λ.*

Demostración. Que $\mathbf{v}$ sea autovector asociado a λ significa $\mathbf{v} \neq \mathbf{0}$ y $T(\mathbf{v}) = \lambda\mathbf{v}$. Como además $\alpha \neq 0$, resulta $\alpha\mathbf{v} \neq \mathbf{0}$. Aplicando T obtenemos $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v}) = \alpha\lambda\mathbf{v} = \lambda(\alpha\mathbf{v})$. Entonces $\alpha\mathbf{v}$ es autovector de T asociado a λ. □

Proposición 2. *El sistema de ecuaciones* (3) (*o equivalentemente,* (2)) *es compatible indeterminado si y sólo si la matriz de coeficientes tiene determinante* 0*; es decir,*

$$\det([T]_{\mathcal{C}} - \lambda I_n) = 0.$$

(No demostraremos esta proposición.)

Debemos entonces encontrar los $\lambda \in \mathbb{R}$ tales que $\det([T]_{\mathcal{C}} - \lambda I_n) = 0$. Después, para cada uno de estos λ, resolvemos el sistema lineal para hallar los autovectores. Esto se detalla a continuación.

La forma de cálculo Primero se hallarán los autovalores y luego los autovectores asociados a cada uno.

1. Consideremos $[T]_{\mathcal{C}} - \lambda I_n$ (es sólo restar λ a lo largo de la diagonal) y calculemos su determinante. En el resultado aparecerá λ. Concretamente, quedará un polinomio en λ

de grado n, que recibe el nombre de *polinomio característico* de T. Por ejemplo, tomemos una transformación lineal $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ cuya matriz sea

$$[T]_{\mathcal{C}} = \begin{bmatrix} 2 & -1 \\ -3 & 0 \end{bmatrix}.$$

Entonces

$$\det([T]_{\mathcal{C}} - \lambda I_2) = \begin{vmatrix} 2-\lambda & -1 \\ -3 & -\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3$$

2. Queremos encontrar los valores de λ que hagan que este determinante sea 0. En otras palabras, estamos buscando las *raíces reales* de este polinomio[3]. Estas raíces serán los autovalores de T. Igual que con raíces de polinomios, podemos hablar de autovalores simples, dobles, etc. Tenemos a nuestra disposición todas las técnicas de cálculo de raíces.

   En nuestro ejemplo, las raíces de este polinomio cuadrático son $\lambda_1 = -1$ y $\lambda_2 = 3$.

3. Para cada autovalor hallado, buscamos soluciones no triviales (distintas al vector nulo) del sistema de ecuaciones lineales (3) (o equivalentemente, (2)). En ese sistema, debemos reemplazar λ por el autovalor que estemos analizando. En nuestro ejemplo tenemos lo siguiente.

   a) Para $\lambda_1 = -1$:

   $$\begin{bmatrix} 2-(-1) & -1 \\ -3 & -(-1) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   Esto significa

   $$\begin{cases} 3x - y = 0 \\ -3x + y = 0 \end{cases}$$

   El conjunto de soluciones de este sistema es $S = \{(x, 3x) \in \mathbb{R}^2 \colon x \in \mathbb{R}\}$. Cualquier elemento no nulo de este conjunto será un autovector de T asociado al autovalor -1. Si quisiéramos mostrar algún autovector concreto, podemos obtener una solución particular asignando a x algún valor (que no produzca la solución nula), por ejemplo, $x = 1$. Así obtenemos el autovector $\mathbf{v}_1 = (1, 3)$.

   En vez de S, es habitual en este contexto denotar al conjunto de soluciones como V_{-1}, que se llama el *autoespacio* asociado al autovalor -1. Observemos que el autoespacio está formado por todos los autovectores asociados a -1, incorporando además al vector $(0, 0)$ (que, reiteramos, no es un autovector).

   b) Para $\lambda_2 = 3$:

   $$\begin{bmatrix} 2-3 & -1 \\ -3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   Esto significa

   $$\begin{cases} -x - y = 0 \\ -3x - 3y = 0 \end{cases}$$

   $$V_3 = \{(x, -x) \in \mathbb{R}^2 \colon x \in \mathbb{R}\}$$

   Un autovector concreto es por ejemplo $\mathbf{v}_2 = (1, -1)$.

Observemos que siempre el sistema lineal a resolver queda compatible *indeterminado*. Precisamente, los valores de λ se hallaron buscando que así sea. Si al resolver obtuviéramos que la única solución es que todas las incógnitas valgan 0, eso indicaría que hay algún error de cuentas.

Proposición 3. *Sea $T\colon V \to V$ una TL, y sean $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ autovectores de T asociados al mismo autovalor λ. Entonces toda combinación lineal $\mathbf{u}$ de $\mathbf{v}_1, \dots, \mathbf{v}_k$ pertenece al autoespacio V_λ. Si $\mathbf{u} \neq \mathbf{0}$, entonces $\mathbf{u}$ es un autovector de T asociado a λ.*

Demostración. Por hipótesis, $T(\mathbf{v}_1) = \lambda\mathbf{v}_1, \ \dots, \ T(\mathbf{v}_k) = \lambda\mathbf{v}_k$. Si consideramos una combinación lineal $\mathbf{u} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k$, entonces, por ser T una transformación lineal, tenemos

$$\begin{aligned} T(\mathbf{u}) &= \alpha_1 T(\mathbf{v}_1) + \cdots + \alpha_k T(\mathbf{v}_k) \\ &= \alpha_1\lambda\mathbf{v}_1 + \cdots + \alpha_k\lambda\mathbf{v}_k \\ &= \lambda(\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k) \\ &= \lambda\mathbf{u}, \end{aligned}$$

y así $\mathbf{u} \in V_\lambda$. Si además $\mathbf{u} \neq \mathbf{0}$, esto completa las condiciones para que $\mathbf{u}$ sea un autovector con autovalor λ. □

Comentario. La proposición anterior dice que los autoespacios son cerrados por combinaciones lineales (al hacer cualquier CL de elementos de V_λ el resultado queda en V_λ). Un conjunto con esta propiedad recibe el nombre de *subespacio vectorial* (del espacio vectorial original V). Así, todo autoespacio de $T\colon V \to V$ es un subespacio vectorial de V.

Bases

Definición 1. Una *base* de un espacio vectorial V es un subconjunto $B \subset V$ tal que todo elemento de V puede escribirse en forma única como combinación lineal de elementos de B.

Destacamos que estamos pidiendo dos cosas: que cualquiera sea el vector de V que consideremos, este se pueda escribir como CL de los vectores de B, y que dicha CL sea única, es decir, que haya una sola elección posible de coeficientes.

Una *base ortonormal* (que abreviamos BON) es una base formada por vectores unitarios (módulo 1) y ortogonales dos a dos. Es decir, los vectores están normalizados y siempre que tomemos dos vectores distintos en dicha base, éstos serán ortogonales.

Componentes y cambio de base

Dada una base ordenada $B \subset V$, sabemos que todo vector $\mathbf{v} \in V$ se escribe como la CL mencionada en la definición de base. Es decir, si $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ es una base ordenada de V, y

$$\mathbf{v} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n$$

entonces los coeficientes $\lambda_1, \lambda_2, \dots, \lambda_n$ se llaman las componentes de $\mathbf{v}$ en la base B. Con las componentes formamos una matriz columna, con la siguiente notación:

$$[\mathbf{v}]_B = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$$

Decimos que esta matriz columna es el vector $\mathbf{v}$ escrito en la base B. Notemos que a ningún vector le pueden corresponder dos matrices columna diferentes, precisamente por la unicidad mencionada en la definición de base.

Recalcamos lo siguiente: Las componentes de un vector dependen de la base elegida. Si cambia la base, cambian las componentes (salvo para algunos vectores, pero hablando generalmente podemos esperar que cambien las componentes). En lo que sigue, estudiaremos cómo es ese cambio.

Fórmulas de cambio de base Mirando detenidamente el último ejemplo, veremos que la matriz de coeficientes del sistema de ecuaciones lineales guarda una relación con la base B. Efectivamente, dicha matriz se obtiene si ponemos los vectores de B como columnas. Llamamos a esta matriz $[B]_{\mathcal{C}}$ (leemos "B en C"):

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

Esta se llama la matriz de cambio de base de B en $\mathcal{C}$. El sistema de ecuaciones lineales se puede escribir en forma matricial como

$$[B]_{\mathcal{C}}[\mathbf{v}]_B = [\mathbf{v}]_{\mathcal{C}} \tag{1}$$

Al ser $[B]_{\mathcal{C}}$ una matriz inversible, tenemos la siguiente forma de obtener la solución (que será única):

$$[\mathbf{v}]_B = [B]_{\mathcal{C}}^{-1}[\mathbf{v}]_{\mathcal{C}} \tag{2}$$

Si aplicáramos (2) al último ejemplo, daría el mismo resultado para $[\mathbf{v}]_B$ que hallamos "a mano", es decir, planteando el sistema de ecuaciones lineales y resolviéndolo.

Las fórmulas (1) y (2) relacionan las componentes de un mismo vector $\mathbf{v}$ en las dos bases B y $\mathcal{C}$. Podemos interpretar estas fórmulas pensando que $[\mathbf{v}]_{\mathcal{C}}$ y $[\mathbf{v}]_B$ son formas de expresar un mismo vector $\mathbf{v}$ en dos "idiomas" distintos (las bases) y las matrices "traducen" de un idioma al otro. Por ejemplo, en (1) la matriz $[B]_{\mathcal{C}}$ traduce de la base B a la canónica, y en (2) la matriz $[B]_{\mathcal{C}}^{-1}$ traduce en sentido contrario, de la base canónica a B. Denotamos entonces esta última matriz como

$$[\mathcal{C}]_B = [B]_{\mathcal{C}}^{-1}$$

ya que es la matriz de cambio de base de $\mathcal{C}$ en B.

Definición general Si se tienen dos bases ordenadas B y B' de un mismo espacio vectorial V de dimensión n, definimos la matriz de cambio de base de B en B', denotada $[B]_{B'}$, escribiendo *como columnas* sucesivas a los vectores de B escritos *en la base* B'. Es decir, si $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, entonces

$$[B]_{B'} = \begin{bmatrix} [\mathbf{b}_1]_{B'} & \dots & [\mathbf{b}_n]_{B'} \end{bmatrix}$$

Aquí, $[\mathbf{b}_1]_{B'}$, etc. son matrices columna de $n \times 1$, por lo que la notación de arriba representa una matriz de $n \times n$.

Esta matriz verifica la siguiente propiedad fundamental:

$$[\mathbf{v}]_{B'} = [B]_{B'}[\mathbf{v}]_B$$

Cambio de base para transformaciones lineales

Cuando vimos transformaciones lineales $T\colon \mathbb{R}^n \to \mathbb{R}^m$ hablamos de la matriz asociada $[T]_{\mathcal{C}_n\mathcal{C}_m}$, donde $\mathcal{C}_n$ y $\mathcal{C}_m$ eran las respectivas bases canónicas. La idea es que esta matriz representa a la TL de la siguiente forma:

$$\begin{bmatrix} \text{salida} \end{bmatrix} = \begin{bmatrix} \text{matriz} \\ \text{de } T \end{bmatrix} \begin{bmatrix} \text{entrada} \end{bmatrix} \tag{3}$$

Diagonalización de matrices

Igual que en el tema anterior (autovalores y autovectores), podemos aplicar todas estas ideas al caso de una matriz M cuadrada, sin pensarla explícitamente como la matriz de una transformación lineal.

Sea M es una matriz cuadrada de orden n. Diagonalizar M consiste en:

- encontrar (si es que existe) una base B de $\mathbb{R}^n$ formada por autovectores de M,
- formar la matriz de cambio de base $[B]_{\mathcal{C}}$ como siempre,
- escribir una matriz diagonal D con los correspondientes autovalores (lo que antes llamábamos $[T]_B$).

Si se puede encontrar una base de autovectores, M se dice diagonalizable. En ese caso, se verificará que

$$M = [B]_{\mathcal{C}} D [B]_{\mathcal{C}}^{-1}.$$

También ocurrirá que M es matriz simétrica si y sólo si es diagonalizable en una base ortonormal.

En resumen, simplemente hay que hacer con la matriz M lo que hacíamos con $[T]_{\mathcal{C}}$.