Universidad Monteávila

25 años hacia la cima

Álgebra Lineal Básica

Ramón Bruzual

Caracas, Venezuela
Septiembre 2025

Estas notas han sido elaboradas para facilitar el aprendizaje
del álgebra lineal a los estudiantes de la Universidad Monteávila

Correo-E: rbruzual@profesor.uma.edu.ve
https://sites.google.com/profesor.uma.edu.ve/prof-rbruzual

Índice general

Introducción

El álgebra lineal es una herramienta matemática fundamental en la ingeniería moderna, con aplicaciones que van desde la resolución de sistemas de ecuaciones hasta el aprendizaje automático y la simulación computacional, razón por la cual su estudio se ha vuelto prácticamente obligatorio en el primer semestre de las carreras modernas de ingeniería.

El álgebra lineal proporciona un marco teórico y práctico para abordar problemas en ingeniería donde las relaciones lineales son predominantes. Su lenguaje (vectores, matrices, transformaciones) es esencial para describir fenómenos físicos, optimizar procesos y diseñar algoritmos eficientes.

Entre las aplicaciones del álgebra lineal destacan, entre muchas otras:

(1) Modelación de problemas en ingeniería mediante sistemas de la forma

$$A\mathbf{x} = \mathbf{b}$$

donde A es una matriz de coeficientes, $\mathbf{x}$ el vector de incógnitas y $\mathbf{b}$ el vector de términos independientes.

(2) Métodos numéricos como la *factorización LU* o *descomposición QR* son cruciales en:
   (a) Análisis de circuitos eléctricos (leyes de Kirchhoff).
   (b) Mecánica de estructuras (equilibrio de fuerzas).

(3) Las matrices representan transformaciones geométricas de la forma

$$\mathbf{y} = T(\mathbf{x}) = A\mathbf{x}.$$

lo que se aplica a
   (a) Gráficos 3D (rotaciones, escalados con OpenGL).
   (b) Procesamiento de imágenes (convoluciones matriciales).

(4) La ecuación característica $A\mathbf{v} = \lambda\mathbf{v}$ aparece en:
   (a) Dinámica estructural (modos de vibración).
   (b) Algoritmos de compresión de datos (PCA).
   (c) Mecánica cuántica (estados estacionarios).

(5) Problemas de optimización tales como

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|_2$$

son la base de:

(a) Aprendizaje automático (regresión lineal).
(b) Logística (optimización de rutas).
(c) Compresión de archivos informáticos.

En conclusión, el álgebra lineal es parte fundamental del lenguaje universal de la ingeniería moderna. Su dominio permite no solo resolver problemas tradicionales, sino también innovar en áreas emergentes como la computación cuántica o la inteligencia artificial. Sin duda, invertir en su comprensión es invertir en las herramientas del futuro.

Estas notas han sido elaboradas especialmente para el curso "Algebra lineal", común a las carreras de Ingeniería que se dictan en la Universidad Monteávila, con la intención de facilitar la comprensión del tema a los bachilleres que, por primera vez se enfrentan a un curso de Matemáticas a nivel universitario.

El lenguaje en que está escrito es sencillo y accesible, se han colocado abundantes ejemplos para facilitar la comprensión del tema y es muy recomendable que el estudiante no se limite a esta guía y revise y aproveche el amplio panorama que brinda la bibliografía señalada al final.

Ramón Bruzual

Capítulo 1

Los números reales

El primer tema a repasar antes de comenzar el estudio del Álgebra Lineal es el que corresponde con el conjunto de los números reales, ya que para tener éxito en el estudio del álgebra es fundamental conocer y manejar correctamente sus propiedades.

El conjunto de los *números reales*, que usualmente se denota con el símbolo $\mathbb{R}$, está formado por la unión del conjunto de los números racionales, con el conjunto de los números irracionales.

El conjunto de los *números racionales*, que usualmente se denota por $\mathbb{Q}$, está formado por los números reales que se pueden expresar como el cociente de dos números enteros, es decir, tienen la forma $\frac{a}{b}$, donde a y b son números enteros y $b \neq 0$.

Por ejemplo $0, -1, \frac{1}{5}, 5, \frac{157}{5}$ son números racionales.

Los *números irracionales* son aquellos números reales que no se pueden expresarse como el cociente de dos números enteros.

Por ejemplo, $\sqrt{2}, \sqrt[3]{8}, \pi$ son números irracionales.

Todo número real tiene un desarrollo decimal, el desarrollo decimal de los números racionales es finito o se repite, el desarrollo decimal de los números irracionales es infinito y no se repite.

Por ejemplo,

$$\frac{1}{8} = 0,125 \quad \frac{1}{3} = 0,333\ldots \quad \frac{9}{4} = 1,285714285714285714\ldots$$

$$\sqrt{2} = 1,41421\ldots \quad \pi = 3,141592\ldots$$

El conjunto de los *números naturales*, que usualmente se denota por $\mathbb{N}$, corresponde con los números que utilizamos para contar. El 0 será incluido en el conjuntos de los números naturales. El conjunto de los *números enteros*, que usualmente se denota por $\mathbb{Z}$, incluye al conjunto de los números naturales y sus opuestos con respecto a la suma.

En notación de conjuntos

$$\mathbb{N} = \{0, 1, 2, \cdots\}, \quad \mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}.$$

Dentro del conjunto de los números naturales destacan el conjunto de los *números pares*, que son los de la for $2n$ para algún número natural n, los *números impares* que son los de la forma $2n+1$ para algún número natural n. También están los *números primos* que son aquellos que son mayores o iguales que 2 y solamente son divisibles por si mismos y por la unidad. Es importante recordar que todo número natural se puede expresar como producto de números primos y que esta descomposición es única.

El siguiente diagrama resume la relación entre los principales subconjuntos de los números reales.

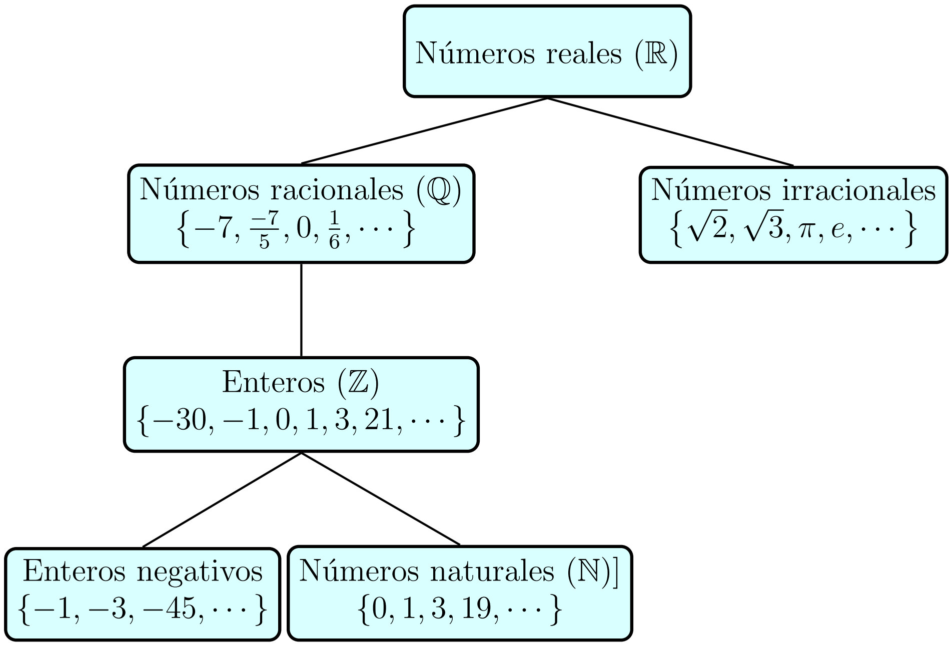


Figura 1.1. Distintas clases de números reales

1.1. Operaciones básicas con los números reales

1.1.1. La suma y el producto. Existen dos operaciones fundamentales: las suma $(+)$ y el producto $(\cdot)$ de números reales, que satisfacen las siguientes propiedades (a, b, c denotan números reales).

Propiedad conmutativa: $a + b = b + a$, $a \cdot b = b \cdot a$ (es usual escribir $a \times b$ o simplemente ab en vez de $a \cdot b$).

Propiedad asociativa: $a + (b + c) = (a + b) + c$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$; por esta propiedad las expresiones $a + b + c$ y $a \cdot b \cdot c$ no son ambiguas.

Propiedad distributiva: $a \cdot (b + c) = a \cdot b + a \cdot c$.

Elemento neutro: Existe un único número real, 0, tal que $a + 0 = 0 + a = a$ y existe un único número real, 1, tal que $a \cdot 1 = 1 \cdot a = a$.

Existencia de inversos: Para cualquier número real a, existe otro número real,$-a$ tal que $a + (-a) = (-a) + a = 0$ y para cualquier número real a, diferente de 0, existe otro número real, a^{-1} tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$. Al número $-a$ se le llama el *inverso aditivo* de a y al número a^{-1} se le llama el *inverso multiplicativo* de a.

Propiedades del factor 0: Para cualquier número real a, $a \cdot 0 = 0$, si $a \cdot b = 0$, entonces $a = 0$ ó $b = 0$.

Es importante destacar que, como al multiplicar un número por 0 el resultado es 0, el número 0 no puede tener un inverso multiplicativo.

Ley de los signos:

- $-(-a) = a$
- $(-a) \cdot (-b) = a \cdot b$
- $-a \cdot b = (-a) \cdot b = a \cdot (-b)$
- $(-1) \cdot a = -a$

De esto último se deduce la conocida *regla de los signos* :

$$(+) \times (+) = +, \quad (+) \times (-) = (-) \times (+) = - \quad (-) \times (-) = +$$

1.1.2. Resta y división.

La resta: Se defina $a - b$ como $a + (-b)$, es decir $a - b = a + (-b)$.

División: Se define $\frac{a}{b}$ ó $(a \div b\ ,\ a/b)$ como $a \cdot (b^{-1})$.

Observación 1.1. Notemos que $\frac{1}{b} = 1 \cdot b^{-1} = b^{-1}$.
Como 0 no tiene inverso multiplicativo, no está definido dividir entre 0.

Se cumplen las siguiente propiedades (b y d deben ser diferentes de 0).

- $-\dfrac{a}{b} = \dfrac{-a}{b} = \dfrac{a}{-b}$
- $\dfrac{-a}{-b} = \dfrac{a}{b}$
- $\dfrac{a}{b} = \dfrac{c}{d}$ si y solo si $a \cdot d = b \cdot c$
- Si $k \neq 0$ entonces $\dfrac{a}{b} = \dfrac{k \cdot c}{k \cdot d}$
- $\left(\dfrac{a}{b}\right) \cdot \left(\dfrac{c}{d}\right) = \dfrac{a \cdot c}{b \cdot d}$
- $\dfrac{a}{b} + \dfrac{c}{b} = \dfrac{a + c}{b}$
- $\dfrac{a}{b} + \dfrac{c}{d} = \dfrac{a \cdot d + b \cdot c}{b \cdot d}$
- $\left(\dfrac{a}{b}\right) \div \left(\dfrac{c}{d}\right) = \dfrac{a \cdot d}{b \cdot c}$

Ejercicios.

(1) Simplificar las siguientes expresiones (a, b y c representan número reales).

(a) $\dfrac{(5a)(b-c)(a-a)(12b)}{5\sqrt{2} - 3\sqrt{6}}$

(b) $\dfrac{2}{\sqrt{2}}$

(c) $\dfrac{1}{2} + \dfrac{2}{5} + \dfrac{1}{3}$

(d) $\dfrac{1}{2} - \dfrac{1}{3} + \dfrac{2}{6}$

(2) ¿Cuáles de las siguientes afirmaciones son verdaderas y cuáles son falsas?

(a) La suma de dos números racionales es un número racional.
(b) La suma de un número racional y un irracional es un número irracional.
(c) La suma de dos números irracionales es un número irracional.
(d) La suma de dos números irracionales puede ser un número racional.
(e) La suma de dos números irracionales puede ser un número irracional.
(f) El producto de un número racional no nulo y un número irracional es un número irracional.
(g) El producto de dos números irracionales es un número irracional.
(h) El producto de dos números irracionales puede ser un número entero.
(i) El producto de un número entero y un número irracional es un número irracional.

(3) Encontrar el error en la siguiente prueba de que $1 = 2$.

Sean a y b números reales tales que $a = b$.

Como $a = b$ se tiene que

$$a^2 = a \cdot a = a \cdot b,$$

restando b^2 a ambos miembros de la desigualdad,

$$a^2 - b^2 = a \cdot b - b^2,$$

por lo tanto,

$$(a - b)(a + b) = (a - b)b,$$

dividiendo entre $(a - b)$,

$$a + b = b.$$

Si tomamos $a = b = 1$, obtenemos

$$2 = 1 + 1 = a + b = b = 1.$$

1.2. Propiedades de orden y la recta real

Los números reales pueden ser representados como puntos en una recta, de manera que a cada número real le corresponde un único punto en la recta y, recíprocamente, cada punto de la recta corresponde con un único número real.

Por ejemplo, en la siguiente figura se observa la ubicación de los números

$$-4, -3, -2, -\frac{3}{2}, -\sqrt{2}, -1, 0, 1, 2, \frac{5}{2}, 3, \pi, 4$$

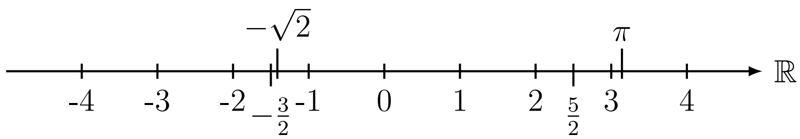


FIGURA 1.2. La recta real

El conjunto de los números reales está ordenado, los números *positivos* son los que quedan a la derecha de 0 y los *negativos* a la izquierda de 0.

Para esta relación de orden se usan los símbolos $<, \leq, >, \geq$.

Cuando a es positivo, se escribe $a > 0$ ó $0 < a$, si a es negativo se escribe $a < 0$ ó $0 > a$, se dice que $a < b$ si $b - a > 0$.

Se dice que $a \leq b$ si $a < b$ ó $a = b$, de manera análoga se define $a \geq b$.

La relación de orden en los números reales satisface las siguientes propiedades.

- Para cualquier número real a, se cumple una y solo una de las siguientes afirmaciones
$$a < 0, \quad a = 0 \quad a > 0$$
- Si a y b son números reales diferentes, se cumple una y solo una de las siguientes afirmaciones
$$a < b \qquad a > b$$
es decir, todos los números reales son comparables.
- Si $a \neq 0$, entonces $a^2 > 0$.
- (transitividad) Si $a < b$ y $b < c$, entonces $a < c$
- Si $a < b$, entonces $a + c < b + c$
- Si $a < b$, entonces $\begin{cases} a \cdot c < b \cdot c & \text{si } c > 0 \\ a \cdot c > b \cdot c & \text{si } c < 0 \end{cases}$

  En palabras: Al multiplicar ambos lados de una desigualdad por un número positivo se mantiene el sentido de la desigualdad, al multiplicarla por un número negativo cambia el sentido de la desigualdad.

  En particular, si $a < b$, entonces $-a > -b$.

1.3. Valor absoluto

El valor absoluto de un número real x ($|x|$) se define de la siguiente manera

$$|x| = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}$$

Se tiene que, para todo número real x, $|x| \geq 0$ y $|x| = 0$ si y solo si $x = 0$.

El valor absoluto de un número corresponde con su distancia al origen (el cero) en la recta real, si $a > 0$ el conjunto

$$\{x \in \mathbb{R} \, : \, |x| < a\} = \{x \in \mathbb{R} \, : \, -a < x < a\}$$

tiene la siguiente representación gráfica en la recta real

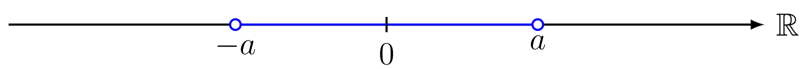


FIGURA 1.3. El intervalo $|x| < a$

Las pequeñas circunferencias sin rellenar en los extremos indican que los puntos a y $-a$ no pertenecen al conjunto, de haber considerado el conjunto $|x| \leq a$ se habrían rellenado estas circunferencias.

Si $a > 0$, se tiene que $|x| < a$ si y solo si $-a < x < a$, igualmente $|x| \leq a$ si y solo si $-a \leq x \leq a$.

También se cumple que $|x| > a$ si y solo si $x > a$ ó $x < -a$ y que $|x| \geq a$ si y solo si $x \geq a$ ó $x \leq -a$.

Si a y b son números reales, $|b - a|$ representa la distancia que hay entre el punto a y el punto b.

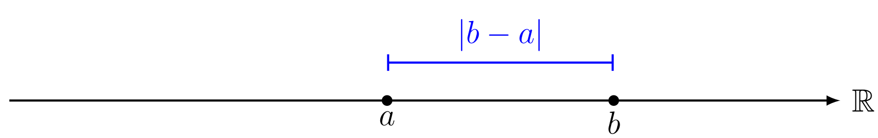


FIGURA 1.4. Distancia entre dos puntos de la recta

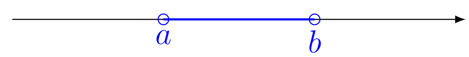
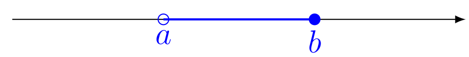
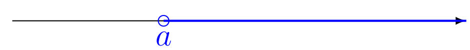
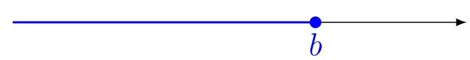
1.4. Intervalos

Un *intervalo* es un subconjunto de los números reales que incluye todos los números entre dos valores dados, pudiendo incluir los extremos o no. Geométricamente corresponden con un segmento, una semirrecta o toda la recta real.

Los intervalos pueden ser de longitud finita, llamados *intervalos finitos* , como por ejemplo el conjunto $\{x \in \mathbb{R} \, : \, 1 < x \leq 2\}$ o de longitud infinita, llamados *intervalos infinitos* , como por ejemplo el conjunto $\{x \in \mathbb{R} \, : \, x > 2\}$. Los intervalos finitos corresponden con segmentos de recta y los infinitos con semirrectas o toda la recta.

A los intervalos se les dice abiertos cuando no incluyen sus extremos , cerrados cuando incluyen sus extremos y semiabiertos o semicerrados cuando incluyen solamente uno de sus extremos.

En la siguiente tabla se muestran ejemplos de diferentes tipos de intervalo, con su representación geométrica, las pequeñas circunferencias sin rellenar indican que el punto correspondiente no pertenece al conjunto.

Notación	**Conjunto correspondiente**	**Nombre**	**Gráfico**
(a, b)	$\{x \in \mathbb{R} \,:\, a < x < b\}$	Intervalo abierto (finito)	
$[a, b]$	$\{x \in \mathbb{R} \,:\, a \leq x \leq b\}$	Intervalo cerrado (finito)	
$(a, b]$	$\{x \in \mathbb{R} \,:\, a < x \leq b\}$	Intervalo semicerrado o semiabierto (finito)	
$(a, +\infty)$	$\{x \in \mathbb{R} \,:\, x > a\}$	Intervalo abierto (infinito)	
$-(\infty, b]$	$\{x \in \mathbb{R} \,:\, x \leq b\}$	Intervalo semiabierto (infinito)	
$-(\infty, +\infty)$	$\mathbb{R}$	Intervalo abierto (infinito)	

Ejercicios.

(1) En los siguientes ejercicios, expresar la afirmación en forma de desigualdad.

(a) $a + 2$ es positivo

(b) $4y$ es negativo

(c) $a + b$ es no negativo

(d) a es menor que -3

(e) $2b + 4$ es mayor que o igual que 100

(f) $c - 1$ es menor que o igual que 5.

(2) En los siguientes ejercicios escriba la desigualdad usando notación de intervalos, y a continuación grafique el intervalo.

(a) $x < 0$

(b) $0 < x < 5$

(c) $x \geq 5$

(d) $-1 \leq x$
(e) $8 < x \leq 10$
(f) $-5 < x \leq -3$
(g) $-2 \leq x \leq 4$
(h) $x > -7$

(3) En los siguientes ejercicios escriba el intervalo en forma de desigualdad.
(a) $[-7, 9]$
(b) $[1, 15)$
(c) $(-\infty, 2)$
(d) $[-5, \infty)$

(4) En los siguientes ejercicios, resuelva la desigualdad lineal, escriba el conjunto solución usando la notación de intervalos y grafique el conjunto solución.
(a) $x + 3 > -2$
(b) $3x - 9 < 6$
(c) $\frac{3}{2}x + 4 \leq 10$
(d) $5 - \frac{5}{4}x \geq -4$
(e) $\frac{3}{2} - x > x$
(f) $-(1 - x) \geq 2x - 1$
(g) $2 + x \geq 3(x - 1)$
(h) $-7x + 3 \leq 4 - x$
(i) $-\frac{29}{5} < \frac{3}{2}x < 4$
(j) $-3 \leq -x < 2$
(k) $-7 < x - 2 < 1$

1.5. Conjuntos acotados, supremo e ínfimo

Sea A un subconjunto del conjunto de los números reales $\mathbb{R}$.

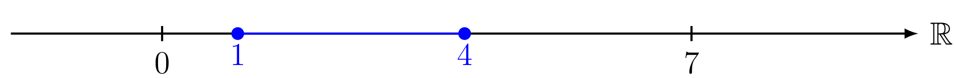
Se dice que A está *acotado superiormente* si existe $K \in \mathbb{R}$ tal que $x \leq K$ para todo $x \in A$, en este caso se dice que K es una cota superior de A.

Se dice que A está *acotado inferiormente* si existe $k \in \mathbb{R}$ tal que $x \geq k$ para todo $x \in A$, en este caso se dice que k es una cota inferior de A.

Se dice que A está *acotado* si A está acotado superior e inferiormente.

EJEMPLO 1.2.

(1) El conjunto $[1, 4]$ está acotado, los números 4, 7, 99 son ejemplos de cota superior y los números 1, 0, -99 son ejemplos de cota inferior.

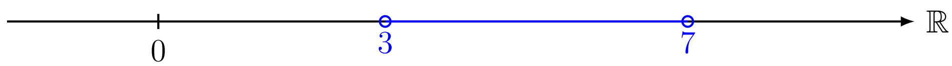


(2) El conjunto de los números naturales $\mathbb{N} = \{0, 1, 2, \dots\}$ está acotado inferiormente y no está acotado superiormente. Cualquier número negativo ó 0 es una cota inferior.

(3) El conjunto de los números enteros $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ no está acotado ni superior ni inferiormente.

(4) El conjunto $(3, 7]$ está acotado, los números 7, 11, 89 son ejemplos de cota superior y los números 1, 3, -99 son ejemplos de cota inferior.

Al observar un conjunto como el intervalo $(3, 7)$



queda claro que existen una cota superior y una cota inferior que destacan, que son 7 y 3 respectivamente.

Este tipo de cotas destacadas se definen de la siguiente manera.

Sea A un subconjunto de $\mathbb{R}$ que está acotado superiormente, se dice que M es el *supremo* de A, denotado por $\sup A$, si:

(i) M es cota superior de A,

(ii) Si K es una cota superior de A entonces $M \leq K$.

En palabras, el supremo de A es la menor de las cotas superiores de A.

Si A no está acotado superiormente, se suele escribir $\sup A = +\infty$.

Por ejemplo $\sup(3, 7) = \sup[3, 7] = 7$, lo que muestra que el supremo de un conjunto puede pertenecer o no al conjunto.

De manera análoga, se define el ínfimo.

Sea A un subconjunto de $\mathbb{R}$ que está acotado inferiormente, se dice que m es el *ínfimo* de A, denotado por $\inf A$, si:

(i) m es cota inferior de A,

(ii) Si k es una cota inferior de A entonces $k \leq m$.

En palabras, el ínfimo de A es la mayor de las cotas inferiores de A.

Si A no está acotado inferiormente, se suele escribir $\inf A = -\infty$.

Volviendo al ejemplo anterior, $\inf(3, 7) = \inf[3, 7] = 3$, lo que muestra que, al igual que el supremo, el ínfimo de un conjunto puede pertenecer o no al conjunto.

1.6. Factorización y productos notables

Al trabajar con expresiones algebraicas, tal como se comenzará a hacer pronto, es muy importante reconocer las siguientes expresiones básicas, conocidas como *productos notables*.

Sean a y b números reales, a continuación se muestra una lista de productos notables básicos.

- $(a + b)^2 = a^2 + 2ab + b^2$

- $(a-b)^2 = a^2 - 2ab + b^2$
- $a^2 - b^2 = (a+b)(a-b)$
- $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$
- $(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$
- $a^3 + b^3 = (a+b)(a^2 - ab + b^2)$
- $a^3 - b^3 = (a-b)(a^2 + ab + b^2)$

Observación 1.3. Es muy importante evitar la siguiente confusión común:

$$\cancel{(a+b)^2 = a^2 + b^2}$$

1.6.1. Racionalización. Como ejemplo de aplicación de los productos notables veamos el siguiente problema:

Eliminar las raíces del denominador de la expresión

$$\frac{x}{\sqrt{x+4}-2}.$$

Sabiendo que $a^2 - b^2 = (a+b)(a-b)$, multiplicamos numerador y denominador por $\sqrt{x+4}+2$ y obtenemos

$$\begin{aligned}\frac{x}{\sqrt{x+4}-2} &= \frac{x(\sqrt{x+4}+2)}{(\sqrt{x+4}-2)(\sqrt{x+4}+2)}\\ &= \frac{x(\sqrt{x+4}+2)}{x+4-4}\\ &= \frac{x(\sqrt{x+4}+2)}{x}\\ &= \sqrt{x+4}+2.\end{aligned}$$

1.7. Desigualdades e inecuaciones

1.7.1. Desigualdades lineales. Encontrar el conjunto solución de una desigualdad de la forma

$$ax + b < c,$$

o similar, donde a, b yc son constantes y x es la variable es bastante sencillo. Basta aplicar las propiedades de las desigualdades ya estudiadas en la Sección 1.2.

Lo ilustremos con algunos ejemplos.

Ejemplo 1.4.

(1) Encontrar el conjunto de los $x \in \mathbb{R}$ tales que $2x + 3 > 7$.

Restando 3 a ambos miembros: $2x > 4$

Dividiendo ambos miembros entre 2: $x > 2$.

El conjunto solución es

$$\{x \in \mathbb{R} \, : \, x > 2\} = (2, +\infty).$$

(2) Resolver $-3x + 5 \leq 3$.

Restando 5 a ambos miembros: $-3x \leq -2$

Dividiendo ambos miembros entre -3: $x \geq \frac{2}{3}$.

(Notar que al dividir entre un número negativo debemos cambiar el sentido de la desigualdad)

El conjunto solución es

$$\left\{x \in \mathbb{R} \, : \, x \geq \frac{2}{3}\right\} = \left[\frac{2}{3}, +\infty\right).$$

(3) Resolver $4x + 2 \geq x - 7$.

Restando 2 a ambos miembros: $4x \geq x - 9$

Restando x a ambos miembros: $3x \geq -9$

Dividiendo ambos miembros entre 3: $x \geq -3$.

El conjunto solución es

$$\{x \in \mathbb{R} \, : \, x \geq -3\} = [-3, +\infty).$$

1.7.2. Desigualdades no lineales. Para resolver una desigualdad del tipo

$$(x-1)(x+3)(x-2) \geq 0$$

conviene elaborar una tabla de signos de los factores, de la siguiente manera.

Conjunto	$(-\infty, -3)$	-3	$(-3, 1)$	1	$(1, 2)$	2	$(2, +\infty)$
$(x-1)$	$-$	-	$-$	0	$+$	$+$	$+$
$(x+3)$	$-$	0	$+$	$+$	$+$	$+$	$+$
$(x-2)$	$-$	-	$-$	-	$-$	0	$+$
$(x-1)(x+3)(x-2)$	$-$	0	$+$	0	$-$	0	$+$

De la tabla de signos se deduce que el conjunto solución es $[-3, 1] \cup [2, +\infty)$.

De la misma tabla se puede deducir que el conjunto solución de la desigualdad $(x-1)(x+3)(x-2) > 0$ es $(-3,1) \cup (2,+\infty)$, y también que el conjunto solución de la desigualdad $(x-1)(x+3)(x-2) < 0$ es $(-\infty,-3) \cup (1,2)$.

Supongamos que toca resolver la siguiente desigualdad

$$\frac{(x-1)(x+3)}{x-2} \geq 0.$$

Tenemos exactamente los mismos factores que antes,sin embargo el factor $x-2$ está en el denominador, por lo que la expresión no está definida para $x=2$, procedemos de la misma manera, pero marcando que en $x=2$ la expresión no está definida. La tabla que corresponde ahora es

Conjunto	$(-\infty,-3)$	-3	$(-3,1)$	1	$(1,2)$	2	$(2,+\infty)$
$(x-1)$	$-$	-	$-$	0	$+$	$+$	$+$
$(x+3)$	$-$	0	$+$	$+$	$+$	$+$	$+$
$(x-2)$	$-$	-	$-$	-	$-$	0	$+$
$\frac{(x-1)(x+3)}{x-2}$	$-$	0	$+$	0	$-$	ND	$+$

(ND se usado como abreviación de “no definida”)

El conjunto solución es $[-3,1] \cup (2,+\infty)$.

Ejercicios.

(1) En los siguientes ejercicios, resuelva la desigualdad no lineal, escriba el conjunto solución usando la notación de intervalos y grafique el conjunto solución.

(a) $x^2-9<0$
(b) $x^2 \geq 16$
(c) $x(x-5) \geq 0$
(d) $4x^2+7x<0$
(e) $x^2-8x+12<0$
(f) $(3x+2)(x-1) \leq 0$
(g) $9x \geq 2x^2-18$
(h) $4x^2 > 9x+9$
(i) $(x+1)(x-2)(x-4)<0$
(j) $(1-x)(x+\frac{1}{2})(x-3) \leq 0$
(k) $(x^2-1)(x^2-4) \leq 0$
(l) $(x-1)^2(x+3)(x-5) \geq 0$
(m) $\frac{5}{x+8}<0$
(n) $\frac{10}{x^2+2}>0$

1.8. Orden de las Operaciones

Existe una regla, llamada *PEMDAS*, que sirve de ayuda para no equivocarnos en el orden en que se deben realizar las operaciones de suma, resta, multiplicación, división. PEMDAS es un acrónimo que se detalla a continuación y el orden de las operaciones corresponde con el orden de las letras en el acrónimo:

- **P** – Paréntesis: Primero, resolver todo lo que esté dentro de los paréntesis u otros signos de agrupación
- **E** – Exponentes: Luego, resolver los exponentes o potencias.
- **MD** – Multiplicación y División: Resolver de **izquierda a derecha**. Estas operaciones tienen la misma prioridad.
- **AS** – Suma (*Adición*) y Resta (*Sustracción*): Resolver de **izquierda a derecha**. Estas operaciones también tienen la misma prioridad.

EJEMPLO 1.5. Resolver la expresión:

$$3 + 6 \times (5 + 4)^2 - 10 \div 2$$

Usando PEMDAS:

(1) **Paréntesis:** Resolver $(5 + 4)$:

$$3 + 6 \times 9^2 - 10 \div 2$$

(2) **Exponentes:** Resolver 9^2:

$$3 + 6 \times 81 - 10 \div 2$$

(3) **Multiplicación y División:** Resolver de izquierda a derecha:

$$3 + 486 - 5$$

(4) **Suma y Resta:** Resolver de izquierda a derecha:

$$484$$

La regla PEMDAS garantiza que las expresiones matemáticas sean evaluadas correctamente, evitando ambigüedades. Es muy importante seguir el orden correcto para obtener los resultados correctos.

Ejercicios.
Calcular el valor de las siguientes expresiones

(1) $13 - 2 \times 6$
(2) $27 + 3 - 45 \div 5 + 16$
(3) $9 \times (15 \div (6 - 1) - (9 - 3) \div 2)$

1.9. Razones, proporciones y porcentaje

1.9.1. Razones. Una *razón* es una comparación entre dos cantidades mediante una división. Se expresa generalmente como una fracción, con dos puntos (:) o en forma de cociente decimal.

EJEMPLO 1.6. Si en una clase hay 12 niñas y 8 niños, la razón de niñas a niños es:

$$\frac{12}{8} = \frac{3}{2} \quad \text{o también} \quad 3:2$$

1.9.2. Proporciones. La *proporcionalidad* es una relación o razón constante entre diferentes magnitudes que se pueden medir.

EJEMPLO 1.7. Si 2 lápices cuestan 4 bolívares, entonces 4 lápices deben costar 8 bolívares. Esto se expresa como:

$$\frac{2}{4} = \frac{4}{8}$$

Ambas razones son equivalentes, por lo tanto, forman una proporción.

1.9.3. Porcentajes. Un *porcentaje* es una razón que compara una cantidad con 100. Se representa con el símbolo % y se utiliza para indicar partes de un todo.

Cálculo del porcentaje:
Para calcular el $x\,\%$ de un número n, usamos la fórmula:

$$x\,\% \text{ de } n = \frac{x}{100} \cdot n$$

EJEMPLO 1.8. ¿Cuánto es el 25 % de 80?

$$\frac{25}{100} \cdot 80 = 20$$

Aumentos y descuentos porcentuales:

- Para aumentar un valor en un $x\,\%$, se multiplica por $1 + \frac{x}{100}$.
- Para reducir un valor en un $x\,\%$, se multiplica por $1 - \frac{x}{100}$.

EJEMPLO 1.9. Un producto cuesta \$100 y se aplica un descuento del 15 %, ¿Cuál es el precio con descuento?

$$100 \cdot (1 - \frac{15}{100}) = 100 \cdot 0.85 = 85$$

El precio con descuento es de \$85.

Ejemplo 1.10. Un propietario de estacionamiento necesita que el costo, incluyendo el IVA del 16 %, de estacionar por un día en su local sea de \$5 ¿Cuál debe ser el precio que debe colocar al día de estacionamiento?

Sea p el precio de la hora, entonces de se debe cumplir

$$\left(1+\frac{16}{100}\right)p=5$$

$$\frac{116}{100}p=5 \qquad \frac{29}{25}p=5 \qquad p=\frac{5\times 25}{29}=4,31$$

El precio debe ser de \$4,31.

Aplicaciones:

Las razones, proporciones y porcentajes se aplican en muchas áreas cotidianas como:

- Finanzas (intereses, impuestos, descuentos)
- Estadísticas (porcentajes de población, encuestas)
- Recetas (proporciones de ingredientes)
- Escalas (mapas, planos)

Ejercicios.

(1) Si 5 kg de arroz cuestan \$75, ¿cuánto costarán 8 kg?

(2) ¿Cuál es el 18 % de 240?

(3) Una prenda tiene un precio de \$120. Si se le aplica un aumento del 10 %, ¿cuál es su nuevo precio?

(4) La razón entre hombres y mujeres en una sala es 2:3. Si hay 20 hombres, ¿cuántas mujeres hay?

(5) El costo de un paquete de harina de maíz incluyendo el IVA del 16 % es de $\$1,3$. ¿Cuál es el precio sin IVA de la harina de maíz?

1.10. Concepto de función y definiciones básicas

Una *función* f es una regla de correspondencia que asigna a cada elemento x de un conjunto A exactamente un elemento, llamado $f(x)$, de un conjunto B.

Para la función f se suele usar la notación $f : A \to B$, el conjunto A se denomina *dominio* de f y al conjunto $\{f(x) \,:\, x \in A\}$ se le llama rango de f, notar que el rango es un subconjunto de B.

Ejemplo 1.11. Si $A = \{1, 2, 3, 4\}$ y $B = \{a, b, c, d, e, f\}$, la siguiente regla de correspondencia define una función de A en B

$$f(1) = a \quad f(2) = b \quad f(3) = c \quad f(4) = d.$$

En este caso el dominio de f es el conjunto $A = \{1,2,3,4\}$ y el rango de f es el conjunto $B = \{a,b,c,d\}$. Es usual una representación gráfica del siguiente tipo para esta clase de funciones

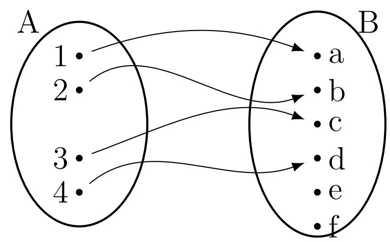


FIGURA 1.5. Representación básica de una función

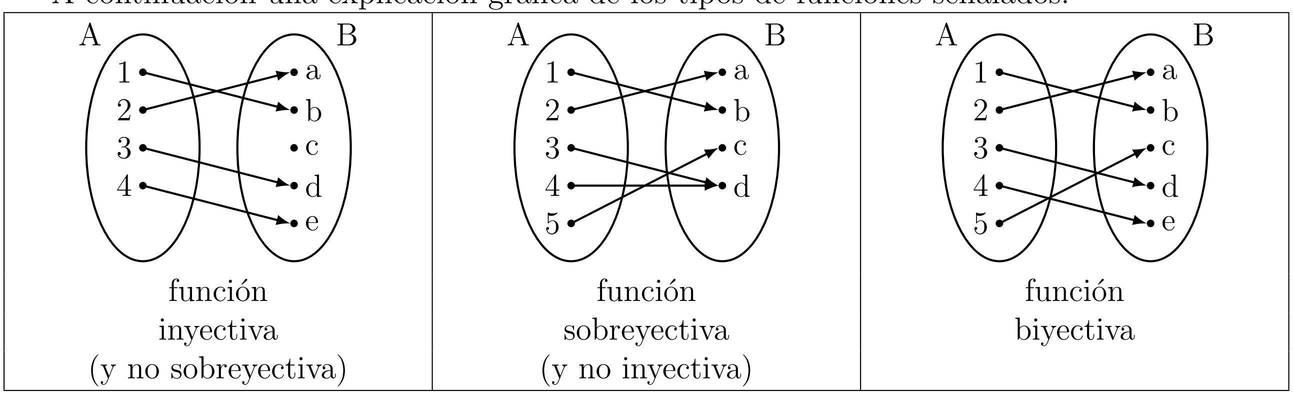
Sea $f : A \to B$ una función.

Se dice que f es *inyectiva* o *uno a uno* si $f(x_1) \neq f(x_2)$ siempre que $x_1 \neq x_2$.

Se dice que f es *sobreyectiva* si el rango de f es todo el conjunto B, es decir si para cada $y \in B$ existe $x \in A$ tal que $f(x) = y$.

Se dice que f es *biyectiva* si si f es inyectiva y sobreyectiva.

A continuación una explicación gráfica de los tipos de funciones señalados.



1.10.1. Gráfico de una función. Lo usual de ahora en adelante será considerar funciones cuyo dominio y rango son subconjuntos del conjunto de los números reales, estas funciones tienen una representación gráfica que corresponde con el subconjunto del plano $\mathbb{R}^2$ definido por

$$\{(x,y) \in \mathbb{R}^2 \ : \ x \in \text{Dominio}(f) \text{ y } y = f(x)\}.$$

El gráfico de una función suele ser una curva en el plano, tal como se ilustra en la siguiente figura.

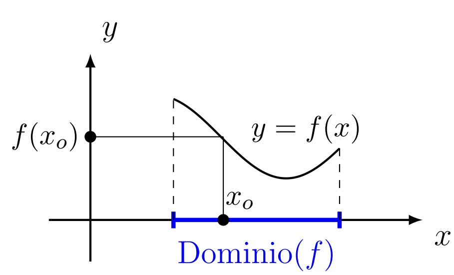
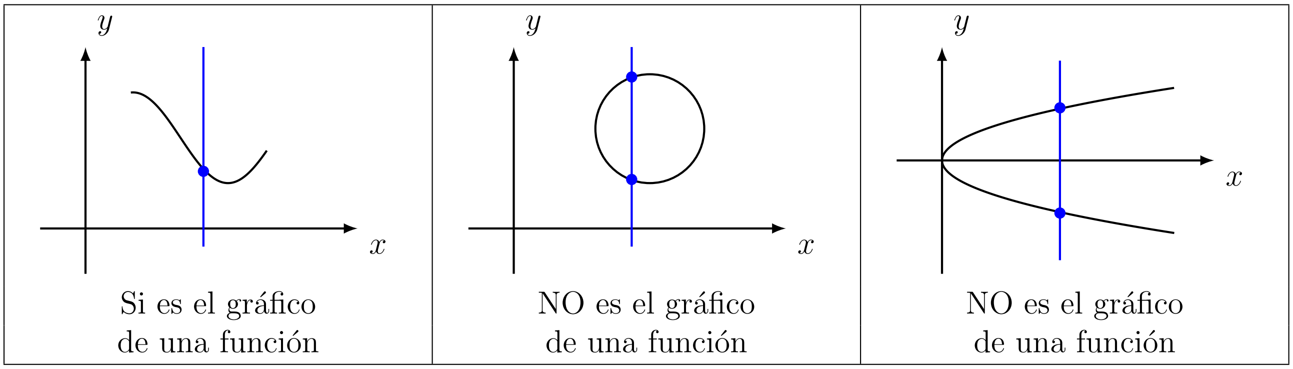


FIGURA 1.6. Gráfico de una función

Si f es una función, a cada x del dominio de f le corresponde un único valor $f(x)$, por lo tanto, una recta vertical solamente puede intersecar el gráfico de f en a lo sumo un punto.



Usualmente las funciones vienen dadas por fórmulas, sin especificar el dominio y se sobrentiende que el dominio de la función debe ser el subconjunto más grande de $\mathbb{R}$ en el cual está definida la fórmula.

Por ejemplo, si $f(x) = x^2 - 4$ el dominio de f es todo $\mathbb{R}$; si $g(x) = \frac{1}{x^2-4}$ el dominio de g es $(-\infty, -2) \cup (-2, 2) \cup (2, +\infty)$; si $h(x) = \sqrt{x-2}$ el dominio de h es $[2, +\infty)$.

A continuación los ejemplos de sus gráficos

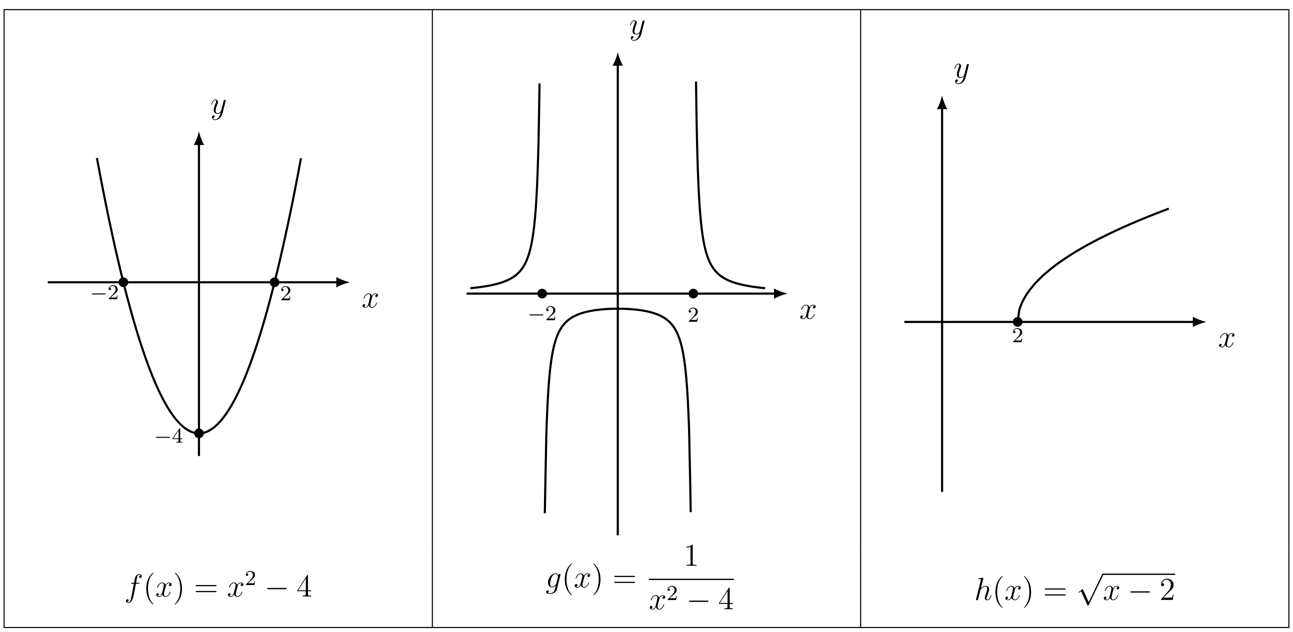


Figura 1.7. Gráficos

Los puntos donde el gráfico de la función corta el eje x se llaman *ceros* o *raíces* de la función y corresponden con la solución de la ecuación $f(x) = 0$, el punto donde el gráfico corta el eje y es el valor de $f(0)$.

Se dice que una función es *creciente* si $f(x_1) < f(x_2)$ si $x_1 < x_2$, se dice que es *decreciente* si $f(x_1) > f(x_2)$ si $x_1 < x_2$. . El gráfico de una función creciente es una curva ascendente y el de una función decreciente es una curva descendente.

Una función es constante cuando existe $C \in \mathbb{R}$ tal que $f(x) = C$ para todo $x \in$ Dominio(f), el gráfico de una función constante es una recta horizontal.

Se dice que una función es *par* si $f(x) = f(-x)$ para todo $x \in$ Dominio(f) y se dice que es impar si $f(x) = -f(-x)$ para todo $x \in$ Dominio(f). El gráfico de una función par es simétrico con respecto al eje y y el gráfico de una función impar es simétrico con respecto al origen.

EJEMPLO 1.12.
En la figura 1.7 se observa lo siguiente:

(1) La función $f(x) = x^2 - 4$ tiene por dominio $(-\infty, +\infty)$ y por rango $[-4, +\infty)$; no es inyectiva, ni sobreyectiva; sus ceros son $x = 2$ y $x = -2$, su gráfico corta el eje y en $y = 4$, es decreciente en el intervalo $(-\infty, 0)$ y creciente en el intervalo $(0, +\infty)$; es par.

(2) La función $f(x) = \frac{1}{x^2-4}$ tiene por dominio $(-\infty, -2) \cup (-2, 2) \cup (2, +\infty)$ y por rango $(-\infty, 0) \cup (0, +\infty)$; no es inyectiva, ni sobreyectiva; no tiene ceros, su gráfico corta el eje y en $y = -\frac{1}{4}$, es creciente en el intervalo $(-\infty, -2) \cup \cup(-2, 0)$ y decreciente en el intervalo $(0, 2) \cup (0, +\infty)$; es par.

(3) La función $f(x) = \sqrt{x-2}$ tiene por dominio $[2, +\infty)$ y por rango $[0, +\infty)$; es inyectiva, no sobreyectiva; tiene un cero en $x = 2$, su gráfico no corta el eje y, es creciente en el intervalo, no es ni par ni impar.

1.11. Ejercicios varios

(1) En los siguientes ejercicios, resuelva la desigualdad lineal, escriba el conjunto solución usando la notación de intervalos y grafique el conjunto solución.
  (a) $3 < x + 4 \leq 10$
  (b) $7 < 3 - \frac{1}{2}x \leq 8$
  (c) $100 + x \leq 41 - 6x \leq 121 + x$
  (d) $-1 \leq \frac{x-4}{4} < \frac{1}{2}$
  (e) $2 \leq \frac{4x+2}{-3} \leq 10$

(2) En los siguientes ejercicios, resuelva la desigualdad no lineal, escriba el conjunto solución usando la notación de intervalos y grafique el conjunto solución.
  (a) $\frac{5}{x} \geq -1$
  (b) $\frac{x-3}{x+2} < 0$
  (c) $\frac{x+1}{x-1} + 2 > 0$
  (d) $\frac{x-2}{x+3} \leq 1$
  (e) $\frac{x(x-1)}{x+5} \geq 0$

(f) $\frac{(1+x)(1-x)}{x} \leq 0$
(g) $\frac{x^2-2x+3}{x+1} \leq 1$
(h) $\frac{x}{x^2-16} > 0$
(i) $\frac{2}{x+3} - \frac{1}{x+1} < 0$
(j) $\frac{4x+5}{x^2} \geq \frac{4}{x+5}$

(3) Un lecho de flores rectangular debe tener una longitud del doble que su ancho. Si el área encerrada debe ser mayor que 98 m^2, ¿qué puede concluir acerca del ancho del lecho de flores?

(4) La relación entre la temperatura T_c en grados Celsius y T_F en grados Fahrenheit es $T_F = \frac{9}{5}T_c + 32$. Se considera que una persona tiene fiebre si su temperatura oral es mayor que 98.6°F. En la escala Celsius, ¿qué temperaturas indican que hay fiebre?

(5) Hallar el dominio de las siguientes funciones
(a) $f(x) = \sqrt{2x-3}$
(b) $h(x) = \sqrt{4-10x}$
(c) $r(x) = \sqrt{x(x-5)}$
(d) $g(x) = \dfrac{1}{\sqrt{x+2}}$

(6) Si $a < b$, entonces explique por qué, en general, $a^2 < b^2$ no es verdadero.

(7) Si a y b son números reales, el número

$$\frac{a+b}{2}$$

se llama *promedio o promedio aritmético* de los números a y b. Explique por qué si a y b son números reales y $a < b$, entonces el promedio de a y b es el punto medio del segmento que une a a con b.

(8) Trace sobre la recta numérica el conjunto de los números reales que satisfacen $2 < |x-3| < 5$, escriba el conjunto solución en notación de intervalos.

(9) Escriba las siguientes expresiones sin los símbolos de valor absoluto.
(a) $|\pi - 4|$
(b) $|\sqrt{5} - 3|$
(c) $|8 - \sqrt{63}|$
(d) $|\sqrt{5} - 2.3|$
(e) $|-6| - |-2|$
(f) $|-3| - |10|$
(g) $|h|$, si h es negativo
(h) $|-h|$, si h es negativo
(i) $|x-6|$, si $x < 6$
(j) $|2x-1|$, si $x \geq \frac{1}{2}$
(k) $|x-y| - |y-x|$

(l) $\dfrac{|x-y|}{|y-x|}$, $x \neq y$

(10) Escriba la expresión $|x-2| + |x-5|$ sin los símbolos de valor absoluto, si x está en el intervalo dado.

(a) $(-\infty, 1)$
(b) $(7, \infty)$
(c) $(3, 4]$
(d) $[2, 5]$

(11) Escriba la expresión $|x+1| - |x-3|$ sin los símbolos de valor absoluto, si x está en el intervalo dado.

(a) $[-1, 3)$
(b) $(0, 1)$
(c) (π, ∞)
(d) $(-\infty, -5)$

Capítulo 2

Introducción a la geometría analítica del plano

2.1. El plano cartesiano

Así como los números reales se encuentran en correspondencia con los puntos en una recta, los puntos del plano se encuentran en correspondencia con pares ordenados de números reales, es decir, los puntos del plano están en correspondencia con el producto cartesiano $\mathbb{R} \times \mathbb{R}$.

Esta correspondencia se hace de la siguiente manera: Se trazan dos rectas perpendiculares, llamadas ejes cartesianos, que se cruzan en el punto correspondiente al número 0 de cada recta. El punto de intersección se llama origen y se representa por el símbolo O. La recta horizontal se llama *eje x* o *eje de abscisas*, la vertical *eje y* o *eje de ordenadas*. Un punto P del plano tiene *coordenadas* (a, b), donde $a, b \in \mathbb{R}$ si una recta vertical trazada desde P al eje x lo corta en el punto a, y una recta horizontal trazada desde P al eje y lo corta en el punto b.

Los ejes de coordenadas dividen el plano en cuatro cuadrantes, la siguiente figura ilustra la ubicación de un punto en el plano y los cuadrantes.

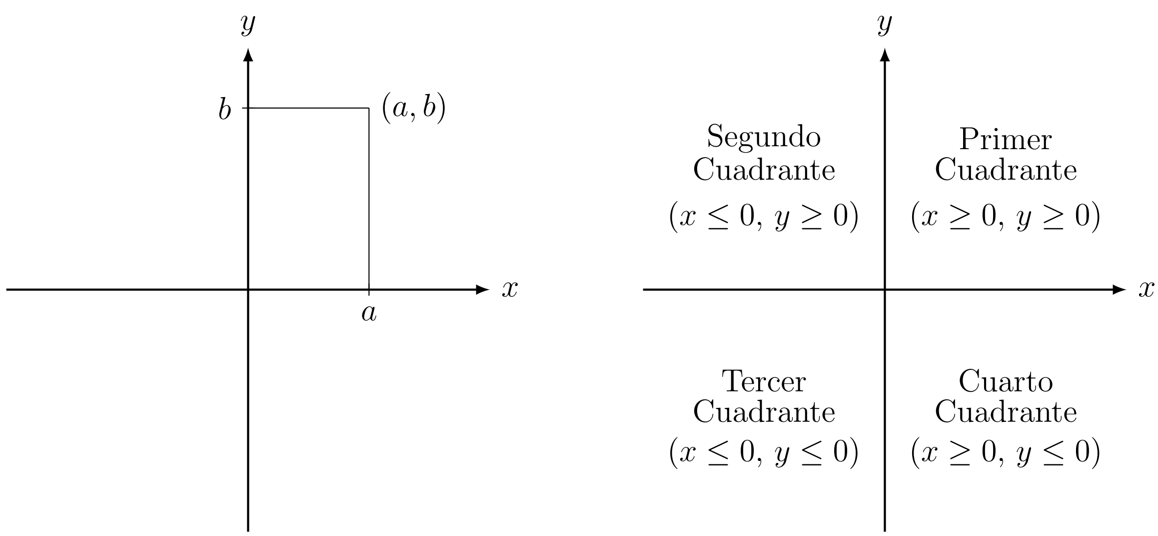


Figura 2.1. Plano cartesiano

La Geometría Analítica emplea métodos algebraicos y ecuaciones para el estudio de problemas geométricos, antes de comenzar su estudio, conviene repasar algunos resultados de Geometría Euclideana.

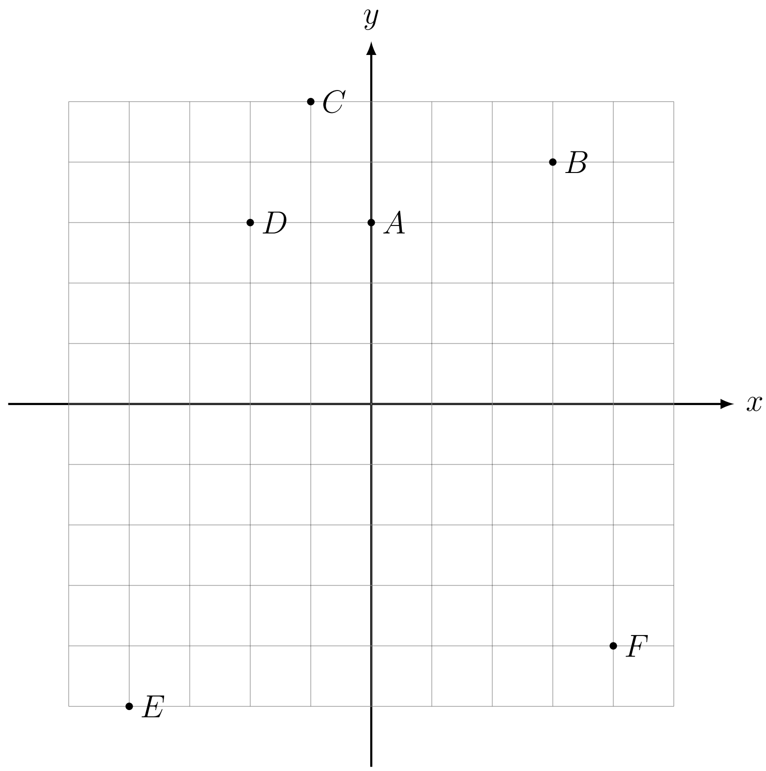
Ejercicios.

(1) Ubicar los siguientes puntos en el plano cartesiano: $(2,3)$, $(4,5)$, $(0,2)$, $(-1,-3)$, $(1,4)$, $(-3,0)$, $(-4,2)$, $(-1,-1)$

(2) Suponiendo que el punto (a,b) está en el cuadrante I, ¿en cuál cuadrante se encuentran los siguientes puntos?
$(-a,b)$, $(a,-b)$, $(-a,-b)$, (b,a), $(-b,a)$, $(-b,-a)$, (a,a), $(b,-b)$, $(-a,-a)$, $(-a,a)$, $(b,-a)$, $(-b,b)$.

(3) Responder la pregunta anterior suponiendo que (a,b) está en el segundo cuadrante.

(4) Indicar las coordenadas de los puntos indicados en la siguiente figura (suponer que la separación de las líneas de la cuadrícula es de una unidad).



2.2. Triángulos semejantes y Teorema de Pitágoras

La Geometría Euclidiana se dedica al estudio de las figuras planas y sólidas basándose en los axiomas y teoremas empleados por el matemático griego Euclides. Nos limitaremos a repasar algunos resultados básicos de geometría euclideana en el plano.

2.2.1. Semejanza de Triángulos. Se dice que dos triángulos son *semejantes* cuando tienen ángulos internos iguales. Como la suma de los ángulos internos de un triángulo es de 180^o, para que dos triángulos sean semejantes basta que tengan dos ángulos iguales.

En la siguiente figura se observan dos triángulos semejantes.

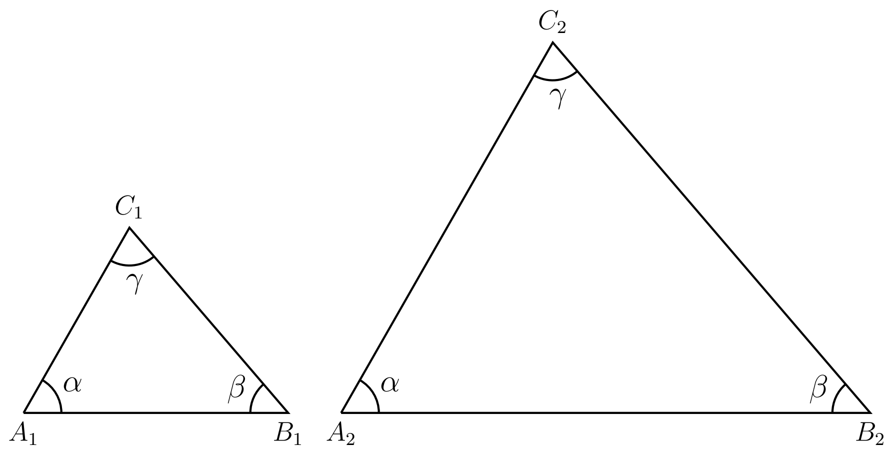


FIGURA 2.2. Triángulos semejantes

Cuando se tienen dos triángulos semejantes, los lados correspondientes son proporcionales, por ejemplo, en la figura 2.2, se tiene que

$$\frac{\overline{A_2B_2}}{\overline{A_1B_1}} = \frac{\overline{A_2C_2}}{\overline{A_1C_1}} = \frac{\overline{B_2C_2}}{\overline{B_1C_1}},$$

donde, si A y B son puntos del plano, $\overline{AB}$ denota la longitud del segmento que une los puntos A y B.

Existe una variedad de criterios para semejanza de triángulos, más información se puede encontrar en [**3**].

Un caso particular que va a resultar de interés es el que se muestra en la siguiente figura

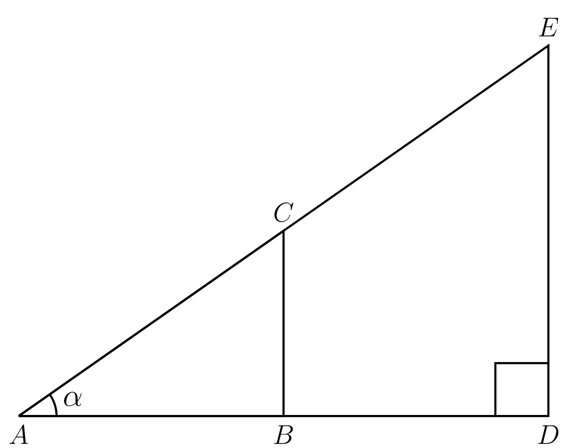


FIGURA 2.3. Semejanza de triángulos rectángulos

En este caso los triángulos $\widehat{ABC}$ y $\widehat{ADE}$ son semejantes, por lo tanto se cumple la siguiente igualdad

$$\frac{\overline{ED}}{\overline{AD}} = \frac{\overline{BC}}{\overline{AB}}.$$

2.2.2. Teorema de Pitágoras.

Supongamos tenemos un triángulo rectángulo $\widehat{ABC}$ como el que se muestra en la siguiente figura y sean a, b y c las longitudes de los segmentos AB, BC y AC, respectivamente.

Los lados AB y BC son llamados *catetos* y el lado AC se llama *hipotenusa*.

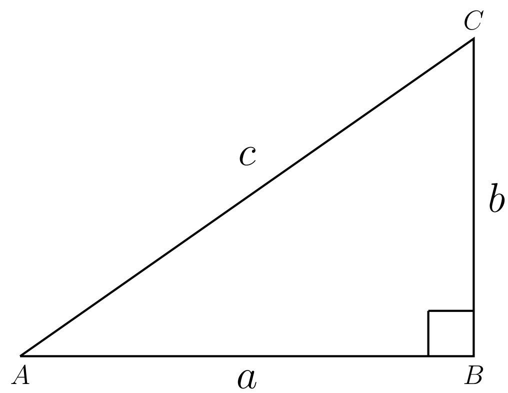


FIGURA 2.4. Teorema de Pitágoras

El Teorema de Pitágoras establece que

$$a^2 + b^2 = c^2,$$

es decir, el cuadrado de la longitud de la hipotenusa es igual a la suma de los cuadrados de las longitudes de los catetos.

2.3. Teorema del coseno

Consideremos un triángulo como el de la siguiente figura

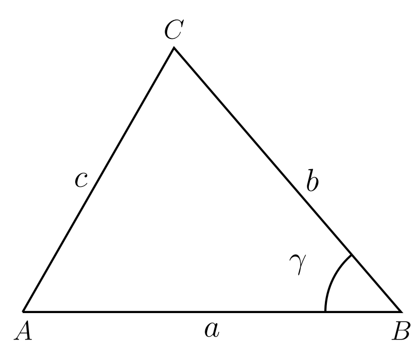


FIGURA 2.5. Teorema del coseno

El Teorema del coseno o ley del coseno establece que

$$c^2 = a^2 + b^2 - 2\,a\,b\,\cos\gamma.$$

El Teorema del coseno es una generalización del Teorema de Pitágoras, ya que al considerar $\gamma = \frac{\pi}{2}$, se obtiene el teorema de Pitágoras.

2.4. Ecuación de la recta

Supongamos que tenemos una recta que pasa por los puntos (x_1, y_1) y (x_2, y_2) y sea (x, y) un punto de la recta, tal como se muestra en la siguiente figura

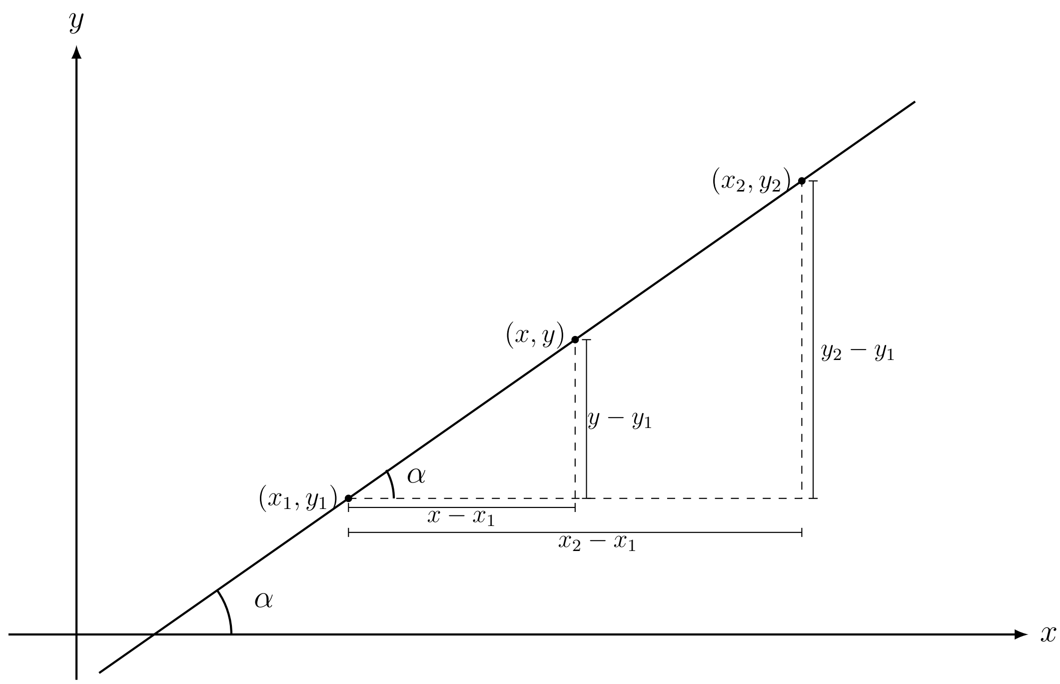


Figura 2.6. Deducción geométrica de la ecuación de la recta

Por semejanza de triángulos tenemos que

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1},$$

despejando y en función de x se obtiene

$$y = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1 = \frac{y_2 - y_1}{x_2 - x_1}\,x - \frac{y_2 - y_1}{x_2 - x_1}\,x_1 + y_1,$$

en resumen, se obtiene una ecuación de la forma

$$y = mx + b,$$

donde $m = \frac{y_2 - y_1}{x_2 - x_1}$ es la llamada *pendiente* de la recta y es la tangente del ángulo que forma la recta con el eje x, el punto $b = -\frac{y_2 - y_1}{x_2 - x_1}\,x_1 + y_1$ se conoce con el nombre de ordenada en el origen y corresponde con el punto de corte de la recta con el eje y.

La pendiente de la recta puede ser positiva, negativa o nula (0). Las rectas de pendiente positiva son aquellas cuya gráfica sube al aumentar x, si la pendiente es negativa la gráfica baja al aumentar x y las de pendiente 0 son rectas horizontales.

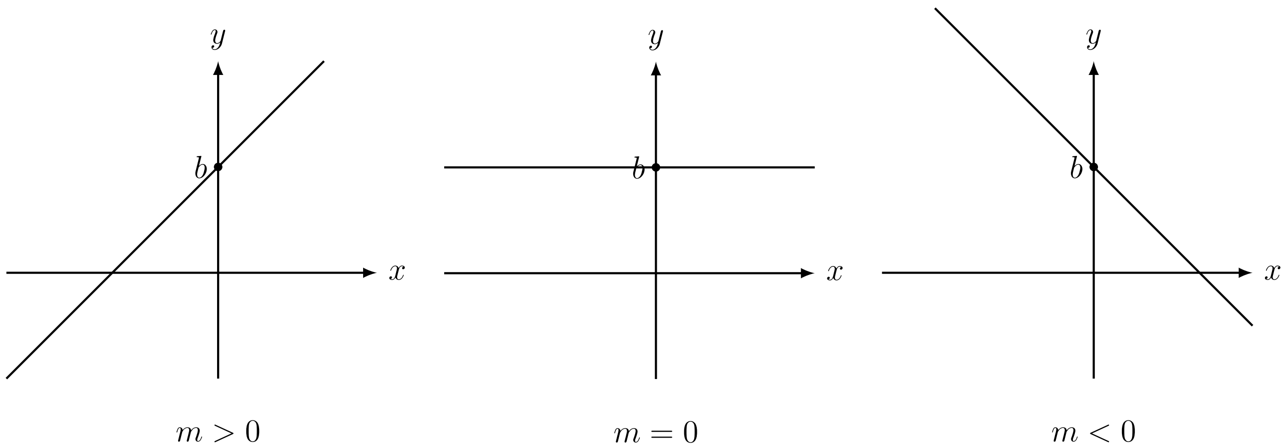


FIGURA 2.7. Aspecto de la recta según su pendiente

OBSERVACIÓN 2.1. Es importante tener en cuenta lo siguiente:

- La ecuación de la recta vertical que para por el punto x_o del eje x es $x = x_o$ (notar que, en este caso, la ecuación no tiene la forma $y = mx + b$) se suele decir que la pendiente de esta recta vertical es "infinita". Por esto es que la ecuación general de la recta es $ax + by + c = 0$ donde a, b y c son constante reales y a y b no pueden ser cero simultáneamente.
- Dos rectas son paralelas si tienen la misma pendiente.
- Dos rectas son perpendiculares si el producto de sus pendientes es -1.

Ejercicios.

(1) Hallar la pendiente de la recta que pasa por los puntos dados y trazar su gráfica.
   (a) $(3, -7)$ y $(1, 0)$.
   (b) $(-4, -1)$ y $(1, -1)$.
   (c) $(5, 2)$ y $(4, -3)$.
   (d) $(1, 4)$ y $(6, -2)$.
   (e) $(8, -\frac{1}{2})$ y $(2, \frac{5}{2})$.

(2) Calcular la pendiente y las intersecciones con los ejes de las siguientes rectas, trazar el gráfico.
   (a) $3x - 4y + 12 = 0$.
   (b) $\frac{1}{2}x - 3y = 3$.
   (c) $-4x - 2y + 6 = 0$.
   (d) $y = 2x + 6$.
   (e) $y + \frac{2}{3}x = 1$.

(3) Hallar la ecuación de la recta que pasa por $(1, 2)$ con la pendiente indicada.
   (a) $\frac{2}{3}$.
   (b) 0.
   (c) $\frac{1}{10}$.
   (d) -2.
(4) Determinar cuáles de las rectas dadas son paralelas y cuáles son perpendiculares.
   (a) $3x - 5y + 9 = 0$.
   (b) $5x = -3y$.
   (c) $-3x + 5y = 2$.
   (d) $3x + 5y + 4 = 0$.
   (e) $-5x - 3y + 8 = 0$.
   (f) $5x - 3y - 2 = 0$.
(5) ¿Cómo probaría o refutaría analíticamente que un triángulo con vértices $(2, 3)$, $(-1, -3)$ y $(4, 2)$ es rectángulo?

2.5. Distancia entre dos puntos del plano cartesiano

Supongamos que tenemos un par de puntos (x_1, y_1) y (x_2, y_2) en el plano cartesiano xy, con estos puntos se forma un triángulo rectángulo, tal como se muestra en la siguiente figura.

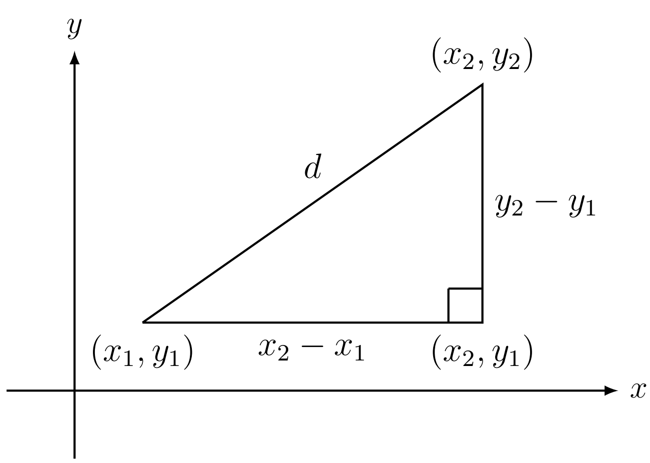


Figura 2.8. Distancia entre dos puntos

Si d es la distancia entre los puntos (x_1, y_1) y (x_2, y_2), por el Teorema de Pitágoras, se tiene que

$$d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2,$$

o lo que es equivalente,

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

2.6. Ecuación de la circunferencia

Supongamos que tenemos una circunferencia con centro (x_o, y_o) y radio $r > 0$ en el plano cartesiano xy.

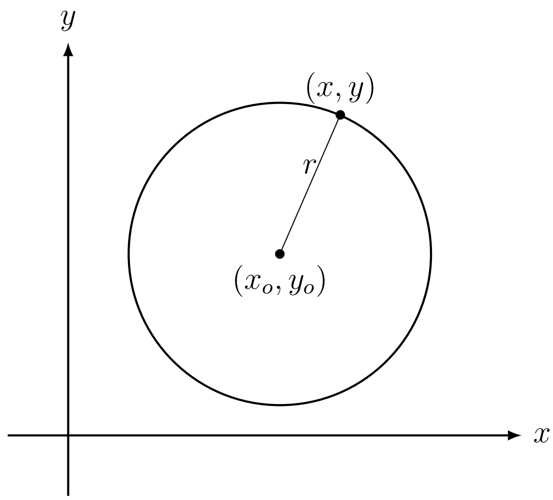


Figura 2.9. Circunferencia de radio r

Sea (x, y) un punto cualquiera de la circunferencia, entonces la distancia del punto (x, y) al punto (x_o, y_o) es r, por lo que por la fórmula de la distancia entre dos puntos se tiene que

$$(x - x_o)^2 + (y - y_o)^2 = r^2.$$

Luego, la ecuación $(x-x_o)^2+(y-y_o)^2 = r^2$ corresponde con la circunferencia con centro (x_o, y_o) y radio r. Además el conjunto $\{(x, y) \in \mathbb{R}^2 \,:\, (x - x_o)^2 + (y - y_o)^2 \leq r^2\}$ es el círculo con centro (x_o, y_o) y radio r.

Ejemplo 2.2. En general la ecuación de una circunferencia no se presenta en su *forma canónica* $(x - x_o)^2 + (y - y_o)^2 = r^2$, sino que aparecen desarrollados los cuadrados, quedando una expresión de la forma $x^2 + y^2 + Ax + By + C = 0$. Para hallar el centro y el radio de una circunferencia cuya ecuación viene dada de esta última forma conviene utilizar la técnica de completación de cuadrados que se detalla a continuación.

Tenemos la ecuación $x^2 + y^2 - 2x - 4y + 1 = 0$ y queremos averiguar si corresponde con una circunferencia y, en caso afirmativo, tenemos que hallar su centro y su radio.

Se procede de la siguiente manera

$$x^2 - 2x = (x - 1)^2 - 1 \qquad \text{y} \qquad y^2 - 4y = (y - 2)^2 - 4,$$

por lo tanto la ecuación dada se transforma en

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

que equivale a

$$(x-1)^2 + (y-2)^2 = 2^2.$$

Por lo tanto, la ecuación dada corresponde con una circunferencia con centro $(1,2)$ y radio 2.

Ejercicios

(1) Determinar el centro y el radio de cada una de las siguientes circunferencias, trazar su gráfico.
   (a) $x^2 + y^2 = 5$.
   (b) $x^2 + y^2 = 9$.
   (c) $x^2 + (y-3)^2 = 49$.
   (d) $(x+2)^2 + y^2 = 36$.
   (e) $(x - \frac{1}{2})^2 + (y - \frac{3}{2})^2 = 1$.
   (f) $(x+3)^2 + (y-5)^2 = 25$.

(2) Hallar el centro y el radio de cada una de las siguientes circunferencias.
   (a) $x^2 + y^2 + 8y = 0$.
   (b) $x^2 + y^2 - 6x = 0$.
   (c) $x^2 + y^2 + 2x - 4y - 4 = 0$.
   (d) $x^2 + y^2 - 18x - 6y - 10 = 0$.
   (e) $x^2 + y^2 - 20x + 16y + 128 = 0$.
   (f) $x^2 + y^2 + 3x - 16y + 63 = 0$.
   (g) $2x^2 + 2y^2 + 4x + 16y + 1 = 0$.

(3) Obtener la ecuación de la circunferencia que satisface las condiciones indicadas.
   (a) Centro $(0,0)$, radio 1.
   (b) Centro $(1,-3)$, radio 5.
   (c) Centro $(0,2)$, radio $\sqrt{2}$.
   (d) Centro $(-9,-4)$, radio $\frac{3}{2}$.
   (e) Extremos de un diámetro en $(-1,4)$ y $(3,8)$.
   (f) Extremos de un diámetro en $(4,2)$ y $(-3,5)$.
   (g) Centro $(0,0)$; la gráfica pasa por $(-1,-2)$.
   (h) Centro $(4,-5)$; la gráfica pasa por $(7,-3)$.
   (i) Centro $(5,6)$; la gráfica es tangente al eje x.
   (j) Centro $(-4,3)$; la gráfica es tangente al eje y.

(4) Trazar la gráfica de la semicircunferencia definida por la ecuación indicada.
   (a) $y = \sqrt{4 - x^2}$
   (b) $x = 1 - \sqrt{1 - y^2}$
   (c) $x = \sqrt{1 - (y-1)^2}$
   (d) $y = -\sqrt{9 - (x-3)^2}$

2.7. La elipse

La ecuación básica de la *elipse* es

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

donde a y b números reales positivos y está representada en la siguiente figura.

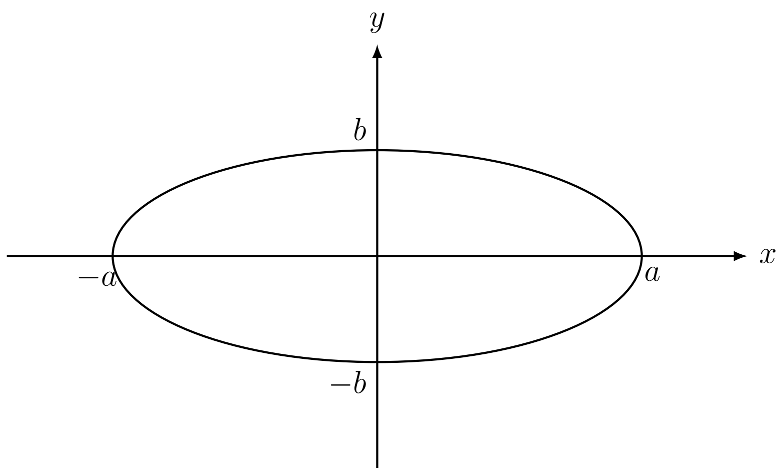


Figura 2.10. Elipse

Una ecuación de la forma

$$\frac{(x - x_o)^2}{a^2} + \frac{(y - y_o)^2}{b^2} = 1$$

también corresponde con una elipse, solo que en vez de tener su centro en $(0, 0)$, lo tiene en (x_o, y_o).

Si $a > b$, el eje mayor de la elipse se ubica en la recta $y = y_o$ y tiene longitud $2a$; el eje menor se ubica en la recta $x = x_o$ y tiene longitud $2b$. Si $a < b$ cambian el eje mayor y el menor. Los vértices de la elipse son los puntos donde interseca los ejes.

Ejemplo 2.3. Identificar las características principales y trazar el gráfico de la elipse de ecuación

$$x^2 + 4y^2 - 2x - 16y + 13 = 0.$$

Completando cuadrados

$$(x - 1)^2 - 1 + 4(y^2 - 4y) + 13 = 0,$$

$$(x - 1)^2 - 1 + 4((y - 2)^2 - 4) + 13 = 0,$$

$$(x - 1)^2 - 1 + 4(y - 2)^2 - 16 + 13 = 0,$$

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

$$(x - 1)^2 + 4(y - 2)^2 = 4,$$

$$\frac{(x - 1)^2}{4} + (y - 2)^2 = 1,$$

$$\frac{(x-1)^2}{2^2} + \frac{(y-2)^2}{1^2} = 1.$$

Tenemos una elipse con centro en el punto $(1,2)$; el eje mayor se ubica sobre la recta $y = 2$ y tiene longitud 4; el eje menor se ubica sobre la recta $x = 1$ y tiene longitud 2. Sus vértices son los puntos $(-1,2), (3,2), (1,1)$ y $(1,3)$.

Podemos representarla gráficamente usando el programa “Symbolab”, de acceso libre en Internet y obtenemos

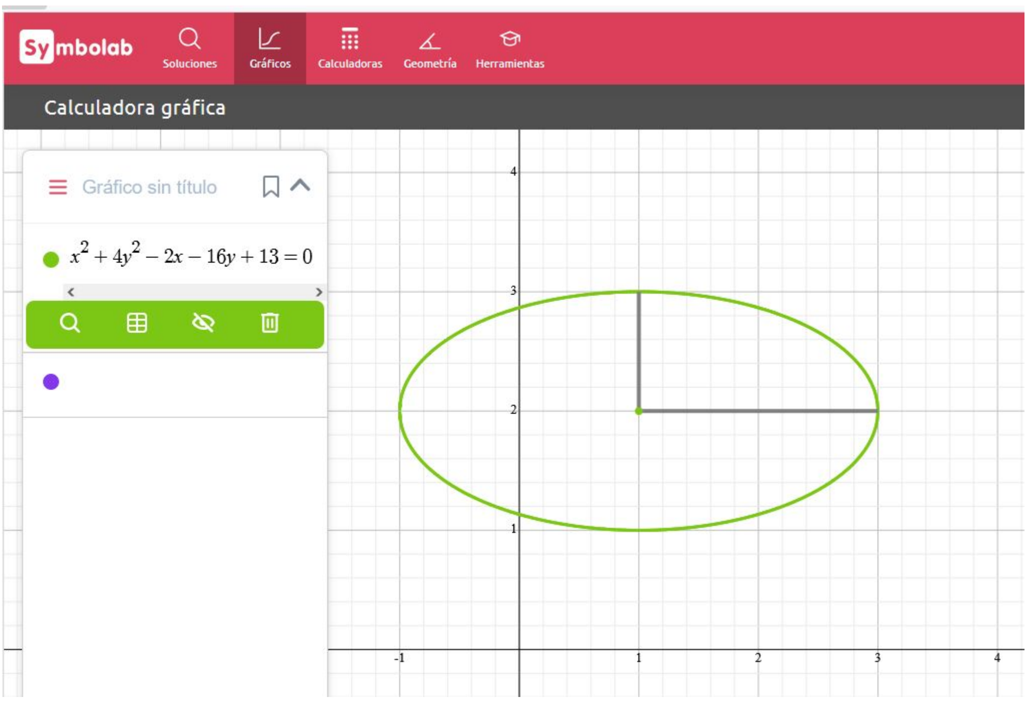


Figura 2.11. Gráfico de la elipse $x^2 + 4y^2 - 2x - 16y + 13 = 0$

Ejercicios.

Hallar el centro, ubicar los ejes mayor y menor, calcular sus longitudes y graficar las siguientes elipses.

(1) $\frac{x^2}{25} + \frac{y^2}{9} = 1.$
(2) $\frac{x^2}{16} + \frac{y^2}{4} = 1.$
(3) $x^2 + \frac{y^2}{16} = 1.$
(4) $\frac{x^2}{4} + \frac{y^2}{10} = 1.$
(5) $9x^2 + 16y^2 = 144.$
(6) $25x^2 + 9y^2 - 100x + 18y - 116 = 0.$
(7) $9x^2 + 5y^2 + 18x - 10y - 31 = 0.$
(8) $x^2 + 3y^2 + 18y + 18 = 0.$

2.8. La parábola

Una parábola es una curva de ecuación

$$y = ax^2 + bx + c$$

donde a, b y c son constantes reales y $a \neq 0$ (notar que si $a = 0$ se trata de una recta).

Dos funciones cuadráticas básicas son $f(x) = x^2$ y $g(x) = -x^2$ y sus gráficos son los siguientes

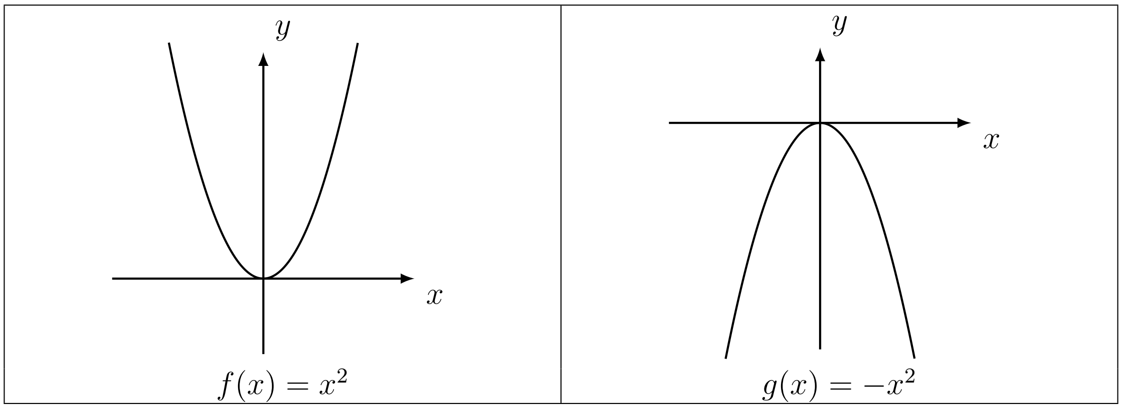


Figura 2.12. Gráficos de $y = x^2$ y $y = -x^2$

El gráfico de $f(x) = x^2$ es una parábola que abre hacia arriba y el de $g(x) = -x^2$ es una parábola que abre hacia abajo. Ambas parábolas tienen por eje de simetría el eje y y por vértice el punto $O = (0, 0)$.

Una función cuadrática de la forma $p(x) = a(x - h)^2 + k$ ($a \neq 0$, h y k constantes) tiene un gráfico similar al de $f(x) = x^2$ si $a > 0$ y, similar al de $g(x) = -x^2$ si $a < 0$. En cualquiera de los dos casos el eje de simetría es la recta $x = h$ y el vértice es el punto (h, k), tal como se ilustra en la siguiente figura.

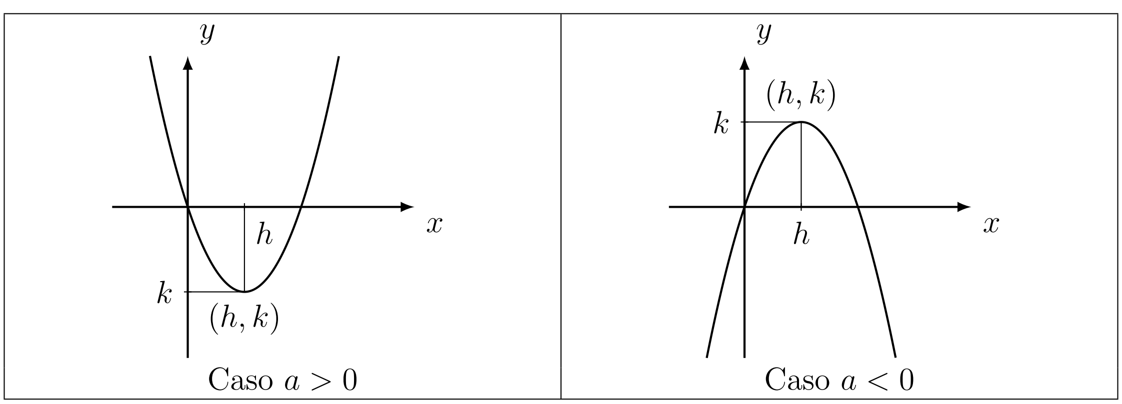


Figura 2.13. Gráficos de $y = a(x - h)^2 + k$

Toda función cuadrática $f(x) = ax^2 + bx + c$, se puede llevar a la forma $a(x - h)^2 + k$ completando cuadrados, de la siguiente manera:

Como $a \neq 0$, f se puede reescribir de la forma

$$f(x) = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right)$$

completando cuadrados

$$x^2 + \frac{b}{a}x = \left(x + \frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 = \left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2},$$

por lo tanto

$$f(x) = a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a}\right) = a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right).$$

De esta última igualdad podemos obtener la fórmula de los ceros de f, si $f(x) = 0$ entonces

$$\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2} = 0,$$

por lo tanto,

$$\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}, \qquad x + \frac{b}{2a} = \pm\frac{\sqrt{b^2 - 4ac}}{2a}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Además,

$$f(\left(-\frac{b}{2a}\right) = -\frac{b^2 - 4ac}{4a^2} = \frac{4ac - b^2}{4a^2}$$

De lo anterior se obtienen las siguientes conclusiones:

Sea $f(x) = ax^2 + bx + c$ una función cuadrática.

(1) El dominio de f es el conjunto de todos los números reales, el gráfico de f es una parábola que abre hacia arriba si $a > 0$ y hacia abajo si $a < 0$.

(2) El eje de simetría de la parábola correspondiente es la recta vertical de ecuación

$$x = -\frac{b}{2a}$$

y el vértice es el punto

$$\left(-\frac{b}{2a}, \frac{4ac - b^2}{4a}\right).$$

(3) Si $a > 0$ el valor mínimo que tomará f es

$$\frac{4ac - b^2}{4a} \text{y lo tomará en } x = -\frac{b}{2a}$$

y el rango de f es el intervalo

$$\left[\frac{4ac - b^2}{4a}, +\infty\right).$$

(4) Si $a < 0$ el valor máximo que tomará f es

$$\frac{4ac-b^2}{4a} \text{y lo tomará en } x = -\frac{b}{2a}$$

y el rango de f es el intervalo

$$\left(-\infty, \frac{4ac-b^2}{4a}\right].$$

(5) Los posibles puntos de corte del gráfico de f con el eje x son los puntos

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a},$$

por lo tanto

- Si $b^2 - 4ac < 0$ el gráfico de f no corta al eje x y el gráfico estará contenido en el semiplano superior o el semiplano inferior.
- Si $b^2 - 4ac = 0$ el gráfico de f corta al eje x en un solo punto y el gráfico estará contenido en el semiplano superior o el semiplano inferior.
- Si $b^2 - 4ac > 0$ el gráfico de f corta al eje x en dos puntos y una parte del gráfico estará contenido en el semiplano superior y otra en el semiplano inferior.

EJEMPLO 2.4. Sea $f(x) = x^2 + x + 1$.
El gráfico, elaborado con "Symbolab" luce así

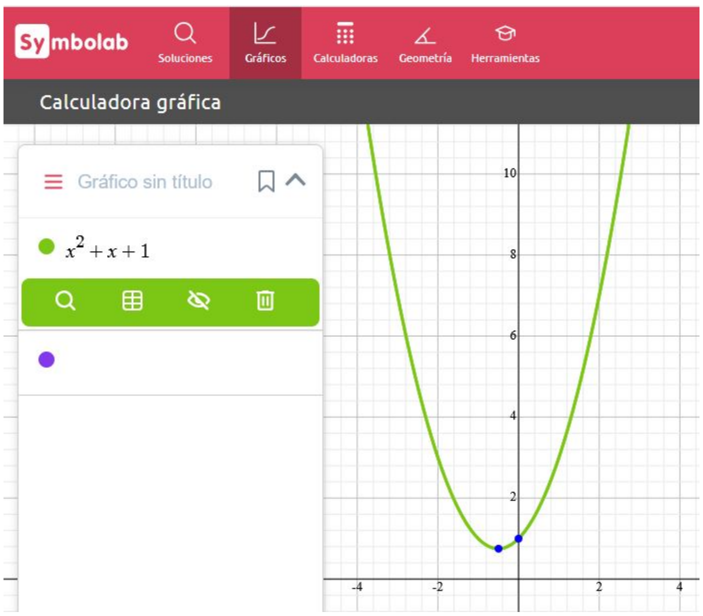


FIGURA 2.14. Gráfico de la parábola $y = x^2 + x + 1$

La justificación analítica del gráfico es la siguiente.
En este caso $a = b = c = 1$, por lo que se tiene que

$$b^2 - 4ac = 1 - 4 = -3 < 0 \qquad -\frac{b}{2a} = -\frac{1}{2}$$

Por lo tanto

(1) El gráfico de f es una parábola que abre hacia arriba y que no corta al eje x.
(2) El eje de simetría de la parábola es $x = -\frac{1}{2}$.
(3) El vértice de la parábola es el punto $\left(-\frac{b}{2a}, \frac{4ac-b^2}{4a}\right) = \left(-\frac{1}{2}, \frac{3}{4}\right)$.
(4) El rango de f es el intervalo $\left[\frac{3}{4}, +\infty\right)$.

A continuación algunos ejemplos gráficos.

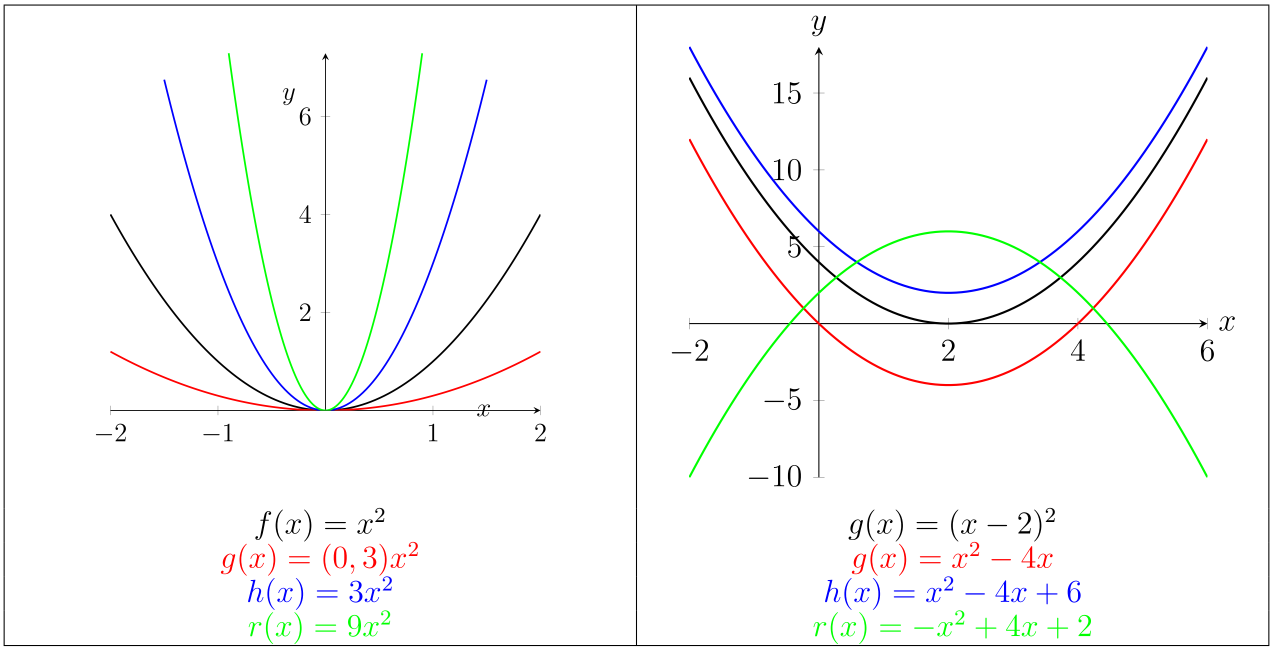


Figura 2.15. Gráficos de diversas funciones cuadráticas

Ejemplo 2.5. Contestar la siguiente pregunta: De todos los rectángulos de perímetro 4 ¿Cuál encierra mayor área?

Supongamos tenemos un rectángulo de perímetro 4 y de lados de longitud x y y, como el perímetro es 4 se debe cumplir que $2x + 2y = 4$, de donde $x + y = 2$ y, por lo tanto $y = 2 - x$.

El área del rectángulo es xy, como $y = x - 2$ el área es la siguiente función de x

$$f(x) = xy = x(2-x) = 2x - x^2$$

La función f es una función cuadrática que alcanza su valor máximo en $x = -\frac{2}{-2} = 1$, por lo que el máximo se alcanza cuando $x = y = 1$.

La respuesta a la pregunta es de todos los rectángulos de perímetro 4, el que encierra mayor área es el cuadrado de lado 1.

Ejercicios.

Hacer un estudio analítico de las parábolas que corresponden con las siguientes ecuaciones, identificar sus características principales, hacer el gráfico a mano y comprobarlo con un programa informático.

(1) $y^2 = -10x$.
(2) $x^2 = -16y$.
(3) $x^2 = \frac{1}{10}y$.
(4) $x^2 = 28y$.
(5) $x^2 = -64y$.
(6) $(y-1)^2 = 16x$.
(7) $(y+3)^2 = -8(x+2)$.
(8) $(x+5)^2 = -4(y+1)$.
(9) $(x-2)^2 + y = 0$.
(10) $y^2 + 12y - 4x + 16 = 0$.
(11) $x^2 + 6x + y + 11 = 0$.
(12) $3(y-1)^2 = 9x$.
(13) $-2x^2 + 12x - 8y - 18 = 0$.
(14) $4y^2 + 16y - 6x - 2 = 0$.
(15) $6y^2 - 12y - 24x - 42 = 0$.
(16) $3x^2 + 30x - 8y + 75 = 0$.

2.9. La hipérbola

La *hipérbola* es la curva cuya ecuación básica (o normal) es de la forma

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

ó

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1.$$

Si despejamos y de la primera ecuación obtenemos

$$y = \pm\frac{b}{a}\sqrt{x^2 - a^2},$$

en el primer cuadrante ($x > 0$ y $y > 0$) tendríamos

$$y = \frac{b}{a}\sqrt{x^2 - a^2} \qquad \text{y} \qquad \frac{y}{x} = \frac{b}{a}\sqrt{1 - \frac{a^2}{x^2}},$$

de manera que para x muy grande tenemos

$$\frac{y}{x} \approx \frac{b}{a},$$

lo que equivale a

$$y \approx \frac{b}{a}x.$$

Esto implica que para x muy grande, el gráfico de la hipérbola está cerca del gráfico de la recta $y = \frac{b}{a}x$; en este caso y otros similares se dice que la recta $y = \frac{b}{a}x$ es una *asíntota* del gráfico de la hipérbola. En el cuarto cuadrante ($x > 0$ y $y < 0$) tendríamos

$$y = \frac{b}{a}\sqrt{x^2 - a^2} \qquad \text{y} \qquad \frac{y}{x} = -\frac{b}{a}\sqrt{1 - \frac{a^2}{x^2}},$$

de manera que para $|x|$ muy grande y $x < 0$ tenemos

$$\frac{y}{x} \approx -\frac{b}{a},$$

lo que equivale a

$$y \approx -\frac{b}{a}x.$$

En este cuadrante la recta $y = -\frac{b}{a}x$ es una asíntota del gráfico de la hipérbola.

A continuación el gráfico de la hipérbola de ecuación

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

con sus asíntotas, sus vértices en $(a, 0)$ y $(-a, 0)$ y sus ejes de simetría que son el eje x y el eje y.

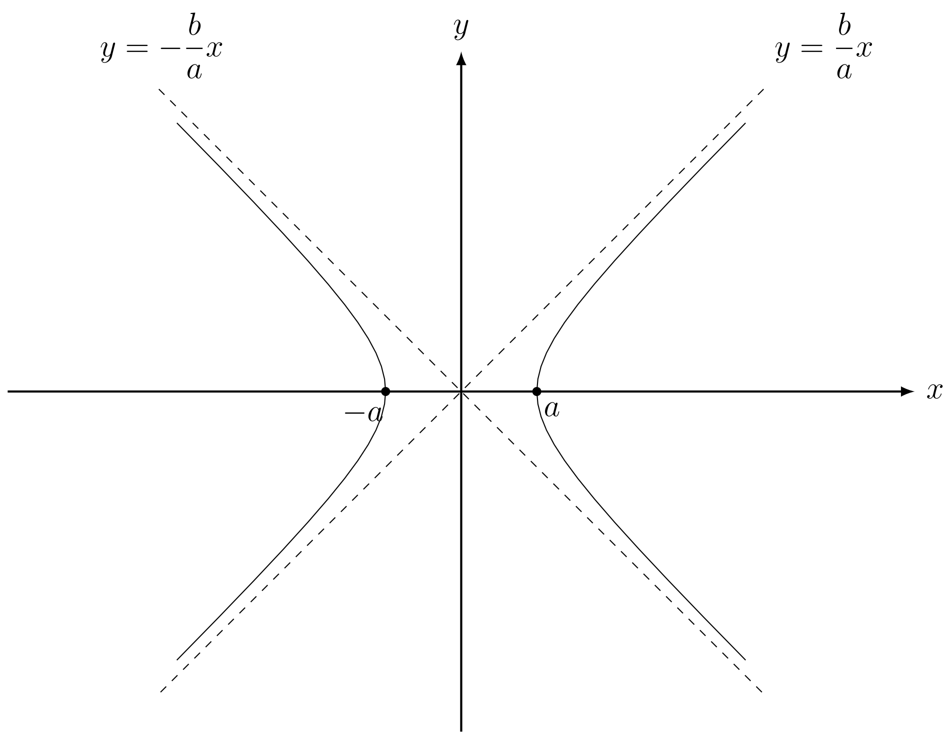


Figura 2.16. Gráfico de la hipérbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

Con un razonamiento similar para la hipérbola de ecuación

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1,$$

podemos obtener que su gráfico es el siguiente

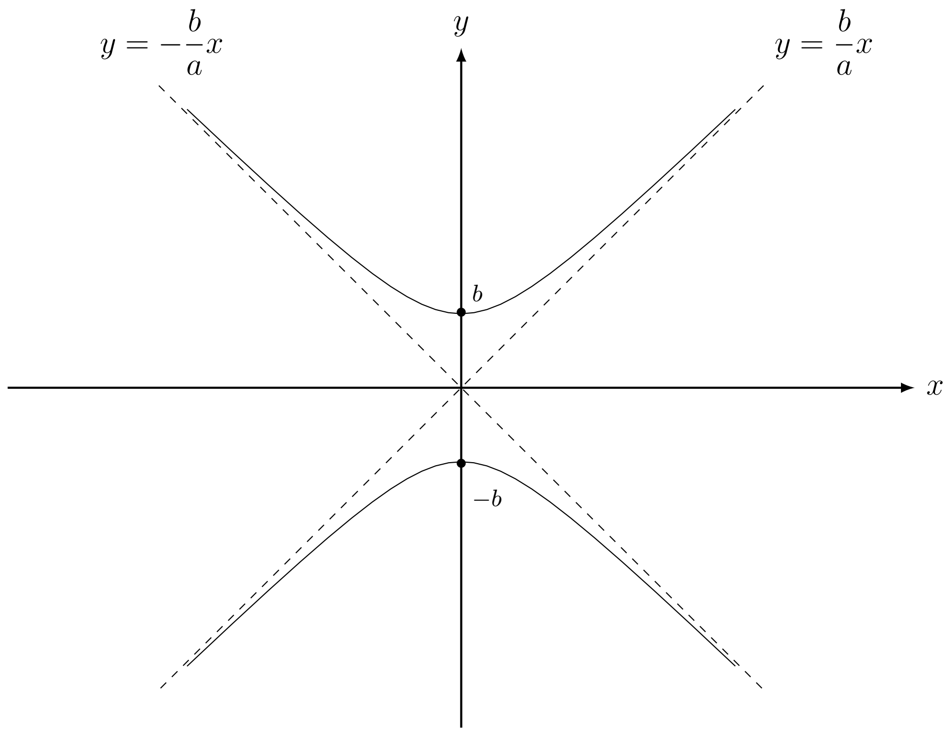


FIGURA 2.17. Gráfico de la hipérbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

EJEMPLO 2.6. Identificar las características principales y trazar el gráfico de la hipérbola de ecuación

$$9x^2 - 4y^2 - 18x - 16y - 43 = 0.$$

Completando cuadrados

$$9(x^2 - 2x) - 4(y^2 + 4y) - 43 = 0,$$

$$9((x-1)^2 - 1) - 4((y+2)^2 - 4) - 43 = 0$$

$$9(x-1)^2 - 4(y+2)^2 - 36 = 0$$

$$\frac{(x-1)^2}{4} - \frac{(y-2)^2}{9} = 1.$$

Los ejes de simetría serían las rectas $x = 1$ y $y = -2$

Las pendiente de las asíntotas serían

$$\frac{\sqrt{9}}{\sqrt{4}} = \frac{3}{2} \qquad \text{y} \qquad -\frac{3}{2},$$

como ambas asíntotas deben pasar por el punto $(1, -2)$ las asíntotas tienen ecuación

$$y = \frac{3}{2}(x-1) - 2 \qquad \text{y} \qquad y = -\frac{3}{2}(x-1) - 2.$$

A continuación el gráfico de la hipérbola elaborado con “Symbolab”.

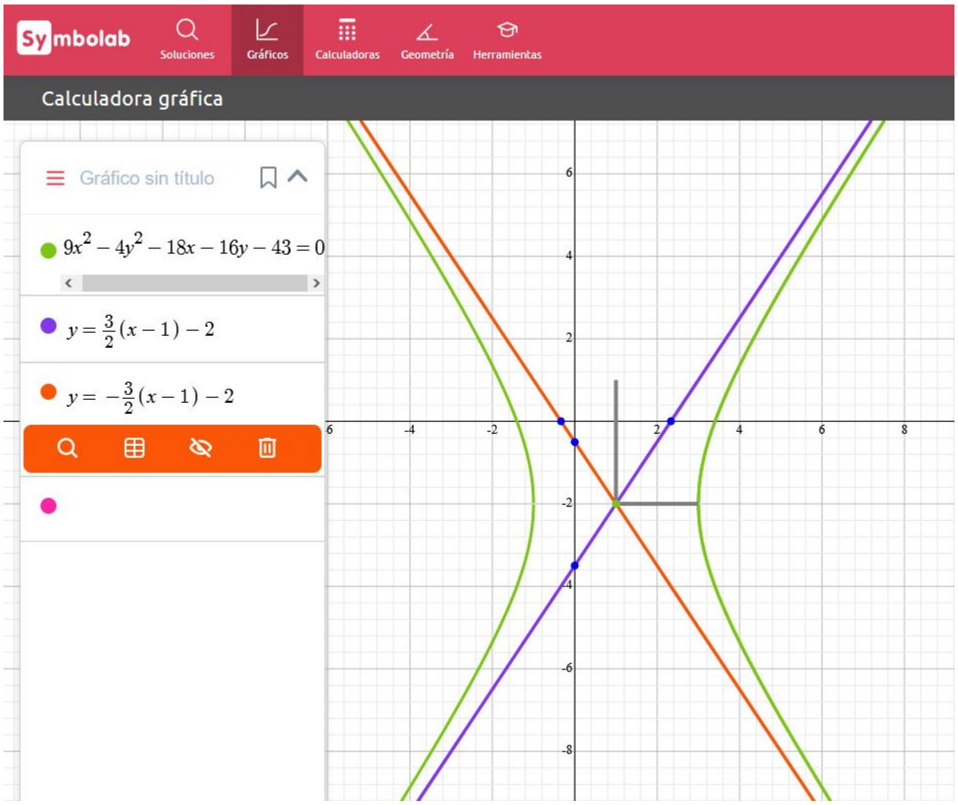


Figura 2.18. Gráfico de la hipérbola $9x^2 - 4y^2 - 18x - 16y - 43 = 0$

Ejercicios.

Hacer un estudio analítico de las hipérbolas que corresponden con las siguientes ecuaciones, identificar sus características principales, hacer el gráfico a mano y comprobarlo con un programa informático.

(1) $\frac{x^2}{16} - \frac{y^2}{25} = 1$
(2) $\frac{x^2}{4} - \frac{y^2}{4} = 1$
(3) $\frac{y^2}{64} - \frac{x^2}{9} = 1$
(4) $\frac{y^2}{6} - 4x^2 = 1$
(5) $9(x-1)^2 - 81(y-2)^2 = 9$
(6) $5x^2 - 6y^2 - 20x + 12y - 16 = 0$
(7) $16x^2 - 25y^2 - 256x - 150y + 399 = 0$
(8) $4x^2 - y^2 - 8x + 6y - 4 = 0$
(9) $2y^2 - 9x^2 - 18x + 20y + 5 = 0$

2.10. Ejercicios varios

(1) Encontrar la ecuación de la recta que satisface las condiciones dadas.
   (a) Pasa por $(2, 3)$ y $(6, -5)$.
   (b) Pasa por $(5, -6)$ y $(4, 0)$.
   (c) Pasa por $(-2, 4)$ y es paralela a la recta de ecuación $3x + y - 5 = 0$.
   (d) Pasa por $(5, -7)$ y es paralela al eje y.
   (e) Pasa por el origen y es paralela a la recta que pasa por $(1, 0)$ y $(-2, 6)$.
   (f) Pasa por $(2, 3)$ y es perpendicular a $x - 4y + 1 = 0$.
   (g) Pasa por $(0, -2)$ y es perpendicular a $3x + 4y + 5 = 0$.
   (h) Pasa por $(-5, -4)$ y es perpendicular a la recta que pasa por $(1, 1)$ y $(3, 11)$.
   (i) Pasa por el origen y es perpendicular a todas las rectas con pendiente 2.

(2) ¿Cómo demostraría o refutaría de manera analítica que el cuadrilátero con vértices $(0, 4)$, $(-1, 3)$, $(-2, 8)$ y $(-3, 7)$ es un paralelogramo?

(3) Halle la ecuación de la mitad superior de la circunferencia $x^2 + (y - 3)^2 = 4$. Repita lo anterior con respecto a la mitad derecha de la circunferencia.

(4) Halle la ecuación de la mitad inferior de la circunferencia $(x - 5)^2 + (y - 1)^2 = 9$. Repita lo anterior con respecto a la mitad izquierda de la circunferencia.

(5) Trazar el conjunto de puntos en el plano xy, cuyas coordenadas satisfagan la desigualdad dada.
   (a) $x^2 + y^2 \geq 9$.
   (b) $(x - 1)^2 + (y + 5)^2 \leq 25$.
   (c) $1 \leq x^2 + y^2 \leq 4$.
   (d) $x^2 + y^2 > 2y$.

(6) Representar gráficamente el conjunto de los puntos del plano xy que satisfacen cada una de las siguientes condiciones.
   (a) $y = -3x$.
   (b) $y - 2x = 0$.
   (c) $-x + 2y = 1$.
   (d) $2x + 3y = 6$.
   (e) $x = y^2$.
   (f) $y = x^3$.
   (g) $y = x^2 - 4$.
   (h) $x = 2y^2 - 4$.
   (i) $y = x^2 - 2x - 2$.
   (j) $y^2 = 16(x + 4)$.
   (k) $y = x(x^2 - 3)$.
   (l) $y = (x - 2)^2(x + 2)^2$.

(m) $x = -\sqrt{y^2 - 16}$.
(n) $y^3 - 4x^2 + 8 = 0$.
(o) $(x-1)^2 + y^2 = 0$.
(p) $y = \sqrt{x-3}$.
(q) $y = 2 - \sqrt{x+5}$.
(r) $y = |x-9|$.
(s) $x = |y| - 4$.
(t) $|x| + |y| = 4$.
(u) $x + 3 = |y-5|$.

(7) Identificar el tipo de curva que corresponde con cada una de las siguientes ecuaciones, encontrar sus características especiales, graficarla a mano y comprobar el resultado con un programa informático.
(a) $\frac{(x-5)^2}{4} - \frac{(y+1)^2}{49} = 1$.
(b) $\frac{(x-1)^2}{49} + \frac{(y-3)^2}{36} = 1$.
(c) $\frac{(x+1)^2}{25} + \frac{(y-2)^2}{36} = 1$.
(d) $x^2 - 2x - 4y + 17 = 0$.
(e) $\frac{(y-4)^2}{36} - x^2 = 1$.
(f) $\frac{(y-\frac{1}{4})^2}{4} - \frac{(x+3)^2}{9} = 1$.
(g) $y^2 - 8y + 2x + 10 = 0$.
(h) $y^2 - 4y - 4x + 3 = 0$.
(i) $(x+5)^2 + \frac{(y+2)^2}{16} = 1$.
(j) $x^2 + 5x - \frac{1}{4}y + 6 = 0$.
(k) $4x^2 = 2y$.
(l) $4x^2 - 16y^2 = 64$.
(m) $5x^2 - 5y^2 = 25$.
(n) $y^2 - 5x^2 = 20$.
(o) $9x^2 - 16y^2 + 144 = 0$.
(p) $\frac{(x+2)^2}{10} - \frac{(y+4)^2}{25} = 1$.

(8) Deducir la ecuación de la hipérbola que satisface cada una de las siguientes condiciones
(a) Vértices en $(\pm 2, 0)$, pasa por $(2\sqrt{3}, 4)$.
(b) Vértices en $(0, \pm 3)$, pasa por $(\frac{5}{5}, 5)$.
(c) Centro en $(-1, 3)$, un vértice en $(-1, 4)$, pasa por $(-5, 3 + \sqrt{5})$.
(d) Centro en $(3, -5)$, un vértice en $(3, -2)$, pasa por $(1, -1)$.

(9) Deducir la ecuación de la elipse que satisface cada una de las siguientes condiciones
(a) Vértices en $(0, \pm 3)$, extremos del eje menor en $(\pm 1, 0)$.
(b) Vértices en $(\pm 4, 0)$, extremos del eje menor en $(0, \pm 2)$.
(c) Vértices en $(-3, -3)$, $(5, -3)$, extremos del eje menor en $(1, -1)$, $(1, -5)$.

(d) Vértices en $(1,-6)$, $(1,2)$, extremos del eje menor en $(-2,-2)$, $(4,-2)$.

(10) Deducir la ecuación de la parábola que satisface cada una de las siguientes condiciones

(a) Vértice en $(0,0)$, pasa por $(-2,8)$, eje a lo largo del eje y

(b) Vértice en $(0,0)$, pasa por $(1,\frac{1}{4})$, eje a lo largo del eje x

(11) **Ingresos** Un fabricante encuentra que el ingreso generado por vender x unidades de cierta mercancía está dado por la función $R(x) = 80x - 0.4x^2$, donde el ingreso $R(x)$ se mide en dólares. ¿Cuál es el ingreso máximo, y cuántas unidades deben fabricarse para obtener este máximo?

(12) **Ventas** Un vendedor de bebidas gaseosas en una conocida playa analiza sus registros de ventas y encuentra que, si vende x latas de gaseosa en un día, su utilidad (en dólares) está dada por

$$P(x) = -0.001x^2 + 3x - 1800$$

¿Cuál es su utilidad máxima por día, y cuántas latas debe vender para obtener una utilidad máxima?

(13) **Publicidad** La efectividad de un anuncio comercial por televisión depende de cuántas veces lo ve una persona. Después de algunos experimentos, una agencia de publicidad encontró que, si la efectividad E se mide en una escala de 0 a 10, entonces

$$E(n) = \frac{1}{n} - \frac{1}{20}n^2$$

donde n es el número de veces que una persona ve un anuncio comercial determinado. Para que un anuncio tenga máxima efectividad, ¿cuántas veces debe verlo una persona?

(14) **Productos farmacéuticos** Cuando cierto medicamento se toma oralmente, la concentración del medicamento en el flujo sanguíneo del paciente después de t minutos está dada por $C(t) = 0.06t - 0.0002t^2$, donde $0 \leq t \leq 240$ y la concentración se mide en mg/L. ¿Cuándo se alcanza la máxima concentración de suero, y cuál es esa máxima concentración?

(15) **Agricultura** El número de manzanas producidas por cada árbol en una huerta de manzanos depende de la densidad con la que estén plantados los árboles. Si n árboles se plantan en un acre de terreno, entonces cada árbol produce $900 - 9n$ manzanas. Por tanto, el número de manzanas producidas por acre es

$$A(n) = n(900 - 9n)$$

¿Cuántos árboles deben plantarse por acre para obtener la máxima producción de manzanas?

Capítulo 3

Funciones algebraicas

3.1. Funciones polinómicas

Una *función polinómica* o *polinomio* es una función de la forma

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_o \tag{3.1}$$

donde n es un entero no negativo.

Las funciones constantes, las de la forma $f(x) = mx + b$ (a veces llamadas lineales) y las cuadráticas son casos particulares de funciones polinómicas y corresponden con los casos $n = 0$, $n = 1$ y $n = 2$, respectivamente.

El *grado* del polinomio es la mayor potencia de x que parece en el polinomio, es decir si $a_n \neq 0$, el grado del polinomio (3.1) es n; el coeficiente a_n se llama *coeficiente principal* o *coeficiente director* y al coeficiente a_o se le conoce como término independiente.

3.1.1. Gráficos de polinomios. Para $|x|$ muy grande el comportamiento del gráfico del polinomio

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_o$$

($a_n \neq 0$) es similar al del gráfico de $a_n x^n$.

En detalle: sea c una constante, si n es par tenemos la siguiente representación gráfica

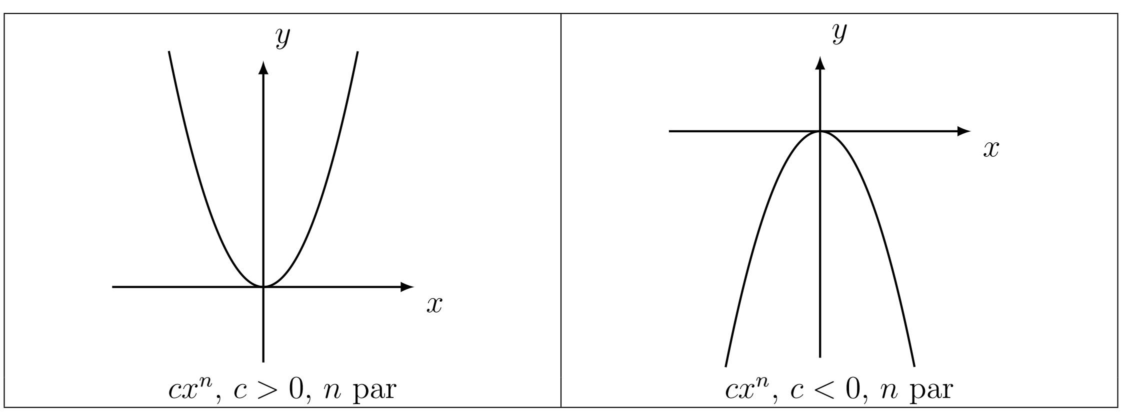
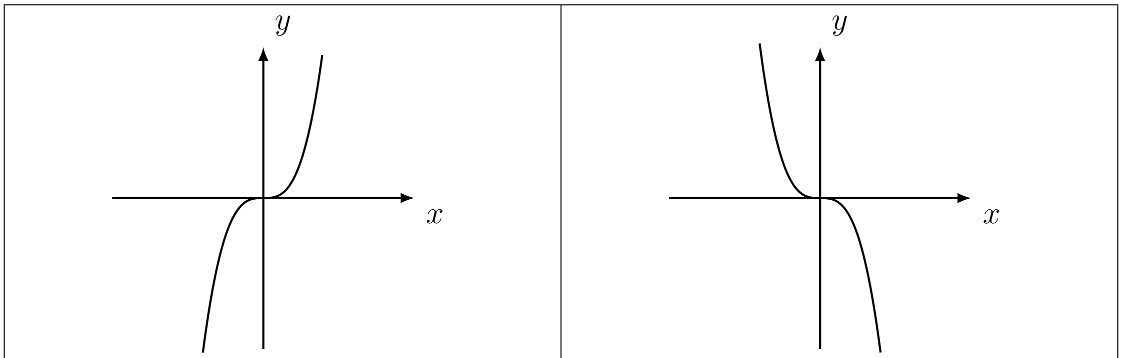


FIGURA 3.1. Gráficos de cx^n, n par

Para $|x|$ muy grande, el gráfico se comporta de manera similar al de $p(x) = x^2$ ó $p(x) = -x^2$, siendo más abierto o cerrado según los valores de c y n.

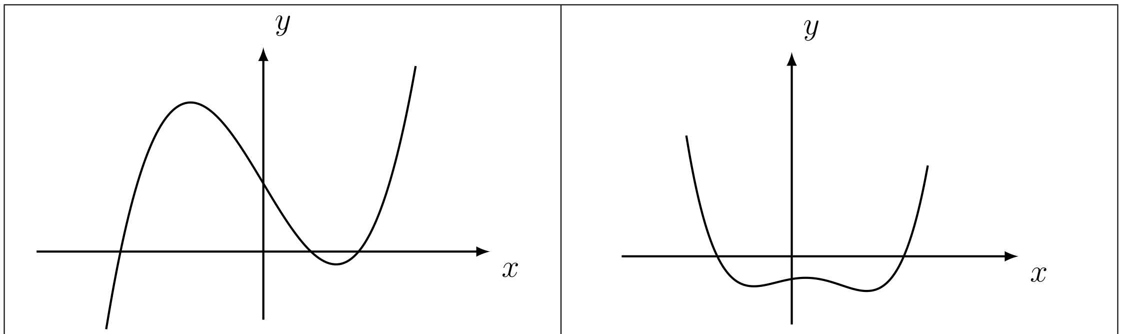
Para n impar la similitud se da con el gráfico de $p(x) = x^3$.



cx^n, $c > 0$, n impar cx^n, $c < 0$, n impar

FIGURA 3.2. Gráficos de cx^n, n impar

EJEMPLO 3.1. Ejemplo de gráficos de polinomios de grado impar y par.



$p(x) = x^3 - 7x + 6$ $p(x) = 2x^4 - 2x^3 - 2x^2 + x - 2$

FIGURA 3.3. Gráficos de polinomios

3.1.2. Operaciones con polinomios. Los polinomios se pueden sumar, restar y multiplicar y el resultado siempre será otro polinomio.

También se puede dividir un polinomio entre otro polinomio de menor grado.

Al dividir el polinomio $D(x)$ (el *dividendo*) entre otro polinomio de menor grado, (el *divisor*), se va a obtener el polinomio *cociente* $c(x)$ y el resto $r(x)$; este último debe tener grado estrictamente menor que el divisor $d(x)$ y se cumple

$$\frac{D(x)}{d(x)} = c(x) + \frac{r(x)}{d(x)} \quad \text{o lo que es equivalente} \quad D(x) = c(x)d(x) + r(x).$$

Para hacer la división de polinomios se procede de manera bastante análoga a como se trabaja con los números. Se ilustrará el procedimiento con ejemplos.

EJEMPLO 3.2. Veamos el cociente del polinomio $x^4 - 3x^3 + 5x^2 - 4x + 2$ entre $x^2 - x + 1$.

$$\begin{array}{r|l}
x^4-3x^3+5x^2-4x+2 & x^2-x+1 \\ \cline{1-2}
-x^4+x^3-x^2 \qquad\qquad & x^2-2x+2 \\ \cline{1-1}
-2x^3+4x^2-4x+2 & \\
2x^3-2x^2+2x \quad & \\ \cline{1-1}
2x^2-2x+2 & \\
-2x^2+2x-2 & \\ \cline{1-1}
0 \quad &
\end{array}$$

En detalle: Primero dividimos x^4, el término de mayor grado del dividendo, entre x^2, el término de mayor grado del divisor, lo que nos da x^2; esto es lo primero que colocamos en el espacio del cociente. Después multiplicamos x^2 por el divisor, lo que nos da $x^4 - x^3 + x^2$, colocamos este polinomio cambiado de signo bajo el dividendo y sumamos. En la misma columna donde está el dividendo obtenemos $-2x^3 + 4x^2 - 4x + 2$, al dividir $-2x^3$ entre x^2 obtenemos $-2x$, colocamos $-2x$ a continuación en el espacio del cociente y repetimos los pasos anteriores hasta llegar a un polinomio de grado menor que 2 en la columna del dividendo.

El resultado obtenido en este ejemplo muestra que

$$x^4 - 3x^3 + 5x^2 - 4x + 2 = (x^2 - 2x + 2)(x^2 - x + 1),$$

es decir, el resto en esta división es 0. En este caso, como es natural, se dice que el polinomio $x^2 - x + 1$ divide al polinomio $x^4 - 3x^3 + 5x^2 - 6x + 2$, o que el polinomio $x^4 - 3x^3 + 5x^2 - 6x + 2$ es múltiplo del polinomio $x^2 - x + 1$.

Ejemplo 3.3. Veamos el cociente del polinomio $x^4 - 3x^3 + 5x^2 - 4x + 2$ entre $x^2 + 1$.
Siguiendo el mismo procedimiento que en el ejemplo anterior.

$$\begin{array}{r|l}
x^4-3x^3+5x^2-4x+2 & x^2+1 \\ \cline{1-2}
-x^4 \qquad -x^2 \qquad\qquad & x^2-3x+4 \\ \cline{1-1}
-3x^3+4x^2-4x+2 & \\
3x^3 \qquad\quad +3x \quad & \\ \cline{1-1}
4x^2-x+2 & \\
-4x^2 \qquad -4 & \\ \cline{1-1}
-x-2 &
\end{array}$$

Por lo tanto,

$$x^4 - 3x^3 + 5x^2 - 4x + 2 = (x^2 - 3x + 4)(x^2 + 1) - x - 2$$

Ejemplo 3.4. Veamos el cociente del polinomio $x^3 - 3x^2 + 2$ entre $x + 1$.
Siguiendo el mismo procedimiento que en el ejemplo anterior.

$$
\begin{array}{r|l}
x^3 - 3x^2 + 2 & x + 1 \\ \hline
-x^4 - x^2 & x^2 - 4x + 4 \\ \cline{1-1}
-4x^2 + 2 & \\
4x^2 + 4x & \\ \cline{1-1}
4x + 2 & \\
-4x - 4 & \\ \cline{1-1}
-2 &
\end{array}
$$

Por lo tanto,

$$x^3 - 3x^2 + 2 = (x^2 - 4x + 4)(x + 1) - 2,$$

el cociente es $x^2 - 4x + 4$ y el resto es -2.

Para el caso en que se divide entre un polinomio de la forma $x - \alpha$ se puede usar el método abreviado o método de Ruffini, que se ilustra a continuación para este caso.

$$
\begin{array}{r|cccc}
 & x^3 & x^2 & x & x^0 \\
 & 1 & -3 & 0 & 2 \\
-1 & & -1 & 4 & -4 \\ \hline
 & 1 & -4 & 4 & -2 \\
 & x^2 & x & x^0 &
\end{array}
$$

Ejemplo 3.5. Veamos el cociente del polinomio $x^3 - 3x^2 + 2$ entre $x - 1$.

Utilizando el método abreviado.

$$
\begin{array}{r|cccc}
 & x^3 & x^2 & x & x^0 \\
 & 1 & -3 & 0 & 2 \\
1 & & 1 & -2 & -2 \\ \hline
 & 1 & -2 & -2 & 0 \\
 & x^2 & x & x^0 &
\end{array}
$$

En este caso el resto es 0 y, por lo tanto,

$$x^3 - 3x^2 + 2 = (x^2 - 2x - 2)(x - 1).$$

Notar que el polinomio $x^3 - 3x^2 + 2$ se anula en $x = 1$.

3.2. Raíces y factorización de polinomios

Sean $p(x)$ un polinomio y a un número real, de acuerdo al algoritmo de la división de polinomios, explicado en la Sección 3.1.2 existen polinomios $q(x)$ y $r(x)$ tales que $p(x) = (x - a)q(x) + r(x)$ y el grado de $r(x)$ es menor que el grado de $x - a$; como el grado de $x - a$ es 1 el grado de $r(x)$ debe ser 0, es decir $r(x)$ es una constante, por lo tanto existe una constante c tal que

$$p(x) = q(x)(x - a) + c.$$

De esta última igualdad se deduce lo siguiente: *Un polinomio $p(x)$ es divisible por $x - a$ si y solo si $p(a) = 0$.*

Supongamos que el polinomio $p(x)$ tiene coeficientes enteros $a_n, \dots, a_o$, que a también es un entero y que $p(a) = 0$, entonces se tiene que

$$a(a_n a^{n-1} + a_{n-1}a^{n-2} + \cdots + a_2 a + a_1) = -a_o,$$

lo que implica que, en este caso, a es un divisor de a_o.

En general, si $f(x)$ es una función, se dice que a es una *raíz* de $f(x)$ si $f(a) = 0$

En resumen:

- El resto que se obtiene al dividir un polinomio $p(x)$ entre $x - a$ es igual a $p(a)$.
- Un polinomio $p(x)$ es divisible entre $x - a$ si y solo si $p(a) = 0$.
- Si $p(x)$ tiene coeficientes enteros y un número entero a es una raíz de $p(x)$, entonces a divide al término independiente de $p(x)$.

Se dice que la raíz a de un polinomio $p(x)$ tiene multiplicidad n, donde n es un entero positivo, si existe otro polinomio $q(x)$ tal que

$$p(x) = (x - a)^n q(x) \quad \text{y} \quad q(a) \neq 0.$$

A las raíces de multiplicidad 1 se les dice *raíces simples* y a las de multiplicidad 2 *raíces dobles.*

Todo polinomio $p(x)$ de grado n tiene a lo sumo n raíces reales contadas de acuerdo a su multiplicidad.

Puede ocurrir que un polinomio no tenga raíces reales, como por ejemplo ocurre con $x^2 + 1$, $x^8 + 9$ o con cualquier polinomio de grado 2 de la forma $ax^2 + bx + c$, donde $a \neq 0$ y $b^2 - 4ac < 0$.

Se dice que un polinomio $p(x)$ es *irreducible* si $p(x)$ no se puede factorizar como el producto de dos polinomios no constantes.

Se cumple lo siguiente:

Todo polinomio irreducible es de la forma $x - a$ ó $ax^2 + bx + c$.

Es decir: Todo polinomio irreducible tiene grado menor o igual que 2.

También se cumple:

Todo polinomio se puede escribir como un producto de polinomios irreducibles.

EJEMPLO 3.6.

(1) $x^2 + 2x + 1 = (x + 1)^2$
(2) $x^4 - 1 = (x - 1)(x + 1)(x^2 + 1)$
(3) $x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$
(4) $(4x - 1)(2x - 2) = 2 - 10x + 8x^2$
(5) $(-2 + x)^2(-1 + x)(3 + x)^3 = -108 + 108x + 45x^2 - 40x^3 - 10x^4 + 4x^5 + x^6$

3.3. Funciones racionales

Una función racional es una función de la forma

$$f(x) = \frac{p(x)}{q(x)},$$

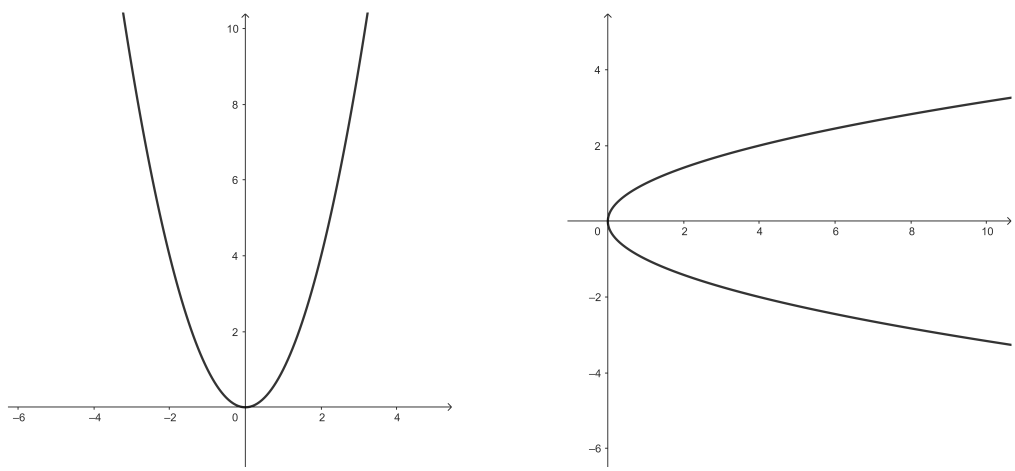
donde $p(x)$ y $q(x)$ son polinomios.

En este caso el dominio de $f(x)$ ($D(f)$) es el conjunto

$$D(f) = \{x \in \mathbb{R} \, : \, q(x) \neq 0\}.$$

3.3.1. La función raíz cuadrada. Es natural pensar en la función raíz cuadrada de x como la inversa de la función $f(x) = x^2$, sin embargo la situación no es tan simple y se deben cuidar muy bien los detalles.

Para cualquier número real x se cumple que $x^2 = (-x)^2$ y $x^2 \geq 0$, por lo tanto la función $f(x) = x^2$ no es inyectiva ni sobreyectiva, gráficamente tenemos



$y = x^2$ $\qquad\qquad$ $x = y^2$

FIGURA 3.4. Representación de las curvas $y = x^2$ y $x = y^2$

Sin embargo, dado un número real $a \geq 0$, existe un único número real $b \geq 0$ tal que $b^2 = a$, usando esta propiedad se define la función *raíz cuadrada* de la siguiente manera: Si $x \geq 0$, $\sqrt{x}$ es el único número real no negativo tal que $(\sqrt{x})^2 = x$.

OBSERVACIÓN 3.7.

- La función $\sqrt{x}$ solamente está definida para $x \geq 0$.
- $\sqrt{x} \geq 0$ si $x \geq 0$ y $\sqrt{x} = 0$ si y solo si $x = 0$.
- Si $x \geq 0$ entonces $(\sqrt{x})^2 = x$.
- $\sqrt{x^2} = |x|$ para todo $x \in \mathbb{R}$.

A continuación el gráfico de la función $\sqrt{x}$

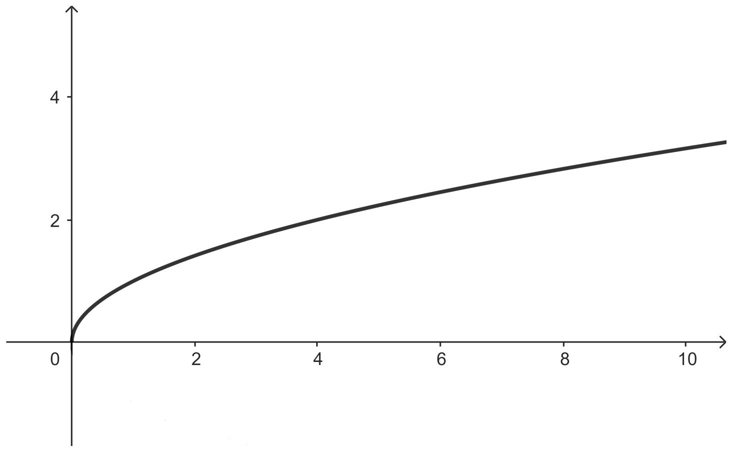
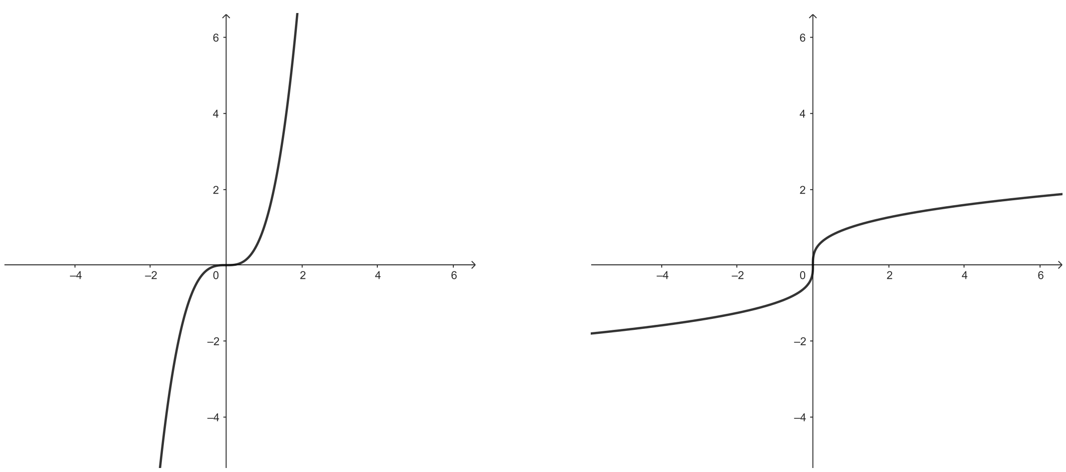


Figura 3.5. Gráfico de la función $\sqrt{x}$

3.3.2. La función raíz cúbica. Como x y x^3 tienen igual signo para todo $x \in \mathbb{R}$, la función $f(x) = x^3$ es una biyección de $\mathbb{R}$ en $\mathbb{R}$, gráficamente



$y = x^3$ $\qquad$ $x = y^3$

Figura 3.6. Representación de las curvas $y = x^3$ y $x = y^3$

La función *raíz cúbica*, $\sqrt[3]{x}$, se define como la inversa de la función $f(x) = x^3$, su gráfico es el de la curva $y = x^3$, ya representada en la Figura 3.6, es decir, dado $x \in \mathbb{R}$, $\sqrt[3]{x}$ es el único número real tal que $(\sqrt[3]{x})^3 = x$.

Observación 3.8.

- La función $\sqrt[3]{x}$ está definida para todo x real.
- $\sqrt[3]{x} > 0$ si $x > 0$, $\sqrt[3]{x} < 0$ si $x < 0$ y $\sqrt[3]{0} = 0$.
- $\sqrt[3]{x^3} = x$ y $(\sqrt[3]{x})^3 = x$ para todo $x \in \mathbb{R}$.

3.3.3. La función raíz n-ésima. Ya que el comportamiento de la función $f(x) = x^n$ es similar al de $g(x) = x^2$ si n es par y al de $h(x) = x^3$ si n es impar, la función *raíz n-ésima*, $\sqrt[n]{x}$, se define de manera análoga a como se definieron $\sqrt{x}$ y $\sqrt[3]{x}$, según n sea par o impar.

Las propiedades de estas funciones y sus gráficos son similares a las de $\sqrt{x}$ o $\sqrt[3]{x}$, según corresponda.

3.3.4. Funciones algebraicas generales. Las *funciones algebraicas* son aquellas que se expresan mediante operaciones algebraicas tales como suma, resta, multiplicación, división, potenciación y extracción de raíces de la variable independiente x. En otras palabras, todas las funciones que se pueden obtener combinando las funciones estudiadas en este capítulo son funciones algebraicas.

EJEMPLO 3.9. Las siguientes fórmulas definen funciones algebraicas:

(1) $\sqrt{x} + x^2 - 2$
(2) $\dfrac{\sqrt[5]{x}}{x^2 - x}$
(3) $|x - \sqrt{x}|$

3.4. Ejercicios varios

(1) En los siguientes problemas use la división larga para determinar el cociente $q(x)$ y el residuo $r(x)$ cuando el polinomio $f(x)$ se divide entre el polinomio indicado $d(x)$. En cada caso, escriba la respuesta en la forma $f(x) = d(x)q(x) + r(x)$ y verifique que se cumple la igualdad.
   (a) $f(x) = 8x^2 + 4x - 7; \quad d(x) = x^2$.
   (b) $f(x) = x^2 + 2x - 3; \quad d(x) = x^2 + 1$.
   (c) $f(x) = 5x^3 - 7x^2 + 4x + 1; \quad d(x) = x^2 + x - 1$.
   (d) $f(x) = 14x^3 - 12x^2 + 6; \quad d(x) = x^2 - 1$.
   (e) $f(x) = 2x^3 + 4x^2 - 3x + 5; \quad d(x) = (x + 2)^2$.
   (f) $f(x) = x^3 + x^2 + x + 1; \quad d(x) = (2x + 1)^2$.
   (g) $f(x) = 27x^3 + x - 2; \quad d(x) = 3x^2 - x$.
   (h) $f(x) = x^4 + 8; \quad d(x) = x^3 + 2x - 1$.
   (i) $f(x) = 6x^5 + 4x^4 + x^3; \quad d(x) = x^3 - 2$.
   (j) $f(x) = 5x^6 - x^5 + 10x^4 + 3x^2 - 2x + 4; \quad d(x) = x^2 + x - 1$.

(2) Aplique el resultado enunciado en la página 48 para determinar si $f(x)$ es divisible entre el polinomio lineal indicado.
   (a) $f(x) = 2x^2 - 4x + 6; \quad x - 2$.
   (b) $f(x) = 3x^2 + 7x - 1; \quad x + 3$.
   (c) $f(x) = x^3 - 4x^2 + 5x + 2; \quad x - \frac{1}{2}$.
   (d) $f(x) = 5x^3 + x^2 - 4x - 6; \quad x + 1$.
   (e) $f(x) = x^4 - x^3 + 2x^2 + 3x - 5; \quad x - 3$.

(f) $f(x) = 2x^4 - 7x^2 + x - 1; \quad x + \frac{3}{2}$.

(3) Use la división sintética para calcular el cociente $q(x)$ y el residuo $r(x)$ cuando se divide $f(x)$ entre el polinomio lineal indicado.

(a) $f(x) = 2x^2 - x + 5; \quad x - 2$.

(b) $f(x) = 4x^2 - 8x + 6; \quad x - \frac{1}{2}$.

(c) $f(x) = x^3 - x^2 + 2; \quad x + 3$.

(d) $f(x) = 4x^3 - 3x^2 + 2x + 4; \quad x - 7$.

(e) $f(x) = x^4 + 16; \quad x - 2$.

(f) $f(x) = 4x^4 + 3x^3 - x^2 - 5x - 6; \quad x + 3$

(g) $f(x) = x^5 + 56x^2 - 4; \quad x + 4$

(4) Calcular el valor de k tal que $f(x)$ sea divisible entre $d(x)$.

(a) $f(x) = kx^4 + 2x^2 + 9k; \quad d(x) = x - 1$.

(b) $f(x) = x^3 + kx^2 - 2kx + 4; \quad d(x) = x + 2$.

(5) Determine si el número real indicado es una raíz de la función polinomial f. En caso de serlo, determine todas las demás raíces y a continuación presente la factorización completa de $f(x)$.

(a) 1; $f(x) = 4x^3 - 9x^2 + 6x - 1$.

(b) $\frac{1}{2}$; $f(x) = 2x^3 - x^2 + 32x - 16$.

(c) 5; $f(x) = x^3 - 6x^2 + 6x + 5$.

(d) 3; $f(x) = x^3 - 3x^2 + 4x - 12$.

(e) $-\frac{2}{3}$; $f(x) = 3x^3 - 10x^2 - 2x + 4$.

(f) -2; $f(x) = x^3 - 4x^2 - 2x + 20$.

Capítulo 4

Vectores en el plano

Muchas magnitudes que pueden ser medidas, tales como la longitud, el área, el volumen, la masa y la temperatura, pueden describirse completamente al especificar su valor en una unidad específica como metros, metros cuadrados, centímetros cúbicos, gramos, grados centígrados, respectivamente; se les conoce con el nombre de *magnitudes escalares*.

Otras magnitudes como la velocidad, la fuerza y la aceleración, requieren tanto de un valor numérico en una unidad adecuada como de una dirección y sentido para su descripción; a éstas se les conoce como *magnitudes vectoriales* o *vectores*. Por ejemplo, la velocidad del viento es un vector que consta de rapidez del viento y su sentido, como 10 km/h suroeste. Es usual representar a los vectores como flechas o segmentos de recta dirigidos.

4.1. Conceptos básicos

Tal como ya se indicó, un *vector* es una magnitud que tiene asociada una dirección, un sentido y un valor numérico, como por ejemplo la fuerza, la velocidad, la aceleración y el desplazamiento.

Consideremos el plano cartesiano con sus ejes x e y. Un vector en el plano es un segmento de recta que va dirigido desde un punto A hasta un punto B, al punto A se le conoce como punto inicial y al punto B como punto final; a este vector se le suele denotar con el símbolo $\overrightarrow{AB}$.

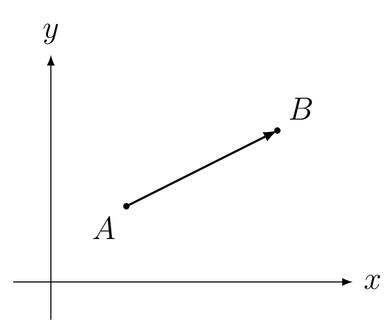


FIGURA 4.1. Representación gráfica del vector $\overrightarrow{AB}$

Con más precisión, sea $\overrightarrow{AB}$ un vector en el plano, su *dirección* es la recta sobre la cual se ubica, su *longitud* es la longitud del segmento $\overline{AB}$ y su *sentido* es el que está determinado por el punto inicial y el punto final.

El conjunto de todos los puntos del plano está en correspondencia natural con el conjunto de todos los vectores cuyo punto inicial es el origen de coordenadas O, a cada punto P del plano le corresponde el vector $\overrightarrow{OP}$, por ejemplo si $P = (x_o, y_o)$ el vector $\overrightarrow{OP}$ corresponde con el punto (x_o, y_o).

La primera parte de la siguiente figura ilustra la correspondencia entre el vector $\overrightarrow{v_o} = \overrightarrow{0P}$ y el punto $P = (x_o, y_o)$ y la segunda parte ilustra la correspondencia entre los vectores $\overrightarrow{OA}$, $\overrightarrow{OB}$ y $\overrightarrow{OC}$ y los puntos A, B y C, respectivamente.

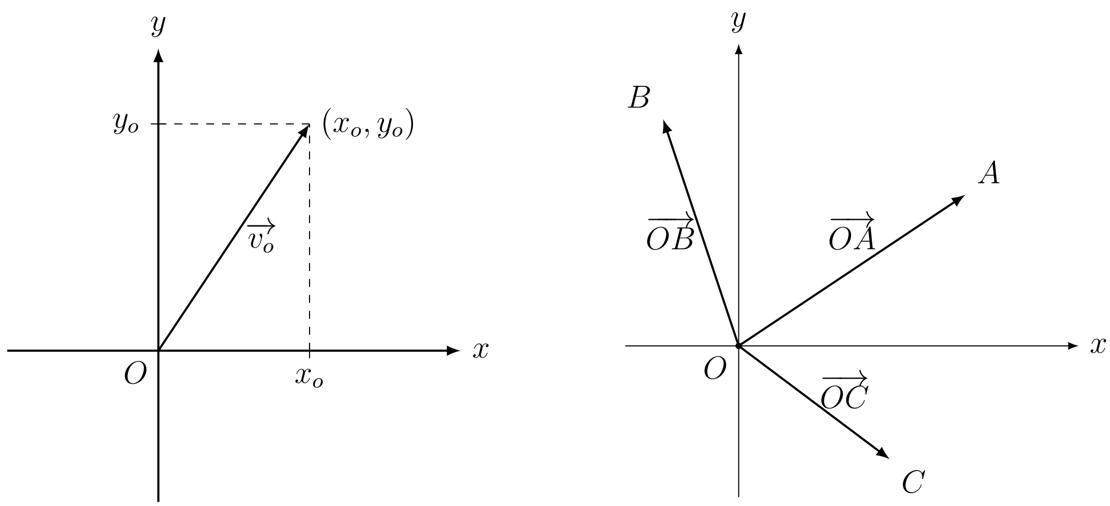


FIGURA 4.2. Correspondencia entre vectores y puntos del plano

Existen dos clases de vectores, los libres y los ligados. Un *vector libre* es una clase de equivalencia de vectores que tienen direcciones paralelas y coinciden en longitud y sentido, sin importar su punto inicial, es decir, se consideran iguales a dos vectores cuyas direcciones son paralelas y que coinciden en longitud y sentido. Geométricamente, dos vectores libres son iguales si uno puede obtenerse al deslizar (o trasladar) el otro, de manera paralela a sí mismo hasta que los dos vectores coincidan. El considerar vectores libres simplifica los cálculos y lo usual es trabajar con el vector libre cuyo punto inicial es el origen de coordenadas O.

Los vectores que no son libres son llamados *vectores ligados*.

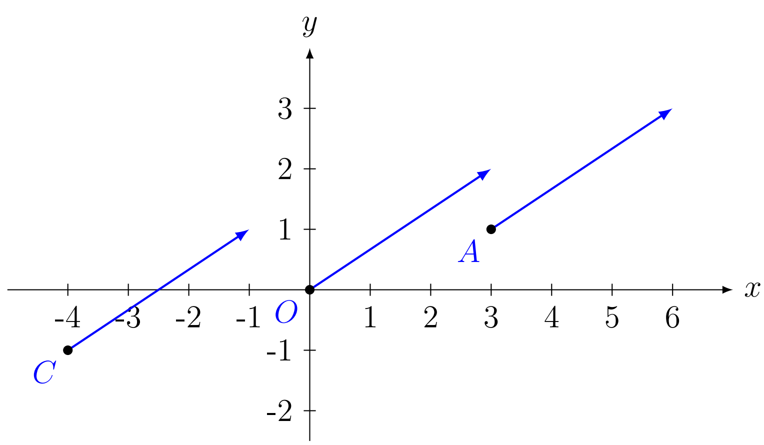


FIGURA 4.3. Vectores libres equivalentes

En álgebra lineal y geometría analítica lo usual es trabajar con vectores libres, a menos que el punto de aplicación sea esencial para el problema. Un vector libre representa una dirección y una magnitud , pero no depende de un punto específico en el espacio, por ejemplo, el vector $(2, 3)$ en un sistema de coordenadas indica un desplazamiento de 2 unidades en la dirección x y 3 unidades en y, considerando el vector como libre, este desplazamiento puede ocurrir desde cualquier punto del plano. Si el vector estuviera atado a un punto fijo, perdería su capacidad de expresar traslaciones generales.

Los vectores libres y el álgebra lineal son utilizados en la elaboración de juegos de video, cualquier movimiento que se hace en un juego es asociado a un vector libre, si este movimiento está asociado a una fuerza (como un golpe) es de importancia la magnitud del vector.

En lo que sigue, siempre se considerarán vectores libres.

Los vectores pueden ser sumados, restados y multiplicados por un número real. En el contexto de los vectores, a los números reales se les conoce como *escalares*. .

Es usual denotar a los vectores con una letra con una flecha en la parte superior y al especificar las coordenadas se usan las notaciones de vector columna o vector fila según convenga, a veces con paréntesis y a veces con corchetes, por ejemplo el vector $\vec{v} = (a, b)$ puede ser denotado como

$$(a \;\; b), \qquad [a \;\; b], \qquad \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{ó} \quad \begin{bmatrix} a \\ b \end{bmatrix}.$$

Sean $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$, $\vec{w} = (x_3, y_3)$ vectores del plano, $\vec{0} = (0, 0)$ el vector nulo y α, β escalares, la suma y el producto por un escalar se hacen de la siguiente manera

$$\vec{u} + \vec{v} = (x_1 + x_2, y_1 + y_2) \qquad \alpha\vec{u} = (\alpha x_1, \alpha y_1),$$

el vector $-\vec{u}$ se define como $-\vec{u} = (-1)\vec{u} = (-x_1, -x_2)$.

Estas operaciones tienen las siguientes propiedades

PROPOSICIÓN 4.1.

$\vec{u} + \vec{v} = \vec{v} + \vec{u}$	$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$	$\vec{v} + \vec{0} = \vec{v}$
$\vec{u} + (-\vec{u}) = \vec{0}$	$\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v}$	$(\alpha + \beta)\vec{u} = \alpha\vec{u} + \beta\vec{u}$
$(\alpha\beta)\vec{u} = \alpha(\beta\vec{u}) = (\alpha\beta)\vec{u}$	$1\vec{u} = \vec{u}$	$0\vec{u} = \vec{0}$

Es importante destacar que los símbolos $+$ y $-$ se usan para dos fines diferentes: suma y resta de escalares, suma y resta de vectores. Esto es usual en el contexto de vectores, por lo que se debe tener muy claro la clase de elementos que involucra cada operación.

Cuando multiplicamos un vector por un escalar α su longitud se modifica, quedando multiplicada por $|\alpha|$; si $\alpha > 0$ se mantiene el sentido y si $\alpha < 0$ cambia el sentido.

Geométricamente la suma de vectores se hace de acuerdo a la ley del paralelogramo, que se detalla en la siguiente figura.

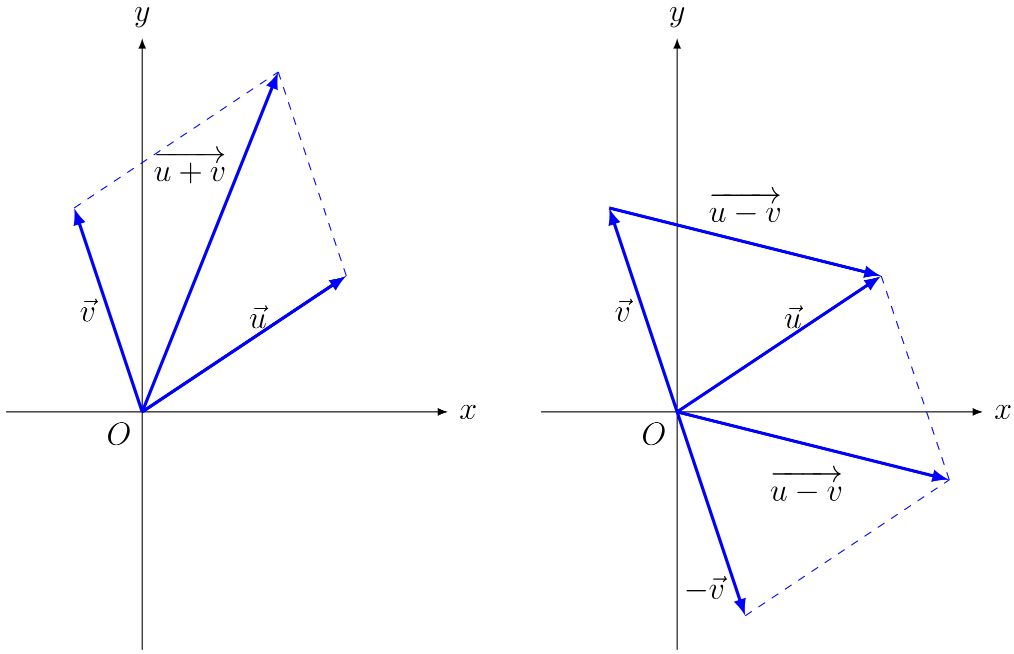


Figura 4.4. Ley del paralelogramo

4.2. Ecuación paramétrica de la recta

Consideremos un vector no nulo $\vec{u} = (u_1, u_2)$ y un punto (o vector) $\vec{p} = (x_o, y_o)$ en el plano $\mathbb{R}^2$.

El conjunto de los vectores de la forma $t\vec{u}$ con $t \in \mathbb{R}$ corresponde con la recta que pasa por el origen y que tiene la dirección del vector $\vec{u}$, por lo tanto, la ecuación *paramétrica o vectorial de la recta* que pasa por el origen en la dirección del vector $\vec{u}$ es

$$\vec{v}(t) = t\vec{u}, \qquad -\infty < t < +\infty.$$

Como los vectores se suman de acuerdo a la ley del paralelogramo, si ahora consideramos los vectores de la forma $\vec{p} + t\vec{u}$ con $t \in \mathbb{R}$, obtenemos la recta que para por el punto $\vec{p}$ y que es paralela al vector $\vec{u}$ (ver Figura 4.5), por lo tanto, la ecuación *paramétrica o vectorial de la recta* que pasa por el punto $\vec{p}$ en la dirección del vector $\vec{u}$ es

$$\vec{v}(t) = \vec{p} + t\vec{u}, \qquad -\infty < t < +\infty. \tag{4.1}$$

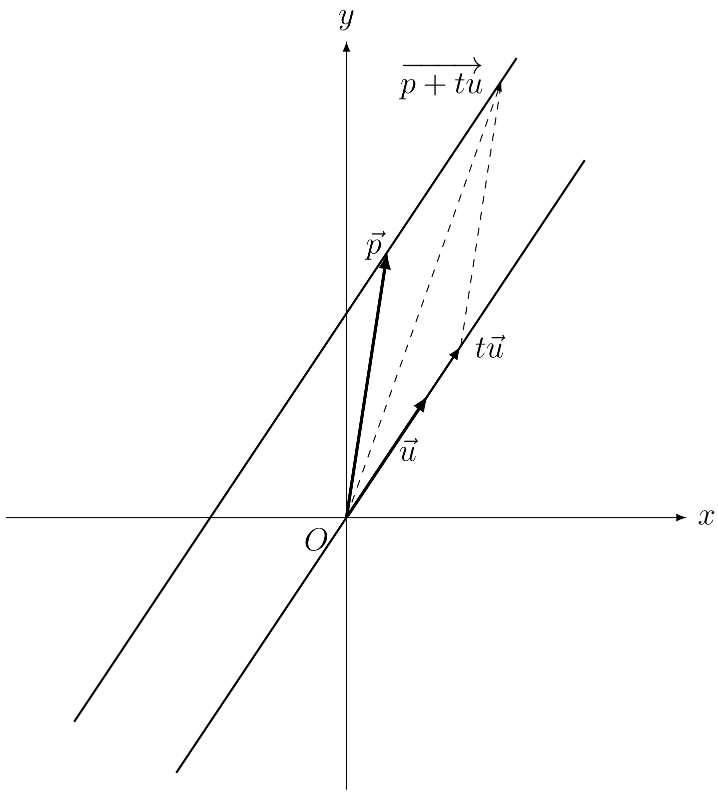


FIGURA 4.5. Ecuación paramétrica de la recta

Si consideramos las coordenadas $(x(t), y(t))$ del vector $\vec{v}(t)$ obtenemos esta otra forma de la ecuación (4.1)

$$\begin{cases} x(t) = x_o + tu_1 \\ y(t) = y_o + tu_2 \end{cases} \qquad -\infty < t < +\infty.$$

Si suponemos $u_1 \neq 0$ y despejamos t en la primera ecuación obtenemos la ecuación *punto-pendiente* de la recta, que tiene la forma

$$y = mx + b \quad \text{donde } m \text{ y } b \text{ son números reales,}$$

si $u_1 = 0$ se trata de una recta vertical, cuya ecuación cartesiana tiene la forma

$$x = c \text{ donde } c \text{ es una constante.}$$

Por esto, la *ecuación cartesiana o general* de la recta en el plano es

$$Ax + By + C = 0 \text{ donde } A,\ B \text{ y } C \text{ son constantes.}$$

Recordemos que dos puntos diferentes del plano (x_o, y_o) y (x_1, y_1) determinan una recta, como vector director de la recta determinada por este pare de puntos se puede escoger

$$\vec{u} = (x_1, y_1) - (x_o, y_o) = (x_1 - x_o, y_1 - y_o),$$

de manera que la ecuación paramétrica nos queda

$$\begin{cases} x(t) = x_o + t(x_1 - x_o) \\ y(t) = y_o + t(y_1 - y_o) \end{cases} \qquad -\infty < t < +\infty,$$

despejando t de ambas ecuaciones e igualando obtenemos la ecuación de la recta que pasa por los puntos (x_o, y_o) y (x_1, y_1)

$$\frac{x - x_o}{x_1 - x_o} = \frac{y - y_o}{y_1 - y_o},$$

válida siempre que $x_1 \neq x_o$ y $y_1 \neq y_o$.

4.3. Norma y producto escalar

La *norma* de un vector $\vec{u} = (x, y)$, que se denota por $\|\vec{u}\|$ es igual a su longitud, como su longitud es igual a la distancia del punto (x, y) al origen O, se tiene que

$$\|\vec{u}\| = \sqrt{x^2 + y^2}.$$

La norma de los vectores satisface las siguientes propiedades:
Sean $\vec{u}$ y $\vec{v}$ vectores en el plano y α un número real, entonces

(1) $\|\vec{u}\| \geq 0$, $\|\vec{u}\| = 0$ si y solo si $\vec{u} = 0$,
(2) $\|\overrightarrow{\alpha\, u}\| = |\alpha|\, \|\vec{u}\|$,
(3) (Desigualdad triangular) $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$.

A manera de ejercicio, interpretar geométricamente cada una de las propiedades anteriores.

El *producto escalar* o *producto interno* de dos vectores es igual al producto de sus longitudes multiplicada por el coseno del ángulo que forman. El producto interno de dos vectores $\vec{u}$ y $\vec{v}$ se denota por $\langle \vec{u}, \vec{v} \rangle$ (algunos autores usan la notación $\vec{u} \cdot \vec{v}$); en la Figura 4.6 los vectores $\vec{u}$ y $\vec{v}$ forman un ángulo γ, por lo que

$$\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\|\, \|\vec{v}\|\, \cos\gamma,$$

y se observa que $\|\vec{u}\| \cos\gamma$ corresponde con la longitud de la proyección de $\vec{u}$ en la dirección de $\vec{v}$.

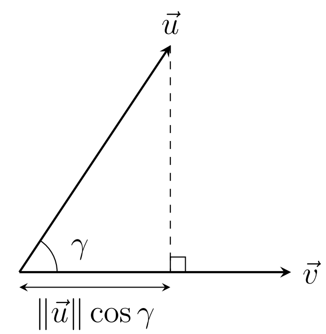


Figura 4.6. Producto escalar de vectores

Es importante tener en cuenta la interpretación geométrica del producto escalar:

(1) El valor absoluto del producto escalar $\langle \vec{u}, \vec{v} \rangle$ es igual a la longitud de la proyección del vector $\vec{u}$ en la dirección del vector $\vec{v}$, multiplicado por la longitud del vector $\vec{v}$.
(2) Si α es el ángulo entre los dos vectores y los vectores no son nulos, el producto escalar será positivo si $0 \leq \alpha < \frac{\pi}{2}$, negativo si $\frac{\pi}{2} < \alpha \leq \pi$ y nulo si $\alpha = \frac{\pi}{2}$.
(3) Si la longitud del vector $\vec{v}$ es igual a 1, el valor absoluto del producto escalar $\langle \vec{u}, \vec{v} \rangle$ es igual a la longitud de la proyección del vector $\vec{u}$ en la dirección del vector $\vec{v}$.
(4) Si dos vectores son ortogonales su producto escalar es igual a 0.
(5) $\langle \vec{u}, \vec{u} \rangle = \|\vec{u}\|^2$, o lo que es equivalente $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}$.

En coordenadas se tiene los siguiente, si $\vec{u} = (u_1, u_2)$ y $\vec{v} = (v_1, v_2)$, entonces

$$\langle \vec{u}, \vec{v} \rangle = u_1\, v_1 + u_2\, v_2.$$

El producto escalar de vectores satisface las siguientes propiedades:
Sean $\vec{u}$, $\vec{v}$, $\vec{w}$ vectores en el plano y α un número real, entonces

(1) $\langle \vec{u}, \vec{u} \rangle \geq 0$ y es igual a 0 si y solo si $\vec{u} = 0$,
(2) $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$,
(3) $\langle \alpha\, \vec{u}, \vec{v} \rangle = \alpha\, \langle \vec{u}, \vec{v} \rangle$,
(4) $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$.

También se cumple la llamada desigualdad de Cauchy-Schwarz

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|.$$

4.4. Ejercicios varios

(1) Sean $\vec{u} = (2, 3)$ y $\vec{v} = (5, 4)$. Hallar $\vec{u} + \vec{v}$, $\vec{u} - \vec{v}$ y $\vec{v} - \vec{u}$. Representar gráficamente (se sugiere usar GeoGebra) y verificar que se cumple la ley del paralelogramo. Escoger otros pares de vectores combinado llos diferentes cuadrantes y repetir el ejercicio.
(2) Encontrar la magnitud y el ángulo que forma con el eje x para cada uno de los siguientes vectores

(a) $(-4, 4)$	(b) $(\sqrt{3}, 2)$	(c) $(2, \sqrt{3})$
(d) $(-\sqrt{3}, -2)$	(e) $(\sqrt{3}, -2)$	(f) $(3, -8)$
(g) $(10, 0)$	(h) $(1, 1)$	(i) $(-5, -5)$

(3) Determinar el ángulo que forman cada par de vectores

(a) $\vec{u} = (0, 1)$, $\vec{v} = (0, -1)$	(b) $\vec{u} = (2, 2)$, $\vec{v} = (4, -5)$
(c) $\vec{u} = (2, -1)$, $\vec{v} = (3, 6)$	(d) $\vec{u} = (0, 1)$, $\vec{v} = (0, 8)$

(4) Determina la longitud de la proyección del vector $4, 5)$ sobre la dirección del vector $(1, 1)$.
(5) Encontrar un vector que tenga la magnitud dada y que forme el ángulo indicado con el eje x.

(a) $\|\vec{v}\| = 1$, $\alpha = \frac{\pi}{4}$
(b) $\|\vec{v}\| = 3$, $\alpha = \frac{-\pi}{3}$
(c) $\|\vec{v}\| = \sqrt{3}$, $\alpha = \frac{\pi}{6}$
(d) $\|\vec{v}\| = \sqrt{2}$, $\alpha = \frac{3\pi}{4}$
(e) $\|\vec{v}\| = 6$, $\alpha = \frac{7\pi}{6}$
(f) $\|\vec{v}\| = 1$, $\alpha = \frac{-\pi}{6}$

(6) ¿Cuáles de los siguientes vectores $\vec{u}_1 = (1,2)$, $\vec{u}_2 = (0,1)$, $\vec{u}_3 = (-2,-4)$, $\vec{u}_4 = (2,4)$, $\vec{u}_5 = (-2,-1)$,$\vec{u}_6 = (-6,3)$, $\vec{u}_7 = (-3,-6)$
  (a) son ortogonales?
  (b) tienen igual dirección y sentido?
  (c) tienen igual dirección y sentidos opuestos?

(7) Determinar cuáles de los siguientes vectores ligados son equivalentes
  (a) $\overrightarrow{A_1B_1}$, donde $A_1 = (1,1)$, $B_1 = (5,5)$.
  (b) $\overrightarrow{A_2B_2}$, donde $A_1 = (3,3)$, $B_1 = (7,7)$.
  (c) $\overrightarrow{A_3B_3}$, donde $A_1 = (0,0)$, $B_1 = (4,4)$.
  (d) $\overrightarrow{A_4B_4}$, donde $A_1 = (-1,-2)$, $B_1 = (1,1)$.
  (e) $\overrightarrow{A_3B_3}$, donde $A_1 = (2,1)$, $B_1 = (4,4)$.

(8) Hallar la ecuación cartesiana y paramétrica de la recta que pasa por los siguientes puntos
  (a) $(0,1)$ y $(2,3)$.
  (b) $(3,1)$ y $(3,-2)$.
  (c) $(-1,1)$ y $(-1,-3)$.
  (d) $(4,5)$ y $(-3,-2)$.

(9) Determinar si los siguientes puntos está sobre una recta
  (a) $(0,2)$, $(2,8)$, $(\frac{1}{2},\frac{7}{2})$.
  (b) $(\frac{1}{2},\frac{7}{2})$, $(3,\frac{1}{2})$, $(0,3)$.

(10) Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^2$ que pasa a por el punto $P = (2,-1)$ y es paralela a la recta con ecuación general $2x - 3y = 1$.

(11) Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^2$ que pasa por el punto $P = (2,-1)$ y es perpendicular a la recta con ecuación general $2x - 3y = 1$.

(12) Sugiera una "justificación vectorial" del hecho de que, en $\mathbb{R}^2$, dos rectas con pendientes m_1 y m_2 son perpendiculares si y sólo si $m_1m_2 = -1$.

Capítulo 5

Vectores en el espacio

5.1. Representación de puntos y vectores en el espacio

Consideremos el espacio tridimensional

$$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) : x, y, z \in \mathbb{R}\}.$$

Al igual que $\mathbb{R}^2$ se identifica con el plano, $\mathbb{R}^3$ se identifica con el espacio ambiente. Para establecer la correspondencia debemos considerar un eje adicional, usualmente llamado eje z, perpendicular al plano formado por el eje x y el eje y. Cada punto P del espacio está en correspondencia con un elemento (x, y, z) de $\mathbb{R}^3$.

La siguiente figura ilustra esta correspondencia, en el mismo se observa, de manera gráfica, como el punto P del espacio corresponde con la terna (x_o, y_o, z_o).

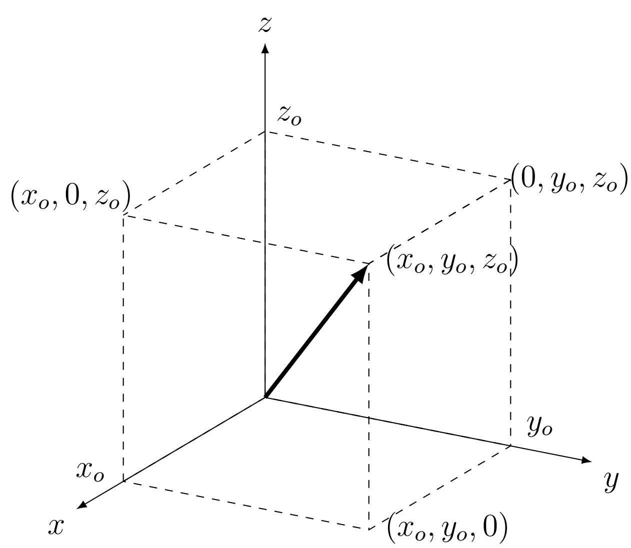


Figura 5.1. Puntos y vectores en el espacio

Al igual que en el plano, al punto $(0, 0, 0)$ se le suele llamar el *origen de coordenadas*, o simplemente, el *origen*. Existen tres planos que se destacan en este espacio, que son: el plano "xy", el plano "yz" y el plano "xz".

Al igual que en el caso bidimensional, existe una identificación natural entre los puntos de $\mathbb{R}^3$ y los vectores en el espacio: Al punto (x, y, z) le hacemos corresponder el vector de extremo inicial el origen y de extremo final el punto (x, y, z). El origen de coordenadas se identifica con el vector $(0, 0, 0)$.

Cuando $x > 0$, $y > 0$, $z > 0$ decimos que el punto (o el vector) (x, y, z) se encuentra en el primer octante.

La suma de vectores y el producto por un escalar se definen de manera natural:

Si $\vec{u} = (x_1, y_1, z_1), \vec{v} = (x_2, y_2, z_2) \in \mathbb{R}^3$ y $\lambda \in \mathbb{R}$,

$$\vec{u} + \vec{v} = (x_1 + x_2, y_1 + y_2, z_1 + z_2),$$

$$\lambda\vec{u} = (\lambda x_1, \lambda y_1, \lambda z_1).$$

Si $\lambda > 0$ entonces $\lambda\vec{u}$ y $\vec{u}$ tienen el mismo sentido. Si $\lambda < 0$ entonces $\lambda\vec{u}$ y $\vec{u}$ tienen sentido contrario.

Se dice que dos vectores no nulos $\vec{u}$ y $\vec{v}$ son *paralelos* cuando existe λ tal que

$$\vec{v} = \lambda\vec{u}.$$

También en el caso tridimensional, la suma y diferencia de vectores se puede hacer, de manera geométrica, siguiendo la ley del paralelogramo.

5.2. Distancia entre dos puntos del espacio y norma de vectores

Supongamos que queremos hallar la distancia d entre dos puntos $(x_1, y_1, z_1), (x_2, y_2, z_2)$ del espacio.

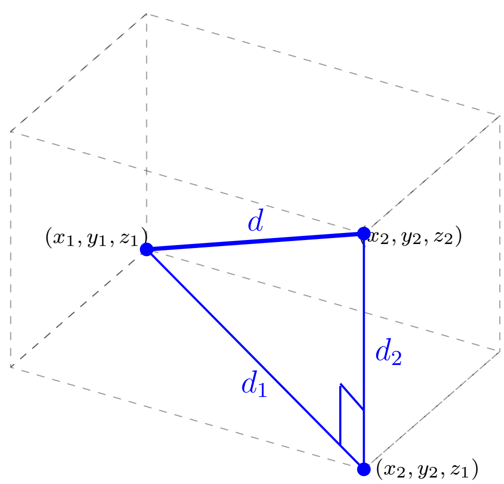


Figura 5.2. Distancia entre dos puntos de $\mathbb{R}^3$

Sea d_1 la distancia entre los puntos (x_1, y_1, z_1) y (x_2, y_2, z_1), como estos puntos está en el plano paralelo al plano xy, que corta al eje z en $z = z_1$, por la fórmula de la distancia en el plano tenemos que

$$d_1^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2.$$

Si d_2 es la diferencia de alturas entre los puntos (x_1, y_1, z_1) y (x_2, y_2, z_2) entonces

$$d_2 = |z_2 - z_1|,$$

Analizando la figura anterior y usando el teorema de Pitágoras obtenemos que

$$d^2 = d_1^2 + d_2^2,$$

es decir,

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

Al igual que en el caso bidimensional, dado un vector $\vec{u} = (x, y, z) \in \mathbb{R}^3$, definimos la *norma* de $\vec{u}$ como

$$\|\vec{u}\| = \sqrt{x^2 + y^2 + z^2}.$$

Se tiene que $\|\vec{u}\|$ es la distancia del punto (x, y, z) al origen, es decir, la longitud del vector $\vec{u}$.

Al igual que en el caso bidimensional, la norma de los vectores satisface las siguientes propiedades:

Sean $\vec{u}$ y $\vec{v}$ vectores en el espacio y α un número real, entonces

(1) $\|\vec{u}\| \geq 0$, $\|\vec{u}\| = 0$ si y solo si $\vec{u} = 0$,
(2) $\|\overrightarrow{\alpha\, u}\| = |\alpha|\, \|\vec{u}\|$,
(3) (Desigualdad triangular) $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$.

Decimos que $\vec{u} \in \mathbb{R}^n$ es unitario si $\|\vec{u}\| = 1$.

Dado $\vec{u} \in \mathbb{R}^n$ no nulo, si consideramos

$$\vec{v} = \frac{\vec{u}}{\|\vec{u}\|},$$

obtenemos que $\vec{v}$ es unitario y que tiene la misma dirección que $\vec{u}$.

Ejercicios.

Describir los subconjuntos de $\mathbb{R}^3$ que satisfacen cada una de las siguientes ecuaciones y usar GeoGebra para obtener una representación gráfica.

(1) $x^2 + y^2 + z^2 = 4$
(2) $x^2 - 2x + 6 + y^2 - 4y + z^2 + 2z \leq 9$

5.3. Producto escalar

Sean $\vec{u} = (x_1, y_1, z_1), \vec{v} = (x_2, y_2, z_2) \in \mathbb{R}^3$. El *producto escalar* de estos vectores se define de manera similar al caso bidimensional, es decir, es igual al producto de las longitudes de los vectores multiplicado por el coseno del ángulo que forman y, en coordenadas se cumple

$$\langle (x_1, y_1, z_1), (x_2, y_2, z_2) \rangle = x_1x_2 + y_1y_2 + z_1z_2,$$

em forma abreviada,

$$\langle \vec{u}, \vec{v} \rangle = x_1x_2 + y_1y_2 + z_1z_2.$$

Es muy importante notar que, tanto en el caso bidimensional como en el caso tridimensional, el producto escalar siempre es igual a la suma del producto de las coordenadas.

Al igual que en el caso de $\mathbb{R}^2$ el producto interno tiene las siguientes propiedades.

Sean $\vec{u}$, $\vec{v}$, $\vec{w}$ vectores en el espacio y α un número real, entonces

(1) $\langle \vec{u}, \vec{u} \rangle \geq 0$ y es igual a 0 si y solo si $\vec{u} = 0$,
(2) $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$,
(3) $\langle \alpha\, \vec{u}, \vec{v} \rangle = \alpha\, \langle \vec{u}, \vec{v} \rangle$,
(4) $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$.

Nuevamente, los vectores $\vec{u}$ y $\vec{v}$ son *perpendiculares* u *ortogonales* si $\langle \vec{u}, \vec{v} \rangle = 0$

5.4. Ecuación de la recta en el espacio.

Al igual que en el plano $\mathbb{R}^2$ una *recta* en $\mathbb{R}^3$ está determinada por un punto $\vec{p}_o$ y una dirección $\vec{u}$ (donde $\vec{u}$ es un vector no nulo). El vector $\vec{u}$ es llamado el *vector director* de la recta.

Los puntos $\vec{p}$ sobre la L que pasa por $\vec{p}_o$ en la dirección de $\vec{u}$ son todos los puntos de la forma

$$\vec{p} = \vec{p}_o + t\, \vec{u},$$

donde $t \in \mathbb{R}$. Esta ecuación se llama *ecuación vectorial de la recta*.

Si $\vec{p}_o = (x_o, y_o, z_o)$, $\vec{u} = (u_1, u_2, u_3)$ y $\vec{p} = (x, y, z)$ tenemos

$$(x, y, z) = (x_o, y_o, z_o) + t(u_1, u_2, u_3).$$

Luego

$$x = x_o + tu_1, \quad y = y_o + tu_2, \quad z = z_o + tu_3,$$

estas últimas son las ecuaciones correspondientes entre las componentes.

Si $u_1 \neq 0$, $u_2 \neq 0$ y $u_3 \neq 0$ se puede eliminar t y la ecuación se expresa en su forma cartesiana

$$\frac{x - x_o}{u_1} = \frac{y - y_o}{u_2} = \frac{z - z_o}{u_3}$$

Al igual que en el plano, una recta está determinada si damos dos puntos por los que pasa. Supongamos que L es una recta que pasa por los puntos (diferentes) $\vec{p}_o = (x_o, y_o, z_o)$ y $\vec{p}_1 = (x_1, y_1, z_1)$.

Sea $\vec{u} = \vec{p}_1 - \vec{p}_o = (x_1 - x_o, y_1 - y_o, z_1 - z_o)$. Entonces L es la recta de dirección $\vec{u}$ que pasa por cualquiera de los puntos $\vec{p}_o = (x_o, y_o, z_o)$ ó $\vec{p}_1 = (x_1, y_1, z_1)$. Por lo tanto la ecuación de L es:

$$\frac{x - x_o}{x_1 - x_o} = \frac{y - y_o}{y_1 - y_o} = \frac{z - z_o}{z_1 - z_o}$$

Dos rectas son *perpendiculares* si sus vectores directores lo son.

Dos rectas son *paralelas* si sus vectores directores son paralelos.

En $\mathbb{R}^3$, si dos rectas son paralelas, entonces son iguales o no se intersectan.

En $\mathbb{R}^3$, si dos rectas no son paralelas, entonces no se cortan o su intersección es un punto.

Ejercicios.

(1) Considere la ecuación vectorial $\vec{x} = \vec{p} + t(\vec{q} - \vec{p})$, donde $\vec{p}$ y $\vec{q}$ corresponden a distintos puntos P y Q en $\mathbb{R}^2$ o $\mathbb{R}^3$.
   (a) Explique porqué esta ecuación describe el segmento de recta PQ cuando t varía de 0 a 1.
   (b) Para cuál valor de t, $\vec{x}$ es el punto medio de PQ, y cuál es $\vec{x}$ en este caso?
   (c) Encuentre el punto medio de PQ cuando $P = (2, -3)$ y $Q = (0, 1)$.
   (d) Encuentre el punto medio de PQ cuando $P = (1, 0, 1)$ y $Q = (4, 1, -2)$.
   (e) Encuentre los dos puntos que dividen PQ del inciso (c) en tres partes iguales.
   (f) Encuentre los dos puntos que dividen PQ del inciso (d) en tres partes iguales.

(2) Encontrar la ecuación vectorial y cartesiana de la recta que pasa por los puntos $P = (0, 1, -1)$ y $Q = (-2, 1, 3)$.

5.5. Ecuación del plano

Una manera de determinar un plano es indicar un punto por donde pasa el plano y un vector perpendicular a él.

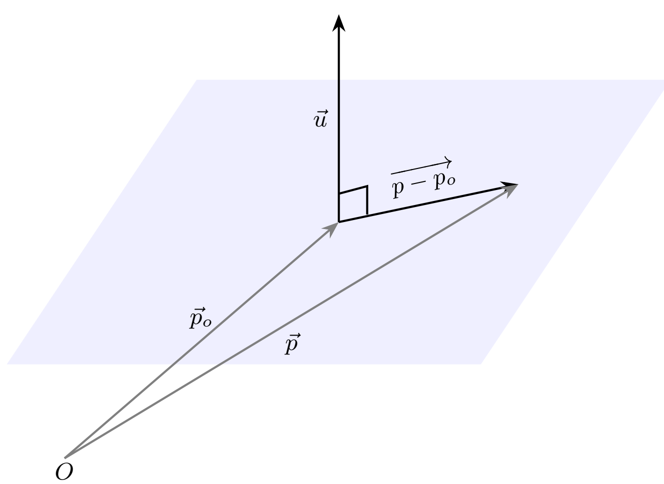


Figura 5.3. Ecuación del plano

Sea $\vec{p}_o = (x_o, y_o, z_o)$un punto del plano y $\vec{u} = (a, b, c)$ un vector perpendicular al plano.

Si $\vec{p} = (x, y, z)$ es otro punto del plano entonces $\vec{p} - \vec{p}_o = (x, y, z) - (x_o, y_o, z_o)$ es perpendicular a $\vec{u} = (a, b, c)$, es decir,

$$\langle (a, b, c), (x, y, z) - (x_o, y_o, z_o) \rangle = 0.$$

Por lo tanto

$$a(x - x_o) + b(y - y_o) + c(z - z_o) = 0.$$

Y así

$$ax + by + cz + d = 0$$

donde $d = -ax_o - by_o - cz_o$. Esta ecuación se llama *ecuación cartesiana del plano* (notar que el vector (a, b, c) es perpendicular al plano).

Un plano también está determinado por dos rectas no paralelas que se cortan.

Sean L_1 y L_2 dos rectas no paralelas de direcciones respectivas $\vec{u}$ y $\vec{v}$ que se cortan en un punto $\vec{p}_o$. Los puntos $\vec{p}$ sobre el plano determinado por L_1 y L_2 son todos los puntos de la forma

$$\vec{p} = \vec{p}_o + t\vec{u} + s\vec{v},$$

donde $t, s \in \mathbb{R}$.

Esta ecuación se llama también *ecuación vectorial del plano* y las ecuaciones correspondientes entre las componentes se llaman las *ecuaciones paramétricas del plano*, éstas son:

$$x = x_o + tu_1 + sv_1, \qquad y = y_o + tu_2 + sv_2, \qquad z = z_o + tu_3 + sv_3.$$

Un plano está determinado si damos tres puntos no alineados por los que pasa el plano.

Sean $\vec{p}_o, \vec{p}_1, \vec{p}_2$ tres puntos diferentes y no alineados por los que pasa el plano. Sean $\vec{u} = \vec{p}_1 - \vec{p}_o$ y $\vec{v} = \vec{p}_2 - \vec{p}_o$. Sean L_1 y L_2 dos rectas de direcciones respectivas $\vec{u}$ y $\vec{v}$ que pasan por $\vec{p}_o$. Entonces L_1 y L_2 se cortan en el punto $\vec{p}_o$. Estas dos rectas no paralelas que se cortan, determinan un plano.

5.6. Ejercicios varios

(1) Hallar la ecuación de la recta que pasa por P con vector director $\vec{d}$ en forma cartesiana y forma paramétrica.

(a) $P = (1, 0, 1), \vec{d} = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}$.

(b) $P = (0, 0, 0), \vec{d} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$.

(c) $P = (3, 0, -2), \vec{d} = \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix}$.

(2) Hallar la ecuación cartesiana del plano que pasa por P con vector normal $\vec{n}$.

(a) $P = (0, 1, 0)$, $\vec{n} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$.

(b) $P = (3, 0, -2)$, $\vec{n} = \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix}$.

(3) Encontrar la ecuación cartesiana del plano que pasa por P, Q y R.

(a) $P = (1, 1, 1), Q = (4, 0, 2), R = (0, 1, -1)$.

(b) $P = (1, 1, 0), Q = (1, 0, 1), R = (0, 1, 1)$.

(4) Determinar cuáles de los siguientes planos son paralelos o perpendiculares.

(a) $2x + 3y - z = 1$.

(b) $4x - y + 5z = 0$.

(c) $x - y - z = 3$.

(d) $4x + 6y - 2z = 0$.

(5) Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^3$ que pasa por el punto $P = (-1, 0, 3)$ y es perpendicular al plano con ecuación general $x - 3y + 2z = 5$.

(6) Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^3$ que pasa por el punto $P = (-1, 0, 3)$ y es paralela a la recta con ecuaciones paramétricas

$$x = 1 - t$$

$$y = 2 + 3t$$

$$z = -2 - t.$$

(7) Encuentre la ecuación del plano que pasa por $P = (0, -2, 5)$ y es paralelo al plano con ecuación general $6x - y + 2z = 3$.

Capítulo 6

Espacios vectoriales

En los dos capítulos anteriores resultó notorio que $\mathbb{R}^2$ y $\mathbb{R}^3$ comparten muchas propiedades, otra gran variedad de estructuras algebraicas también comparten propiedades similares; este hecho motiva a hacer abstracción y, en vez de estudiar $\mathbb{R}^2$, $\mathbb{R}^3$ y otras estructuras por separado, definir lo que se conoce como espacio vectorial.

6.1. Definición y ejemplos

Definición 6.1. Un *espacio vectorial* (o un *espacio lineal*) consta de:

(1) Un conjunto V de objetos, que llamaremos vectores.

(2) Una operación llamada *suma de vectores*, que asigna a cada par de vectores $x, y \in V$ un vector $x + y \in V$, llamado suma de x y de y, de manera que se cumplan las siguientes propiedades:

  (a) La suma es conmutativa, esto es, $x + y = y + x$, para todo $x, y \in V$,

  (b) la suma es asociativa, esto es, $x + (y + z) = (x + y) + z$, para todo $x, y, z \in V$,

  (c) existe un único vector $\mathbf{0}$ en V, llamado el *vector cero*, tal que $x + \mathbf{0} = x$ para todo $x \in V$,

  (d) para cada vector $x \in X$ existe un único vector $-x \in V$ tal que $x + (-x) = 0$.

(3) Una operación llamada *multiplicación por un escalar*, que asigna a cada escalar $\lambda \in \mathbb{R}$ y a cada vector $x \in V$ un vector λx, llamado producto de λ y de x, de manera que se cumplan las siguientes propiedades:

  (a) $1x = x$ para todo $x \in V$,

  (b) $(\lambda_1\lambda_2)\, x = \lambda_1(\lambda_2 x)$, para todo $\lambda_1, \lambda_2 \in \mathbb{R}$, $x \in V$,

  (c) $\lambda\,(x + y) = \lambda x + \lambda y$, para todo $\lambda \in \mathbb{R}$, $x, y \in V$,

  (d) $(\lambda_1 + \lambda_2)\, x = \lambda_1 x + \lambda_2 x$, para todo $\lambda_1, \lambda_2 \in \mathbb{R}$, $x \in V$.

Observación 6.2. En este contexto, a los números reales se les suele llamar escalares y no se le suele poner flechas en la parte de arriba a los símbolos que denotan vectores.

Proposición 6.3. *En todo espacio vectorial* V

$$0u = \mathbf{0} \text{ para todo } u \in V,$$

en palabras:

El número cero multiplicado por cualquier vector es igual al vector nulo

Demostración. Sea $u \in V$

$$0u = (0+0)u = 0u + 0u,$$

sumando $-0u$ a ambos miembros

$$\mathbf{0} = 0u.$$

□

Ejemplo 6.4.

(1) $\mathbb{R}$, $\mathbb{R}^2$ y $\mathbb{R}^3$ son espacios vectoriales con las operaciones usuales.

(2) Sea n un entero positivo y sea

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, x_2, \ldots, x_n) : x_k \in \mathbb{R},\ k = 1, \ldots n\}$$

es un espacio vectorial con las operaciones

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$
$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Al elemento de la forma $(x_1, x_2, \ldots, x_n)$ se le suele llamar n-upla.

(3) El conjunto $\{0\}$ es un espacio vectorial.

(4) $\emptyset$ no es un espacio vectorial ¿por qué?

(5) El conjunto de los polinomios con las operaciones usuales es un espacio vectorial.

(6) Sea n un entero positivo y sea $\mathcal{P}_n$ el conjunto de los polinomios de grado menor o igual que n, entonces $\mathcal{P}_n$ es un espacio vectorial.

(7) Sea n un entero positivo, el conjunto de los polinomios de grado n NO es un espacio vectorial, ¿por qué?

(8) Sea A un subconjunto de $\mathbb{R}$, el conjunto de las funciones de A en $\mathbb{R}$, con las operaciones usuales de suma y producto por un escalar, es un espacio vectorial.

Sea V un espacio vectorial y sean u y $u_1, \ldots, u_n$ elementos de V, se dice que u es una *combinación lineal* de $u_1, \ldots, u_n$ si existen escalares $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n = \sum_{k=1}^{n} \lambda_k u_k.$$

Ejemplo 6.5.

(1) Todo vector de $\mathbb{R}^2$ es combinación lineal de los vectores $(1,0)$ y $(0,1)$, ya que

$$(x, y) = x(1,0) + y(0,1) \quad \text{si } x, y \in \mathbb{R}.$$

(2) Todo vector de $\mathbb{R}^3$ es combinación lineal de los vectores $(1,0,0)$, $(0,1,0)$ y $(0,0,1)$, ya que

$$(x, y, z) = x(1,0,0) + y(0,1,0) + z(0,0,1) \quad \text{si } x, y, z \in \mathbb{R}.$$

(3) Todo vector de $\mathbb{R}^2$ es combinación lineal de los vectores $(1,1)$ y $(0,1)$, ya que

$$(x,y) = x(1,1) + (y-x)(0,1) \quad \text{si } x,y \in \mathbb{R}.$$

(4) Todo polinomio es combinación lineal de los polinomios (vectores)

$$\{1, x, x^2, \dots, x^n, \dots\}.$$

6.2. Subespacios

DEFINICIÓN 6.6. Sea V un espacio vectorial y W un subconjunto de V, se dice que W es un *subespacio* de V, si W en si mismo es un espacio vectorial con las mismas operaciones de suma de vectores y multiplicación por un escalar de V

EJEMPLO 6.7.

(1) El conjunto $\{0\}$ (el conjunto cuyo único elemento es el vector nulo) siempre es un subespacio.
(2) El conjunto $\{(x_1, x_2, 0) \, : \, x_1, x_2 \in \mathbb{R}\}$ es un subespacio de $\mathbb{R}^3$, es de hacer notar que este subespacio se identifica de manera natural con $\mathbb{R}^2$.
(3) Sea n un entero positivo, el conjunto de los polinomios de grado menor o igual que n es un subespacio del espacio de los polinomios.
(4) El conjunto $\{(x,y) \in \mathbb{R}^2 \, : \, y = 2x\}$ es un subespacio de $\mathbb{R}^2$, más generalmente, cualquier recta que pasa por el origen es un subespacio de $\mathbb{R}^2$.
(5) El conjunto $\{(x,y) \in \mathbb{R}^2 \, : \, y = 2x - 1\}$ NO es un subespacio de $\mathbb{R}^2$, más generalmente, cualquier recta que no pasa por el origen no es un subespacio de $\mathbb{R}^2$.
(6) Todo plano que contiene al origen es un subespacio de $\mathbb{R}^3$.

OBSERVACIÓN 6.8. La intersección de subespacios es un subespacio, la unión de subespacios no necesariamente es un subespacio.

A manera de ejercicio, demostrar el siguiente resultado que sigue de la definición de subespacio.

TEOREMA 6.9. *Sea V un espacio vectorial y sea W un subconjunto de V, entonces W es un subespacio de V si y solo si W es distinto de $\emptyset$ y $u + \alpha v \in W$ para todos $u, v \in W$ y $\alpha \in \mathbb{R}$.*

DEFINICIÓN 6.10. Sea V un espacio vectorial y sea A un subconjunto de V. El *subespacio generado* por A, que se denotará por $\mathrm{Span}(A)$, es el subespacio más pequeño de V, que contiene al conjunto A.

OBSERVACIÓN 6.11. Hay ciertos detalles que se deben tomar en cuenta:

- Si V es un espacio vectorial y $A \subseteq V$, el subespacio generado por el conjunto A es igual a la intersección de todos los subespacios de V que contienen a A. Como la intersección de subespacios es un subespacio, este conjunto es un subespacio y, por construcción, cualquier subespacio que contenga a A lo contiene; por eso es el subespacio más pequeño que contiene a A

- El subespacio más pequeño de un espacio vectorial es el conjunto $\{0\}$.
- El subespacio generado por el conjunto $\emptyset$ es $\{0\}$.
- Si $A \neq \emptyset$ el subespacio generado por A es el conjunto de las combinaciones lineales de los elementos de A.

Ejemplo 6.12.

(1) El subespacio de $\mathbb{R}^2$ generado por el conjunto $\{(1,-1)\}$ es la recta $y = -x$.
(2) El subespacio de $\mathbb{R}^2$ generado por el conjunto $\{(1,-1),(0,5)\}$ es $\mathbb{R}^2$.
(3) El subespacio de $\mathbb{R}^2$ generado por el conjunto $\{(1,-1),(0,1),(4,6),(6,-1)\}$ es $\mathbb{R}^2$.
(4) El subespacio de $\mathbb{R}^3$ generado por el conjunto $\{(1,1,1)\}$ es la recta que pasa por el origen en la dirección del vector $(1,1,1)$.
(5) El subespacio de $\mathbb{R}^3$ generado por el conjunto $\{(1,0,0),(0,0,9)\}$ es el plano xz.

Ejercicios.

(1) Determinar si el conjunto dado es un espacio vectorial, con las operaciones de suma de vectores y producto por un escalar usuales en el contexto correspondiente.
 (a) $\{(x,y) \in \mathbb{R}^2 \,:\, y = 0\}$.
 (b) $\{(x,y) \in \mathbb{R}^2 \,:\, y \leq 0\}$.
 (c) El conjunto de vectores en $\mathbb{R}^3$ de la forma (x,x,x).
 (d) El conjunto de polinomios de grado 4.
 (e) El conjunto de polinomios de grado menor o igual que 4.
 (f) $\emptyset$.
 (g) Los vectores del plano que están en el primer cuadrante.
 (h) El conjunto de polinomios de grado $\leq n$ con término constante cero.
 (i) El conjunto de polinomios de grado $\leq n$ con término constante a_0 positivo.
 (j) El conjunto de polinomios de grado $\leq n$ con término constante a_0 negativo.
(2) Sea $H = \{(x,y,z,w) \in \mathbb{R}^4 \,:\, ax + by + cz + dw = 0\}$, donde a, b, c y d son números reales, no todos cero. Demuestre que H es un subespacio propio de $\mathbb{R}^4$. A H se le conoce como hiperplano en $\mathbb{R}^4$ que pasa por el origen.

6.3. Bases, dimensión y coordenadas

Sea V un espacio vectorial.

Definición 6.13. Sea $A \subset V$, decimos que A es *linealmente independiente* si dados un número natural n, $u_1, \dots, u_n \in A$ y escalares $\lambda_1, \dots, \lambda_n$ se tiene que

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$$

implica que

$$\lambda_1 = \cdots = \lambda_n = 0.$$

En caso contrario se dice que el conjunto es *linealmente dependiente*.

EJEMPLO 6.14.

(1) Todo subconjunto de un espacio vectorial al que pertenece el vector 0 es linealmente dependiente.
(2) Un subconjunto de un espacio vectorial formado por un solo vector es linealmente independiente si y solo si el vector no es nulo.
(3) Los siguientes subconjuntos de $\mathbb{R}^2$ son linealmente independientes:

$$\{(1,0),(0,2)\} \qquad \{(-1,-2),(3,2)\}.$$

(4) El conjunto $\{(2,0),(0,3),(1,4)\}$ es linealmente dependiente.
(5) El conjunto $\{1, 1-x^3, 1+x^3\}$ es linealmente dependiente.

DEFINICIÓN 6.15. Sea V un espacio vectorial, se dice que un subconjunto $\mathfrak{B}$ de V es una *base* de V si:

(1) $\mathfrak{B}$ es linealmente independiente,
(2) El subespacio generado por $\mathcal{B}$ es V.

EJEMPLO 6.16.

(1) Los conjuntos $\{(1,0),(0,1)\}$, $\{(2,2),(3,2)\}$, $\{(5,0),(0,5)\}$ son bases de $\mathbb{R}^2$.
(2) El conjunto $\{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}$ es una base de $\mathbb{R}^4$.
(3) El conjunto $\{1, x, x^2, \dots, x^n, \dots\}$ es una base del conjunto de los polinomios.

Se cumple el siguiente resultado.

TEOREMA 6.17. *Sea V un espacio vectorial, que tiene una base con n elementos, donde n es un entero positivo, entonces:*

(1) *Cualquier otra base de V tiene exactamente n elementos.*
(2) *Todo subconjunto de V con más de n elementos es linealmente dependiente.*
(3) *Todo subconjunto de V linealmente independiente con n vectores es una base de V.*
(4) *Todo subconjunto de V que genere todo el espacio V y que tenga n elementos es linealmente independiente y, por lo tanto, es una base de V.*

Tomando en cuenta el resultado anterior, se da la siguiente definición.

DEFINICIÓN 6.18. Sea V un espacio vectorial, si V tiene una base con n elementos (n un entero positivo) se dice que la *dimensión* de V es n, en caso de que V no tenga base finita, se dice que V tiene dimensión infinita.

Observación 6.19. El caso del subespacio $\{0\}$ merece especial atención en lo que se refiere a bases y dimensión. A este espacio se le asigna dimensión 0, podemos aceptar esto como una convención, pero lo que ocurre es que el conjunto vacío ($\emptyset$) genera este subespacio y es linealmente independiente. Como la cantidad de elementos del conjunto vacío es 0, la dimensión del espacio $\{0\}$ es 0.

Si V es un espacio vectorial que contiene vectores no nulos (es decir $\{0\} \subsetneqq V$) el conjunto $\mathfrak{B}$ es una base de V si es linealmente independiente y todo elemento del espacio V es una combinación lineal de elementos de $\mathfrak{B}$.

Ejemplo 6.20.

(1) La dimensión de $\mathbb{R}^2$ es 2 y la de $\mathbb{R}^3$ es 3.
(2) La dimensión del espacio de los polinomios es infinita.
(3) La dimensión del espacio de los polinomios de grado menor o igual que n es $n+1$.
(4) La dimensión de $\mathbb{R}^n$ es n.

También se cumple el siguiente resultado.

Teorema 6.21. *Todo subconjunto linealmente independiente de un espacio vectorial está contenido en una base.*

En palabras: todo subconjunto linealmente independiente se puede "completar" hasta obtener una base.

6.3.1. Coordenadas. En el espacio $\mathbb{R}^2$ se tiene que si $(x_1, x_2) \in \mathbb{R}^2$,

$$(x_1, x_2) = x_1(1,0) + x_2(0,1),$$

en $\mathbb{R}^3$ se tiene que si $(x_1, x_2, x_3) \in \mathbb{R}^3$,

$$(x_1, x_2, x_3) = x_1(1,0,0) + x_2(0,1,0) + x_3(0,0,1),$$

más generalmente, si e_k es el vector de $\mathbb{R}^n$ que tiene 0 en todas sus coordenadas, menos en la coordenada k donde tiene el número 1 ($e_k = (0,\dots,0,1,0\dots,0)$, el 1 en el lugar k), se tiene que

$$(x_1, x_2, \dots, x_n) = x_1 e_1 + x_2 e_2 + \cdots x_n e_n = \sum_{k=1}^{n} x_k e_k.$$

Al conjunto $\{e_1, \dots, e_k\}$ se le suele llamar la *base canónica o estándar* de $\mathbb{R}^n$ y, en general, se dice que x_k es la *k-ésima coordenada* del vector si $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

Un espacio vectorial tiene diferentes bases y las expresiones de un vector en términos de esas bases varían, por ejemplo, los conjuntos $\{(1,0),(0,1)\}$ y $\{(1,1),(0,1)\}$ son bases de $\mathbb{R}^2$, si $u = (x_1, x_2) \in \mathbb{R}^2$ se tiene que

$$u = x_1(1,0) + x_2(0,1) = x_1(1,1) + (x_2 - x_1)(0,1).$$

Por esto se dice que el vector $u = (x_1, x_2)$ tiene coordenadas (x_1, x_2) en la base canónica y coordenadas $(x_1, x_2 - x_1)$ en la base $\{(1,1),(0,1)\}$ y se denota así

$$u_{\mathfrak{B}} = [x_1 \ \ x_2 - x_1] \qquad \text{ó} \qquad u_{\mathfrak{B}} = \begin{bmatrix} x_1 \\ x_2 - x_1 \end{bmatrix} \qquad \text{donde } \mathfrak{B} = \{(1,1),(0,1)\}.$$

En general, si V es un espacio vectorial de dimensión n, $\mathfrak{B} = \{u_1, \dots, u_n\}$ es una base de V y $u = x_1u_1 + x_2u_2 + \cdots x_nu_n$ se dice que las coordenadas de u en la base $\mathfrak{B}$ son $x_1, x_2, \dots, x_n$ y se denota por

$$u_{\mathfrak{B}} = [x_1 \ \ x_2 \ \cdots \ x_n] \qquad \text{ó} \qquad u_{\mathfrak{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Ejercicios.

(1) Determinar si el conjunto de vectores dado es linealmente dependiente o independiente.

(a) $\begin{bmatrix} 9 \\ -8 \end{bmatrix}, \begin{bmatrix} -11 \\ -3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \end{bmatrix}$ (c) $\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 7 \end{bmatrix}$

(d) $\begin{bmatrix} -6 \\ 1 \end{bmatrix}, \begin{bmatrix} 12 \\ -2 \end{bmatrix}$ (e) $\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 8 \end{bmatrix}$ (f) $\begin{bmatrix} -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 7 \end{bmatrix}$

(g) $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ (h) $\begin{bmatrix} -10 \\ -6 \end{bmatrix}, \begin{bmatrix} 10 \\ -6 \end{bmatrix}, \begin{bmatrix} 5 \\ 9 \end{bmatrix}$ (i) $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

(j) $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

(2) Determinar si el conjunto dado de vectores genera el espacio vectorial especificado.

(a) $\mathbb{R}^2 : \begin{bmatrix} 10 \\ 9 \end{bmatrix}, \begin{bmatrix} 18 \\ 6 \end{bmatrix}$ (b) $\mathbb{R}^2 : \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ (c) $\mathbb{R}^2 : \begin{bmatrix} 11 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

(d) $\mathbb{R}^2 : \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ -8 \end{bmatrix}$ (f) $\mathbb{R}^4 : \begin{bmatrix} 1 \\ -12 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 0 \\ -8 \end{bmatrix}$ (h) $\mathbb{R}^3 : \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(i) $\mathbb{R}^3 : \begin{bmatrix} -6 \\ 5 \\ 9 \end{bmatrix}, \begin{bmatrix} 7 \\ -12 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ (j) $\mathbb{R}^3 : \begin{bmatrix} -6 \\ 5 \\ 9 \end{bmatrix}, \begin{bmatrix} 7 \\ -12 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ -7 \\ 15 \end{bmatrix}$ (k) $\mathbb{R}^3 : \begin{bmatrix} -5 \\ 6 \\ 8 \end{bmatrix}, \begin{bmatrix} -3 \\ 7 \\ -1 \end{bmatrix}$

(3) Sea $\mathcal{P}_2$ el conjunto de los polinomios de grado menor o igual que 2, determinar si el conjunto B es una base para el espacio vectorial $\mathcal{P}_2$.

(a) $B = \{x, 1 + x, x - x^2\}$ (b) $B = \{1 - x, 1 - x^2, x - x^2\}$
(c) $B = \{1, 1 + 2x + 3x^2\}$ (d) $B = \{1, 2 - x, 3 - x^2, x + 2x^2\}$
(e) $B = \{x^2, 2 + x, x - x^2\}$ (f) $B = \{x^2, 2 + x\}$

(4) Encuentre una base para $\text{Span}\{-2x, 2x - x^2, 1 - x^2, 1 + x^2\}$ en $\mathcal{P}_2$.

(5) Encuentre una base para $\text{Span}\{1 - x, x - x^2, 1 - x^2, 1 - 2x + x^2\}$ en $\mathcal{P}_2$.

(6) ¿Por qué dos polinomios de grado menor o igual a dos, no pueden generar P_2?.

(7) Para qué valor(es) de α serán linealmente dependientes los vectores

$$\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \\ -2 \end{bmatrix}, \begin{bmatrix} \alpha \\ 1 \\ 2 \end{bmatrix}?$$

(8) Encuentre una base en R^3 para el conjunto de los vectores contenidos en el plano

$$3x - 2y + 5z = 0.$$

(9) Encuentre una base en R^3 para el conjunto de los vectores contenidos en el plano

$$3x - 2y + z = 0.$$

6.4. Espacios con producto interno

Definición 6.22. Sea V un espacio vectorial, un *producto interno* o *producto escalar* en V es una función

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

que satisface las siguientes propiedades para $u, v, w \in V$ y $\alpha \in \mathbb{R}$

(1) $\langle u, u\rangle \geq 0$ y es igual a 0 si y solo si $u = 0$,
(2) $\langle u, v\rangle = \langle v, u\rangle$,
(3) $\langle \alpha\, u, v\rangle = \alpha\, \langle u, v\rangle$,
(4) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.

Proposición 6.23. *Sea $\langle\ ,\ \rangle$ un producto interno en el espacio vectorial V, entonces*

$$\langle \mathbf{0}, u\rangle = 0 \text{ para todo } u \in V.$$

La demostración se deja como ejercicio.

Ya en los capítulos anteriores vimos ejemplos de producto interno, por ejemplo también es un producto interno en $\mathbb{R}^3$ el dado por

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3)\rangle = x_1y_1 + 3x_2y_2 + 10x_3y_3.$$

Definición 6.24. Sea V un espacio vectorial con producto interno $\langle\ ,\ \rangle$, la *norma asociada* a este producto interno se define como

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

Proposición 6.25 (Desigualdad de Cauchy-Schwarz). *Sea $\langle , \rangle$ un producto interno en el espacio vectorial V. Entonces*

$$|\langle u, v\rangle| \leq \|u\|\ \|v\|$$

para todo $u, v \in V$.

Además se cumple la igualdad si y sólo si $u = \lambda v$ para algún $\lambda \in \mathbb{R}$, es decir, u y v son colineales.

DEMOSTRACIÓN. Sean $u, v \in V$, entonces

$$\langle xv - u, xv - u\rangle \geq 0 \text{ para todo } x \in \mathbb{R},$$

por lo tanto

$$\langle v, v\rangle x^2 - 2\langle u, v\rangle x + \langle u, u\rangle \geq 0 \text{ para todo } x \in \mathbb{R},$$

es decir,

$$\|v\|^2 x^2 - 2\langle u, v\rangle x + \|u\|^2 \geq 0 \text{ para todo } x \in \mathbb{R}.$$

Si $\|v\| = 0$ entonces $v = 0$ y la desigualdad de Cauchy-Schwarz es trivialmente cierta.

Si $\|v\| > 0$, tenemos una parábola que se abre hacia arriba. Usando el discriminante se concluye que

$$4\langle u, v\rangle^2 - 4\|u\|^2\|v\|^2 \leq 0$$

y de esto último se deduce inmediatamente la desigualdad.

En caso de que se de la igualdad se tiene

$$\langle xv - u, xv - u\rangle = 0 \text{ para algún } x \in \mathbb{R},$$

de donde sigue que

$$u = xv \text{ para algún } x \in \mathbb{R}.$$

□

TEOREMA 6.26. *Sea V un espacio vectorial con producto interno $\langle , \rangle$ entonces la norma asociada $\| \ \|$ satisface las siguientes propiedades:*

Si u y v son elementos de V y α es un número real, se cumple

(1) $\|u\| \geq 0$, $\|u\| = 0$ *si y solo si* $u = 0$,
(2) $\|\alpha\, u\| = |\alpha|\, \|u\|$,
(3) *(Desigualdad triangular)* $\|u + v\| \leq \|u\| + \|v\|$.

En general una función que satisface las propiedades enunciadas en el teorema anterior se le conoce como *norma*, es importante destacar que existen normas no asociadas con un producto interno.

DEMOSTRACIÓN. Se probará la tercera propiedad, quedando las dos primeras de ejercicio.

Sean u y v elementos de V, entonces

$$\begin{aligned}\|u+v\|^2 =&\langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle = \\ =&\|u\|^2 + 2\langle u,v\rangle + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = \\ =&(\|u\| + \|v\|)^2.\end{aligned}$$

□

Definición 6.27. Sea V un espacio vectorial con producto interno $\langle\ ,\ \rangle$.

(1) Si $u, v \in V$, se dice que u y v son *ortogonales* (abreviado $u \perp v$) si

$$\langle u, v\rangle = 0.$$

(2) Si $\mathcal{M} = \{u_1, \ldots, u_n\}$ es un conjunto de vectores, se dice que $\mathcal{M}$ es *ortogonal* si sus elementos son ortogonales dos a dos, es decir

$$\langle u_i, u_j\rangle = 0 \text{ si } i \neq j.$$

(3) Se dice que $\mathcal{M}$ es *ortonormal* si $\mathcal{M}$ es ortogonal y cada vector u_i tiene norma 1.

Observación 6.28. De la igualdad

$$\|u+v\|^2 = \|u\|^2 + 2\langle u,v\rangle + \|v\|^2$$

sigue que dos vectores u y v son ortogonales si y solo si

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

Proposición 6.29. *Sean V un espacio vectorial con producto interno $\langle\ ,\ \rangle$, $\mathcal{M} = \{u_1, \ldots, u_n\}$ un conjunto ortogonal de vectores y $\alpha_1, \ldots, \alpha_n$ una colección de escalares, entonces*

(1) *Si $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$ entonces*

$$\|u\| = \sqrt{\alpha_1^2\|u_1\|^2 + \cdots + \alpha_n^2\|u_n\|^2}.$$

(2) *Si $\mathcal{M}$ no contiene al vector nulo, entonces $\mathcal{M}$ es linealmente independiente.*

(3) *(Generalización del teorema de Pitágoras) Si $\mathcal{M}$ es ortonormal y $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$, entonces*

$$\|u\| = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}.$$

Demostración.

(1) Si $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$ entonces

$$\begin{aligned}\|u\|^2 =& \langle u, u\rangle = \left\langle \sum_{i=1}^{n} \alpha_i u_i, \sum_{j=1}^{n} \alpha_j u_j \right\rangle = \sum_{i,j=1}^{n} \langle \alpha_i u_i, \alpha_j u_j\rangle \\ =& \sum_{i,j=1}^{n} \alpha_i \alpha_j \langle u_i, u_j\rangle = \sum_{i=1}^{n} \alpha_i^2 \|u_i\|^2.\end{aligned}$$

(2) Si $\alpha_1 u_1 + \cdots + \alpha_n u_n = \mathbf{0}$, entonces

$$\sqrt{\alpha_1^2\|u_1\|^2 + \cdots + \alpha_n^2\|u_n\|^2} = 0,$$

como los vectores no son nulos se tiene que cumplir

$$\alpha_1 = \cdots = \alpha_n = 0.$$

(3) Sigue de (1).

□

COROLARIO 6.30. *Todo conjunto ortonormal es linealmente independiente.*

Ejercicios.

(1) Demuestre que para cualquier número real t, los vectores

$$\begin{bmatrix} \operatorname{sen} t \\ \cos t \end{bmatrix}, \quad \begin{bmatrix} \cos t \\ -\operatorname{sen} t \end{bmatrix}$$

forman una base ortonormal de $\mathbb{R}^2$.

(2) Explicar en detalle la demostración del Corolario 6.30.

(3) Dar un ejemplo de una base ortonormal de $\mathbb{R}^3$ que no sea la base canónica.

6.5. Proceso de ortonormalización de Gram-Schmidt

Supongamos que tenemos el siguiente problema:

Tenemos dos vectores, u_1 y u_2, linealmente independientes en $\mathbb{R}^3$ y queremos hallar una base ortonormal del subespacio $\operatorname{Span}\{u_1, u_2\}$.

El primer paso sería construir un vector unitario en la misma dirección que el vector u_1, este vector es

$$v_1 = \frac{u_1}{\|u_1\|}.$$

La proyección del vector u_2 en la dirección de v_1 es

$$\langle u_2, v_1\rangle v_1.$$

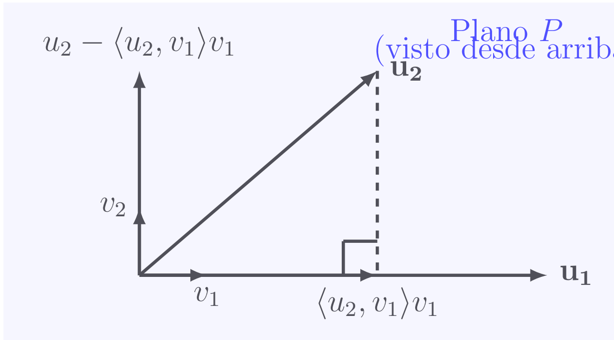


Figura 6.1. Ortonormalización de dos vectores

El vector $w = u_2 - \langle u_2, v_1\rangle v_1$ satisface

$$\begin{aligned}\langle w, v_1\rangle &= \langle u_2 - \langle u_2, v_1\rangle v_1, v_1\rangle = \langle u_2, v_1\rangle - \langle\langle u_2, v_1\rangle v_1, v_1\rangle \\ &= \langle u_2, v_1\rangle - \langle u_2, v_1\rangle \langle v_1, v_1\rangle = \langle u_2, v_1\rangle - \langle u_2, v_1\rangle \|v_1\|^2 \\ &= \langle u_2, v_1\rangle - \langle u_2, v_1\rangle \\ &= 0,\end{aligned}$$

es decir, w es ortogonal a v_1, además $w \neq 0$ porque u_1 y u_2 son linealmente independientes y

$$\operatorname{Span}\{w, v_1\} = \operatorname{Span}\{u_1, u_2\}.$$

Tomando

$$v_2 = \frac{w}{\|w\|},$$

obtenemos

$$\{v_1, v_2\} \text{ es ortonormal y } \operatorname{Span}\{v_1, v_2\} = \operatorname{Span}\{u_1, u_2\}.$$

Observación 6.31. Es muy importante hacer notar que de todos los vectores que son múltiplos de v_1 (o, lo que es igual, múltiplos de u_1) el que mejor aproxima a u_2 es $\langle u_2, v_1\rangle v_1$.

Gráficamente resulta claro, si lo queremos ver analíticamente, consideremos la función cuadrática en α

$$f(\alpha) = \|u_2 - \alpha v_1\|^2 = \|u_2\|^2 - 2\langle u_2, v_1\rangle\alpha + \alpha^2\|v_1\|^2 = \|u_2\|^2 - 2\langle u_2, v_1\rangle\alpha + \alpha^2,$$

esta función alcanza su valor mínimo en

$$\alpha_o = -\frac{(-2\langle u_2, v_1\rangle)}{2} = \langle u_2, v_1\rangle,$$

(ver (3) en página 35).

Por lo tanto, la mejor aproximación se obtiene al tomar el vector $\alpha_o v_1 = \langle u_2, v_1\rangle v_1$

No existe limitación para realizar un proceso de ortonormalización como el ya descrito en un espacio con producto interno de dimensión arbitraria.

En detalle, supongamos que V es un espacio con producto interno y $\{u_1, \dots, u_n\}$ es un subconjunto linealmente independiente de V, se procede de la siguiente manera:

$$
\begin{aligned}
v_1 =& \frac{u_1}{\|u_1\|} \qquad \operatorname{Span}\{v_1\} = \operatorname{Span}\{u_1\}, \\
v_2 =& \frac{1}{\|u_2 - \langle u_2, v_1\rangle v_1\|} \left(u_2 - \langle u_2, v_1\rangle v_1\right) \qquad \operatorname{Span}\{v_1, v_2\} = \operatorname{Span}\{u_1, u_2\} \\
& \qquad \text{y } v_2 \perp \operatorname{Span}\{u_1\}, \\
v_3 =& \frac{1}{\|\langle u_3, v_1\rangle v_1 - \langle u_3, v_2\rangle v_2\|} \left(u_3 - (\langle u_3, v_1\rangle v_1 - \langle u_3, v_2\rangle v_2\right) \\
& \qquad \operatorname{Span}\{v_1, v_2, v_3\} = \operatorname{Span}\{u_1, u_2, u_3\} \\
& \qquad \text{y } v_3 \perp \operatorname{Span}\{u_1, u_2\}, \\
v_4 =& \frac{1}{\|u_4 - (\langle u_4, v_1\rangle v_1 - \langle u_4, v_2\rangle v_2 - \langle u_4, v_1\rangle v_1\|} \left(u_4 - (\langle u_4, v_1\rangle v_1 - \langle u_4, v_2\rangle v_2 - \langle u_4, v_1\rangle v_1)\right) \\
& \qquad \operatorname{Span}\{v_1, v_2, v_3, v_4\} = \operatorname{Span}\{u_1, u_2, u_3, u_4\}, \\
& \qquad \text{y } v_4 \perp \operatorname{Span}\{u_1, u_2, u_3\}, \\
& \vdots \\
& \vdots
\end{aligned}
$$

El proceso anterior se conoce como *método de ortonormalización de Gram-Schmidt* y es un algoritmo que permite transformar un conjunto de vectores linealmente independientes en un conjunto ortonormal.

Observación 6.32. Si en el proceso de Gram-Schmidt se omite el paso de dividir ente la norma del vector, se obtiene un conjunto ortogonal.

Ejemplo 6.33 (En $\mathbb{R}^3$).
Ortonormalizar los siguientes vectores linealmente independientes:

$$
u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}
$$

Primer vector:

$$
v_1 = \frac{u_1}{||u_1||} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}
$$

Segundo vector:

$$v_2 = \frac{1}{\|u_2 - \langle u_2, v_1\rangle\, v_1\|}\,(u_2 - \langle u_2, v_1\rangle\, v_1).$$

$$\begin{aligned}
u_2 - \langle u_2, v_1\rangle\, v_1 &= \begin{bmatrix}1\\0\\1\end{bmatrix} - \left\langle \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix} \right\rangle \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix}\\
&= \begin{bmatrix}1\\0\\1\end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix} = \begin{bmatrix}1\\0\\1\end{bmatrix} - \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}\\
&= \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix}.
\end{aligned}$$

$$\|u_2 - \langle u_2, v_1\rangle\, v_1\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 1^2} = \sqrt{\frac{1}{4} + \frac{1}{4} + 1} = \sqrt{\frac{3}{2}}$$

$$v_2 = \frac{1}{\sqrt{\frac{3}{2}}} \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}\\\sqrt{\frac{2}{3}}\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix}$$

Tercer vector:

$$v_3 = \frac{1}{\|\langle u_3, v_1\rangle\, v_1 - \langle u_3, v_2\rangle\, v_2\|}\,(u_3 - (\langle u_3, v_1\rangle\, v_1 - \langle u_3, v_2\rangle\, v_2)$$

$$\begin{aligned}
\langle u_3, v_1\rangle\, v_1 &= \left\langle \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix} \right\rangle \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix}\\
&= \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\langle u_3, v_2\rangle\, v_2 &= \left\langle \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix} \right\rangle \begin{bmatrix}\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix}\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix}\\
&= \begin{bmatrix}\frac{1}{6}\\-\frac{1}{6}\\\frac{1}{3}\end{bmatrix}
\end{aligned}$$

$$u_3 - \langle u_3, v_1\rangle\, v_1 - \langle u_3, v_2\rangle\, v_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{6} \\ -\frac{1}{6} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} - \frac{1}{6} \\ 1 - \frac{1}{2} + \frac{1}{6} \\ 1 - \frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} = \frac{2}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$\| u_3 - \langle u_3, v_1\rangle\, v_1 - \langle u_3, v_2\rangle\, v_2 \| = \frac{2}{3}\sqrt{1+1+1} = \frac{2\sqrt{3}}{3}$$

$$v_3 = \left(\frac{3}{2\sqrt{3}}\right)\left(\frac{2}{3}\right) \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

Después de ortonormalizar el sistema, se obtienen los vectores ortonormales

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix} \qquad \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Existen programas informáticos que permiten ortonormalizar conjuntos de vectores, por ejemplo "Mathematica" de Wolfram Labs; a medida que aumenta el número de vectores el procedimiento se complica.

EJEMPLO 6.34. Hallar una base ortonormal del subespacio de $\mathbb{R}^5$ generado por los vectores

$$u_1 = \begin{bmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}$$

Primer vector:

$$v_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{3^2+4^2}} \begin{bmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ 0 \\ \frac{4}{5} \\ 0 \\ 0 \end{bmatrix}$$

Segundo vector:

$$v_2 = \frac{1}{\|u_2 - \langle u_2, v_1\rangle\, v_1\|}\,(u_2 - \langle u_2, v_1\rangle\, v_1).$$

$$
\begin{aligned}
u_2 - \langle u_2, v_1 \rangle \, v_1 &= \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}, \frac{1}{5} \begin{bmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \end{bmatrix} \right\rangle \frac{1}{5} \begin{bmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix} - \frac{15}{25} \begin{bmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{9}{5} \\ 0 \\ \frac{12}{5} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{16}{5} \\ 0 \\ -\frac{12}{5} \\ 3 \\ 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 16 \\ 0 \\ -12 \\ 15 \\ 0 \end{bmatrix}
\end{aligned}
$$

$$
\| u_2 - \langle u_2, v_1 \rangle \, v_1 \| = \frac{1}{5} \sqrt{16^2 + 12^2 + 15^2} = \frac{1}{5} \sqrt{625} = \frac{25}{5} = 5.
$$

$$
v_2 = \frac{1}{5} \begin{bmatrix} \frac{16}{5} \\ 0 \\ -\frac{12}{5} \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{16}{25} \\ 0 \\ -\frac{12}{25} \\ \frac{3}{5} \\ 0 \end{bmatrix}
$$

Una base ortonormal del espacio es el conjunto

$$
\left\{ \begin{bmatrix} \frac{3}{5} \\ 0 \\ \frac{4}{5} \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} \frac{16}{25} \\ 0 \\ -\frac{12}{25} \\ \frac{3}{5} \\ 0 \end{bmatrix} \right\}.
$$

Una consecuencia del proceso de ortonormalización de Gram-Schmidt es la siguiente.

Teorema 6.35. *Todo espacio vectorial de dimensión finita y con producto interno posee una base ortogonal y una base ortonormal.*

Observación 6.36. Sea V un espacio vectorial de dimensión finita con producto interno $\langle \, , \, \rangle$ y sea $\{v_1, \dots, v_n\}$ una base ortonormal de V. Entonces

(1) Todo vector v de V se escribe en términos de la base de la siguiente manera

$$
v = \langle v, v_1 \rangle v_1 + \cdots + \langle v, v_n \rangle v_n = \sum_{i=1}^{n} \langle v, v_i \rangle v_i,
$$

es decir, si

$$
v = \alpha_1 v_1 + \cdots + \alpha_n v_n,
$$

entonces

$$
\alpha_i = \langle v, v_i \rangle v_i.
$$

(2) Se cumple que
$$\|v\| = \sqrt{\langle v, v_1\rangle^2 + \cdots + \langle v, v_n\rangle^2}\,:$$
(3) La proyección del vector v sobre el espacio generado por $\{v_1, \dots, v_k\}$ $(1 \leq k \leq n)$ es el vector
$$\langle v, v_1\rangle v_1 + \cdots + \langle v, v_k\rangle v_k = \sum_{i=1}^{k} \langle v, v_i\rangle v_i.$$
(4) El vector del subespacio generado por $\{v_1, \dots, v_k\}$ $(1 \leq k \leq n)$ que mejor aproxima a V es
$$\langle v, v_1\rangle v_1 + \cdots + \langle v, v_k\rangle v_k = \sum_{i=1}^{k} \langle v, v_i\rangle v_i,$$
es decir, su proyección ortogonal sobre el subespacio generado por $\{v_1, \dots, v_k\}$.

6.6. Complemento ortogonal y suma directa

Sea V un espacio vectorial con producto interno $\langle\ ,\ \rangle$.

Si A es un subconjunto de V se define
$$A^{\perp} = \{v \in V \ :\ \langle v, u\rangle = 0 \text{ para todo } u \in A\}.$$

Se cumple que, cualquiera que sea el subconjunto A de V, el conjunto $A^{\perp}$ es un subespacio de V.

Si W es un subespacio de W se define su complemento ortogonal como $W^{\perp}$, se tiene que todo vector v de V se puede escribir, de manera única como
$$v = w + z, \text{ donde } w \in W \text{ y } z \in W^{\perp},$$
en este caso se dice que V es la suma directa de los subespacios W y $W^{\perp}$ y se abrevia
$$V = W \oplus W^{\perp}.$$

6.7. Ejercicios varios

(1) Sea $H = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n \ :\ a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}$, donde $a_1, a_2, \dots, a_n$ son números reales no todos cero. Demuestre que H es un subespacio propio de $\mathbb{R}^n$. A H se le conoce como hiperplano en $\mathbb{R}^n$ que pasa por el origen.

(2) Sean v_1 y v_2 dos vectores en $\mathbb{R}^2$. Demuestre que $H = \{v \ :\ v = av_1 + bv_2;\ a, b \in \mathbb{R}\}$ es un subespacio de $\mathbb{R}^2$.

(3) Sean H_1 y H_2 subespacios de un espacio vectorial V y sea
$$H_1 + H_2 = \{v \ :\ v = v_1 + v_2 \text{ con } v_1 \in H_1 \text{ y } v_2 \in H_2\}.$$
Demuestre que $H_1 + H_2$ es un subespacio de V.

(4) Determinar si el conjunto de vectores dado es linealmente dependiente o independiente.

(a) $\begin{bmatrix}8\\-7\\-8\end{bmatrix},\begin{bmatrix}-11\\-12\\-7\end{bmatrix},\begin{bmatrix}12\\-3\\7\end{bmatrix}$ (b) $\begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}-1\\1\\-1\end{bmatrix},\begin{bmatrix}4\\-1\\1\end{bmatrix}$

(c) $\begin{bmatrix}-3\\4\\2\end{bmatrix},\begin{bmatrix}7\\-1\\3\end{bmatrix},\begin{bmatrix}1\\1\\8\end{bmatrix}$ (d) $\begin{bmatrix}-1\\0\\11\end{bmatrix},\begin{bmatrix}7\\-20\\-29\end{bmatrix},\begin{bmatrix}1\\-5\\1\end{bmatrix}$

(5) Sea $\mathcal{P}_n$ el conjunto de los polinomios de grado menor o igual que n, determinar si el conjunto de elementos de $\mathcal{P}_2$ dado es linealmente dependiente o independiente.
  (a) $4-3x+3x^2, 4-2x-2x^2$
  (b) $P_2 : 1-x, x$
  (c) $P_2 : -x, x^2-2x, 3x+5x^2$
  (d) $x-1, x-2, x-3.x-4$
(6) Determinar si el conjunto dado de vectores genera el espacio vectorial dado.

  (a) $\mathbb{R}^3 : \begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}-1\\2\\3\end{bmatrix}$ (b) $\mathbb{R}^3 : \begin{bmatrix}0\\5\\1\end{bmatrix},\begin{bmatrix}0\\-1\\3\end{bmatrix}$
  (c) $P_2 : 1-x, 3-x^2$ (d) $P_2 : -10+3x+11x^2, 10+9x-4x^2, 5+x+4x^2$
(7) Encuentre una base para $\text{Span}\{1, 1+x, 2x\}$ en $\mathcal{P}_1$.
(8) Encuentre una base para $\text{Span}\{1-2x, 2x-x^2, 1-x^2, 1+x^2\}$ en $\mathcal{P}_2$.
(9) Encuentre una base para $\text{Span}\{1-x, x-x^2, 1-x^2, 1-2x+x^2\}$ en $\mathcal{P}_2$.
(10) Sea $\{v_1, \dots, v_n\}$ una base para un espacio vectorial V y sean $c_1, c_2, \dots, c_n$ escalares distintos de cero. Demuestre que $\{c_1, c_2, \dots, c_n, v_n\}$ también es una base para V.
(11) Determine una condición sobre los números a, b, c y d para que

$$\begin{bmatrix}a\\b\end{bmatrix} y \begin{bmatrix}c\\d\end{bmatrix}$$

sean vectores linealmente dependientes.
(12) Para cuál(es) valor(es) de α serán linealmente dependientes los vectores

$$\begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}2\\-1\\4\end{bmatrix},\begin{bmatrix}3\\\alpha\\4\end{bmatrix}?$$

(13) Encuentre una base en $\mathbb{R}^3$ para el conjunto de vectores en la recta $x = 2, y = -2t, z = 3t$.
(14) Encuentre una base en $\mathbb{R}^3$ para el conjunto de vectores en la recta $x = 3t, y = -2t, z = t$.
(15) Determinar si el conjunto dado es una base para el espacio vectorial a que se refiere.

  (a) $\mathcal{P}_2 : -2-11x+7x^2, -5-x-5x^2$ (b) $\mathcal{P}_2 : 1-x^2, x$
  (c) $\mathcal{P}_2 : -3x, 1+x^2, x^2-5$ (d) $\mathcal{P}_2 : 1+3x+7x^2, 5+12x+35x^2, 8+5x-12x^2$
  (e) $\mathcal{P}_3 : 1, 1+x, 1+x^2, 1+x^3$ (f) $\mathcal{P}_3 : 3, x^3-4x+6, x^2$

(16) Expresar el vector $\begin{bmatrix}x\\y\end{bmatrix} \in \mathbb{R}^2$ en términos de la base dada.

(a) $\left\{\begin{bmatrix}7\\-12\end{bmatrix}, \begin{bmatrix}3\\6\end{bmatrix}\right\}$ (b) $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix}2\\-3\end{bmatrix}, \begin{bmatrix}3\\-2\end{bmatrix}\right\}$ (d) $\left\{\begin{bmatrix}0\\-7\end{bmatrix}, \begin{bmatrix}-8\\7\end{bmatrix}\right\}$

(e) $\left\{\begin{bmatrix}5\\7\end{bmatrix}, \begin{bmatrix}3\\-4\end{bmatrix}\right\}$ (f) $\left\{\begin{bmatrix}a\\c\end{bmatrix}, \begin{bmatrix}b\\d\end{bmatrix}\right\}$, donde $ad - bc \neq 0$

(17) Expresar el vector $\begin{bmatrix}x\\y\\z\end{bmatrix} \in \mathbb{R}^3$ en términos de la base dada.

(a) $\left\{\begin{bmatrix}-5\\0\\3\end{bmatrix}, \begin{bmatrix}1\\2\\-2\end{bmatrix}, \begin{bmatrix}5\\2\\0\end{bmatrix}\right\}$ (b) $\left\{\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$ (d) $\left\{\begin{bmatrix}-4\\-1\\2\end{bmatrix}, \begin{bmatrix}-1\\-1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$

(18) Expresar el polinomio $a_0 + a_1x + a_2x^2$ en $\mathcal{P}_2$ en términos de la base dada.

(a) $\{1, x-1, x^2-1\}$

(b) $\{1 + x + 4x^2, -3 + 4x - 2x^2, 3 - 2x + 4x^2\}$

(c) $\{-2 - 4x - x^2, -4 + 4x - 4x^2, -1 + 5x + 5x^2, -1 + 5x + 15^2\}$

(19) En $\mathbb{R}^2$ suponga que $x_{B_1} = \begin{bmatrix}4\\0\\-1\end{bmatrix}$, donde $B_1 = \left\{\begin{bmatrix}3\\-4\\3\end{bmatrix} \begin{bmatrix}0\\1\\-4\end{bmatrix} \begin{bmatrix}-3\\2\\-5\end{bmatrix}\right\}$.

Escriba x en términos de la base $B_2 = \left\{\begin{bmatrix}-4\\-2\\-4\end{bmatrix} \begin{bmatrix}-3\\1\\1\end{bmatrix} \begin{bmatrix}-2\\-2\\-5\end{bmatrix}\right\}$.

(20) En $\mathcal{P}_2, setienequeq_{B_1} = \begin{bmatrix}2\\1\\3\end{bmatrix}$, donde $B_1 = \{1 - x, 3x, x^2 - x - 1\}$.

Exprese q en términos de la base $B_2 = \{3 - 2x, 1 + x, x + x^2\}$.

(21) Construir una base ortonormal de $\mathbb{R}^2$ cuyo primer vector tenga la dirección y el sentido del vector

(a) $\begin{bmatrix}1\\3\end{bmatrix}$ (b) $\begin{bmatrix}-1\\-2\end{bmatrix}$ (c) $\begin{bmatrix}3\\0\end{bmatrix}$

(22) Construir una base ortonormal para el espacio o subespacio vectorial dado.

(a) $H = \{(x, y) \in \mathbb{R}^2 : x + y = 0\}$.

(b) $H = \{(x, y) \in \mathbb{R}^2 : 2x + y = 0\}$.

(c) $H = \{(x, y, z) \in \mathbb{R}^3 : 2x + y = 0\}$.

(d) $H = \{(x, y, z) \in \mathbb{R}^3 : x = 3t, y = 4t, z = 0; t \in \mathbb{R}\}$.

(e) $L = \{(x, y, z) \in \mathbb{R}^3 : x = t, y = 2t, z = -2t; t \in \mathbb{R}\}$.

(23) Sea $\{v_1, v_2, \dots, v_k\}$ un conjunto de vectores linealmente independientes en un espacio vectorial V. Pruebe que $v_i \neq 0$ para $i = 1, 2, \dots, k$. (Sugerencia: Si $v_i = 0$, entonces es sencillo encontrar constantes $c_1, c_2, \dots, c_k$ con $c_i \neq 0$ tales que $c_1v_1 + c_2v_2 + \dots + c_kv_k = 0$.)

(24) En los siguientes problemas se da un subespacio H de un espacio vectorial V y un vector v. Debe hallar la proyección ortogonal de v sobre H y escribir v como $v_1 + v_2$, donde $v_1 \in H$ y $v_2 \in H^{\perp}$

(a) $H = \{(x, y) \in \mathbb{R}^2 \,:\, x + y = 0\}$; $v = (-1, 2)$.
(b) $H = \{(x, y) \in \mathbb{R}^2 \,:\, x - y = 0\}$; $v = (2, -1)$.
(c) $H = \{(x, y) \in \mathbb{R}^2 \,:\, x + 2y = 0\}$; $v = (2, -1)$.
(d) $H = \{(x, y, z) \in \mathbb{R}^3 \,:\, \frac{x}{2} = \frac{y}{3} = \frac{z}{4} = 0\}$; $v = (-1, 2, 0)$.
(e) $H = \{(x, y, z) \in \mathbb{R}^3 \,:\, x - y + z = 0\}$; $v = (1, 2, 0)$.
(f) $H = \{(x, y, z) \in \mathbb{R}^3 \,:\, x - y + z = 0\}$; $v = (2, 3, 1)$.

(25) Encuentre una condición sobre los números a, b para que el conjunto

$$\left\{ \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} b \\ -a \end{pmatrix} \right\}$$

sea una base ortonormal de $\mathbb{R}^2$.

Capítulo 7

Sistemas de ecuaciones lineales y matrices

7.1. Ecuaciones lineales en dos variables

Una *ecuación lineal* en las variables x, y es una ecuación de la forma

$$a_1x + a_2y = b$$

donde a_1, a_2 y $b$son números reales.

Si $b = 0$ se dice que la ecuación es *homogénea* , en otro caso se dice que es *no homogénea.*

Ejemplo 7.1. Las siguientes ecuaciones en las variables x e y son lineales

$$2x - y = 0$$

$$2x - y = 6$$

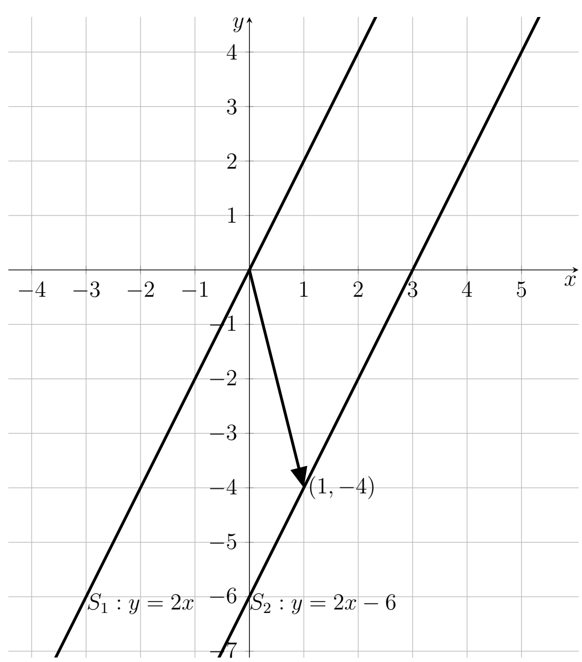
La solución de la primera es el conjunto

$$S_1 = \{(x, y) \in \mathbb{R}^2 \ : \ y = 2x\},$$

y la solución de la segunda es el conjunto

$$S_2 = \{(x, y) \in \mathbb{R}^2 \ : \ y = 2x - 6\}.$$

Gráficamente



El conjunto S_1 es un subespacio del espacio vectorial $\mathbb{R}^2$, ya que $(0,0) \in S_1$ y si los vectores $(x_1, y_1), (x_2, y_2) \in S_1$ y $\alpha \in \mathbb{R}$ se tiene que

$$y_1 + \alpha y_2 = 2x_1 + 2\alpha x_2 = 2(x_1 + \alpha x_2),$$

lo que implica que $(x_1, y_1) + \alpha(x_2, y_2) = (x_1 + \alpha x_2, y_1 + \alpha y_2) \in S_1$.

La solución de la segunda es el conjunto

$$S_2 = \{(x, y) \in \mathbb{R}^2 \ : \ y = 2x - 1\},$$

El conjunto S_2 **NO** es un subespacio del espacio vectorial $\mathbb{R}^2$, ya que $(0,0) \notin S_2$, sin embargo, es notorio que si $(x, y) \in S_1$ entonces $(x, y) + (1, -4) \in S_2$.

Ambas ecuaciones tienen infinitas soluciones y si (x_1, y_1) satisface la primera ecuación y (x_2, y_2) satisface la segunda, entonces $(x_1 + x_2, y_1 + y_2) = (x_1, y_1) + (x_2, y_2)$ satisface la segunda ecuación.

Es importante destacar que el sistema de ecuaciones

$$\begin{cases} 2x - y = 0 \\ 2x - y = 6 \end{cases}$$

no tiene solución.

En general, el conjunto de soluciones de una ecuación lineal en dos variables corresponde con una recta en el plano; si la ecuación es homogénea, la recta pasa por el origen y es un subespacio; si la ecuación no es homogénea, la solución es una recta paralela a la recta que corresponde con la homogénea y no es un subespacio (es lo que se suele llamar un *subespacio afín*).

Ejemplo 7.2. Consideremos ahora el siguiente sistema de ecuaciones

$$\begin{cases} x + y = 4 \\ 2x - 5y = -6 \end{cases}$$

Una forma de resolverlo es multiplicar la primera ecuación por -2 y sumarla a la segunda:

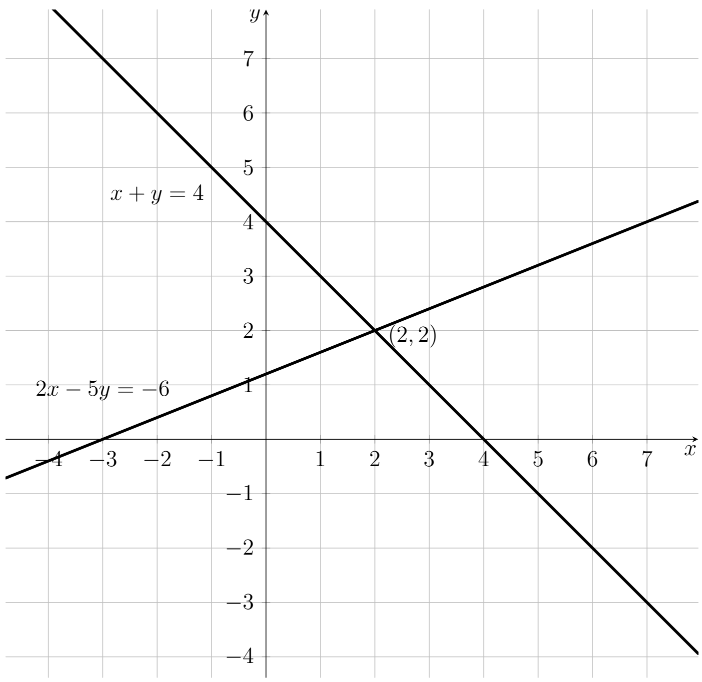
$$-2(x+y) + 2x - 5y = (-2)(4) - 6 \quad -2x - 2y + 2x - 5y = -8 - 6 \quad -7y = -14 \quad y = 2,$$

sustituyendo en la primera

$$x + 2 = 4 \qquad x = 4 - 2 = 2.$$

La solución del sistema es única y corresponde con el punto $(2,2)$ del plano.

La solución de cada una de las ecuaciones corresponde con una recta en el plano y, la solución del sistema corresponde con la intersección del par de rectas que corresponden a cada ecuación, gráficamente:



Ejemplo 7.3. Consideremos el sistema de ecuaciones

$$\begin{cases} x - y = 3 \\ 3x - 3y = 9 \end{cases}$$

En este caso ambas ecuaciones son equivalente (la segunda es igual a la primera multiplicada por 3), por lo tanto, el sistema tiene infinitas soluciones y equivale a la ecuación $x - y = 3$.

Observación 7.4. Generalizando lo que acabamos de observar, tenemos lo siguiente:

(1) Una ecuación lineal en dos variables posee infinitas soluciones, el conjunto de las soluciones corresponde con una recta en el plano.
   (a) Si la ecuación es homogénea, la recta que corresponde con la solución pasa por el origen y el conjunto de las soluciones es un subespacio de $\mathbb{R}^2$.
   (b) Si la ecuación no es homogénea, la recta que corresponde con la solución no pasa por el origen y el conjunto de las soluciones no es un subespacio de $\mathbb{R}^2$.
(2) Un sistema de dos ecuaciones en dos variables puede
   (a) Tener infinitas soluciones, en este caso las rectas asociadas a cada ecuación coinciden.
   (b) Tener una solución única, en este caso las rectas asociadas a la ecuaciones se intersecan en un punto.
   (c) No tener solución, en este caso las rectas asociadas a la ecuaciones son paralelas.

Ejemplo 7.5. Un problema de navegación.

Un bote navega a velocidad constante por un río. Recorre 15 kilómetros en una hora y media a favor de la corriente y recorre 12 kilómetros en 2 horas contra la corriente. Hallar la velocidad del bote en agua tranquila y la velocidad del río.

Solución:

Sean

$$\begin{aligned} x &= \text{velocidad del bote} \\ y &= \text{velocidad del río} \\ v_f &= \text{velocidad del bote a favor de la corriente} \\ v_c &= \text{velocidad del bote contra la corriente} \end{aligned}$$

entonces

$$\begin{cases} x + y = v_f, \\ x - y = v_c. \end{cases}$$

Mediremos el tiempo en horas y la velocidad en kilómetros por hora. Tenemos que

$$v_f = \frac{15}{1,5} = 10$$

$$v_c = \frac{12}{2} = 6$$

Luego

$$\begin{cases} x + y = 10, \\ x - y = 6. \end{cases}$$

Resolviendo el sistema hallamos $x = 8$, $y = 2$.

Luego la velocidad del bote en agua tranquila es 8 kilómetros por hora y la velocidad del río es 2 kilómetros por hora.

Ejercicios.

(1) Encontrar la(s) solución(es) (si las hay) de los siguientes sistemas.

(a) $\begin{cases} x + y = 3 \\ x + 2y = -8 \end{cases}$ (b) $\begin{cases} 2x + 3y = 3 \\ -2x - 3y = -3 \end{cases}$ (c) $\begin{cases} 4x + 5y = 0 \\ -2x - y = 3 \end{cases}$

(d) $\begin{cases} -2x = 1 \\ 4x - 3y = 0 \end{cases}$ (e) $\begin{cases} y = -3 \\ -2x + 4y = 8 \end{cases}$ (f) $\begin{cases} 7x + 4y = 0 \\ 2x + 4y = 0 \end{cases}$

(2) En un zoológico hay aves (de dos patas) y cuadrúpedos (de cuatro patas). Si el zoológico contiene 60 cabezas y 200 patas, ¿cuántas aves y cuantos cuadrúpedos viven en él?

(3) La compañía Sunrise Porcelain fabrica tazas y platos de cerámica. Para cada taza o plato un trabajador mide una cantidad fija de material y la pone en la máquina que los forma, de donde pasa al vidriado y secado automático. En promedio, un trabajador necesita tres minutos para iniciar el proceso de una taza y dos minutos para el de un plato. El material para una taza cuesta $\$0,25$ y el material para un plato cuesta $\$0,20$. Si se asignan $\$44$ diarios para la producción de tazas y platos, ¿cuántos deben fabricarse de cada uno en

un día de trabajo de 8 horas, si un trabajador se encuentra trabajando cada minuto y se gastan exactamente \$44 en materiales?

7.2. Ecuaciones lineales en varias variables

Lo estudiado en dos variables en la sección anterior se generaliza, de manera natural, a varias variables.

Una *ecuación lineal* en las variables $x_1, \dots, x_n$ es una ecuación que puede escribirse en la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{7.1}$$

donde los coeficientes $a_1, \dots, a_n$ y b son números reales, que generalmente se conocen de antemano.

Si $b = 0$ se dice que la ecuación (7.1) es *homogénea* , en otro caso se dice que es *no homogénea*.

El subíndice n puede ser cualquier entero positivo. En los ejemplos a estudiar n variará está entre 2 y 5. En problemas de la vida real, n podría ser 50 o 5000, o incluso mucho mayor.

EJEMPLO 7.6. Las ecuaciones

$$5x_1 - 3x_2 + 2 = x_1 \quad \text{y} \quad x_1 = 2(\sqrt{5} - x_2) + x_3$$

son lineales porque se pueden reordenar algebraicamente en la forma de la ecuación (7.1):

$$4x_1 - 3x_2 = -2 \quad \text{y} \quad x_1 + 2x_2 - x_3 = 2\sqrt{5}$$

Las ecuaciones

$$2x_1 - 5x_2 = x_1^2x_2 \quad \text{y} \quad x_2 = 2\sqrt{x_2} - 6$$

no son lineales debido a la presencia de $x_1^2x_2$ en la primera ecuación y de $\sqrt{x_2}$ en la segunda.

Un sistema de ecuaciones lineales (o sistema lineal) es una colección de una o más ecuaciones lineales que contienen las mismas variables, por ejemplo, $x_1, \dots, x_n$.

Por ejemplo, el sistema

$$\begin{cases} 2x_1 - x_2 + 1.5x_3 = 8 \\ x_1 - 4x_3 = -7 \end{cases}$$

es un sistema de dos ecuaciones en las variables x_1, x_2, x_3 (3 incógnitas).

Más precisamente.

DEFINICIÓN 7.7. Sean m y n enteros positivos. Un *sistema de m ecuaciones lineales* en n variables (o incógnitas) $x_1, x_2, \dots x_n$ es un conjunto de ecuaciones de la forma

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1, \\ \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{7.2}$$

donde a_{ij} $1 \leq i \leq m$, $1 \leq j \leq n$, b_k, $1 \leq k \leq m$ son números reales.

Una *solución* del sistema es una lista de números $(s_1, s_2, \dots, s_n)$ que da validez a cada ecuación cuando se utilizan los valores $s_1, \dots, s_n$ en lugar de $x_1, \dots, x_n$, respectivamente.

El conjunto de todas las posibles soluciones se llama *conjunto solución* del sistema lineal. Se dice que dos sistemas lineales son *equivalentes* si tienen el mismo conjunto solución. Es decir, cada solución del primer sistema es una solución del segundo sistema, y cada solución del segundo sistema también es una solución del primero.

El sistema se dice *homogéneo* si $b_1 = b_2 = \cdots = b_m = 0$, en otro caso se dice *no homogéneo.*

Al igual que en el caso de dos variables, un sistema de ecuaciones lineales puede tener

(1) ninguna solución, o
(2) exactamente una solución, o
(3) un número infinito de soluciones.

Se dice que un sistema de ecuaciones lineales es *consistente* si tiene una solución o un número infinito de soluciones; un sistema es *inconsistente* cuando no tiene ninguna solución.

7.3. Matrices y su relación con los sistemas de ecuaciones lineales

Definición 7.8. Sean m y n enteros mayores o iguales que 1.

Una *matriz* $m \times n$ es un arreglo rectangular de $m.n$ números dispuestos en m filas y n columnas

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

La igualdad anterior la abreviaremos mediante $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ o simplemente $A = (a_{ij})$.

A los números a_{ij} se les llama *entradas* o *coeficientes* de la matriz A.

Observación 7.9. Formalmente una *matriz* es una función de

$$A : \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{R}.$$

La imagen de (i, j), $A(i, j)$ se denota mediante a_{ij}.

Dos matrices A y B son iguales si son del mismo tamaño y sus componentes correspondientes son iguales. Es decir, $A = B$ cuando

$$a_{ij} = b_{ij}$$

para todo $i = 1, \dots, m$, $j = 1, \dots, n$.

Una *matriz cuadrada* es una matriz en la que el número de filas es igual al número de columnas $(m = n)$.

La matriz que tiene todas sus componentes iguales a cero, se denomina *matriz nula* .

Cuando tenemos la matriz

$$\begin{array}{cc}
 & \begin{array}{ccccc} \phantom{a_{11}} & & j & & \phantom{a_{1n}} \end{array} \\
\begin{array}{c} \\ \\ i \\ \\ \\ \end{array} &
\begin{pmatrix}
a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\
\vdots & \vdots\vdots\vdots & \vdots & \vdots\vdots\vdots & \vdots \\
a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\
\vdots & \vdots\vdots\vdots & \vdots & \vdots\vdots\vdots & \vdots \\
a_{m1} & \cdots & a_{mj} & \cdots & a_{mn}
\end{pmatrix}
\end{array}$$

Al vector

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

se le denomina *fila* i o *renglón* i.

Al vector

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

se le denomina *columna* j.

Los siguientes casos particulares destacan:

Un *vector fila* (o *vector renglón*) es una matriz de la forma

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix}$$

Un *vector columna* es una matriz de la forma

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}$$

EJEMPLO 7.10.

(1) Un vector fila 1×3 es una matriz 1×3:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix}.$$

(2) Un vector columna 3×1 es una matriz 3×1:

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}.$$

(3) La matriz nula 3×2 es:

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

(4) La matriz nula 4×4 es:

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

La información esencial de un sistema de ecuaciones lineales puede registrarse de forma compacta en una matriz.

Dado el sistema

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases} \tag{7.3}$$

con los coeficientes de cada variable alineados en columnas, la matriz

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

se denomina *matriz de coeficientes* del sistema (7.3), y la matriz

$$\begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ -4 & 5 & 9 & | & -9 \end{bmatrix}$$

se denomina *matriz aumentada* del sistema (7.3). (Aquí la segunda fila contiene un cero porque la segunda ecuación podría escribirse como $0 \cdot x_1 + 2x_2 - 8x_3 = 8$).

Notar que la matriz aumentada de un sistema consiste es la matriz de coeficientes con una columna adicional que contiene las constantes de los miembros derechos de las ecuaciones.

En este ejemplo, la matriz aumentada tiene 3 filas y 4 columnas, por lo que es una matriz 3×4.

La notación matricial es muy útil para simplificar cálculos, tal como se verá en lo que sigue.

Ejemplo 7.11. Volvamos a la ecuación sencilla del Ejemplo 7.2

La ecuación era

$$\begin{cases} x + y = 4 \\ 2x - 5y = -6 \end{cases}$$

La matriz aumentada del sistema es

$$\begin{bmatrix} 1 & 1 & | & 4 \\ 2 & -5 & | & -6 \end{bmatrix}$$

La forma en que se resolvió el sistema fue la siguiente:

Se multiplicó la primera ecuación por -2 y s se sumó a la segunda:

$$-2(x+y) + 2x - 5y = (-2)(4) - 6 \quad -2x - 2y + 2x - 5y = -8 - 6 \quad -7y = -14 \quad y = 2,$$

sustituyendo en la primera

$$x + 2 = 4 \qquad x = 4 - 2 = 2.$$

Este procedimiento se ha podido realizar en la matriz aumentada del sistema, lo que ahorra tiempo, espacio y simplifica las operaciones al no tener que escribir cada vez las incógnitas. En detalle, en la matriz del sistema multiplicar pr -2 la primera fila y sumarla a la fila 2, se obtiene la matriz

$$\begin{bmatrix} 1 & 1 & | & 4 \\ 0 & -7 & | & -14 \end{bmatrix}$$

después dividir fila 2 entre -7, para obtener

$$\begin{bmatrix} 1 & 1 & | & 4 \\ 0 & 1 & | & 2 \end{bmatrix} \tag{7.4}$$

Esta matriz corresponde al sistema equivalente

$$\begin{cases} x + y = 4 \\ \quad\;\; y = 2 \end{cases}$$

De aquí ya se deduce que $y = 2$ y, sustituyendo en la primera ecuación, $x + 2 = 2$, de donde $x = 2$.

También es posible simplificar aún más el trabajo de despeje usando matrices, si en la matriz (7.4) restamos la fila 2 a la fila 1, obtenemos la matriz

$$\begin{bmatrix} 1 & 0 & | & 2 \\ 0 & 1 & | & 2 \end{bmatrix}$$

que corresponde al sistema equivalente

$$\begin{cases} x \qquad = 2 \\ \quad\;\; y = 2 \end{cases}$$

de donde es inmediato que $x = y = 2$.

Las matrices se pueden utilizar en la forma anterior para resolver un sistema de ecuaciones, gracias a las siguientes reglas que aplican a todo sistema de ecuaciones:

(1) Si se intercambian dos ecuaciones no se altera la solución.
(2) Cuando se multiplican o se dividen ambos lados de una ecuación por un número distinto de cero, se obtiene una ecuación equivalente.
(3) Si se suma el múltiplo de una ecuación a otra del mismo sistema se obtiene un sistema ecuaciones equivalente.

Ejemplo 7.12. Veamos un ejemplo algo más complejo, consideremos el sistema

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ \qquad 2x_2 - 8x_3 = 8 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases}$$

Se va a mostrar el procedimiento de eliminación, con y sin notación matricial, y los resultados se colocarán uno al lado del otro para facilitar la comparación.

La matriz aumentada del sistema es

$$\begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ -4 & 5 & 9 & | & -9 \end{bmatrix}$$

Colocaremos cada sistema equivalente a la izquierda y la matriz correspondiente a la derecha

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ -4 & 5 & 9 & | & -9 \end{bmatrix}$$

Paso 1: Mantener x_1 en la primera ecuación y eliminarlo en las otras ecuaciones. Para hacerlo, sumar la ecuación 1 multiplicada por 4 a la ecuación 3.

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ 0 & -3 & 13 & | & -9 \end{bmatrix}$$

Multiplicar la ecuación 2 por $\frac{1}{2}$

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 1 & -4 & | & 4 \\ 0 & -3 & 13 & | & -9 \end{bmatrix}$$

Utilizar x_2 de la ecuación 2 para eliminar $-3x_2$ en la ecuación 3: multiplicar fila 2 por 3 y sumar a fila 3.

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 1 & -4 & | & 4 \\ 0 & 0 & 1 & | & 3 \end{bmatrix}$$

Aquí ya tenemos $x_3 = 3$, $x_2 = 4 + 4x_3 = 16$, $x_1 = 2x_2 - x_3 = 29$.

Pero también es posible continuar simplificando:

Utilizar x_2 de la ecuación 2 para eliminar $-2x_2$ en la ecuación 1: multiplicar fila 2 por 2 y sumar a fila 1.

$$\begin{aligned} x_1 - 7x_3 &= 8 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & 0 & -7 & | & 8 \\ 0 & 1 & -4 & | & 4 \\ 0 & 0 & 1 & | & 3 \end{bmatrix}$$

Utilizar x_3 de la ecuación 3 para eliminar $-7x_3$ en la ecuación 1 y $-4x_3$ en la ecuación 2: Multiplicar fila 3 por 7 y sumar a fila 1, multiplicar ecuación 3 por 4 y sumar a fila 2

$$x_1 = 29$$
$$x_2 = 16$$
$$x_3 = 3$$

$$\begin{bmatrix} 1 & 0 & 0 & | & 29 \\ 0 & 1 & 0 & | & 16 \\ 0 & 0 & 1 & | & 3 \end{bmatrix}$$

De donde $x_1 = 29$, $x_2 = 16$ y $x_3 = 29$.

7.4. Matrices escalonadas y matrices reducidas

En la sección anterior vimos lo útil que resulta utilizar matrices, operando sobre sus filas para resolver sistemas de ecuaciones lineales.

En esta sección vamos a precisar los conceptos referentes a las operaciones que se pueden realizar en las filas de las matrices para resolver ecuaciones lineales.

Definición 7.13. Las *operaciones elementales fila o renglón* para matrices son las siguientes:

(1) Multiplicar (o dividir) una fila por un número diferente de cero.
(2) Sumar un múltiplo de una fila a otra fila.
(3) Intercambiar dos filas.

Para $c \neq 0$, estas operaciones se denotarán de la siguiente manera:

(1) $R_i \to cR_i$ significa "se sustituye la i-ésima fila por esa misma fila multiplicada por c".
(2) $R_i \to R_i + cR_j$ significa "se sustituye la i-ésima fila por la suma de la fila i más la fila j multiplicada por c".
(3) $R_i \leftrightarrow R_j$ significa "se intercambian las filas i y j".

El proceso de aplicar estas operaciones elementales con filas para simplificar una matriz se llama *reducción por filas*.

Definición 7.14. Sean A y B dos matrices $m \times n$. Se dice que B es *equivalente por filas* a la matriz A (o simplemente *equivalente* a A), si B se puede obtener de la matriz A a través de una sucesión finita de operaciones elementales de renglón. Si B es equivalente a A se abrevia $B \sim A$.

Ejemplo 7.15. Si

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

y

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -7 & -7 & -8 & -9 \end{bmatrix},$$

$B \sim A$; pues B se obtiene de A mediante la operación de renglón

$$R_2 \to -2R_1 + R_2$$

La relación entre matrices $\sim$ es una relación de equivalencia, es decir, si A, B y C son matrices $m \times n$, se cumple:

(1) $A \sim A$. (Reflexividad)
(2) Si $A \sim B$ entonces $B \sim A$. (Simetría)
(3) Si $A \sim B$ y $B \sim C$ entonces $A \sim C$. (Transitividad)

DEFINICIÓN 7.16. Se dice que una matriz es *escalonada por filas* si satisface las siguientes condiciones:

(1) Todas las filas cuyos elementos son todos cero (si las hay) aparecen en la parte inferior de la matriz.
(2) El primer número diferente de cero (a partir de la izquierda) en cualquier fila, que no contiene sólo ceros, es 1.
(3) Si dos filas sucesivas no contienen solamente ceros, entonces el primer 1 en la fila inferior ocurre más a la derecha que el primer 1 en el renglón superior.

Se dice que una matriz es *escalonada reducida* si es escalonada por filas y además satisface la siguiente condición adicional:

(4) Cualquier columna que contenga el primer 1 de un renglón tiene ceros en las demás posiciones.

El primer número diferente de cero de una fila (si lo hay) se llama pivote de esa fila.

A través de operaciones elementales fila se puede llevar una matriz a una matriz equivalente que es escalonada por filas y también a una matriz equivalente que es escalonada reducida.

EJEMPLO 7.17. Las siguientes matrices se encuentran en la forma escalonada por renglones:

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

OBSERVACIÓN 7.18. Por lo general, la forma escalonada por renglones de una matriz no es única. Es decir, una matriz puede ser equivalente a más de una matriz en forma escalonada por renglones.

Por ejemplo, las siguientes matrices son diferentes y son equivalentes por renglones

$$\begin{bmatrix} 1 & 3 & 2 & 5 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \to R_1 - 3R_2} \begin{bmatrix} 1 & 0 & -7 & -13 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

EJEMPLO 7.19. Las siguientes matrices están en la forma escalonada reducida por renglones:

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

Las dos primeras tienen tres pivotes; las otras tres tienen dos pivotes.

7.4.1. Ejercicios.

(1) Determine si la matriz dada se encuentra en la forma escalonada por renglones (pero no en la forma escalonada reducida por renglones), en la forma escalonada reducida por renglones o en ninguna de las dos.

(a) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 5 & 2 \\ 0 & -2 & 5 \\ 0 & 0 & 3 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 1 & 4 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & 4 \end{pmatrix}$ (i) $\begin{pmatrix} 1 & 5 & 2 \\ 0 & 1 & 5 \end{pmatrix}$

(2) Utilizar operaciones elementales con renglones para reducir las matrices dadas a la forma escalonada por renglones y a la forma escalonada reducida por renglones.

(a) $\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 6 \\ 4 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & -1 & 1 \\ 2 & 4 & 3 \\ 5 & 6 & -2 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \\ -1 & 1 & 1 \end{pmatrix}$ (e) $\begin{pmatrix} 2 & -4 & 8 \\ 3 & 5 & 8 \\ -6 & 0 & 4 \end{pmatrix}$ (f) $\begin{pmatrix} 2 & -4 & -2 \\ 3 & 1 & 6 \end{pmatrix}$

(g) $\begin{pmatrix} 3 & -6 & -3 \\ 5 & 10 & 5 \end{pmatrix}$ (h) $\begin{pmatrix} 2 & -7 \\ 3 & 5 \\ 4 & -3 \end{pmatrix}$ (i) $\begin{pmatrix} 1 & 5 & 2 \\ -3 & -14 & -1 \end{pmatrix}$

7.5. Métodos de Gauss y de Gauss-Jordan

7.5.1. Método de Gauss. El método de Gauss se utiliza para obtener una matriz escalonada equivalente a una matriz dada.

A continuación ilustramos el método de Gauss mediante un ejemplo

EJEMPLO 7.20. Sea

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

Obtener una matriz escalonada equivalente a A

Paso 1: Se busca una fila en A que tenga su primer elemento distinto de cero y se intercambia (si es necesario) con la primera fila de la matriz A; si no existe una fila de A que tenga su primer elemento no nulo, entonces se busca una fila de la matriz A que tenga el segundo elemento distinto de cero y se intercambia (si es necesario) con la primera fila de la matriz A; de no suceder así, se busca una fila de A que tenga el tercer elemento distinto de cero y se intercambia (si es necesario) con la primera fila de A, etc.; obteniendo finalmente una matriz $B_1 \sim A$ con un primer elemento no nulo en la primera fila que es el pivote. Dividiendo esta fila entre el pivote la fila queda con un 1 al principio.

Para este caso particular la operación a realizar es $R_1 \leftrightarrow R_3$, resultando la equivalencia de matrices

$$\begin{array}{ccc} A & \sim & B_1 \\ \begin{bmatrix} 0 & -4 & -1 & 3 \\ 3 & 4 & 0 & 7 \\ 1 & 1 & 3 & 5 \end{bmatrix} & \sim & \begin{bmatrix} 1 & 1 & 3 & 5 \\ 3 & 4 & 0 & 7 \\ 0 & -4 & -1 & 3 \end{bmatrix} \end{array}.$$

El pivote de la primera fila de la matriz B_1 es $b^1_{11} = 1$.

Paso 2: Con el pivote de la primera fila de B_1 se transforman en ceros los elementos que están por debajo de él mediante la operación suma de filas, obteniendo una matriz $B_2 \sim B_1 \sim A$, que tendrá todas las componentes nulas debajo del pivote de la primera fila.

Para este caso particular la operación a realizar es $R_2 \to -3R_1 + R_2$, resultando la equivalencia de matrices

$$\begin{array}{ccc} B_1 & \sim & B_2 \\ \begin{bmatrix} 1 & 1 & 3 & 5 \\ 3 & 4 & 0 & 7 \\ 0 & -4 & -1 & 3 \end{bmatrix} & \sim & \begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & -4 & -1 & 3 \end{bmatrix} \end{array}$$

En caso de no haber obtenido un 1 en el segundo lugar de la fila 2, debemos hacer la división que corresponda para obtener una matriz equivalente con 1 en ese lugar.

Paso 3: Ahora se repiten los pasos 1 y 2 con la segunda fila de la matriz B_2, produciendo una matriz $B_3 \sim B_2 \sim B_1 \sim A$ cuyas componentes serán nulas debajo del pivote de su segunda fila.

Para el caso particular ilustrado, el pivote de la segunda fila de la matriz B_2 es $b^2_{22} = 1$. Se pueden hacer ceros los elementos debajo del mismo mediante la operación $R_3 \to 4R_2 + R_3$, esto es

$$\begin{array}{ccc} B_2 & \sim & B_3 \\ \begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & -4 & -1 & 3 \end{bmatrix} & \sim & \begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & 0 & -37 & -29 \end{bmatrix} \end{array}$$

Paso 4: Continuamos de manera análoga con las filas restantes.

Para el caso particular considerado, se debe dividir la tercera fila entre -37

$$\begin{array}{ccc} B_2 & \sim & B_3 \\ \begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & -4 & -1 & 3 \end{bmatrix} & \sim & \begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & 0 & 1 & \frac{29}{37} \end{bmatrix} \end{array}$$

Para el caso particular ilustrado, el pivote de la segunda fila de la matriz B_2 es $b^2_{22} = 1$. Se pueden hacer ceros los elementos debajo del mismo mediante la operación $R_3 \leftrightarrow 4R_2 + R_3$, esto es

$$\begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & -4 & -1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & 0 & -37 & -29 \end{bmatrix}$$

Para el caso particular considerado, la matriz B_3 ya está en forma escalonada; con lo que terminaría el proceso para este ejemplo particular.

Para el caso general, se repiten los pasos 1, 2 y 3 con las filas subsecuentes de las matrices equivalentes que resulten, hasta obtener una matriz H en forma escalonada.

Observación 7.21. Debe quedar claro que el propósito fundamental del método de Gauss es obtener una matriz en forma escalonada equivalente a una matriz dada, mediante el uso de las operaciones elementales de renglón en cualquier combinación. Así que el algoritmo anterior sólo es una guía para este propósito. Cualquier modificación es válida siempre y cuando se empleen únicamente las operaciones de renglón para matrices y se alcance el objetivo de obtener una matriz en forma escalonada equivalente por filas a la matriz inicial.

A continuación otro ejemplo, presentado en forma más esquemática

Ejemplo 7.22. Usar el método de Gauss para obtener una matriz escalonada por filas equivalente a la matriz

$$\begin{bmatrix} 3 & 6 & 12 \\ 5 & 11 & 23 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 6 & 12 \\ 5 & 11 & 23 \end{bmatrix} \xrightarrow{R_1 \to \frac{1}{3} R_1} \begin{bmatrix} 1 & 2 & 4 \\ 5 & 11 & 23 \end{bmatrix} \xrightarrow{R_2 \to R_2 - 5R_1} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \end{bmatrix}$$

7.5.2. Método de Gauss-Jordan. El método de Gauss se utiliza para obtener una matriz escalonada reducida equivalente a una matriz dada y se procede de la siguiente manera.

(1) Se lleva la matriz a forma escalonada mediante el método de Gauss.
(2) Se hacen ceros todos los elementos arriba de cada pivote utilizando operaciones de renglón.

Lo ilustraremos mediante un ejemplo.

Ejemplo 7.23. Utilizar el método de Gauss-Jordan para obtener una matriz escalonada reducida equivalente a la matriz

$$\begin{bmatrix} 1 & 3 & 2 & 25 \\ 1 & 4 & 1 & 20 \\ 2 & 5 & 5 & 55 \end{bmatrix}$$

$$\begin{bmatrix}1&3&2&25\\1&4&1&20\\2&5&5&55\end{bmatrix}\xrightarrow{R_2\to R_2-R_1}\begin{bmatrix}1&3&2&25\\0&1&-1&-5\\2&5&5&55\end{bmatrix}$$

$$\xrightarrow{R_3\to R_3-2R_1}\begin{bmatrix}1&3&2&25\\0&1&-1&-5\\0&-1&1&5\end{bmatrix}$$

$$\xrightarrow{R_1\to R_1-3R_2}\begin{bmatrix}1&0&5&40\\0&1&-1&-5\\0&-1&1&5\end{bmatrix}$$

$$\xrightarrow{R_3\to R_3+R_2}\begin{bmatrix}1&0&5&40\\0&1&-1&-5\\0&0&0&0\end{bmatrix}$$

7.5.3. Aplicación a la resolución de sistemas de ecuaciones lineales. Los métodos de Gauss y de Gauss-Jordan se aplican a la resolución de sistemas de ecuaciones lineales, transformando la matriz aumentada del sistema en una matriz escalonada por filas o en una matriz reducida por filas, respectivamente.

Los métodos serán explicados en base a ejemplos.

Ejemplo 7.24. Consideremos el sistema

$$\begin{cases}2x_1+4x_2+6x_3=18\\4x_1+5x_2+6x_3=24\\3x_1+\ x_2\ -2x_3=\ 4\end{cases}$$

La matriz aumentada del sistema es

$$\begin{bmatrix}2&4&6&|&18\\4&5&6&|&24\\3&1&-2&|&4\end{bmatrix}$$

Primero vamos resolver utilizando el método de Gauss, es decir, buscaremos una matriz escalonada por filas que sea equivalente a la matriz aumentada del sistema.

$$\begin{bmatrix}2&4&6&|&18\\4&5&6&|&24\\3&1&-2&|&4\end{bmatrix}\xrightarrow{R_1\to\frac{1}{2}R_1}\begin{bmatrix}1&2&3&|&9\\4&5&6&|&24\\3&1&-2&|&4\end{bmatrix}$$

$$\xrightarrow{R_2\to R_2-4R_1}\begin{bmatrix}1&2&3&|&9\\0&-3&-6&|&-12\\3&1&-2&|&4\end{bmatrix}$$

$$\xrightarrow{R_3\to R_3-3R_1}\begin{bmatrix}1&2&3&|&9\\0&-3&-6&|&-12\\0&-5&-11&|&-23\end{bmatrix}$$

$$
\xrightarrow{R_2 \to -\frac{1}{3}R_2} \begin{bmatrix} 1 & 2 & 3 & | & 9 \\ 0 & 1 & 2 & | & 4 \\ 0 & -5 & -11 & | & -23 \end{bmatrix}
\xrightarrow{R_1 \to R_1 - 2R_2} \begin{bmatrix} 1 & 0 & -1 & | & 1 \\ 0 & 1 & 2 & | & 4 \\ 0 & -5 & -11 & | & -23 \end{bmatrix}
\xrightarrow{R_3 \to R_3 + 5R_2} \begin{bmatrix} 1 & 0 & -1 & | & 1 \\ 0 & 1 & 2 & | & 4 \\ 0 & 0 & -1 & | & -3 \end{bmatrix}
\xrightarrow{R_3 \to -R_3} \begin{bmatrix} 1 & 0 & -1 & | & 1 \\ 0 & 1 & 2 & | & 4 \\ 0 & 0 & 1 & | & 3 \end{bmatrix}
$$

Ya tenemos una matriz escalonada de la que se deduce el siguiente sistema equivalente

$$
\begin{cases} x_1 \quad\; - x_3 = 1 \\ \quad x_2 + 2x_3 = 4 \\ \qquad\quad x_3 = 3 \end{cases}
$$

Por lo tanto, $x_3 = 3$, $x_2 = 4 - 2x_3 = -2$, $x_1 = 1 + x_3 = 4$.

Si queremos resolver el sistema por el método de Gauss-Jordan, deberíamos continuar de la siguiente manera

$$
\begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}
\xrightarrow{R_1 \to R_1 + R_3} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}
\xrightarrow{R_2 \to R_2 - 2R_3} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{bmatrix}
$$

de donde se obtiene inmediatamente la solución, ya que esta última matriz corresponde son el sistema equivalente

$$
\begin{cases} x_1 = 4 \\ x_2 = -2 \\ x_3 = 3 \end{cases}
$$

Ejemplo 7.25. Consideremos el sistema

$$
\begin{cases} 2x_1 + 4x_2 + 6x_3 = 18 \\ 4x_1 + 5x_2 + 6x_3 = 24 \\ 2x_1 + 7x_2 + 12x_3 = 30 \end{cases}
$$

La matriz aumentada del sistema es

$$
\begin{bmatrix} 2 & 4 & 6 & | & 18 \\ 4 & 5 & 6 & | & 24 \\ 2 & 7 & 12 & | & 30 \end{bmatrix}
$$

Por el método de Gauss:

$$
\begin{aligned}
\left[\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 30 \end{array}\right]
&\xrightarrow{R_1 \to \frac{1}{2}R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 30 \end{array}\right] \\
&\xrightarrow{R_2 \to R_2 - 4R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 2 & 7 & 12 & 30 \end{array}\right] \\
&\xrightarrow{R_3 \to R_3 - 2R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & 3 & 6 & 12 \end{array}\right] \\
&\xrightarrow{R_2 \to -\frac{1}{3}R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 6 & 12 \end{array}\right] \\
&\xrightarrow{R_1 \to R_1 - 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 6 & 12 \end{array}\right] \\
&\xrightarrow{R_3 \to R_3 - 3R_2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]
\end{aligned}
$$

Esto es equivalente al sistema de ecuaciones

$$
\begin{cases} x_1 - x_3 = 1 \\ x_2 + 2x_3 = 4 \end{cases}
$$

Hasta aquí se puede llegar. Se tienen sólo dos ecuaciones para las tres incógnitas x_1, x_2, x_3 y existe un número infinito de soluciones. Para comprobar esto se elige un valor de x_3. Entonces $x_2 = 4 - 2x_3$ y $x_1 = 1 + x_3$. Ésta será una solución para cualquier número x_3.

El conjunto solución es un subespacio de $\mathbb{R}^3$, dado por

$$S = \{(1 + t, 4 - 2t, t) \,:\, t \in \mathbb{R}\}.$$

En cuanto al método de Gauss-Jordan, en este caso los dos métodos coinciden.

Ejemplo 7.26. Consideremos el sistema

$$
\begin{cases} \phantom{2x_1 - {}} 2x_2 + 3x_3 = 4 \\ 2x_1 - 6x_2 + 7x_3 = 15 \\ x_1 - 2x_2 + 5x_3 = 10 \end{cases}
$$

La matriz aumentada para este sistema es

$$
\left[\begin{array}{ccc|c} 0 & 2 & 3 & 4 \\ 2 & -6 & 7 & 15 \\ 1 & -2 & 5 & 10 \end{array}\right]
$$

Por el método de Gauss:

$$\begin{bmatrix} 0 & 2 & 3 & | & 4 \\ 2 & -6 & 7 & | & 15 \\ 1 & -2 & 5 & | & 10 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & -2 & 5 & | & 10 \\ 2 & -6 & 7 & | & 15 \\ 0 & 2 & 3 & | & 4 \end{bmatrix}$$
$$\xrightarrow{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & -2 & 5 & | & 10 \\ 0 & -2 & -3 & | & -5 \\ 0 & 2 & 3 & | & 4 \end{bmatrix}$$

Aquí ya se observa que el sistema es inconsistente, ya que las dos últimas filas corresponden con el sistema de ecuaciones
$$\begin{cases} -2x_2 - 3x_3 = -5 \\ 2x_2 + 3x_3 = 4 \end{cases}$$
que claramente no tiene solución.

Sigamos con el método de Gauss a ver que ocurre
$$\begin{bmatrix} 1 & -2 & 5 & | & 10 \\ 0 & -2 & -3 & | & -5 \\ 0 & 2 & 3 & | & 4 \end{bmatrix} \xrightarrow{R_2 \to -\frac{1}{2} R_2} \begin{bmatrix} 1 & -2 & 5 & | & 10 \\ 0 & 1 & \frac{3}{2} & | & \frac{5}{2} \\ 0 & 2 & 3 & | & 4 \end{bmatrix}$$
$$\xrightarrow{R_3 \to R_3 - 2R_2} \begin{bmatrix} 1 & -2 & 5 & | & 10 \\ 0 & 1 & \frac{3}{2} & | & \frac{5}{2} \\ 0 & 0 & 0 & | & -1 \end{bmatrix}$$
$$\xrightarrow{R_1 \to R_1 + 2R_2} \begin{bmatrix} 1 & 0 & 8 & | & 15 \\ 0 & 1 & \frac{3}{2} & | & \frac{5}{2} \\ 0 & 0 & 0 & | & -1 \end{bmatrix}$$

Ahora la última ecuación es $0x_1 + 0x_2 + 0x_3 = -1$, lo cual también es imposible ya que $0 \neq -1$. Por lo tanto, el sistema no tiene solución, es decir es inconsistente.

Ahora surge la pregunta natural ¿Cuál de los métodos es mejor? ¿El de Gauss o el de Gauss-Jordan?

Depende; al resolver sistemas de ecuaciones en una computadora es preferible el método de eliminación gaussiana porque significa menos operaciones elementales por renglones.

De hecho, para resolver un sistema de n ecuaciones con n incógnitas usando la eliminación de Gauss-Jordan se requieren aproximadamente de $\frac{n^3}{2}$ sumas y multiplicaciones, mientras que la eliminación gaussiana requiere sólo $\frac{n^3}{2}$ n3 sumas y multiplicaciones.

Si es esencial obtener la forma escalonada reducida por renglones de una matriz corresponde el método de Gauss-Jordan

EJEMPLO 7.27. Resolver el sistema
$$\begin{cases} x_1 + 3x_2 - 5x_3 + x_4 = 4 \\ 2x_1 + 5x_2 - 2x_3 + 4x_4 = 6 \end{cases}$$

La matriz aumentada del sistema es
$$\begin{bmatrix} 1 & 3 & -5 & 1 & | & 4 \\ 2 & 5 & -2 & 4 & | & 6 \end{bmatrix}$$

Reduciendo por renglones

$$\begin{aligned}
\begin{bmatrix} 1 & 3 & -5 & 1 & 4 \\ 2 & 5 & -2 & 4 & 6 \end{bmatrix} &\xrightarrow{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & 3 & -5 & 1 & 4 \\ 0 & -1 & 8 & 2 & -2 \end{bmatrix} \\
&\xrightarrow{R_2 \to -R_2} \begin{bmatrix} 1 & 3 & -5 & 1 & 4 \\ 0 & 1 & -8 & -2 & 2 \end{bmatrix} \\
&\xrightarrow{R_1 \to R_1 - 3R_2} \begin{bmatrix} 1 & 0 & 19 & 7 & -2 \\ 0 & 1 & -8 & -2 & 2 \end{bmatrix}
\end{aligned}$$

Hasta aquí se puede llegar. La matriz de coeficientes se encuentra en forma escalonada reducida por renglones. Es evidente que existe un número infinito de soluciones. Los valores de las variables x_3 y x_4 se pueden escoger de manera arbitraria, y entonces $x_2 = 2+8x_3+2x_4$ y $x_1 = -2-19x_3-7x_4$.

Por lo tanto, todas las soluciones se representan en la forma $(-2-19x_3-7x_4, 2+8x_3+2x_4, x_3, x_4)$ donde $x - 3$ y x_4 son números reales.

Al resolver muchos sistemas, es evidente que los cálculos se vuelven fastidiosos. Un buen método práctico es usar una calculadora o computadora siempre que las fracciones se compliquen. Debe hacerse notar, sin embargo, que si los cálculos se llevan a cabo en una computadora o calculadora pueden introducirse errores de “redondeo”.

EJEMPLO 7.28 (Un problema de administración de recursos.).

Un departamento de pesca y caza del estado proporciona tres tipos de comida a un lago que alberga a tres especies de peces.

(i) Cada pez de la especie 1 consume cada semana un promedio de 1 unidad del alimento 1, 1 unidad del alimento 2 y 2 unidades del alimento 3.
(ii) Cada pez de la especie 2 consume cada semana un promedio de 3 unidades del alimento 1, 4 unidades del alimento 2 y 5 unidades del alimento 3.
(iii) Para un pez de la especie 3, el promedio semanal de consumo es 2 unidades del alimento 1, 1 unidad del alimento 2 y 5 unidades del alimento 3.
(iv) Cada semana se proporcionan al lago 25.000 unidades de alimento 1, se proporcionan 20.000 unidades del alimento 2 y 55.000 del alimento 3.

Si se supone que los peces se comen todo el alimento, ¿cuántos peces de cada especie pueden coexistir en el lago?

Solución:

Sean x_1, x_2 y x_3 el número de peces de cada especie que hay en el ambiente del lago. Utilizando la información del problema se observa que:

(i) x_1 peces de la especie 1 consumen x_1 unidades de alimento 1.
(ii) x_2 peces de la especie 2 consumen $3x_2$ unidades de alimento 1.
(iii) x_3 peces de la especie 3 consumen $2x_3$ unidades de alimento 1.

Entonces

$$x_1 + 3x_2 + 2x_3 = 25.000.$$

Análogamente se pueden obtenerlas ecuaciones para los otros dos alimentos.

Se llega al siguiente sistema de ecuaciones:

$$\begin{cases} x_1 + 3x_2 + 2x_3 = 25.000 \\ x_1 + 4x_2 + x_3 = 20.000 \\ 2x_1 + 5x_2 + 5x_3 = 55.000 \end{cases}$$

La matriz ampliada del sistema es

$$A = \begin{bmatrix} 1 & 3 & 2 & | & 25.000 \\ 1 & 4 & 1 & | & 20.000 \\ 2 & 5 & 5 & | & 55.000 \end{bmatrix}$$

Usamos el método de Gauss-Jordan para obtener la matriz escalonada reducida.

$$\begin{bmatrix} 1 & 3 & 2 & | & 25.000 \\ 1 & 4 & 1 & | & 20.000 \\ 2 & 5 & 5 & | & 55.000 \end{bmatrix} \xrightarrow{R_2 \to R_2 - R_1} \begin{bmatrix} 1 & 3 & 2 & | & 25.000 \\ 0 & 1 & -1 & | & -5.000 \\ 2 & 5 & 5 & | & 55.000 \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 - 2R_1} \begin{bmatrix} 1 & 3 & 2 & | & 25.000 \\ 0 & 1 & -1 & | & -5.000 \\ 0 & -1 & 1 & | & 5.000 \end{bmatrix}$$

$$\xrightarrow{R_1 \to R_1 - 3R_2} \begin{bmatrix} 1 & 0 & 5 & | & 40.000 \\ 0 & 1 & -1 & | & -5.000 \\ 0 & -1 & 1 & | & 5.000 \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 + 3R_2} \begin{bmatrix} 1 & 0 & 5 & | & 40.000 \\ 0 & 1 & -1 & | & -5.000 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

Hemos logrado una reducción del sistema de ecuaciones:

$$\begin{cases} x_1 + 5x_3 = 40.000 \\ x_2 - x_3 = -5.000 \end{cases}$$

Si x_3 se elige arbitrariamente se tiene un número infinito de soluciones dadas por:

$$\begin{cases} x_1 = 40.000 - 5x_3 \\ x_2 = x_3 - 5.000 \end{cases}$$

Pero debemos tomar en cuenta otras restricciones: Por supuesto se debe tener

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0.$$

De $x_2 \geq 0$ se sigue que $x_3 \geq 5.000$.

De $0 \leq x_1 = 40.000 - 5x_3$ se sigue que $x_3 \leq 8.000$.

Esto significa que las poblaciones que pueden convivir en el lago con todo el alimento consumido deben satisfacer:

$$\begin{cases} x_1 = 40.000 - 5x_3 \\ x_2 = x_3 - 5.000 \\ 5.000 \leq x_3 \leq 8.000 \end{cases}$$

Notar que el sistema de ecuaciones lineales que hemos resuelto tiene un número infinito de soluciones. Sin embargo el problema de administración de recursos tiene sólo un número finito de soluciones porque x_1, x_2 y x_3 deben ser enteros positivos y existen nada más 3.001 enteros en el intervalo [5.000, 8.000].

Ejemplo 7.29 (Análisis de redes).

Muchas situaciones prácticas originan redes: redes de transporte, redes de comunicaciones y redes económicas, por nombrar algunas. De particular interés son los posibles flujos a través de las redes. Por ejemplo, los vehículos que fluyen a través de una red de carreteras, la información que fluye a través de una red de datos, y los bienes y servicios que fluyen a través de una red económica.

En este texto, una red consistirá de un número finito de nodos (también llamados uniones o vértices) conectados mediante una serie de aristas dirigidas llamadas ramas o arcos. Cada rama se marcará con un flujo que representa la cantidad de algún objeto que puede fluir a lo largo o a través de cada rama en la dirección indicada. (Piense en los automóviles que viajan a lo largo de una red de calles de un sentido.) La regla fundamental que gobierna el flujo a través de una red es la conservación del flujo:

En cada nodo, el flujo de entrada es igual al flujo de salida.

Se puede analizar el flujo a través de toda una red planteando las ecuaciones adecuadas y resolviendo el sistema resultante de ecuaciones lineales.

El problema que se va a resolver es el siguiente:

Describir los posibles flujos a través de la red de tuberías de agua que se muestra en la Figura 7.1, donde el flujo se mide en litros por minuto.

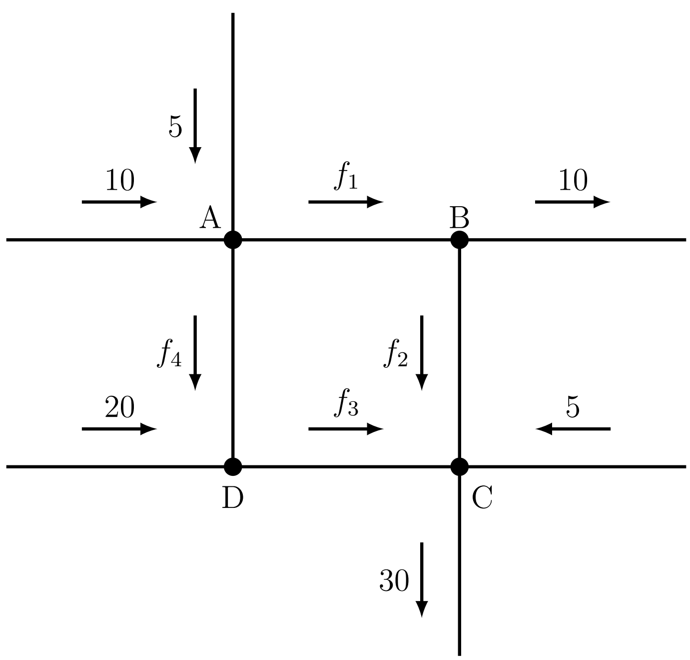


Figura 7.1. Flujo en tuberías

En cada nodo, se escribe la ecuación que representa la conservación del flujo. Luego se reescribe cada ecuación con las variables a la izquierda y la constante a la derecha, para obtener un sistema lineal en forma estándar.

Nodo A:	$15 = f_1 + f_4$	$\Rightarrow f_1 + f_4 = 15$
Nodo B:	$f_1 = f_2 + 10$	$\Rightarrow f_1 - f_2 = 10$
Nodo C:	$f_2 + f_3 + 5 = 30$	$\Rightarrow f_2 + f_3 = 25$
Nodo D:	$f_4 + 20 = f_3$	$\Rightarrow f_3 - f_4 = 20$

Usando eliminación de Gauss-Jordan, se reduce la matriz aumentada:

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 15 \\ 1 & -1 & 0 & 0 & 10 \\ 0 & 1 & 1 & 0 & 25 \\ 0 & 0 & 1 & -1 & 20 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 15 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & -1 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

(Comprobarlo a manera de ejercicio.) Se ve que existe una variable libre, f_4, de modo que se tiene un número infinito de soluciones. Al hacer $f_4 = t$ y expresar las variables pivote en términos de f_4, se obtiene

$$f_1 = 15 - t$$

$$f_2 = 5 - t$$

$$f_3 = 20 + t$$

$$f_4 = t$$

Estas ecuaciones describen todos los flujos posibles y permiten analizar la red. Por ejemplo, se ve que, si se controla el flujo en la rama AD de modo que $t = 5$ L/min, entonces los otros flujos son $f_1 = 10$, $f_2 = 0$, y $f_3 = 25$.

Puede hacerlo todavía mejor. Se pueden encontrar los posibles flujos mínimo y máximo en cada rama. Cada uno de los flujos debe ser no negativo. Al examinar la primera y segunda ecuaciones a la vez, se ve que $t \leq 15$ (de otro modo f_1 sería negativo) y $t \leq 5$ (de otro modo f_2 sería negativo). La segunda de estas desigualdades es más restrictiva que la primera, de modo que debe usarla. La tercera ecuación no aporta más restricciones al parámetro t, así se deduce que $0 \leq t \leq 5$. Al combinar este resultado con las cuatro ecuaciones, se ve que

$$10 \leq f_1 \leq 15$$

$$0 \leq f_2 \leq 5$$

$$20 \leq f_3 \leq 25$$

$$0 \leq f_4 \leq 5$$

Ahora se tiene una descripción completa de los posibles flujos a través de esta red.

7.6. Ejercicios varios

(1) Utilizar los métodos de eliminación de Gauss y de Gauss-Jordan para encontrar, si existen, todas las soluciones de los sistemas dados. Verificar con un programa informático.

(a) $\begin{cases} 9x_1 + 9x_2 - 7x_3 = 6 \\ -7x_1 - x_3 = -10 \\ 9x_1 + 6x_2 + 8x_3 = 45 \end{cases}$

(b) $\begin{cases} x_1 - 2x_2 + 3x_3 = 11 \\ 4x_1 + x_2 - x_3 = 4 \\ 2x_1 - x_2 + 3x_3 = 10 \end{cases}$

(c) $\begin{cases} 9x_2 - 7x_3 = 2 \\ -x_3 = -2 \\ -3x_1 + 6x_2 + 8x_3 = 1 \end{cases}$

(d) $\begin{cases} x_1 - 5x_2 = 0 \\ -x_1 + 5x_2 = 0 \end{cases}$

(e) $\begin{cases} 2x_1 - 3x_2 - 4x_3 + 5x_4 = 0 \\ 7x_2 + 3x_4 = 0 \end{cases}$

(f) $\begin{cases} 2x_1 - x_2 = 0 \\ 3x_1 + 5x_2 = 0 \\ 7x_1 - 3x_2 = 0 \\ -2x_1 + 3x_2 = 0 \end{cases}$

(g) $\begin{cases} -2x_1 + x_2 + 6x_3 = 18 \\ 5x_1 + 8x_3 = -16 \\ 3x_1 + 2x_2 - 10x_3 = -3 \end{cases}$

(h) $\begin{cases} 3x_1 + 6x_2 - 6x_3 = 9 \\ 2x_1 - 5x_2 + 4x_3 = 6 \\ 5x_1 + 28x_2 - 26x_3 = -8 \end{cases}$

(i) $\begin{cases} x_1 + x_2 - x_3 = 7 \\ 4x_1 - x_2 + 5x_3 = 4 \\ 2x_1 + 2x_2 - 3x_3 = 0 \end{cases}$

(j) $\begin{cases} x_1 + 2x_2 - 2x_3 - x_4 = 1 \\ -3x_1 + 4x_2 + x_3 - 2x_4 = 4 \\ -3x_1 + 14x_2 + 4x_3 - 7x_4 = 3 \\ 6x_1 + 12x_2 - 12x_3 - 6x_4 = 5 \end{cases}$

(k) $\begin{cases} x_1 + 2x_2 - 4x_3 = 4 \\ -2x_1 - 4x_2 + 8x_3 = -9 \end{cases}$

(l) $\begin{cases} x_1 - 2x_2 + x_3 + x_4 = 2 \\ 3x_1 + 2x_3 - 2x_4 = -8 \\ 4x_2 - x_3 - x_4 = 1 \\ 5x_1 + 3x_3 - x_4 = 0 \end{cases}$

(m) $\begin{cases} x_1 + x_2 = 4 \\ 2x_1 - 3x_2 = 7 \\ 3x_1 - 2x_2 = 11 \end{cases}$

(2) Un viajero que acaba de regresar de Europa gastó \$30 diarios en Inglaterra, \$20 diarios en Francia y \$20 diarios en España por concepto de hospedaje. En comida gastó \$20 diarios en Inglaterra, \$30 diarios en Francia y \$20 diarios en España. Sus gastos adicionales fueron de \$10 diarios en cada país. Los registros del viajero indican que gastó un total de \$340 en hospedaje, \$320 en comida y \$140 en gastos adicionales durante su viaje por estos tres países. Calcule el número de días que pasó el viajero en cada país o muestre que los registros son incorrectos debido a que las cantidades gastadas no son compatibles una con la otra.

(3) Una embotelladora de refrescos desea cotizar la publicidad de sus productos en televisión, radio y revista, se tienen tres propuestas del plan de medios de acuerdo con el presupuesto asignado acerca de la cantidad de anuncios por medio en el transcurso de un mes. En el primer presupuesto cada anuncio en televisión tiene un coste de \$250 000, en radio \$5 000 y en revista \$30 000. En el segundo presupuesto \$310 000, \$4 000 y \$15 000 y en el último presupuesto \$560 000, \$10 000 y \$35 000. Los totales por presupuesto son los siguientes: \$21 795 000, \$31 767 000 y \$61 225 000. Determine la cantidad de anuncios cotizados por cada medio.

Capítulo 8

Álgebra matricial

8.1. El espacio vectorial de las matrices m x n

Recordemos que si m y n son enteros mayores o iguales que 1, una *matriz* $m \times n$ es un arreglo rectangular de $m.n$ números dispuestos en m filas y n columnas

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

El conjunto de las matrices $m \times n$ se suele denotar por $\mathbb{R}^{m \times n}$.

Es de esperar que el conjunto de las matrices sea un espacio vectorial, ya que una matriz es básicamente un vector "muy largo" ordenado de manera adecuada.

8.1.1. Suma de matrices. La suma de matrices se realiza sumando componente a componente.

Si $A = (a_{ij})$ y $B = (b_{ij})$ son dos matrices $m \times n$ entonces la suma de A y B es una matriz $m \times n$ dada por

$$A + B = (a_{ij} + b_{ij})$$

es decir

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \vdots\vdots\vdots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

Ejemplo 8.1.

$$\begin{bmatrix} 3 & -1 & 1/2 \\ 6 & \sqrt{2} & 0 \end{bmatrix} + \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{2} & 4 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 1 \\ 6 & 2\sqrt{2} & 4 \end{bmatrix}$$

8.1.2. Producto de una matriz por un escalar. El producto de una matriz por un escalar se realiza multiplicando el escalar por cada componente.

Si $\lambda \in \mathbb{R}$ y $A = (a_{ij})$ es una matriz $m \times n$ entonces

$$\lambda A = (\lambda a_{ij})$$

es decir

$$\lambda \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \vdots\vdots\vdots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}$$

Ejemplo 8.2.

$$4 \begin{bmatrix} 3 & -1 & 1/2 \\ 6 & \sqrt{2} & 0 \end{bmatrix} = \begin{bmatrix} 12 & -4 & 2 \\ 24 & 4\sqrt{2} & 0 \end{bmatrix}$$

Si A y B son matrices $m \times n$ entonces, se define la matriz $-A$ por

$$-A = (-1)A$$

y se define

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

Ejemplo 8.3. Si

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

entonces

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

8.1.3. Propiedades de la suma y la multiplicación escalar. Todas las propiedades algebraicas de la suma y la multiplicación escalar para vectores (Proposición 4.1) se trasladan a las matrices. Para completar, se resumen dichas propiedades en el siguiente teorema.

La matriz nula, que denotaremos por O, es la matriz que tiene 0 en todas sus posiciones.

En detalle

Teorema 8.4. *Sean A, B y C matrices del mismo tamaño, y sean α y β escalares. Entonces*

(1) $A + B = B + A$ *(Conmutatividad)*
(2) $(A + B) + C = A + (B + C)$ *(Asociatividad)*
(3) $A + O = A$
(4) $A + (-A) = O$
(5) $\alpha(A + B) = \alpha A + \beta B$ *(Distributividad)*
(6) $(\alpha + \beta)A = \alpha A + \beta A$ *(Distributividad)*
(7) $\alpha(\beta A) = (\alpha\beta)A$
(8) $1A = A$
(9) $0A = O$

Es decir, el conjunto de las matrices $m \times n$ es un espacio vectorial.

Ejemplo 8.5.
Sean $A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, y $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

(1) ¿ Es $B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ una combinación lineal de A_1, A_2, y A_3?

(2) ¿ Es $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ una combinación lineal de A_1, A_2, y A_3?

(1) Se quieren encontrar escalares c_1, c_2, c_3 tales que $c_1A_1 + c_2A_2 + c_3A_3 = B$. Por tanto,

$$c_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$$

El lado izquierdo de esta ecuación puede reescribirse como

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix}$$

Al comparar entradas y usar la definición de igualdad matricial, se tienen cuatro ecuaciones lineales:

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 + c_3 &= 4 \\ -c_1 + c_3 &= 2 \\ c_2 + c_3 &= 1 \end{aligned}$$

La eliminación de Gauss-Jordan produce fácilmente

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

(hacerlo como ejercicio) de modo que $c_1 = 1, c_2 = -2, c_3 = 3$.

En consecuencia, $B = A_1 - 2A_2 + 3A_3$, lo que puede comprobarse fácilmente.

(2) Esta vez se debe resolver

$$c_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Al proceder como en el inciso (a), se obtiene el sistema lineal:

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 + c_3 &= 2 \\ -c_1 + c_3 &= 3 \\ c_2 + c_3 &= 4 \end{aligned}$$

La reducción por filas o renglones produce:

$$\left[\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 4 \end{array}\right] \xrightarrow{R_4 - R_1} \left[\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 0 & 3 \end{array}\right]$$

No es necesario avanzar más: el último renglón implica que no hay solución. Por tanto, en este caso, C no es una combinación lineal de A_1, A_2, y A_3.

EJEMPLO 8.6. Describir el subespacio generado por las matrices A_1, A_2, y A_3 del Ejemplo 8.5. Una forma de hacer esto es simplemente escribir una combinación lineal general de A_1, A_2, y A_3. Por tanto,

$$c_1 A_1 + c_2 A_2 + c_3 A_3 = c_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix}$$

(que es análogo a la representación paramétrica de un plano).

Supongamos que queremos saber cuándo la matriz $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ está en $\operatorname{Span}\{A_1, A_2, A_3\}$.

Por la representación anterior, se debe cumplir

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

para alguna elección de escalares c_1, c_2, c_3. Esto da origen a un sistema de ecuaciones lineales cuyo lado izquierdo es exactamente igual que el del Ejemplo 8.5, pero cuyo lado derecho es general. La matriz aumentada de este sistema es

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right]$$

y la reducción por renglones produce

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}x - \frac{1}{2}y \\ 0 & 1 & 0 & -\frac{1}{2}x - \frac{1}{2}y + w \\ 0 & 0 & 1 & \frac{1}{2}x + \frac{1}{2}y \\ 0 & 0 & 0 & w - z \end{array}\right]$$

(comprobar a manera de ejercicio). La única restricción proviene del último renglón, donde claramente se debe tener $w - z = 0$ para tener una solución. Por tanto, $\operatorname{Span}\{A_1, A_2, A_3\}$ es el espacio de todas las matrices $\begin{bmatrix} w & x \\ y & w \end{bmatrix}$ para las que $w = z$. En notación de conjuntos

$$\operatorname{Span}\{A_1, A_2, A_3\} = \left\{ \begin{bmatrix} w & x \\ y & w \end{bmatrix} \ : \ x, y, w \in \mathbb{R} \right\}$$

EJEMPLO 8.7. Determinar si las matrices A_1, A_2, y A_3 del Ejemplo 8.5 son linealmente independientes.

Se debe resolver la ecuación $c_1A_1 + c_2A_2 + c_3A_3 = O$.

En detalle:

$$c_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} c_2 + c_1 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

llegamos a un sistema de ecuaciones cuya matriz aumentada es

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

Al reducir por renglones la matriz aumentada se obtiene

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

Por tanto, $c_1 = c_2 = c_3 = 0$, y se concluye que las matrices A_1, A_2, y A_3 son linealmente independientes.

Ejercicios.

(1) Sean $A = \begin{bmatrix} 5 & -9 \\ 3 & 0 \\ 6 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 5 & -9 \\ 3 & 0 \\ 6 & 1 \end{bmatrix}$ y $C = \begin{bmatrix} 4 & 3 \\ 1 & 0 \\ 5 & 1 \end{bmatrix}$

  (a) Hallar $6B - 7A + 5C$

  (b) Encuentre una matriz E tal que $A + 2B - 3C + E$ es la matriz cero de 2×3.

  (c) Encuentre una matriz G tal que $A+B+G$ es la matriz de 2×3 con todos sus elementos iguales a 1.

(2) Dados $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}$, resuelva la siguiente ecuación para X:

$$3(2A + B + X) = 5(X - A + B)$$

8.2. Producto de matrices.

Bajo ciertas condiciones se puede definir un producto para matrices.

Sean m, n y p enteros mayores o iguales que 1 y sea $A = (a_{ij})$ una matriz $m \times n$ y $B = (b_{ij})$ una matriz $n \times p$. Entonces el producto de A y B es una matriz $C = (c_{ij})$ de tamaño $m \times p$ donde

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}.$$

En palabras: fijamos una fila en la primera matriz y una columna en la segunda matriz. Se multiplican las entradas de la fila de la primera matriz por las entradas de la columna de la segunda matriz y luego se suman. El valor obtenido se coloca en la entrada de la matriz producto correspondiente a esa fila y esa columna. Es importante notar que esta operación se puede hacer solamente cuando el número de columnas de la primera matriz debe ser igual al número de filas de la segunda matriz.

A continuación se explica la forma de multiplicar matrices con algunos ejemplos

Ejemplo 8.8.

(1) Una matriz 1×1 por una matriz 1×1 da una matriz 1×1.

$$\begin{bmatrix} a_{11} \end{bmatrix} \begin{bmatrix} b_{11} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} \end{bmatrix}$$

(2) Una matriz 1×2 por una matriz 2×1 da una matriz 1×1.

$$\begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} \end{bmatrix}$$

(3) Una matriz 1×3 por una matriz 3×1 da una matriz 1×1.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \end{bmatrix}$$

(4) Una matriz $1 \times n$ por una matriz $n \times 1$ da una matriz 1×1.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots a_{1n}b_{n1} \end{bmatrix}$$

Como es natural, una matriz 1×1 se identifica con el número que corresponde a su única entrada, por lo que el producto de un vector fila por un vector columna es igual al producto escalar de los vectores.

(5) Un ejemplo numérico

$$\begin{bmatrix} 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} 8 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \cdot 8 + 2 \cdot 3 + (-1)5 \end{bmatrix} = \begin{bmatrix} 25 \end{bmatrix} = 25$$

Ejemplo 8.9. Producto de un vector demanda y un vector de precios.

Suponga que un fabricante produce cuatro artículos. Su demanda está dada por el vector de demanda

$$d = \begin{bmatrix} 30 & 20 & 40 & 10 \end{bmatrix}$$

(una matriz 1×4). El precio por unidad que recibe el fabricante por los artículos está dado por el vector de precios

$$p = \begin{bmatrix} 2000Bs \\ 1500Bs \\ 1800Bs \\ 4000Bs \end{bmatrix}$$

(una matriz 4×1). Si se cumple la demanda, ¿cuánto dinero recibirá el fabricante?

Solución: La demanda del primer artículo es 30 y el fabricante recibe 2000 Bs por cada artículo vendido. Entonces recibe 60000 Bs de las ventas del primer artículo. Si se sigue este razonamiento, se ve que la cantidad total de bolívares que recibe es

$$30 \cdot 2000 + 20 \cdot 1500 + 40 \cdot 1800 + 10 \cdot 4000 = 202000$$

Este resultado se puede escribir como

$$\begin{bmatrix} 30 & 20 & 40 & 10 \end{bmatrix} \begin{bmatrix} 2000 \\ 1500 \\ 1800 \\ 4000 \end{bmatrix} = 202000Bs.$$

Es decir, se multiplicó un vector fila por un vector columna y se obtuvo un escalar.

EJEMPLO 8.10. Una matriz $n \times n$ por una matriz $n \times n$ da una matriz $n \times n$, tal como se indica en los siguientes casos:

(1) Una matriz 2×2 por una matriz 2×2 da una matriz 2×2.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

(2) Una matriz 3×3 por una matriz 3×3 da una matriz 3×3.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{bmatrix}$$

(3) Dadas

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

Tenemos que

$$AB = \begin{bmatrix} 2+1-1 & 0+1-1 & 2+1+1 \\ 2+3-4 & 0+3-4 & 2+3+4 \\ -1-1+2 & 0-1+2 & -1-1-2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 4 \\ 1 & -1 & 9 \\ 0 & 1 & -4 \end{bmatrix}$$

En general, si $A = (a_{ij})$ es una matriz $m \times n$ y $B = (b_{ij})$ es una matriz $n \times p$ y $C = AB$, entonces C es una matriz $n \times p$ y c_{ij} es igual al producto escalar del vector fila

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

por el vector columna

$$\begin{bmatrix} \mathbf{b_{1j}} \\ \mathbf{b_{2j}} \\ \vdots \\ \mathbf{b_{nj}} \end{bmatrix}$$

En grande:

$$\begin{array}{cccccc}
 & & & \begin{matrix} & & \mathbf{j} & & \end{matrix} & & \begin{matrix} & & \mathbf{j} & & \end{matrix} \\
\mathbf{i} & \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \vdots\,\vdots\,\vdots & \vdots & \vdots\,\vdots\,\vdots & \vdots \\ \mathbf{a_{i1}} & \cdots & \mathbf{a_{ij}} & \cdots & \mathbf{a_{in}} \\ \vdots & \vdots\,\vdots\,\vdots & \vdots & \vdots\,\vdots\,\vdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} & & \begin{pmatrix} b_{11} & \cdots & \mathbf{b_{1j}} & \cdots & b_{1p} \\ \vdots & \vdots\,\vdots\,\vdots & \vdots & \vdots\,\vdots\,\vdots & \vdots \\ b_{i1} & \cdots & \mathbf{b_{ij}} & \cdots & b_{in} \\ \vdots & \vdots\,\vdots\,\vdots & \vdots & \vdots\,\vdots\,\vdots & \vdots \\ b_{n1} & \cdots & \mathbf{b_{nj}} & \cdots & b_{np} \end{pmatrix} & = \ \mathbf{i} & \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\ \vdots & \vdots\,\vdots\,\vdots & \vdots & \vdots\,\vdots\,\vdots & \vdots \\ c_{i1} & \cdots & \mathbf{c_{ij}} & \cdots & c_{ip} \\ \vdots & \vdots\,\vdots\,\vdots & \vdots & \vdots\,\vdots\,\vdots & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mp} \end{pmatrix}
\end{array}$$

A continuación una serie de ejemplos.

EJEMPLO 8.11.

(1) Una matriz 2×1 por una matriz 1×2 da una matriz 2×2.

$$\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{bmatrix}$$

(2) Una matriz 3×1 por una matriz 1×3 da una matriz 3×3 (**note la importancia del orden de la operación**).

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} \\ a_{31}b_{11} & a_{31}b_{12} & a_{31}b_{13} \end{bmatrix}$$

(3) Una matriz 3×2 por una matriz 2×3 da una matriz 3×3.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} \end{bmatrix}$$

(4) Una matriz 2×3 por una matriz 3×2 da una matriz 2×2.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}$$

(5) Un ejemplo con números

$$\begin{bmatrix} 8 \\ \pi \\ \sqrt{5} \end{bmatrix} \begin{bmatrix} 3 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 8 \cdot 3 & 8 \cdot 2 & 8 \cdot (-1) \\ \pi \cdot 3 & \pi \cdot 2 & \pi \cdot (-1) \\ \sqrt{5} \cdot 3 & \sqrt{5} \cdot 2 & \sqrt{5} \cdot (-1) \end{bmatrix} = \begin{bmatrix} 24 & 16 & -8 \\ 3\pi & 2\pi & -\pi \\ 3\sqrt{5} & 2\sqrt{5} & -\sqrt{5} \end{bmatrix}$$

(6) Otro ejemplo numérico

$$\begin{bmatrix} 6 & 2 & \sqrt{2} \\ 1 & e & -4 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 \cdot (-2) + 2 \cdot 1 + \sqrt{2} & 6 \cdot 0 + 2 \cdot 0 + \sqrt{2} \cdot 1 \\ 1 \cdot (-2) + e \cdot 1 + (-4) \cdot 1 & 1 \cdot 0 + e \cdot 0 + (-4) \cdot 1 \end{bmatrix}$$
$$= \begin{bmatrix} -10 + \sqrt{2} & \sqrt{2} \\ -6 + e & -4 \end{bmatrix}$$

En general, el producto de matrices no es conmutativo, es decir, puede ocurrir que si A y B son matrices, entonces $AB \neq BA$, mostraremos esto mediante un ejemplo.

EJEMPLO 8.12.
Sean
$$A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \qquad \text{y} \qquad B = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix},$$
entonces

$$AB = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 3 \cdot 1 + 1 \cdot 2 & 3 \cdot (-1) + 1 \cdot 0 \\ 5 \cdot 1 + 2 \cdot 2 & 5 \cdot (-1) + 2 \cdot 0 \end{bmatrix} = \begin{bmatrix} 5 & -3 \\ 9 & -5 \end{bmatrix}$$
$$BA = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} 1 \cdot 3 + (-1) \cdot 5 & 1 \cdot 1 + (-1) \cdot 2 \\ 2 \cdot 3 + 0 \cdot 5 & 2 \cdot 1 + 0 \cdot 2 \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ 6 & 2 \end{bmatrix}.$$

También puede ocurrir que el producto de dos matrices sea igual a la matriz nula, sin que ninguna de las dos matrices sea nula, tal como muestra el siguiente ejemplo.

EJEMPLO 8.13.

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

Sin embargo el producto de matrices sí es asociativo, es decir:

PROPOSICIÓN 8.14. *Sean A una matriz $m \times n$, B una matriz $n \times p$ y C una matriz $p \times q$ entonces*

$$A(BC) = (AB)C$$

y esta matriz es $m \times q$.

Ejercicios.

(1) Dados $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 0 \\ 0 & 6 \end{bmatrix}$ y $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, encuentre una matriz X tal que $AX + XB = C$.

(2) Realizar las operaciones indicadas

(a) $\begin{bmatrix} 3 & -2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -5 & 6 \\ 1 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} -4 & 5 & 1 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} 3 & -1 & 1 \\ 5 & 6 & 4 \\ 0 & 1 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 7 & 1 & 4 \\ 2 & -3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 4 \\ -2 & 3 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 4 & 6 \\ -2 & 3 & 5 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 2 & -3 & 5 \\ 1 & 0 & 6 \\ 2 & 3 & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 3 & 2 & 1 & -2 \\ -6 & 4 & 0 & 3 \\ 2 & 4 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 0 \\ 2 \end{bmatrix}$

(f) $\begin{bmatrix} 3 & -2 & 1 \\ 4 & 0 & 6 \\ 5 & 1 & 9 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

8.3. Inversa de una matriz y la matriz identidad

En esta sección consideraremos solamente matrices cuadradas.

Entre los tipos sencillos de matrices están las diagonales y las triangulares.

Se dice que una matriz A es *diagonal* si $a_{ij} = 0$ para $i \neq j$, es decir, cuando todos sus elementos son cero, excepto los de la diagonal que va desde el vértice superior izquierdo al inferior derecho.

Se dice que la matriz A es *triangular superior* si $a_{ij} = 0$ para $i > j$ y se dice que es *triangular inferior* si $a_{ij} = 0$ para $i < j$.

EJEMPLO 8.15. Las matrices

$$A = \begin{bmatrix} 2 & 1 & 7 \\ 0 & 2 & -5 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{y} \qquad B = \begin{bmatrix} -2 & 3 & 0 & 1 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

son triangulares superiores.

Las matrices

$$C = \begin{bmatrix} 5 & 0 & 0 \\ 2 & 3 & 0 \\ -1 & 2 & 4 \end{bmatrix} \qquad \text{y} \qquad D = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

son triangulares inferiores.

Las matrices

$$E = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & -4 \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

son diagonales.

Observen que toda matriz diagonal es también triangular superior y triangular inferior.

Entre las matrices diagonales hay unas más sencillas, que son las que tienen 1 en todas las componentes de la diagonal principal.

Observemos que para matrices 2×2:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Análogamente para matrices 3×3 tenemos:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

A continuación definimos la *matriz identidad* , que puede ser de diferentes tamaños, pero siempre es una matriz cuadrada.

Para matrices 1×1 la matriz identidad es:

$$I_1 = I = \begin{bmatrix} 1 \end{bmatrix}$$

Para matrices 2×2 la matriz identidad es:

$$I_2 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Para matrices 3×3 la matriz identidad es:

$$I_3 = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Para matrices $n \times n$ la matriz identidad es

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

es decir, es la matriz que tiene todas sus entradas iguales a 0, salvo en la diagonal que tiene solamente el número 1.

Si A es una matriz $n \times n$, se cumple que

$$AI_n = I_nA = A.$$

Si no se presta a confusión se suele escribir simplemente I en vez de I_n.

DEFINICIÓN 8.16. Sea A una matriz cuadrada $n \times n$.
Se dice que A es invertible si existe una matriz cuadrada B, también $n \times n$, tal que

$$AB = BA = I$$

La matriz B de existir es única, se le llama la inversa de A y se denota por A^{-1}.

EJEMPLO 8.17. Sean

$$A = \begin{bmatrix} 2 & 4 & 0 \\ 0 & 2 & 1 \\ 3 & 0 & 2 \end{bmatrix} \qquad \text{y} \qquad B = \begin{bmatrix} 4/20 & -8/20 & 4/20 \\ 3/20 & 4/20 & -2/20 \\ -6/20 & 12/20 & 4/20 \end{bmatrix}$$

es fácil verificar que (hacerlo a manera de ejercicio)

$$AB = BA = I_3.$$

Por lo tanto,

$$A^{-1} = \begin{bmatrix} 4/20 & -8/20 & 4/20 \\ 3/20 & 4/20 & -2/20 \\ -6/20 & 12/20 & 4/20 \end{bmatrix} = \frac{1}{20}\begin{bmatrix} 4 & -8 & 4 \\ 3 & 4 & -2 \\ -6 & 12 & 4 \end{bmatrix}$$

OBSERVACIÓN 8.18. No todas matriz no nula es invertible:
En el ejemplo 8.2 vimos que

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

lo que implica que ninguna de estas matrices es invertible.

El producto de matrices invertible también es invertible.

PROPOSICIÓN 8.19. *Sean A y B matrices $n \times n$, invertibles. Entonces la matriz AB es invertible y*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

DEMOSTRACIÓN.

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I,$$

de igual manera se prueba que

$$(B^{-1}A^{-1})(AB) = I.$$

□

8.3.1. Cálculo de la inversa de una matriz. Existen caso en los que es sencillo calcular la inversa de una matriz, por ejemplo si

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

es una matriz diagonal, entonces A es invertible si y solo si todos los elementos de la diagonal son diferentes de 0 y, en este caso

$$A^{-1} = \begin{bmatrix} \dfrac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \dfrac{1}{a_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{a_{nn}} \end{bmatrix}$$

Si A es una matriz 2×2 dada por

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

y $ad - bc \neq 0$ su inversa es la matriz dada por

$$A^{-1} = \begin{bmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{8.1}$$

8.3.2. Método de Gaus-Jordan para calcular la matriz inversa. Dada una matriz cuadrada A de tamaño $n \times n$, el método de Gauss-Jordan para encontrar su inversa A^{-1} consiste en:

(1) Construir la matriz aumentada $[A|I_n]$, donde I_n es la matriz identidad de tamaño n.
(2) Aplicar operaciones elementales de fila en la matriz aumentada para transformar A en I_n.
(3) La matriz resultante en el lado derecho será A^{-1}.

Si no es posible transformar A en la identidad, entonces la matriz no es invertible.

Se ilustra el método con ejemplos

EJEMPLO 8.20.

(1) ¿Es la matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ invertible? En caso de ser invertible, hallar su inversa.

La matriz aumentada es

$$[A|I] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \xrightarrow{R_2 \to R_2 - 3R_1} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right] \xrightarrow{R_2 \to -\frac{1}{2}R_2} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right] \xrightarrow{R_1 \to R_1 - 2R_2} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right]$$

Luego

$$A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

Verificación:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} -2 + 2\left(\dfrac{3}{2}\right) & 1 + 2\left(\dfrac{-1}{2}\right) \\ (-2)(3) + 4\left(\dfrac{3}{2}\right) & 3 + 4\left(\dfrac{-1}{2}\right) = \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Igualmente se tiene

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

También se ha podido usar la fórmula (8.1):

$$A^{-1} = \frac{1}{(1)(4) - (2)(3)} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

(2) ¿Es la matriz $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$ invertible? En caso de ser invertible, hallar su inversa.

La matriz aumentada es

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3 \to R_3 - 4R_1]{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right] \xrightarrow{R_2 \to -R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 \to -R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

$$\xrightarrow[R_2 \to R_2 - R_3]{R_1 \to R_1 - 2R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

La matriz es invertible y

$$A^{-1} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}$$

(3) ¿Es la matriz $A = \begin{pmatrix} 1 & -2 & 2 & 2 \\ 0 & 4 & -2 & 1 \\ 1 & -2 & 4 & 0 \\ 1 & -1 & 2 & 2 \end{pmatrix}$ invertible? En caso de ser invertible, hallar su inversa.

La matriz aumentada es

$$[A|I] = \left[\begin{array}{cccc|cccc} 1 & -2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 4 & -2 & 1 & 0 & 1 & 0 & 0 \\ 1 & -2 & 4 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 2 & 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc|cccc} 1 & -2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 4 & -2 & 1 & 0 & 1 & 0 & 0 \\ 1 & -2 & 4 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 2 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_4 \to R_4 - R_1]{R_3 \to R_3 - R_1} \left[\begin{array}{cccc|cccc} 1 & -2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 4 & -2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_4 \to R_4 - 4R_2]{R_2 \leftrightarrow R_4} \left[\begin{array}{cccc|cccc} 1 & -2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 2 & -2 & -1 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 & 4 & 1 & 0 & -4 \end{array}\right]$$

$$\xrightarrow[R_4 \to R_4 + 2R_3]{R_3 \to \frac{1}{2}R_3} \left[\begin{array}{cccc|cccc} 1 & -2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -1 & 3 & 1 & 1 & -4 \end{array}\right]$$

$$\xrightarrow{R_4 \to -R_4} \left[\begin{array}{cccc|cccc} 1 & -2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & -3 & -1 & -1 & 4 \end{array}\right]$$

$$\xrightarrow[R_1 \to R_1 - 2R_4]{R_3 \to R_3 + R_4} \left[\begin{array}{cccc|cccc} 1 & -2 & 2 & 0 & 7 & 2 & 2 & -8 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -1 & -\frac{1}{2} & 4 \\ 0 & 0 & 0 & 1 & -3 & -1 & -1 & 4 \end{array}\right]$$

$$\xrightarrow[R_1\to R_1+2R_2]{R_1\to R_1-2R_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 12 & 4 & 3 & -14 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -1 & -\frac{1}{2} & 4 \\ 0 & 0 & 0 & 1 & -3 & -1 & -1 & 4 \end{array}\right]$$

La matriz A es invertible y

$$A^{-1} = \begin{pmatrix} 12 & 4 & 3 & -14 \\ -1 & 0 & 0 & 1 \\ -\frac{7}{2} & -1 & -\frac{1}{2} & 4 \\ -3 & -1 & -1 & 4 \end{pmatrix}$$

(4) ¿Es la matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ invertible? En caso de ser invertible, hallar su inversa.

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3\to R_3-7R_1]{R_2\to R_2-4R_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -6 & -12 & -7 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right]$$

Obtuvimos un renglón con tres ceros en lado izquierdo, por lo tanto, la matriz no es invertible.

OBSERVACIÓN 8.21.

- El método de Gauss-Jordan es sistemático pero puede volverse computacionalmente intensivo para matrices grandes.
- La aparición de una fila de ceros en el lado izquierdo durante el proceso indica que la matriz no es invertible.
- Las fracciones en los resultados son normales y pueden evitar errores de redondeo.
- Se recomienda verificar siempre el resultado multiplicando $A\,A^{-1}$ para confirmar que da la identidad.

8.3.3. Propiedades de las matrices invertibles.

DEFINICIÓN 8.22. Si A es una matriz y n es un entero positivo, como es natural , se define

$$A^n = \overbrace{A\,A \cdots A}^{\text{n veces}}.$$

Si A es invertible, se define

$$A^{-n} = (A^{-1})^n = (A^n)^{-1}.$$

El siguiente resultado resume las propiedades más importantes de las matrices invertibles.

Proposición 8.23.

(1) *Si A es una matriz invertible, entonces A^{-1} es invertible y $(A^{-1})^{-1} = A$*

(2) *Si A es una matriz invertible y c es un escalar distinto de cero, entonces cA es una matriz invertible y $(cA)^{-1} = \frac{1}{c}A^{-1}$*

(3) *Si A y B son matrices invertibles del mismo tamaño, entonces AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$*

(4) *Si A es una matriz invertible, entonces A^n es invertible para todo entero n no negativo y $(A^n)^{-1} = (A^{-1})^n$*

Ejercicios.

(1) Determine si la matriz dada es invertible; de ser invertible, calcule la inversa.

(a) $\begin{bmatrix} 3 & -2 \\ 5 & -4 \end{bmatrix}$ (b) $\begin{bmatrix} 4 & -7 \\ -8 & 14 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

(g) $\begin{bmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{bmatrix}$ (h) $\begin{bmatrix} 3 & 4 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 3 & 4 \end{bmatrix}$ (i) $\begin{bmatrix} 0 & 0 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 2 & 3 & 0 & 0 \\ 3 & 4 & 0 & 0 \end{bmatrix}$

(2) Muestre que si A, B y C son matrices invertibles, entonces ABC es invertible y $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

8.4. Propiedades de la multiplicación de matrices

La multiplicación de matrices, no comparte todas sus propiedades con el producto de los números reales; aunque en muchos aspectos lo hace, existen algunas diferencias significativas.

Ya hemos visto que el producto de dos matrices diferentes de O puede dar la matriz O y que el producto no es conmutativo. A manera de refuerzo, consideremos el siguiente ejemplo.

Ejemplo 8.24. Consideremos las matrices

$$A = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Al multiplicar, se obtiene

$$AB = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -3 & -2 \end{bmatrix} \quad \text{y} \quad BA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

Por tanto, $AB \neq BA$. Así que, en contraste con la multiplicación de números reales, la multiplicación de matrices no es conmutativa, es decir, el orden de los factores en un producto sí importa.

También se tiene que

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

Por lo tanto, para matrices, la ecuación $A^2 = 0$ no implica que $A = 0$ (a diferencia de la situación para números reales, donde la ecuación $x^2 = 0$ sólo tiene $x = 0$ como solución).

El siguiente resultado resume las principales propiedades de la multiplicación de matrices.

TEOREMA 8.25. *Sean A, B y C matrices (cuyos tamaños son tales que pueden realizarse las operaciones indicadas) y sea k un escalar. Entonces*

(1) $A(BC) = (AB)C$ *(Asociatividad)*
(2) $A(B + C) = AB + AC$ *(Distributividad izquierda)*
(3) $(A + B)C = AC + BC$ *(Distributividad derecha)*
(4) $k(AB) = (kA)B = A(kB)$
(5) $I_m A = A = AI_n$ *si A es $m \times n$* *(Identidad multiplicativa)*

EJEMPLO 8.26. Si A y B son matrices cuadradas del mismo tamaño.

$$¿\text{Es } (A + B)^2 = A^2 + 2AB + B^2?$$

Veamos

$$\begin{aligned} (A + B)^2 &= (A + B)(A + B) \\ &= (A + B)A + (A + B)B \quad \text{por distributividad izquierda} \\ &= A^2 + BA + AB + B^2 \quad \text{por distributividad derecha} \end{aligned}$$

Por tanto, $(A + B)^2 = A^2 + 2AB + B^2$ si y sólo s $BA = AB$.

8.5. Ecuaciones matriciales

EJEMPLO 8.27. Supongamos que tenemos que resolver el siguiente sistema de ecuaciones:

$$\begin{cases} x_1 \quad + \quad 2x_3 = 3 \\ 2x_1 - x_2 + 3x_3 = 4 \\ 4x_1 + x_2 + 8x_3 = 1 \end{cases}$$

En forma matricial esta ecuación queda de la siguiente manera:

$$\begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 1] \end{bmatrix}$$

En el Ejemplo 8.20 vimos que la matriz

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

es invertible y que su inversa es la matriz

$$A^{-1} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}$$

Nuestro sistema tiene la forma

$$AX = X_o$$

donde

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \text{y} \qquad X_o = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix},$$

por lo tanto, la solución la podemos obtener de la siguiente manera

$$X = A^{-1}X_o,$$

es decir,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -23 \\ -11 \\ 13 \end{bmatrix},$$

de donde

$$x_1 = -23 \qquad x_2 = -11 \qquad x_3 = 13.$$

EJEMPLO 8.28. Resolver la siguiente ecuación matricial para encontrar X, suponiendo que las matrices involucradas son tales que todas las operaciones indicadas están definidas:

$$A^{-1}(BX)^{-1} = (A^{-1}B^3)^2$$

Existen muchas formas de proceder aquí, una manera es

$$\begin{aligned} A^{-1}(BX)^{-1} = (A^{-1}B^3)^2 &\Rightarrow ((BX)A)^{-1} = (A^{-1}B^3)^2 \\ &\Rightarrow [((BX)A)^{-1}]^{-1} = [(A^{-1}B^3)^2]^{-1} \\ &\Rightarrow (BX)A = [(A^{-1}B^3)(A^{-1}B^3)]^{-1} \\ &\Rightarrow (BX)A = B^{-3}(A^{-1})^{-1}B^{-3}(A^{-1})^{-1} \\ &\Rightarrow BXA = B^{-3}AB^{-3}A \\ &\Rightarrow B^{-1}BXAA^{-1} = B^{-1}B^{-3}AB^{-3}AA^{-1} \\ &\Rightarrow IXI = B^{-4}AB^{-3}I \\ &\Rightarrow X = B^{-4}AB^{-3} \end{aligned}$$

8.6. La matriz transpuesta

Definición 8.29. Sea A un matriz $m \times n$, la *matriz transpuesta* de A, que se denota por A^T es la matriz $n \times m$ definida por

$$(A^T)_{ij} = A_{ji}.$$

En palabras, la transpuesta de A es la matriz que se obtiene al intercambiar filas y columnas de A, la i-ésima columna de A^T es la i-ésima fila de A.

Ejemplo 8.30. Si

$$A = \begin{bmatrix} 1 & 5 & 3 \\ 2 & -2 & 4 \\ 9 & 8 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad C = \begin{bmatrix} 5 \\ 3 \\ \sqrt{2} \end{bmatrix}$$

entonces

$$A^T = \begin{bmatrix} 1 & 2 & 9 \\ 5 & -2 & 8 \\ 3 & 4 & 6 \end{bmatrix} \qquad B^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \qquad C^T = \begin{bmatrix} 5 & 3 & \sqrt{2} \end{bmatrix}$$

El siguiente resultado resume algunas de las propiedades más importantes de la transpuesta.

Proposición 8.31. *Sean A y B matrices (cuyos tamaños son tales que pueden realizarse las operaciones indicadas) y sea k un escalar. Entonces*

(1) $(A^T)^T = A$.
(2) $(A + B)^T = A^T + B^T$.
(3) $(kA)^T = k(A^T)$.
(4) $(AB)^T = B^T A^T$.
(5) $(A^n)^T = (A^T)^n$ *para todos los enteros n no negativos.*

Observación 8.32. Las propiedades (b) y (d) del Teorema 8.31 pueden generalizarse a sumas y productos de un número finito de matrices:

$$(A_1 + A_2 + \cdots + A_k)^T = A_1^T + A_2^T + \cdots + A_k^T \quad \text{y} \quad (A_1 A_2 \cdots A_k)^T = A_k^T \cdots A_2^T A_1^T$$

si supone que los tamaños de las matrices son tales que todas las operaciones pueden realizarse.

Se dice que una matriz A es *simétrica* si $A = A^T$.

Observación 8.33. Sean

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

Entonces $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, de modo que $A + A^T = \begin{bmatrix} 2 & 5 \\ 5 & 8 \end{bmatrix}$, una matriz simétrica.
Se tiene

$$B^T = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix}$$

de modo que

$$BB^T = \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 5 \\ 5 & 14 \end{bmatrix}$$

$$\text{y} \quad B^T B = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 2 & 2 \\ 2 & 10 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

En consecuencia, tanto BB^T y B^TB son simétricas, ¡aun cuando B ni siquiera es cuadrada. (Comprobar que AA^T y A^TA también son simétricas.)

El siguiente teorema dice que los resultados análogos a los del ejemplo anterior se cumplen en condiciones generales.

Proposición 8.34.

(1) *Si A es una matriz cuadrada, entonces $A + A^T$ es una matriz simétrica.*
(2) *Para cualquier matriz A, AA^T y A^TA son matrices simétricas.*

8.7. Ejercicios varios

(1) Si $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, yB = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, encuentre las condiciones para a, b, c, d tal que $AB = BA$.
(2) Una matriz A de $n \times n$ tal que $A^2 = I_n$ se llama involutiva. Pruebe que la siguiente matriz es involutiva:

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

(3) Dada la siguiente matriz pruebe que $A^2 = A$:

$$A = \begin{bmatrix} -1 & 3 & 5 \\ 1 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix}.$$

(4) Un fabricante de joyería de diseño tiene órdenes por dos anillos, tres pares de aretes, cinco prendedores y un collar. El fabricante estima que le llevará 1 hora de mano de obra hacer un anillo, $1\frac{1}{2}$ horas hacer un par de aretes, $\frac{1}{2}$ hora para un prendedor y 2 horas para un collar.

a) Exprese las órdenes del fabricante como un vector renglón. b) Exprese los requerimientos en horas para los distintos tipos de joyas como un vector columna. c) Utilice el producto escalar para calcular el número total de horas que requerirá para terminar las órdenes.

(5) Determinar si la matriz dada en invertible; en caso de serlo, halle su inversa.

(a) $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \\ 1 & 3 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 2 & 3 \\ -1 & 1 & 0 & 4 \\ 2 & 1 & -1 & 3 \\ -1 & 0 & 5 & 7 \end{bmatrix}$

(6) Calcular $(A^T)^{-1}$ y $(A^{-1})^T$ y verifique que son iguales.

(a) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & 0 \\ 6 & 3 \end{bmatrix}$ (c) $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$

(d) $A = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{bmatrix}$ (e) $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{bmatrix}$ (f) $A = \begin{bmatrix} 9 & 0 & 12 \\ 0 & 2 & 0 \\ -4 & 0 & -5 \end{bmatrix}$

(7) Resolver los siguientes sistemas de ecuaciones mediante el método de la matriz inversa

(a) $\begin{cases} x_1 - 3x_2 = -7 \\ 2x_1 - x_2 = 4 \end{cases}$

(b) $\begin{cases} 2x_1 - 3x_2 + x_3 = 4 \\ 4x_1 - 5x_2 + 3x_3 = 4 \\ -3x_1 + 2x_2 - 3x_3 = 1 \end{cases}$

(c) $\begin{cases} x_1 - x_2 + x_3 = -2 \\ 2x_1 - x_2 + 3x_3 = -9 \\ 5x_1 - 3x_2 + 4x_3 = -11 \end{cases}$

(8) Sea $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \\ 2 & 1 & 0 \end{bmatrix}$. Hallar, si es que existe, una matriz B tal que $A^2 - A = AB$.

(9) Sea $A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$. Resolver la ecuación matricial

$$A^2X - A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}.$$

8.8. Lectura adicional: Factorización LU

Así como es natural, y muchas veces necesario para resolver un problema, factorizar un número natural en un producto de otros números naturales más simples (por ejemplo, $30 = 2 \cdot 3 \cdot 5$), frecuentemente también es útil factorizar matrices como productos de otras matrices. Cualquier representación de una matriz como producto de dos o más matrices se llama factorización de matrices.

Ejemplo 8.35.

$$\begin{bmatrix} 3 & -1 \\ 9 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}$$

es una factorización de una matriz.

Dependiendo del problema a resolver algunos tipos de factorizaciones pueden ser más útiles que otras; en esta sección, se presenta una factorización matricial que surge en la solución de sistemas de ecuaciones lineales mediante eliminación gaussiana y es particularmente adecuada para su implementación en computadora.

Se trata de la descomposición LU, que factoriza una matriz cuadrada A como el producto de dos matrices

$$A = LU$$

donde:

- L es una matriz triangular inferior (Lower) con 1's en la diagonal principal
- U es una matriz triangular superior (Upper)

Más adelante se estudiarán otras factorizaciones matriciales, también útiles. El tema es muy extenso y se le han dedicado libros enteros.

Se ilustrará la utilidad del método con un ejemplo

EJEMPLO 8.36. El siguiente ejemplo ilustra la idea básica.

Queremos resolver el sistema $Ax = b$, donde la incógnita es $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ y

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} 1 \\ -4 \\ 9 \end{bmatrix}$$

La matriz A tiene la siguiente factorización LU.

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

por lo que

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} \qquad \text{y} \qquad U = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix}$$

Resolvemos la ecuación en dos pasos, primero resolvemos la ecuación $Ly = b$:

$$\begin{cases} y_1 = 1 \\ 2y_1 + y_2 = -4 \\ -y_1 - 2y_2 + y_3 = 9 \end{cases}$$

se obtiene $y_1 = 1,\ y_2 = -6,\ y_3 = -2$

Por lo tanto, la solución de $Ly = b$ es

$$y_o = \begin{bmatrix} 1 \\ -6 \\ -2 \end{bmatrix}$$

Resolvemos $Ux = y_o$:

$$\begin{cases} 2x_1 + x_2 + 3x_3 = 1 \\ -3x_2 - 3x_3 = -6 \\ 2x_3 = -2 \end{cases}$$

se obtiene $x_3 = -1,\ x_2 = 3,\ x_1 = \frac{1}{2}$, por lo tanto la solución es

$$x_o = \begin{bmatrix} \frac{1}{2} \\ 3 \\ -1 \end{bmatrix},$$

ya que

$$Ax_o = LUx_o = L(Ux_o) = L(y_o) = b.$$

En el próximo capítulo se verá que esta descomposición puede ser muy útil para el cálculo de determinantes.

A continuación se explicará un método para obtener la descomposición LU, que solamente es adecuado para matrices cuadradas que pueden reducirse a forma escalonada sin usar intercambios de renglón, en otro caso es posible que no se pueda hacer la descomposición.

Los pasos básicos para obtener la descomposición LU son:

(1) Aplicar operaciones elementales de fila para convertir A en U
(2) Los multiplicadores utilizados forman los elementos de L

EJEMPLO 8.37 (Descomposición LU para una matriz 2×2). Sea $A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$.

Paso 1: Hacer cero en posición (2,1), el multiplicador es 2.

$$\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \xrightarrow{R_2 : R_2 - 2R_1} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

Ya tenemos

$$R = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix},$$

para obtener L colocamos el multiplicador (2) inmediatamente debajo del primer 1 de la matriz identidad.

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

Verificación:

$$LR = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 + 1 \cdot 0 & 1 \cdot 1 + 1 \cdot 0 \\ 2 \cdot 2 + 1 \cdot 0 & 2 \cdot 1 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} = A$$

EJEMPLO 8.38 (Descomposición LU para una matriz 3×3). Sea

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

La reducción por renglones de A se realiza del modo siguiente:

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \xrightarrow[\text{multiplicadores } 2 \text{ y } -1]{\substack{R_2 \to R_2 - 2 \cdot R_1 \\ R_3 \to R_3 - (-1) \cdot R_1}} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix}$$

$$\xrightarrow[\text{multiplicador } (-2)]{R_3 \to R_3 - (-2)} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix}$$

Ya tenemos

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

Ahora se toma la matriz identidad

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y los dos primeros multiplicadores se colocan en orden en la primera columna, debajo del 1; el segundo multiplicador se coloca en la segunda columna, debajo del 1:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}$$

Ya tenemos

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}$$

Verificación:

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

EJEMPLO 8.39 (Descomposición LU para una matriz 4×4).
Sea

$$A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix}$$

La reducción por renglones de A se realiza del modo siguiente:

$$\begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix} \xrightarrow{R_2-2R_1}\xrightarrow{R_3-R_1}\xrightarrow[\text{multiplicadores } 2,2,-3]{R_4+3R_1} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 8 & 7 & -16 \end{bmatrix}$$

$$\xrightarrow{R_3-\frac{1}{2}R_2}\xrightarrow[\text{multiplicadores } \frac{1}{2},8]{R_4-4R_2} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & -1 & -8 \end{bmatrix}$$

$$\xrightarrow[\text{multiplicador } -1]{R_4+R_3} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix}$$

Ya tenemos

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

Nuevamente tomamos la matriz identidad

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

y se colocan los multiplicadores en sus posiciones correspondientes y se obtiene

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

Verificación:

$$LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix} = A.$$

8.9. Lectura adicional: Factorización QR

Si A es una matriz $n \times n$ con columnas linealmente independientes entonces, la aplicación del proceso de Gram-Schmidt a dichas columnas produce una factorización de A muy útil, que permite escribirla como el producto de una matriz Q con columnas ortonormales y una matriz triangular superior R. Esta es la *factorización QR* y tiene aplicaciones en la aproximación numérica de autovalores y de matrices.

El proceso de la factorización QR es como sigue.

Sea

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

un matriz $n \times n$ tal que todas sus columnas

$$a_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad a_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \quad \cdots \quad a_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}$$

son linealmente independientes.

Sean

$$q_1 = \begin{bmatrix} q_{11} \\ q_{21} \\ \vdots \\ q_{n1} \end{bmatrix}, \quad q_2 = \begin{bmatrix} q_{12} \\ q_{22} \\ \vdots \\ q_{n2} \end{bmatrix}, \quad \cdots \quad q_n = \begin{bmatrix} q_{1n} \\ q_{2n} \\ \vdots \\ q_{nn} \end{bmatrix}$$

los vectores ortonormales obtenidos al aplicar a $a_1, a_2, \cdots, a_n$ el proceso de Gram-Schmidt.

Por el proceso de Gram-Schmidt se tiene que, para $1 \leq i \leq n$,

$$\operatorname{Span}\{a_1, \cdots, a_i\} = \operatorname{Span}\{q_1, \ldots, q_i\},$$

por lo tanto, existen escalares $r_{1i}, r_{2i}, \cdots, r_{ii}$ tales que

$$a_i = r_{1i}q_1 + r_{2i}q_2 + \cdots + r_{ii}q_i \qquad \text{para } i = 1, 2, \cdots, n,$$

en detalle

$$\begin{aligned}
\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} = a_1 &= r_{11}q_1 = \begin{bmatrix} r_{11}q_{11} \\ r_{11}q_{21} \\ \vdots \\ r_{11}q_{n1} \end{bmatrix} \\
\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} = a_2 &= r_{12}q_1 + r_{22}q_2 = \begin{bmatrix} r_{12}q_{11} + r_{22}q_{12} \\ r_{12}q_{21} + r_{22}q_{22} \\ \vdots \\ r_{12}q_{n1} + r_{22}q_{n2} \end{bmatrix} \\
&\vdots \\
\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = a_n &= r_{1n}q_1 + r_{2n}q_2 + \cdots + r_{nn}q_n = \begin{bmatrix} r_{1n}q_{11} + r_{2n}q_{12} + \cdots r_{nn}q_{1n} \\ r_{1n}q_{21} + r_{2n}q_{22} + \cdots r_{nn}q_{2n} \\ \vdots \\ r_{1n}q_{n1} + r_{2n}q_{n2} + \cdots r_{nn}q_{nn} \end{bmatrix}
\end{aligned}$$

y en forma matricial

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = QR,$$

donde

$$Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix} \qquad \text{y} \qquad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}.$$

donde

$$Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} \qquad \text{y} \qquad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}.$$

Por construcción la matriz Q tiene columnas ortonormales, por lo tanto se acaba de establecer el siguiente resultado

TEOREMA 8.40. *Sea A una matriz $n \times n$ con columnas linealmente independientes. Entonces A puede factorizarse como $A = QR$, donde Q es una matriz $n \times n$ con columnas ortonormales y R es una matriz $n \times n$ triangular superior.*

OBSERVACIÓN 8.41. Como la matriz Q tiene columnas ortonormales, se tiene que $Q^T Q = I$, la justificación de esta igualdad es como sigue.

Sea

$$q_j = \begin{bmatrix} q_{1j} \\ q_{2j} \\ \vdots \\ q_{nj} \end{bmatrix}$$

como los vectores $q_1, q_2, \cdots, q_n$ son ortonormales, se tiene que

$$\begin{bmatrix} q_{i1} & q_{i2} & \cdots & q_{in} \end{bmatrix} \begin{bmatrix} q_{1j} \\ q_{2j} \\ \vdots \\ q_{nj} \end{bmatrix} = \langle q_i, q_j \rangle = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases}$$

en términos de matrices

$$Q^T Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

EJEMPLO 8.42 (Ejemplo práctico de como obtener una factorización QR, caso 2×2).

Hallar la descomposición QR de la matriz

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}.$$

Lo primero es aplicar el proceso de ortonormalización de Gram-Schmidt a los vectores

$$a_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{y} \qquad a_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

al aplicarlo se obtienen los vectores

$$q_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} \qquad \text{y} \qquad q_2 = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

La matriz Q es

$$Q = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

$$R = Q^T A = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} \sqrt{5} & \frac{7}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} \end{bmatrix}$$

Verificación:

$$QR = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{5} & \frac{7}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$$

Ejemplo 8.43 (Ejemplo práctico de como obtener una factorización QR, caso 3×3).
Hallar la descomposición QR de la matriz

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

Al aplicar el proceso de ortonormalización de Gram-Schmidt a los vectores

$$a_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad a_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

se obtienen los vectores

$$q_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \qquad q_2 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} \end{bmatrix} \quad \text{y} \quad q_3 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

La matriz Q es

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

La matriz R es

$$R = Q^T A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} - \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}$$
$$= \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}$$

$$R = Q^T A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} - \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}.$$

Verificación:

$$QR = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} - \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} + \frac{\sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}}}{\sqrt{6}} & -\frac{1}{6} + \frac{\sqrt{\frac{2}{3}} - \frac{1}{\sqrt{6}}}{\sqrt{6}} \\ 1 & \frac{1}{2} - \frac{\sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}}}{\sqrt{6}} & \frac{7}{6} - \frac{\sqrt{\frac{2}{3}} - \frac{1}{\sqrt{6}}}{\sqrt{6}} \\ 0 & \sqrt{\frac{2}{3}}\left(\sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}}\right) & \frac{2}{3} + \sqrt{\frac{2}{3}}\left(\sqrt{\frac{2}{3}} - \frac{1}{\sqrt{6}}\right) \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = A.$$

Ejemplo 8.44 (Ejemplo práctico de como obtener una factorización QR, caso 4×4).
Hallar la factorización QR de la matriz

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

Al aplicar el proceso de ortonormalización de Gram-Schmidt a los vectores

$$a_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \end{bmatrix}, \qquad a_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \quad \text{y} \quad a_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

se obtienen los vectores

$$q_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \qquad q_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix} \qquad q_3 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\sqrt{\frac{2}{3}} \\ 0 \end{bmatrix} \quad \text{y} \quad q_4 = \begin{bmatrix} \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ -\frac{\sqrt{3}}{2} \end{bmatrix}.$$

La matriz Q es

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & 0 & -\sqrt{\frac{2}{3}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & 0 & 0 & -\frac{\sqrt{3}}{2} \end{bmatrix}$$

La matriz R es

$$R = Q^T A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\sqrt{\frac{2}{3}} & 0 \\ \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & 0 & \frac{1}{2\sqrt{3}} + \frac{\sqrt{3}}{2} \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \sqrt{\frac{3}{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}$$

$$R = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \sqrt{\frac{2}{3}} + \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & 0 & \frac{1}{2\sqrt{3}} + \frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \sqrt{\frac{3}{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}$$

Verificación:

$$Q\,R = \begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & 0 & -\sqrt{\frac{2}{3}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & 0 & 0 & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \sqrt{\frac{3}{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}.$$

Capítulo 9

Determinantes

9.1. Determinante de matrices 2 x 2 y 3 x 3

Ya vimos que si

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{9.1}$$

es una matriz 2×2 y si $ad - bc \neq 0$, entonces A es invertible y

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

Es muy natural hacerse la siguiente pregunta ¿Qué ocurre si $ad - bc = 0$?

Tratemos de aplicar un procedimiento similar a Gauss-Jordan en este caso:

$$\begin{aligned} \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] & \xrightarrow{R_1 \to c\,R_1} \xrightarrow{R_1 \to a\,R_1} \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ ac & ad & 0 & a \end{array}\right] \\ & \xrightarrow{R_2 \to R_2 - R_1} \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ 0 & ad - bc & -c & a \end{array}\right] \\ & = \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ 0 & 0 & -c & a \end{array}\right] \end{aligned}$$

El haber obtenido del lado izquierdo una matriz con la segunda fila nula implica que, en este caso, la matriz no es invertible.

Por lo tanto, la matriz (9.1) es invertible si y solo si $ad - bc \neq 0$.

Definición 9.1. Sea

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

una matriz 2×2.

El *determinante* de A se define por

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

Es usual denotar a $\det A$ por

$$|A| \quad \text{o} \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

(no confundir con valor absoluto)

Para el caso 2×2, se demostró que A es invertible si y sólo si $\det A \neq 0$.

Definición 9.2. Sea

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

una matriz 3×3.

El *determinante* de A se define por

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

Ejemplo 9.3 (Cálculo de un determinante 3×3). Sea

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

$$\begin{aligned} \det A &= 4 \det \begin{bmatrix} -5 & 1 \\ 6 & 9 \end{bmatrix} - 7 \det \begin{bmatrix} 3 & 1 \\ -8 & 9 \end{bmatrix} + (-2) \det \begin{bmatrix} 3 & -5 \\ -8 & 6 \end{bmatrix} \\ &= 4[(-5)(9) - (6)(1)] - 7[(3)(9) - (-8)(1)] - 2[(3)(6) - (-8)(-5)] \\ &= 4(-51) - 7(35) - 2(-22) \\ &= -204 - 245 + 44 \\ &= -405 \end{aligned}$$

Observación 9.4 (Otro método para calcular determinantes 3×3). De la definición de determinante 3×3 podemos deducir otra forma de calcularlo, por la definición

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} =& a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ =& a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ =& a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \end{aligned}$$

Una manera de recordar este método es la siguiente: Escribimos la matriz y copiamos a su lado derecho las dos primeras filas, respetando el orden, tal como se muestra en la siguiente figura

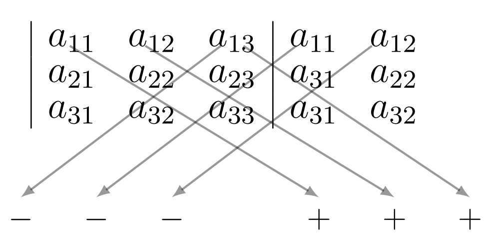


Figura 9.1. Cálculo determinante 3×3

Se calculan los seis productos indicados por las flechas, poniendo signo mas a los que corresponden con las flechas que apuntan hacia la derecha y signo menos a los que corresponden con las fechas que apuntan hacia la izquierda.

Ejemplo 9.5. Sea

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

usando el nuevo método.

$$\begin{aligned} \det A &= (3)(2)(4) + (5)(3)(-1) + (2)(4)(2) - (-1)(2)(2) - (2)(3)(3) - (4)(4)(5) \\ &= 24 - 15 + 16 + 4 - 18 - 80 \\ &= -69 \end{aligned}$$

Ejercicios.

Hallar los determinantes de las siguientes matrices

(a) $\begin{bmatrix} 7 & 9 & -5 \\ 9 & 3 & 1 \\ -8 & -8 & 10 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 2 & 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} -1 & 1 & 0 \\ 2 & 1 & 4 \\ 1 & 5 & 6 \end{bmatrix}$

(d) $\begin{bmatrix} 10 & 10 & -8 \\ -7 & 0 & -2 \\ 10 & 6 & 9 \end{bmatrix}$ (e) $\begin{bmatrix} 3 & -1 & 4 \\ 6 & 3 & 5 \\ 2 & -1 & 6 \end{bmatrix}$ (f) $\begin{bmatrix} -1 & 0 & 6 \\ 0 & 2 & 4 \\ 1 & 2 & -3 \end{bmatrix}$

(g) $\begin{bmatrix} 6 & -10 & 4 \\ 10 & 7 & 5 \\ 3 & 9 & 5 \end{bmatrix}$ (h) $\begin{bmatrix} -2 & 3 & 1 \\ 4 & 6 & 5 \\ 0 & 2 & 1 \end{bmatrix}$ (i) $\begin{bmatrix} 5 & -2 & 1 \\ 6 & 0 & 3 \\ -2 & 1 & 4 \end{bmatrix}$

9.2. Determinante de matrices n x n

Recordemos que el determinante de una matriz 3×3 fue definido en términos del determinante de una matriz 2×2.

El determinante de una matriz $n \times n$ se definirá de manera inductiva, es decir el determinante de una matriz 4 lo definiremos en términos del determinante de una matriz 3×3, el determinante de una matriz 5×5 se definirá en términos del determinante de una matriz 4×4 y así sucesivamente.

Definición 9.6. Sea A una matriz $n \times n$, el *menor* ij (que denotaremos por M_{ij}) de la matriz A es la matriz $(n-1) \times (n-1)$ que se obtiene de A al eliminar el renglón i y la columna j.

El *cofactor* ij de A, que se denotará por $A(i|j)$ se define por

$$A(i|j) = (-1)^{i+j} \det(M_{ij}).$$

Observación 9.7. No confundir A_{ij} (entrada ij de la matriz A) con $A(i|j)$ (el cofactor ij de la matriz A).

Ejemplo 9.8. Sea

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

Entonces

$$M_{13} = \begin{bmatrix} 0 & 1 \\ 6 & 3 \end{bmatrix} \qquad M_{32} = \begin{bmatrix} 2 & 4 \\ 0 & 5 \end{bmatrix}$$

y

$$A(1|3) = -6 \qquad A(3|2) = -10.$$

DEFINICIÓN 9.9. Sea $A = (a_{ij}$ una matriz $n \times n$.
El determinante de A se define por

$$\det A = a_{11}A_{(1|1)} + a_{12}A_{(1|2)} + \cdots + a_{1n}A_{(1|n)} = \sum_{k=1}^{n} a_{1k}A_{(1|k)}$$

La expresión anterior se conoce como *expansión por cofactores.*

EJEMPLO 9.10. Cálculo de $\det A$, donde

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

$$\begin{aligned} \det \begin{bmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 2 & 1 & 9 & 6 \\ 3 & 2 & 4 & 8 \end{bmatrix} &= a_{11}A(1|1) + a_{12}A(1|2) + a_{13}A(1|3) + a_{14}A(1|4) \\ &= 1 \det \begin{bmatrix} -1 & 3 & 4 \\ 1 & 9 & 6 \\ 2 & 4 & 8 \end{bmatrix} - 3 \det \begin{bmatrix} 0 & 3 & 4 \\ 2 & 9 & 6 \\ 3 & 4 & 8 \end{bmatrix} + 5 \det \begin{bmatrix} 0 & -1 & 4 \\ 2 & 1 & 6 \\ 3 & 2 & 8 \end{bmatrix} - 2 \begin{bmatrix} 0 & -1 & 3 \\ 2 & 1 & 9 \\ 3 & 2 & 4 \end{bmatrix} \\ &= 1(-92) - 3(-70) + 5(2) - 2(-16) = 160 \\ &= 160 \end{aligned}$$

Resulta claro que el cálculo del determinante de una matriz $n \times n$ puede ser laborioso; para calcular un determinante 4×4 deben calcularse cuatro determinantes 3×3. Para calcular un determinante de 5×5 deben calcularse cinco determinantes 4×4, lo que equivale a calcular veinte determinantes de 3×3. Más adelante veremos que existen algunas técnicas que permiten simplificar estos cálculos.

9.3. Propiedades del determinante

Veamos un ejemplo en el que es fácil calcular el determinante de una matriz.

EJEMPLO 9.11 (Determinante de una matriz triangular 4×4). Sea

$$A = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

una matriz 4×4 triangular inferior.

Se tiene

$$\begin{aligned}
\det A &= a_{11}A(1|1) + 0\,A(1|2) + 0\,A(1|3) + 0\,A(1|4) = a_{11}A_{11} \\
&= a_{11} \det \begin{bmatrix} a_{22} & 0 & 0 \\ a_{32} & a_{33} & 0 \\ a_{42} & a_{43} & a_{44} \end{bmatrix} \\
&= a_{11}a_{22} \det \begin{bmatrix} a_{33} & 0 \\ a_{43} & a_{44} \end{bmatrix} \\
&= a_{11}\, a_{22}\, a_{33}\, a_{44}.
\end{aligned}$$

En general se cumple.

PROPOSICIÓN 9.12. *Sea $A = (a_{ij})$ una matriz $n \times n$ triangular superior o inferior. Entonces:*

$$\det A = a_{11}\, a_{22}\, a_{33} \cdots a_{nn}$$

Para calcular un determinante, la expansión por cofactores se puede hacer por cualquier renglón o columna, mas precisamente.

PROPOSICIÓN 9.13.
Si $A = (a_{ij})$ es una matriz $n \times n$, entonces

$$\det A = \sum_{k=1}^{n} a_{ik} A_{(i|k)} \quad \textit{(para cualquier renglón } i\textit{)}$$

$$\det A = \sum_{k=1}^{n} a_{kj} A_{(k|j)} \quad \textit{(para cualquier columna } j\textit{)}$$

Además se tiene la siguiente propiedad fundamental.

TEOREMA 9.14. *Sean A y B dos matrices $n \times n$. Entonces:*

$$\det AB = \det A \det B$$

Además se cumplen las siguientes propiedades fundamentales.

TEOREMA 9.15 (Propiedades fundamentales del determinante). *Sean A y B matrices cuadradas, entonces*

(1) *Si cualquier renglón o columna de A es el vector cero, entonces* $\det A = 0$.
(2) *Si cualquier renglón (columna) de A se multiplica por un escalar c, entonces* $\det A$ *se multiplica por c.*

(3) *Si A y B son iguales excepto en la columna j, y C tiene en la columna j la suma de las columnas correspondientes de A y B, entonces* $\det C = \det A + \det B$, *en detalle*

$$\det \begin{bmatrix} a_{11} & \dots & a_{1j}+b_{1j} & \cdots & a_{1n} \\ a_{21} & \dots & a_{22}+b_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{n2}+b_{n2} & \cdots & a_{nn} \end{bmatrix} = \det \begin{bmatrix} a_{11} & \dots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \dots & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \det \begin{bmatrix} a_{11} & \dots & b_{1j} & \cdots & a_{1n} \\ a_{21} & \dots & b_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & b_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Tomando en cuenta la propiedad (2), esto quiere decir que, como función de las columnas de la matriz, la función determinante es lineal en cada columna. Lo mismo ocurre con las filas.

Por esto se dice que la función determinate es multilineal.

(4) *El intercambio de dos renglones o columnas distintos multiplica* $\det A$ *por* -1.

(5) *Si un renglón (columna) es múltiplo de otro, entonces* $\det A = 0$.

(6) *Sumar un múltiplo escalar de un renglón a otro no cambia* $\det A$.

(7) $\det A = \det A^T$.

Ejercicios.

Hallar los determinantes de las siguientes matrices

(a) $\begin{bmatrix} -2 & -10 & 7 & 0 \\ 0 & -5 & 4 & -1 \\ 0 & -10 & 0 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 3 & 1 \\ 0 & 1 & 4 & 2 \\ 0 & 0 & 1 & 5 \\ 1 & 2 & 3 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} -3 & 0 & 0 & 0 \\ -4 & 7 & 0 & 0 \\ 5 & 8 & -1 & 0 \\ 2 & 3 & 0 & 6 \end{bmatrix}$

(d) $\begin{bmatrix} -6 & 8 & -5 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 \\ -5 & 0 & -5 & -6 & 0 \\ 0 & -8 & 0 & 0 & 2 \\ 0 & -7 & 0 & 2 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} 2 & 3 & -1 & 4 & 5 \\ 0 & 1 & 7 & 8 & 2 \\ 0 & 0 & 4 & -1 & 5 \\ 0 & 0 & 0 & -2 & 8 \\ 0 & 0 & 0 & 0 & 6 \end{bmatrix}$ (f) $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 \\ -9 & 1 & 0 & 0 & 0 \\ -6 & 0 & 0 & -2 & -7 \\ 9 & -9 & 0 & -5 & 0 \end{bmatrix}$

(g) $\begin{bmatrix} -8 & 0 & 0 & -10 & 0 \\ 0 & -8 & 0 & -7 & 1 \\ 0 & 9 & -3 & 0 & -4 \\ -5 & 0 & 7 & 5 & 5 \\ -2 & 0 & -10 & 3 & -7 \end{bmatrix}$

9.4. Determinantes y matrices inversas

A partir de las propiedades ya enunciadas del determinante, podemos establecer el siguiente resultado.

PROPOSICIÓN 9.16. *Sea A una matriz cuadrada.*
Si A es invertible, entonces $\det A \neq 0$ *y*

$$\det A^{-1} = \frac{1}{\det A}.$$

DEMOSTRACIÓN. Como A es invertible, existe una matriz A^{-1} tal que

$$A\,A^{-1} = I,$$

luego

$$\det\left(A\,A^{-1}\right) = \det I.$$

Como el determinante del producto es el producto de los determinantes y el determinante de la identidad es igual a 1, entonces

$$(\det A)\left(\det A^{-1}\right) = 1,$$

por lo tanto,

$$\det A \neq 0 \qquad \text{y} \qquad \det A^{-1} = \frac{1}{\det A}.$$

□

Más adelante veremos que la condición $\det A \neq 0$ es suficiente para que A sea invertible.

DEFINICIÓN 9.17 (Adjunta). Sea A una matriz $n \times n$. La *adjunta* de A, denotada por adj A, es la transpuesta de la matriz de cofactores:

$$\text{adj } A = (A(j|i))$$

Se cumplen los siguientes resultados.

TEOREMA 9.18.
Sea A una matriz $n \times n$, entonces

(1) $A \cdot \text{adj}A = (\det A)I$
(2) *A es invertible si y solo si $\det A \neq 0$ y en este caso,*

$$A^{-1} = \frac{1}{\det A}\text{adj}A.$$

EJEMPLO 9.19 (Inversa de una matriz 2×2).
Sea

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

una matriz 2×2 tal que $\det A = ad - bc \neq 0$.

Los menores son

$$M_{11} = d \qquad M_{12} = c \qquad M_{21} = b \qquad M_{22} = a$$

Recordemos que el cofactor ij es $A(i|j) = (-1)^{i+j}\det M_{ij}$, luego

$$A(1|1) = d \qquad A(1|2) = -c \qquad A(2|1) = -b \qquad A(2|2) = a$$

La inversa es

$$\frac{1}{\det A}\text{adj}A = \frac{1}{\det A}\begin{bmatrix} A(1|1) & A(2|1) \\ A(1|2) & A(2|2) \end{bmatrix} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

que coincide con la fórmula (8.1).

Ejemplo 9.20. En el Ejemplo 8.17 vimos que la inversa de la matriz es

$$A = \begin{bmatrix} 2 & 4 & 0 \\ 0 & 2 & 1 \\ 3 & 0 & 2 \end{bmatrix} \quad \text{es la matriz} \quad \frac{1}{20}\begin{bmatrix} 4 & -8 & 4 \\ 3 & 4 & -2 \\ -6 & 12 & 4 \end{bmatrix}$$

Ahora vamos a calcular la inversa usando el método recién estudiado.
El determinante de A es

$$\det A = (2)(2)(2) + (4)(1)(3) + 0 - 0 - 0 - 0 = 8 + 12 = 20$$

$$M_{11} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad M_{12} = \begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix} \quad M_{13} = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}$$

$$M_{21} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \quad M_{22} = \begin{bmatrix} 2 & 0 \\ 3 & 2 \end{bmatrix} \quad M_{23} = \begin{bmatrix} 2 & 4 \\ 3 & 0 \end{bmatrix}$$

$$M_{31} = \begin{bmatrix} 4 & 0 \\ 2 & 1 \end{bmatrix} \quad M_{32} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad M_{33} = \begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}$$

La matriz adjunta de A es

$$\text{adj}A = (A(j|i)) = (\,(-1)^{i+j}\det M_{ji}\,) = \begin{bmatrix} 4 & -8 & 4 \\ -(-3) & 4 & -2 \\ -6 & -(-12) & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -8 & 4 \\ 3 & 4 & -2 \\ -6 & 12 & 4 \end{bmatrix}$$

de donde se obtiene

$$A^{-1} = \frac{1}{20}\begin{bmatrix} 4 & -8 & 4 \\ 3 & 4 & -2 \\ -6 & 12 & 4 \end{bmatrix}.$$

El siguiente teorema es muy importante y resume varios de los resultados obtenidos.

Teorema 9.21. *Sea A una matriz $n \times n$.*
Las siguientes afirmaciones son equivalentes:

(1) *A es invertible.*
(2) *La única solución a $Ax = 0$ es $x = 0$.*
(3) *$Ax = b$ tiene solución única para cada b.*
(4) *A es equivalente por renglones a I_n.*
(5) *La forma escalonada de A tiene n pivotes.*
(6) $\det A \neq 0$.

Ejercicios.

Utilizar los métodos de este capítulo para determinar si la matriz dada es invertible, de ser así, calcule la inversa.

(a) $\begin{bmatrix} 7 & -8 \\ 9 & 9 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 6 \\ -4 & -8 \end{bmatrix} 3$ (c) $\begin{bmatrix} -3 & 9 \\ 7 & -21 \end{bmatrix}$

(d) $\begin{bmatrix} -3 & 9 \\ 7 & -21 \end{bmatrix}$ (e) $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 1 & -1 \end{bmatrix}$ (f) $\begin{bmatrix} 9 & 3 & 9 \\ 6 & -10 & 4 \\ 10 & 7 & 5 \end{bmatrix}$

9.5. Regla de Cramer

TEOREMA 9.22 (Regla de Cramer).
Sea A una matriz $n \times n$ tal que $\det A \neq 0$.
Entonces la solución única al sistema $Ax = b$ está dada por:

$$x_j = \frac{\det A_j}{\det A} \quad \textit{para } j = 1, 2, \dots, n$$

donde A_j es la matriz obtenida al reemplazar la columna j de A por b.

EJEMPLO 9.23. Utilizar la regla de Cramer para resolver el sistema

$$\begin{cases} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{cases}$$

$$D = \det \begin{bmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{bmatrix} = 6 \qquad D_1 = \det \begin{bmatrix} 18 & 4 & 6 \\ 24 & 5 & 6 \\ 4 & 1 & -2 \end{bmatrix} = 24$$

$$D_2 = \det \begin{bmatrix} 2 & 18 & 6 \\ 4 & 24 & 6 \\ 3 & 4 & -2 \end{bmatrix} = -12 \qquad D_3 = \begin{bmatrix} 2 & 4 & 18 \\ 4 & 5 & 24 \\ 3 & 1 & 4 \end{bmatrix} = 18$$

9.6. Ejercicios varios

(1) Para $A = \begin{bmatrix} 1 & -1 & 3 \\ 4 & 1 & 6 \\ 2 & 0 & -2 \end{bmatrix}$ verifique que $\det A^{-1} = \frac{1}{\det A}$.

(2) Utilizar los métodos de este capítulo para determinar si la matriz dada es invertible, de ser así, calcule la inversa.

(a) $\begin{bmatrix} 3 & -5 \\ -8 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 2 & 1 & 4 \end{bmatrix}$

(d) $\begin{bmatrix} 5 & 0 & 0 & 0 \\ -2 & -10 & 7 & 0 \\ 0 & -5 & 4 & -1 \\ 0 & -10 & 0 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 0 & 1 & 7 \\ 0 & 0 & -8 & -5 \\ 0 & 0 & -7 & 0 \\ 0 & 10 & 0 & -5 \end{bmatrix}$

(3) ¿Para cuáles valores de α la matriz $\begin{bmatrix} \alpha+1 & -3 \\ 5 & 1-\alpha \end{bmatrix}$ no esinvertible?

(4) ¿Para qué valores de α la matriz $\begin{bmatrix} -\alpha & \alpha-1 & \alpha+1 \\ 1 & 2 & 3 \\ 2-\alpha & \alpha+3 & \alpha+7 \end{bmatrix}$ no tiene inversa?

(5) Justificar la siguiente afirmación: Si la matriz $n \times n$ A de no es invertible, entonces $(A)(\text{adj } A)$ es la matriz cero.

(6) Muestre que si A es triangular, entonces $\det A \neq 0$ si y sólo si todos los elementos en la diagonal de A son diferentes de cero.

(7) Resolver el sistema dado usando la regla de Cramer.

(a) $\begin{cases} 7x_1 - 8x_2 = 3 \\ 9x_1 + 9x_2 = -8 \end{cases}$

(b) $\begin{cases} 3x_1 - x_2 = 0 \\ 4x_1 + 2x_2 = 5 \end{cases}$

(c) $\begin{cases} 2x_1 + 2x_2 + x_3 = 7 \\ x_1 + 2x_2 + x_3 = 0 \\ -x_1 + x_2 + 3x_3 = 1 \end{cases}$

(d) $\begin{cases} 2x_1 + 5x_2 - x_3 = -1 \\ 4x_1 + x_2 + 3x_3 = 3 \\ -2x_1 + 2x_2 = 0 \end{cases}$

(e) $\begin{cases} 6x_1 - 10x_2 + 4x_3 = -2 \\ 10x_1 + 7x_2 + 5x_3 = 3 \\ 3x_1 + 9x_2 + 5x_3 = -7 \end{cases}$

(f) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 6 \\ 2x_1 - x_3 - x_4 = 4 \\ 3x_3 + 6x_4 = 3 \\ x_1 - x_4 = 5 \end{cases}$

(g) $\begin{cases} x_1 - x_4 = 7 \\ 2x_2 + x_3 = 2 \\ 4x_1 + x_2 = -3 \\ 3x_3 - 5x_4 = 2 \end{cases}$

Capítulo 10

Transformaciones lineales

10.1. Conceptos básicos y ejemplos

Definición 10.1. Sean V y W espacios vectoriales.
Una *transformación lineal* (o *aplicación lineal* u *operador lineal*) es una función

$$T : V \to W$$

tal que, para todo par de vectores $x, y \in V$ y todo escalar $\alpha \in \mathbb{R}$ se tiene que

(1) $T(x + y) = T(x) + T(y)$
(2) $T(\alpha x) = \alpha T(x)$

Veamos algunos ejemplos

Ejemplo 10.2.

(1) Sea $T : \mathbb{R} \to \mathbb{R}$ una transformación lineal, entonces

$$T(x) = T(x \cdot 1) = x \cdot T(1),$$

por lo tanto, toda transformación lineal $T : \mathbb{R} \to \mathbb{R}$ tiene la forma

$$T(x) = ax \qquad \text{donde } a = T(1).$$

(2) Sea $T : \mathbb{R}^2 \to \mathbb{R}$ una transformación lineal, entonces

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = T\left(x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = x_1 T\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = ax_1 + bx_2,$$

donde $a = T\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $b = T\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Notar que

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(3) Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal, entonces

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = T\left(x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = x_1 T\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Como $T\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $T\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ son vectores de $\mathbb{R}^2$, existen números reales a_{11}, a_{12}, a_{21} y a_{22} tales que

$$T\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}a_{11}\\a_{21}\end{bmatrix}\quad \text{y}\quad T\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}a_{12}\\a_{22}\end{bmatrix},$$

luego

$$\begin{aligned}T\begin{bmatrix}x_1\\x_2\end{bmatrix} &= x_1\begin{bmatrix}a_{11}\\a_{21}\end{bmatrix}+x_2\begin{bmatrix}a_{12}\\a_{22}\end{bmatrix}\\ &=\begin{bmatrix}a_{11}x_1+a_{12}x_2\\x_2a_{21}+x_2a_{22}\end{bmatrix}\\ &=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}.\end{aligned}$$

En conclusión:

$$T\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

es el producto de la matriz 2×2 $\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}$, por el vector $\begin{bmatrix}x_1\\x_2\end{bmatrix}$, además las columnas de la matriz son los vectores

$$T\begin{bmatrix}1\\0\end{bmatrix}\quad \text{y}\quad T\begin{bmatrix}0\\1\end{bmatrix}.$$

(4) Si $\mathcal{P}$ es el espacio de los polinomios, la función $T:\mathcal{P}\to\mathcal{P}$ por definida por

$$T(a_o+a_1x+a_2x^2+\cdots a_nx^n)=a_1+2a_2x+\cdots na_nx^{n-1}$$

es una transformación lineal.

(5) Si $\mathcal{C}^\infty$ es el conjunto de las funciones de $\mathbb{R}$ en $\mathbb{R}$ que son diferenciables un número infinito de veces, entonces $\mathcal{C}^\infty$ es un espacio vectorial y la derivada

$$D:\mathcal{C}^\infty\to\mathcal{C}^\infty \qquad Df=f'$$

es una transformación lineal.

(6) Sea A una matriz $n\times m$, la función

$$T:\mathbb{R}^m\to\mathbb{R}^n$$

definida por

$$Tx=Ax,$$

es decir

$$T\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots\vdots\vdots&\vdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}a_{11}x_1&+&a_{12}x_2&+&\cdots&+&a_{1n}x_n\\a_{21}x_1&+&a_{22}x_2&+&\cdots&+&a_{2n}x_n\\\vdots&+&\vdots\vdots\vdots&+&\vdots&+&\vdots\\a_{m1}x_1&+&a_{m2}x_2&+&\cdots&+&a_{mn}x_n\end{bmatrix}$$

es una transformación lineal.

Más adelante veremos que toda transformación lineal $T:\mathbb{R}^n\to\mathbb{R}^m$ es de esta forma.

Es muy importante tener en cuenta el siguiente resultado, que sigue de que si T es lineal, entonces

$$T(0) = T(0+0) = T(0) + T(0)$$

PROPOSICIÓN 10.3. *Si T es una transformación lineal, entonces*

$$T(0) = 0.$$

OBSERVACIÓN 10.4. En la proposición anterior el símbolo 0 se utilizó para dos vectores no necesariamente iguales, el primer 0 se refiere al vector nulo del dominio y el segundo al vector nulo del espacio de llegada.

Ejercicios.

Determinar su la transformación indicada es lineal, justifique su respuesta.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^1;\ T\begin{bmatrix} x \\ y \end{bmatrix} = x$

(b) $T : \mathbb{R}^2 \to \mathbb{R}^2;\ T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$

(c) $T : \mathbb{R}^2 \to \mathbb{R}^2;\ T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ y \end{bmatrix}$

(d) $T : \mathbb{R}^2 \to \mathbb{R}^1;\ T\begin{bmatrix} x \\ y \end{bmatrix} = x + 1$

(e) $T : \mathbb{R}^3 \to \mathbb{R}^2;\ T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$

(f)$T : \mathbb{R}^3 \to \mathbb{R}^2;\ T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$

(g) $T : \mathbb{R}^n \to \mathbb{R}^2;\ T\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} |x_4| \\ x_1 \end{bmatrix}$

(h) $T : \mathbb{R}^4 \to \mathbb{R}^{2\times 2};\ T\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x & z \\ y & w \end{bmatrix}$

(i) $T : \mathcal{P}_2 \to \mathcal{P}_1;\ T(a_0 + a_1x + a_2x^2) = a_1 + a_2x$

(j) $T : \mathbb{R} \to \mathbb{R}^n;\ T(x) = \begin{bmatrix} x \\ x \\ \vdots \\ x \end{bmatrix}$

10.2. Matriz asociada a una transformación lineal

Supongamos que $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal, vamos a proceder de manera análoga a como se hizo en (3) del Ejemplo 10.2 para ver que forma tiene T.

Tenemos que

$$\begin{aligned} T\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} &= T\left(x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right) \\ &= x_1 T\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 T\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n T\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \end{aligned}$$

también tenemos que

$$T\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}=\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{bmatrix}\quad T\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}=\begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{bmatrix}\quad\cdots\quad T\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}=\begin{bmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{bmatrix}$$

sustituyendo

$$T\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}a_{11}x_1&+&a_{12}x_2&+&\cdots&+&a_{1n}x_n\\a_{21}x_1&+&a_{22}x_2&+&\cdots&+&a_{2n}x_n\\\vdots&+&\vdots\vdots\vdots&+&\vdots&+&\vdots\\a_{m1}x_1&+&a_{m2}x_2&+&\cdots&+&a_{mn}x_n\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots\vdots\vdots&\vdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

Hemos probado

PROPOSICIÓN 10.5. *A cada transformación lineal $T:\mathbb{R}^n\to\mathbb{R}^m$ le corresponde una y solo una matriz $m\times n$ y recíprocamente, a cada matriz $m\times n$ le corresponde una y solo una transformación lineal $T:T:\mathbb{R}^n\to\mathbb{R}^m$.*

Ahora pasaremos a una situación más general, hemos venido trabajando solamente con la bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ (ver página 74 para la definición de base canónica), para ser más precisos, debemos indicar que la matriz a la que se refiere la Proposición 10.5 es la matriz de la transformación en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. No hay nada que impida considerar otras bases y, tal como se verá más adelante, se presentan situaciones en la que es conveniente considerar otras bases.

Supongamos que tenemos un espacio vectorial V de dimensión n, un espacio vectorial W de dimensión m y $T:V\to W$ una transformación lineal.

Sean $\mathfrak{B}_1=\{v_1,\dots,v_n\}$ una base de V y $\mathfrak{B}_2=\{w_1,\dots,w_m\}$ una base de W, en este contexto es usual referirse a estas bases como *bases ordenadas*, ya que aunque en un conjunto no importa el orden en que coloquemos los elementos, el orden en que se colocan los elementos de la base será relevante en lo que sigue.

Sea $x\in V$, entonces existen escalares $x_1,x_2,\dots,x_n$ tales que

$$x=x_1v_1+x_2v_2+\cdots x_nv_n$$

y, para cada j, $1\le j\le n$ existen escalares $a_{1j},a_{2j},\dots,a_{mj}$ tales que

$$Tx_j=a_{1j}w_1+a_{2j}w_2+\cdots+a_{mj}w_m,$$

en la notación referente a bases (ver página 75)

$$x_{\mathfrak{B}_1}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\quad\text{y}\quad(Tx_j)_{\mathfrak{B}_2}=\begin{bmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{bmatrix}$$

por ser T lineal tenemos que

$$\begin{aligned} Tx =& x_1 Tv_1 + x_2 Tv_2 + \cdots + x_n Tv_n \\ =& x_1(a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m) + \cdots + x_n(a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m \\ =& (a_{11}x_1 + \cdots + a_{1n}x_n)w_1 + (a_{21}x_1 + \cdots + a_{2n}x_n)w_2 + \cdots + (a_{m1}x_1 + \cdots + a_{mn}x_n)w_m \end{aligned}$$

lo que implica que

$$\begin{aligned} (Tx)_{\mathfrak{B}_2} =& \begin{bmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ \vdots & + & \vdots\vdots\vdots & + & \vdots & + & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots\vdots\vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ =& A\, x_{\mathfrak{B}_1}, \end{aligned}$$

donde

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

es la matriz cuyas columnas son las coordenadas de la imagen bajo T de los vectores de la base de V con respecto a la base de W.

El cálculo anterior nos lleva al siguiente resultado.

Teorema 10.6. *Sean V un espacio vectorial de dimensión n, W un espacio vectorial de dimensión m, $T : V \to W$ una transformación lineal y sean $\mathfrak{B}_1 = \{v_1, \dots, v_n\}$ una base ordenada de V y $\mathfrak{B}_2 = \{w_1, \dots, w_m\}$ una base ordenada de W. Entonces existe una matriz $m \times n$, A tal que*

$$(Tx)_{\mathfrak{B}_2} = A\, x_{\mathfrak{B}_1}.$$

Además, fijadas las bases $\mathfrak{B}_1$ y $\mathfrak{B}_2$, la matriz A es única.

La matriz A que aparece en el teorema anterior la denotaremos por $[T]_{\mathfrak{B}_1\mathfrak{B}_2}$ y se le suele llamar matriz de T son respecto a las bases $\mathfrak{B}_1$ $\mathfrak{B}_2$; es importante notar que la $j-$*ésima* columna de esta matriz es $(Tv_j)_{\mathfrak{B}_2}$.

Cuando se trata de una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ y hay una referencia a la matriz de T, sin especificar las bases, se debe entender que se están usando las bases canónicas y la matriz se denota simplemente por $[T]$ y se le suele llamar la *matriz estándar* de T.

También se cumple el siguiente resultado, queda como ejercicio su justificación.

Teorema 10.7. *Sean V un espacio vectorial de dimensión n, W un espacio vectorial de dimensión m, A una matriz $m \times n$ y sean $\mathfrak{B}_1 = \{v_1, \dots, v_n\}$ una base ordenada de V y $\mathfrak{B}_2 = \{w_1, \dots, w_m\}$*

una base ordenada de W. Entonces existe una única transformación lineal $T : V \to W$ tal que

$$[T]_{\mathfrak{B}_1 \mathfrak{B}_2} = A.$$

En resumen, existe una correspondencia biunívoca entra las matrices $m \times n$ y las transformaciones lineales de un espacio vectorial de dimensión n en un espacio vectorial de dimensión n.

Ejemplo 10.8. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ el operador lineal definido por

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ 2x - 5y \end{bmatrix}$$

Hallar la matriz de estándar de T.

Se tiene que

$$T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{y} \quad T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$$

por lo tanto, la matriz estándar de T es

$$[T] = \begin{bmatrix} 1 & 1 \\ 2 & -5 \end{bmatrix}.$$

Ejemplo 10.9. Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ el operador lineal definido por

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + \sqrt{3}y - z \\ 2x - 5y \\ y - z \end{bmatrix}$$

Hallar la matriz de estándar de T.

Se tiene que

$$T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad \text{y} \quad T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{3} \\ -5 \\ 1 \end{bmatrix} \quad \text{y} \quad T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}$$

por lo tanto, la matriz estándar de T es

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

Ejemplo 10.10. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ el operador lineal definido por

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ 2x - 5y \end{bmatrix}$$

Hallar $[T]_{\mathcal{B}\mathcal{B}}$ donde

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\}.$$

Para esto calculamos

$$T \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Por otra parte,

$$T\begin{bmatrix}0\\2\end{bmatrix}=\begin{bmatrix}2\\-10\end{bmatrix}=2\begin{bmatrix}1\\1\end{bmatrix}+(-6)\begin{bmatrix}0\\2\end{bmatrix}$$

En consecuencia,

$$[T]_{\mathcal{B}\mathcal{B}}=\begin{bmatrix}2 & 2\\-\frac{5}{2} & -6\end{bmatrix}$$

EJEMPLO 10.11. Sea $T:\mathbb{R}^4\to\mathbb{R}^4$ la transformación lineal definida por

$$T\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}x_1+3x_4\\-\frac{1}{2}x_1+\frac{5}{2}x_2\\-10x_3+x_4\\5x_4\end{bmatrix}$$

Hallar la matriz de estándar de T.

Se tiene que Se tiene que

$$T\begin{bmatrix}1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}1\\-\frac{1}{2}\\0\\0\end{bmatrix}\quad\text{y}\quad T\begin{bmatrix}0\\1\\0\\0\end{bmatrix}=\begin{bmatrix}0\\\frac{5}{2}\\0\\0\end{bmatrix}\quad\text{y}\quad T\begin{bmatrix}0\\0\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\-10\\0\end{bmatrix}\quad\text{y}\quad T\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}3\\0\\1\\5\end{bmatrix}$$

por lo tanto, la matriz estándar de T es

$$\begin{bmatrix}1 & 0 & 0 & 3\\-\frac{1}{2} & \frac{5}{2} & 0 & 0\\0 & 0 & -10 & 1\\0 & 0 & 0 & 5\end{bmatrix}.$$

EJEMPLO 10.12. Sea $T:\mathbb{R}\to\mathbb{R}^4$ el operador lineal definido por

$$Tu=\begin{bmatrix}2u\\-5u\\3u\\0\end{bmatrix}$$

Hallar la matriz estándar de T.

En este caso la matriz es

$$[T]=\begin{bmatrix}2\\-5\\3\\0\end{bmatrix}.$$

EJEMPLO 10.13. Sea $T:\mathbb{R}^2\to\mathbb{R}$ el operador lineal definido por

$$T\begin{bmatrix}u\\v\end{bmatrix}=5u-3v$$

Hallar la matriz estándar de T.

Se tiene que

$$T\begin{bmatrix}1\\0\end{bmatrix} = 5 \quad \text{y} \quad T\begin{bmatrix}0\\1\end{bmatrix} = -3$$

por lo tanto, la matriz de T con respecto a la base canónica es

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

EJEMPLO 10.14. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal definida por

$$T\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}\frac{5}{2}x_1 - \frac{1}{2}x_2\\ -\frac{1}{2}x_1 + \frac{5}{2}x_2\end{bmatrix}$$

y sean $\mathfrak{B}_1$ la base canónica de $\mathbb{R}^2$ y $\mathfrak{B}_2$ la base formada por los vectores $\begin{bmatrix}1\\1\end{bmatrix}$ y $\begin{bmatrix}1\\-1\end{bmatrix}$

Hallar $[T]_{\mathfrak{B}_1\mathfrak{B}_1}$, $[T]_{\mathfrak{B}_1\mathfrak{B}_2}$, $[T]_{\mathfrak{B}_2\mathfrak{B}_1}$ y $[T]_{\mathfrak{B}_2\mathfrak{B}_2}$.

Recordemos que si $\mathfrak{B}_1 = \{v_1, v_2\}$ y $\mathfrak{B}_2 = \{w_1, w_2\}$, entonces $[T]_{\mathfrak{B}_1\mathfrak{B}_2}$ es la matriz cuyas columnas son los vectores $T(v_1)_{\mathfrak{B}_2}$ y $T(v_2)_{\mathfrak{B}_2}$.

En este caso

$$v_1 = \begin{bmatrix}1\\0\end{bmatrix} \quad v_2 = \begin{bmatrix}0\\1\end{bmatrix} \quad w_1 = \begin{bmatrix}1\\1\end{bmatrix} \quad w_2 = \begin{bmatrix}1\\-1\end{bmatrix}$$

y

$$Tv_1 = \begin{bmatrix}\frac{5}{2}\\-\frac{1}{2}\end{bmatrix} \quad Tv_2 = \begin{bmatrix}-\frac{1}{2}\\\frac{5}{2}\end{bmatrix} \quad Tw_1 = \begin{bmatrix}2\\2\end{bmatrix} \quad Tw_2 = \begin{bmatrix}3\\-3\end{bmatrix}$$

(1) $[T]_{\mathfrak{B}_1\mathfrak{B}_1}$:

Esta es la más fácil de todas, ya que se trata de la matriz estándar

$$(v_1)_{\mathfrak{B}_1} = \begin{bmatrix}1\\0\end{bmatrix} \quad (v_2)_{\mathfrak{B}_1} = \begin{bmatrix}0\\1\end{bmatrix} \qquad (Tv_1)_{\mathfrak{B}_1} = \begin{bmatrix}\frac{5}{2}\\-\frac{1}{2}\end{bmatrix} \quad (Tv_2)_{\mathfrak{B}_1} = \begin{bmatrix}-\frac{1}{2}\\\frac{5}{2},\end{bmatrix}$$

por lo tanto,

$$[T]_{\mathfrak{B}_1\mathfrak{B}_1} = \begin{bmatrix}\frac{5}{2} & -\frac{1}{2}\\-\frac{1}{2} & \frac{5}{2}.\end{bmatrix}$$

(2) $[T]_{\mathfrak{B}_1\mathfrak{B}_2}$:

Debemos hallar $(Tv_1)_{\mathfrak{B}_2}$ y $(Tv_2)_{\mathfrak{B}_2}$.

Si $(Tv_1)_{\mathfrak{B}_2} = \begin{bmatrix}a\\b\end{bmatrix}$ y $(Tv_2)_{\mathfrak{B}_2} = \begin{bmatrix}c\\d\end{bmatrix}$ se debe cumplir $Tv_1 = aw_1 + bw_2$ y $Tv_2 = cw_1 + dw_2$, es decir,

$$\begin{bmatrix}\frac{5}{2}\\-\frac{1}{2}\end{bmatrix} = a\begin{bmatrix}1\\1\end{bmatrix} + b\begin{bmatrix}1\\-1\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}-\frac{1}{2}\\\frac{5}{2}\end{bmatrix} = c\begin{bmatrix}1\\1\end{bmatrix} + d\begin{bmatrix}1\\-1\end{bmatrix},$$

de donde

$$\begin{cases} a+b &= \frac{5}{2}\\ a-b &= -\frac{1}{2}\end{cases} \quad \text{y} \quad \begin{cases} c+d &= -\frac{1}{2}\\ c-d &= \frac{5}{2}\end{cases}$$

usando el método de Gauss-Jordan "por duplicado"

$$\left[\begin{array}{cc|cc} 1 & 1 & \frac{5}{2} & -\frac{1}{2} \\ 1 & -1 & -\frac{1}{2} & \frac{5}{2} \end{array}\right]$$

$$\begin{aligned} \left[\begin{array}{cc|cc} 1 & 1 & \frac{5}{2} & -\frac{1}{2} \\ 1 & -1 & -\frac{1}{2} & \frac{5}{2} \end{array}\right] & \xrightarrow{R_2 \to R_2 - R_1} \left[\begin{array}{cc|cc} 1 & 1 & \frac{5}{2} & -\frac{1}{2} \\ 0 & -2 & -3 & 3 \end{array}\right] \\ & \xrightarrow{R_2 \to -\frac{1}{2} R_2} \left[\begin{array}{cc|cc} 1 & 1 & \frac{5}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{3}{2} & -\frac{3}{2} \end{array}\right] \\ & \xrightarrow{R_1 \to -R_1 - R_2} \left[\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{3}{2} \end{array}\right] \end{aligned}$$

luego,

$$(Tv_1)_{\mathfrak{B}_2} = \begin{bmatrix} 1 \\ \frac{3}{2} \end{bmatrix} \quad \text{y} \quad (Tv_2)_{\mathfrak{B}_2} = \begin{bmatrix} 1 \\ -\frac{3}{2} \end{bmatrix}$$

y

$$[T]_{\mathfrak{B}_1 \mathfrak{B}_1} = \begin{bmatrix} 1 & 1 \\ \frac{3}{2} & -\frac{3}{2} \end{bmatrix}$$

Verificación: Para verificar que todos los cálculos estuvieron correctos debemos comprobar que se cumple la igualdad

$$Tx_{\mathfrak{B}_2} = [T]_{\mathfrak{B}_1 \mathfrak{B}_1} x_{\mathfrak{B}_1} = \begin{bmatrix} 1 & 1 \\ \frac{3}{2} & -\frac{3}{2} \end{bmatrix} x_{\mathfrak{B}_1}$$

Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ un vector de $\mathbb{R}^2$, entonces

$$x_{\mathfrak{B}_1} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

y

$$Tx_{\mathfrak{B}_2} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{si} \quad \begin{bmatrix} \frac{5}{2}x_1 - \frac{1}{2}x_2 \\ -\frac{1}{2}x_1 + \frac{5}{2}x_2 \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} a+b \\ a-b \end{bmatrix}$$

Resolviendo el sistema (con a y b como incógnitas)

$$\begin{cases} a+b &= \frac{5}{2}x_1 - \frac{1}{2}x_2 \\ a-b &= -\frac{1}{2}x_1 + \frac{5}{2}x_2 \end{cases}$$

se obtiene

$$a = x_1 + x_2 \qquad b = \frac{3}{2}x_1 - \frac{3}{2}x_2,$$

por lo tanto,

$$Tx_{\mathfrak{B}_2} = \begin{bmatrix} x_1 + x_2 \\ \frac{3}{2}x_1 - \frac{3}{2}x_2 \end{bmatrix}$$

y

$$[T]_{\mathfrak{B}_1 \mathfrak{B}_1} x_{\mathfrak{B}_1} = \begin{bmatrix} 1 & 1 \\ \frac{3}{2} & -\frac{3}{2} \end{bmatrix} x_{\mathfrak{B}_1} = \begin{bmatrix} 1 & 1 \\ \frac{3}{2} & -\frac{3}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ \frac{3}{2}x_1 - \frac{3}{2}x_2 \end{bmatrix}$$

quedando así comprobado que se cumple la igualdad.

(3) $[T]_{\mathfrak{B}_2\mathfrak{B}_1}$:

Se deja como ejercicio, la respuesta es

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

(4) $[T]_{\mathfrak{B}_2\mathfrak{B}_2}$:

Las columnas de esta matriz son

$$(Tw_1)_{\mathfrak{B}_2} \quad \text{y} \quad (Tw_2)_{\mathfrak{B}_2}$$

$$Tw_1 = T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{y} \quad Tw_2 = T\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

por lo tanto,

$$(Tw_1)_{\mathfrak{B}_2} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad \text{y} \quad (Tw_2)_{\mathfrak{B}_2} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

y

$$[T]_{\mathfrak{B}_2\mathfrak{B}_2} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Ejemplo 10.15. Sean $\mathcal{P}_2$ el conjunto de los polinomios de grado menor a 2 y $D : \mathcal{P}_2 \to \mathcal{P}_2$ el operador lineal definido por

$$D(a_o + a_1x + a_2x^2) = a_1 + 2a_2x \qquad (\text{ la derivada})$$

Sea $\mathfrak{B} = \{1, x, x^2\}$, es claro que $\mathfrak{B}$ es una base de $\mathcal{P}_2$, vamos a hallar la matriz de D tomando esta base en dominio y rango.

El polinomio 1 tiene coordenadas $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, el polinomio x tiene coordenadas $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ y el polinomio x^2 tiene coordenadas $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, además $D1 = 0$, $Dx = 1$ y $Dx^2 = 2x$, por lo tanto, en coordenadas

$$D\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad D\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad D\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix},$$

la matriz es

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

Ejercicios.

(1) Sea T una transformación lineal de $\mathbb{R}^2 \to \mathbb{R}^3$ tal que $T\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}2\\3\\3\end{bmatrix}$ y $T\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}-4\\0\\5\end{bmatrix}$. Hallar el valor de

$$T\begin{bmatrix}2\\4\end{bmatrix} \qquad \text{y} \qquad T\begin{bmatrix}-3\\7\end{bmatrix}.$$

(2) Expresar en coordenadas las transformaciones lineales T: $\mathbb{R}^2 \to \mathbb{R}^2$ que tienen la representación matricial estándar A_T.

(a) $A_T = \begin{bmatrix}4 & 0\\0 & 1\end{bmatrix}$ (b) $A_T = \begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}$

(c) $A_T = \begin{bmatrix}1 & 0\\0 & \frac{1}{4}\end{bmatrix}$ (d) $A_T = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$

(3) Encontrar la representación matricial $[T]$ de la transformación lineal T, a menos que se especifique lo contrario, suponga que $\mathfrak{B}_1$ y $\mathfrak{B}_2$ son las bases canónicas.

(a) $T : \mathbb{R}^2 \to \mathbb{R}$; $T\begin{bmatrix}x\\y\end{bmatrix} = 3x - 2y$

(b) $T : \mathbb{R}^2 \to \mathbb{R}^2$; $T\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x - 2y\\-x + y\end{bmatrix}$

(c) $T : \mathbb{R}^2 \to \mathbb{R}^3$; $T\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x + y\\x - y\\2x + 3y\end{bmatrix}$

(d) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}3x + 2y\\-5x - 4y\end{bmatrix}$, $\mathfrak{B}_1 = \mathfrak{B}_2 = \begin{bmatrix}3\\-2\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}$

10.3. Transformaciones lineales y su relación con sus valores en una base

Supongamos que tenemos dos espacios vectoriales V y W, que la dimensión de V es n y que $\mathfrak{B} = \{v_1, v_2, \dots, v_n\}$ es una base de V.

Si v es un elemento de V, entonces existen escalares únicos $x_1, x_2, \dots, x_n \in \mathbb{R}$ tales que

$$v = x_1v_1 + x_2v_2 + \cdots + x_nv_n.$$

Si $T : V \to W$ es una transformación lineal, entonces

$$Tv = x_1Tv_1 + x_2Tv_2 + \cdots + x_nTv_n = x_1w_1 + x_2w_2 + \cdots + x_nw_n,$$

donde $w_1 = Tv_1, w_2 = Tv_2, \dots, w_n = Tv_n$, lo que implica que la transformación lineal T está determinada por sus valores en los elementos de la base $\mathfrak{B}$.

Por lo tanto, se cumplen los siguientes resultados.

TEOREMA 10.16. *Sean V y W dos espacios vectoriales y sea $T : V \to W$ una transformación lineal, entonces T está determinada por sus valores en una base de V.*

Corolario 10.17. *Sean V y W dos espacios vectoriales y sean $S, T : V \to W$ dos transformaciones lineales. Si S y T coinciden en una base de V, entonces*

$$S = T.$$

Nuevamente, supongamos que tenemos dos espacios vectoriales V y W, que la dimensión de V es n, que $\mathfrak{B} = \{v_1, v_2, \dots, v_n\}$ es una base de V y agreguemos que tenemos un conjunto de n vectores $\{w_1, w_2, \dots, w_n\}$ si definimos la función $T : V \to W$ por

$$T(x_1 v_1 + x_2 v_2 + \cdots + x_n v_n) = x_1 w_1 + x_2 w_2 + \cdots + x_n w_n,$$

para $x_1, x_2, \dots, x_n \in \mathbb{R}$, se tiene que T es lineal y que

$$T v_i = w_i \quad \text{para } k = 1, 2, \dots, n.$$

Por lo tanto, también se cumple el siguiente resultado.

Teorema 10.18. *Sean V un espacio vectorial de dimensión n y W otro espacio vectorial.*

Si $\mathfrak{B} = \{v_1, v_2, \dots, v_n\}$ una base de V y $\{w_1, w_2, \dots, w_n\}$ un subconjunto de W que tiene exactamente n elementos, entonces existe una única transformación lineal $T : V \to W$ tal que

$$T v_i = w_i \quad \textit{para } k = 1, 2, \dots, n.$$

Adicionalmente, esta transformación T está dada por la fórmula

$$T(x_1 v_1 + x_2 v_2 + \cdots + x_n v_n) = x_1 w_1 + x_2 w_2 + \cdots + x_n w_n,$$

para $x_1, x_2, \dots, x_n \in \mathbb{R}$.

Ejemplo 10.19. Hallar la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ que corresponde con la rotación en 90° en sentido antihorario.

Observación: De que la suma y la diferencia de vectores se hace de acuerdo a la ley del paralelogramo, se deduce que una rotación es una transformación lineal.

Por el Teorema 10.18 basta calcular

$$T \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad T \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Como estamos rotando 90° en sentido antihorario, se tiene que

$$T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{y} \quad T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

por lo tanto,

$$T \begin{bmatrix} x \\ y \end{bmatrix} = xT \begin{bmatrix} 1 \\ 0 \end{bmatrix} + yT \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x \begin{bmatrix} 0 \\ 1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$$

Ejercicios.

Encontrar la matriz estándar de la transformación lineal que corresponde con cada una de las siguientes operaciones.

a) Rotación en sentido contrario al de las manecillas del reloj de 120° en torno al origen.
b) Rotación en el sentido de las manecillas del reloj 30° en torno al origen.
c) Proyección sobre la recta $y = 2x$
d) Proyección sobre la recta $y = -x$
e) Reflexión en la recta $y = x$
f) Rotación en sentido contrario al de las manecillas del reloj de 60°, seguida por reflexión en la recta, $y = x$
g) Reflexión en el eje y, seguida por rotación de 30° en sentido de las manecillas del reloj.

10.4. Operaciones básicas con transformaciones lineales

Ya hemos visto que existe una correspondencia entre transformaciones lineales y matrices, que las matrices se pueden sumar, multiplicar por escalar y, en algunos casos especiales, se pueden multiplicar entre ellas. Tomando en cuenta esta correspondencia es natural pensar que podamos realizar operaciones similares con las transformaciones lineales. Tomando en cuenta esta correspondencia, es muy natural pensar que podamos realizar operaciones similares con las transformaciones lineales.

Sean V y W dos espacios vectoriales y $S, T : V \to W$ dos transformaciones lineales, si $v_1, v_2 \in V$ y λ es un escalar se tiene que

$$\begin{aligned}(S+T)(v_1+v_2) =& S(v_1+v_2) + T(v_1+v_2) = Sv_1 + Sv_2 + Tv_1 + Tv_2\\ =&(S+T)v_1 + (S+T)v_2\end{aligned}$$

y

$$\begin{aligned}(\lambda T)(v_1+v_2) =&(\lambda)T(v_1+v_2) = \lambda(Tv_1 + Tv_2) = \lambda Tv_1 + \lambda Tv_2 =\\ =&(\lambda T)v_1 + (\lambda T)v_2\end{aligned}$$

Con $\mathcal{L}(V, W)$ denotaremos al conjunto de las transformaciones lineales de V en W, como es de esperarse, se cumple el siguiente resultado.

TEOREMA 10.20. *Sean V y W dos espacios vectoriales entonces $\mathcal{L}(V, W)$, con las operaciones usuales de suma y producto por un escalar, también es un espacio vectorial.*

Además si $\mathfrak{B}_1$ es una base de V, $\mathfrak{B}_2$ es una base de W y λ es un escalar se cumple que

$$[\lambda S + T]_{\mathfrak{B}_1\mathfrak{B}_2} = \lambda[S]_{\mathfrak{B}_1\mathfrak{B}_2} + [T]_{\mathfrak{B}_1\mathfrak{B}_2}.$$

Por supuesto, el elemento neutro para la suma en $\mathcal{L}(V, W)$ es la transformación lineal nula, que manda todos los vectores al vector 0. Los cálculos que faltarían para establecer el teorema son análogos a los realizados al inicio de esta sección y se dejan como ejercicio.

Como consecuencia, tenemos que existe una correspondencia, que preserva la suma y el producto por un escalar, entre el conjunto de las matrices $m \times n$ y el conjunto de las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$ ($\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$).

Al igual que todas las funciones, se pueden componer transformaciones lineales y el resultado es otra transformación lineal, en detalle, se cumple el siguiente resultado.

TEOREMA 10.21. *Sean V, W y X espacios vectoriales y $T : V \to W$, $S : W \to X$ dos transformaciones lineales, entonces la composición $S \circ T : V \to X$ definida por*

$$(S \circ T)v = S(Tv)$$

también es una transformación lineal.

Además si $\mathfrak{B}_1$ es una base de V, $\mathfrak{B}_2$ es una base de W y $\mathfrak{B}_3$ es una base de X se tiene que

$$[S \circ T]_{\mathfrak{B}_1\mathfrak{B}_3} = [S]_{\mathfrak{B}_2\mathfrak{B}_3}[T]_{\mathfrak{B}_1\mathfrak{B}_2}.$$

DEMOSTRACIÓN. Veamos primero la linealidad.

Sean $v_1, v_2 \in V$ y $\lambda \in \mathbb{R}$, entonces

$$(S \circ T)(v_1 + v_2) = S(T(v_1 + v_2) = S(Tv_1 + Tv_2) = S(Tv_1) + S(Tv_2) = (S \circ T)(v_1) + (S \circ T)(v_2)$$

y

$$(S \circ T)(\lambda v_1) = S(T(\lambda v_1) = S(\lambda T v_1) = \lambda S(Tv_1) = \lambda(S \circ T)(v_1).$$

Para demostrar que

$$[S \circ T]_{\mathfrak{B}_1\mathfrak{B}_3} = [S]_{\mathfrak{B}_2\mathfrak{B}_3}[T]_{\mathfrak{B}_1\mathfrak{B}_2}$$

basta demostrar que

$$[S \circ T]_{\mathfrak{B}_1\mathfrak{B}_3} v_{\mathfrak{B}_1} = [S]_{\mathfrak{B}_2\mathfrak{B}_3}[T]_{\mathfrak{B}_2\mathfrak{B}_3} v_{\mathfrak{B}_1}$$

para todo vector $v \in V$.

Tenemos las siguientes igualdades

$$[S \circ T]_{\mathfrak{B}_1\mathfrak{B}_3} v_{\mathfrak{B}_1} = ((S \circ T)v)_{\mathfrak{B}_3}$$

$$[T]_{\mathfrak{B}_1\mathfrak{B}_2} v_{\mathfrak{B}_1} = (Tv)_{\mathfrak{B}_2}$$

y

$$[S]_{\mathfrak{B}_2\mathfrak{B}_3}(Tv)_{\mathfrak{B}_2} = (S(Tv)))_{\mathfrak{B}_3} = ((S \circ T)v)_{\mathfrak{B}_3},$$

por lo tanto, se cumple la igualdad.

□

Para el caso de transformaciones lineales es usual escribir ST en vez de $S \circ T$.

OBSERVACIÓN 10.22. Para los operadores lineales se cumplen propiedades análogas a las enunciadas para matrices en el Teorema 8.25.

El espacio $\mathcal{L}(V) = \mathcal{L}(V, V)$ (el conjunto de las transformaciones lineales de V en V) es de particular interés. Si la dimensión de V es n, este espacio está en correspondencia con el espacio de las matrices $n \times n$.

El operador identidad es el operador $I : V \to V$ definido por

$$Iv = v \quad \text{para } v \in V.$$

Al operador identidad le corresponde la matriz identidad (en cualquier base).

Se cumplen las siguientes propiedades, que son fácilmente verificables.

Teorema 10.23. *Sean V un espacio vectorial, T, T_1, T_2 operadores lineales de V en V y λ un escalar, entonces*

(1) $IT = TI = T$.
(2) $T(T_1 + T_2) = T_1T + T_2T$.
(3) $(T_1 + T_2)T = T_1T + T_2T$.
(4) $\lambda(T_1T_2) = (\lambda T_1)T_2) = T_1(\lambda T_2)$.
(5) $T(T_1T_2) = (TT_1)T_2$.

Al igual que con las matrices, la composición de operadores no es una operación conmutativa.

Definición 10.24. Sean V un espacio vectorial y $T : V \to V$ un operador lineal, se dice que T es *invertible* si existe un operador lineal $S : V \to V$ tal que

$$ST = TS = I.$$

Al operador S se le llama el *inverso* de T y se le denota por T^{-1}.

Al igual que con las matrices, no todo operador lineal no nulo es invertible.

Por la correspondencia entre operadores y matrices se tiene el siguiente resultado.

Teorema 10.25. *Sea V un espacio vectorial y $T : V \to V$ un operador lineal, entonces T es invertible si y sólo si su matriz con respecto a cualquier base es invertible.*

Fijada una base, la matriz de la inversa de T es la inversa de la matriz de T

Ejemplo 10.26. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ el operador lineal definido por

$$T(x, y, z) = (x + 2z, 2x - y + 3z, 4x + y + 8z).$$

- ¿Es T invertible?
- En caso de que T sea invertible, hallar su inverso.

Existen muchas manera de resolver este problema, veamos algunas.

Método 1: Resolviendo la ecuación $T(x, y, z) = (u, v, w)$.

En esta ecuación la incógnita es el vector (x, y, z) y se debe despejar este vector en función del vector (u, v, w).

El sistema de ecuaciones es

$$\begin{cases} 2x \quad + \quad 2z = u \\ 2x - y + 3z = v \\ 4x + y + 8z = w \end{cases}$$

La matriz aumentada del sistema es

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & u \\ 2 & -1 & 3 & v \\ 4 & 1 & 8 & w \end{array}\right]$$

utilizando el método de Gauss-Jordan (hacerlo a manera de ejercicio) se obtiene

$$\begin{cases} x & = -11u + 2v + 2w \\ y & = -4u + w \\ z & = 6u - v - w \end{cases}$$

De manera que se tiene que el operador es invertible y su inverso es

$$T^{-1}(u, v, w) = (-11u + 2v + 2w, -4u + w, 6u - v - w).$$

Verifiquemos que efectivamente hallamos el operador inverso, tenemos que

$$\begin{aligned} T^{-1}T(x, y, z) =& T^{-1}(x + 2z, 2x - y + 3z, 4x + y + 8z) \\ =& (-11(x + 2z) + 2(2x - y + 3z) + 2(4x + y + 8z), -4(x + 2z) + 4x + y + 8z, \\ & \qquad 6(x + 2z) - (2x - y + 3z) - (4x + y + 8z)) \\ =& (-11x - 22z + 4x - 2y + 6z + 8x + 2y + 16z, -4x - 8z + 4x + y + 8z, \\ & \qquad 6x + 12z - 2x + y - 3z - 4x - y - 8z) \\ =& (x, y, z) \end{aligned}$$

Como ejercicio, verificar que

$$TT^{-1}(u, v, w) = (u, v, w).$$

Método 2: Utilizando la representación matricial de T.

La matriz estándar de T es

$$[T] = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$$

Para saber si esta matriz es invertible, podemos calcular su determinate, tenemos que

$$\det[T] = -8 + 0 + 4 - 0 - 3 - 8 = 1$$

por lo tanto, como la matriz de T es invertible, se tiene que T es invertible.

Ahora podemos utilizar el método de Gauss-Jordan para hallar la inversa de $[T]$.

La inversa de esta matriz fue calculada en el Número 2 del Ejemplo 8.20 (página 127) y es

$$\begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}.$$

por lo tanto,

$$T^{-1}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} -11u + 2v + 2w \\ -4u + w \\ 6u - v - w \end{bmatrix},$$

lo que es equivalente a

$$T^{-1}(u, v, w) = (-11u + 2v + 2w, -4u + w, 6u - v - w).$$

Ejercicios.

(1) Defina las transformaciones lineales $S : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ y $T : \mathbb{R}^2 \to \mathbb{R}^2$ por

$$S : \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b & b \\ 0 & a-b \end{bmatrix} \quad \text{y} \quad T\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 2c+d \\ -d \end{bmatrix}$$

Calcule $(S \circ T)\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $(S \circ T)\begin{bmatrix} x \\ y \end{bmatrix}$. ¿Puede calcular $(T \circ S)\begin{bmatrix} x \\ y \end{bmatrix}$? Si es así, calcúlelo.

(2) Defina las transformaciones lineales $S : \mathcal{P}_1 \to \mathcal{P}_2$ y $T : \mathcal{P}_2 \to \mathcal{P}_1$ por

$$S(a + bx) = a + (a + b)x + 2bx^2$$

y

$$T(a + bx + cx^2) = b + 2cx$$

Calcule $(S \circ T)(3 + 2x - x^2)$ y $(S \circ T)(a + bx + cx^2)$. ¿Puede calcular $(T \circ S)(a + bx)$? Si es así, calcúlelo.

(3) Defina las transformaciones lineales $S : \mathcal{P}_n \to \mathcal{P}_n$ y $T : \mathcal{P}_n \to \mathcal{P}_n$ mediante

$$S(p(x)) = p(x + 1) \qquad \text{y} \qquad T(p(x)) = xp'(x)$$

encuentre $(S \circ T)(p(x))$ y $(T \circ S)(p(x))$.

10.5. Núcleo y rango de una transformación lineal

Definición 10.27. Sean V y W espacios vectoriales y sea $T : V \to W$ una transformación lineal.

El *núcleo* de T es el conjunto

$$\text{Núcleo}(T) = \{v \in V \ : \ Tv = 0\}.$$

Como es natural, la *imagen* de T es el conjunto

$$\text{Imagen}(T) = \{Tv \ : \ v \in V\}.$$

Observación 10.28. Muchos autores se refieren al núcleo de T como *kernel* de T y a la dimensión de la imagen de T le llaman rango de T, para evitar confusiones con la nomenclatura habitual del cálculo se mantendrá la notación y la nomenclatura establecida en la definición anterior.

Es importante notar que el núcleo es un subconjunto de V y la imagen es un subconjunto de W, además tenemos el siguiente resultado fundamental.

Teorema 10.29. *Sean V y W espacios vectoriales de dimensión finita y sea $T : V \to W$ una transformación lineal. Entonces*

(1) *El núcleo de T es un subespacio de V.*
(2) *El rango de T es un subespacio de W.*

Demostración.

(1) Como $T(0) = 0$ se tiene que Núcleo$(T) \neq \emptyset$, sean $v_1, v_2 \in$ Núcleo(T) y sea $\lambda \in \mathbb{R}$, entonces

$$T(v_1 + \lambda v_2) = Tv_1 + \lambda T v_2 = 0 + \lambda 0 = 0,$$

lo que implica que $v_1 + \lambda v_2 \in$ Núcleo(T).

(2) Sean $w_1, w_2 \in$ Imagen(T) y sea $\lambda \in \mathbb{R}$. Como $w_1, w_2 \in$ Imagen(T) existen $v_1, v_2 \in V$ tales que

$$Tv_1 = w_1 \quad \text{y} \quad Tv_2 = w_2,$$

tenemos que

$$w_1 + \lambda w_2 = Tv_1 + \lambda T v_2 = T(v_1 + \lambda v_2),$$

lo que implica que $w_1 + \lambda w_2 \in$ Imagen(T).

$\square$

Teorema 10.30 (Teorema del núcleo y la imagen). *Sean V y W espacios vectoriales de dimensión finita y sea $T : V \to W$ una transformación lineal, entonces*

$$\dim \operatorname{Imagen}(T) + \dim \operatorname{Núcleo}(T) = \dim V.$$

Demostración. Sea n la dimensión de V y k la dimensión de Núcleo(T), como Núcleo(T) es un subespacio de V debe ser $k \leq n$.

Si $k = n$, entonces Núcleo$(T) = V$, por lo que T es la transformación lineal nula, la dimensión del rango es 0 y, por lo tanto, se cumple la igualdad.

Supongamos $k < n$, sea $\{v_1, \ldots, v_k\}$ una base de Núcleo(T), entonces existen $n - k$ vectores $\{v_{k+1}, \ldots, v_n\}$ tales que el conjunto $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ es una base de V.

Vamos a probar que el conjunto $\{Tv_{k+1}, \ldots, Tv_n\}$ es una base de Imagen(T). Veamos primero que genera Imagen(T); sea $w \in$ Imagen(T), entonces existe $v \in V$ tal que $w = Tv$ y existen escalares $\alpha_1, \ldots, \alpha_n$ tales que

$$v = \alpha_1 v_1 + \cdots + \alpha_k v_k + \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n$$

luego

$$\begin{aligned}
Tv =&\alpha_1 Tv_1 + \cdots + \alpha_k Tv_k + \alpha_{k+1} Tv_{k+1} + \cdots + \alpha_n Tv_n \\
=&0 + \alpha_{k+1} Tv_{k+1} + \cdots + \alpha_n Tv_n \\
=&\alpha_{k+1} Tv_{k+1} + \cdots + \alpha_n Tv_n
\end{aligned}$$

lo que implica que el conjunto $\{Tv_{k+1}, \ldots, Tv_n\}$ genera Imagen(T); falta ver que es linealmente independiente.

Supongamos que $\beta_{k+1}, \ldots, \beta_n$ son escalares tales que

$$\beta_{k+1} Tv_{k+1} + \cdots + \beta_n Tv_n = 0,$$

entonces

$$T(\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n) = 0,$$

es decir, $\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n \in \text{Núcleo}(T)$, por lo tanto, existen escalares $\beta_1, \ldots, \beta_k$ tales que

$$\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n = \beta_1 v_1 + \cdots + \beta_k v_k,$$

o lo que es equivalente,

$$\beta_1 v_1 + \cdots + \beta_k v_k - \beta_{k+1} v_{k+1} - \cdots - \beta_n v_n = 0,$$

de donde se concluye que

$$\beta_1 = \beta_2 = \cdots = \beta_n = 0.$$

Luego, la dimensión de Imagen(T) es $n - k$ y, por lo tanto

$$\dim \text{Imagen}(T) + \dim \text{Núcleo}(T) = n - k + k = n = \dim V.$$

$\square$

También tenemos lo siguiente.

PROPOSICIÓN 10.31. *Sean V y W espacios vectoriales de dimensión finita y sea $T : V \to W$ una transformación lineal. Entonces T es inyectiva si y solo si* Núcleo$(T) = \{0\}$.

Dicho de otra manera, T es inyectiva si y solo si $\dim \text{Núcleo}(T) = 0$.

DEMOSTRACIÓN. Ya sabemos que $T0 = 0$.

Supongamos T inyectiva, entonces el único elemento cuya imagen es 0 puede ser 0, por lo tanto Núcleo$(T) = \{0\}$.

Supongamos Núcleo$(T) = \{0\}$ y sean $u, v \in V$ tales que $Tu = Tv$, entonces $T(u - v) = 0$, por lo que $u - v \in \text{Núcleo}(T)$, como el núcleo de T es $\{0\}$, se cumple $u - v = 0$, es decir $u = v$. $\square$

EJEMPLO 10.32. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal definida por

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ 0 \\ y+z \end{bmatrix}.$$

Vamos a hallar su núcleo, su rango, una base del núcleo y una base del rango.

Si $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \text{Núcleo}(T)$, entonces $2x = 0$ y $y + z = 0$, de donde $x = 0$ y $y = -z$; por lo tanto un vector del núcleo de T tiene la forma

$$\begin{bmatrix} 0 \\ y \\ -y \end{bmatrix} = y \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

por lo tanto el núcleo de T tiene dimensión 1 y el conjunto $\left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ es una base del núcleo de T.

Supongamos que el vector $\begin{bmatrix} u \\ v \\ w \end{bmatrix} \in \text{Imagen}(T)$, entonces existen $x, y, z \in \mathbb{R}$ tales que

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2x \\ 0 \\ y+z \end{bmatrix} = 2x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (y+z) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

por lo tanto el rango de T tiene dimensión 2 y el conjunto $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ es una base del rango de T.

Se cumple la igualdad $\dim \text{Imagen}(T) + \dim \text{Núcleo}(T) = 3 = \dim \mathbb{R}^3$, además en este caso podemos visualizar el rango y el núcleo. El rango es el plano xz y el núcleo es la recta $y = -z$ en el plano xz.

EJEMPLO 10.33. Sea $T : \mathbb{R}^2 \to \mathbb{R}^5$ la transformación lineal definida por

$$T(x_1, x_2) = (2x_1 + x_2, 5x_1 - x_2, 10x_2, 0, 4x_1).$$

Hallar la dimensión del núcleo y la dimensión del rango de T.

Si $(x_1, x_2) \in \text{Núcleo}(T)$, entonces $(2x_1 + x_2, 5x_1 - x_2, 10x_2, 0, 4x_1) = (0, 0, 0, 0, 0)$, lo que implica que $x_1 = x_2 = 0$, por lo tanto $\text{Núcleo}(T) = \{(0, 0)\}$.

La dimensión del núcleo es 0, por lo que la dimensión del rango tiene que ser igual a 2.

EJEMPLO 10.34. Para un entero positivo n, sea $\mathcal{P}_n$ el espacio vectorial de los polinomios de grado menor o igual que n. Encontrar la dimensión del núcleo y del rango de la transformación lineal $T : \mathcal{P}_2 \to \mathcal{P}_3$ definida por

$$T(p(x)) = xp(x).$$

En detalle, se tiene

$$T(a + bx + cx^2) = ax + bx^2 + cx^3$$

Si $p(x) = a + bx + cx^2 \in \text{Núcleo}(T)$, entonces $ax + bx^2 + cx^3 = 0$ para todo $x \in \mathbb{R}$, lo que implica que $a = b = c = 0$ y, por lo tanto, $p(x) = 0$.

Luego $\text{Núcleo}(T) = \{0\}$ y $\dim \text{Núcleo}(T) = 0$, como la dimensión de $\mathcal{P}_2$ es 3, se tiene que la dimensión del rango es 3.

En este ejemplo habría sido igualmente sencillo encontrar primero el rango de T, pues es fácil ver que $\{x, x^2, x^3\}$ es una base para el rango de T. Sin embargo, por lo general, uno de los dos (el rango o el núcleo de una transformación lineal) es el más sencillo de calcular.

EJEMPLO 10.35. Sea W el espacio vectorial de todas las matrices 2×2 simétricas y sea $T : W \to \mathcal{P}_2$ la transformación lineal definida por

$$T \begin{bmatrix} a & b \\ b & c \end{bmatrix} = (a - b) + (b - c)x + (c - a)x^2$$

Hallar la dimensión del núcleo y la dimensión del rango de T.

Comencemos por el núcleo. Si $T \begin{bmatrix} a & b \\ b & c \end{bmatrix} = 0$ (el polinomio nulo), entonces se debe cumplir $(a - b) + (b - c)x + (c - a)x^2 = 0$ para todo $x \in \mathbb{R}$, lo que implica que $a = b$, $b = c$ y $c = a$, es decir, $a = b = c$.

Por lo tanto, tenemos que

$$\text{Núcleo}(T) = \left\{ \begin{bmatrix} a & a \\ a & a \end{bmatrix} : a \in \mathbb{R} \right\} = \left\{ a \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} : a \in \mathbb{R} \right\}.$$

lo que implica que el conjunto $\left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ es una base para el núcleo de T y $\dim \text{Núcleo}(T) = 1$.

La dimensión de W es 3, ya que el conjunto $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$ es una base de W.

Finalmente, $\dim \text{Imagen}(T) = \dim W - \dim \text{Núcleo}(T) = 3 - 1 = 2$.

10.5.1. Condiciones equivalentes a la biyectividad. Recordemos que si V es un espacio vectorial de dimensión finita y V_1 es un subespacio de V cuya dimensión es igual a la de V, entonces $V_1 = V$.

Supongamos que V y W son dos espacios vectoriales de igual dimensión finita y sea $T : V \to W$ una transformación lineal. Si n es la dimensión de los espacios V y W, entonces, por el Teorema del núcleo y la imagen (Teorema 10.30) tenemos que

$$\dim \text{Imagen}(T) + \dim \text{Núcleo}(T) = n.$$

Si T es inyectiva, entonces $\dim \text{Núcleo}(T) = 0$, lo que implica que $\dim \text{Imagen}(T) = n = \dim W$; por lo tanto $\text{Imagen}(T) = W$, es decir T es sobreyectiva.

Si T es sobreyectiva, entonces $\text{Imagen}(T) = W$, lo que implica que $\dim \text{Imagen}(T) = \dim W = n$; de aquí se deduce que $\dim \text{Núcleo}(T) = 0$ y, por lo tanto, T es inyectiva.

De lo anterior también se puede concluir que T es biyectiva si y solo si la imagen de una base de V por T es una base de W, es decir T es biyectiva si y solo si dada una base $\mathfrak{B} = \{v_1, \dots, v_n\}$ de V entonces $\{Tv_1, \dots, Tv_n\}$ es una base de W.

Hemos probado el siguiente resultado.

TEOREMA 10.36. *Sean V y W dos espacios vectoriales de igual dimensión finita y sea $T : V \to W$ una transformación lineal. Entonces las siguientes condiciones son equivalentes:*

(1) *T es sobreyectiva.*
(2) *T es inyectiva.*
(3) *T es biyectiva.*
(4) *La imagen de una base de V por T es una base de W.*
(5) $\text{Núcleo}(T) = \{0\}$

10.6. Ejercicios varios

(1) Sea T una transformación lineal de $\mathbb{R}^2 \to \mathbb{R}^3$ tal que $T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}$ y $T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 0 \\ 5 \end{bmatrix}$. Hallar el valor de
$$T \begin{bmatrix} 2 \\ 4 \end{bmatrix} \qquad \text{y} \qquad T \begin{bmatrix} -3 \\ 7 \end{bmatrix}.$$

(2) Encontrar la representación matricial $[T]$ de la transformación lineal T, a menos que se especifique lo contrario, suponga que $\mathfrak{B}_1$ y $\mathfrak{B}_2$ son las bases canónicas. Hallar también su núcleo, su imagen, la dimensión del núcleo y la dimensión de la imagen. Indicar cuáles so inyectiva y cuáles son sobreyectivas.
   (a) $T : \mathcal{P}_2 \to \mathcal{P}_3$, $T(a_0 + a_1x + a_2x^2) = a_1 - a_1x + a_0x^3$
   (b) $T : \mathbb{R}^2 \to \mathbb{R}^2$; $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$
   (c) $T : \mathbb{R}^3 \to \mathbb{R}^2$; $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ -2x + 2y - 2z \end{bmatrix}$
   (d) $T : \mathbb{R}^2 \to \mathbb{R}^2$; $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$
   (e) $T : \mathbb{R}^2 \to \mathbb{R}^3$; $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ 3x - 2y \\ y - x \end{bmatrix}$
   (f) $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + 2z \\ 3x + y + 4z \\ 5x - y + 8z \end{bmatrix}$

(g) $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x+2y+z \\ 2x-4y-2z \\ -3x+6y+3z \end{bmatrix}$

(h) $T : \mathbb{R}^4 \to \mathbb{R}^2$, $T \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x+z \\ 5w-4y \end{bmatrix}$

(i) $T : \mathbb{R}^4 \to \mathbb{R}^4$, $T \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x-y+2z+w \\ -x+2w \\ x-2y+5z+4w \\ 2x-y+z-w \end{bmatrix}$

(j) $T : \mathbb{R}^4 \to \mathbb{R}^2$, $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} aw+bx \\ cy+dz \end{bmatrix}$

(k) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4x-y \\ 3x+2y \end{bmatrix}$, $\mathfrak{B}_1 = \mathfrak{B}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix}$

(l) $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x+y+z \\ y-3z \end{bmatrix}$, $\mathfrak{B}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathfrak{B}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

(m) $T : \mathbb{R}^2 \to \mathbb{R}^4$, $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x+3y \\ -5x-4y \\ -6x-9y \\ x+y \end{bmatrix}$, $\mathfrak{B}_1 = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathfrak{B}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(n) $T : \mathbb{R} \to \mathcal{P}_3$, $T(a) = a + ax + ax^2 + ax^3$

(o) $T : \mathcal{P}_2 \to \mathbb{R}^2$, $T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_0 + a_1 \\ a_1 + a_2 + a_3 \end{bmatrix}$

(p) $T : \mathcal{P}_3 \to \mathcal{P}_1$, $T(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_1 + a_3)x - a_2$

(q) $T : \mathcal{P}_4 \to \mathcal{P}_4$, $T(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4) = a_4x^4 + a_2x^2 + a_0$

(r) $T : \mathcal{P}_3 \to \mathcal{P}_2$, $T(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_0 - a_1 + 2a_2 + 3a_3) + (a_1 + 4a_2 + 3a_3)x + (a_0 + 6a_2 + 5a_3)x^2$

(s) $T : \mathbb{R}^{2\times2} \to \mathbb{R}^{2\times2}$, $T \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a-b+2c+d & -a+2c+2d \\ a-2b+5c+4d & 2a-b+c-d \end{bmatrix}$

(t) $T : \mathcal{P}_4 \to \mathcal{P}_3$, $T(a_o + a_1x + a_2x^2 + a_3x^3 + a_4x^4) = a_3x^3 + a_1x$

(u) $T : \mathbb{R}^{2\times3} \to \mathbb{R}^3$, $T \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} a+e \\ b+f \\ c+d \end{bmatrix}$

(v) $T : \mathcal{P}_2 \to \mathcal{P}_3$, $T(p(x)) = xp(x)$, $\mathfrak{B}_1 = \{1, x, x^2\}$, $\mathfrak{B}_2 = \{1, (1+x), (1+x)^2, (1+x)^3\}$

(w) $T : \mathcal{P}_2 \to \mathcal{P}_3$, $T(p(x)) = xp(x) + p(x)$, $\mathfrak{B}_1 = \{1, x, x^2\}$, $\mathfrak{B}_2 = \{1, (x-1), (x-1)(x-2), (x-1)(x-2)(x-3)\}$

(3) Sea α un número real. Demuestre que $\begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$ es invertible y encuentre su inversa. Interprete geométricamente.

(4) Encontrar la matriz estándar de la transformación lineal que corresponde con cada una de las siguientes operaciones.

(a) Rotación de 45° en sentido de las manecillas del reloj, seguida por proyección sobre el eje y, seguida por rotación de 45° en el sentido de las manecillas del reloj.

(b) Reflexión en la recta $y = x$, seguida por rotación de 30° en sentido contrario al de las manecillas del reloj, seguida por reflexión en la recta $y = -x$

Capítulo 11

Autovalores, autovectores y diagonalización

11.1. Definiciones y resultados básicos

Sea V un espacio vectorial y $T : V \to V$ una transformación lineal.

Se dice que un escalar $\lambda \in \mathbb{R}$ es un *autovalor* de T si existe un vector **no nulo** $v \in V$ tal que

$$Tv = \lambda v.$$

En este caso, se dice que v es una *autovector* asociado al autovalor λ.

Observación 11.1. Es muy importante destacar que en la definición de autovalor, el vector que debe existir y que satisface $Tv = \lambda v$ debe ser no nulo, es decir, diferente del vector 0. Si no pedimos esto, la definición pierde sentido, ya que siempre se cumple $T0 = 0$.

Recordemos que, en el caso de dimensión finita, existe una correspondencia natural entre matrices y transformaciones lineales. Si V tiene dimensión n y $\mathfrak{B}$ es una base de V, por $[T]_{\mathfrak{B}}$ denotaremos la matriz de T con respecto a la base $\mathfrak{B}$ (en la notación ya establecida $[T]_{\mathfrak{B}} = [T]_{\mathfrak{B}\mathfrak{B}}$).

Si λ es un autovalor de T, v es una autovector asociado al autovalor λ y $A = [T]_{\mathfrak{B}}$ se tendría que

$$Av_{\mathfrak{B}} = \lambda v_{\mathfrak{B}}.$$

Por esta razón los conceptos de autovalor y autovector se usan de manera indistinta, tanto para matrices como para transformaciones lineales.

Precisando: Sea n un entero positivo y sea

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots\vdots\vdots & \vdots & \vdots \\ a_{n1} & a_{m2} & \cdots & a_{nn} \end{bmatrix}$$

una matriz $n \times n$.

Se tiene que λ es una autovalor de A si existe un vector no nulo

$$v = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

tal que

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots\vdots\vdots & \vdots & \vdots \\ a_{n1} & a_{m2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Como es natural al vector no nulo v se le conoce como autovector asociado al autovalor λ de la matriz A.

Surge ahora la siguiente pregunta ¿Cómo podemos hallar los autovalores y autovectores de una transformación lineal o de una matriz cuadrada?

Debemos notar varias cosas:

(1) Si por I denotamos al operador identidad y λ es una autovalor de la transformación lineal T, entonces existe un vector no nulo v tal que

$$(T - \lambda I)v = 0,$$

es decir, λ es un autovalor de T si y solo si $\text{Núcleo}(T - \lambda I) \neq \{0\}$.

Además, todo elemento, salvo el vector nulo, de $\text{Núcleo}(T - \lambda I)$ es una autovector de T asociado al autovalor λ. Por esto es común referirse al subespacio $\text{Núcleo}(T - \lambda I)$ como *autoespacio* asociado al autovalor λ.

(2) Por lo anterior y el Teorema 10.36 tenemos que λ es un autovalor de T si y solo si la transformación lineal $T - \lambda I$ no es biyectiva y, por lo tanto no es invertible.

(3) Tomando en cuenta la correspondencia entre matrices y transformaciones lineales tenemos que si A es una matriz $n \times n$ y I_n es la matriz identidad $n \times n$, entonces λ es un autovalor de A si y solo si la matriz $A - \lambda I_n$ no es invertible.

Por lo tanto, λ es un autovalor de A si y solo si $\det(A - \lambda I_n) = 0$.

Este último resultado nos da un método para calcular los autovalores y autovectores de una matriz: Lo primero es hallar las soluciones de la ecuación $\det(A - \lambda I_n) = 0$ y, después resolver las ecuaciones $Av = \lambda v$ para cada autovalor λ.

OBSERVACIÓN 11.2. Si no se presta a confusión es usual utilizar la letra I para la matriz identidad $n \times n$ y para el operador identidad.

A continuación unos ejemplos para ilustrar todo lo anterior.

EJEMPLO 11.3. Sea $P : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal definida por

$$P(x, y) = (x, 0).$$

Hallar los autovalores y los autovectores de P, en caso de que existan.

La matriz estándar de P es

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

y

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 0 \\ 0 & -\lambda \end{bmatrix} = -\lambda(\lambda - 1)$$

por lo tanto, los autovalores de A son $\lambda_1 = 0$ y $\lambda_2 = 1$.

Para hallar los autovectores, debemos resolver las ecuaciones

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

La primera igualdad equivale a la ecuación $x = 0$, la incognita y no aparece en la ecuación; podemos escoger $y = 1$. El vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (y cualquier múltiplo no nulo de este vector) es un autovector asociado al autovalor 0.

La segunda igualdad equivale a las ecuaciones $x = x$, $y = 0$; podemos escoger $x = 1$, $y = 0$. El vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (y cualquier múltiplo no nulo de este vector) es un autovector asociado al autovalor 1.

Podemos verificar directamente que los cálculos son correctos, efectuando las operaciones correspondientes con la matriz:

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

También podemos verificar los cálculos a partir de la definición de la transformación lineal

$$P(0,1) = (0,0) = 0\,(0,1) \qquad \text{y} \qquad P(1,0) = (1,0) = 1\,(1,0).$$

En este caso, que es bastante sencillo, podemos interpretar geométricamente la acción de la transformación lineal P y hasta se pueden encontrar los autovalores y autovectores a partir de esta interpretación.

Analizando la siguiente figura

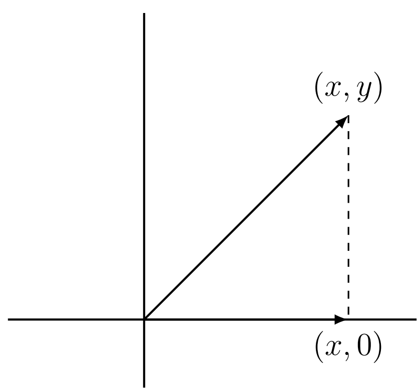


Figura 11.1. Proyección de un vector en el plano sobre el eje x

queda claro que la acción de P sobre un vector es proyectarlo sobre el eje x.

Por lo tanto los vectores verticales (los múltiplos del vector $(0,1)$) tienen por imagen bajo P el vector nulo y los vectores que está alineados con el eje x (los múltiplos del vector $(1,0)$) tienen por imagen bajo P el mismo vector.

Es importante notar que los vectores $(-1,0)$, $(10,0)$ o cualquier vector de la forma $(a,0)$ con $a \neq 0$ es una autovalor asociado al autovalor 1. Igualmente cualquier vector de la forma $(0,b)$ con $b \neq 0$ es una autovalor asociado al autovalor 0.

Ejemplo 11.4. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal definida por

$$R(x,y) = (-y,x).$$

Hallar los autovalores y los autovectores de T, en caso de que existan.

La matriz estándar de R es

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

y

$$\det(B - \lambda I) = \det \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$$

Este polinomio no tiene raíces reales, por lo tanto la transformación R y la matriz asociada B no poseen autovectores ni autovalores.

Este ejemplo también lo podemos interpretar geométricamente, ya esta transformación la habíamos estudiado en el Ejemplo 10.19 y corresponde con la rotación en 90° en sentido antihorario, gráficamente

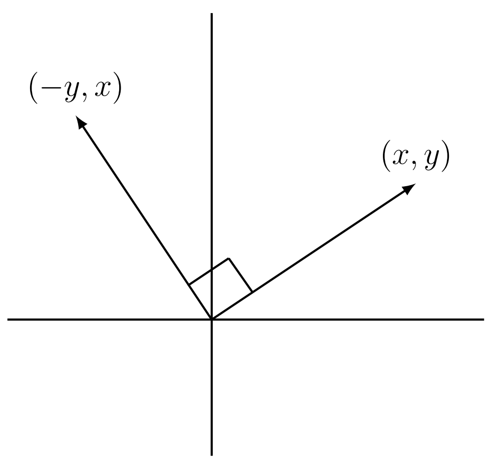


Figura 11.2. Rotación de 90° en el plano

La transformación R no tiene autovectores porque ningún vector no nulo puede ser un múltiplo escalar de si mismo rotado 90°.

Observación 11.5. Los autovectores de una transformación lineal o de una matriz también suelen ser llamados *vectores propios*, *vectores característicos* o *eigenvectores*. Análogamente se usan los términos *valores propios*, *valores característicos* o *eigenvalores* para los autovalores. .

Definición 11.6. Si A es una matriz $n \times n$, al polinomio $\rho(\lambda) = \det(A - \lambda I)$ se le llama *polinomio característico* de A.

Por lo tanto, los autovalores de una matriz A, o de la transformación lineal asociada, son las raíces de su polinomio característico.

Seguimos con otros ejemplos.

EJEMPLO 11.7. Sea A la siguiente matriz 2×2,

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}.$$

Entonces $\det(A - \lambda I) = \det \begin{bmatrix} 4-\lambda & 0 \\ 0 & 4-\lambda \end{bmatrix} = (\lambda - 4)^2 = 0.$

De manera que $\lambda = 4$ es una raíz de multiplicidad 2 del polinomio característico de A.

Como $A = 4I$, se tiene que $Av = 4v$ para todo vector $v \in \mathbb{R}^2$, por lo tanto, los vectores

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

son autovectores de A.

Hemos visto que A tiene un solo autovalor y dos autovectores linealmente independientes asociados a este autovalor.

EJEMPLO 11.8. Sea A la siguiente matriz 2×2

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

Entonces $\det(A - \lambda I) = \det \begin{bmatrix} 4-\lambda & 1 \\ 0 & 4-\lambda \end{bmatrix} = (\lambda - 4)^2 = 0.$

De manera que $\lambda = 4$ es una raíz de multiplicidad 2 del polinomio característico de A.

Si $v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, se tiene que $Av = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4x_1 + x_2 \\ 4x_2 \end{bmatrix}$, por lo tanto

$$Av = 4v \text{ si y solo si } \begin{bmatrix} 4x_1 + x_2 \\ 4x_2 \end{bmatrix} = \begin{bmatrix} 4x_1 \\ 4x_2 \end{bmatrix},$$

para que se cumpla la igualdad debe ser $x_2 = 0$.

Por lo tanto, el vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (o cualquier múltiplo no nulo de este vector) es un autovalor de A asociado al autovalor 4 (y todos los autovectores de A son múltiplos de este vector).

EJEMPLO 11.9. Consideremos la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4x + 2y \\ 3x + 3y \end{bmatrix}$$

La matriz estándar de T es

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

El polinomio característico de A es

$$\begin{aligned}\rho(\lambda) = \det \begin{bmatrix} 4-\lambda & 2 \\ 3 & 3-\lambda \end{bmatrix} = (4-\lambda)(3-\lambda) - 6 = \lambda^2 - 7\lambda + 6 \\ =& (\lambda - 1)(\lambda - 6),\end{aligned}$$

por lo tanto, los autovalores de A son $\lambda_1 = 1$ y $\lambda_2 = 6$.

Consideremos ahora la transformación lineal $U : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$U \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \qquad \text{y} \qquad U \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

notemos que U está siendo definida en una base, por linealidad

$$U \begin{bmatrix} x \\ y \end{bmatrix} = x\, U \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\, U \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2x \\ -3x \end{bmatrix} + \begin{bmatrix} y \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ -3x + y \end{bmatrix}$$

Los dos autovectores de A son linealmente independientes, por lo tanto la transformación U es invertible y se cumple

$$U^{-1} \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \qquad \text{y} \qquad U^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Veamos como actúa la transformación $U^{-1}\, T\, U$ sobre un vector $\begin{bmatrix} x \\ y \end{bmatrix}$.

$$\begin{aligned} U^{-1}\, T\, U \begin{bmatrix} x \\ y \end{bmatrix} =& U^{-1}\, T\, U x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + U^{-1}\, T\, U y \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x\, U^{-1}\, T\, U \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\, U^{-1}\, T\, U \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ =& x\, U^{-1}\, T \begin{bmatrix} 2 \\ -3 \end{bmatrix} + y\, U^{-1}\, T \begin{bmatrix} 1 \\ 1 \end{bmatrix} = x\, U^{-1} \begin{bmatrix} 2 \\ -3 \end{bmatrix} + y\, U^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} \\ =& x\, U^{-1} \begin{bmatrix} 2 \\ -3 \end{bmatrix} + 6\, y\, U^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6\, y \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ =& \begin{bmatrix} x \\ 6y \end{bmatrix} \end{aligned}$$

Es notorio que este operador es particularmente simple, además su matriz estándar es

$$\begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix},$$

y, por construcción

$$[U^{-1}]\, [T]\, [U] = [U^{-1}\, T\, U] = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$$

La matriz estándar de U es

$$[U] = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix}$$

y su inversa (ejercicio) es

$$[U^{-1}] = [U]^{-1} = \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

Si efectuamos las operaciones de multiplicación, obtenemos

$$\begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 2 & 6 \\ -3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix},$$

de donde se obtiene que

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

Resumiendo, hemos obtenido que existe una matriz invertible P y una matriz diagonal D tal que

$$A = P^{-1}DP.$$

Ejercicios.

(1) Para cada una de las siguientes matrices, hallar su polinomio característico, sus autovalores y sus autoespacios.

(a) $\begin{bmatrix} 1 & 3 \\ -2 & 6 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 3 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

(g) $\begin{bmatrix} 4 & 0 & 1 & 0 \\ 0 & 4 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 \end{bmatrix}$ (h) $\begin{bmatrix} 1 & 0 & 2 \\ 3 & -1 & 3 \\ 2 & 0 & 1 \end{bmatrix}$ (i) $\begin{bmatrix} 1 & -1 & -1 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{bmatrix}$

(2) De un ejemplo de una matriz 2×2, no triangular y con autovalores 2 y 5.

11.2. Semejanza de matrices

DEFINICIÓN 11.10. Sean A y B dos matrices cuadradas $n \times n$, se dice que A y B son *semejantes* si existe una matriz $n \times n$ invertible, P, tal que

$$A = P^{-1}BP.$$

OBSERVACIÓN 11.11. En el Ejemplo 11.9 vimos que la matriz

$$\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}$$

es semejante a la matriz

$$\begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}.$$

La semejanza de matrices tiene una cantidad de propiedades que resultarán muy importantes.

PROPOSICIÓN 11.12. *La semejanza de matrices es una relación de equivalencia, es decir, si A, B y C son matrices $n \times n$, se cumple*

(1) *A es semejante a A.*
(2) *Si A es semejante a B entonces B es semejante a A.*
(3) *Si A es semejante a B y B es semejante a C entonces A es semejante a C.*

DEMOSTRACIÓN.

(1) La inversa de la matriz identidad I_n es ella misma, por lo tanto

$$A = I_n^{-1} A I_n.$$

(2) Si $A = P^{-1}BP$, entonces multiplicando por P el lado izquierdo de ambos miembros de la igualdad obtenemos

$$PA = P(P^{-1}BP) = (PP^{-1})BP = I_n BP = BP,$$

ahora multiplicamos el lado derecho por P^{-1} y obtenemos

$$PAP^{-1} = B,$$

de donde

$$B = (P^{-1})^{-1} A P^{-1}.$$

(3) Si $A = P^{-1}BP$ y $B = Q^{-1}CQ$, entonces

$$A = P^{-1}Q^{-1}CQP = (QP)^{-1}C(QP).$$

□

También tenemos el siguiente resultado, que puede ser muy útil al momento de efectuar cálculos.

PROPOSICIÓN 11.13. *Si A y B son matrices $n \times n$ y A es semejante a B, entonces A^k es semejante a B^k para $k = 1, 2, \ldots$ y si $A = P^{-1}BP$ entonces*

$$A^k = P^{-1}B^k P.$$

IDEA DE LA DEMOSTRACIÓN.

$$A^2 = (P^{-1}BP)(P^{-1}BP) = P^{-1}B(PP^{-1})BP = P^{-1}BI_nBP = P^{-1}B^2P,$$

de igual manera

$$A^3 = (P^{-1}BP)(P^{-1}B^2P) = P^{-1}B^3P$$

y así sucesivamente. □

EJEMPLO 11.14. A manera de ilustración, supongamos que tenemos que calcular A^4 donde

$$A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}.$$

En el Ejemplo 11.9 vimos que

$$\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix},$$

por lo tanto,

$$\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}^4 = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}^4 \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1^4 & 0 \\ 0 & 6^4 \end{bmatrix}^4 \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

aunque no vamos a completar el cálculo, es claro que la semejanza puede facilitar operaciones, sobre todo en caso de potencias más altas y matrices más grandes.

También tenemos este otro resultado.

TEOREMA 11.15. *Sean A y B dos matrices $n \times n$ que son semejantes. Entonces*

(1) $\det A = \det B$.
(2) *A es invertible si y solo si B es invertible.*
(3) *A y B tienen el mismo polinomio característico y, por lo tanto, los mismos autovalores.*
(4) *Si A es semejante a B, $A = P^{-1}BP$ y $v \in \mathbb{R}^n$ es un autovector de B con autovalor λ, entonces $P^{-1}v$ es una autovector de A con autovalor λ.*

DEMOSTRACIÓN.

(1) Si $A = P^{-1}BP$, entonces

$$\det A = \det(P^{-1}BP) = (\det(P^{-1})(\det B)(\det P) = \frac{1}{\det(P^{-1}}(\det B)(\det P) = \det B.$$

(2) Esto es consecuencia de que una matriz es invertible si y solo si su determinate es diferente de 0, también se puede probar directamente, si $A = P^{-1}BP$ y B es invertible, entonces

$$AP^{-1}B^{-1}P = P^{-1}B(PP^{-1})B^{-1}P = P^{-1}(BB^{-1})P = P^{-1}P = I,$$

igualmente

$$P^{-1}B^{-1}PA = I.$$

(3) Sean $\rho_A(\lambda)$ y $\rho_B(\lambda)$ los polinomios característicos de A y B, respectivamente, si $A = P^{-1}BP$ entonces

$$\begin{aligned}\rho_A(\lambda) &= \det(A - \lambda I) = \det(P^{-1}BP - \lambda I) = \det(P^{-1}BP - \lambda P^{-1}P)\\ &= \det(P^{-1}(B - \lambda I)P) = (\det P^{-1})(\det(B - \lambda I))(\det P)\\ &= \frac{1}{\det P}(\det(B - \lambda I))(\det P) = \det(B - \lambda I)\\ &= \rho_B(\lambda).\end{aligned}$$

(4) Si $A = P^{-1}BP$ y $v \in \mathbb{R}^n$ es tal que $Bv = \lambda v$, entonces

$$A(P^{-1}v) = (P^{-1}BP)(P^{-1}v) = P^{-1}B(PP^{-1}v) = P^{-1}Bv = P^{-1}(\lambda v) = \lambda(P^{-1}v).$$

$\square$

Ejercicios.

Demostrar que las siguientes matrices A y B no son semejantes.

a) $A = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

b) $A = \begin{bmatrix} 2 & 1 \\ -4 & 6 \end{bmatrix}, B = \begin{bmatrix} 3 & -1 \\ -5 & 7 \end{bmatrix}$

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

d) $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$

11.3. Diagonalización

DEFINICIÓN 11.16. Sea A una matriz $n \times n$, se dice que A es *diagonalizable* si existe una matriz diagonal D tal que A es semejante a D.

Una vez dada esta definición, surgen las siguientes preguntas: ¿Cómo saber si una matriz es diagonalizable o no? Si la matriz es diagonalizable ¿Cómo hallar la matriz D?

A continuación razonaremos sobre estas preguntas, supongamos que D es una matriz $n \times n$ diagonal, que los números que aparecen en la diagonal son $\lambda_1, \lambda_2, \dots, \lambda_k$ ($\lambda_i \neq \lambda_j$ si $i \neq j$) y que cada elemento λ_i aparece r_i veces en la diagonal, es decir A tiene la forma

$$\begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_k \end{bmatrix}$$

donde hemos ordenado los números en la diagonal para que los que son iguales aparezcan uno a continuación del otro.

Como el determinante de una matriz diagonal es el producto de los elementos de la diagonal, tenemos que el polinomio característico de D es

$$\rho_D(\lambda) = (\lambda_1 - \lambda)^{r_1}(\lambda_2 - \lambda)^{r_2} \cdots (\lambda_k - \lambda)^{r_k}.$$

Por lo tanto, cada λ_i es una raíz de multiplicidad r_i del polinomio $\rho_D(\lambda)$.

Notemos que se debe cumplir $1 \leq r_i \leq n$ y $r_1 + r_2 + \cdots + r_k = n$.

Además cada vector de la forma $e_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, donde el 1 está en el lugar j es una autovector asociado al autovalor que ocupa el lugar jj en la matriz D, por lo que $\dim \text{Núcleo}(D - \lambda_k I) = r_k$.

Se ha probado el siguiente resultado.

TEOREMA 11.17. *Si D es una matriz diagonal y $\lambda_1, \lambda_2, \ldots, \lambda_k$ son los elementos de su diagonal, entonces el polinomio característico de D tiene la forma*

$$\rho_D(\lambda) = (\lambda_1 - \lambda)^{r_1}(\lambda_2 - \lambda)^{r_2} \cdots (\lambda_k - \lambda)^{r_k},$$

además $\dim \text{Núcleo}(D - \lambda_k I) = r_k$ *y* $r_1 + r_2 + \cdots + r_k = n$.

EJEMPLO 11.18.

(1) Si $A = \begin{bmatrix} 3 & 0 \\ 0 & \sqrt{2} \end{bmatrix}$ entonces el polinomio característico de A es $\rho(\lambda) = (3-\lambda)(\sqrt{2}-\lambda)$, los autovalores de A son 3 y $\sqrt{2}$, el vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ es una autovector asociado al autovalor 3 y el vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ es un autovector asociado al autovalor $\sqrt{2}$.

(2) Si $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$ entonces el polinomio característico de A es $\rho(\lambda) = (3-\lambda)^2(\sqrt{2}-\lambda)$, los autovalores de A son 3 y $\sqrt{2}$, los vectores $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ son autovectores asociado al autovalor 3 y el vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ es un autovector asociado al autovalor 2.

Notemos que, en este caso $\dim \text{Núcleo}(A - 3I) = 2$ y $\dim \text{Núcleo}(A - \sqrt{2}I) = 1$.

(3) Si $A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, el polinomio característico de A es $\rho(\lambda) = (5-\lambda)(4-\lambda)(\sqrt{5}-\lambda)\lambda$, los autovalores de A son $5, 4, \sqrt{5}$ y 0. El vector $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ es un autovector asociado al autovalor 5, el vector $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ es un autovector asociado al autovalor 4, el vector $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ es un autovector asociado al autovalor $\sqrt{5}$ y el vector $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ es un autovector asociado al autovalor 0.

Cada uno de los autoespacios tiene dimensión 1.

DEFINICIÓN 11.19. Sea A una matriz $n \times n$ y sea λ un autovalor de A.

Se define la *multiplicidad geométrica* del autovalor λ como $\dim \text{Núcleo}(A - \lambda I)$ y se define la *multiplicidad algebraica* del autovalor λ como la multiplicidad de λ como raíz del polinomio característico de A.

OBSERVACIÓN 11.20. Tanto la multiplicidad geométrica como la algebraica son enteros comprendidos entre 1 y n. En las matrices diagonales ambas coinciden y la suma de las multiplicidades es n.

Supongamos que la matriz $n \times n$, A, es diagonalizable y sean P una matriz $n \times n$ invertible y D una matriz diagonal tal que

$$A = P^{-1}\, D\, P.$$

Por el Teorema 11.15 tenemos que A y D tienen los mismos autovalores, el mismo polinomio característico y que, si λ es una autovalor de A, entonces

$$P^{-1}(\text{Núcleo}(D - \lambda I)) = \text{Núcleo}(A - \lambda I),$$

lo que implica que

$$\dim \text{Núcleo}(D - \lambda I) = \dim \text{Núcleo}(A - \lambda I).$$

De esto último se deduce el siguiente resultado.

TEOREMA 11.21. *Sea A una matriz $n \times n$. Las siguientes condiciones son equivalentes:*

(1) *A es diagonalizable.*
(2) *A posee n autovectores linealmente independientes.*
(3) *Existe una base de $\mathbb{R}^n$ cuyo elementos son autovectores de A.*
(4) *El polinomio característico de A se puede descomponer como producto de factores lineales y la multiplicidad algebraica de cada autovalor coincide con su multiplicidad geométrica.*

COROLARIO 11.22. *Si una matriz $n \times n$ posee n autovalores diferentes entonces es diagonalizable.*

OBSERVACIÓN 11.23. Supongamos que sabemos que la matriz $n \times n$, A, es diagonalizable y que ya hallamos una base $\mathfrak{B} = \{v_1, v_2, \ldots, v_n\}$ de autovectores de A tales que $Av_i = \lambda_i v_i$, $i = 1, \ldots, n$

Ya sabemos que A es semejante a una matriz diagonal D, en cuya diagonal aparecen los autovalores de A, contados de acuerdo a su multiplicidad (algebraica o geométrica, porque en este caso son iguales).

Además $De_i = \lambda_i e_i$, donde e_i es el vector con ceros en todos sus lugares, salvo en el lugar i, donde tiene un 1.

¿Cómo hallamos la matriz P tal que

$$D = P^{-1}\, A\, P,$$

o lo que es equivalente

$$P\,D\,P^{-1} = A.$$

Sea P la matriz cuya i-ésima columna es el vector v_i, entonces $Pe_i = v_i$, como la imagen bajo P de una base es una base, se tiene que P es invertible y $P^{-1}v_i = e_i$; por lo tanto,

$$PDP^{-1}v_i = PDe_i = P\lambda_i e_i = \lambda_i Pe_i = \lambda_i v_i,$$

lo que implica que A y PDP coinciden en una base y de aquí se deduce que

$$A = PDP^{-1}.$$

EJEMPLO 11.24. Sea A la matriz 2×2 definida por

$$A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}.$$

Notar que esta matriz corresponde con el Ejemplo 11.9, para una mejor comprensión, detallamos todos los cálculos de nuevo.

Entonces

$$\det(A - \lambda I) = \det \begin{bmatrix} 4-\lambda & 2 \\ 3 & 3-\lambda \end{bmatrix} = (4-\lambda)(3-\lambda) - 6 = \lambda^2 - 7\lambda + 6 = (\lambda-1)(\lambda-6).$$

Por lo tanto, A tiene dos autovalores que son 1 y 6. Como A es una matriz 2×2 con 2 autovalores diferentes, podemos asegurar que A es diagonalizable.

Procedamos a hallar los autovectores, debemos resolver la ecuaciones

$$\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 6 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

que equivalen a

$$4x_1 + 2x_2 = x_1\,, 3x_1 + 3x_2 = x_2 \qquad \text{y} \qquad 4x_1 + 2x_2 = 6x_1\,, 3x_1 + 3x_2 = 6x_2,$$

y se reducen a

$$3x_1 + 2x_2 = 0 \qquad \text{y} \qquad x_1 = x_2.$$

Como autovector asociado a 1 podemos escoger al vector $\begin{bmatrix} 2 \\ -3 \end{bmatrix}$ y como autovector asociado a 6 podemos escoger al vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

De acuerdo a la Observación 11.23, la matriz P es

$$P = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix}.$$

Verifiquemos que $P^{-1}AP$ es una matriz diagonal:

$$P^{-1} = \frac{1}{\det P} \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}.$$

$$P^{-1}AP = \frac{1}{5}\begin{bmatrix}1 & -1\\ 3 & 2\end{bmatrix}\begin{bmatrix}4 & 2\\ 3 & 3\end{bmatrix}\begin{bmatrix}2 & 1\\ -3 & 1\end{bmatrix} = \frac{1}{5}\begin{bmatrix}1 & -1\\ 3 & 2\end{bmatrix}\begin{bmatrix}2 & 6\\ -3 & 6\end{bmatrix} = \frac{1}{5}\begin{bmatrix}5 & 0\\ 0 & 30\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 6\end{bmatrix}$$

que es la matriz cuyas componentes en la diagonal son los autovalores de A.

EJEMPLO 11.25. Sea A la matriz 3×3 definida por

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

Vamos a averiguar si A es diagonalizable y, en caso de que lo sea, vamos a hallar una matriz P tal que $P^{-1}AP$ es diagonal.

El polinomio característico de A es

$$\begin{aligned}
\det(A-\lambda I) =& \det\begin{bmatrix}1-\lambda & -1 & 4\\ 3 & 2-\lambda & -1\\ 2 & 1 & -1-\lambda\end{bmatrix}\\
=&(1-\lambda)(2-\lambda)(-1-\lambda) + 2 + 12 - (8(2-\lambda) - (1-\lambda) - 3(-1-\lambda))\\
=&-\lambda^3 + 2\lambda^2 + \lambda - 2 + 14 - 16 + 8\lambda + 1 - \lambda - 3 - 3\lambda\\
=&-6 + 5\lambda + 2\lambda^2 - \lambda^3\\
=&-((\lambda-3)(\lambda-1)(\lambda+2))
\end{aligned}$$

Por lo tanto, los autovalores de A son $\lambda_1 = 1$, $\lambda_2 = -2$ y $\lambda_3 = 3$; como la matriz tiene tres autovalores diferentes, ya podemos asegurar que es diagonalizable.

Hallemos los autovectores.

Para $\lambda_1 = 1$, debemos resolver

$$\begin{bmatrix}1 & -1 & 4\\ 3 & 2 & -1\\ 2 & 1 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix},$$

que equivale a

$$\begin{cases} x_1 - x_2 + 4x_3 = x_1\\ 3x_1 + 2x_2 - x_3 = x_2\\ 2x_1 + x_2 - x_3 = x_3 \end{cases},$$

simplificando,

$$\begin{cases} -x_2 + 4x_3 = 0\\ 3x_1 + x_2 - x_3 = 0\\ 2x_1 + x_2 - 2x_3 = 0 \end{cases},$$

Con el método de Gauss-Jordan se obtiene que $x_1 + x_3 = 0$, $-x_2 + 4x_3 = 0$; tomando $x_3 = 1$ obtenemos el siguiente autovector asociado al autovalor 1

$$v_1 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$$

Procediendo de manera análoga se obtienen los autovectores

$$v_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \qquad \text{y} \qquad v_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},$$

asociados a los autovalores $\lambda_2 = -2$ y $\lambda_3 = 3$, respectivamente.

Sea

$$P = \begin{bmatrix} -1 & 1 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{bmatrix},$$

utilizando el método de Gauss-Jordan se obtiene que

$$P^{-1} = \begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & -\frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

Verificamos

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & -\frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & -\frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -1 & -2 & 3 \\ 4 & 2 & 6 \\ 1 & 2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \end{aligned}$$

OBSERVACIÓN 11.26. Como existe un número infinito de maneras en las cuales se puede elegir un vector característico, existe un número infinito de formas para seleccionar una matriz de diagonalización P. El único consejo es elegir los vectores característicos y la matriz P que permitan un manejo aritmético lo más sencillo posible. En términos generales, esto quiere decir que debe insertarse el mayor número de ceros y unos posible.

EJEMPLO 11.27. Sea A la matriz 3×3 definida por

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

El polinomio característico de A es

$$\begin{aligned}\rho(\lambda) =& \det(A-\lambda I) = \det\begin{bmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{bmatrix} \\ =& -\lambda^3 + 6\lambda^2 + 15\lambda + 8 \\ =& (\lambda+1)^2(8-\lambda)\end{aligned}$$

Los autovalores son $\lambda_1 = -1$ y $\lambda_2 = 8$

A diferencia del ejemplo anterior, todavía no sabemos si la matriz es diagonalizable ya que tenemos un autovalor con multiplicidad algebraica 2.

Debemos hallar los autovectores.

Para $\lambda = -1$, debemos resolver

$$\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = - \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

es decir

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 = -x_1 \\ 2x_1 + 2x_3 = -x_2 \\ 4x_1 + 2x_2 + 3x_3 = -x_3. \end{cases},$$

que equivale a

$$\begin{cases} 4x_1 + 2x_2 + 4x_3 = 0 \\ 2x_1 + x_2 + 2x_3 = 0 \\ 4x_1 + 2x_2 + 4x_3 = 0. \end{cases}.$$

Todas las ecuaciones obtenidas son equivalentes y equivalen a

$$2x_1 + x_2 + 2x_3 = 0.$$

Podemos tomar como autovectores

$$v_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$$

que son linealmente independientes; por lo tanto, el autovalor 2 tiene multiplicidad geométrica 2 y ya podemos garantizar que la matriz es diagonalizable.

Para el autovalor 8 se obtiene el autovector

$$v_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}.$$

Sea $P = \begin{bmatrix} 1 & 0 & 2 \\ -2 & -2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$, por el método de Gauss Jordan, se obtiene

$$P^{-1} = \frac{1}{9}\begin{bmatrix} 5 & -2 & -4 \\ -4 & -2 & 5 \\ 2 & 1 & 2 \end{bmatrix}.$$

Verificamos

$$\begin{aligned} P^{-1}AP =& \frac{1}{9}\begin{bmatrix} 5 & -2 & -4 \\ -4 & -2 & 5 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 2 \\ -2 & -2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \\ =& \frac{1}{9}\begin{bmatrix} 5 & -2 & -4 \\ -4 & -2 & 5 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 & 16 \\ 2 & 2 & 8 \\ 0 & -1 & 16 \end{bmatrix} \\ =& \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{bmatrix} \end{aligned}$$

EJEMPLO 11.28. Sea A la matriz 3×3 definida por

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

El polinomio característico de A es

$$\begin{aligned} \rho(\lambda) =& \det(A - \lambda I) = \det\begin{bmatrix} 4-\lambda & 1 & 0 \\ 0 & 4-\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix} \\ =& -\lambda(4-\lambda)^2 \end{aligned}$$

Los autovalores son $\lambda_1 = 4$ y $\lambda_2 = 0$; $\lambda_1 = 4$ cin multiplicidad algebraica igual a 2.

Veamos cuáles son los autovectores asociados a $\lambda_1 = 4$

$$\begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

La ecuación correspondiente es

$$\begin{cases} 4x_1 + x_2 = 4x_1 \\ 4x_2 = 4x_2 \\ x_3 = 0 \end{cases}$$

que equivalen a $x_2 = x_3 = 0$, por lo que una base del autoespacio correspondiente al autovalor 4 es el vector

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

lo que implica que la multiplicidad geométrica de este autovalor es 1; por lo tanto, esta matriz no es diagonalizable.

Ejercicios

(1) A continuación se da una diagonalización de la matriz A en la forma $P^{-1}AP = D$. Diga cuáles son los autovalores de A y las bases para los autoespacios correspondientes.

(a) $\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \\ 1 & -1 & 1 \\ 2 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{4} & -\frac{3}{4} & \frac{1}{4} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{3}{8} \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 \\ 2 & 0 & 2 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 1 \\ 2 & 1 & 0 \\ 3 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

(2) Determinar si cada una de las siguientes matrices A son diagonalizables; si es diagonalizable, si es diagonalizable encuentre una matriz invertible P y una matriz diagonal D tales que $P^{-1}AP = D$ y calcules A^3 y A^4.

(a) $A = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$ (b) $A = \begin{bmatrix} -3 & 4 \\ -1 & 1 \end{bmatrix}$ (c) $A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

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

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

(3) Justificar la siguiente afirmación: Si A una matriz invertible diagonalizable, entonces A^{-1} también es diagonalizable.

11.4. Diagonalización ortogonal de matrices simétricas

DEFINICIÓN 11.29. Sea n un entero positivo y A una matriz $n \times n$. Se dice que la matriz A es *ortogonal* si sus columnas constituyen una base ortonormal de $\mathbb{R}^n$.

EJEMPLO 11.30. La matriz

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

es ortogonal.

Se tiene lo siguiente

$$A^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

y

$$A^T A = A A^T = I.$$

Más generalmente, se cumple el siguiente resultado.

TEOREMA 11.31. *Sea Q una matriz $n \times n$, entonces Q es ortogonal si y solo si $Q^T Q = QQ^T = I$.*

DEMOSTRACIÓN. Sea q_i la i-ésima columna de Q (y, en consecuencia, el i-ésimo renglón o fila de Q^T).

Como la entrada (i, j) de $Q^T Q$ es el producto escalar del i-ésimo renglón de Q^T y la j-ésima columna de Q, se tiene que

$$(Q^T Q)_{ij} = \langle q_i, q_j \rangle,$$

por ser Q ortogonal

$$(Q^T Q)_{ij} = \begin{cases} 0 & \text{si i } \neq j \\ 1 & \text{si i } = j \end{cases},$$

lo que implica que $Q^T Q = I$.

De igual manera se prueba $QQ^T = I$. □

COROLARIO 11.32. *Una matriz cuadrada Q es ortogonal si y solo si $Q^{-1} = Q^T$.*

DEFINICIÓN 11.33. Sea A una matriz cuadrada. Se dice que A es *diagonalizable ortogonalmente* si existen una matriz ortogonal Q y una matriz diagonal D tales que

$$D = Q^T A\, Q.$$

PROPOSICIÓN 11.34. *Si A es diagonalizable ortogonalmente, entonces A es simétrica.*

DEMOSTRACIÓN. Supongamos se cumple que

$$D = Q^T A\, Q,$$

donde D es diagonal y Q es ortogonal, entonces

$$A = Q\, D\, Q^T,$$

por lo tanto

$$A^T = (Q\, D\, Q^T)^T = (Q^T)^T D^T Q^T = Q\, D\, Q^T = A.$$

□

Observación 11.35. El recíproco del resultado anterior es cierto, toda matriz simétrica es diagonalizable ortogonalmente. El resultado se aceptará sin demostración, por la complejidad de su prueba.

El procedimiento para encontrar la matriz ortogonal Q que diagonaliza una matriz simétrica A sigue estos tres pasos:

i) Encontrar una base para cada espacio característico de A.
ii) Encontrar una base ortonormal para cada espacio característico de A usando el proceso de Gram-Schmidt o algún otro.
iii) Escribir Q como la matriz cuyas columnas son los vectores característicos ortonormales obtenidos en el inciso ii).

A continuación unos ejemplos referentes a diagonalización ortogonal.

Ejemplo 11.36. Consideremos la matriz

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

la matriz A es simétrica y su ecuación característica es

$$\det A - \lambda I) = \det \begin{bmatrix} 1-\lambda & -2 \\ -2 & 3-\lambda \end{bmatrix} = \lambda^2 - 4\lambda - 1,$$

sus raíces son

$$\frac{(4 \pm \sqrt{20})}{2} = \frac{(4 \pm 2\sqrt{5})}{2} = 2 \pm \sqrt{5}.$$

Autovectores para $\lambda_1 = 2 - \sqrt{5}$:

Debemos resolver la ecuación

$$\begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (2 - \sqrt{5}) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

que equivale a

$$\begin{cases} x_1 - 2x_2 = (2 - \sqrt{5})x_1 \\ -2x_1 + 3x_2 = (2 - \sqrt{5})x_2 \end{cases}$$

$$\begin{cases} (\sqrt{5} - 1)x_1 - 2x_2 = 0 \\ -2x_1 + (1 + \sqrt{5})x_2 = 0. \end{cases}$$

Si multiplicamos la primera ecuación por $1 + \sqrt{5}$ obtenemos

$$4x_1 - 2(1 + \sqrt{5})x_2 = 0, \qquad 2x_1 - (1 + \sqrt{5})x_2 = 0,$$

que equivale a la segunda ecuación, tomando $x_1 = 1$, se obtiene

$$x_2 = \frac{2}{1+\sqrt{5}} = -1+\sqrt{5},$$

por lo tanto,

$$v_1 = \begin{bmatrix} 2 \\ -1+\sqrt{5} \end{bmatrix}$$

es un autovector asociado al autovalor $2-\sqrt{5}$.

$$|v_1| = \sqrt{2^2 + (-1+\sqrt{5})^2} = \sqrt{10-2\sqrt{5}},$$

lo que implica que

$$u_1 = \frac{1}{\sqrt{10-2\sqrt{5}}} \begin{bmatrix} 2 \\ -1+\sqrt{5} \end{bmatrix}$$

es un autovector de norma 1 asociado al autovalor $2-\sqrt{5}$.

De manera análoga se obtiene que

$$u_2 = \frac{1}{\sqrt{10-2\sqrt{5}}} \begin{bmatrix} 1-\sqrt{5} \\ 2 \end{bmatrix}$$

es un autovector de norma 1 asociado al autovalor $2+\sqrt{5}$.

(Notemos que $\langle v_1, v_2 \rangle = 0$)

Finalmente

$$Q = \frac{1}{\sqrt{10-2\sqrt{5}}} \begin{bmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2 \end{bmatrix}$$

$$Q^T = \frac{1}{\sqrt{10-2\sqrt{5}}} \begin{bmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{bmatrix},$$

verificamos

$$\begin{aligned} Q^T A Q &= \frac{1}{10-2\sqrt{5}} \begin{bmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2 \end{bmatrix} \\ &= \frac{1}{10-2\sqrt{5}} \begin{bmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{bmatrix} \begin{bmatrix} 4-2\sqrt{5} & -3-\sqrt{5} \\ -7+3\sqrt{5} & 4+2\sqrt{5} \end{bmatrix} \\ &= \frac{1}{10-2\sqrt{5}} \begin{bmatrix} 30-14\sqrt{5} & 0 \\ 0 & 10+6\sqrt{5} \end{bmatrix} = \begin{bmatrix} 2-\sqrt{5} & 0 \\ 0 & 2+\sqrt{5}. \end{bmatrix} \end{aligned}$$

EJEMPLO 11.37. Consideremos la matriz

$$A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix},$$

la matriz A es simétrica y

$$\det(A-\lambda I)=\det\begin{bmatrix}5-\lambda & 4 & 2\\ 4 & 5-\lambda & 2\\ 2 & 2 & 2-\lambda\end{bmatrix}=-(\lambda-1)^2(\lambda-10).$$

Los vectores

$$\begin{bmatrix}-1\\1\\0\end{bmatrix} \qquad \text{y} \qquad v_2=\begin{bmatrix}-1\\0\\2\end{bmatrix}$$

son autovectores correspondientes al autovalor $\lambda=1$ y el vector

$$v_3=\begin{bmatrix}2\\2\\1\end{bmatrix}$$

es un autovector correspondientes al autovalor $\lambda=10$

Para encontrar Q se aplica el proceso de Gram-Schmidt a $\{v_1,v_2\}$y se obtienen los vectores

$$u_1=\begin{bmatrix}\frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\\ 0\end{bmatrix} \qquad \text{y} \qquad u_2=\begin{bmatrix}-\frac{1}{3\sqrt{2}}\\ -\frac{1}{3\sqrt{2}}\\ \frac{4}{3\sqrt{2}}\end{bmatrix}.$$

Normalizando v_3, se obtiene

$$u_3=\frac{\mathbf{v}_3}{|\mathbf{v}_3|}=\frac{1}{3}\mathbf{v}_3=\begin{bmatrix}\frac{2}{3}\\ \frac{2}{3}\\ \frac{1}{3}\end{bmatrix}.$$

$$v_2'=v_2-(v_2\cdot u_1)u_1=\begin{bmatrix}-1\\0\\2\end{bmatrix}-\frac{1}{\sqrt{2}}\begin{bmatrix}\frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\\ 0\end{bmatrix}=\begin{bmatrix}-1\\0\\2\end{bmatrix}-\frac{1}{\sqrt{2}}\begin{bmatrix}\frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\\ 0\end{bmatrix}=\begin{bmatrix}-1\\0\\2\end{bmatrix}.$$

Por lo tanto,

$$Q=\begin{bmatrix}-\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3}\\ -\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3}\\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3}\end{bmatrix}.$$

Verificamos

$$\begin{aligned}Q^TAQ&=\begin{bmatrix}-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ -\frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}}\\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}5&4&2\\4&5&2\\2&2&2\end{bmatrix}\begin{bmatrix}-\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3}\\ -\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3}\\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3}\end{bmatrix}\\ &=\begin{bmatrix}-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ -\frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}}\\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}-\frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{20}{3}\\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{20}{3}\\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3}\end{bmatrix}\\ &=\begin{bmatrix}1&0&0\\0&1&0\\0&0&10\end{bmatrix}\end{aligned}$$

Capítulo 12

Aproximación por mínimos cuadrados

En este capítulo se ofrece una pequeña introducción al tema de aproximación de soluciones aproximadas, utilizando técnicas de álgebra lineal. Detalles sobre las demostraciones de los resultados presentados y ampliación del tema se pueden encontrar en el libro de David Poole [8].

En las ciencias es usual basarse en datos experimentales, para establecer una relación matemática entre variables que podemos observar.

Por ejemplo, podemos medir la altura de un árbol a diferentes edades y tratar de deducir una función que exprese la altura h del árbol de su edad. También podemos tratar de encontrar una regla que establezca el tamaño de una población en función del tiempo, en base a la observación de su crecimiento.

Las relaciones entre variables también son de interés en los negocios; por ejemplo, una compañía que produce algunos objetos puede estar interesada en conocer la relación entre sus costos totales c y el número n de objetos producidos.

En cada uno de estos ejemplos, los datos vienen en forma de dos mediciones: una para la variable independiente y uno para la (supuesta) variable dependiente.

Por lo tanto, se tiene un conjunto de puntos de datos de la forma (x_i, y_i) y se busca una función que aproxime de la manera la relación entre la variable independiente x y la variable dependiente y.

Supongamos que los puntos de datos (1, 2), (2, 2) y (3, 4) surgieron de mediciones tomadas durante algún experimento y supongamos también que existen razones para creer que los valores x e y se relacionan mediante una función lineal; es decir, se espera que los puntos se encuentren en alguna recta con ecuación $y = a + bx$. Si las mediciones son precisas, los tres puntos deberían satisfacer esta ecuación y se tendría

$$2 = a + b \cdot 1 \quad 2 = a + b \cdot 2 \quad 4 = a + b \cdot 3.$$

Por lo que tenemos el siguiente sistema de ecuaciones lineales en dos variables a yb

$$\begin{cases} a + b = 2 \\ a + 2b = 2 \\ a + 3b = 4 \end{cases}$$

que equivale a

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

Este sistema es inconsistente, ya que los tres puntos no se encuentran sobre una misma recta. Lo que queda por hacer es buscar una recta que pase “tan cerca como sea posible” por los tres puntos.

Dada cualquier recta, se medirá la distancia vertical desde cada punto de datos a la recta (lo que representa los errores en la dirección vertical) y luego se tratará de elegir la recta que minimice el error total.

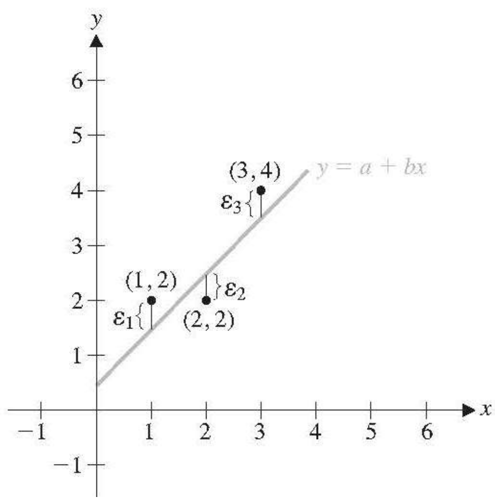


FIGURA 12.1. Recta que minimiza el error de aproximación

Si los errores se denotan mediante $\varepsilon_1, \varepsilon_2, y\varepsilon_3$, entonces puede formar el *vector de error*

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

La pregunta que debemos hacernos es ¿Cómo encontrar la recta que minimiza $\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2$

Se quiere que ε sea tan pequeño como sea posible, de modo que $\|\varepsilon\|$ debe estar tan cerca de cero como sea posible. Existen muchas maneras (normas) para medir el tamaño del vector ε, la elección de la norma euclideana es la que conduce a fórmulas más fáciles de trabajar, tal como se verá a continuación.

Debemos minimizar

$$\|\varepsilon\| = \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2} \quad o, \text{de manera equivalente} \quad \|\varepsilon\|^2 = \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2.$$

Como es natural, de aquí proviene el término “mínimos cuadrados”: es necesario encontrar la menor suma de cuadrados, en el sentido de la ecuación anterior. Al número $\|\varepsilon\|$ se llama *error de mínimos cuadrados de la aproximación*.

Formalizando las ideas ya presentadas.

Definición 12.1. Si A es una matriz $m \times n$ y b un vector en $\mathbb{R}^m$, una *solución de mínimos cuadrados* de la ecuación

$$Ax = b$$

es un vector x_o en $\mathbb{R}^n$ tal que

$$|b - Ax_o| \leq |b - Ax|$$

para todo x en $\mathbb{R}^n$.

Se tiene el siguiente resultado.

Teorema 12.2 (Teorema de mínimos cuadrados). *Sea A una matriz $m \times n$ y sea b un vector en R^m. Entonces la ecuación*

$$Ax = b$$

siempre tiene al menos una solución de mínimos cuadrados x_o. Más aún,

a) El vectorx_o es una solución en mínimos cuadrados de $Ax = b$ si y sólo si x_o es una solución de la ecuación

$$A^T Ax = A^T b,$$

b) La matriz A tiene columnas linealmente independientes si y sólo si $A^T A$ es invertible. En este caso, la solución de mínimos cuadrados de

$$Ax = b$$

es única y está dada por

$$x_o = (A^T A)^{-1} A^T b$$

Ejemplo 12.3. Volviendo al ejemplo inicial,encuentrar la aproximación lineal por mínimos cuadrados para los puntos de datos $(1, 2)$, $(2, 2)$ y $(3, 4)$.

Ya se vio que el sistema correspondiente es $Ax = b$ donde

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

donde $y = a + bx$ es la recta que se busca. Puesto que las columnas de A son linealmente independientes, habrá una solución única de mínimos cuadrados.

Se tiene que

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \quad y \quad A^T b = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 18 \end{bmatrix}$$

La ecuación $A^T A x_o = A^T b$ es

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} 8 \\ 18 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{2}{3} \\ 1 \end{bmatrix},$$

utilizando el procedimiento de Gauss-Jordan se obtiene

$$x_o = \begin{bmatrix} \frac{2}{3} \\ 1 \end{bmatrix}.$$

Por lo tanto, la recta de aproximación lineal por mínimos cuadrados es

$$y = \frac{2}{3} + x.$$

A manera de ejercicio, graficar la recta y representar los puntos con GeoGebra.

Bibliografía

[1] Apostol, T. *Calculus Volumen I. Segunda edición* Reverté (1999), (España).

[2] Baldor, A. *Álgebra, Tercera edición* Grupo Editorial Patria (2016), México.

[3] Baldor, A. *Geometría y trigonometría. Tercera edición* Grupo Editorial Patria (2011), México.

[4] Del Valle, J. *Álgebra lineal para estudiantes de Ingeniería y Ciencias* Mc Graw Hill (2011), (México).

[5] Grossman, S. *Álgebra lineal. Séptima edición.* Mc Graw Hill (2012), (México).

[6] Lay, D. *Álgebra lineal y sus aplicaciones. Cuarta edición.* Pearson (2012), (México).

[7] Olver, P. and Shakiban, C. *Applied linear algebra, Second Edition* Springer (2018), (Suiza).

[8] Poole, D. *Álgebra lineal, una introducción moderna. Cuarta edición.* Cengage (2017), (México).

[9] Zill, D. *Precálculo con avances de cálculo, quinta edición* Mc Graw Hill, (2008), China.

[10] Zill, D. *Álgebra, trigonometría y geometría analítica* Mc Graw Hill, (2012), México.

Índice de figuras

Índice alfabético