Héctor Ramón Tarifa - Patricio Omar Condorí
José Luis Medina y Víctor Eduardo Mérida

ÁLGEBRA LINEAL PARA INGENIERÍAS
Álgebra y Geometría Analítica



Tarifa - Condorí - Medina - Mérida

ÁLGEBRA LINEAL PARA INGENIERÍAS

UNIVERSIDAD NACIONAL DE JUJUY
Jujuy - Argentina
2025

Libros de cátedra
ediunju

ÁLGEBRA LINEAL PARA INGENIERÍAS

Álgebra y Geometría Analítica

ÁLGEBRA LINEAL PARA INGENIERÍAS

Álgebra y Geometría Analítica

Héctor Ramón Tarifa

Patricio Omar Condorí

José Luis Medina

Víctor Eduardo Mérida

Universidad Nacional de Jujuy

2025

Prohibida la reproducción total o parcial del material contenido en esta publicación por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, sin permiso expreso del Editor.

Álgebra lineal para Ingenierías : álgebra y geometría analítica / Héctor Ramón Tarifa ...
[et al.]. - 1a ed. - San Salvador de Jujuy : Editorial de la Universidad Nacional de Jujuy - EDIUNJU, 2025.
Libro digital, PDF

Archivo Digital: descarga y online
ISBN 978-950-721-634-3

1. Álgebra. I. Tarifa, Héctor Ramón
CDD 512



Diseño de Tapa e Interior: Edgardo Gutiérrez
Corrección y cuidado de edición: Silvina Campo

© 2025 Héctor Ramón Tarifa, Patricio Omar Condorí, José Luis Medina y Víctor Eduardo Mérida
© 2025 Editorial de la Universidad Nacional de Jujuy
Avda. Bolivia 1685 - CP 4600
San Salvador de Jujuy - Pcia. de Jujuy - Argentina
Tel. (0388) 4221511- e-mail: editorial@unju.edu.ar

2025 1ra Edición

Queda hecho el depósito que previene la Ley 11.723

ÍNDICE

INTRODUCCIÓN

CAPÍTULO 1……19

VECTORES EN R^n 19

1. Introducción 19

1.1. Conceptos básicos 21

2. Operaciones con vectores 25

2.1. Suma de vectores 25

2.2. Producto por un escalar 27

2.3. Resta de vectores 35

2.4. Vector determinado por dos puntos de $\boldsymbol{Rn}$ 37

3. Vectores paralelos 38

4. Otras operaciones con vectores 41

4.1. Producto escalar 41

4.2. Versores fundamentales 61

4.3. Producto vectorial 62

4.4. Producto Mixto 76

CAPÍTULO 2 85

NÚMEROS COMPLEJOS 85

1. Introducción 85

1.1. Conceptos básicos..... 87

2. Representación gráfica de números complejos..... 88

3. Otras formas de expresar un número complejo 91

3.1. Números complejos expresados en forma binómica..... 91

3.2. Números complejos expresados en forma polar o trigonométria……94

3.3. Equivalencia entre los sistemas de representaciones 96

4. Igualdad de números complejos..... 97

5. Operaciones con números complejos..... 98

5.1. Suma de números complejos..... 98

5.2. Producto de un número complejo por un escalar..... 99

5.3. Diferencia de números complejos 108

5.4. Producto de números complejos 109

5.5. Conjugación de números complejos 115

5.6. Cociente de números complejos 122

5.7. Potenciación de números complejos 125

5.8. Radicación de números complejos 130

6. Forma exponencial de los números complejos 133

CAPÍTULO 3 141

POLINOMIOS 141

1. Introducción 141

1.1. Conceptos básicos 142

2. Operaciones con polinomios 146

2.1. Suma de polinomios 146

2.2. Producto de un polinomio por un escalar 147

2.3. Producto de polinomios 151

2.4. Raíces de un polinomio 155

3. Divisibilidad de polinomios 163

3.1. El algoritmo de división para polinomios 166

3.2. Teorema del resto 170

4. Raíces Racionales 172

4.1. Criterio de Gauss 172

4.2. Regla de Descartes de los signos 178

5. Multiplicidad de las raíces 179

6. Teorema Fundamental del Álgebra 183

7. Raíces complejas de polinomios reales 186

8. Algoritmo de Euclides 189

8.1. Admisión de raíces múltiples..... 195

9. Relaciones entre coeficientes y raíces..... 197

CAPÍTULO 4..... 201

MATRICES..... 201

1. Introducción..... 201

1.1. Conceptos básicos..... 204

2. Operaciones con matrices..... 209

2.1. Suma de matrices..... 209

2.2. Producto por una matriz por un escalar..... 209

3. Otras operaciones con matrices..... 215

3.1. Producto de matrices..... 216

3.2. Potencia de matrices..... 223

4. Trasposición de matrices..... 224

5. Matrices simétricas y antisimétricas..... 227

6. Matrices triangulares..... 229

7. Matrices diagonales..... 231

8. Matrices idempotentes e involutivas..... 233

9. Inversión de matrices..... 235

10. Matrices ortogonales..... 240

11. Transformaciones elementales: matrices equivalentes, escalonadas y reducidas 241

11.1. Obtención de la forma escalonada reducida por filas: algoritmo de eliminación gaussiana 245

11.2. Obtención de la forma escalonada reducida por filas: algoritmo de Gauss–Jordan 250

12. Condición de existencia y cálculo de la matriz inversa 253

13. Submatrices 258

14. Ecuaciones matriciales 258

CAPÍTULO 5 266

DETERMINANTES 267

1. Introducción 267

1.1. Conceptos básicos 271

2. Matriz de cofactores y Matriz adjunta 278

3. Desarrollo de un determinante por los elementos de una línea 281

4. Propiedades de los determinantes 286

5. Relación entre los determinantes de matrices equivalentes 299

6. Cálculo de determinantes 300

6.1. Método por triangulación o eliminación Gaussiana 301

6.2. Método de Chío 302

7. Cálculo de la inversa a partir de la matriz adjunta 305

8. Determinante, inversa, rango y equivalencia a la identidad..... 309

9. Determinantes y rango 311

9.1. Método del orlado para calcular el rango de una matriz 313

CAPÍTULO 6 316

SISTEMAS DE ECUACIONES LINEALES 317

1. Introducción 317

1.1. Conceptos básicos..... 320

2. Distintas formas de expresar un sistema de ecuaciones lineales... 332

3. Análisis de sistemas de ecuaciones lineales 335

4. Resolución de sistemas de ecuaciones lineales 339

4.1. Método de eliminación de Gauss 340

4.2. Método de Gauss–Jordan 348

4.3. Método de la matriz inversa 352

4.4. Regla de Cramer 356

5. Sistemas de ecuaciones lineales homogéneos 360

5.1. Relación entre los sistemas no homogéneos y su sistema homogéneo asociado..... 363

6. Sistema de ecuaciones lineales con igual número de ecuaciones e incógnitas 372

7. Sistema de ecuaciones lineales con parámetros 375

8. Relación entre las soluciones de un sistema y sus matrices 379

8.1. Sistema con solución 380

8.2. Sistema sin solución 383

8.3. Sistema con menos ecuaciones que incógnitas 384

CAPÍTULO 7 386

ESPACIOS VECTORIALES 387

1. Introducción 387

1.1. Conceptos básicos 390

2. Espacio Vectorial 399

3. Subespacio Vectorial 405

3.1. Criterio de subespacio 409

4. Combinación lineal de vectores 411

4.1. Subespacio generado 418

5. Dependencia lineal 422

5.1. Propiedades de la dependencia lineal 431

6. Sistema de generadores 442

7. Base de un espacio vectorial 448

7.1. Coordenadas respecto a una base 455

7.2. Dimensión de un espacio vectorial..... 461

CAPÍTULO 8 466

TRANSFORMACIONES LINEALES..... 467

1. Introducción..... 467

1.1. Conceptos básicos..... 470

2. Transformaciones Lineales..... 473

2.1. Propiedades de las transformaciones lineales..... 482

3. Núcleo e imagen de una transformación lineal 489

4. Representación matricial de una transformación lineal 497

5. Diagonalización..... 502

5.1. Matrices semejantes y diagonalización..... 516

INTRODUCCIÓN

El Álgebra Lineal se ha convertido en una herramienta fundamental en la ingeniería moderna, puesto que permite modelar, analizar y resolver problemas en una amplia variedad de campos. En esta era gobernada por la tecnología y la innovación, esta rama de la matemática permite no solo abordar problemas en sistemas complejos, sino también brindar soluciones efectivas y eficientes.

En esta línea, las aplicaciones en la ingeniería son numerosas y su importancia sigue en aumento. En el desarrollo de software, el análisis de datos permite procesar grandes volúmenes de información y así obtener, por ejemplo, patrones de comportamiento. En este caso, el Álgebra Lineal proporciona las herramientas necesarias para realizar operaciones de transformación y manipulación de datos como el cálculo de vectores y matrices. Por medio de la formulación de ecuaciones lineales es posible modelar matemáticamente el comportamiento de un proceso químico y simular su evolución en el tiempo. Esto permite llevar a cabo un análisis de sensibilidad, optimización y control de los procesos, punto clave en el diseño y operación de plantas químicas. Un ingeniero industrial a cargo del área logística debe prever el abastecimiento, movimiento, distribución y venta de un producto o servicio. Haciendo uso de matrices de incidencia de una red de transporte es posible determinar las diferentes rutas que pueden conectar un destino con otro u otros, de esta manera se puede

programar la logística de distribución de una empresa a sus consumidores, optimizando así las variables que determinan una ventaja competitiva.

De esta manera, "Álgebra Lineal para Ingenierías" conjuga y sintetiza distintos escritos –sobre los temas fundamentales de esta área de la matemática– que se fueron elaborando a través de los años para colaborar con el proceso de enseñanza–aprendizaje de los alumnos que debían cursar la asignatura Álgebra y Geometría Analítica de las carreras de ingeniería de la Facultad de Ingeniería de la UNJu.

A partir del año 2023 en la, ahora, Cátedra Álgebra y Geometría Analítica se dictan las asignaturas: Geometría Analítica, Álgebra Lineal, Elementos de Álgebra y Geometría Analítica y Álgebra y Geometría Analítica. Es justamente para estas tres últimas que está dedicada la presente obra. Si bien existen diversos textos de variados autores que responden a los temas que competen a esta materia, la obra propuesta se destaca por estar pensada y escrita para los alumnos de la Facultad de Ingeniería de la UNJu, pues responde a los contenidos propuestos para este espacio curricular. No obstante esto, es una obra que puede ser utilizada por cualquier persona que requiera estudiar y/o profundizar temas de Álgebra Lineal ya que fue escrita en un lenguaje accesible y de fácil entendimiento. Es así que la meta que se persigue es proporcionar una base sólida para que el lector pueda comprender los conceptos

relacionados al Álgebra Lineal y estar mejor preparado para abordar temas más avanzados en las ciencias exactas.

El libro está estructurado en ocho capítulos y cada uno de ellos inicia con una introducción que destaca la importancia, desde distintos puntos de vistas, del motivo por el cual es fundamental estudiar el tema que se propone. Seguidamente, y de manera breve, se describen y relacionan los temas que se abordarán en cada uno de los apartados. Antes del desarrollo de estos temas se presenta, bajo el título Conceptos básicos, todas aquellas nociones elementales que permiten un mejor entendimiento de los temas a desarrollar.

A medida que se avanza en la lectura del libro se van haciendo explícitas las relaciones que poseen cada una de las partes que lo componen, a modo de ejemplo, cuando se trata el tema Representación gráfica de números complejos se pone en evidencia la relación directa que tiene con Vectores en $\mathbb{R}^2$ (n=2) del capítulo primero, o bien, al estudiar las distintas formas de expresar un sistema de ecuaciones lineales se resalta el vínculo directo que posee esto con matrices y, a su vez, estas con determinantes. Por lo tanto podría decirse que el libro fue escrito de manera inductiva, pues se desarrolla en primer lugar estructuras particulares (Vectores en $\mathbb{R}^n$, Números Complejos, Polinomios, Matrices) para llegar luego a una general (Espacios Vectoriales).

Aproximadamente en el año 2019 se comenzó con la concreción de los escritos que componen esta obra y, desde entonces, se compartieron con

sus destinatarios principales, quienes de manera desinteresada no solo hicieron llegar sus aportes, sino que también se encargaron de compartir estos escritos con otras personas por medio de los muy variados canales y formas que nos brinda esta era de las nuevas tecnologías que vivimos. Es por esto que se dedica e invita particularmente a los alumnos de la Facultad de Ingeniería de la UNJu a seguir participando y a disfrutar de esta obra que no solo enriquecerá su formación académica, sino que también les proporcionará una base sólida para su carrera profesional en ingeniería.

Los autores

CAPÍTULO 1

VECTORES EN R^n

1. Introducción

La geometría euclídea y el álgebra renacentista estaban predestinadas a entenderse y fusionarse, de hecho, lo hicieron por medio de la obra de Descartes (1596 – 1650) y Fermat (1601 – 1665) en una nueva geometría que se denominó Geometría Analítica. El alcance de esta geometría se vio reforzado por la aparición del cálculo infinitesimal, que se aplicó al estudio de los problemas geométricos.

En los siglos XVII y XVIII, el dominio de los métodos analíticos en las matemáticas era tal que la geometría euclídea se estudiaba olvidándose del rigor euclídeo y colocando en su lugar ecuaciones que representaban curvas, pares de números que representaban puntos y utilizando métodos de resolución de ecuaciones, al principio, y del cálculo infinitesimal después, como metodología de resolución de problemas geométricos.

A finales del siglo XVIII había unos pocos matemáticos que permanecían ligados al uso exclusivo de métodos denominados tradicionalmente geométricos, la gran mayoría utilizaba los métodos algebraicos enriquecidos por los métodos infinitesimales.

Del último grupo de matemáticos, es William K. Clifford (1845 – 1879) el primero en dar una formulación moderna de la definición de los vectores,

sus operaciones y los productos escalar y vectorial en su obra *Elementos de Dinámica.* Más precisamente, los métodos vectoriales se impusieron en la geometría entre el 1830 y el 1880.

En la actualidad, no solo partes de las matemáticas como: álgebra lineal, geometría analítica, cálculo avanzado, sino también, la teoría electromagnética, el análisis vectorial, entre otras, utilizan vectores a la hora de entender el proceso de matematización de algunos fenómenos de la naturaleza.

Las variables físicas que tienen magnitud, dirección y sentido se representan matemáticamente con vectores. La posición, velocidad y aceleración de una partícula son vectores, así como las fuerzas.

De este modo, es de gran importancia el estudio de la Geometría Analítica y, particularmente, del tema vectores, más aún si está destinado para estudiantes de ingeniería.

En este primer capítulo estudiaremos los conceptos más importantes de vectores en R^n y ejemplificaremos, en la gran mayoría, los casos para R^2 y R^3.

Luego de ver la definición formal, igualdad y vector posición de un punto, se realizará el estudio de las operaciones más comunes entre vectores: suma, producto por un escalar y resta.

En el apartado "Otras operaciones con vectores", se presentará uno de los tres productos entre vectores: el producto escalar y los diversos conceptos

que esta operación permite establecer entre vectores (perpendicularidad, módulo, ángulo, etc.).

Finalmente, se hace un estudio detallado del producto vectorial y mixto entre vectores; enfatizando en que, en la mayoría de las operaciones, se destaca tanto las aplicaciones en física, como su interpretación geométrica.

1.1. Conceptos básicos

El ser humano siempre se preocupó por cuantificar las cosas, es decir, asignar un valor cuantitativo o cualitativo a las cosas. Por ejemplo, podemos medir, de manera precisa, la distancia entre dos puntos, el peso de un insecto, la temperatura actual en una habitación, etc. De igual manera podemos medir, pero de forma menos precisa y más subjetivamente, la belleza de un cuadro artístico, el sabor de una comida.

Una magnitud es toda cualidad relativa a un proceso natural o un cuerpo que puede medirse y, consecuentemente, expresarse con un valor numérico.

Las magnitudes que pueden definirse mediante un número y las unidades utilizadas para su medida se denominan **magnitudes escalares**; en cambio, si además se debe especificar dirección y sentido, a esas magnitudes se las llama **vectoriales**.

Con esta diferencia intuitiva veamos la definición de vector y sus operaciones básicas.

Vector de R^n: *Llamaremos n–vector a una n–tupla ordenada de n números reales, así* $R^n = \{(x_1,\ x_2, \dots, x_n) \mid x_i \in R, 1 \leq i \leq n\}$

Los elementos de R^n se suelen denotar como $\overrightarrow{x} = (x_1,\ x_2, \dots, x_n)$, en donde x_1 es la primera componente, x_2 la segunda componente y en general x_i la i–ésima componente.

Cualquier vector de R^n cuyas componentes sean cero se denomina vector nulo y lo denotamos como $\overrightarrow{0}$.

Ejemplo 1: $\overrightarrow{u} = (-1, 2)$ es un vector de R^2 cuya primera y segunda componente son –1 y 2 respectivamente; en cambio $\overrightarrow{v} = (1, 2, -5, 0)$ es un vector de R^4 cuya primera, segunda, tercera y cuarta componente son 1, 2, –5 y 0, respectivamente.

Ejemplo 2: los vectores $\overrightarrow{e_1} = (1, 0, \dots, 0)$, $\overrightarrow{e_2} = (0, 1, \dots, 0)$, …, $\overrightarrow{e_n} = (0, 0, \dots, 1)$ son vectores de R^n y se los llama vectores canónicos. Vale decir, un vector canónico de R^n es aquel en donde solo una componente vale 1 y todas las demás, cero.

Geométricamente, un n–vector $\overrightarrow{x} = (x_1,\ x_2, \dots, x_n)$ es un segmento de recta orientado que tiene por punto inicial el origen $O = (0, 0, \dots, 0)$ y extremo un punto cualquiera distinto del origen, $X = (x_1,\ x_2, \dots, x_n)$. Por lo tanto, los elementos de R^n pueden pensarse como puntos o vectores geométricos de acuerdo con lo que se requiera en el contexto de trabajo.

Dado un segmento de recta orientado (vector geométrico) con punto inicial $P(p_1, p_2, \dots, p_n)$ y punto final $Q(q_1, q_2, \dots, q_n)$ existe un n–vector $\overrightarrow{u}$

con la misma longitud y dirección que $\overrightarrow{PQ}$: basta tomar $\vec{u} = \overrightarrow{PQ}$. En consecuencia:

A cada n–vector es posible asociarle un segmento de recta orientado y, recíprocamente, a cada segmento de recta orientado se le puede asociar un n–vector con la misma longitud y dirección, pero con punto inicial en el origen.

Vector posición de un punto: *Sea el punto* $P(p_1, p_2, \dots, p_n)$. *El punto P determina con el origen de coordenadas un vector* $\overrightarrow{OP}$ *que llamaremos vector posición del punto P.*

Ejemplo 1: en R^2 el punto $P(2,3)$ tiene por vector posición $\overrightarrow{OP} = (2,3)$. Figura 1.1

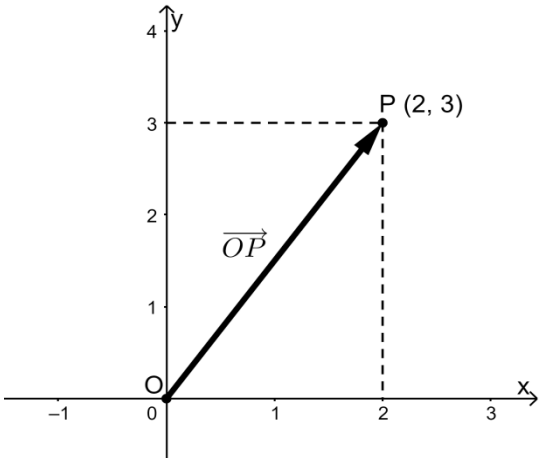


Figura 1.1. Vector posición en R^2.

Ejemplo 2: en R^3 el punto $Q(2,3,4)$ tiene por vector posición $\overrightarrow{OQ} = (2,3,4)$. Figura 1.2.

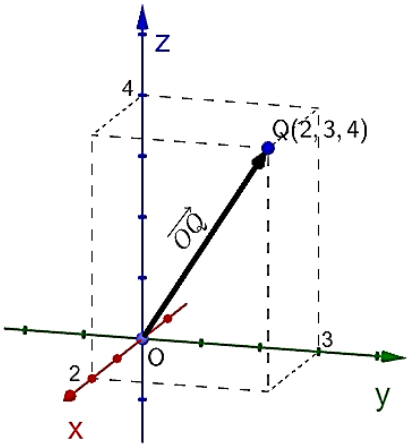


Figura 1.2. Vector posición R^3.

Igualdad de vectores: *Sean* $\vec{u} = (u_1, u_2, \dots, u_n)$ *y* $\vec{v} = (v_1, v_2, \dots, v_n)$ *vectores de* R^n*, se dice que son iguales si y solo si tienen todas las componentes correspondientes, iguales.*

Sean $\vec{u} = (u_1, u_2 \dots u_n)$ *y* $\vec{v} = (v_1, v_2, \dots, v_n) \in R^n$, $\vec{u} = \vec{v} \Leftrightarrow u_i = v_i \; \forall \; 1 \leq i \leq n$

Ejemplo 1: para que los vectores de R^3 (a, b, c) *y* $(2, -3, 0)$ sean iguales, debe ser $a = 2$, $b = -3$ *y* $c = 0$. Por razones similares $(-1, 4, 6)$ no es igual a $(4, -1, 6)$.

Ejemplo 2: determinar $x, y \in R$ para los cuales son iguales los siguientes vectores de R^2 $\vec{u} = (x - 1, 3)$ *y* $\vec{v} = (y - 3, x + 1)$.

Entonces: $\begin{cases} x - 1 = y - 3 \\ 3 = x + 1 \end{cases} \Rightarrow \begin{cases} x = 2 \\ y = 4 \end{cases}$

Verificación: si $x = 2$ *e* $y = 4 \Rightarrow \vec{u} = (1, 3)$ *y* $\vec{v} = (1, 3)$.

2. Operaciones con vectores

Dadas unas cantidades vectoriales, es posible obtener otras a partir de ellas si hacemos uso de las operaciones básicas entre vectores: la suma y la multiplicación por escalares que definiremos a continuación.

2.1. Suma de vectores

La suma de dos vectores, $\vec{u} = (u_1, u_2, \dots, u_n)$ *y* $\vec{v} = (v_1, v_2, \dots, v_n)$ *de* R^n *es otro vector,* $\vec{u} + \vec{v}$*, cuyas componentes son la suma de las respectivas componentes de los vectores* $\vec{u}$ *y* $\vec{v}$*.*

Sean $\vec{u} = (u_1, u_2, \dots, u_n)$ y $\vec{v} = (v_1, v_2, \dots, v_n) \in R^n \Longrightarrow$

$\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$

Ejemplo 1: sean los vectores de R^2 $\vec{u} = (1, 3)$ y $\vec{v} = (-3, 2)$ $\Longrightarrow$ $\vec{u} + \vec{v} = (-2, 5)$. Ver Figura 1.3.

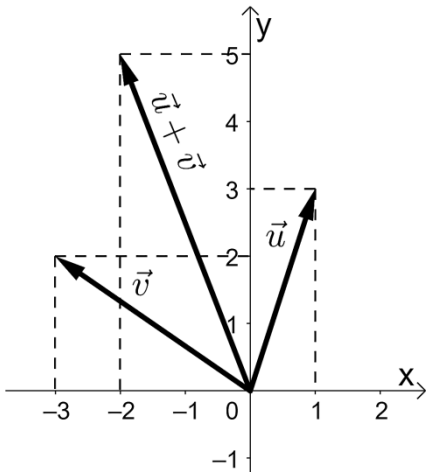


Figura 1.3. Suma de vectores en R^2.

Ejemplo 2: sean los vectores de R^3 $\vec{p} = (1, 3, -6)$ y $\vec{q} = (-3, 2, 4)$ al realizar su suma obtendremos: $\vec{p} + \vec{q} = (-2, 5, -2)$. Esto se muestra en la figura 1.4.

Ejemplo 3: sean los vectores de R^4: $\vec{r} = (3, -3, -4, -4)$ y $\vec{s} = (-3, -2, 2, -1)$, al realizar su suma obtendremos: $\vec{r} + \vec{s} = (0, -5, -2, -5)$.

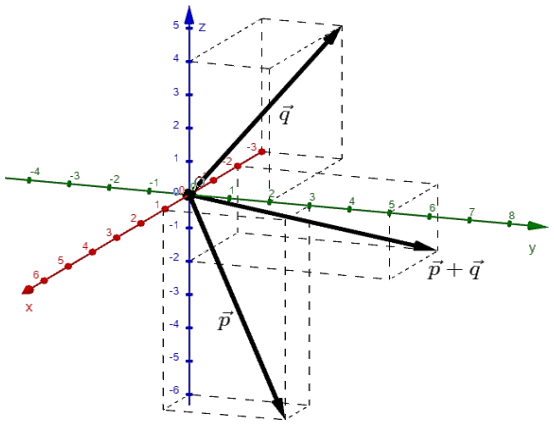


Figura 1.4. Suma de vectores en R^3.

Gráficamente, la suma en R^2 se puede realizar mediante la regla del paralelogramo, que consiste en construir un paralelogramo que tenga por lados los dos vectores dados, la diagonal de dicho paralelogramo representa la suma de los vectores.

La figura 1.5. muestra la suma del ejemplo 1.

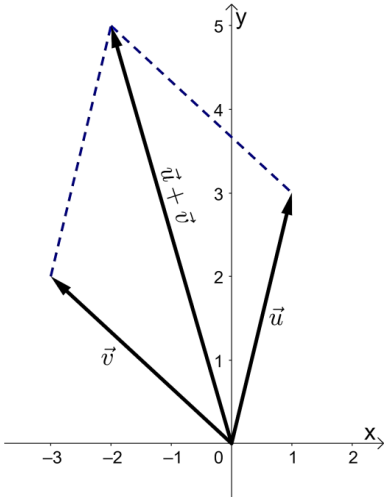


Figura 1.5. Suma de vectores, regla del paralelogramo.

2.2. Producto por un escalar

El producto de un vector $\overrightarrow{u} = (u_1, u_2, \dots, u_n)$ $de\ R^n$ *por un escalar* $\lambda \in R$*, es el vector* $\overrightarrow{\lambda u}$ *cuyas componentes son el producto del escalar por las respectivas componentes del vector dado.*

Sea $\overrightarrow{u} = (u_1, u_2, \dots, u_n) \in R^n \wedge \lambda \in R \implies \overrightarrow{\lambda u} = (\lambda u_1, \lambda u_2, \dots, \lambda u_n)$.

Ejemplo 1: sea el vector $\overrightarrow{u} = (1, -4, 6, 2)$ $de\ R^4$ y el escalar $\lambda = 3$, entonces el vector $\overrightarrow{\lambda u} = (3, -12, 18, 6)$.

En términos geométricos, el vector $\overrightarrow{\lambda u}$ es el vector que tiene igual dirección que $\overrightarrow{u}$ y que dependiendo del signo de λ tiene igual sentido ($\lambda > 0$) o sentido opuesto ($\lambda < 0$) a la del vector $\overrightarrow{u}$ y cuya longitud es igual al valor absoluto de λ por la longitud del vector $\overrightarrow{u}$.

Ejemplo 2: sea el vector $\overrightarrow{u} = (2, 3)\ de\ R^2\ \ y\ \lambda = 3,$ entonces el vector $\overrightarrow{\lambda u} = (6, 9)$ tendrá la misma dirección que $\overrightarrow{u}$ y una longitud igual al triple de la de $\overrightarrow{u}$.

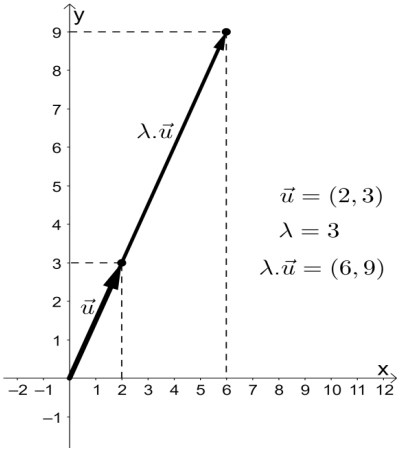


Figura 1.6. Producto de un vector por escalar en R².

Ejemplo 3: sea el vector $\overrightarrow{v} = (2, 4, 2)\ de\ R^3$ y el escalar $\lambda = -\frac{1}{2} \in \mathbb{R}$, entonces el vector $\overrightarrow{\lambda v} = (-1, -2, -1)$ tendrá dirección opuesta a la de $\overrightarrow{v}$ y una longitud igual a la mitad de la de $\overrightarrow{v}$. Figura 1.7.

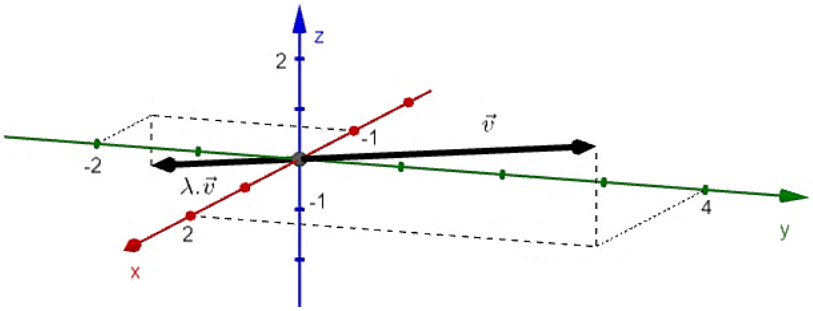


Figura 1.7. Producto de un vector por escalar en R³.

De acuerdo con la definición del producto de un vector por un escalar, si el vector de R^n es $\overrightarrow{u} = (u_1, u_2, \dots, u_n)$ y el escalar es:

a) $\lambda = -1$ entonces $\overrightarrow{\lambda u} = -\overrightarrow{u}$ vector opuesto a $\overrightarrow{u}$.

b) $\lambda = 0$ entonces $\overrightarrow{\lambda u} = \overrightarrow{0}$ vector nulo.

Aplicación del producto de un vector por un escalar: la Segunda Ley de Newton (Ley Fundamental de la Dinámica) relaciona el concepto de aceleración con la fuerza. Si una fuerza neta que actúa sobre un cuerpo hace que este acelere en la misma dirección que la fuerza neta, la magnitud de la aceleración es directamente proporcional a la magnitud de la fuerza neta que actúa sobre él, siendo sus direcciones y sentidos iguales. Esta ley puede enunciarse de la siguiente manera:

Si una fuerza [1] neta no nula $\left(\overrightarrow{F_N}\right)$ *actúa sobre un cuerpo, este cuerpo se acelera* $(\overrightarrow{a})$*. El vector aceleración tiene la misma dirección y sentido que la fuerza neta o resultante y es inversamente proporcional a la masa del cuerpo:* $\overrightarrow{a} = \frac{1}{m}\overrightarrow{F_N}$ *donde m es la masa del cuerpo, siendo una magnitud escalar.*

Esta relación también puede escribirse en su forma más conocida: $\overrightarrow{F_N} = m \cdot \overrightarrow{a}$ (1).

La ecuación (1) nos indica dos cuestiones:

a) Al ser la masa la medida de la inercia del cuerpo, que es un escalar, por lo tanto, en la segunda Ley de Newton se aplica la operación producto de un vector ($\overrightarrow{a}$) por un escalar (m) y se obtiene otro vector $\left(\overrightarrow{F_N}\right)$.

[1] Fuerza es una magnitud vectorial que se puede simbolizar como F.

b) Los vectores fuerza neta (o resultante) y aceleración son paralelos[2], ya que uno se obtiene a partir del otro multiplicado por un escalar.

Propiedades de la suma de vectores y producto por un escalar

Veremos, a continuación, cuáles son las propiedades algebraicas que poseen las dos operaciones básicas antes definidas. Estas propiedades son similares a las que poseen las operaciones de suma y producto de números reales.

P1) La suma es ley de composición interna en R^n.

$\forall \vec{x}, \vec{y} \in R^n \Longrightarrow (\vec{x} + \vec{y}) \in R^n$

Esto significa que la suma, como está definida, es una operación que a cada par de elementos de R^n le hace corresponder un único elemento de R^n.

D) Es inmediata por la definición de suma en R^n.

P2) La suma es asociativa en R^n.

$\forall \vec{x}, \vec{y}, \vec{z} \in R^n \Longrightarrow (\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$

D) Sean x_i, y_i, z_i las i–ésimas componentes de los vectores $\vec{x}, \vec{y}, \vec{z}$ respectivamente. Por lo tanto $x_i + y_i$ es la i–ésima componente de $\vec{x} + \vec{y}$así que $(x_i + y_i) + z_i$ es la i–ésima componente de $(\vec{x} + \vec{y}) + \vec{z}$. Como $(x_i + y_i) + z_i = x_i + (y_i + z_i)\ \forall\ i = 1, 2, \ldots, n$ por la propiedad asociativa de números reales, las componentes respectivas de $(\vec{x} + \vec{y}) + \vec{z}$ y de $\vec{x} + (\vec{y} + \vec{z})$ son iguales. Concluimos que $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$.

[2] Ver el apartado “Vectores paralelos”.

P3) La suma es conmutativa en R^n.

$\forall \vec{x}, \vec{y} \in R^n \Longrightarrow \vec{x} + \vec{y} = \vec{y} + \vec{x}$.

D) Sean x_i, y_i las i–ésimas componentes de los vectores $\vec{x}\, e\, \vec{y}$, respectivamente. Por lo tanto $x_i + y_i$ es la i–ésima componente de $\vec{x} + \vec{y}$. Como se cumple $x_i + y_i = y_i + x_i$ con $i = 1, 2 \ldots n$, por la propiedad conmutativa de números reales, las componentes respectivas de $\vec{x} + \vec{y}$ y de $\vec{y} + \vec{x}$ son iguales. Concluimos que $\vec{x} + \vec{y} = \vec{y} + \vec{x}$.

P4) Existe un único elemento neutro para la suma en R^n.

El elemento neutro se denota con $\vec{e}$.

$\exists \vec{e} \in R^n$ *único* $/ \forall \vec{x} \in R^n : \vec{x} + \vec{e} = \vec{e} + \vec{x} = \vec{x}$.

D) Para demostrar la existencia proponemos el elemento neutro: $\vec{e} = (e_1, e_2, \ldots, e_n)$, por lo tanto, debe cumplirse que:

$\vec{x} + \vec{e} = \vec{x}$

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

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

$x_1 + e_1 = x_1 \Longrightarrow e_1 = 0$

$x_2 + e_2 = x_2 \Longrightarrow e_2 = 0$

... $\Longrightarrow$ ...

$x_n + e_n = x_n \Longrightarrow e_n = 0$

$\therefore \vec{e} = (0, 0, \ldots, 0)$

La otra parte de la igualdad: $\vec{e} + \vec{x} = \vec{x}$ se demuestra de manera similar. Todo esto significa que el elemento neutro de la suma de vectores existe y es el vector nulo ya que se definió a este como aquel vector cuyas componentes son 0, es decir que $\vec{e} = \vec{0}$.

Para demostrar la unicidad del neutro vamos a suponer que existen dos neutros para la suma: $\overrightarrow{0_1}$ y $\overrightarrow{0_2}$, por lo tanto, si:

$\overrightarrow{0_1}$ es neutro, entonces $\overrightarrow{0_1} + \overrightarrow{0_2} = \overrightarrow{0_2}$

$\overrightarrow{0_2}$ es neutro, entonces $\overrightarrow{0_2} + \overrightarrow{0_1} = \overrightarrow{0_1}$

De donde se deduce, por ser los primeros miembros de ambas igualdades iguales, que $\overrightarrow{0_1} = \overrightarrow{0_2}$ y consecuentemente el neutro es único.

P5) Todo elemento en R^n admite un único inverso aditivo u opuesto en R^n.

El elemento opuesto a uno cualquiera $\overrightarrow{x}$, lo denotaremos por $\overrightarrow{o}$.

$\forall\, \overrightarrow{x} \in R^n, \exists\, \overrightarrow{o} \in R^n$ $\acute{u}nico$ $/\ \overrightarrow{x} + \overrightarrow{o} = \overrightarrow{o} + \overrightarrow{x} = \overrightarrow{0}$.

D) Para demostrar la existencia del opuesto se propone $\overrightarrow{o} = (o_1, o_2, \dots, o_n)$, por lo tanto, debe cumplirse que:

$\overrightarrow{x} + \overrightarrow{o} = \overrightarrow{0}$

$(x_1, x_2, \dots, x_n) + (o_1, o_2, \dots, o_n) = (0, 0, \dots, 0)$

$(x_1 + o_1, x_2 + o_2, \dots, x_n + o_n) = (0, 0, \dots, 0)$

$x_1 + o_1 = 0 \implies o_1 = -x_1$

$x_2 + o_2 = 0 \implies o_2 = -x_2$

... $\implies$ ...

$x_n + o_n = 0 \implies o_n = -x_n$

$\therefore\ \overrightarrow{o} = (-x_1, -x_2, \dots, -x_n)$

La otra parte de la igualdad: $\overrightarrow{o} + \overrightarrow{x} = \overrightarrow{0}$ se demuestra de manera similar. Todo esto significa que el inverso aditivo u opuesto a un elemento $\overrightarrow{x} \in R^n$ existe y es el vector opuesto al dado ya que cuando se definió producto por un escalar, si el escalar es –1 al multiplicarse por cualquier

vector, se obtiene el opuesto de dicho vector, es decir que $\overrightarrow{o} = \overrightarrow{-x}\ \forall\ \overrightarrow{x} \in R^n$.

Para demostrar la unicidad del elemento opuesto vamos a suponer que, para un elemento $\overrightarrow{x}\ de\ R^n$ existen dos opuestos $\overrightarrow{x_1}\ y\ \overrightarrow{x_2}$, por lo tanto se debe cumplir que $\overrightarrow{x} + \overrightarrow{x_1} = \overrightarrow{0}$ y también $\overrightarrow{x} + \overrightarrow{x_2} = \overrightarrow{0}$ de donde se deduce que $\overrightarrow{x_1} = \overrightarrow{x_2}$. Consecuentemente, el inverso aditivo u opuesto de cada elemento de R^n es único.

P6) El producto es ley de composición externa en R^n con escalares en R.

$\forall \alpha \in R, \forall\ \overrightarrow{x} \in\ R^n \Longrightarrow \alpha\overrightarrow{x} \in\ R^n$.

Esto significa que el producto de un vector de R^n por un escalar, como está definido, es una operación que, a cada par vector–escalar le hace corresponder un único elemento de R^n.

D) La existencia de tal vector es inmediata, por la definición de producto de un vector de R^n por un escalar.

P7) El producto por un escalar satisface la asociatividad mixta.

$\forall \alpha, \beta \in R, \forall\ \overrightarrow{x} \in\ R^n: \alpha(\beta\ \overrightarrow{x}) = (\alpha\beta)\ \overrightarrow{x}$.

D) Sea x_i la i–ésima componente del vector $\overrightarrow{x}$. Por lo tanto, βx_i es la i–ésima componente de $\overrightarrow{\beta x}$ y, por lo tanto, $\alpha(\beta x_i)$ la i–ésima componente de $\overrightarrow{\alpha(\beta x)}$. Como sabemos que $\alpha(\beta x_i) = (\alpha\beta)x_i\ \forall\ i = 1, 2, \dots, n$ por la propiedad asociativa del producto de números reales, las componentes respectivas de $\alpha(\beta\ \overrightarrow{x})$ y de $(\alpha\beta)\overrightarrow{x}$ son iguales. Concluimos entonces que $\alpha(\beta\overrightarrow{x}) = (\alpha\beta)\overrightarrow{x}$.

P8) El producto por un escalar es distributivo respecto a la suma en R.

$\forall \alpha, \beta \in R, \forall \overrightarrow{x} \in R^n: (\alpha + \beta)\overrightarrow{x} = \alpha\overrightarrow{x} + \beta\overrightarrow{x}$.

D) Sea x_i la i–ésima componente del vector $\overrightarrow{x}$. Por lo tanto, $(\alpha + \beta)x_i$ es la i–ésima componente de $(\alpha + \beta)\overrightarrow{x}$. Como $(\alpha + \beta)x_i = \alpha x_i + \beta x_i \ \forall\, i = 1, 2, \dots, n$ por la propiedad distributiva del producto de números reales, respecto a la suma, las componentes respectivas de: $(\alpha + \beta)\overrightarrow{x}$ y de $\alpha\overrightarrow{x} + \beta\overrightarrow{x}$ son iguales. Entonces, por igualdad de vectores, se concluye: $(\alpha + \beta)\overrightarrow{x} = \alpha\overrightarrow{x} + \beta\overrightarrow{x}$.

P9) El producto por un escalar es distributivo respecto a la suma en R^n.

$\forall \alpha \in R, \forall \overrightarrow{x}, \overrightarrow{y} \in R^n : \alpha(\overrightarrow{x} + \overrightarrow{y}) = \alpha\overrightarrow{x} + \alpha\overrightarrow{y}$.

D) Sean x_i, y_i las i–ésimas componentes de los vectores $\overrightarrow{x}\, e\, \overrightarrow{y}$ respectivamente. Por lo tanto $x_i + y_i$ es la i–ésima componente de $\overrightarrow{x} + \overrightarrow{y}$, como también $\alpha(x_i + y_i)$ es la i–ésima componente de $\alpha(\overrightarrow{x} + \overrightarrow{y})$. Como $\alpha(x_i + y_i) = \alpha\, x_i + \alpha\, y_i \ \forall\, i = 1, 2, \dots, n$ por la propiedad distributiva del producto de números reales, respecto a la suma, las componentes respectivas de $\alpha(\overrightarrow{x} + \overrightarrow{y})$ y de $\alpha\,\overrightarrow{x} + \alpha\,\overrightarrow{y}$ son iguales. Concluimos entonces que $\alpha(\overrightarrow{x} + \overrightarrow{y}) = \alpha\,\overrightarrow{x} + \alpha\overrightarrow{y}$.

P10) La unidad para los R es el neutro para el producto por un escalar.

$\forall \overrightarrow{x} \in R^n: 1\,\overrightarrow{x} = \overrightarrow{x}$.

D) Siendo $\overrightarrow{x} = (x_1, x_2, \dots, x_n)$ tenemos:

$1\,\overrightarrow{x} = 1\,(x_1, x_2, \dots, x_n)$

$1\,\overrightarrow{x} = (1\,x_1, 1\,x_2, \dots, 1\,x_n)$

$1\,\vec{x} = (x_1, x_2, \dots, x_n)$

$1\,\vec{x} = \vec{x}$

El conjunto de vectores de R^n entre los cuales se definieron las operaciones de suma y producto por un escalar con las respectivas propiedades, determinan un espacio vectorial y constituyen una estructura de las más importantes en matemática.

La importancia de esta definición axiomática es que amplía la idea geométrica intuitiva de vector. Cualquier conjunto de elementos entre los cuales puedan definirse las operaciones anteriores y que cumplan con las mismas propiedades podrá considerarse como un espacio vectorial y sus elementos como vectores, aunque a veces pueda resultar difícil interpretarlos como segmentos orientados.

2.3. Resta de vectores

La resta de dos vectores, $\vec{u} = (u_1, u_2, \dots, u_n)$ *y* $\vec{v} = (v_1, v_2, \dots, v_n)$ *de* R^n *es otro vector,* $\vec{u} - \vec{v}$*, que se obtiene sumando al vector minuendo el opuesto del vector sustraendo, es decir,* $\vec{u} - \vec{v} = \vec{u} + (-\vec{v})$.

Sean $\vec{u} = (u_1, u_2, \dots, u_n)$ y $\vec{v} = (v_1, v_2, \dots, v_n) \in R^n \implies \vec{u} - \vec{v} = \vec{u} + (-\vec{v})$.

Geométricamente, el vector opuesto a uno dado es aquel que tiene la misma dirección y longitud, pero sentido contrario.

Ejemplo 1: sean los vectores de R^2 $\vec{u} = (1, 3)$ y $\vec{v} = (-3, 2) \implies$ $\vec{u} - \vec{v} = \vec{u} + (-\vec{v}) = (1, 3) + (3, -2) = (4, 1)$. Figura 1.8.

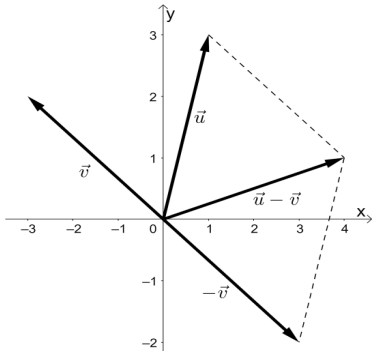


Figura 1.8. Resta de vectores en R^2.

Geométricamente, se puede interpretar que la resta $\vec{u} - \vec{v}$ es la diagonal del paralelogramo de lados $\vec{u}\ y\ \vec{v}$ que va desde el punto final de vector $\vec{v}$ hasta el punto final del vector $\vec{u}$. Tal como se muestra en la Figura 1.9.

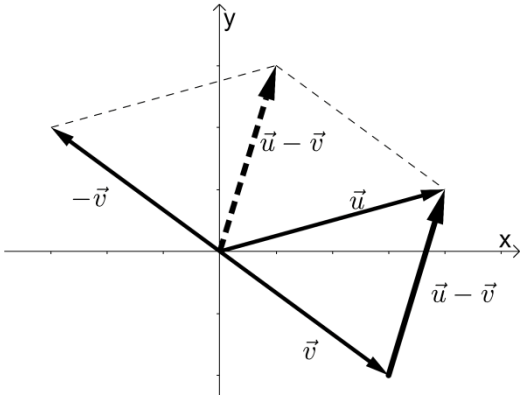


Figura 1.9. Resta de vectores en R^2.

Ejemplo 2: sean los vectores de R^3 $\vec{p} = (2,3,2)\ y\ \vec{q} = (-2,2,4) \implies \vec{p} - \vec{q} = (4,1,-2)$. Esto se muestra en la figura 1.10.

2.4. Vector determinado por dos puntos de R^n

Las componentes de un vector determinado por dos puntos P y Q de R^n son iguales a las diferencias entre las coordenadas respectivas del punto P (punto extremo) y Q (punto origen). Es decir: $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ siendo $\overrightarrow{OQ}$ y $\overrightarrow{OP}$ el vector posición del punto Q y P respectivamente.

Si $P(p_1, p_2, \dots, p_n)$ y $Q(q_1, q_2, \dots, q_n) \in R^n \Longrightarrow \overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$

$\overrightarrow{PQ} = (q_1, q_2, \dots, q_n) - (p_1, p_2, \dots, p_n)$

$\overrightarrow{PQ} = (q_1 - p_1, q_2 - p_2, \dots, q_n - p_n)$

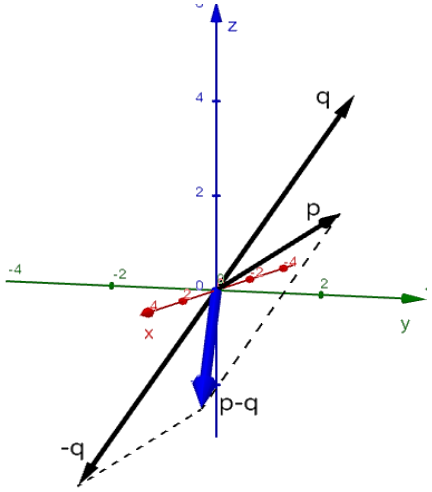


Figura 1.10. Resta de vectores en R^3.

Ejemplo 1: sean los puntos $A(4, 1)$ y $B(6, 4)$, entonces los vectores posición son, respectivamente, $\overrightarrow{OA} = (4, 1)$ y $\overrightarrow{OB} = (6, 4)$.

$\Longrightarrow \overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (6, 4) - (4, 1) = (2, 3)$.

Debemos tener en cuenta que $\overrightarrow{AB} \neq \overrightarrow{BA}$ ya que $\overrightarrow{AB} = (2, 3)$, Figura 1.11a, mientras que $\overrightarrow{BA} = (-2, -3)$, Figura 1.11b.

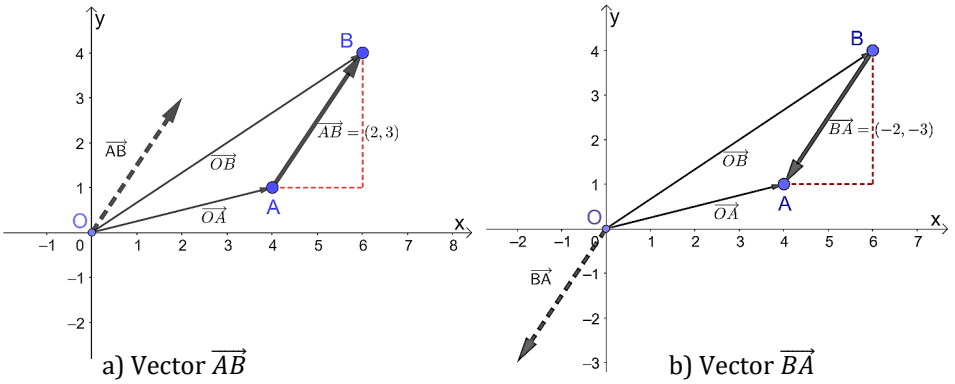


Figura 1.11. Vectores determinados por dos puntos en R^2.

Ejemplo 2: sean los puntos P y Q, los vectores posición de estos puntos son:

$\overrightarrow{OP} = (-2, 3, 2)\ y\ \overrightarrow{OQ} = (2, 2, 5) \in R^3$, respectivamente.

$\Longrightarrow \overrightarrow{PQ} = (2, 2, 5) - (-2, 3, 2) = (4, -1, 3)$. Esto se muestra en la figura 1.12.

En R^3 ocurre exactamente lo mismo que en R^2 respecto a los vectores $\overrightarrow{PQ}\ y\ \overrightarrow{QP}$.

3. Vectores paralelos

Dos vectores no nulos $\overrightarrow{u} = (u_1, u_2, \dots, u_n)\ y\ \overrightarrow{v} = (v_1, v_2, \dots, v_n)$ *de* R^n *se dicen paralelos si existe* $\lambda \in R,\ \lambda \neq 0$ *tal que* $\overrightarrow{u} = \lambda\overrightarrow{v}$.

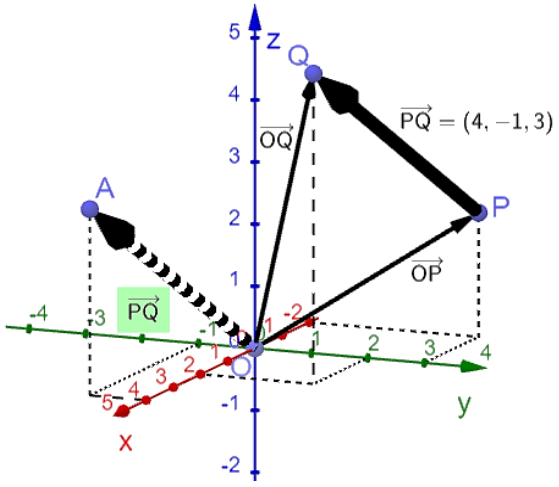


Figura 1.12. Vector determinado por dos puntos en R^3 .

La definición de vectores paralelos nos indica que dos vectores no nulos son paralelos si uno se puede obtener a partir del otro, multiplicado por un escalar.

Ejemplo 1: los vectores $\vec{u} = (6, 9)$ y $\vec{v} = (2, 3)$ de R^2 son paralelos ya que $\vec{u} = 3\vec{v}$. Figura 1.13.

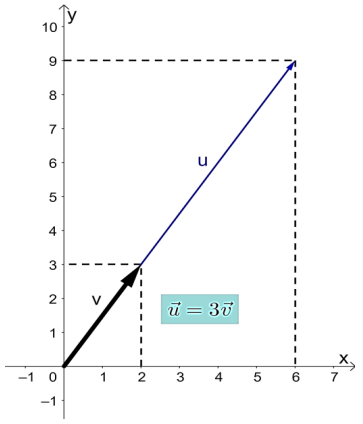


Figura 1.13. Vectores paralelos en R^2.

Veamos una característica importante de los vectores paralelos, que se deduce a partir de la definición.

$\overrightarrow{u} = \lambda\overrightarrow{v}$

$(u_1, u_2, \dots, u_n) = \lambda(v_1, v_2, \dots, v_n)$

$(u_1, u_2, \dots, u_n) = (\lambda v_1, \lambda v_2, \dots, \lambda v_n)$

$$\begin{cases} u_1 = \lambda v_1 \implies \lambda = \dfrac{u_1}{v_1} \\ u_2 = \lambda v_2 \implies \lambda = \dfrac{u_2}{v_2} \\ \vdots \\ u_n = \lambda v_n \implies \lambda = \dfrac{u_n}{v_n} \end{cases}$$

Por lo tanto $\implies \lambda = \dfrac{u_1}{v_1} = \dfrac{u_2}{v_2} = \dots = \dfrac{u_n}{v_n}$

Esto significa que la *condición necesaria y suficiente* para que dos vectores sean paralelos es que sus componentes homólogas sean proporcionales.

Si el valor de la constante de proporcionalidad es positivo, los vectores tienen el mismo sentido, y si es negativo tienen sentido opuesto.

Ejemplo 2: los vectores $\overrightarrow{u} = (4, -6, 6)$ y $\overrightarrow{v} = (-2, 3, -3)$ de R^3 son paralelos y de sentidos opuestos ya que $\overrightarrow{u} = -2\overrightarrow{v}$. Figura 1.14.

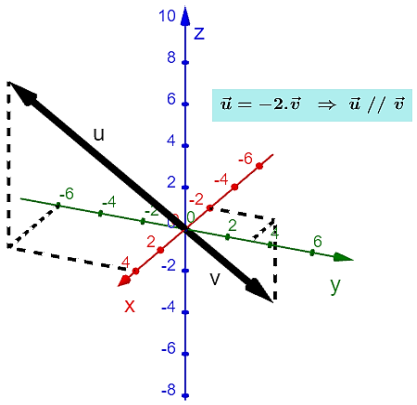


Figura 1.14. Vectores paralelos en R^3.

4. Otras operaciones con vectores

Se mencionó que una de las características de un vector es la longitud que tiene este, concepto que intuitivamente entendemos pero que deberíamos formalizar. Para este efecto, estudiaremos otra operación: el producto escalar entre dos vectores, la cual se relaciona directamente con el concepto de longitud de un vector y la noción de ángulo entre vectores.

Cabe mencionar que el producto escalar no es el único producto que se realiza entre vectores, también existen el producto vectorial y el producto mixto.

4.1. Producto escalar[3]

Dados dos vectores, $\overrightarrow{u} = (u_1, u_2, \dots, u_n)$ *y* $\overrightarrow{v} = (v_1, v_2, \dots, v_n)$ *de* R^n *definimos producto escalar,* $\overrightarrow{u} \cdot \overrightarrow{v}$*, como el escalar que se obtiene al sumar los productos de las componentes de* $\overrightarrow{u}$ *con las respectivas componentes de* $\overrightarrow{v}$

Dados $\overrightarrow{u} = (u_1, u_2, \dots, u_n)$ *y* $\overrightarrow{v} = (v_1, v_2, \dots, v_n) \in R^n \Longrightarrow$

$\overrightarrow{u} \cdot \overrightarrow{v} = u_1v_1 + u_2v_2 + \dots + u_nv_n$*, escrito de manera resumida* $\overrightarrow{u} \cdot \overrightarrow{v} = \sum_{i=1}^{n} u_iv_i$

Ejemplo 1: dados los vectores $\overrightarrow{u} = (1, -2), \overrightarrow{v} = (-2, 5) \in \mathbb{R}^2$ y los vectores de espacio $\mathbb{R}^3$ $\overrightarrow{x} = (3, -2, 1)$, $\overrightarrow{y} = (3, 0, -1)$. Calcular los siguientes productos escalares: $\overrightarrow{u} \cdot \overrightarrow{v}$, $\overrightarrow{x} \cdot \overrightarrow{y}$, $\overrightarrow{u} \cdot \overrightarrow{x}$.

$\overrightarrow{u} \cdot \overrightarrow{v} = 1(-2) + (-2)5 = -12.$

$\overrightarrow{x} \cdot \overrightarrow{y} = 3\,.3 + (-2)0 + 1(-1) = 8.$

[3] Esta operación también suele llamarse producto punto o producto interno.

$\vec{u} \cdot \vec{x}$ evidentemente no está definido ya que son vectores de distintos espacios.

Al ser el producto escalar una nueva operación entre vectores cuyo resultado es un número real, nos podemos preguntar ¿cuáles de las propiedades del producto entre números reales se satisfacen para el producto escalar entre vectores?

Propiedades del producto escalar

P1) El producto escalar es conmutativo

$\forall \vec{x}, \vec{y} \in R^n \Longrightarrow \vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$

D) Sean x_i, y_i las i–ésimas componentes de los vectores $\vec{x}$ e $\vec{y}$ respectivamente. Por lo tanto $x_i y_i$ es la i–ésima componente de $\vec{x} \cdot \vec{y}$. Como $x_i\, y_i = y_i\, x_i \; \forall\, i = 1, 2, \ldots, n$ por la propiedad conmutativa de números reales, las componentes respectivas de $\vec{x} \cdot \vec{y}$ e $\vec{y} \cdot \vec{x}$ son iguales. Concluimos entonces que $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$

P2) El producto escalar no es asociativo

$\forall \vec{x}, \vec{y}, \vec{z} \in R^n \Longrightarrow (\vec{x} \cdot \vec{y}) \cdot \vec{z} \neq \vec{x} \cdot (\vec{y} \cdot \vec{z})$

D) Para comprobar que esto es verdad, razonamos de la siguiente manera. Al realizar el producto escalar $\vec{x} \cdot \vec{y}$ obtenemos un escalar λ, que al multiplicar al vector $\vec{z}$ da por resultado un vector en la misma dirección que este último, mientras que al hacer el producto escalar $\vec{y} \cdot \vec{z}$ se obtiene un escalar, ρ, que al multiplicar al vector $\vec{x}$ da por resultado un vector en la misma dirección que este último. Por lo tanto, los vectores que se obtienen en ambos miembros no necesariamente tienen las mismas

direcciones, quedando así comprobado que el producto escalar no es asociativo.

El razonamiento anterior se puede resumir esquemáticamente de la siguiente manera:

$$\underbrace{\underbrace{(\vec{x}\cdot\vec{y})}_{\lambda}\cdot\vec{z}}_{\overrightarrow{\lambda z}} \neq \underbrace{\vec{x}\cdot\underbrace{(\vec{y}\cdot\vec{z})}_{\rho}}_{\overrightarrow{\rho x}}$$

P3) El producto escalar es distributivo respecto a la suma de vectores:

$\forall \vec{x},\vec{y},\vec{z} \in R^n: \vec{x}\cdot(\vec{y}+\vec{z}) = \vec{x}\cdot\vec{y} + \vec{x}\cdot\vec{z} \ \wedge \ (\vec{y}+\vec{z})\cdot\vec{x} = \vec{y}\cdot\vec{x} + \vec{z}\cdot\vec{x}$

D) Sean x_i, y_i, z_i las i–ésimas componentes de los vectores $\vec{x},\vec{y},\vec{z}$ respectivamente. Por lo tanto $y_i + z_i$ es la i–ésima componente de $\vec{y}+\vec{z}$, también $x_i(y_i + z_i)$ será la i–ésima componente de $\vec{x}(\vec{y}+\vec{z})$. Como $x_i(y_i + z_i) = x_i y_i + x_i z_i \ \forall\, i = 1, 2, \ldots, n$ por la propiedad distributiva del producto de números reales, respecto a la suma, las componentes respectivas de $\vec{x}\cdot(\vec{y}+\vec{z})$ y de $\vec{x}\cdot\vec{y} + \vec{x}\cdot\vec{z}$ son iguales. Concluimos entonces que $\vec{x}(\vec{y}+\vec{z}) = \vec{x}\cdot\vec{y} + \vec{x}\cdot\vec{z}$

Para la segunda parte de la propiedad, se razona de manera similar.

P4) El producto escalar es asociativo respecto al producto por un número real:

$\forall\, \vec{x},\vec{y} \in R^n, \forall\, \lambda \in R : (\lambda\vec{x})\cdot\vec{y} = \vec{x}\cdot(\lambda\vec{y}) = \lambda(\vec{x}\cdot\vec{y})$

D) Sean x_i, y_i las i–ésimas componentes de los vectores $\vec{x}\, e\, \vec{y}$ respectivamente. Por lo tanto $\lambda\, x_i\, y_i$ es la i–ésima componente de $(\lambda\vec{x})\cdot\vec{y}$. Como sabemos $\lambda\, x_i\, y_i = x_i\, \lambda y_i$ para todo $i = 1, 2, \ldots, n$ por la

propiedad conmutativa del producto de números reales, las componentes respectivas de $(\lambda\vec{x})\cdot\vec{y}$ y de $\vec{x}\cdot(\lambda\vec{y})$ son iguales. Concluimos entonces que $(\lambda\vec{x})\cdot\vec{y} = \vec{x}\cdot(\lambda\vec{y})$

Para la segunda parte de la propiedad, se razona de manera similar.

P_5) El producto escalar de un vector por sí mismo es un número real mayor o igual que cero

$\forall\, \vec{x} \in R^n : \vec{x}\cdot\vec{x} \geq 0$

D) Sea $\vec{x} = (x_1, x_2, \dots, x_n) \in R^n$ entonces:

$\vec{x}\cdot\vec{x} = (x_1x_1 + x_2x_2 + \dots + x_nx_n) = (x_1^2 + x_1^2 + \dots + x_n^2)$ y la suma de cuadrados siempre es positiva o cero.

P_6) El producto escalar de un vector por sí mismo es cero si y solo si el vector es nulo.

$\vec{x}\cdot\vec{x} = 0 \Leftrightarrow \vec{x} = \vec{0}$

$\vec{x}\cdot\vec{x} = (x_1^2 + x_2^2 + \dots + x_n^2) = 0 \Leftrightarrow x_i^2 = 0\ \forall\, i = 1, \dots, n$ De donde se puede concluir que $\vec{x}\cdot\vec{x} = 0 \Leftrightarrow \vec{x} = \vec{0}$

Sabemos que si el producto de dos números reales x e y es cero, necesariamente es x o y igual a cero. Esto no ocurre con el producto escalar de vectores, como lo muestra el siguiente ejemplo.

Sean los vectores $\vec{u} = (2, 1, -5)\ y\ \vec{v} = (-2, -1, -1)\ de\ R^3$ entonces

$\vec{u}\cdot\vec{v} = 2(-2) + 1(-1) + (-5)(-1) = 0$

Otro concepto importante en el cual se hace uso del producto escalar es el de perpendicularidad entre vectores.

Vectores perpendiculares

Dos vectores no nulos $\vec{u} = (u_1, u_2, \dots, u_n)$ *y* $\vec{v} = (v_1, v_2, \dots, v_n)$ *de* R^n *son perpendiculares si y solo si su producto escalar es nulo.*

$\forall\, \vec{u},\ \vec{v} \in R^n :\ \vec{u} \perp \vec{v} \Leftrightarrow \vec{u} \cdot \vec{v} = 0$

Ejemplo 1: en R^2 los vectores $\vec{u} = (-2, 3)$ *y* $\vec{v} = (-3, -2)$ son perpendiculares ya que $\vec{u} \cdot \vec{v} = -2(-3) + 3(-2) = 0$. Esto se muestra en la figura 1.15.

Ejemplo 2: determinar los valores de $k \in \mathbb{R}$ para los cuales los siguientes vectores de R^3 $\vec{x} = (-3k, -4, 1)$ *e* $\vec{y} = (-k, k, 1)$ son perpendiculares.

Para que los vectores sean perpendiculares, debe cumplirse que:

$\vec{x} \cdot \vec{y} = 0$. O sea que: $(-3k, -4, 1) \cdot (-k, k, 1) = 3k^2 - 4k + 1 = 0$

Las raíces de esta ecuación son $k_1 = 1$ *y* $k_2 = \frac{1}{3}$

Los pares de vectores de R^3 que satisfacen la condición son:

$\vec{x_1} = (-3, -4, 1)$ *e* $\vec{y_1} = (-1, 1, 1)$ y $\vec{x_2} = (-1, -4, 1)$ *e* $\vec{y_2} = (-\frac{1}{3}, \frac{1}{3}, 1)$.

Figura 1.16 a) y b)

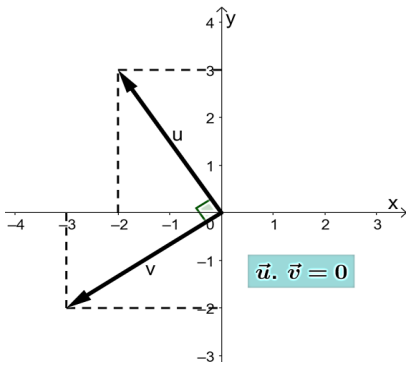
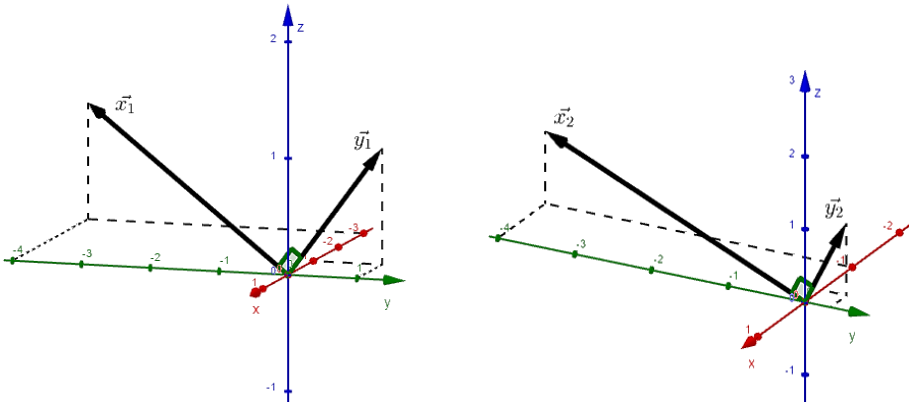


Figura 1.15. Vectores perpendiculares en R^2.

Las figuras correspondientes al ejemplo 2 son las siguientes:



a) $k_1 = 1$ b) $k_2 = 1/3$

Figura 1.16. Vectores perpendiculares en R^3.

Ejemplo 3: comprobar que los vectores $\check{\imath} = (1,0)\ y\ \check{\jmath} = (0,1) de\ R^2$ son perpendiculares.

$(1,0) \cdot (0,1) = 0 + 0 = 0$

Una vez que se analice el tema de ángulo entre vectores, se podrá demostrar con más facilidad la condición necesaria y suficiente de perpendicularidad entre vectores.

Una de las utilidades del producto escalar entre vectores es que permite definir módulo[4] de un vector. Antes de estudiar esta definición, haremos referencia a dos casos particulares y luego generalizaremos el concepto.

[4] En la literatura aparecen otros nombres como magnitud, longitud o norma para referirse a módulo de un vector.

Sea P un punto del plano (R^2) cuyas coordenadas son (x, y) o sea $P(x, y)$. Representando P en un sistema de coordenadas cartesianas ortogonales y calculando (mediante el Teorema de Pitágoras) la longitud del segmento comprendido entre el origen del sistema y el punto P, obtenemos $\sqrt{x^2 + y^2}$

Esto se muestra en la Figura 1.17.

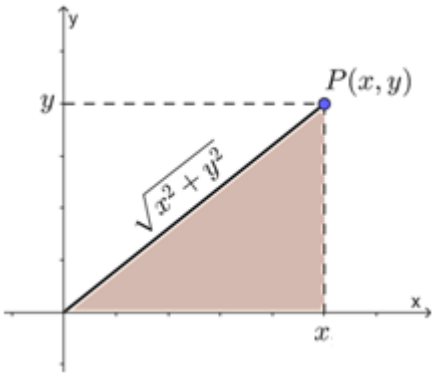


Figura 1.17. Longitud del segmento de recta en R^2.

De la misma manera, si estamos en el espacio (R^3), el punto será $P(x, y, z)$, aplicando el Teorema de Pitágoras dos veces obtenemos la longitud del segmento comprendido entre O y P e igual a $\sqrt{x^2 + y^2 + z^2}$, como se muestra en la Figura 1.18.

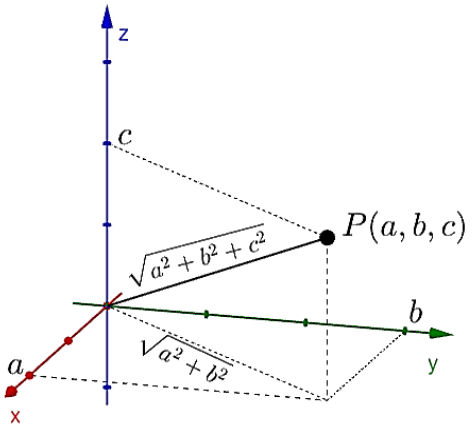


Figura 1.18. Longitud del segmento de recta R^3.

Por otro lado, sabemos que el vector $\overrightarrow{OP}$ es el vector posición del punto P, de esta manera la longitud del vector $\overrightarrow{OP}$ *en* R^2 ó *en* R^3 es la raíz cuadrada de la suma de los cuadrados de las componentes del vector. Esta observación se puede generalizar y nos induce a la definición de módulo de un vector.

Módulo de un vector: *Se define módulo de un vector* $\vec{u} = (u_1, u_2, \dots, u_n)$ *de* R^n, $\|\vec{u}\|$, *como la raíz cuadrada de* $\vec{u} \cdot \vec{u}$

Sea $\vec{u} = (u_1, u_2, \dots, u_n) \in R^n \Longrightarrow \|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2}$

Ejemplo 1: sea el vector $\vec{u} = (-3, 4) \Longrightarrow \|\vec{u}\| = \sqrt{(-3)^2 + 4^2} = \sqrt{25} = 5$

Ejemplo 2: el concepto de módulo de un vector puede usarse para demostrar el Teorema de Pitágoras.

Si $\vec{x}$ *e* $\vec{y}$ son dos vectores perpendiculares, entonces $\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$

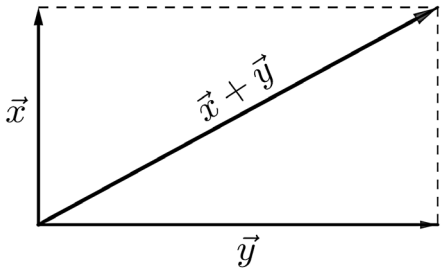


Figura 1.19. Suma de dos vectores perpendiculares.

En efecto:

$\|\vec{x}+\vec{y}\|^2 = (\vec{x}+\vec{y}\,)\cdot(\vec{x}+\vec{y}\,) = \vec{x}\cdot\vec{x}+\vec{x}\cdot\vec{y}+\vec{y}\cdot\vec{x}+\vec{y}\cdot\vec{y}$

$\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + 0 + 0 + \|\vec{y}\|^2$

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

Podemos observar que $\vec{x}\cdot\vec{x} = \|\vec{x}\|^2$ es una de las propiedades del módulo de un vector que se demostrarán a continuación.

Propiedades del módulo de un vector

P_1) El módulo del producto de un escalar por un vector es igual al producto del valor absoluto del escalar por el módulo del vector.

$\forall\, \vec{u} \in R^n\ \forall \alpha \in R:\ \|\alpha\,\vec{u}\| = |\alpha|\,\|\vec{u}\|$

D) $\|\alpha\,\vec{u}\| = \sqrt{\alpha\,\vec{u}\cdot\alpha\,\vec{u}} = \sqrt{\alpha^2(\,\vec{u}\cdot\vec{u}\,)} = \sqrt{\alpha^2}\,\sqrt{\vec{u}\cdot\vec{u}} = |\alpha|\,\|\vec{u}\|$

P_2) El módulo de un vector es nulo si y solo si el vector es nulo.

$\|\vec{u}\| = 0 \Leftrightarrow \vec{u}\cdot\vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}$

P_3) El producto escalar de un vector por sí mismo es igual al cuadrado de su módulo.

$\forall\, \vec{u} \in R^n: \vec{u}\cdot\vec{u} = \|\vec{u}\|^2$

D) $\overrightarrow{u} \cdot \overrightarrow{u} = u_1^2 + u_2^2 + \ ... + u_n^2$ como $\|\overrightarrow{u}\| = \sqrt{u_1^2 + u_2^2 + \ ... + u_n^2}$ se sigue que

$\overrightarrow{u} \cdot \overrightarrow{u} = \|\overrightarrow{u}\|^2$

P_4) El cuadrado del módulo de la suma de dos vectores es igual al cuadrado del módulo del primer vector más el cuadrado del módulo del segundo vector, más el doble de su producto escalar.

$\forall\, \overrightarrow{u},\ \overrightarrow{v} \in\ R^n$: $\|\vec{u} + \overrightarrow{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2(\ \overrightarrow{u} \cdot \overrightarrow{v}\ )$

D) $\|\vec{u} + \overrightarrow{v}\|^2 = (\vec{u} + \overrightarrow{v}) \cdot (\vec{u} + \overrightarrow{v}\ )$

$\|\vec{u} + \overrightarrow{v}\|^2 = \vec{u}\ (\vec{u} + \overrightarrow{v}\ ) + \overrightarrow{v}(\vec{u} + \overrightarrow{v}\ )$

$\|\vec{u} + \overrightarrow{v}\|^2 = \vec{u} \cdot \vec{u} + \vec{u} \cdot \overrightarrow{v} + \overrightarrow{v} \cdot \vec{u} + \overrightarrow{v} \cdot \overrightarrow{v}$

$\|\vec{u} + \overrightarrow{v}\|^2 = \vec{u} \cdot \vec{u} + 2(\ \vec{u} \cdot \overrightarrow{v}\ ) + \overrightarrow{v} \cdot \overrightarrow{v}$

$\|\vec{u} + \overrightarrow{v}\|^2 = \|\vec{u}\|^2 + \|\overrightarrow{v}\|^2 + 2(\ \vec{u} \cdot \overrightarrow{v}\ )$

P_5) El cuadrado del módulo de la diferencia de dos vectores es igual al cuadrado del módulo del primer vector más el cuadrado del módulo del segundo vector menos el doble de su producto escalar.

$\forall\, \overrightarrow{u},\ \overrightarrow{v} \in\ R^n$: $\|\vec{u} - \overrightarrow{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2(\overrightarrow{u} \cdot \overrightarrow{v}\ )$

D) Esta igualdad se demuestra de manera similar a la anterior, considerando la resta como suma entre el minuendo y el opuesto del sustraendo.

El concepto de módulo de un vector es útil para definir y poder calcular la distancia entre dos puntos de R^n.

Distancia entre dos puntos: *Se define la distancia entre dos puntos* $P = (p_1, p_2, \dots, p_n)$ *y* $Q = (q_1, q_2, \dots, q_n)$ *de* R^n *como la longitud del vector* $\overrightarrow{OQ} - \overrightarrow{OP}$

Dados $P = (p_1, p_2, \dots, p_n)$ *y* $Q = (q_1, q_2, \dots, q_n) \in R^n \implies$

$$dist(P, Q) = \|\overrightarrow{OQ} - \overrightarrow{OP}\| = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

Debemos recordar que el vector $\overrightarrow{OP}$ es el vector posición del punto P, lo mismo ocurre con el $\overrightarrow{OQ}$ y el punto Q

Ejemplo 1: calcular la distancia entre los puntos $P(-3, 2)$ *y* $Q(-1, -4)$

$dist(P, Q) = \|\overrightarrow{OQ} - \overrightarrow{OP}\| = \sqrt{(-1 + 3)^2 + (-4 - 2)^2} = \sqrt{38}$. Figura 1.20.

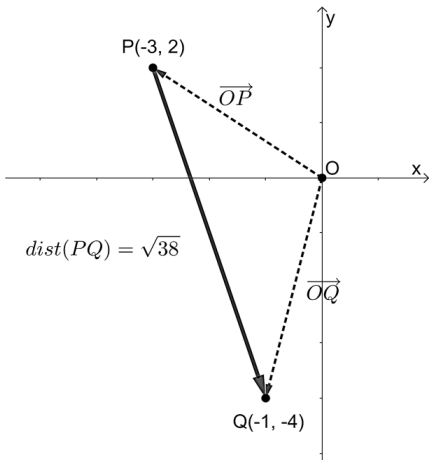


Figura 1.20. Distancia entre dos puntos en R^2.

Ejemplo 2: calcular la distancia entre los puntos $R(1, -2, 2)$ *y* $S(3, 2, 2)$

$dist(R, S) = \|\overrightarrow{OS} - \overrightarrow{OR}\| = \sqrt{(3 - 1)^2 + (2 + 2)^2 + (2 - 2)^2} = \sqrt{20}$. Esto se muestra en la Figura 1.21.

Vector unitario según una dirección dada: *El producto de un vector no nulo por el recíproco de su módulo o lo que es lo mismo, el cociente entre un vector no nulo y su módulo, es un vector unitario (módulo 1).*

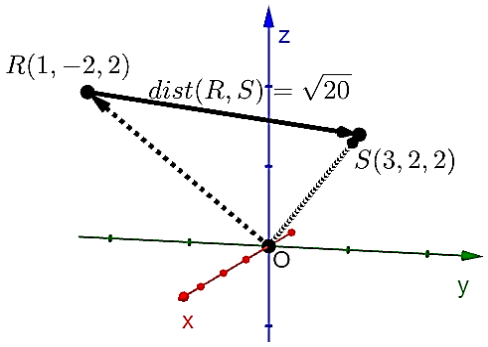


Figura 1.21. Distancia entre dos puntos en R^3.

Simbólicamente se expresa un vector unitario de la siguiente manera: sea $\vec{x} \in R^n \ y \ \vec{x} \neq \vec{0}$, si consideramos $\frac{\vec{x}}{\|\vec{x}\|}$ se verifica

$$\left\| \frac{\vec{x}}{\|\vec{x}\|} \right\| = \left\| \frac{1}{\|\vec{x}\|} \vec{x} \right\| = \left| \frac{1}{\|\vec{x}\|} \right| \|\vec{x}\| = \frac{1}{\|\vec{x}\|} \|\vec{x}\| = 1$$

Ejemplo: determinar un vector $\vec{y}$ de módulo 6 en la dirección de $\vec{x} = (1, -2, 2) \in R^3$.

Como $\vec{x} = (1, -2, 2) \implies \|\vec{x}\| = 3$ y por lo tanto $\frac{\vec{x}}{\|\vec{x}\|} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$ es un vector unitario. A partir de allí solo resta multiplicar este vector por el escalar 6 para que su módulo sea también 6. Por lo tanto:

$\vec{y} = 6 \frac{\vec{x}}{\|\vec{x}\|} \iff \vec{y} = 6\left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right) \iff \vec{y} = (2, -4, 4)$. Figura 1.22.

Gráficamente, se puede comprobar que ambos vectores, $\vec{x}$ e $\vec{y}$ tienen la misma dirección y al calcular el módulo de $\vec{y}$ resulta $\|\vec{y}\| = \sqrt{2^2 + (-4)^2 + 4^2} = \sqrt{36} = 6$.

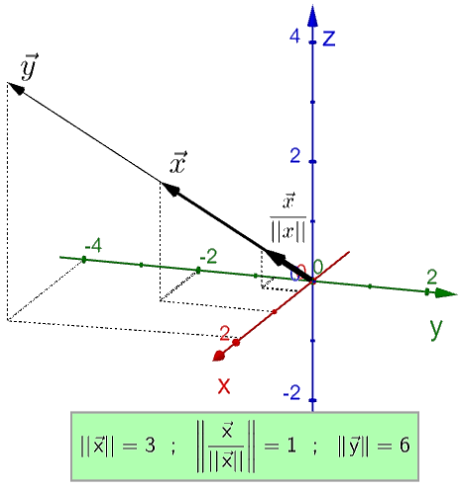


Figura 1.22. Vectores paralelos en R^3.

Veamos ahora cómo el producto escalar también nos permite determinar el ángulo entre vectores no nulos de R^n, para lo cual en primera instancia definimos ángulo entre vectores y luego aplicamos el Teorema de los cosenos[5].

[5] Este teorema plantea que, en un triángulo cualquiera, el cuadrado de la longitud de un lado es igual a la suma de los cuadrados de las longitudes de los lados restantes menos el duplo del producto de dichas longitudes multiplicado por el coseno del ángulo opuesto al lado en cuestión.

Ángulo entre vectores: *Dados dos vectores no nulos* $\vec{u}$ *y* $\vec{v}$ *de* R^n *se define el ángulo entre estos dos vectores como el menor ángulo positivo* φ $(0 \leq \varphi \leq \pi)$ *que gira uno de ellos para coincidir con la dirección del otro.* Figura 1.23.

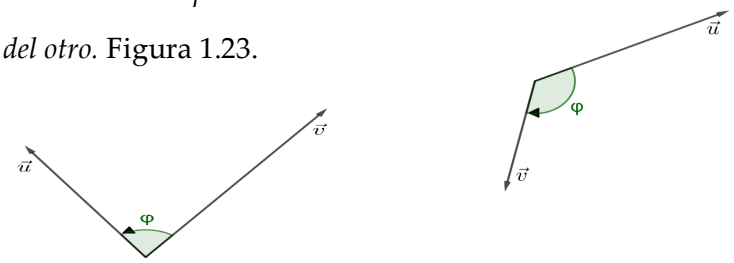


Figura 1.23. Ángulos entre dos vectores.

Ahora bien, dados dos vectores $\vec{u}$ *y* $\vec{v}$ *de* R^n no nulos, siempre podemos construir un triángulo en donde el tercer lado es la diferencia de los vectores dados, tal como lo muestra la Figura 1.24.

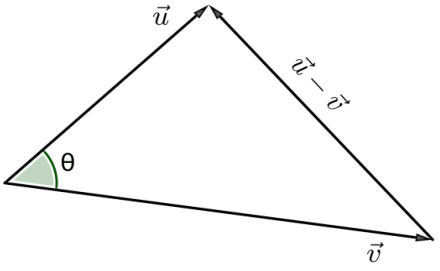


Figura 1.24. Construcción de triángulos por tres vectores.

Al aplicar el Teorema de los cosenos a este triángulo, tenemos:

$$\|\vec{u} - \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\,\|\vec{v}\|\cos\varphi$$

Aplicando propiedades del valor absoluto en el primer miembro, tenemos:

$$\|\vec{u}\|^2 + \|\vec{v}\|^2 - 2(\vec{u}\cdot\vec{v}) = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\,\|\vec{v}\|\cos\varphi$$

Suprimiendo términos semejantes, obtenemos:

$\vec{u} \cdot \vec{v} = \|\vec{u}\| \, \|\vec{v}\| \, cos \, \varphi$

Podemos establecer dos conclusiones a partir de esta última expresión:

a) En concordancia con lo que estamos estudiando, esta igualdad nos brinda una forma de calcular la medida del ángulo entre dos vectores no nulos de R^n, tomando a φ como el ángulo tal que: $cos \, \varphi = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|}$

b) Esta igualdad también es la expresión del producto escalar en función del ángulo que forman los vectores y de los módulos de estos. Es decir, se puede *definir* el producto escalar de dos vectores como el producto de sus módulos por el coseno del ángulo que forman dichos vectores.

Esta última observación nos permite hacer la interpretación geométrica del producto escalar.

Interpretación geométrica del producto escalar: el producto escalar de dos vectores no nulos es igual al módulo de uno de ellos por la proyección del otro sobre él.

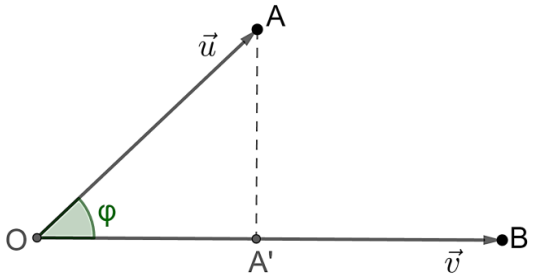


Figura 1.25. Interpretación geométrica del producto escalar.

En la Figura 1.25 se representan dos vectores $\overrightarrow{u}\ y\ \overrightarrow{v}$. Al proyectar [6] el vector $\overrightarrow{u}$ sobre la dirección del vector $\overrightarrow{v}$ se obtiene el vector $\overrightarrow{OA'}$ cuyo módulo coincide con la medida del segmento $\underline{OA'}$. A la longitud de este segmento se la suele llamar proyección escalar de $\overrightarrow{u}$ sobre $\overrightarrow{v}$, es decir, $\underline{OA'} = |proy\ de\ \overrightarrow{u}\ sobre\ \overrightarrow{v}| = |\vec{u}_{\vec{v}}|$

Teniendo en cuenta que $cos\ \varphi = \frac{OA\prime}{\|\vec{u}\|} \implies \underline{OA'} = \|\vec{u}\| cos\ \varphi$, por lo tanto

$\overrightarrow{u} \cdot \overrightarrow{v} = \|\vec{u}\|\ \|\vec{v}\|\ cos\ \varphi$

$\overrightarrow{u} \cdot \overrightarrow{v} = \|\vec{v}\|\ |\vec{u}_{\vec{v}}|$ que es la interpretación geométrica enunciada.

De esta última expresión se puede despejar y calcular la proyección escalar de un vector sobre otro: $|\vec{u}_{\vec{v}}| = \frac{\overrightarrow{u} \cdot \overrightarrow{v}}{\|\vec{v}\|}$

Ejemplo 1: calcular el ángulo entre los vectores $\overrightarrow{u} = (-2, 4)\ y\ \overrightarrow{v} = (3, -1) \in R^2$.

Si llamamos φ al ángulo entre los vectores, tenemos: $cos\ \varphi = \frac{\overrightarrow{u} \cdot \overrightarrow{v}}{\|\overrightarrow{u}\|\ \|\overrightarrow{v}\|}$

$\overrightarrow{u} \cdot \overrightarrow{v} = -10,\ \|\overrightarrow{u}\| = \sqrt{20}\ \ y\ \ \|\overrightarrow{v}\| = \sqrt{10}$

$cos\ \varphi = \frac{\overrightarrow{u} \cdot \overrightarrow{v}}{\|\overrightarrow{u}\|\ \|\overrightarrow{v}\|} \Longrightarrow \varphi = arcos \frac{\overrightarrow{u} \cdot \overrightarrow{v}}{\|\overrightarrow{u}\|\ \|\overrightarrow{v}\|}$

$\varphi = arcos \frac{-10}{\sqrt{20}\sqrt{10}}$

$\varphi = 135°.$

Los vectores y el ángulo que forman se muestran en la Figura 1.26.

[6] Significa trazar, por el extremo de uno de los vectores, una perpendicular a la dirección del otro, hasta que la intercepte esta.

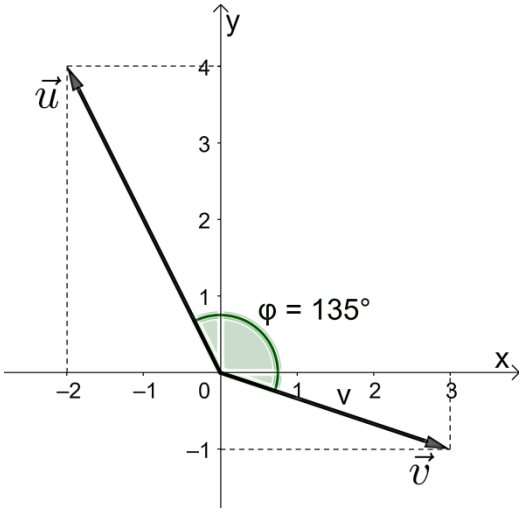


Figura 1.26. Ángulo entre dos vectores en R^2.

Ejemplo 2: calcular el ángulo, φ, los vectores $\vec{v} = (-2, 4, 1)$ y $\vec{u} = (3, -1, 5)$

$\vec{v} \cdot \vec{u} = -2\,.3 + 4(-1) + 1\,.5 = -5$

$\|\vec{v}\| = \sqrt{(-2)^2 + 4^2 + 1^2} = \sqrt{21}$

$\|\vec{u}\| = \sqrt{3^2 + (-1)^2 + 5^2} = \sqrt{35}$

Por lo tanto $\varphi = arcos\frac{-5}{\sqrt{21}\sqrt{35}} \cong 100° 37' 40''$, como lo muestra la Figura 1.27.

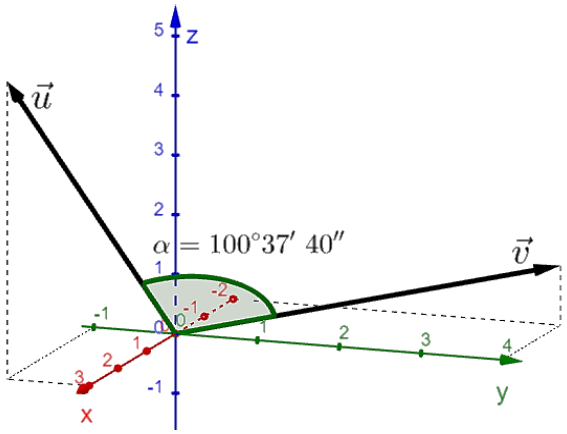


Figura 1.27. Ángulo entre vectores en R^3.

Ejemplo 3: hallar la proyección escalar del vector $\vec{u} = (2, 1)$ sobre el vector $\vec{v} = (-3, 4)$

$|\vec{u}_{\vec{v}}| = \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|}$

$\vec{u} \cdot \vec{v} = -6 + 4 = -2$

$\|\vec{v}\| = 5$

$|\vec{u}_{\vec{v}}| = \frac{2}{5}$

La ventaja de considerar el producto escalar de dos vectores no nulos, $\vec{u}\ y\ \vec{v}\ de\ R^n$ como $\vec{u} \cdot \vec{v} = \|\vec{u}\|\ \|\vec{v}\|\ cos\ \varphi$, es que se puede demostrar con facilidad la condición necesaria y suficiente para que dos vectores no nulos sean perpendiculares.

La condición necesaria y suficiente para que dos vectores no nulos $\vec{u}\ y\ \vec{v}\ de\ R^n$ *sean perpendiculares es que el producto escalar de los mismos sea nulo.*

Sean los vectores $\vec{u}\ y\ \vec{v}\ no\ nulos, \vec{u}\ \perp\ \vec{v}\ \Leftrightarrow\ \vec{u} \cdot \vec{v} = 0$

D)

$\Longrightarrow$) Si $\overrightarrow{u} \perp \overrightarrow{v} \Rightarrow \overrightarrow{u} \cdot \overrightarrow{v} = 0$

$\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \, \|\overrightarrow{v}\| cos\, \varphi$ siendo φ el ángulo que forman los vectores. Como por hipótesis se tiene que $\overrightarrow{u} \perp \overrightarrow{v} \Rightarrow \varphi = 90°$

$\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \, \|\overrightarrow{v}\| cos\, 90°$

$\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \, \|\overrightarrow{v}\| \, 0$

$\overrightarrow{u} \cdot \overrightarrow{v} = 0$

$\Longleftarrow$) Si $\overrightarrow{u} \cdot \overrightarrow{v} = 0 \Rightarrow \overrightarrow{u} \perp \overrightarrow{v}$

Si el producto escalar es 0 y los vectores son no nulos, necesariamente debe ser $\varphi = 90°$, o sea el ángulo que forman los vectores es 90°, es decir que $\overrightarrow{u} \perp \overrightarrow{v}$, quedando demostrada de esta manera la condición necesaria y suficiente.

Aplicaciones del producto escalar: en física, se define el trabajo (W) realizado por una fuerza F que actúa sobre un objeto como el producto de la fuerza por la distancia que recorre su punto de aplicación en dirección de la fuerza. Si la fuerza es constante y el desplazamiento es rectilíneo, entonces el trabajo de la fuerza es el producto escalar entre la fuerza aplicada y el vector desplazamiento $\left(\overrightarrow{\Delta r}\right)$, es decir:

$W = \overrightarrow{F} \cdot \overrightarrow{\Delta r}$ *o* $W = \|\overrightarrow{F}\| \cdot \|\overrightarrow{\Delta r}\| cos\, \varphi$ (1). Tal como se indica en la Figura 1.28.

Las ecuaciones (1) indican que el trabajo realizado por una fuerza es una magnitud escalar.

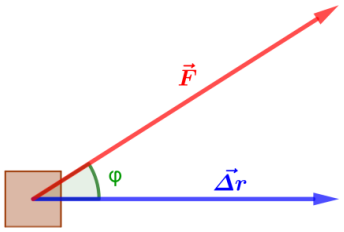


Figura 1.28. Trabajo mecánico.

Analizando la Figura 1.28 podemos concluir que si el ángulo $\hat{\varphi}$ entre los vectores es agudo, el trabajo será positivo, si $\hat{\varphi} = 90°$, el trabajo será nulo; y si es obtuso, el trabajo de la fuerza será negativo.

Como se dijo anteriormente, el producto escalar no es el único producto que se puede efectuar entre vectores, existe también el producto vectorial, que se diferencia del anterior porque en este caso se obtiene, por resultado, otro vector.

Cabe destacar que el producto vectorial tiene diversas aplicaciones y solo es posible en el espacio de tres dimensiones (R^3).

En primer lugar, veremos que todo vector de R^3 puede ser expresado, según tres direcciones dadas, como la suma de productos de un escalar por un vector unitario en las direcciones establecidas.

Particularmente, en un sistema de coordenadas cartesianas ortogonales, se representan los vectores $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$, tal como se muestra en la Figura 1.29.

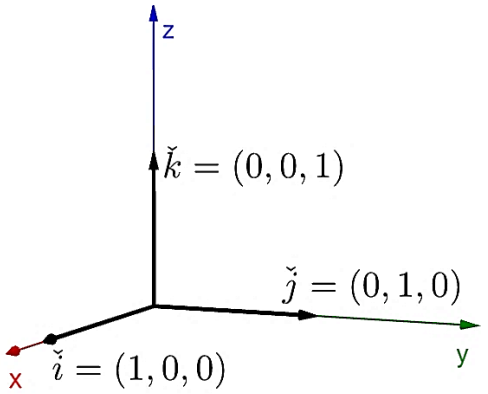


Figura 1.29. Versores fundamentales de R^3.

4.2. Versores fundamentales[7]

El vector $(1, 0, 0)$ *al que se denomina* $\check{\imath}$ *tiene la dirección del eje* $\overrightarrow{ox}$ *su sentido corresponde al sentido positivo de dicho eje y su módulo es 1. Por lo tanto* $\check{\imath} = (1, 0, 0)$ *es el versor fundamental en la dirección del eje* $\overrightarrow{ox}$.

De manera similar:

$\check{\jmath} = (0, 1, 0)$ *es el versor fundamental en la dirección del eje* $\overrightarrow{oy}$

$\check{k} = (0, 0, 1)$ *es el versor fundamental en la dirección del eje* $\overrightarrow{oz}$

A partir de esto, todo vector $\overrightarrow{u} = (u_1, u_2, u_3)$ *de* R^3 puede ser expresado como la suma de productos de un escalar por un versor, ya que:

$\overrightarrow{u} = (u_1, u_2, u_3)$

$\overrightarrow{u} = (u_1, 0, 0\ ) + (0, u_2, 0\ ) + (0, 0, u_3\ )$

$\overrightarrow{u} = u_1(1, 0, 0\ ) + u_2(0{,}1\ , 0\ ) + u_3(0, 0, 1)$

$\overrightarrow{u} = u_1\check{\imath} + u_2\check{\jmath} + u_3\check{k}$

[7] También denominados, vectores canónicos.

Ejemplo: el vector $\vec{u} = (2, 3, 1)$ *de* R^3 puede ser expresado como $\vec{u} = 2\check{\imath} + 3\check{\jmath} + \check{k}$ es decir dos veces el versor en la dirección x, tres veces el versor en la dirección y, finalmente, una vez el versor en la dirección z. Figura 1.30.

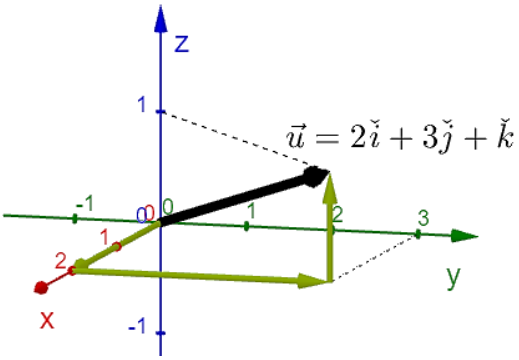


Figura 1.30. Vector $\vec{u}$.

Corresponde ahora definir la nueva operación entre vectores que se realiza con exclusividad en R^3: el producto vectorial.

4.3. Producto vectorial[8]

Dados dos vectores $\vec{u} = (u_1, u_2, u_3)$ *y* $\vec{v} = (v_1, v_2, v_3)$ *de* R^3 *se define producto vectorial entre esos dos vectores como el vector* $\vec{u} \times \vec{v}$ *definido por:*

$\vec{u} \times \vec{v} = (u_2v_3 - u_3v_2)\check{\imath} + (u_3v_1 - u_1v_3)\check{\jmath} + (u_1v_2 - u_2v_1)\check{k}$.

Observemos que el vector anterior se puede escribir simbólicamente en forma de determinante como sigue:

$$\vec{u} \times \vec{v} = \begin{vmatrix} \check{\imath} & \check{\jmath} & \check{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

[8] Esta operación también se conoce como Producto cruz.

Al desarrollar este determinante se obtiene el vector anterior, destacamos que el determinante es simbólico ya que su desarrollo es un vector y no un escalar.

Veamos qué módulo, dirección y sentido tiene el nuevo vector definido mediante esta operación.

Módulo del producto vectorial: *El módulo del producto vectorial de dos vectores cualesquiera* $\overrightarrow{u} = (u_1, u_2, u_3)$ y $\overrightarrow{v} = (v_1, v_2, v_3)$ $de\ R^3$ *es el producto de sus módulos por el seno del ángulo que forman.*

Sean $\overrightarrow{u} = (u_1, u_2, u_3)$ y $\overrightarrow{v} = (v_1, v_2, v_3) \in R^3 \Longrightarrow \|\overrightarrow{u}\ x\ \overrightarrow{v}\| = \|\overrightarrow{u}\|\ \|\overrightarrow{v}\|\ sen\ \varphi$ *siendo* φ *el ángulo entre* $\overrightarrow{u}$ y $\overrightarrow{v}$

D) Sabemos que por propiedades del módulo de un vector:

$\|\overrightarrow{u} \times \overrightarrow{v}\|^2 = (\overrightarrow{u} \times \overrightarrow{v}) \cdot (\overrightarrow{u} \times \overrightarrow{v})$

$= [(u_2v_3 - u_3v_2), (u_3v_1 - u_1v_3), (u_1v_2 - u_2v_1)] \cdot [(u_2v_3 - u_3v_2), (u_3v_1 - u_1v_3), (u_1v_2 - u_2v_1)]$

$= (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2$

$= u_2^2v_3^2 + u_3^2v_2^2 + u_3^2v_1^2 + u_1^2v_3^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2v_3u_3v_2 - 2u_3v_1u_1v_3 - 2u_1v_2u_2v_1$

$= u_1^2v_1^2 + u_1^2v_2^2 + u_1^2v_3^2 + u_2^2v_1^2 + u_2^2v_2^2 + u_2^2v_3^2 + u_3^2v_1^2 + u_3^2v_2^2 + u_3^2v_3^2 - u_1^2v_1^2 - u_2^2v_2^2 - u_3^2v_3^2 - 2u_2v_3u_3v_2 - 2u_3v_1u_1v_3 - 2u_1v_2u_2v_1$

$= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1^2v_1^2 + u_2^2v_2^2 + u_3^2v_3^2 + 2u_1v_1u_2v_2 + 2u_1v_1u_3v_3 + 2u_2v_2u_3v_3)$

$= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2$

$= \|\overrightarrow{u}\|^2\ \|\overrightarrow{v}\|^2 - (\overrightarrow{u} \cdot \overrightarrow{v})^2$

Sabemos que $\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\|\ \|\overrightarrow{v}\| cos\ \varphi$, entonces:

$\|\overrightarrow{u}\|^2\ \|\overrightarrow{v}\|^2 - (\overrightarrow{u} \cdot \overrightarrow{v})^2 = \|\overrightarrow{u}\|^2\ \|\overrightarrow{v}\|^2 - (\|\overrightarrow{u}\|\ \|\overrightarrow{v}\|\ cos\ \varphi)^2$

$= \|\overrightarrow{u}\|^2 \, \|\overrightarrow{v}\|^2 - \|\overrightarrow{u}\|^2\|\overrightarrow{v}\|^2 \cos^2 \varphi$

$= \|\overrightarrow{u}\|^2 \, \|\overrightarrow{v}\|^2(1 - \cos^2 \varphi)$

$\|\overrightarrow{u} \, x \, \overrightarrow{v}\|^2 = \|\overrightarrow{u}\|^2 \, \|\overrightarrow{v}\|^2 sen^2 \varphi$

$\|\overrightarrow{u} \, x \, \overrightarrow{v}\| = \|\overrightarrow{u}\| \, \|\overrightarrow{v}\| \, sen \, \varphi$

De esta última expresión, se obtiene una nueva definición de vectores paralelos en R^3, que se expresa en el siguiente corolario:

Vectores paralelos: *Dos vectores no nulos* $\overrightarrow{u} = (u_1, u_2, \, u_3) \; y \; \overrightarrow{v} = (v_1, v_2, v_3) \; de \; R^3$ *se dicen paralelos si y solo si su producto vectorial es nulo.*

Dados $\overrightarrow{u} = (u_1, u_2, \, u_3) \; y \; \overrightarrow{v} = (v_1, v_2, v_3) \; de \; R^3$, $\overrightarrow{u} \; // \; \overrightarrow{v} \iff \overrightarrow{u} \, x \, \overrightarrow{v} = 0$

D)

$\Longrightarrow$) Si $\overrightarrow{u} \; // \; \overrightarrow{v}$ entonces $\overrightarrow{v} = \lambda \, \overrightarrow{u}$ por lo tanto

$\overrightarrow{u} \; \times \; \overrightarrow{v} = \overrightarrow{u} \; \times \; (\lambda \, \overrightarrow{u} \,)$

$\overrightarrow{u} \; \times \; \overrightarrow{v} = \lambda (\, \overrightarrow{u} \; \times \; \overrightarrow{u} \,)$ por propiedades del producto vectorial

$\overrightarrow{u} \; \times \; \overrightarrow{v} = \lambda \, .0$ por propiedades del producto vectorial[9]

$\overrightarrow{u} \; \times \; \overrightarrow{v} = 0$

$\Longleftarrow$) $\overrightarrow{u} \; \times \; \overrightarrow{v} = 0 \implies \overrightarrow{u} \; // \; \overrightarrow{v}$

$\overrightarrow{u} \; \times \; \overrightarrow{v} = 0 \implies \|\overrightarrow{u} \; \times \; \overrightarrow{v}\| = 0$ pero

$\|\overrightarrow{u} \; \times \; \overrightarrow{v}\| = \|\overrightarrow{u}\| \, \|\overrightarrow{v}\| \, sen \, \varphi$

Como los vectores $\overrightarrow{u} \; y \; \overrightarrow{v}$ son no nulos entonces $sen \, \varphi = 0$

Si $sen \, \varphi = 0$ entonces $\varphi = 0°$ ó $\varphi = 180°$ lo que implica que los vectores $\overrightarrow{u} \; y \; \overrightarrow{v}$ son paralelos.

Dirección del producto vectorial: *La dirección del vector* $\overrightarrow{u} \; \times \; \overrightarrow{v}$ *es la de la recta perpendicular al plano determinado por los vectores* $\overrightarrow{u} \; y \; \overrightarrow{v}$

[9] Estas propiedades se demostrarán en el apartado correspondiente.

Sean $\overrightarrow{u} = (u_1, u_2, u_3)\ y\ \overrightarrow{v} = (v_1, v_2, v_3)\ \in\ R^3 \Longrightarrow (\overrightarrow{u}\ \times\ \overrightarrow{v}) \perp \overrightarrow{u}\ \wedge\ (\overrightarrow{u}\ \times\ \overrightarrow{v}) \perp \overrightarrow{v}$

D) Para que dos vectores sean perpendiculares, su producto escalar debe ser nulo. Multipliquemos escalarmente $\overrightarrow{u}\ \times\ \overrightarrow{v}$ con $\overrightarrow{u}$.

$(\overrightarrow{u}\ \times\ \overrightarrow{v}) \cdot \overrightarrow{u} = [(u_2v_3 - u_3v_2), (u_3v_1 - u_1v_3), (u_1v_2 - u_2v_1)] \cdot (u_1, u_2, u_3)$

$(\overrightarrow{u}\ \times\ \overrightarrow{v}) \cdot \overrightarrow{u} = (u_2v_3 - u_3v_2)u_1 + (u_3v_1 - u_1v_3)u_2 +\ (u_1v_2 - u_2v_1)u_3$

$(\overrightarrow{u}\ \times\ \overrightarrow{v}) \cdot \overrightarrow{u} = u_1u_2v_3 - u_1u_3v_2 + u_2u_3v_1 - u_1u_2v_3 + u_1u_3v_2 - u_2u_3v_1$

$(\overrightarrow{u}\ \times\ \overrightarrow{v}) \cdot \overrightarrow{u} = 0$

Este resultado indica que $(\overrightarrow{u}\ \times\ \overrightarrow{v}) \perp \overrightarrow{u}$.

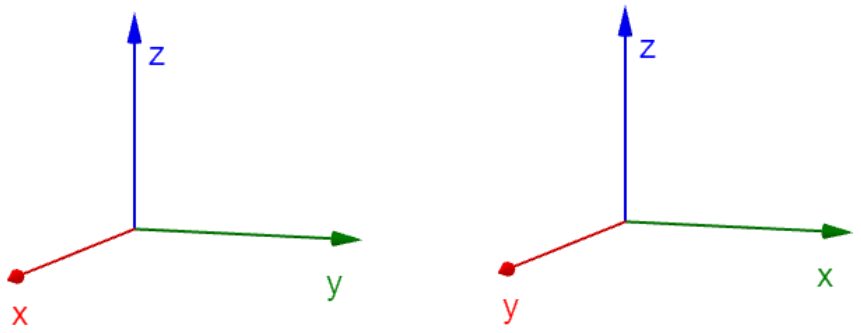
De manera similar se puede demostrar que $(\overrightarrow{u}\ \times\ \overrightarrow{v}) \perp \overrightarrow{v}$.

Por lo tanto $(\overrightarrow{u}\ \times\ \overrightarrow{v}) \perp \overrightarrow{u}\ \wedge\ (\overrightarrow{u}\ \times\ \overrightarrow{v}) \perp \overrightarrow{v}$, lo que implica que $\overrightarrow{u}\ \times\ \overrightarrow{v}$ es perpendicular al plano determinado por los vectores $\overrightarrow{u}\ y\ \overrightarrow{v}$.

Sentido del producto vectorial: antes de poder determinar el sentido del producto vectorial, tengamos en cuenta un concepto previo: el de triedros[10].

Podemos observar, en las figuras 1.31a) y 1.31b), que no existe rotación alguna de uno de los triedros que hagan coincidir todos los ejes del mismo nombre –inclusive su dirección– con el otro triedro.

[10] Figura geométrica formada por tres semirrectas que parten del mismo origen, pero que no están situadas en un mismo plano.



a) b)

Figura 1.31. Triedros no equivalentes.

En efecto, si se hacen coincidir los orígenes y los ejes x e y de modo que se superpongan las partes positivas de ambos triedros, los sentidos de los ejes z resultan opuestos. Se dice que los triedros tienen distinta *orientación*.

Para distinguir las orientaciones, consideremos un triedro como el de la Figura 1.31a), e imaginemos que la parte positiva del eje x gira hacia la parte positiva del eje y. Un tornillo colocado perpendicularmente en el plano XY, que gira del mismo modo, avanza hacia la parte positiva del eje z. Se dice que este triedro es positivo o directo. En caso contrario, el triedro de la figura 1.31b), se dice que es negativo o inverso.

Existen vectores en cuya definición interviene la orientación del espacio, de tal manera que al cambiar ésta, cambia el sentido del vector, es decir que estos vectores no quedan definidos independientemente del sistema

de coordenadas al que está referido el espacio. Por tal motivo se los denomina *pseudovectores*. Esto ocurre con el producto vectorial.

El sentido del producto vectorial $\vec{u} \times \vec{v}$ es tal que el triedro ($\vec{u}, \vec{v}, \vec{u} \times \vec{v}$) sea positivo o directo, si es directo el sistema de coordenadas elegido. Figura 1.32.

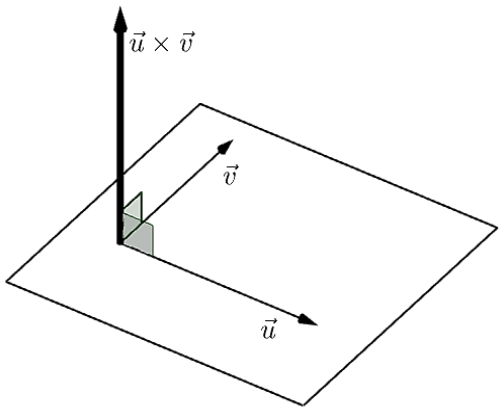
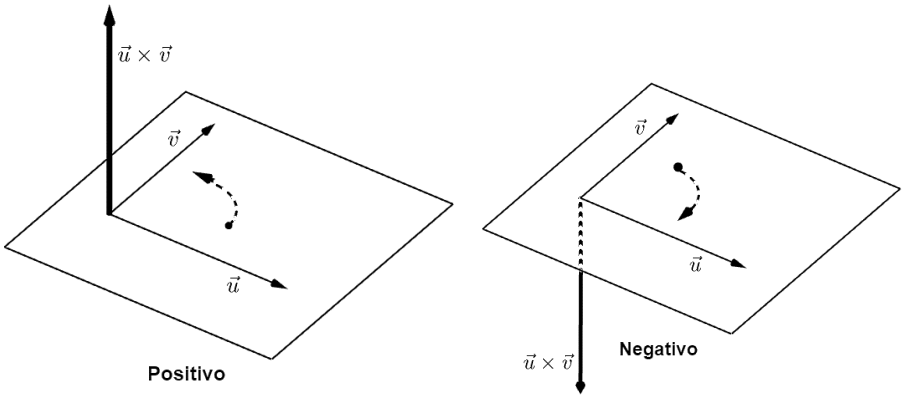


Figura 1.32. Triedro positivo o directo.

Conceptualmente podemos relacionar el sentido del producto vectorial con el avance de un tornillo que se mantiene perpendicular al plano XY, tal como se muestran en las Figuras 1.33a) y 1.33b).



a) b)

Figura 1.33. Producto vectorial.

Ejemplo 1.

a) Calcular el producto vectorial de $\overrightarrow{u} = (3, -1, 1)$ y $\overrightarrow{v} = (1, 1, 1)$

$$\overrightarrow{u} \times \overrightarrow{v} = \begin{vmatrix} i & j & k \\ 3 & -1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -2i - 2j + 4k = (-2, -2, 4)$$

b) Comprobar que el vector $\overrightarrow{u}\, x\, \overrightarrow{v}$ es perpendicular a los vectores dados.

$(-2, -2, 4) \cdot (3, -1, 1) = -6 + 2 + 4 = 0$

$(-2, -2, 4) \cdot (1, 1, 1) = -2 - 2 + 4 = 0.$

Todo lo realizado se muestra en la Figura 1.34.

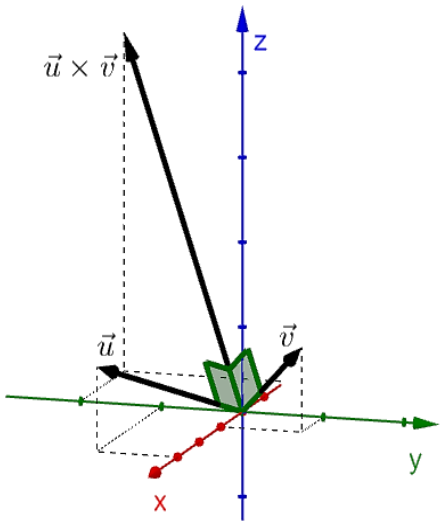


Figura 1.34. Producto $\vec{u} \times \vec{v}$.

c) Hallar un vector unitario perpendicular a $\vec{u}\ y\ \vec{v}$

El vector $\vec{u}\ x\ \vec{v}$ es perpendicular a $\vec{u}\ y\ \vec{v}$ pero no es unitario. Por lo tanto:

$\frac{(-14,8,-9)}{\sqrt{341}} = \left(-\frac{14}{\sqrt{341}}, \frac{8}{\sqrt{341}}, -\frac{9}{\sqrt{341}}\right)$ que es el vector solicitado.

Ejemplo 2: realizar los siguientes productos vectoriales: $\check{\imath} \times \check{\jmath}$; $\check{\jmath} \times \check{k}$; $\check{k} \times \check{\imath}$;

$$\check{\imath} \times \check{\jmath} = \begin{vmatrix} \check{\imath} & \check{\jmath} & \check{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = 0\check{\imath} + 0\check{\jmath} + \check{k} = (0,0,1) = \check{k}$$

$$\check{\jmath} \times \check{k} = \begin{vmatrix} \check{\imath} & \check{\jmath} & \check{k} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \check{\imath} + 0\check{\jmath} + 0\check{k} = (1,0,0) = \check{\imath}$$

$$\check{k} \times \check{\imath} = \begin{vmatrix} \check{\imath} & \check{\jmath} & \check{k} \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = 0\check{\imath} + \check{\jmath} + 0\check{k} = (0,1,0) = \check{\jmath}$$

De manera similar se puede demostrar que:

$$\check{\imath} \times \check{\imath} = \check{\jmath} \times \check{\jmath} = \check{k} \times \check{k} = \vec{0}$$

Finalmente:

$\check{\imath} \times \check{\imath} = \check{\jmath} \times \check{\jmath} = \check{k} \times \check{k} = \overrightarrow{0}$

$\check{\imath} \times \check{\jmath} = \check{k}$ $\qquad\qquad$ $\check{\jmath} \times \check{\imath} = -\check{k}$

$\check{\jmath} \times \check{k} = \check{\imath}$ $\qquad\qquad$ $\check{k} \times \check{\jmath} = -\check{\imath}$

$\check{k}\, x\, \check{\imath} = \check{\jmath}$ $\qquad\qquad$ $\check{\imath} \times \check{k} = -\check{\jmath}$

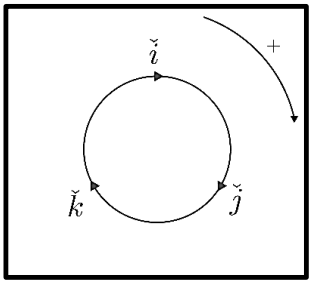


Figura 1.35. Producto vectorial de versores fundamentales.

De donde debemos notar que:

El producto vectorial de cualquier vector canónico consigo mismo da como resultado el vector nulo (en las propiedades del producto vectorial esto se generaliza para cualquier vector de R^3).

El producto vectorial de dos vectores canónicos consecutivos en el sentido de las agujas del reloj, como se indica en el esquema en la parte superior derecha, tiene por resultado el siguiente vector canónico y cuando se considera el sentido contrario de las agujas del reloj, el resultado es el siguiente vector canónico con signo negativo.

Propiedades del producto vectorial

P1) El producto vectorial es anticonmutativo

$\forall\, \overrightarrow{u},\ \overrightarrow{v} \in R^3 : \overrightarrow{u} \times \overrightarrow{v} = -\overrightarrow{v} \times \overrightarrow{u}$

$\overrightarrow{u} \times \overrightarrow{v} = (u_2v_3 - u_3v_2)\check{\imath} + (u_3v_1 - u_1v_3)\check{\jmath} + (u_1v_2 - u_2v_1)\check{k}.$

$\overrightarrow{u} \times \overrightarrow{v} = -(u_3v_2 - u_2v_3)\check{\imath} - (u_1v_3 - u_3v_1)\check{\jmath} - (u_2v_1 - u_1v_2)\check{k}.$

$\overrightarrow{u} \times \overrightarrow{v} = -\left[(u_3v_2 - u_2v_3)\check{\imath} + (u_1v_3 - u_3v_1)\check{\jmath} + (u_2v_1 - u_1v_2)\check{k}\right].$

$\overrightarrow{u} \times \overrightarrow{v} = -\overrightarrow{v} \times \overrightarrow{u}$

P2) El producto vectorial es distributivo a derecha y a izquierda, respecto de la suma de vectores

$\forall\, \overrightarrow{u},\ \overrightarrow{v}, \overrightarrow{w} \in R^3 : \overrightarrow{u} \times (\overrightarrow{v} + \overrightarrow{w}) = \overrightarrow{u} \times \overrightarrow{v} + \overrightarrow{u} \times \overrightarrow{w}$

$\overrightarrow{u} \times (\overrightarrow{v} + \overrightarrow{w}) = \left[\big(u_2(v_3 + w_3) - u_3(v_2 + w_2)\big)\check{\imath} + \big(u_3(v_1 + w_1) - u_1(v_3 + w_3)\big)\check{\jmath} + \big(u_1(v_2 + w_2) - u_2(v_1 + w_1)\big)\check{k}\right]$

$\overrightarrow{u} \times (\overrightarrow{v} + \overrightarrow{w}) = \left[\big((u_2v_3 + u_2w_3) - (u_3v_2 + u_3w_2)\big)\check{\imath} + \big((u_3v_1 + u_3w_1) - (u_1v_3 + u_1w_3)\big)\check{\jmath} + ((u_1v_2 + u_1w_2) - (u_2v_1 + u_2w_1))\check{k}\right]$

$\overrightarrow{u} \times (\overrightarrow{v} + \overrightarrow{w}) = \left[(u_2v_3 - u_3v_2)\check{\imath} + (u_3v_1 - u_1v_3)\check{\jmath} + (u_1v_2 - u_2v_1)\check{k}\right] + \left[(u_2w_3 - u_3w_2)\hat{\imath} + (u_3w_1 - u_1w_3)\hat{\jmath} + (u_1v_2 - u_2w_1)\check{k}\right]$

$\overrightarrow{u} \times (\overrightarrow{v} + \overrightarrow{w}) = \overrightarrow{u} \times \overrightarrow{v} + \overrightarrow{u} \times \overrightarrow{w}$

Para el caso de la distributividad, a derecha, respecto de la suma de vectores se procede de manera similar.

Una manera alternativa al procedimiento anterior y para la distributividad a derecha, es usar las propiedades ya estudiadas del producto vectorial.

$(\overrightarrow{v} + \overrightarrow{w}) \times \overrightarrow{u} = \overrightarrow{v} \times \overrightarrow{u} + \overrightarrow{w} \times \overrightarrow{u}$

$(\overrightarrow{v} + \overrightarrow{w}) \times \overrightarrow{u} = -[\overrightarrow{u} \times (\overrightarrow{v} + \overrightarrow{w})]$ por anticonmutatividad

$(\overrightarrow{v} + \overrightarrow{w}) \times \overrightarrow{u} = -[\overrightarrow{u} \times \overrightarrow{v} + \overrightarrow{u} \times \vec{w}]$ por distributividad a izquierda

$(\overrightarrow{v} + \overrightarrow{w}) \times \overrightarrow{u} = -\overrightarrow{u} \times \overrightarrow{v} - \overrightarrow{u} \times \vec{w}$

$(\overrightarrow{v} + \overrightarrow{w}) \times \overrightarrow{u} = \overrightarrow{v} \times \overrightarrow{u} + \overrightarrow{w} \times \overrightarrow{u}$ Por anticonmutatividad.

Quedando así demostrada la propiedad.

P3) Asociativa respecto al producto por un escalar.

$\forall\, \overrightarrow{u},\ \overrightarrow{v} \in\ R^3, \forall \alpha \in\ R: \alpha(\overrightarrow{u} \times \overrightarrow{v}) = (\alpha\overrightarrow{u}) \times \overrightarrow{v} = \overrightarrow{u} \times (\alpha\overrightarrow{v})$

$\overrightarrow{u} \times \overrightarrow{v} = (u_2v_3 - u_3v_2)\hat{\imath} + (u_3v_1 - u_1v_3)\check{\jmath} + (u_1v_2 - u_2v_1)\check{k}.$

$\alpha(\overrightarrow{u} \times \overrightarrow{v}) = \alpha\left[(u_2v_3 - u_3v_2)\check{\imath} + (u_3v_1 - u_1v_3)\check{\jmath} + (u_1v_2 - u_2v_1)\check{k}\right]$

$\alpha(\overrightarrow{u} \times \overrightarrow{v}) = \alpha(u_2v_3 - u_3v_2)\check{\imath} + \alpha(u_3v_1 - u_1v_3)\check{\jmath} + \alpha(u_1v_2 - u_2v_1)\check{k}$ (1)

$(\alpha\overrightarrow{u}) \times \overrightarrow{v} = (\alpha u_2v_3 - \alpha u_3v_2)\check{\imath} + (\alpha u_3v_1 - \alpha u_1v_3)\check{\jmath} + (\alpha u_1v_2 - \alpha u_2v_1)\check{k}$

$(\alpha\overrightarrow{u}) \times \overrightarrow{v} = \alpha(u_2v_3 - u_3v_2)\check{\imath} + \alpha(u_3v_1 - u_1v_3)\check{\jmath} + \alpha(u_1v_2 - u_2v_1)\check{k}$ (2)

$\overrightarrow{u} \times (\alpha\overrightarrow{v}) = (\alpha u_2v_3 - \alpha u_3v_2)\check{\imath} + (\alpha u_3v_1 - \alpha u_1v_3)\check{\jmath} + (\alpha u_1v_2 - \alpha u_2v_1)\check{k}$

$\overrightarrow{u} \times (\alpha\overrightarrow{v}) = \alpha(u_2v_3 - u_3v_2)\check{\imath} + \alpha(u_3v_1 - u_1v_3)\check{\jmath} + \alpha(u_1v_2 - u_2v_1)\check{k}$ (3)

Se puede observar que las expresiones (1), (2) y (3) son iguales, con lo que queda demostrada la propiedad.

P4) El producto vectorial a derecha e izquierda por el vector nulo, es el vector nulo.

$\forall\, \overrightarrow{u} \in\ R^3: \overrightarrow{u} \times \overrightarrow{0} = \overrightarrow{0} \times \overrightarrow{u} = \overrightarrow{0}$

$\overrightarrow{0} \times \overrightarrow{u} = (u_20 - u_30)\check{\imath} + (u_30 - u_10)\check{\jmath} + (u_10 - u_20)\check{k} = \overrightarrow{0}$ (1)

$\overrightarrow{u} \times \overrightarrow{0} = (u_20 - u_30)\check{\imath} + (u_30 - u_10)\check{\jmath} + (u_10 - u_20)\check{k} = \overrightarrow{0}$ (2)

Se puede observar que las expresiones (1) y (2) son iguales, con lo que queda demostrada la propiedad.

P5) El producto vectorial de cualquier vector por sí mismo es el vector nulo.

$\forall\, \overrightarrow{u} \in\ R^3: \overrightarrow{u} \times \overrightarrow{u} = \overrightarrow{0}$

$\overrightarrow{u} \times \overrightarrow{u} = (u_2u_3 - u_3u_2)\check{\imath} + (u_3u_1 - u_1u_3)\check{\jmath} + (u_1u_2 - u_2u_1)\check{k} = \overrightarrow{0}$

Interpretación geométrica del producto vectorial

El área del paralelogramo (A_P) cuyos lados no paralelos están dados por los vectores $\vec{u}\ y\ \vec{v}\ de\ R^3$, está dada por el módulo del producto vectorial de dichos vectores.

Dados $\vec{u}\ ,\ \vec{v}\ \in\ R^3\ \Longrightarrow A_P = \|\vec{u} \times \vec{v}\|$.

D) Se debe considerar el paralelogramo cuyos lados no paralelos son los vectores $\vec{u}\ ,\vec{v}\ de\ R^3\ y\ \varphi$ el ángulo entre ellos, como muestra la Figura 1.36.

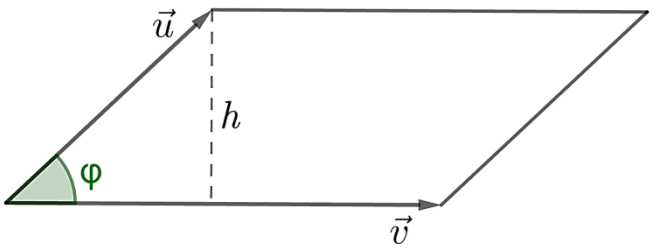


Figura 1.36. Paralelogramo definido por dos vectores.

$A_P = Base\ .\ altura$

$A_P = \ \|\vec{v}\|\ .\ h$

$sen\varphi = \frac{h}{\|\vec{u}\|} \Longrightarrow h = \|\vec{u}\|\ sen\varphi$ Por lo tanto

$A_P = \ \|\vec{v}\|\ .\ \|\vec{u}\|\ sen\varphi$

$A_P = \ \|\vec{u} \times \vec{v}\|$

A su vez, este resultado permite calcular el área del triángulo determinado por los vectores $\vec{u}\ y\ \vec{v}$.

El área del triángulo (A_T) determinado por los vectores $\vec{u}\ y\ \vec{v}\ de\ R^3$ es igual a la mitad del módulo del producto vectorial de dichos vectores. Figura 1.37.

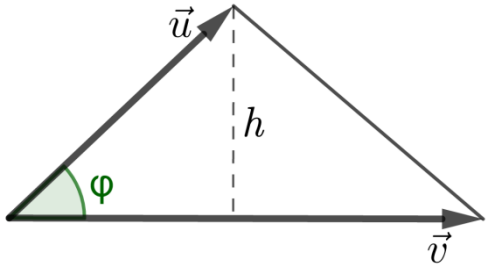


Figura 1.37. Triángulo definido por dos vectores.

$$A_T = \frac{base\ .altura}{2}$$

$$A_T = \frac{\|\vec{v}\|\ .\|\vec{u}\|\ sen\varphi}{2}$$

$$A_T = \frac{\|\vec{u}\times\vec{v}\|}{2}$$

Ejemplo: dados los puntos $A(1,3,1), B(1,-1,3) y\ C(-2,3,2)$, calcular el área del triángulo $\overset{\Delta}{\widetilde{ABC}}$.

Una de las formas de resolver la situación es utilizando vectores y pensando en el área del triángulo como la mitad del área del paralelogramo que tiene por lados no paralelos dos vectores que se pueden formar con los puntos dados. Figura 1.38.

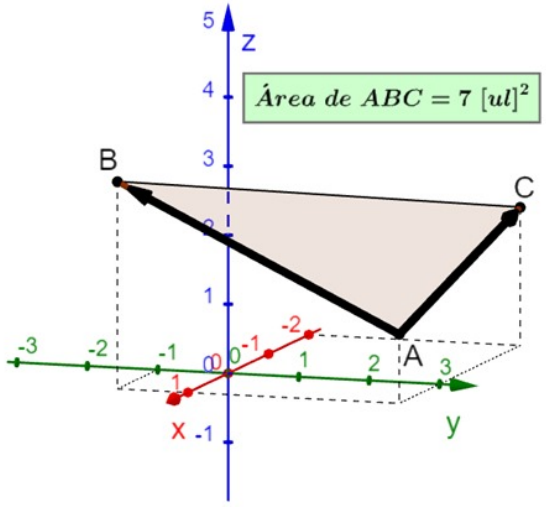


Figura 1.38. Triángulo definido por tres puntos.

$\overrightarrow{AB} = (0, -4, 2)\ y\ \overrightarrow{AC} = (-3, 0, 1)$

$\overrightarrow{AB} \times \overrightarrow{AC} = (-4, -6, -12)$

$\|\overrightarrow{AB}\ x\ \overrightarrow{AC}\| = \sqrt{16 + 36 + 144} = 14$

Área del triángulo: $\frac{1}{2}\ 14 = 7\ [\ ul]^2$

Aplicaciones del producto vectorial: Torque, Torca, Momento de fuerza o Momento de torsión es una magnitud vectorial que se simboliza con la letra griega $\overrightarrow{\tau}$ (tau).

Una fuerza $\overrightarrow{F}$ actúa en un punto de un cuerpo que está posicionado por el vector $\overrightarrow{r}$ respecto a un origen O sobre el eje de rotación, entonces el Torque es un vector que resulta del producto vectorial entre la fuerza y el vector posición: $\overrightarrow{\tau} = \overrightarrow{r} \times \overrightarrow{F}$

El módulo del torque estará dado por: $\|\vec{\tau}\| = \|\vec{r}\| \cdot \|\vec{F}\| \cdot sen\varphi$, siendo φ el ángulo entre los dos vectores. De esta expresión se deduce que cuando se desea hacer girar un cuerpo, por ejemplo, desenroscar un bulón, una forma sencilla de hacerlo es tomar la llave desde un punto lo más alejado del eje de giro, aumentando $\|\vec{r}\|$ y aplicar una fuerza $\vec{F}$ que en lo posible sea perpendicular a $\vec{r}$ de manera tal que $\varphi = 90°$. Figura 1.39.



Figura 1.39. Torque de una fuerza.

Con los productos estudiados, se puede formar un tercer producto entre vectores, producto mixto cuyo resultado es un escalar.

4.4. Producto Mixto[11]

Se llama producto mixto de tres vectores $\vec{u}$, $\vec{v}$ *y* $\vec{w}$, *y se simboliza como* $(\vec{u},\vec{v},\vec{w})$, *al número real que se obtiene al realizar* $(\vec{u} \times \vec{v}) \cdot \vec{w}$

Dados $\vec{u}$, $\vec{v}$, $\vec{w} \in R^3 \Longrightarrow (\vec{u},\vec{v},\vec{w}) = (\vec{u} \times \vec{v}) \cdot \vec{w}$

Si los vectores son $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$ y $\vec{w} = (w_1, w_2, w_3)$ y se resuelven los productos en el orden indicado, se obtiene

[11] Esta operación también es conocida como Triple producto escalar.

$$\vec{u} \times \vec{v} = \begin{vmatrix} \check{\imath} & \check{\jmath} & \check{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

$$\vec{u} \times \vec{v} = (u_2v_3 - u_3v_2)\check{\imath} + (u_3v_1 - u_1v_3)\check{\jmath} + (u_1v_2 - u_2v_1)\check{k}.$$

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = \left[(u_2v_3 - u_3v_2)\check{\imath} + (u_3v_1 - u_1v_3)\check{\jmath} + (u_1v_2 - u_2v_1)\check{k}\right] \cdot (w_1, w_2, w_3)$$

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = [w_1(u_2v_3 - u_3v_2) + w_2(u_3v_1 - u_1v_3) + w_3(u_1v_2 - u_2v_1)]$$

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = (u_2v_3w_1 - u_3v_2w_1 + u_3v_1w_2 - u_1v_3w_2 + u_1v_2w_3 - u_2v_1w_3)$$

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = (u_1v_2w_3 + u_2v_3w_1 + u_3v_1w_2 - u_3v_2w_1 - u_1v_3w_2 - u_2v_1w_3)$$

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

Es decir que el producto mixto se resuelve mediante un determinante de orden 3 cuyas filas son las componentes de los vectores dados.

Ejemplo 1: calcular, por dos procedimientos, el producto mixto de los siguientes vectores $\vec{u} = (0, 1, 1)$, $\vec{v} = (0, 3, 0)$ y $\vec{w} = (5, 0, 2)$

a) Aplicamos la definición, es decir, calculamos el producto vectorial $\vec{u}\ x\ \vec{v}$ y a ese vector lo multiplicamos escalarmente por $\vec{w}$

$\vec{u} \times \vec{v} = (-3, 0, 0)$

$(\vec{u} \times \vec{v}) \cdot \vec{w} = (-3, 0, 0) \cdot (5, 0, 2)$

$(\vec{u} \times \vec{v}) \cdot \vec{w} = -15$

b) Calculamos directamente el producto mixto, resolviendo el determinante de orden 3 cuyas filas son las componentes de los vectores dados.

$$(\overrightarrow{u} \times \overrightarrow{v}) \cdot \overrightarrow{w} = \begin{vmatrix} 0 & 1 & 1 \\ 0 & 3 & 0 \\ 5 & 0 & 2 \end{vmatrix} = -15$$

Propiedades del producto mixto

P1) El producto mixto es cíclico. Esto significa que puede tomarse como primer factor otro vector, con tal de que no se altere el orden cíclico (colocar los tres vectores en una circunferencia cuyo recorrido pase primero por $\overrightarrow{u}$, luego por $\overrightarrow{v}$ y finalmente por $\overrightarrow{w}$, ver esquema de la Figura 1.40).

$\forall\, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w} \in R^3 \Longrightarrow (\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = (\overrightarrow{v}, \overrightarrow{w}, \overrightarrow{u}) = (\overrightarrow{w}, \overrightarrow{u}, \overrightarrow{v})$

D) Sean los vectores:

$\overrightarrow{u} = (u_1, u_2, u_3),\ \overrightarrow{v} = (v_1, v_2, v_3)\ \ y\ \ \overrightarrow{w} = (w_1, w_2, w_3)$ entonces:

$$(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

$$(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = (u_1v_2w_3 + u_2v_3w_1 + u_3v_1w_2 - u_3v_2w_1 - u_1v_3w_2 - u_2v_1w_3)\ \ (1)$$

$$(\overrightarrow{v}, \overrightarrow{w}, \overrightarrow{u}) = \begin{vmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \end{vmatrix}$$

$$(\overrightarrow{v}, \overrightarrow{w}, \overrightarrow{u}) = (u_1v_2w_3 + u_2v_3w_1 + u_3v_1w_2 - u_3v_2w_1 - u_1v_3w_2 - u_2v_1w_3)\ (2)$$

$$(\overrightarrow{w}, \overrightarrow{u}, \overrightarrow{v}) = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

$$(\overrightarrow{w}, \overrightarrow{u}, \overrightarrow{v}) = (u_1v_2w_3 + u_2v_3w_1 + u_3v_1w_2 - u_3v_2w_1 - u_1v_3w_2 - u_2v_1w_3)\ (3)$$

Como (1), (2) y (3) son iguales, queda demostrada la propiedad.

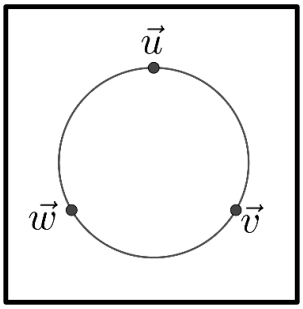


Figura 1.40. Propiedad cíclica del producto mixto.

Se debe tener en cuenta que en el desarrollo del segundo y tercer determinante se aplicó directamente la propiedad conmutativa tanto de la suma como del producto de números reales.

P_2) El producto mixto cambia de signo al intercambiar dos de los vectores (nótese que esto cambia el orden cíclico).

Por ejemplo $\forall\, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w} \in R^3 \Longrightarrow (\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = -(\overrightarrow{v}, \overrightarrow{u}, \overrightarrow{w})$

D) En este caso trabajando el segundo miembro de la igualdad se demostrará que es igual al primer miembro.

$-(\overrightarrow{v}, \overrightarrow{u}, \overrightarrow{w}) = -(\overrightarrow{v} \times \overrightarrow{u}) \cdot \vec{w}$ por definición de producto mixto

$-(\overrightarrow{v}, \overrightarrow{u}, \overrightarrow{w}) = (\overrightarrow{u} \times \overrightarrow{v}) \cdot \vec{w}$ por anticonmutatividad del producto vectorial

$-(\overrightarrow{v}, \overrightarrow{u}, \overrightarrow{w}) = (\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w})$ por definición de producto mixto

De manera similar se puede demostrar que: $(\overrightarrow{v}, \overrightarrow{w}, \overrightarrow{u}) = -(\overrightarrow{u}, \overrightarrow{w}, \overrightarrow{v})$ o también que $(\overrightarrow{w}, \overrightarrow{u}, \overrightarrow{v}) = -(\overrightarrow{w}, \overrightarrow{v}, \overrightarrow{u})$

P_3) El producto mixto es distributivo respecto a la suma de vectores.

$\forall\, \overrightarrow{u}, \overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w} \in R^3 \Longrightarrow (\overrightarrow{u} + \overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w}) = (\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) + (\overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w})$

D) Sean los vectores

$\overrightarrow{u} = (u_1, u_2, u_3),\ \overrightarrow{x} = (x_1, x_2, x_3),\ \overrightarrow{v} = (v_1, v_2, v_3)\ y\ \overrightarrow{w} = (w_1, w_2, w_3)$

entonces:

$(\overrightarrow{u} + \overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w}) = \begin{vmatrix} u_1 + x_1 & u_2 + x_2 & u_3 + x_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$ Por propiedades de los determinantes[12]

$(\overrightarrow{u} + \overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} + \begin{vmatrix} x_1 & x_2 & x_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$

$(\overrightarrow{u} + \overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w}) = (\overrightarrow{u} \times \overrightarrow{v}) \cdot \overrightarrow{w} + (\overrightarrow{x} \times \overrightarrow{v}) \cdot \overrightarrow{w}$

$(\overrightarrow{u} + \overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w}) = (\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) + (\overrightarrow{x}, \overrightarrow{v}, \overrightarrow{w})$

P4) Si cualquiera de los vectores esta multiplicado por un escalar, el producto mixto queda multiplicado por dicho escalar.

$\forall\, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w} \in R^3, \forall \alpha \in R \Longrightarrow (\alpha\, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = \alpha\, (\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w})$

D) Sean los vectores

$\overrightarrow{u} = (u_1, u_2, u_3),\ \overrightarrow{v} = (v_1, v_2, v_3)\ y\ \overrightarrow{w} = (w_1, w_2, w_3)$ y el escalar α, entonces:

$(\alpha\, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = \begin{vmatrix} \alpha u_1 & \alpha u_2 & \alpha u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$

[12] Si cada elemento de una fila o columna de una matriz cuadrada se puede escribir como la suma de dos o más términos, su determinante se puede escribir como la suma de dos o más determinantes.

$(\alpha\,\overrightarrow{u},\overrightarrow{v},\overrightarrow{w}) = \alpha \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$ Por propiedades de los determinantes[13]

$(\alpha\,\overrightarrow{u},\overrightarrow{v},\overrightarrow{w}) = \alpha\,(\overrightarrow{u},\overrightarrow{v},\overrightarrow{w})$

De manera similar puede demostrarse para:

$(\overrightarrow{u},\alpha\,\overrightarrow{v},\overrightarrow{w}) = \alpha\,(\overrightarrow{u},\overrightarrow{v},\overrightarrow{w})$ y $(\overrightarrow{u},\overrightarrow{v},\alpha\,\overrightarrow{w}) = \alpha\,(\overrightarrow{u},\overrightarrow{v},\overrightarrow{w})$

Interpretación geométrica del producto mixto

Considérese tres vectores de R^3 y constrúyase un paralelepípedo (cuerpo cuyas seis caras son paralelogramos), tal como se muestra en la Figura 1.41.

Si se desea calcular el volumen de este paralelepípedo (V_P) se deberá hacer:

$V_P = \text{Área de la base} \,.\, altura$

$\text{Área de la base} = \|\overrightarrow{u} \times \overrightarrow{v}\|$ y $cos\,\varphi = \frac{h}{\|\overrightarrow{w}\|}$ $\Longrightarrow$ $h = \|\overrightarrow{w}\| \cos\varphi$. Por lo tanto

$V_P = \|\overrightarrow{u} \times \overrightarrow{v}\|.\,\|\overrightarrow{w}\| \cos\varphi$

$V_P = (\overrightarrow{u} \times \overrightarrow{v}) \cdot \overrightarrow{w}$

$V_P = (\overrightarrow{u},\overrightarrow{v},\overrightarrow{w})$

[13] Si se multiplican todos los elementos de una fila o columna de una matriz por un mismo número, el determinante correspondiente queda multiplicado por dicho número.

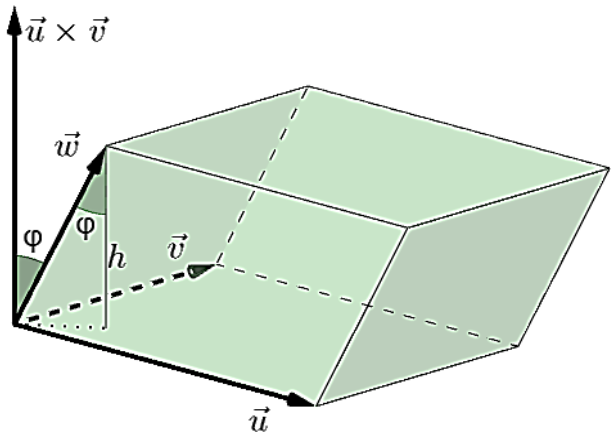


Figura 1.41. Paralelepípedo definido por tres vectores.

Esto significa que *geométricamente el producto mixto de tres vectores* $\overrightarrow{u}, \overrightarrow{v}$ *y* $\overrightarrow{w}$ *es, en valor absoluto, igual al volumen del paralelepípedo que tiene por aristas esos vectores llevados a un origen común.*

Ejemplo: determinar el volumen del paralelepípedo que tiene por aristas los vectores $\overrightarrow{u} = (1, 4, 0), \overrightarrow{v} = (2, 1, 5)$ *y* $\overrightarrow{w} = (4, 2, -1)$.

$$(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = \begin{vmatrix} 1 & 4 & 0 \\ 2 & 1 & 5 \\ 4 & 2 & -1 \end{vmatrix} = 77$$

Esto significa que $V_P = 77\ [ul]^3$, tal como se muestra en la Figura 1.42.

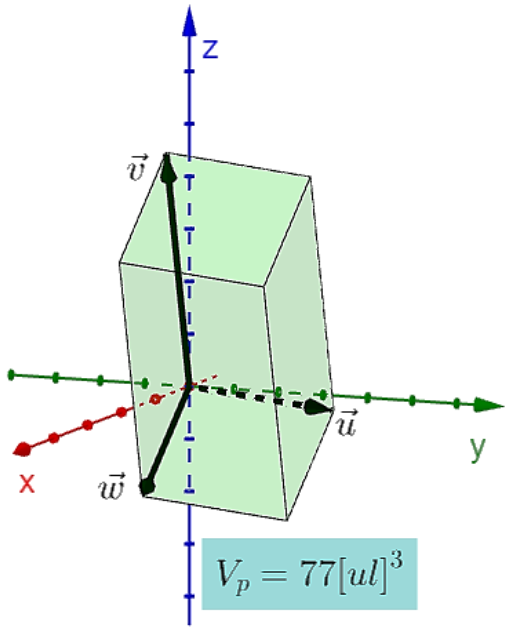


Figura 1.42. Interpretación geométrica del producto mixto.

Si el volumen del paralelepípedo es 0 significa que los tres vectores están en el mismo plano: vectores coplanares. Esta consecuencia se enuncia de la siguiente manera:

La condición necesaria y suficiente para que tres vectores $\vec{u}, \vec{v}$ *y* $\vec{w}$ *sean coplanares es que su producto mixto sea cero.*

$(\vec{u}, \vec{v}, \vec{w}) = 0 \Leftrightarrow \vec{u}, \vec{v}$ *y* $\vec{w}$ *son coplanares.*

D)

$\Rightarrow$) Si el producto mixto de los tres vectores es cero, el volumen del paralelepípedo es cero y los vectores deben ser coplanares.

⟸) Si los tres vectores son coplanares, el paralelepípedo formado por ellos está en un plano, la altura es nula y su volumen también, luego el producto mixto es cero.

Ejemplo: sean los vectores $\overrightarrow{u} = (1, 0, 4), \overrightarrow{v} = (2, 0, 2) y\ \overrightarrow{w} = (2, 0, 3)$ al calcular su producto mixto, obtenemos:

$$(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) = \begin{vmatrix} 1 & 0 & 4 \\ 2 & 0 & 2 \\ 2 & 0 & 3 \end{vmatrix} = 0$$

Esto significa que los tres vectores son coplanares, sus direcciones están todos sobre el mismo plano, en este caso $y = 0$. Figura 1.43.

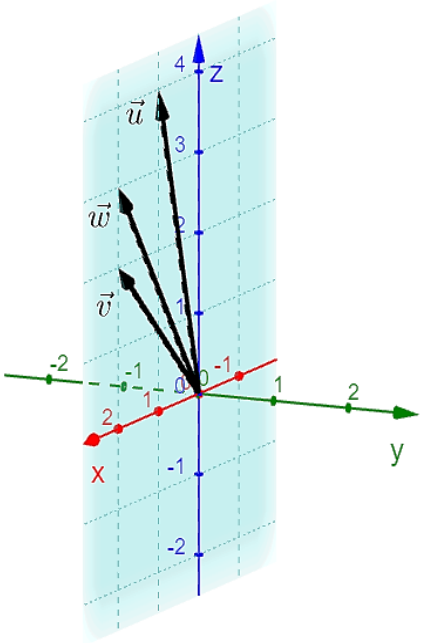


Figura 1.43. Vectores coplanares.

CAPÍTULO 2

NÚMEROS COMPLEJOS

1. Introducción

En el campo de los números racionales ($\mathbb{Q}$) son siempre posibles las cuatro operaciones racionales: adición y, su inversa, sustracción; multiplicación y, su inversa, división, siempre que el divisor sea distinto de cero. En consecuencia, la potenciación con exponente natural como caso particular de la multiplicación es siempre posible. Pero veamos qué ocurre con sus operaciones inversas.

Sean $a, b \in \mathbb{Q}$ y $n \in \mathbb{N}$. En la relación $\sqrt[n]{a} = b \iff b^n = a$, el número b recibe el nombre de raíz n-ésima de a, si conocemos a y b, el exponente n, al que hay que elevar b para obtener el número a se llama logaritmo de a respecto de la base b. Tenemos, por lo tanto, dos operaciones inversas a la potenciación: la radicación y la logaritmación.

La extracción de raíces no siempre es posible ya que, si el índice es par y el radicando es negativo, no existe ningún número real cuyas potencias de exponente par sean negativas. Asimismo, carecen de logaritmo en el sistema real los números negativos, dado que, si la base es positiva sus potencias de exponente real cualquiera, son siempre positivas, y, por lo tanto, no pueden coincidir con el número dado si este es negativo.

Todavía carecen de sentido, expresiones tales como: $\sqrt{-3}, \sqrt[4]{-\frac{1}{16}}, \log(-4)$, ...

Muchos otros problemas quedan sin solución en el campo de números reales; a modo de ejemplo, resolver ecuaciones de la forma $x^n = a$ con $n \in \mathbb{N}$ y $a \in \mathbb{R}$. Si n es impar, la ecuación tiene exactamente una solución real. Por ejemplo, si $x^3 = 8 \implies x = 2$. Si n es par y a es positivo, existen exactamente dos soluciones reales. Por ejemplo, si $x^4 = 16 \implies x = \pm 2$. Sin embargo, si n es par y a es negativo, la ecuación no tiene soluciones en $\mathbb{R}$ ya que, como se dijo anteriormente, ningún número real elevado a una potencia par es negativo. Particularmente la ecuación $x^2 = -1$ no tiene solución en los números reales.

Se hace necesaria la ampliación de $\mathbb{R}$ a un conjunto en donde puedan resolverse situaciones como las enunciadas anteriormente, de manera que estos nuevos números definidos comprendan como caso particular a los números reales. Tal conjunto es el de los números complejos.

En este segundo capítulo estudiaremos las distintas formas en las que se puede expresar un mismo número complejo y la equivalencia entre ellas; las operaciones (incluyendo la conjugación) que se pueden realizar entre estos números y en las diversas formas en que se expresan; en todos los casos, la correspondiente interpretación geométrica; también se demostrarán las propiedades de las operaciones.

1.1. Conceptos básicos

En vista de la insuficiencia del campo de números reales, lo ampliaremos creando nuevos números, cada uno de los cuales se compone de un par de números reales de igual modo que todo número racional se compone de un par de números enteros.

Número complejo: *llamaremos número complejo a todo par ordenado de números reales.*

El conjunto de números complejos se designa con $\mathbb{C}$, es decir que $\mathbb{C} = \mathbb{R}^2$ o sea que se puede caracterizar $\mathbb{C}$ como $\mathbb{C} = \{(a,b)/a \in \mathbb{R} \wedge b \in \mathbb{R}\}$.

La notación usual para los números complejos es $z = (a,b)$ conocida como forma Cartesiana. Son ejemplos: $z_1 = (2,3)$, $z_2 = \left(-\sqrt{3},2\right)$, $z_3 = \left(-\frac{1}{2},-3\right)$, $z_4 = (-1,1)$.

Parte real de un número complejo es su primera componente, mientras que **parte imaginaria** es su segunda componente. Esto se simboliza como: Sea $z = (a,b) \Longrightarrow Re(z) = a \ \wedge Im(z) = b$.

Para que los nuevos números definidos comprendan como caso particular a los números reales, *convendremos en adoptar para cada número real "a" la expresión compleja* $(a,0)$.

Particularmente el par ordenado (1,0) identifica al número real 1 y el (0,0) al 0. Los números complejos no reales se llaman $imaginarios$, en particular los números $(0,b)$ que tienen nula su primera componente se llaman $imaginarios\ puros$. El más sencillo es $(0,1)$ llamada unidad imaginaria y lo designaremos siempre por la letra i.

Un número complejo *es real si y solo si su parte imaginaria es cero.*

Un número complejo *es imaginario puro si y solo si su parte real es cero.*

Ejemplos: $z_5 = (-2, 0) = -2,\ z_6 = (0, -2) = 2i$

2. Representación gráfica de números complejos

Considerando un sistema de coordenadas cartesianas ortogonales, los números complejos se corresponden con los puntos del plano. La abscisa de cada punto es la parte real y la ordenada es la parte imaginaria.

Un complejo $z = (a, b)$ se representa entonces geométricamente por el punto del plano –que indicamos con la misma letra z– de coordenadas cartesianas a y b o bien por el vector asociado al complejo que va desde el origen del sistema hasta el punto z (vector posición del punto), tal como se muestra en la Figura 2.1.

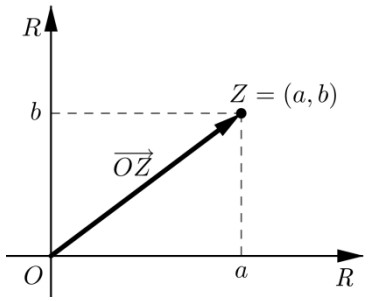


Figura 2.1. Representación geométrica de z.

Luego:

A cada complejo le corresponde un punto del plano y a cada punto del plano le corresponde un único número complejo. Un número complejo puede tomarse como representante de un vector y un vector puede tomarse como representante de un número complejo.

Esto significa que geométricamente los números complejos se pueden interpretar como un caso particular de los *n–vectores*, $\mathbb{R}^{\boldsymbol{n}}$, en donde $n = 2$, por lo tanto, son válidos muchos de los conceptos estudiados en el Capítulo 1: Vectores de $\mathbb{R}^{\boldsymbol{n}}$.

Ejemplo 1: representar gráficamente $z_1 = (2, 3)$, $z_2 = \left(-\sqrt{3}, 3\right)$, $z_3 = \left(-\frac{3}{2}, -2\right)$, $z_4 = (-1, 1)$, $z_5 = (-3, 0) = -3$, $z_6 = (0, -2) = -2i$

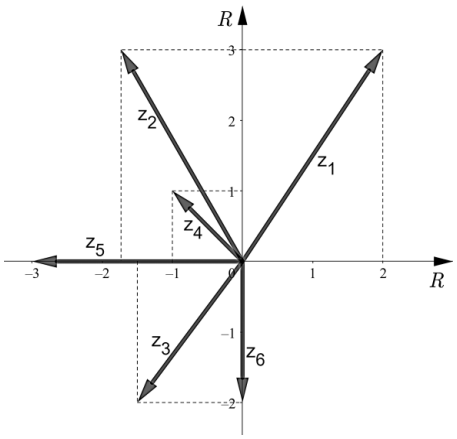


Figura 2.2. Vectores asociados a números complejos.

Ejemplo 2: determinar analítica y gráficamente los complejos $z = (x, y)$ tales que:

a) $Re(z) = -3$

Todos los pares ordenados que tengan como primera componente igual a –3, son los que tienen la forma $(-3, y)$. La ecuación $x = -3$ representa una recta paralela al eje de las ordenadas que pasa por el punto de abscisa –3. Esto significa que todos los números complejos cuya parte real es igual a

–3, están representados gráficamente en la recta de ecuación $x = -3$. Todo esto se muestra en la Figura 2.3.

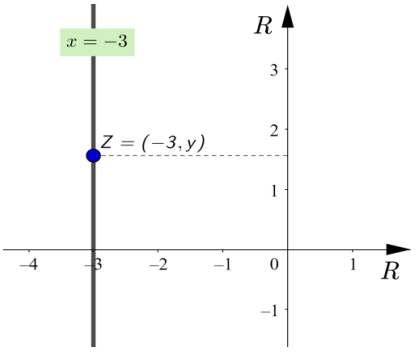


Figura 2.3. Conjunto de números complejos.

b) *Im (z)* ≤ 2

Esta condición se traduce como $y \leq 2$, la segunda componente de los números complejos es igual o menor que 2. La ecuación $y \leq 2$ corresponde al semiplano que contiene al origen y cuyo borde es la recta de ecuación $y = 2$, tal cual se muestra en la Figura 2.4.

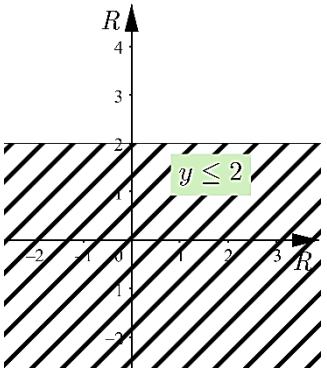


Figura 2.4. Conjunto de números complejos.

c) $Re(z) + Im(z) = 2$

La condición se traduce como $x + y = 2$, ecuación en dos variables que representa una recta, cuya ecuación explícita es $y = -x + 2$. Esto se representa en la Figura 2.5.

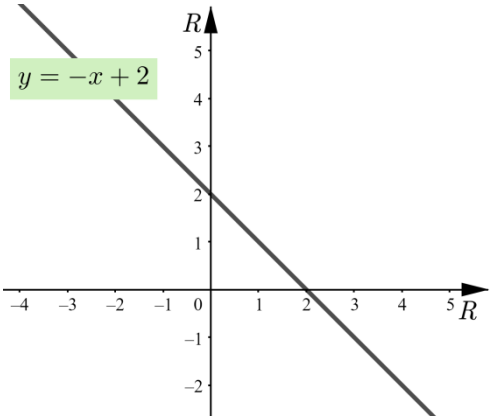


Figura 2.5. Conjunto de números complejos.

3. Otras formas de expresar un número complejo

3.1. Números complejos expresados en forma binómica

Por el momento, en $\mathbb{C} = \mathbb{R}^2$ vamos a definir la adición y multiplicación de la siguiente manera:

Sean $(a, b)\ y\ (c, d)\ \in \mathbb{C}\ \Longrightarrow$

$(a, b) + (c, d) = (a + c, b + d)\ \wedge\ (a, b).(c, d) = (ac - bd,\ ad + bc)$

Ejemplo: sean los complejos $z_1 = (1, -6)\ y\ \ z_2 = (-2, 4)$ entonces:

$z_1 + z_2 = (1, -6) + (-2, 4) = (-1, -2)$

$z_1.z_2 = (1, -6)(-2, 4) = \big(1(-2) - (-6)4, 1.4 + (-6)(-2)\big) = (22, 16)$

Veamos qué ocurre cuando multiplicamos el número real b, expresado como complejo $b = (b, 0)$ por la unidad imaginaria i que expresada como complejo es $(0, 1)$ es decir $i = (0,1)$.

$(b, 0)(0, 1) = (b.0 - 0.1, b.1 + 0.0) = (0, b)$

Es decir que al multiplicar un número real (expresado como complejo) por la unidad imaginaria, se convierte a este en un número imaginario puro, ya que $(0, b)$ tiene su primer componente igual a cero.

$(b, 0)(0, 1) = (0, b)$ reemplazando $(b, 0) = b$ y $i = (0,1) = i$ obtenemos

$bi = (0, b)$ o lo que es lo mismo

$(0, b) = bi = ib$[14]

Ahora bien, si realizamos la siguiente suma:

$(a, 0) + (0, b) = (a + 0, 0 + b) = (a, b)$

Pero $(a, 0) = a \ \ y \ \ (0, b) = bi$ por lo tanto

$(a, b) = a + bi$ o lo que es lo mismo:

$z = a + bi$ conocida como expresión binómica de un número complejo.

Ejemplo: expresar los números complejos del ejemplo 1, en forma binómica.

$z_1 = (2, 3) = 2 + 3i$ $\qquad$ $z_3 = \left(-\frac{1}{2}, -3\right) = -\frac{1}{2} - 3i$

$z_2 = \left(-\sqrt{3}, 2\right) = -\sqrt{3} + 2i$ $\qquad$ $z_4 = (-1, 1) = -1 + i$

A partir de esto, nos interesan las potencias sucesivas de la unidad imaginaria, para lo cual hacemos:

[14] En el apartado correspondiente, se demostrará la propiedad conmutativa del producto de números complejos.

$i^0 = \mathbf{1}$ $\quad i^4 = i^2\, i^2 = -1(-1) = \mathbf{1}$ $\quad i^8 = \mathbf{1}$

$i^1 = \boldsymbol{i}$ $\quad i^5 = i^2\, i^3 = -1(-i) = \boldsymbol{i}$ $\quad i^9 = \boldsymbol{i}$

$i^2 = -\mathbf{1}$ $\quad i^6 = i^2\, i^4 = -1\,.1 = -\mathbf{1}$

$i^{10} = -\mathbf{1}$

$i^3 = i^2\, i = -1.\, i = -\boldsymbol{i}$ $\quad i^7 = i^4\, i^3 = 1(-i) = -\boldsymbol{i}$

$i^{11} = -\boldsymbol{i}$

La regularidad que se puede observar en los resultados nos conduce a enunciar el siguiente teorema que permite calcular cualquier potencia de la unidad imaginaria.

Teorema de las potencias n–ésimas de i: Sean $n, k\ y\ r$ tres números naturales tales que $n = 4k + r$ con $0 \leq r \leq 3$, entonces $i^n = i^r$

D)

$i^n = i^{4\,k+r}$

$i^n = i^{4\,k}\, i^r$

$i^n = (i^4)^k\, i^r$

$i^n = (1)^k\, i^r$

$i^n = 1\ i^r$

$i^n = \ i^r$

Es decir que podemos calcular cualquier potencia de i con exponente natural n, considerando como exponente el resto r de la división de n en cuatro.

Ejemplos

Calcular $i^{253}\ e\ i^{757}$

Como $253 = 4$ x $63 + 1 \implies i^{253} = i^1 = i$

Como $747 = 4$ x $186 + 3 \implies i^{747} = i^3 = -i$

Finalmente, nótese que hasta el momento no hemos definido ninguna relación de la forma $z_1 < z_2$ con $z_1, z_2 \in \mathbb{C}$, esto se debe a que la diferencia esencial que presentan los complejos respecto al cuerpo[15] de los números reales, consiste en que es no ordenado.

En efecto, si el conjunto de números complejos fuera ordenado, como $i \neq 0$ se presentarían dos casos: $i > 0$ ó $i < 0$.

Si $i > 0$ entonces $i.i > 0.i$ es decir $i^2 > 0$ o sea que $-1 > 0$, lo cual es un absurdo. De manera similar, si suponemos $i < 0$. Por lo tanto, el cuerpo de los números complejos no puede ser ordenado.

Las dos formas de expresar un complejo, cartesiana y binómica, no son las únicas; existe una tercera (también una cuarta) forma llamada polar (porque utiliza un sistema de coordenadas polares para ubicar el punto en el plano, que representa el número complejo) o trigonométrica.

3.2. Números complejos expresados en forma polar o trigonométrica

Como se dijo, las coordenadas cartesianas (a, b) del punto z determinan el vector $\overrightarrow{oz}$. Podemos fijar la posición de z, en el plano, de otra forma.

Elegimos una semirrecta $\overrightarrow{ox}$ a la que llamaremos *eje polar* y al origen "o" lo llamaremos *polo*. Elegimos además una *unidad de medida* y un sentido de

[15] Un cuerpo o campo es un conjunto K con dos operaciones binarias, usualmente llamadas suma "+" y producto ".". La suma es asociativa, conmutativa y su elemento neutro es 0, además para cada elemento existe un inverso. El producto es asociativo, conmutativo, su elemento neutro es 1 y todo elemento distinto de cero tiene inverso, además el producto se distribuye respecto de la suma.

giro en el plano, considerando *sentido positivo*, el sentido contrario al del movimiento de las agujas del reloj. La posición de z queda determinada por el módulo del vector $\overrightarrow{oz}$, es decir por el número $\rho = \|z\|$ y por el ángulo que debe girar el eje polar $\overrightarrow{ox}$ alrededor del polo, o, en sentido positivo hasta superponerse a la semirrecta de origen o, que contiene al vector $\overrightarrow{oz}$ y cuya medida en radianes o grados sexagesimales llamaremos φ.

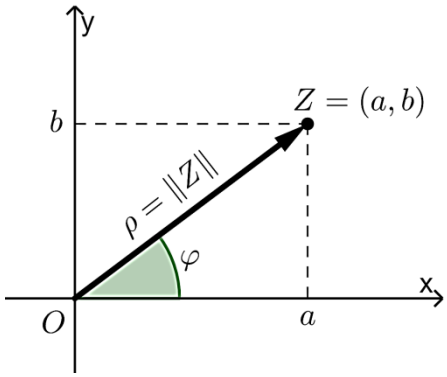


Figura 2.6. Forma polar o trigonométrica.

El conjunto formado por el polo, el eje polar, la unidad de medida y el sentido positivo para los giros recibe el nombre de Sistema de Coordenadas Polares, tal como se muestra en la Figura 2.6.

Los números: ρ llamado módulo o radio vector y φ llamado argumento, reciben el nombre de coordenadas polares del punto z o del vector $\overrightarrow{oz}$.

Las coordenadas polares suelen escribirse como: (ρ, φ) o en forma resumida como ρ_{φ}.

El número ρ, por definición, es positivo o nulo, no negativo. El número real φ se llama argumento principal del complejo no nulo si $0 \leq \varphi < 2\pi$

El número complejo de módulo ρ y argumento φ suele designarse como ρ_{φ}

Para un punto dado z existen infinitos argumentos. Si uno de ellos es φ se deducen de él todos los demás con la fórmula: $\varphi + 2k\pi$ con $k \in \mathbb{Z}$

De la figura 2.6 se puede deducir que:

$a = \rho\, cos\, \hat{\varphi}$ y $b = \rho\, sen\, \hat{\varphi}$ con $0 \leq \hat{\varphi} < 2\pi$ por lo tanto, se tiene que:

$z = a + bi = \rho\, cos\, \hat{\varphi} + \rho\, i\, sen\, \hat{\varphi}$ es decir que:

$z = \rho\,(cos\, \hat{\varphi} + i\, sen\, \hat{\varphi})$ que es la forma polar o trigonométrica del complejo z.

Donde $\rho\, y\, \hat{\varphi}$ definen unívocamente a z, mientras que z caracteriza unívocamente a ρ pero no a $\hat{\varphi}$.

3.3. Equivalencia entre los sistemas de representaciones

De la figura anterior puede deducirse que:

$\rho = \|z\| = \sqrt{a^2 + b^2}$ y $\hat{\varphi} = arctg\frac{b}{a}$ con $a \neq 0$

Fórmulas que permiten pasar del sistema cartesiano al polar.

De la misma figura puede deducirse que:

$a = \rho\, cos\, \hat{\varphi}$ y $b = \rho\, sen\, \hat{\varphi}$ con $0 \leq \hat{\varphi} < 2\pi$

Fórmulas que permiten pasar del sistema polar al cartesiano.

Ejemplo 1: determinar la forma polar del complejo $z = 1 + \sqrt{3}\, i$

En este caso $a = 1$ y $b = \sqrt{3}$, aplicando las fórmulas correspondientes

$\rho = \sqrt{1^2 + \left(\sqrt{3}\right)^2} = \sqrt{4} = 2$ y $\hat{\varphi} = arctg\frac{\sqrt{3}}{1}$

Esta última ecuación da dos valores para el argumento $\hat{\varphi} = 60°$ ó $\hat{\varphi} = 240°$, sin embargo, uno solo corresponde a z. Lo elegimos teniendo en

cuenta los signos de las coordenadas cartesianas en cada uno de los cuadrantes.

Por ser en este caso $a\ y\ b$ positivos, resulta $\hat{\varphi}$ del primer cuadrante e igual a 60°.

Por lo tanto $z = 2\ (cos\ 60° + i\ sen\ 60°) = 2_{60°}$

Ejemplo 2: determinar la forma cartesiana del complejo $z = 3_{210°}$

En este caso $\rho = 3\ \ y\ \ \hat{\varphi} = 210°$, aplicando las fórmulas correspondientes

$a = 3\ cos\ 210° = 3\ \left(-\frac{\sqrt{3}}{2}\right) = -\frac{3\sqrt{3}}{2}$

$b = 3\ sen\ 210° = 3\ \left(-\frac{1}{2}\right) = -\frac{3}{2}$

Por lo tanto $z = \left(-\frac{3\sqrt{3}}{2}, -\frac{3}{2}\right)$ ó $z = -\frac{3\sqrt{3}}{2} - \frac{3}{2}i$

4. Igualdad de números complejos

Dos números complejos $z_1 = a + bi,\ z_2 = c + di$ serán iguales si y solo si tienen sus componentes reales y sus componentes imaginarias respectivamente iguales

Si $z_1 = a + bi\ \in \mathbb{C}\ y\ \ z_2 = c + di\ \in \mathbb{C}$, entonces $a + bi = c + di\ \Leftrightarrow a = c\ \wedge b = d$

D)

$a + bi = c + di$

$\Rightarrow\ a - c = (d - b)i$

$\Rightarrow (a - c)^2 = (d - b)^2 i^2$

$\Rightarrow (a - c)^2 = -(d - b)^2$

Como $(a - c)^2 \geq 0\ \ y\ -(d - b)^2 \leq 0$ se tiene que $(a - c)^2 = 0 = (d - b)^2$ esto es $a - c = 0 = d - b$ y por lo tanto $a = c\ \ y\ \ d = b$

Si los números complejos están expresados en forma polar, se tiene:

La condición necesaria y suficiente para que dos números complejos $\rho_{1_{\varphi_1}}\ y\ \rho_{2_{\varphi_2}}$ *sean iguales, es que sean iguales sus módulos y que sus argumentos difieran en múltiplos de* $2\,\pi$.

Si $\rho_{1_{\varphi_1}} \in \mathbb{C}\ y\ \rho_{2_{\varphi_2}} \in \mathbb{C}$, *entonces*

$\rho_{1_{\varphi_1}} = \rho_{2_{\varphi_2}} \Leftrightarrow \rho_1 = \rho_2\ \wedge\ \varphi_1 = \varphi_2 + 2k\pi\ \ con\ \ k \in \mathbb{Z}$

Veamos el porqué de esta condición.

Sean $\rho_{1_{\widehat{\varphi_1}}} = \rho_1\,(cos\,\widehat{\varphi_1} + i\ sen\,\widehat{\varphi_1})$ y $\rho_{2_{\widehat{\varphi_2}}} = \rho_2\,(cos\,\widehat{\varphi_2} + i\ sen\,\widehat{\varphi_2})$,

como $\rho_{1_{\varphi_1}} = \rho_{2_{\varphi_2}}$ entonces

$\rho_1\,(cos\,\widehat{\varphi_1} + i\ sen\,\widehat{\varphi_1}) = \rho_2\,(cos\,\widehat{\varphi_2} + i\ sen\,\widehat{\varphi_2})$ de donde

$\rho_1 = \rho_2$ y $\begin{cases} cos\,\widehat{\varphi_1} = cos\,\widehat{\varphi_2} \\ sen\,\widehat{\varphi_1} = sen\,\widehat{\varphi_2} \end{cases}$ de estas dos últimas igualdades se deduce que los ángulos deben diferir en $2k\pi\ \ con\ \ k \in \mathbb{Z}$ o sea $\varphi_1 = \varphi_2 + 2k\pi\ \ con\ \ k \in \mathbb{Z}$.

Por ejemplo, las expresiones $3_{75°}$, $3_{435°}$ y $3_{795°}$ representan el mismo número complejo.

5. Operaciones con números complejos

5.1. Suma de números complejos

La suma de dos números complejos es otro complejo cuya componente real es la suma de las componentes reales de los complejos dados y cuya componente imaginaria es la suma de sus componentes imaginarias.

Si $z_1 = a_1 + b_1 i\ \in \mathbb{C}\ y\ z_2 = a_2 + b_2 i\ \in \mathbb{C}$ entonces

$z_1 + z_2 = a_1 + b_1 i\ +\ \ a_2 + b_2 i$

$z_1 + z_2 = (a_1 + a_2) + (\, b_1 \; + \; b_2)\, i$

Ejemplo:

$z_1 = 2 - 5i \;\; y \;\; z_2 = -4 - 3i$ entonces $z_1 + z_2 = -2 - 8i$

Interpretación geométrica de la suma de números complejos: geométricamente, la suma de números complejos se efectúa por la regla del paralelogramo, es decir por la regla de la suma de vectores que parten del origen de coordenadas, según lo que vimos en el Capítulo 1. Tal como se muestra en la Figura 2.7.

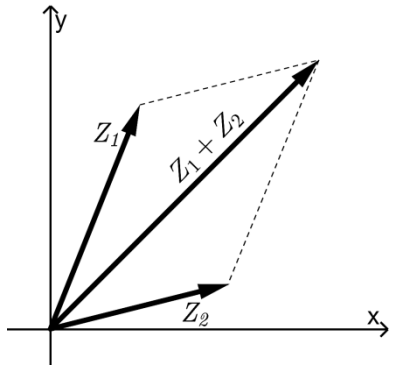


Figura 2.7. Suma de números complejos.

5.2. Producto de un número complejo por un escalar

Dado un número complejo $z = a + bi$ y un número real λ, el producto de λ por z es otro complejo cuyas componentes se obtienen multiplicando las componentes real e imaginaria del complejo dado por el número real λ .

Sean $z = a + bi \;\in \mathbb{C} \;\; y \;\; \lambda \in \mathbb{R}$ entonces $\lambda\, z = \lambda\, a + (\lambda b)i$

Gráficamente, el complejo λ z está sobre la recta que pasa por el origen de coordenadas y el punto que representa al complejo z.

Si el número complejo está expresado en forma polar, $z = \rho_{\varphi}$ entonces $\lambda\, z = \rho *_{\varphi *}$ donde

$$\varphi^* = arc\, tg\, \frac{\lambda\, b}{\lambda\, a} = \frac{b}{a} = \begin{cases} \varphi\ si\ \lambda > 0 \\ \varphi + \pi\ si\ \lambda < 0 \end{cases}$$

Esto quiere decir, como mencionamos anteriormente, que los complejos $z\ y\ \lambda\, z$ están sobre la recta que une el origen de coordenadas y el punto que representa al complejo z. Esto se muestra en la Figura 2.8.

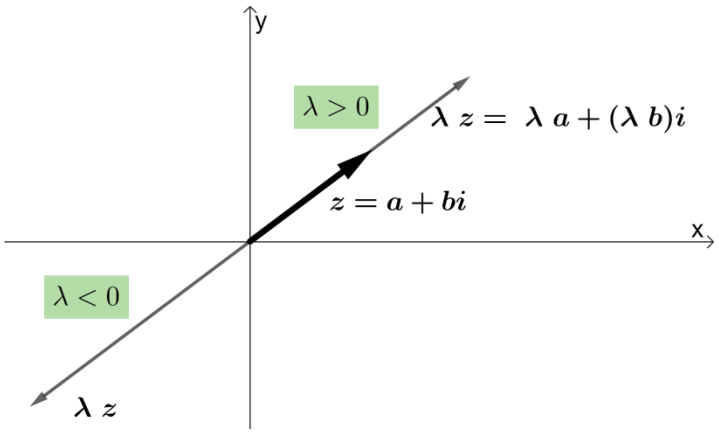


Figura 2.8. Producto de escalar por un complejo.

$$\rho^* = \|\lambda\, z\| = \sqrt{(\lambda\, a)^2 + (\lambda\, b)^2}$$

$$\rho^* = \sqrt{\lambda^2 a^2 + \lambda^2 b^2}$$

$$\rho^* = \sqrt{\lambda^2 (a^2 + b^2)}$$

$$\rho^* = |\, \lambda\, |\ \ \|z\|$$

El módulo del producto de un número real por un complejo z es igual al valor absoluto del real multiplicado por el módulo del complejo.

Ejemplo 1: sea el complejo $z = -2 + 3i$ y el escalar $\lambda = 3$ entonces $\lambda\, z = -6 + 9i$.

En términos geométricos, el complejo λz es el vector que tiene igual dirección que z y que dependiendo del signo de λ tiene igual sentido ($\lambda > 0$) o sentido opuesto ($\lambda < 0$) a la del vector que representa a z y cuya longitud es igual al valor absoluto de λ por la longitud del vector que representa a z.

Ejemplo 2: sea el vector $z = 2 + 3i \;\; y \; \lambda = 3$ entonces el vector que representa al complejo $\lambda z = 6 + 9i$ tendrá la misma dirección que el vector que representa a z y una longitud igual al triple del mismo, como se puede ver en la Figura 2.9.

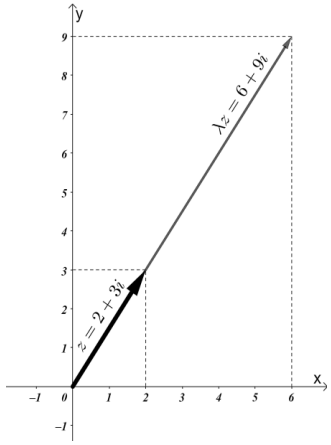


Figura 2.9. Producto de un escalar por un número complejo.

Ejemplo 3: sea el complejo $z = 2 + 4i \;\; y \; \lambda = -\frac{1}{2}$ entonces el complejo $\lambda z = -1 - 2i$ será representado por un vector cuya dirección es opuesta al vector que representa a z y una longitud igual a la mitad del mismo, estos complejos se representan en la Figura 2.10.

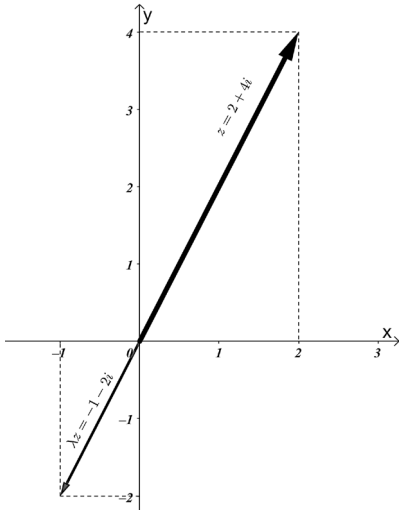


Figura 2.10. Producto de un escalar por un número complejo.

De acuerdo con la definición de producto de un número complejo por un escalar, si el complejo es $z = a + bi$ y el escalar es:

a) $\lambda = -1$ entonces $\lambda\, z = -z$ complejo opuesto a z.

b) $\lambda = 0$ entonces $\lambda\, z = 0$ complejo nulo.

Propiedades de la suma de números complejos y producto por un escalar

Veremos a continuación cuáles son las propiedades algebraicas que poseen las dos operaciones básicas antes definidas, estas propiedades son similares a las que estudiamos en el Capítulo 1 para vectores en $\mathbb{R}^{\boldsymbol{n}}$, con la salvedad de que en este caso $n = 2$.

P_1) La suma es ley de composición interna en $\mathbb{C}$

Si $z_1 = a_1 + b_1 i \in \mathbb{C}$ y $z_2 = a_2 + b_2 i \in \mathbb{C}$ entonces $(z_1 + z_2) \in \mathbb{C}$

Esto significa que la suma, como está definida, es una operación que a cada par de números complejos le hace corresponder un único elemento de $\mathbb{C}$.

D) Es inmediata por la definición de suma de números complejos.

P2) La suma es asociativa en $\mathbb{C}$.

$\forall z_1, z_2, z_3 \in \mathbb{C} : z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$

Esta propiedad ya fue demostrada para $\mathbb{R}^{\boldsymbol{n}}$ y, por lo tanto, es válida para el caso particular de $\mathbb{R}^{\mathbf{2}}$. Sin embargo, se brindará una forma alternativa de demostrar esta y otras propiedades, considerando la forma binómica de los números complejos.

D) Sean $z_1 = a_1 + b_1 i,\ z_2 = a_2 + b_2 i,\ z_3 = a_3 + b_3 i$ entonces

$z_1 + (z_2 + z_3) = (a_1 + b_1 i) + [(a_2 + b_2 i) + (a_3 + b_3 i)]$

$= (a_1 + b_1 i) + [(a_2 + a_3) + (b_2 + b_3) i\ ]$

$= [a_1 + (a_2 + a_3)] + [b_1 + (b_2 + b_3)] i$

$= [(a_1 + a_2) + a_3] + [(b_1 + b_2) + b_3] i$

$= [(a_1 + a_2) + (b_1 + b_2) i\ ] + (a_3 + b_3 i)$

$= (z_1 + z_2) + z_3$ Por lo tanto, la suma de números complejos es asociativa.

P3) La suma es conmutativa en $\mathbb{C}$.

$\forall z_1, z_2 \in \mathbb{C} : z_1 + z_2 = z_2 + z_1$

D) Sean $z_1 = a_1 + b_1 i,\ z_2 = a_2 + b_2 i$ entonces

$z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2) i$

$= (a_2 + a_1) + (b_2 + b_1) i$

$= a_2 + a_1 + b_2 i + b_1 i$

$= a_2 + b_2 i + a_1 + b_1 i$

$= (a_2 + b_2 i) + (a_1 + b_1 i)$

$= z_2 + z_1$

P4) Existe un único elemento neutro para la suma en $\mathbb{C}$.

El elemento neutro se denota con e.

$\exists e \in \mathbb{C}$ *único* $/ \forall z \in \mathbb{C} : z + e = e + z = z$.

D) Para demostrar la existencia proponemos el elemento neutro: $e = (e_1, e_2)$, y debe cumplirse que:

$z + e = z$

$(z_1, z_2) + (e_1, e_2) = (z_1, z_2)$

$(z_1 + e_1, z_2 + e_2) = (z_1, z_2)$

$z_1 + e_1 = z_1 \implies e_1 = 0$

$z_2 + e_2 = z_2 \implies e_2 = 0$

$\therefore\ e = (0, 0)$

La otra parte de la igualdad: $e + z = z$ se demuestra de manera similar.

Todo esto significa que el elemento neutro de la suma de números complejos existe y es el complejo nulo ya que se definió a este como aquel complejo cuyas componentes son todas 0.

Para demostrar la unicidad del neutro vamos a suponer que existen dos neutros para la suma: 0_1 *y* 0_2, por lo tanto, si:

0_1 es neutro entonces $0_1 + 0_2 = 0_2$

0_2 es neutro entonces $0_2 + 0_1 = 0_1$

De donde se deduce, apelando a la propiedad conmutativa de la suma, los dos primeros miembros de ambas igualdades son iguales entre sí,

entonces se concluye que $0_1 = 0_2$ y consecuentemente el neutro para la suma, es único.

P5) Cada elemento en $\mathbb{C}$ admite un único inverso aditivo u opuesto en $\mathbb{C}$.

El elemento opuesto a un z cualquiera, lo denotaremos por o.

$\forall\, z = (z_1, z_2) \in \mathbb{C} : \exists\, o \in \mathbb{C}$ *único* $/\ z + o = o + z = (0,0)$

Para demostrar la existencia del opuesto se propone $o = (o_1, o_2)$, que debe cumplir con:

$z + o = (0,0)$

$(z_1, z_2) + (o_1, o_2) = (0,0)$

$(z_1 + o_1, z_2 + o_2) = (0,0)$

$z_1 + o_1 = 0 \implies o_1 = -z_1$

$z_2 + o_2 = 0 \implies o_2 = -z_2$

$\therefore\ o = (-z_1, -z_2)$

La otra parte de la igualdad: $o + z = (0,0)$ se demuestra de manera similar.

Todo esto significa que el inverso aditivo u opuesto a un elemento $z \in \mathbb{C}$ existe y es el complejo opuesto al dado ya que cuando se definió producto por un escalar, si el escalar es –1 al multiplicarse por cualquier complejo se obtiene el opuesto de dicho complejo; es decir que $o = -z\ \forall\ z \in \mathbb{C}$.

Para demostrar la unicidad del elemento opuesto vamos a suponer que para un complejo z existen dos opuestos $z_1\ y\ z_2$ y que los mismos son iguales.

$z_1 + z = z + z_1 = 0$ y $z_2 + z = z + z_2 = 0$

$z_1 + z = 0$

$z_1 + z + z_2 = 0 + z_2$ sumando miembro a miembro z_2

$z_1 + (z + z_2) = z_2$ por asociatividad de la suma y elemento neutro

$z_1 + 0 = z_2$ por ser z_2 opuesto de z

$z_1 = z_2$

Consecuentemente, el inverso aditivo u opuesto de cada elemento $z \in \mathbb{C}$ es único.

P6) El producto es ley de composición externa en $\mathbb{C}$ con escalares en $\mathbb{R}$.

$\forall\, z \in \mathbb{C},\ \forall \lambda \in \mathbb{R} \implies \lambda\, z \in \mathbb{C}$

Esto significa que el producto de un número complejo por un escalar, como está definido, es una operación que a cada par complejo–escalar, le hace corresponder un único elemento de $\mathbb{C}$.

D) La existencia de tal complejo es inmediata, por la definición de producto de un número complejo por un escalar.

P7) El producto por un escalar satisface la asociatividad mixta.

$\forall\, z \in \mathbb{C}\ \forall \alpha, \beta \in \mathbb{R} :\ \alpha(\beta\, z) = (\alpha\, \beta)z$

D) Sea $z = a + bi$ entonces

$\alpha(\beta\, z) = \alpha[\beta(a + bi)]$

$= \alpha\, (\beta\, a + \beta\, bi)$

$= \alpha\, \beta\, a + \alpha\, \beta\, bi$

$= (\alpha\, \beta)a + (\alpha\, \beta)\, bi$

$= (\alpha\, \beta)(a + bi)$

$= (\alpha\, \beta)z$

Concluimos entonces que $\alpha(\beta\, z) = (\alpha\, \beta)z$.

P8) El producto por un escalar es distributivo respecto a la suma en $\mathbb{R}$.

$\forall\, z\ \in \mathbb{C}\ \forall \alpha, \beta\ \in \mathbb{R} : (\alpha + \beta)z = \alpha\, z + \beta\, z$

D) Sea $z = a + bi$ entonces

$(\alpha + \beta)z = (\alpha + \beta)(a + bi)$

$= (\alpha + \beta)a + (\alpha + \beta)bi$

$= \alpha\, a + \beta\, a + \alpha\, bi + \beta\, bi$

$= \alpha\, a + \alpha\, bi + \beta\, a + \beta\, bi$

$= \alpha\, (a + \ bi) + \beta(\ a + \ bi)$

$= \alpha\, z + \beta\, z$ con lo que se demuestra que el producto de un complejo por un escalar es distributivo respecto de la suma de escalares.

P9) El producto por un escalar es distributivo respecto a la suma en $\mathbb{C}$.

$\forall\, z_1, z_2\ \in \mathbb{C}\ \forall \alpha\ \in \mathbb{R} :\ \alpha\, (z_1 + z_2) = \alpha\, z_1 + \alpha\, z_2$

D) Sean $z_1 = a_1 + b_1 i,\ z_2 = a_2 + b_2 i$ entonces

$\alpha\, (z_1 + z_2) = \alpha\, [(a_1 + b_1 i) + (a_2 + b_2 i)]$

$= \alpha\, [(a_1 + a_2) + (b_1 + b_2)i]$

$= \alpha\, (a_1 + a_2) + \alpha\, (b_1 + b_2)i$

$= \ \alpha\, a_1 + \alpha\, a_2 + \ \alpha\, b_1 i + \alpha\, b_2 i$

$= \ \alpha\, a_1 + \alpha\, b_1 i + \alpha\, a_2 + \alpha\, b_2 i$

$= \ \alpha\, (a_1 + b_1 i) + \alpha(\ a_2 + \ b_2)i$

$= \alpha\, z_1 + \alpha\, z_2$ por lo que se concluye que el producto por un escalar es distributivo respecto a la suma de números complejos.

P10) La unidad para los $\mathbb{R}$ es el neutro para el producto por un escalar.

$\forall\, z\ \in\ \mathbb{C}\ :\ 1.z = z$

Como el conjunto de los números complejos con las operaciones definidas: suma y producto por un escalar cumple con las propiedades enunciadas, se dice que este conjunto tiene estructura de espacio vectorial.

5.3. Diferencia de números complejos

La diferencia de números complejos significa sumar al minuendo el opuesto del sustraendo.

La diferencia de dos números complejos es otro número complejo cuya componente real es la diferencia de las componentes reales de los complejos dados y cuya componente imaginaria es la diferencia de las componentes imaginarias

Si $z_1 = a_1 + b_1 i \ \in \mathbb{C} \ y \ z_2 = a_2 + b_2 i \ \in \mathbb{C}$ entonces

$z_1 - z_2 = z_1 + (-z_2)$

$z_1 - z_2 = (a_1 + b_1 i) + (-a_2 - b_2 i)$

$z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2)\, i$

Ejemplo

$z_1 = 2 - 5i \ y \ z_2 = -4 - 3i$ entonces $z_1 - z_2 = 6 - 2i$

En cuanto a la resta, teniendo en cuenta que el número opuesto al número $z = a + bi$ es el punto del plano complejo que es simétrico al punto z con respecto al origen de coordenadas, la interpretación geométrica de la resta se reduce al caso de la suma. Esto se muestra en la Figura 2.11.

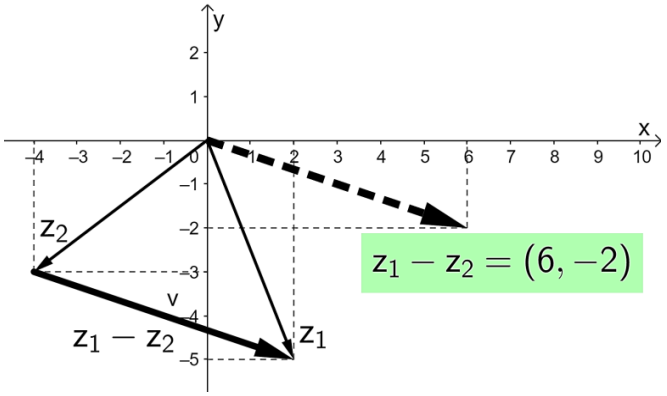


Figura 2.11. Resta de números complejos.

5.4. Producto de números complejos

El producto de dos números complejos es otro número complejo cuya componente real es igual al producto de las componentes reales menos el producto de las componentes imaginarias de los complejos dados y cuya componente imaginaria es igual a la suma de los productos cruzados.

Si $z_1 = a_1 + b_1 i \ \in \mathbb{C} \ y \ z_2 = a_2 + b_2 i \ \in \mathbb{C}$ entonces

$z_1 \, z_2 = (a_1 + b_1 i \,) \, (\ a_2 + b_2 i)$

$z_1 \, z_2 = a_1 \, a_2 + a_1 \, b_2 i + a_2 b_1 i + b_1 \, b_2 i^2$

$z_1 \, z_2 = a_1 \, a_2 + a_1 \, b_2 i + a_2 b_1 i - b_1 \, b_2$

$z_1 \, z_2 = (a_1 \, a_2 - b_1 \, b_2) + (a_1 \, b_2 + a_2 b_1) \, i$

Ejemplo

$z_1 = 2 - 5i \ \ y \ \ z_2 = -4 - 3i$ entonces

$z_1 \, z_2 = [2 \, (-4) - (-5)(-3)] + [2(-3) + (-4)(-5)] \, i$

$z_1 \, z_2 = (-8 - 15) + (-6 + 20) \, i$

$z_1 \, z_2 = -23 + 14 \, i$

Si los números complejos están expresados en forma polar o trigonométrica entonces:

El producto de dos números complejos expresados en formar polar es otro número complejo cuyo módulo es el producto de los módulos de los complejos dados y cuyo argumento es la suma de los argumentos.

Si $z_1 = \rho_1(cos\ \varphi_1 + i\ sen\ \varphi_1)$ y $z_2 = \rho_2(cos\ \varphi_2 + i\ sen\ \varphi_2)$

$z_1\ z_2 = \rho_1\ \rho_2(cos\ \varphi_1 + i\ sen\ \varphi_1)\ (cos\ \varphi_2 + \ i\ sen\ \varphi_2)$

$z_1\ z_2 = \rho_1\ \rho_2(cos\ \varphi_1\ cos\ \varphi_2 + i\ cos\ \varphi_1\ sen\ \varphi_2 + \ i\ sen\ \varphi_1\ cos\ \varphi_2 + i^2 sen\ \varphi_1 sen\ \varphi_2)$

$z_1\ z_2 = \rho_1\ \rho_2[cos\ \varphi_1\ cos\ \varphi_2 - sen\ \varphi_1 sen\ \varphi_2 + (cos\ \varphi_1\ sen\ \varphi_2 + sen\ \varphi_1\ cos\ \varphi_2)\ i\ ]$

$z_1\ z_2 = \rho_1\ \rho_2[cos(\ \varphi_1 + \varphi_2) + i\ sen\ (\varphi_1 + \varphi_2)]$

Ejemplo

Si $z_1 = 3_{25°}$ y $z_2 = 5_{30°}$ entonces $z_1\ z_2 = 15_{55°}$

Interpretación geométrica del producto: dados los complejos $z_1 = \rho_{1_{\varphi_1}}\ y\ z_2 = \rho_{2_{\varphi_2}}$ representados por los vectores $\overrightarrow{oz_1}\ y\ \overrightarrow{oz_2}$ respectivamente según la unidad de medida $U = (1, 0)$ tomada, para hallar el $\overrightarrow{oz} = z_1.z_2$ procedemos de la siguiente manera (Figura 2.12).

Sobre el segmento $\overline{oz_2}$, como homólogo de $\overline{oU}$, construimos el triángulo $\overset{\Delta}{\widehat{oz_2z}}$ semejante al $\overset{\Delta}{\widehat{oUz_1}}$ de tal modo que $\widehat{z_2oz} = \varphi_1$, siendo $\widehat{oz_2z} = \widehat{oUz_1}$. Entonces el vector $\overline{oz}$ representa el producto $z_1.z_2$. Tal como lo muestra la figura 2.12.

Por construcción, el argumento de z es $\varphi_1 + \varphi_2$ en cuanto a su módulo que es la longitud de $\overrightarrow{oz}$, por la semejanza de triángulos, se tiene:

$\frac{\overline{oz}}{\overline{oz_2}} = \frac{\overline{oz_1}}{\overline{oU}} \implies \frac{\overline{oz}}{\rho_2} = \frac{\rho_1}{1}$ por lo tanto $\overline{oz} = \rho_1\rho_2$

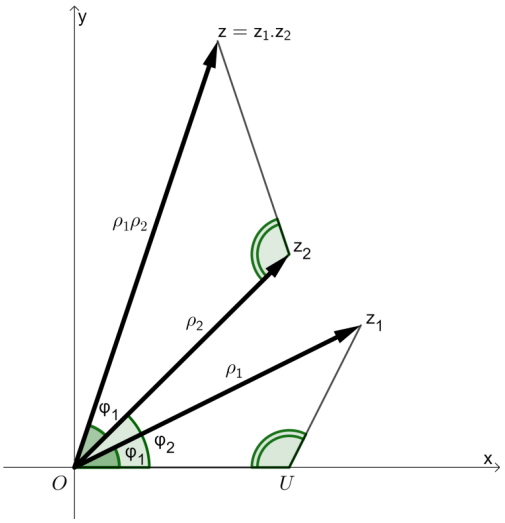


Figura 2.12. Producto de números complejos.

Propiedades del producto de números complejos

P1) El producto es ley de composición interna en $\mathbb{C}$

Si $z_1 = a_1 + b_1 i \in \mathbb{C}$ y $z_2 = a_2 + b_2 i \in \mathbb{C}$ entonces $(z_1 . z_2) \in \mathbb{C}$

Esto significa que el producto de números complejos, como está definido, es una operación que a cada par de números complejos le hace corresponder un único elemento de $\mathbb{C}$.

D) Es inmediata por la definición de producto de números complejos.

P2) El producto es asociativo en $\mathbb{C}$.

$\forall z_1, z_2, z_3 \in \mathbb{C} : z_1\ (z_2\ z_3) = (z_1\ z_2)\ z_3$

D) Sean $z_1 = a_1 + b_1 i,\ z_2 = a_2 + b_2 i,\ z_3 = a_3 + b_3 i$ entonces

$z_1\ (z_2\ z_3) = (a_1 + b_1 i)\ [(a_2 + b_2 i)\ (a_3 + b_3 i)]$

$= (a_1 + b_1 i)\ [(a_2 a_3 - b_2 b_3) + (a_2 b_3 + a_3 b_2) i\ ]$

$= [a_1(a_2 a_3 - b_2 b_3) - b_1(a_2 b_3 + a_3 b_2)] + [a_1(a_2 b_3 + a_3 b_2) + b_1(a_2 a_3 - b_2 b_3)] i$

$= (a_1 a_2 a_3 - a_1 b_2 b_3 - a_2 b_1 b_3 - a_3 b_1 b_2) + [(a_1 a_2 b_3 + a_1 a_3 b_2) + (a_2 a_3 b_1 - b_1 b_2 b_3)] i$ (1)

$(z_1\ z_2)\ z_3 = [(a_1 + b_1 i)(a_2 + b_2 i)]\ (a_3 + b_3 i)$

$= [(a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1) i]\ (a_3 + b_3 i)$

$= [a_3(a_1 a_2 - b_1 b_2) - b_3(a_1 b_2 + a_2 b_1)] + [b_3(a_1 a_2 - b_1 b_2) + a_3(a_1 b_2 + a_2 b_1)] i$

$= (a_1 a_2 a_3 - a_3 b_1 b_2 - a_1 b_2 b_3 - a_2 b_1 b_3) + (a_1 a_2 b_3 - b_1 b_2 b_3 + a_1 a_3 b_2 + a_2 a_3 b_1) i$ (2)

Por ser las expresiones (1) y (2) iguales, se concluye que $z_1\ (z_2\ z_3) = (z_1\ z_2)\ z_3$.

P3) El producto es conmutativo en $\mathbb{C}$.

$\forall z_1, z_2 \in \mathbb{C} : z_1\ z_2 = z_2\ z_1$

D) Sean $z_1 = a_1 + b_1 i,\ z_2 = a_2 + b_2 i$ entonces

$z_1 z_2 = (a_1 + b_1 i)(a_2 + b_2 i)$

$= (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1) i$ (1)

$z_2 z_1 = (a_2 + b_2 i)(a_1 + b_1 i)$

$= (a_1 a_2 - b_1 b_2) + (a_2 b_1 + a_1 b_2) i$ (2)

Por ser las expresiones (1) y (2) iguales, se concluye que $z_1\ z_2 = z_2\ z_1$.

P4) Existe un único elemento neutro para el producto en $\mathbb{C}$.

El elemento neutro para el producto es el complejo $1 + 0i = (1, 0)$

$\forall\, z = a + bi \in \mathbb{C},\ z \neq 0 + 0i : \exists\ único\ n = x + y\, i \in \mathbb{C} / z\, n = n\, z = z$

D) Para la existencia hacemos:

$z\, n = n\, z = z$

$(a + bi)(x + yi) = (x + yi)(a + bi) = a + bi$

$(ax - by) + (ay + bx)i = a + b\, i$

$\begin{cases} ax - by = a \\ ay + bx = b \end{cases}$

$x = \frac{\begin{vmatrix} a & -b \\ b & a \end{vmatrix}}{\begin{vmatrix} a & -b \\ b & a \end{vmatrix}} = \frac{a^2+b^2}{a^2+b^2} = 1$

$y = \frac{\begin{vmatrix} a & a \\ b & b \end{vmatrix}}{\begin{vmatrix} a & -b \\ b & a \end{vmatrix}} = \frac{ab-ab}{a^2+b^2} = 0$

Por lo tanto $n = 1 + 0i = (1, 0)$

Para demostrar la unicidad del neutro vamos a suponer que existen dos neutros para el producto: $n_1\ y\ n_2$, por lo tanto, si:

n_1 es neutro entonces $n_1 . n_2 = n_2$

n_2 es neutro entonces $n_2 . n_1 = n_1$

De donde se deduce, por ser los primeros miembros de ambas igualdades iguales, que $n_1 = n_2$ y, consecuentemente, el neutro para el producto es único.

Ejemplo: sea $z = 2 + 3i$ entonces

$(2 + 3i)(1 + 0i) = (1 + 0i)(2 + 3i) = 2 + 3i$

$(2.1 - 3.0) + (2.0 + 3.1)i = (1.2 - 0.2) + (1.3 + 0.3)i = 2 + 3i$

$2 + 3i = 2 + 3i = 2 + 3i$

P_5) Todo elemento no nulo en $\mathbb{C}$ admite un único inverso multiplicativo.

El inverso multiplicativo a uno cualquiera z, se nota por z^{-1}.

$\forall\, z = a + bi \in \mathbb{C},\ z \neq 0 + 0i : \exists\, único\ z^{-1} = x + y\, i \in \mathbb{C} / z\, z^{-1} = z^{-1}\, z = n$

Para la existencia hacemos:

Sea el complejo no nulo z. Si existe $z^{-1} = \frac{1}{z} = x + yi$, debe cumplirse

$z\, z^{-1} = z^{-1}\, z = n$

$(a + bi)(x + yi) = (x + yi)(a + bi) = 1 + 0i$

$(ax - by) + (ay + bx)i = 1 + 0i$

$\begin{cases} ax - by = 1 \\ bx + ay = 0 \end{cases}$

$x = \frac{\begin{vmatrix} 1 & -b \\ 0 & a \end{vmatrix}}{\begin{vmatrix} a & -b \\ b & a \end{vmatrix}} = \frac{a}{a^2+b^2}$

$y = \frac{\begin{vmatrix} a & 1 \\ b & 0 \end{vmatrix}}{\begin{vmatrix} a & -b \\ b & a \end{vmatrix}} = \frac{-b}{a^2+b^2}$

Por lo tanto $z^{-1} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}\, i = \left(\frac{a}{a^2+b^2},\ \frac{-b}{a^2+b^2}\right)$

Para demostrar la unicidad del elemento inverso vamos a suponer que, para un complejo z existen dos inversos $z_1^{-1}\ y\ z_2^{-1}$, por lo tanto se debe cumplir que $z\, z_1^{-1} = n$ y también que $z\, z_2^{-1} = n$ de donde se deduce que $z_1^{-1} = z_2^{-1}$. Consecuentemente, el inverso multiplicativo de cada elemento $z \in \mathbb{C}$ es único.

Otra forma de demostrar esta propiedad es llevar a cabo un procedimiento similar a la propiedad P_5) de la suma de números complejos.

Ejemplo: si z = 2 + 3i entonces $z^{-1} = \frac{2}{13} - \frac{3}{13}$i, para comprobar que es el inverso multiplicativo hacemos:

$(2+3\mathrm{i})\left(\frac{2}{13}-\frac{3}{13}\mathrm{i}\right)=\left(2\frac{2}{13}-3\left(-\frac{3}{13}\right)\right)+\left(2\left(-\frac{3}{13}\right)+3\frac{2}{13}\right)i$

$=\left(\frac{4}{13}+\frac{9}{13}\right)+\left(-\frac{6}{13}+\frac{6}{13}\right)i$

$=1+0i$

Al conmutar el producto, también se obtiene el mismo resultado.

P_6) El producto es distributivo respecto a la suma en $\mathbb{C}$.

$\forall z_1, z_2, z_3 \in \mathbb{C} : z_1\,(z_2+\,z_3)=z_1\,z_2+z_1\,z_3$

D) Sean $z_1=a_1+b_1i,\ z_2=a_2+b_2i,\ z_3=a_3+b_3i$ entonces

$z_1\,(z_2+\,z_3)=(a_1+b_1i)[(a_2+b_2i)+(a_3+b_3i)]$

$=(a_1+b_1i)[(a_2+a_3)+(b_2+b_3)\,i\,]$

$=[a_1(a_2+a_3)-b_1(b_2+b_3)]+[a_1(b_2+b_3)+b_1(a_2+a_3)]i$

$(a_1a_2+a_1a_3-b_1b_2-b_1b_3)+(a_1b_2+a_1b_3+a_2b_1+a_3b_1)i$ (1)

$\mathrm{z}_1\,\mathrm{z}_2+\mathrm{z}_1\,\mathrm{z}_3=[(a_1+b_1i)(a_2+b_2i)]+[(a_1+b_1i)(a_3+b_3i)]$

$=[(a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)i\,]+[(a_1a_3-b_1b_3)+(a_1b_3+a_3b_1)i]$

$=(a_1a_2-b_1b_2+a_1a_3-b_1b_3)+(a_1b_2+a_2b_1+a_1b_3+a_3b_1)i$ (2)

Por ser las expresiones (1) y (2) iguales se concluye que $\mathrm{z}_1\,(\mathrm{z}_2+\,\mathrm{z}_3)=\mathrm{z}_1\,\mathrm{z}_2+\mathrm{z}_1\,\mathrm{z}_3$.

5.5. Conjugación de números complejos

La conjugación es una operación unaria ya que al conjugar un número complejo se obtiene otro complejo que difiere, únicamente del primero, en su parte imaginaria.

Dado el complejo $z=a+bi$ se llama conjugado de z, y se simboliza como $\overline{z}$, al complejo $a-bi$.

Si $z=a+bi\in\mathbb{C}\ \implies\ \overline{z}=a-bi$

Ejemplo

Si $z_1 = -2 - 3i \implies \overline{z_1} = -2 + 3i$

Si $z_2 = \left(\frac{1}{2}, \sqrt{3}\right) \implies \overline{z_2} = \left(\frac{1}{2}, -\sqrt{3}\right)$

Dos complejos son conjugados si y solo si tienen la misma parte real y sus partes imaginarias son números opuestos.

Sean $z_1, z_2 \in \mathbb{C},\ z_1\ y\ z_2\ son\ conjugados \iff Re(z_1) = Re(z_2) \wedge Im(z_1) = -Im(z_2)$

Dos complejos conjugados caracterizan dos puntos simétricos respecto del eje real.

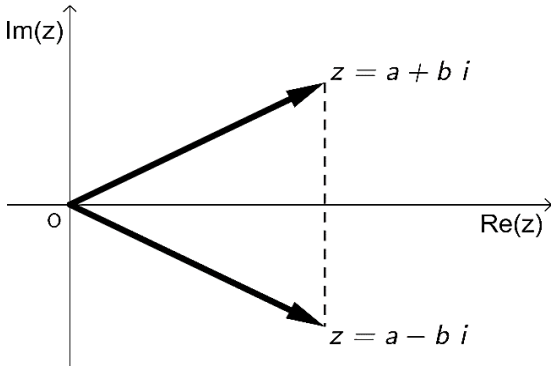


Figura 2.13. Conjugación de números complejos.

De la figura 2.13 se puede deducir que, si los complejos están expresados en forma polar sus módulos serán iguales mientras que sus argumentos serán opuestos. Entonces, si el número complejo está expresado en forma polar, $z = \rho_{\varphi}$ y se desea hallar su conjugado, expresado también en forma polar $\overline{z} = \rho^*_{\varphi^*}$

Si $z = a + b\ i$ entonces $\overline{z} = a^* + b^* i$ de donde se sabe que $\begin{cases} a^* = a \\ b^* = -b \end{cases}$ (1)

Si $z = \rho_{\varphi}$ se sabe que $\begin{cases} a = \rho\, cos\, \varphi \\ b = \rho\, sen\, \varphi \end{cases}$, entonces $\overline{z} = \rho^{*}_{\varphi^{*}}$ de donde se sabe que $\begin{cases} a^{*} = \rho^{*}\, cos\, \varphi^{*} \\ b^{*} = \rho^{*}\, sen\, \varphi^{*} \end{cases}$ (2)

Para hallar ρ^{*} sabemos que

$\rho^{*} = \sqrt{(a^{*})^{2} + (b^{*})^{2}}$

$\rho^{*} = \sqrt{a^{2} + (-b)^{2}}$ *Por (1)*

$\rho^{*} = \sqrt{a^{2} + b^{2}}$

$\rho^{*} = \rho$

Para hallar φ^{*} se establece por (1) y (2) que

$a^{*} = a$ y $b^{*} = -b$

$\rho^{*}\, cos\, \varphi^{*} = \rho\, cos\, \varphi$ y $\rho^{*}\, sen\, \varphi^{*} = -\rho\, sen\, \varphi$ por ser $\rho^{*} = \rho$

$\cos \varphi^{*} = \cos \varphi$ y $sen\, \varphi^{*} = -\, sen\, \varphi$

En consecuencia $\varphi^{*} = 2\,\pi - \varphi$ ó $\varphi^{*} = -\varphi$

Por lo tanto, la expresión polar del conjugado de $z = \rho_{\varphi}$ es $\overline{z} = \rho_{2\pi-\varphi}$ ó $\overline{z} = \rho_{-\varphi}$

Ejemplo: sea el complejo $z = 2_{30^{\circ}} \implies \overline{z} = 2_{-30^{\circ}} = 2_{330^{\circ}}$

Propiedades de la conjugación de números complejos

P_1) Si un número complejo es igual a su conjugado, su componente imaginaria vale cero.

Sea $z = a + bi \in \mathbb{C}$. Si $z = \overline{z} \quad \Rightarrow \quad Im(z) = 0$

D) Si $z = a + bi \implies \overline{z} = a - bi$

Como $z = \overline{z}$ resulta $b = -b \Leftrightarrow 2b = 0 \Leftrightarrow b = 0$

P_2) El conjugado del conjugado de un número complejo es igual a este; esta propiedad también se conoce con el nombre de involución.

Si $z = a + bi \in \mathbb{C} \Rightarrow \overline{\overline{(z)}} = z$

D) Si $z = a + bi \Longrightarrow \overline{z} = a - bi$ por lo tanto $\overline{\overline{z}} = \overline{(a - bi)} = a + bi = z$

Ejemplo: sea el complejo $z = 2 + 3i \Longrightarrow \overline{z} = 2 - 3i \Longrightarrow \overline{\overline{z}} = 2 + 3i = z$

P_3) El conjugado del producto de un escalar por un número complejo es igual al escalar multiplicado por el conjugado del complejo.

Si $z = a + bi \in \mathbb{C}$ y $\alpha \in \mathbb{R} \Longrightarrow \overline{\alpha\, z} = \alpha\, \overline{z}$

D) $\overline{\alpha\, z} = \overline{\alpha\,(a + bi)}$

$= \overline{\alpha\, a + \alpha bi}$

$= \alpha\, a - \alpha bi$

$= \alpha\,(a - bi)$

$= \alpha\, \overline{z}$

Ejemplo: sea el complejo $z = 1 + i$ y $\alpha = -2$ entonces

$\overline{\alpha\, z} = \overline{-2(1 + i)} = \overline{-2 - 2i} = -2 + 2i$

$\alpha\, \overline{z} = -2\, \overline{(1 + i)} = -2(1 - i) = -2 + 2i$

P_4) El conjugado del producto de dos números complejos es igual al producto de sus respectivos conjugados.

Si $z_1 = a_1 + b_1 i \in \mathbb{C}$ y $z_2 = a_2 + b_2 i \in \mathbb{C} \Longrightarrow \overline{z_1\, z_2} = \overline{z_1}\ \overline{z_2}$

D)

$\overline{z_1\, z_2} = \overline{(a_1 + b_1 i)(a_2 + b_2 i)}$

$= \overline{(a_1\, a_2 - b_1\, b_2) + (a_1 b_2 + a_2 b_1) i}$

$= (a_1\, a_2 - b_1\, b_2) - (a_1 b_2 + a_2 b_1) i$

$= [a_1 a_2 - (-b_1)(-b_2)] + [a_1(-b_2) + a_2(-b_1)] i$

$= (a_1 - b_1 i)(a_2 - b_2\, i)$

$= \overline{z_1}\ \overline{z_2}$

Ejemplo: sean los complejos $z_1 = -2 + i\ \ y\ \ z_2 = 1 - i$, entonces

$z_1 z_2 = (-2 + i\ )\ (1 - i)$

$z_1 z_2 = -1 + 3i$ por lo tanto

$\overline{z_1 z_2} = -1 - 3i$

Por otra parte

$\overline{z_1} = -2 - i\ \ y\ \ \overline{z_2} = 1 + i$ entonces

$\overline{z_1}\ \overline{z_2} = (-2 - i)(1 + i)$

$\overline{z_1}\ \overline{z_2} = -1 - 3i$

P5) El conjugado de la suma de dos números complejos es igual a la suma de sus respectivos conjugados.

Si $z_1 = a_1 + b_1 i\ \in \mathbb{C}\ \ y\ \ z_2 = a_2 + b_2 i\ \in \mathbb{C}\ \implies\ \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$

D)

$\overline{z_1 + z_2} = \overline{(a_1 + b_1 \imath) + (a_2 + b_2 \imath)}$

$= \overline{(a_1 + a_2) + (b_1 + b_2)}\ i$

$= (a_1 + a_2) - (b_1 + b_2)i$

$= a_1 + a_2 - b_1 i - b_2 i$

$= a_1 - b_1 i + a_2 - b_2 i$

$= \overline{z_1} + \overline{z_2}$

Ejemplo: sean los complejos $z_1 = -2 + i\ \ y\ \ z_2 = 1 - 3i$, entonces

$z_1 + z_2 = (-2 + i\ ) +\ (1 - 3i)$

$z_1 + z_2 = -1 - 2i$ por lo tanto

$\overline{z_1 + z_2} = -1 + 2i$

Por otra parte

$\overline{z_1} = -2 - i \;\; y \;\; \overline{z_2} = 1 + 3i$ entonces

$\overline{z_1} + \overline{z_2} = (-2 - i) + (1 + 3i)$

$\overline{z_1}\,\overline{z_2} = -1 + 2i$

P6) El producto de un complejo por su conjugado es igual a un número real no negativo, igual al cuadrado de su módulo.

Sea $z = a + bi \in \mathbb{C} \implies z\,\overline{z} \in \mathbb{R},\; z\,\overline{z} \geq 0 \;\wedge\; z\,\overline{z} = \|z\|^2$

D) $z\,\overline{z} = (a + bi)(a - bi)$

$= a^2 - (bi)^2 - abi + abi$

$= a^2 + b^2$

$= \|z\|^2$

Como $a \; y \; b$ son números reales, resulta $z\,\overline{z} \in \mathbb{R},\; z\,\overline{z} \geq 0 \;\wedge\; z\,\overline{z} = \|z\|^2$

Ejemplo: sea el complejo $z = 2 + 3i$ entonces $\overline{z} = 2 - 3i$. Por lo tanto

$z\,\overline{z} = (2 + 3i)(2 - 3i)$

$z\,\overline{z} = 4 - 6i + 6i - 9i^2$

$z\,\overline{z} = 4 - 6i + 6i + 9$

$z\,\overline{z} = 13$

Por otra parte

$\|z\| = \sqrt{2^2 + 3^2} = \sqrt{13}$ por lo tanto

$\|z\|^2 = 13$

P7) La suma de dos complejos conjugados es igual al duplo de la parte real.

Sea $z = a + bi \in \mathbb{C} \implies z + \overline{z} = 2\,Re(z)$

D) Si $z + \overline{z} = (a + bi) + (a - bi)$

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

$= 2a$

$= 2\ Re(z)$

Ejemplo: sea el complejo $z = 2 + 3i$ entonces $\overline{z} = 2 - 3i$. Por lo tanto

$z + \overline{z} = (2 + 3i) + (2 - 3i)$

$z + \overline{z} = 4$

Por otra parte

$Re(z) = 2$ por lo tanto

$2\ Re(z) = 4$

La operación conjugación puede ser útil para calcular el inverso multiplicativo en el caso en que el complejo z está expresado en forma polar.

La expresión: $z^{-1} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}\ i$ puede escribirse de la siguiente manera:

$z^{-1} = \frac{1}{a^2+b^2}(a - bi)$ de donde se sabe que $a^2 + b^2 = \|z\|^2 \quad y \quad a - b\ i = \overline{z}$

$z^{-1} = \frac{1}{\| z \|^2}\ \overline{z}$ (1)

Ahora bien, si $z = \rho_{\varphi}$ entonces $z^{-1} = \rho^*_{\varphi^*}$ en donde el argumento φ^* es igual al argumento de $\overline{z}$ ya que en (1), $\frac{1}{\| z \|^2}$ es un número real positivo.

Esto significa que el argumento $\varphi^* = 2\pi - \varphi$

$\rho^* = \left| \frac{1}{\|z\|^2}\ \overline{z} \right| = \left| \frac{1}{\|z\|^2} \right| \ \| \overline{z} \| = \frac{1}{\|z\|^2}\ \|z\| = \frac{1}{\|z\|} = \frac{1}{\rho}$

Finalmente, si $z = \rho_{\varphi} \quad \Longrightarrow z^{-1} = \left(\frac{1}{\rho}\right)_{2\pi-\varphi} = \left(\frac{1}{\rho}\right)_{-\varphi}$

Ejemplo

Si $z = 2_{30°} \implies z^{-1} = \left(\frac{1}{2}\right)_{330°}$

5.6. Cociente de números complejos

El cociente de dos números complejos $z_1 = a_1 + b_1 i \;\; y \;\; z_2 = a_2 + b_2 i$ con $z_2 \neq 0 + 0i$ es otro complejo que se obtiene al multiplicar z_1 por el inverso multiplicativo de z_2

Si $z_1 = a_1 + b_1 i \in \mathbb{C} \; y \; z_2 = a_2 + b_2 i \in \mathbb{C}$ con $z_2 \neq 0 + 0i \implies \frac{z_1}{z_2} = z_1 \; z_2^{-1}$

$$\frac{z_1}{z_2} = z_1 \; z_2^{-1}$$

$$\frac{z_1}{z_2} = (a_1 + b_1 i)(a_2 + b_2 i)^{-1}$$

$$\frac{z_1}{z_2} = (a_1 + b_1 i)\left(\frac{a_2}{a_2^2 + b_2^2} - \frac{b_2}{a_2^2 + b_2^2} i\right)$$

$$\frac{z_1}{z_2} = \left(\frac{a_1 a_2}{a_2^2 + b_2^2} + \frac{b_1 b_2}{a_2^2 + b_2^2}\right) + \left(-\frac{a_1 b_2}{a_2^2 + b_2^2} + \frac{a_2 b_1}{a_2^2 + b_2^2}\right) i$$

Si se desea evitar el cálculo del inverso multiplicativo del divisor, es conveniente aplicar la propiedad de la conjugación de números complejos que establece: *El producto de dos complejos conjugados es igual a un número real no negativo, igual al cuadrado de su módulo.*

Esto significa que, para calcular el cociente de dos números complejos, se debe multiplicar y dividir por el conjugado del divisor. Ya que:

$$\frac{z_1}{z_2} = \frac{z_1}{z_2} \frac{\overline{z_2}}{\overline{z_2}}$$

$$\frac{z_1}{z_2} = \frac{z_1 \; \overline{z_2}}{\| z_2 \|^2}$$

Otra alternativa con el mismo resultado es pensar que si $z \in \mathbb{C} \implies z^{-1} = \frac{1}{\|z\|^2} \; \overline{z}$.

$\frac{z_1}{z_2} = z_1 \ z_2^{-1}$ como $z_2^{-1} = \frac{1}{\|z_2\|^2} \ \overline{z_2}$

$\frac{z_1}{z_2} = z_1 \ \frac{1}{\|z_2\|^2} \ \overline{z_2}$

$\frac{z_1}{z_2} = \frac{z_1 \ \overline{z_2}}{\| z_2\|^2}$

Ejemplo: calcular el cociente entre $z_1 = 3 + 3i \ \ y \ \ z_2 = 1 - 2i$

a) Calculando el inverso multiplicativo de z_2. Si $z_2 = 1 - 2i \implies z_2^{-1} = \frac{1}{5} + \frac{2}{5}i$ Por lo tanto

$\frac{z_1}{z_2} = (3 + 3i)\left(\frac{1}{5} + \frac{2}{5}i\right)$

$\frac{z_1}{z_2} = \left(3 \ \frac{1}{5} - 3\frac{2}{5}\right) + \left(3\frac{2}{5} + 3\frac{1}{5}\right)i$

$\frac{z_1}{z_2} = \left(\frac{3}{5} - \frac{6}{5}\right) + \left(\frac{6}{5} + \frac{3}{5}\right)i$

$\frac{z_1}{z_2} = -\frac{3}{5} + \frac{9}{5}i$

b) Aplicando la propiedad de la conjugación de números complejos.

$\frac{z_1}{z_2} = \frac{3+3i}{1-2i}$

$\frac{z_1}{z_2} = \frac{(3+3i)}{(1-2i)} \ \frac{(1+2i)}{(1+2i)}$

$\frac{z_1}{z_2} = \frac{3+6i+3i-6}{1^2+2^2}$

$\frac{z_1}{z_2} = -\frac{3}{5} + \frac{9}{5}i$

Si los números complejos están expresados en forma polar o trigonométrica, entonces:

El cociente de dos números complejos expresados en forma polar y siendo el segundo distinto del nulo, es otro complejo cuyo módulo es el cociente de los

módulos y cuyo argumento es la diferencia entre los argumentos del dividendo y divisor.

Si $z_1 = \rho_1(cos\,\varphi_1 + i\,sen\,\varphi_1)$ y $z_2 = \rho_2(cos\,\varphi_2 + i\,sen\,\varphi_2)$

$$\frac{z_1}{z_2} = \frac{\rho_1(cos\,\varphi_1 + i\,sen\,\varphi_1)}{\rho_2(cos\,\varphi_2 + i\,sen\,\varphi_2)}$$

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}\left[\frac{(cos\,\varphi_1 + i\,sen\,\varphi_1)}{(cos\,\varphi_2 + i\,sen\,\varphi_2)}\ \frac{(cos\,\varphi_2 - i\,sen\,\varphi_2)}{(cos\,\varphi_2 - i\,sen\,\varphi_2)}\right]$$

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}\left[\frac{(cos\,\varphi_1\,cos\,\varphi_2 + sen\,\varphi_1 sen\,\varphi_2) + (-cos\,\varphi_1 sen\,\varphi_2 + sen\,\varphi_1\,cos\,\varphi_2)\,i}{cos^2\varphi_2 + sen^2\varphi_2}\right]$$

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}cos(\varphi_1 - \varphi_2) + i\ sen\,(\varphi_1 - \varphi_2)$$

Ejemplo: si $z_1 = 3_{60°}$ y $z_2 = 4_{28°}$ $\Longrightarrow$ $\frac{z_1}{z_2} = \left(\frac{3}{4}\right)_{32°}$

Interpretación geométrica del cociente: razonando sobre la figura 2.14, idéntica a la figura 2.12 del producto de números complejos, suponemos ahora que tenemos como datos $\overrightarrow{oz}$ y $\overrightarrow{oz_2}$ (este último, no nulo) para hallar el cociente $\frac{\overrightarrow{oz}}{\overrightarrow{oz_2}} = \overrightarrow{oz_1}$, basta construir sobre el segmento $\overline{oU}$ como homólogo del $\overline{oz_2}$, el triángulo $\overset{\Delta}{\overbrace{oUz_1}}$ semejante al $\overset{\Delta}{\overbrace{oz_2z}}$. El vector $\overrightarrow{oz_1}$ representa el cociente.

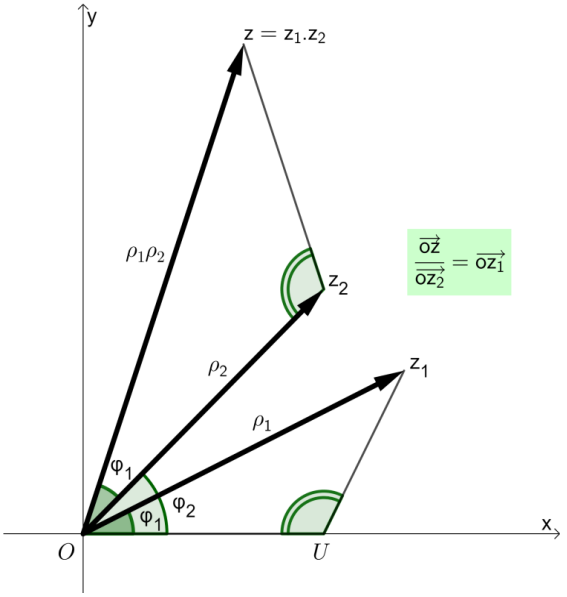


Figura 2.14. Cociente de números complejos.

Por construcción, el argumento de z_1 es φ_1 en cuanto a su módulo que es la longitud de $\overrightarrow{oz_1}$, por la semejanza de triángulos, se tiene:

$\frac{\overline{oz}}{\overline{oz_2}} = \frac{\overline{oz_1}}{\overline{oU}} \implies \frac{\rho_1\rho_2}{\rho_2} = \frac{\overline{oz_1}}{1}$ por lo tanto

$\overline{oz_1} = \rho_1$

5.7. Potenciación de números complejos

Sean los complejos $z_1 = \rho_1(cos\,\varphi_1 + i\,sen\,\varphi_1)$ y $z_2 = \rho_2(cos\,\varphi_2 + i\,sen\,\varphi_2)$ un caso particular del producto de los complejos z_1 y z_2 se presenta cuando ambos son iguales. En tal situación, la expresión $z_1\,z_2 = \rho_1\,\rho_2[cos(\varphi_1 + \varphi_2) + i\,sen\,(\varphi_1 + \varphi_2)]$ se puede escribir como $z_1\,z_2 = \rho_1^2\,[cos(2\varphi_1) + i\,sen\,(2\varphi_1)]$.

Este resultado se generaliza mediante el siguiente teorema de De Moivre.

La potencia n–ésima ($n \in \mathbb{N}$*) de un complejo dado en forma polar es otro complejo cuyo módulo es la potencia n–ésima del módulo del complejo dado y cuyo argumento es* n *veces el argumento del complejo dado.*

Si $z = \rho(cos\,\varphi + i\,sen\,\varphi) \in \mathbb{C} \implies z^n = \rho^n(cos\,n\varphi + i\,sen\,n\varphi)$ (1) para todo $n \in \mathbb{N}$

Esta expresión es conocida como fórmula de De Moivre y permite calcular trigonométricamente las potencias de exponente natural de los números complejos.

Demostraremos este teorema por inducción matemática.

1°) Debemos verificar que la proposición (1) es válida para el primer elemento del conjunto con el que se trabaja. En nuestro caso son los $\mathbb{N}$, cuyo primer elemento es 1.

Para $n = 1$

$Si\ z^n = \rho^n(cos\,n\varphi + i\,sen\,n\varphi) \implies$

$z^1 = \rho^1(cos\,1\varphi + i\,sen\,1\varphi)$

$z = \rho(cos\,\varphi + i\,sen\,\varphi)$

Por lo tanto (1) es válida para $n = 1$.

2°) Debemos suponer que la proposición es verdadera para $n = h$, esto es lo que constituye la *Hipótesis Inductiva*. Es decir que:

$z^h = \rho^h(cos\,h\,\varphi + i\,sen\,h\,\varphi)$ *es verdadera.*

3°) Debemos demostrar que la proposición es verdadera para $n = h + 1$. Es decir que:

$z^{h+1} = \rho^{h+1}[cos(h+1)\varphi + i\,sen\,(h+1)\,\varphi]$ *es verdadera.*

$z^{h+1} = z^h\ z$ *por propiedades de la potenciación*

Dado que $z^h = \rho^h(cos\ h\ \varphi + i\ sen\ h\ \varphi)$ por hipótesis inductiva y además sabemos que $z^1 = \rho(cos\ \varphi + i\ sen\ \varphi)$ entonces:

$z^{h+1} = \rho^h(cos\ h\ \varphi + i\ sen\ h\ \varphi)\ \rho(cos\ \varphi + i\ sen\ \varphi)$

$z^{h+1} = \rho^h\rho\ [cos(h\varphi + \varphi) + i\ sen\ (h\varphi + \varphi)\ ]$

$z^{h+1} = \rho^{h+1}\ [cos(h+1)\varphi + i\ sen\ (h+1)\ \varphi]$

4°) Concluimos que $z^n = \rho^n(cos\ n\varphi + i\ sen\ n\varphi)$ es válida para todo $n \in \mathbb{N}$

Ejemplo: sea $z = \sqrt{2}(\cos 45° + i\ sen\ 45°)$, calcular z^8

Si $z = \sqrt{2}(\cos 45° + i\ sen\ 45°) \implies z^8 = \left(\sqrt{2}\right)^8(\cos 8\,.\,45° + i\ sen\ 8\,.\,45°)$

$z^8 = 2^4(\cos 360° + isen\ 360°)$

$z^8 = 16$

Se puede trabajar, también, de la siguiente manera.

$z = \sqrt{2}(\cos 45° + i\ sen\ 45°)$

$z = \sqrt{2}_{45°}$

Si $z = \sqrt{2}_{45°} \implies z^8 = \left(\sqrt{2}_{45°}\right)^8$

$z^8 = \left(\sqrt{2}^8\right)_{8\,.45°}$

$z^8 = 16_{360°}$

$z^8 = 16_{0°}$

La fórmula de De Moivre es válida, aunque sea nulo o negativo el exponente n. Es decir que la fórmula vale para todo exponente entero ($n \in \mathbb{Z}$).

Lo anterior se puede demostrar de la siguiente manera:

Si $z = \rho(cos\ \varphi + i\ sen\ \varphi) \implies z^{-1} = [\rho(cos\ \varphi + i\ sen\ \varphi)]^{-1}$

$$z^{-1} = \frac{1}{\rho(\cos\varphi + i\ sen\ \varphi)}$$

$$z^{-1} = \frac{1}{\rho}\left(\frac{\cos\varphi - i\ sen\ \varphi}{(\cos\varphi + i\ sen\ \varphi)(\cos\varphi - i\ sen\ \varphi)}\right)$$

$$z^{-1} = \frac{1}{\rho}\left(\frac{\cos\varphi - i\ sen\ \varphi}{\cos^2\varphi + sen^2\varphi}\right)$$

$$z^{-1} = \frac{1}{\rho}(\cos\varphi - i\ sen\ \varphi)$$

$$z^{-1} = \rho^{-1}(\cos\varphi - i\ sen\ \varphi)$$

Pero sabemos por las relaciones entre las funciones trigonométricas de ángulos opuestos que $\cos\varphi = \cos(-\varphi)\ y\ \ sen\ \varphi = -sen\ (-\varphi)$. Reemplazando en la última expresión

$z^{-1} = \rho^{-1}[\cos(-\varphi) + i\ sen\ (-\varphi))$ elevando ambos miembros a la n–ésima potencia

$$(z^{-1})^n = \{\rho^{-1}[\cos(-\varphi) + i\ sen\ (-\varphi))\ \}^n$$

$$z^{-n} = \rho^{-n}(\cos(-n\varphi) + i\ sen\ (-n\varphi)$$

Ejemplo: sea $z = -1 + \sqrt{3}\ i$, calcular z^{-6}.

Expresamos el complejo z, en forma polar.

$\rho = \sqrt{(-1)^2 + \left(\sqrt{3}\right)^2} = 2$ y $\varphi = arcotg - \frac{\sqrt{3}}{1} = -60° = 300°$ pero z caracteriza a un punto del II cuadrante, por lo tanto $\varphi = 120°$. Luego $z = 2_{120°}$

Realizamos la potenciación propuesta.

$$z^{-6} = (2_{120°})^{-6}$$

$$z^{-6} = 2^{-6}_{-6.120°}$$

$$z^{-6} = \left(\frac{1}{64}\right)_{-720°}$$

$$z^{-6} = \left(\frac{1}{64}\right)_{0°}$$

Interpretación geométrica de la potenciación de números complejos

Reiterando idéntico procedimiento explicado para representar gráficamente el producto de dos complejos, en las potencias de exponente positivo de un complejo ρ_{φ} se van obteniendo valores cuyos módulos van en progresión geométrica[16] $\rho, \rho^2, \rho^3, \ldots, \rho^n$ y sus argumentos en progresión aritmética $\varphi, 2\varphi, 3\varphi, \ldots, n\varphi$.

En la Figura 2.15 hemos representado las ocho primeras potencias del número complejo $z = (1,2)_{20°}$.

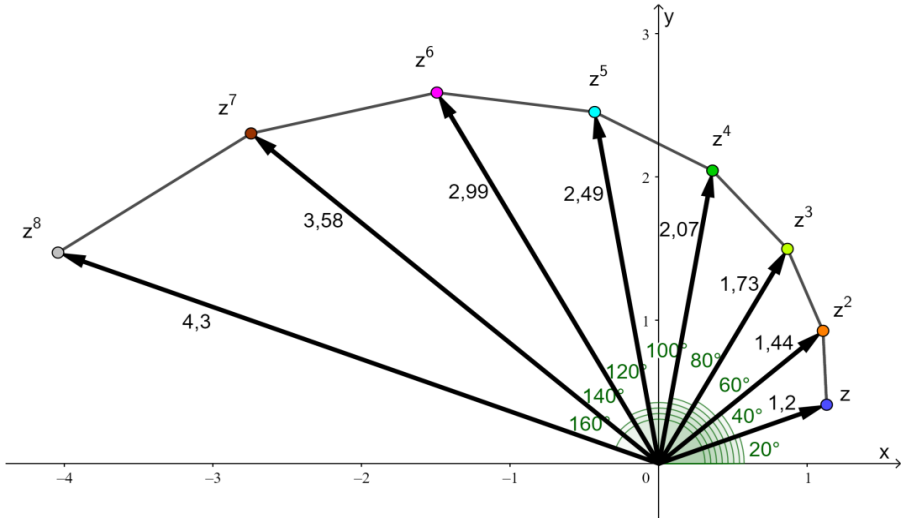


Figura 2.15. Potencia de números complejos.

[16] Sucesión es una función cuyo dominio son los ℕ y su codominio es cualquier otro conjunto de números de diferente naturaleza, generalmente los ℝ. Una sucesión se dice aritmética si la diferencia entre dos términos consecutivos cualesquiera es una constante y se dice geométrica si cada término es igual al anterior multiplicado por una constante.

5.8. Radicación de números complejos

Recordemos que en los números reales, la raíz n–ésima ($n \in \mathbb{N}$) de un número a es otro número b tal que la n–ésima potencia de b es a. Es decir que $\sqrt[n]{a} = b \iff b^n = a$. Por ejemplo $\sqrt[3]{64} = 4$ ya que $4^3 = 64$. De manera similar en los complejos, la raíz n–ésima de z_1 es otro complejo z_2 tal que $(z_2)^n = z_1$. Esto nos conduce a la siguiente

Definición: sean los n^{os} complejos expresados como $z_1 = \rho_1(cos\,\varphi_1 + i\,sen\,\varphi_1) \in \mathbb{C}$ y $z_2 = \rho_2(cos\,\varphi_2 + i\,sen\,\varphi_2) \in \mathbb{C}\ y\ n \in \mathbb{N}$ se dice que z_2 es raíz n–ésima de z_1 si y solo si $(z_2)^n = z_1$

Si $(z_2)^n = z_1$ entonces

$[\rho_2(cos\,\varphi_2 + i\,sen\,\varphi_2)]^n = \rho_1(cos\,\varphi_1 + i\,sen\,\varphi_1)$

$\rho_2^n(cos\ n\,\varphi_2 + i\,sen\,n\,\varphi_2) = \rho_1(cos\,\varphi_1 + i\,sen\,\varphi_1)$ por lo tanto

$\rho_2^n = \rho_1 \implies \rho_2 = \sqrt[n]{\rho_1}$

$n\,\varphi_2 = \varphi_1 + 2\,k\,\pi \implies \varphi_2 = \frac{\varphi_1 + 2\,k\,\pi}{n} = \frac{\varphi_1}{n} + \frac{2\,k\,\pi}{n} \quad con\ k \in \mathbb{Z}$

En consecuencia

$$\sqrt[n]{\rho(cos\,\varphi + i\,sen\,\varphi)} = \sqrt[n]{\rho}\left(cos\,\frac{\varphi+2\,k\,\pi}{n} + i\,sen\,\frac{\varphi+2\,k\,\pi}{n}\right) \quad con\ k \in \mathbb{Z}$$

En la fórmula anterior, el módulo de la raíz buscada está perfectamente determinado, hallando la raíz n–ésima de ρ, pero no ocurre lo mismo para el argumento, pues como está en función de k ∈ ℤ resultarían infinitos valores para la raíz. Sin embargo, esto no ocurre ya que el número de raíces queda perfectamente determinado si damos a k los n primeros valores:

$k = 0 \implies \frac{\varphi}{n}$

$k = 1 \implies \frac{\varphi}{n} + \frac{2\pi}{n}$

$k = 2 \implies \frac{\varphi}{n} + 2\frac{2\pi}{n}$

$k = 3 \implies \frac{\varphi}{n} + 3\frac{2\pi}{n}$

. . .

$k = (n-1) \implies \frac{\varphi}{n} + (n-1)\frac{2\pi}{n}$

$k = n \implies \frac{\varphi}{n} + 2\pi$ que es congruente con el valor obtenido en k = 0

$k = n+1 \implies \frac{\varphi}{n} + (n+1)\frac{2\pi}{n}$ que es congruente con el valor obtenidos en k = 1

Así sucesivamente,

Todo número complejo no nulo $\rho(\cos\varphi + i\,sen\,\varphi)$ *tiene n raíces n–ésimas de módulo* $\sqrt[n]{\rho}$ *y cuyos argumentos mínimos se obtienen de la expresión* $\varphi_k = \frac{\varphi}{n} + \frac{2\,k\,\pi}{n}$ *al dar valores a* k *desde* 0 *hasta* $(n-1)$.

$\sqrt[n]{\rho(\cos\varphi + i\,sen\,\varphi)} = \sqrt[n]{\rho}\left(\cos\frac{\varphi+2\,k\,\pi}{n} + i\,sen\,\frac{\varphi+2\,k\,\pi}{n}\right)$ $con\ k = 0, 1. \ . \ .(n-1)$

Ejemplo: calcular $\sqrt[3]{8(\cos 30° + i\,sen\,30°)}$

Nos piden calcular $\sqrt[3]{8_{30°}}$, como $2\pi \equiv 360°$, entonces:

$z_k = \left(\sqrt[3]{8}\right)_{\frac{30°}{3}+\frac{k\,360°}{3}}$ para $k = 0, 1, 2$

$k = 0 \implies z_0 = 2_{10°}$

$k = 1 \implies z_1 = 2_{10°+\frac{360°}{3}} = 2_{10°+120°} = 2_{130°}$

$k = 2 \implies z_2 = 2_{10°+\frac{720°}{3}} = 2_{10°+240°} = 2_{250°}$

Interpretación geométrica de las raíces

Para representar gráficamente las raíces, trazamos una circunferencia de centro o y radio $\sqrt[n]{\rho}$, a esta circunferencia pertenecerán las imágenes de las n raíces. Una de las imágenes (punto z_0) se fija considerando a partir del origen $\frac{\varphi}{n}$ (argumento dado dividido en n) y los restantes puntos $z_1, z_2, \ldots, z_{n-1}$ se obtienen dividiendo la circunferencia a partir de z_0 en n partes iguales. Es decir que las n raíces n–ésimas, distintas de un complejo no nulo ($n \geq 3$), se identifican con los vértices de un polígono regular de n lados inscriptos en la circunferencia de radio $\sqrt[n]{\rho}$

Continuando con el ejemplo anterior, la representación gráfica de las raíces calculadas se muestra en la Figura 2.16.

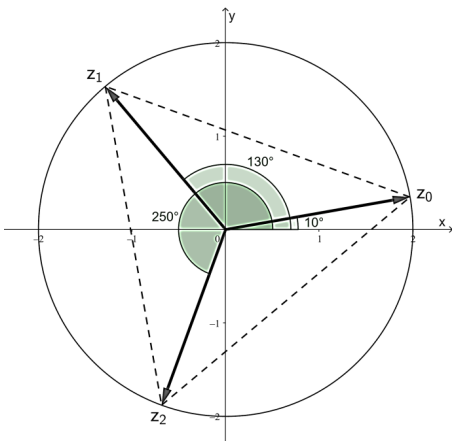


Figura 2.16. Raíces de un número complejo.

6. Forma exponencial de los números complejos

Como se dijo anteriormente, existe una cuarta forma de expresar un número complejo. Veamos cuál es esta forma.

Leonhard Euler, padre de la unidad imaginaria y creador de notaciones matemáticas modernas, introdujo la expresión: $e^{i\varphi} = cos\,\varphi + i\,sen\,\varphi$ [17], llamada *Fórmula de Euler*, la cual define el símbolo $e^{i\varphi}$ para cualquier valor real de φ, es decir φ debe estar expresado en radianes. Esta nueva forma permite expresar a cualquier número complejo $z = a + bi$ en una manera mucho más compacta, que es conocida como forma exponencial $z = \rho e^{i\varphi}$, donde ρ es el módulo del complejo z y φ es el argumento de z.

Por convención establecemos que la forma exponencial para el complejo nulo es 0. Destacando que, particularmente, la expresión $e^{i\varphi}$ o bien su equivalente $cos\,\varphi + i\,sen\,\varphi$ representa un complejo de módulo unitario y dirección φ.

La ventaja de la notación introducida por Euler es que, al realizar las operaciones: producto, cociente y potencias enteras de complejos expresados en forma exponencial, se preservan las propiedades de la potenciación (producto y cociente de potencias de igual base y potencia de potencia).

Lo primero que podemos plantearnos es cómo determinar si dos complejos no nulos y expresados en forma exponencial son iguales.

[17] e: número irracional, base de los logaritmos neperianos.

Igualdad de complejos: evidentemente dos números complejos son iguales si tienen igual módulo y sus argumentos deben diferir en un múltiplo de 2π. Para evidenciar esto, notemos primero que para cualquier $\varphi \in \mathbb{R}$, se tiene que $e^{i\,\varphi} = 1$ y esto es si y solo si $\exists k \in \mathbb{Z} \,/\, \varphi = 2k\pi$. Por lo tanto, sean los números complejos no nulos, denominados: $z_1 = \rho_1 e^{i\varphi_1}\ y\ \ z_2 = \rho_2 e^{i\varphi_2}$

$z_1 = z_2 \Leftrightarrow \rho_1 e^{i\varphi_1} = \rho_2 e^{i\varphi_2} \Leftrightarrow \rho_1 = \rho_2 \wedge\ e^{i\varphi_1} = e^{i\varphi_2} \Leftrightarrow \rho_1 = \rho_2 \wedge$
$e^{i\,(\varphi_1 - \varphi_2)} = 1$

$z_1 = z_2 \Leftrightarrow \rho_1 e^{i\varphi_1} = \rho_2 e^{i\varphi_2} \Leftrightarrow \rho_1 = \rho_2 \wedge \exists k \in \mathbb{Z} \,/\, (\varphi_1 - \varphi_2) = 2k\pi$

Ejemplo: $3e^{i\frac{1}{6}\pi},\ 3\,e^{i\frac{13}{6}\pi}\ \ y\ \ 3\,e^{i\frac{25}{6}\pi}$ representan el mismo número complejo.

Como las formas trigonométricas y exponencial son equivalentes, las operaciones de números complejos que pueden realizarse con la forma trigonométrica son las mismas que en la forma exponencial; la diferencia radica en que en esta última forma se utilizan las propiedades de la potenciación para calcular los argumentos y que es indispensable utilizar radianes en vez de grados sexagesimales.

Producto: sean dos n^{os} complejos cualesquiera $z_1 = \rho_1 e^{i\varphi_1} = \rho_1(cos\,\varphi_1 + i\ sen\,\varphi_1)$ y $z_2 = \rho_2 e^{i\varphi_2} = \rho_2(cos\,\varphi_2 + i\ sen\,\varphi_2)$ no nulos, resulta:

$z_1 z_2 = \rho_1 e^{i\varphi_1}\ \rho_2 e^{i\varphi_2} = \rho_1(cos\,\varphi_1 + i\ sen\,\varphi_1)\ \rho_2(cos\,\varphi_2 + i\ sen\,\varphi_2)$
$= \rho_1\ \rho_2((cos\,\varphi_1\,cos\,\varphi_2 - sen\,\varphi_1 sen\,\varphi_2) + (\cos\varphi_1\ sen\,\varphi_2 +$
$\cos\varphi_2\ sen\,\varphi_1)i)$
$= \rho_1\ \rho_2\ (\cos(\varphi_1 + \varphi_2) + sen\,(\varphi_1 + sen\,\varphi_2)i)$
$= \rho_1\ \rho_2\ e^{i\,(\varphi_1 + \varphi_2)}$

El producto de dos complejos expresados en forma exponencial es otro complejo que tiene por módulo el producto de los módulos de los complejos dados y por argumento la suma de los argumentos de los factores.

Si alguno de los complejos es el nulo, se obtiene como resultado el complejo nulo.

Ejemplo 1: sean los complejos $z_1 = 6\, e^{i\,\frac{1}{3}\pi}$ y $z_2 = 3\, e^{i\,\frac{5}{3}\pi}$, entonces

$z_1 z_2 = 6\, e^{i\,\frac{1}{3}\pi}\ 3\, e^{i\,\frac{5}{3}\pi}$

$z_1 z_2 = 6\,.\,3\, e^{i\left(\frac{1}{3}+\frac{5}{3}\right)\pi}$

$z_1 z_2 = 18\, e^{i\,2\pi}$

Ejemplo 2: realizar la siguiente multiplicación $\left(2\, e^{i\frac{1}{9}\pi}\right)\left(e^{i\frac{1}{6}\pi}\right)\left(2\, e^{i\frac{7}{18}\pi}\right) =$

$\left(2\, e^{i\frac{1}{9}\pi}\right)\left(e^{i\frac{1}{6}\pi}\right)\left(2\, e^{i\frac{7}{18}\pi}\right) =$

$= 2\,.1\,.2\, e^{i\left(\frac{1}{9}+\frac{1}{6}+\frac{7}{18}\right)\pi}$

$= 4\, e^{i\frac{2}{3}\pi}$

Conjugado: sea el complejo $z = \rho e^{i\,\varphi} = \rho(cos\,\varphi + i\ sen\,\varphi)$, entonces:

$\overline{z} = \rho(\cos\varphi - i\ sen\,\varphi)$

$\overline{z} = \rho(\cos(-\varphi) + i\ sen\,(-\varphi))$ de donde se puede ver que

$\overline{\rho e^{i\,\varphi}} = \rho e^{i\,(-\varphi)}$

Del mismo modo que en el apartado "Conjugación de números complejos", se evidencia aquí que el módulo del conjugado es igual al módulo del complejo dado y el argumento del conjugado es el opuesto del argumento.

Ejemplo: sea el complejo $z = 6\ cos^{i\,\frac{1}{3}\pi} \implies \overline{z} = 6\ cos^{i\left(-\frac{1}{3}\pi\right)} = 6\, e^{i\,\frac{5}{3}\pi}$

Inverso multiplicativo: sea el complejo no nulo $z = \rho e^{i\,\varphi} = \rho(cos\,\varphi + i\,sen\,\varphi)$. Recordemos que en el apartado "Conjugación de números complejos" se demostró la propiedad: "El producto de dos complejos conjugados es igual a un número real no negativo, igual al cuadrado de su módulo", es decir que si $z \in \mathbb{C}$ se cumple que el producto $z\,\overline{z} = \|z\|^2$. Por lo tanto

$z\,\frac{\overline{z}}{\|z\|^2} = 1$ el inverso multiplicativo de todo número complejo es único, por lo tanto

$z^{-1} = \frac{\overline{z}}{\|z\|^2}$

Como el complejo está dado en forma exponencial, su inverso multiplicativo será

$z^{-1} = \frac{\rho e^{i\,(-\varphi)}}{\rho^2}$

$z^{-1} = \rho^{-1} e^{i\,(-\varphi)}$

Nuevamente, como en el apartado "Multiplicación de números complejos" se puede observar que el módulo del inverso de z es el inverso del módulo de z, el inverso tiene la misma dirección que el conjugado. Este hecho permite ubicar geométricamente el punto que representa a z^{-1} a partir del punto que representa a z.

Ejemplo: si $z = 2\,e^{i\,\frac{1}{4}\pi} \implies z^{-1} = \frac{1}{2} e^{i\left(-\frac{1}{4}\pi\right)} = \frac{1}{2}\,e^{i\frac{7}{4}\pi}$

Representamos $z\,(2_{45°})$, representamos $\overline{z}\,(2_{315°})$ punto simétrico respecto al $\overrightarrow{ox}$. Multiplicamos $\overline{z}$ por el escalar $\lambda = \frac{1}{\rho^2}\left[\frac{1}{4}.\,2_{315°} = \left(\frac{1}{2}\right)_{315°}\right]$. El

resultado es la representación de z^{-1}. Tal como se muestra en la Figura 2.17.

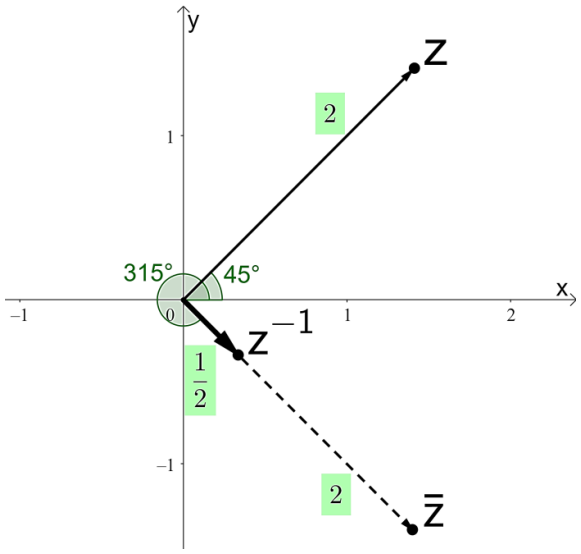


Figura 2.17. Inverso multiplicativo.

División: teniendo las expresiones para el producto y el inverso en forma exponencial, la división es directa. Es decir, sean los números complejos $z_1 = \rho_1 e^{i\varphi_1}$ y $z_2 = \rho_2 e^{i\varphi_2}$ este último no nulo, resulta:

$$\frac{z_1}{z_2} = z_1\, z_2^{-1} = \rho_1 e^{i\varphi_1} \rho_2^{-1} e^{i(-\varphi_2)} = \frac{\rho_1}{\rho_2}\, e^{i\,(\varphi_1 - \varphi_2)}$$

Por lo tanto, el cociente de dos números complejos expresados en forma exponencial tiene como módulo el cociente de los módulos y como argumento la diferencia de los argumentos de los números dados (si el complejo del numerador es 0, el argumento queda indeterminado).

Ejemplo: sean los complejos $z_1 = 6\, e^{i\frac{1}{3}\pi}$ y $z_2 = 3\, e^{i\frac{5}{3}\pi}$, entonces

$$\frac{z_1}{z_2} = \frac{6\, e^{i\frac{1}{3}\pi}}{3\, e^{i\frac{5}{3}\pi}} = \frac{6}{3} e^{i\left(\frac{1}{3}-\frac{5}{3}\right)\pi} = 2e^{i\left(-\frac{4}{3}\right)\pi} = 2e^{i\frac{2}{3}\pi}$$

Veamos, ahora, cómo la forma exponencial permite realizar fácilmente la potenciación de números complejos, cuando el exponente es un número entero.

Potencia entera: sea el complejo $z = \rho e^{i\,\varphi}$ y n un número entero ($con\ z \neq 0\ si\ n \leq 0$).

a) Para valores positivos de n, considerando la expresión para el producto en forma exponencial y de forma análoga a la realizada en el apartado "Potenciación de números complejos", se comprueba fácilmente por inducción que $z^n = \rho^n\, e^{i\,n\,\varphi}$.

b) Para valores de n, enteros negativos, es decir $n = -k$ con $k > 0$, es:

$$z^n = z^{-k} = (z^{-1})^k = \left(\rho^{-1}\, e^{-i\,\varphi}\right)^k = \rho^{-k}\, e^{i(-k)\varphi} = \rho^n\, e^{i\,n\,\varphi}$$

c) Si $n = 0$, entonces $z^0 = 1 = \rho^0\, e^{i\,0}$

En síntesis, para cualquier exponente entero, se verifica: $z^n = \rho^n\, e^{i\,n\,\varphi}$.

La radicación se realiza con las mismas consideraciones realizadas en el apartado "Radicación de números complejos" teniendo en cuenta que cada número complejo tiene n raíces n-ésimas.

Radicación: sean $z_1 = \rho_1 e^{i\varphi_1} \in \mathbb{C}\ \ y\ \ \rho_2 e^{i\varphi_2} z_1 \in \mathbb{C}\ \ y\ \ n \in \mathbb{N}$ se dice que z_2 es raíz n–ésima de z_1 si $(z_2)^n = z_1$ es decir que:

$$\sqrt[n]{z} = \sqrt[n]{\rho\, e^{i\,\varphi}} = \sqrt[n]{\rho}\sqrt[n]{e^{i\,\varphi}} = \sqrt[n]{\rho}\, e^{i\frac{\varphi+2k\pi}{n}}\ \ \forall\ k = 0, 1, 2, \ldots, (n-1)$$

Debemos recordar nuevamente que, en el caso de la radicación en forma exponencial también el ángulo debe ser expresado e radianes, φ debe ser un número real.

Ejemplo 1: sea $z = 2\ e^{i\frac{1}{2}\pi}$, calcular z^6

$z^6 = \left(2\ e^{i\frac{1}{2}\pi}\right)^6$

$z^6 = 2^6 e^{i\ 6\left(\frac{1}{2}\pi\right)}$

$z^6 = 64\ e^{i\ 3\pi}$

$z^6 = 64\ e^{i\ \pi}$

Ejemplo 2: sea $z = 243\ e^{i\frac{3}{2}\pi}$, calcular $\sqrt[5]{z}$

$\sqrt[5]{z} = \sqrt[5]{243\ e^{i\frac{3}{2}\pi}}$

$\sqrt[5]{z} = \sqrt[5]{243}\ e^{i\frac{1}{5}\left(\frac{3}{2}\pi+2k\pi\right)}$

$\sqrt[5]{z} = 3\ e^{i\frac{3+4k}{10}\pi}$

$z_0 = 3\ e^{\frac{3}{10}\pi},\ z_1 = 3\ e^{\frac{7}{10}\pi},\ z_2 = 3\ e^{\frac{11}{10}\pi},\ z_3 = 3\ e^{\frac{15}{10}\pi},\ z_4 = 3\ e^{\frac{19}{10}\pi}$

Ejemplo 3: sea el complejo $z = cos\ 120° + i\ sen\ 120°$, calcular $z^{\frac{3}{5}}$

$z = 1\ e^{i\frac{2}{3}\pi}$ nos piden calcular $\sqrt[5]{z^3}$

$w = \sqrt[5]{z^3}$

$w = \sqrt[5]{\left(1\ e^{i\frac{2}{3}\pi}\right)^3}$

$w = \sqrt[5]{1^3\ e^{i\ 3\frac{2}{3}\pi}}$

$w = \sqrt[5]{1\ e^{i\ 2\pi}}$

$w = \sqrt[5]{1}$

$$w = \sqrt[5]{1\, e^{i\, 0\pi}}$$

$$w = \sqrt[5]{1} e^{i\frac{1}{5}(0\pi+2k\pi)}$$

$$w = 1e^{i\,\frac{2k}{5}\pi}$$

$$w_0 = 1\, e^0 = 1,\ \ w_1 = 1\, e^{i\frac{2}{5}\pi},\ \ w_2 = 1\, e^{i\frac{4}{5}\pi},\ \ w_3 = 1\, e^{i\frac{6}{5}\pi},\ \ w_4 = 1\, e^{i\frac{8}{5}\pi}$$

CAPÍTULO 3

POLINOMIOS

1. Introducción

La resolución de ecuaciones algebraicas, o la determinación de las raíces de un polinomio, está entre los problemas más antiguos de la matemática. Se han encontrado problemas relacionados con este tema en el libro chino *Chu Chang Suan Shu* o *Arte Matemático en Nueve Secciones* que data del año 200 a.C.

Puesto que las ecuaciones polinómicas aparecen en un amplio rango de áreas de la ciencia, desde química y física básica hasta economía, el problema de determinar raíces de polinomios es, frecuentemente, un paso obligado en la resolución de problemas. Por ejemplo, la *optimización* es una técnica matemática que permite usar, de manera eficiente, recursos escasos como el tiempo, la energía o el dinero, bajo la guía de determinados objetivos. Las empresas la emplean para tomar decisiones respecto a la conveniencia o no de, por ejemplo, contratar más empleados, remodelar la empresa, comprar más productos para su posterior venta o simplemente dejar el dinero en el banco. Para establecer la estrategia óptima, se resuelve una serie de ecuaciones polinómicas que reflejan cuánta inversión y cuánto beneficio se asocia a cada acción. Las estrategias

óptimas se asocian, generalmente, con las raíces de las ecuaciones propuestas.

Cuando se trata de hallar las raíces de un polinomio, se debe resolver la ecuación correspondiente (el polinomio igualado a cero). En algunos casos, esta tarea es sencilla (polinomios lineales o cuadráticos), pero en otros no es nada fácil. Para estos últimos, existen métodos útiles que permiten, si no determinar, *aproximar* las raíces buscadas.

Desde el siglo XVI se conocen fórmulas para determinar las raíces de polinomios de hasta cuarto grado en términos de los coeficientes del polinomio, con la utilización de operaciones algebraicas o la extracción de raíces –a partir de los de segundo grado, las fórmulas se complican un poco– pero, las fórmulas para polinomios de quinto grado fueron irresolubles para los investigadores durante mucho tiempo. Recién, en 1824, Niels Henrik Abel demostró que no puede haber fórmulas generales para los polinomios de quinto grado o mayores. Este resultado marcó el camino hacia el teorema de Galois.

En este tercer capítulo estudiaremos, luego de las generalidades de polinomios, las operaciones que se pueden realizar entre ellos y sus correspondientes propiedades. Se demostrarán teoremas y analizarán distintas situaciones que permitirán determinar las raíces de polinomios.

1.1. Conceptos básicos

Si bien los polinomios pueden ser considerados como un tipo especial de funciones, estos tienen una estructura que permite realizar operaciones

entre ellos y tratarlos como objetos algebraicos. Particularmente estudiaremos los polinomios de una sola indeterminada, y designaremos a la misma con la letra x.

Definición: *un polinomio, $P(x)$, en una indeterminada, x, con coeficientes en $\mathbb{C}$ es una expresión de la forma:*

$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ ... + a_2 x^2 + a_1 x + a_0$ ó escrito en su forma sintética

$P(x) = \sum_{i=0}^{n} a_i x^i$

Los a_i, llamados coeficientes del polinomio $P(x)$, son números complejos. Si los a_i son todos números reales, se dice que $P(x)$ es un polinomio con coeficientes en $\mathbb{R}$; y si todos los a_i son racionales, se dice que $P(x)$ es un polinomio con coeficientes en $\mathbb{Q}$, de igual manera para el caso de coeficientes enteros.

Al coeficiente a_n se le da el nombre de coeficiente principal (y al término que lo contiene: término principal), mientras que al a_0 el de término independiente. Cuando el coeficiente principal es 1, es decir $a_n = 1$, se dice que $P(x)$ es un polinomio mónico.

El conjunto de todos los polinomios con coeficientes complejos se denota $\mathbb{C}[x]$, el de los polinomios con coeficientes reales, $\mathbb{R}[x]$, el de polinomios con coeficientes racionales, $\mathbb{Q}[x]$. Obviamente, $\mathbb{Q}[x] \subset \mathbb{R}[x] \subset \mathbb{C}[x]$.

Vamos a usar la letra $\mathbb{K}$ para representar a cualquiera de los conjuntos numéricos $\mathbb{Q}$, $\mathbb{R}$ o $\mathbb{C}$, de forma que $\mathbb{K}[x]$ representará el conjunto de todos los polinomios con coeficientes en $\mathbb{K}$. El beneficio de hacer esto es que

vamos a poder enunciar (usando $\mathbb{K}$) resultados de manera más simple en lugar de tener que escribir un enunciado para cada conjunto numérico.

Ejemplos:

$$S(x) = 3x^3 - \sqrt{5}x^2 + 1$$
$$Q(x) = (2 - 3i)x^5 + 3x^2 + (1 - 4i)$$
$$R(x) = 2x^4 - 3x^3 + x^2 - 5x + 2$$

El polinomio $S(x)$ es de coeficientes reales, mientras que el polinomio $Q(x)$ es de coeficientes complejos y el $R(x)$ es de coeficientes enteros.

Sea el polinomio $P(x) = \sum_{i=0}^{n} a_i x^i$, se considerará $a_i = 0$ si el término en x^i no se encuentra en $P(x)$.

En $S(x)$ del ejemplo anterior $a_1 = 0$ y en $Q(x)$ $a_4 = a_3 = a_1 = 0$

En general pueden presentarse dos casos:

a) Existe al menos un coeficiente a_i no nulo. Por lo tanto, llamaremos grado del polinomio $P(x)$ y se simbolizará $gr(P(x))$, al mayor entero i tal que a_i es no nulo.

b) Todos los coeficientes son nulos. En ese caso el polinomio $P(x)$ es el polinomio nulo.

Continuando con el ejemplo anterior: $gr\big(\mathrm{S}(x)\big) = 3$, $gr\big(\mathrm{Q}(x)\big) = 5$ y $gr\big(\mathrm{R}(x)\big) = 4$.

Otros polinomios importantes son los polinomios constantes, de la forma a_0x^0 donde $a_0 \in \mathbb{K}$, son de grado 0 si $a_0 \neq 0$. Si identificamos el elemento $a_0 \in \mathbb{K}$ con el polinomio contante $a_0x^0 \in \mathbb{K}[x]$, podemos pensar que $\mathbb{K} \subset \mathbb{K}[x]$. Por ejemplo, $T(x) = 15$.

Convenimos en que el polinomio nulo carece de grado, aunque algunos autores le atribuyen grado –1 y en otros casos le asignan grado infinito.

¿Cuándo dos polinomios son iguales? Consideramos que dos polinomios son iguales cuando tienen el mismo grado y sus coeficientes coinciden término a término, pero también los consideraremos iguales a polinomios que difieran en términos con coeficientes nulos.

Igualdad de polinomios: *los polinomios* $P(x) = \sum_{i=0}^{n} a_i x^i$ *y* $Q(x) = \sum_{i=0}^{m} b_i x^i$, *son iguales si* $a_i = b_i$ para $0 \leq i \leq mín(n, m)$ y si $a_i = 0$ *para* $i = m+1, m+2, \dots, n$ *en el caso que* $n > m$, ó si $b_i = 0$ *para* $i = n+1, n+2, \dots, m$ *cuando* $n < m$.

Esto significa, como se dijo anteriormente que, por ejemplo, son iguales los polinomios:

$P(x) = 3x^2 + 2x - 4$ *y* $Q(x) = 3x^2 + 2x - 4$ pero también vamos a considerar iguales a los polinomios:

$R(x) = 0x^4 + 0x^3 + 5x^2 + 8x - 2$ *y* $T(x) = 5x^2 + 8x - 2$

Esto nos permite afirmar que, en el caso del polinomio nulo, al considerar todos los $a_i = 0$, existe un único polinomio nulo en $\mathbb{K}[x]$.

Un punto importante sobre los polinomios es que se puede reemplazar la indeterminada x por un elemento de $\mathbb{K}$ y realizar las operaciones indicadas. A esto se lo llama especializar o evaluar el polinomio.

Evaluación de polinomios: *sea el polinomio* $P \in \mathbb{K}[x]$ *y el número* $b \in \mathbb{C}$, *definimos* $P(b) = \sum_{i=0}^{n} a_i b^i$ *como el elemento de* $\mathbb{K}$ *que obtenemos al reemplazar la indeterminada* x *por el número* b *y efectuar el cálculo expresado en* P.

Evidentemente $P(b) \in \mathbb{K}$, esto significa que al evaluar un polinomio se obtiene por resultado un número.

Ejemplo: sea el polinomio $P(x) = 3x^2 - 4x + 1 \ \in \mathbb{Z}[x] \ y \ b = 2$ entonces: $P(2) = 3 \,.\, 2^2 - 4 \,.2 + 1 = 5$ es decir $P(2) = 5 \ y \ P(2) \in \mathbb{Z}$.

2. Operaciones con polinomios

Como se dijo anteriormente, con los polinomios se pueden realizar operaciones tales como suma, producto e inclusive división. La realización de estas operaciones es muy sencilla. Por ejemplo, para realizar la suma (o resta) de polinomios se procede a sumar (restar) respectivamente los términos correspondientes a la misma potencia de la indeterminada.

2.1. Suma de polinomios

La suma de polinomios es una operación que se realiza en forma intuitiva, para formalizarla acudimos a la siguiente

Definición: *sean los polinomios* $P(x) = \sum_{i=0}^{n} a_i x^i \ \ y \ \ Q(x) = \sum_{i=0}^{m} b_i x^i$ *definimos la suma de estos polinomios como:* $(P + Q) = \sum_{i=0}^{máx(n,m)} (a_i + b_i) x^i$ *donde, de igual manera que la igualdad de polinomios, convenimos en que* $b_i = 0$ *para* $i = m + 1, m + 2, \dots, n$ *en el caso que* $n > m$, ó si $a_i = 0$ *para* $i = n + 1, n + 2, \dots, m$ *cuando* $n < m$.

Ejemplo: si $P(x) = 2x^2 + 3x - 5 \ y \ Q(x) = 4x^4 + x^3 - 4x^2 - 6x + 5$ entonces:

$P(x) + Q(x) = (0 + 4)x^4 + (0 + 1)x^3 + (2 - 4)x^2 + (3 - 6)x + (-5 + 5)$

$P(x) + Q(x) = 4x^4 + x^3 - 2x^2 - 3x$

De manera análoga podemos definir la resta de los polinomios $P(x)\ y\ Q(x)$ por:

$(P-Q)=\sum_{i=0}^{máx(n,m)}(a_i-b_i)x^i$ con idéntico convenio.

Ejemplo: si $P(x)=2x^2+3x-5\ y\ Q(x)=4x^4+x^3-4x^2-6x+5$ entonces:

$P(x)-Q(x)=(0-4)x^4+(0-1)x^3+(2+4)x^2+(3+6)x+(-5-5)$

$P(x)-Q(x)=-4x^4-x^3+6x^2+9x-10$

Destacamos dos cuestiones:

a) estas operaciones están definidas de tal forma que se verifica:

$(P+Q)(b)=P(b)+Q(b)\ \forall P,Q\in\mathbb{K}[x]\ \forall b\in\mathbb{K}$ y

$(P-Q)(b)=P(b)-Q(b)\ \forall P,Q\in\mathbb{K}[x]\ \forall b\in\mathbb{K}$.

b) el grado de la suma o resta de dos polinomios P y Q es menor o igual que el máximo de los grados de dichos polinomios.

$gr(P\pm Q)\leq máx\big(gr(P),gr(Q)\big)$

2.2. Producto de un polinomio por un escalar

Dado un polinomio $P\in\mathbb{K}[x]$, el producto de P por un escalar λ, es otro polinomio $\lambda\,P(x)$, el cual se obtiene multiplicando cada coeficiente de P por el escalar λ.

Sea el polinomio $P(x)=\sum_{i=0}^{n}a_ix^i$ y el escalar λ entonces:

$\lambda\,P(x)=\lambda\ \sum_{i=0}^{n}a_ix^i=\sum_{i=0}^{n}\lambda\,a_ix^i$

Ejemplo: sea $P(x)=3x^3-7x^2+x-8\ \ y\ \ \lambda=-4$ entonces:

$-4P(x)=-12x^3+28x^2-4x+32$

De acuerdo a la definición del producto de un polinomio por un escalar, si el polinomio es $P(x)$ y el escalar es:

a) $\lambda = -1$ entonces $\lambda P(x) = -P(x)$ polinomio opuesto a P.

b) $\lambda = 0$ entonces $\lambda P(x) = N$ polinomio nulo.

Propiedades de la suma y producto por un escalar

Veremos a continuación cuáles son las propiedades algebraicas que poseen las dos operaciones básicas antes definidas con polinomios. Estas propiedades son similares a las que poseen las operaciones de suma y producto de números reales.

P_1) La suma es ley de composición interna en $\mathbb{K}[x]$.

$\forall\, P, Q \in \mathbb{K}[x] \Longrightarrow (P + Q) \in \mathbb{K}[x]$

Esto significa que la suma, como está definida, es una operación que a cada par de polinomios de $\mathbb{K}[x]$ le hace corresponder un único polinomio de $\mathbb{K}[x]$.

D) Es inmediata por la definición de suma en $\mathbb{K}[x]$.

P_2) La suma es asociativa en $\mathbb{K}[x]$.

$\forall\, P, Q, R \in \mathbb{K}[x] \Longrightarrow (P + Q) + R = P + (Q + R)$

D) Sean a_i, b_i, c_i los i–ésimos coeficientes de los polinomios P, Q, R respectivamente. Por lo tanto $a_i + b_i$ es el i–ésimo coeficiente de $P + Q$ así que $(a_i + b_i) + c_i$ es el i–ésimo coeficiente de $(P + Q) + R$. Como $(a_i + b_i) + c_i = a_i + (b_i + c_i)\ \forall\, i = 1, 2, \ldots, n$ por la propiedad asociativa en $\mathbb{K}$, los coeficientes de $(P + Q) + R$ y de $P + (Q + R)$ son iguales. Concluimos entonces que $(P + Q) + R = P + (Q + R)$.

P_3) La suma es conmutativa en $\mathbb{K}[x]$.

$\forall\, P, Q \in \mathbb{K}[x] \Longrightarrow P + Q = Q + P$.

D) Sean a_i, b_i los i–ésimos coeficientes de los polinomios $P\ y\ Q$ respectivamente. Por lo tanto $a_i + b_i$ es el i–ésimo coeficiente de $P + Q$. Como $a_i + b_i = b_i + a_i\ \forall\ i = 1, 2, \dots, n$ por la propiedad conmutativa en $\mathbb{K}$, los coeficientes respectivos de $P + Q$ y de $Q + P$ son iguales. Concluimos entonces que $P + Q = Q + P$.

P_4) Existe un único elemento neutro para la suma en $\mathbb{K}[x]$.

El elemento neutro se denota con $N(x)$.

$\exists N \in \mathbb{K}[x]\ \acute{u}nico / \forall\ P \in \mathbb{K}[x] : P + N = N + P = P.$

D) La existencia es evidente, puesto que se definió el polinomio nulo como aquel polinomio cuyos coeficientes son cero.

Respecto de la unicidad del polinomio nulo, es una consecuencia de la igualdad de polinomios que se destacó en ese apartado.

P_5) Todo polinomio de $\mathbb{K}[x]$ admite un único inverso aditivo u opuesto en $\mathbb{K}[x]$.

El polinomio opuesto a uno cualquiera P, se nota por $-P$.

$\forall\ P \in \mathbb{K}[x], \exists(-P\ ) \in \mathbb{K}[x]\ \acute{u}nico\ /\ P + (-P) = -P + P = N.$

La existencia de tal polinomio es evidente puesto que, cuando se definió producto por un escalar, si el escalar es –1 al multiplicarse por cualquier polinomio, se obtiene el opuesto de dicho polinomio.

Para demostrar la unicidad del elemento opuesto vamos a suponer que para un polinomio $P\ de\ \mathbb{K}[x]$ existen dos opuestos $P_1\ \ y\ \ P_2$ por lo tanto se debe cumplir que $P + P_1\ = N$ y también que $P + P_2\ = N$ de donde se

deduce que $P_1 = P_2$. Consecuentemente, el inverso aditivo u opuesto de cada elemento de $\mathbb{K}[x]$ es único.

P6) El producto es ley de composición externa en $\mathbb{K}[x]$ con escalares en $\mathbb{C}$.
$\forall \alpha \in \mathbb{K}, \forall P \in \mathbb{K}[x] \Longrightarrow \alpha P \in \mathbb{K}[x]$.
Esto significa que el producto de un polinomio de $\mathbb{K}[x]$ por un escalar, como está definido, es una operación que a cada par polinomio–escalar, le hace corresponder un único elemento de $\mathbb{K}[x]$.
D) La existencia de tal polinomio es inmediata, por la definición de producto de un polinomio de $\mathbb{K}[x]$ por un escalar.

P7) El producto por un escalar satisface la asociatividad mixta.
$\forall \alpha, \beta \in \mathbb{K}, \forall\, P \in \mathbb{K}[x]: \alpha(\beta\, P) = (\alpha\beta)\, P$.
D) Sea a_i el i–ésimo coeficiente del polinomio P. Por lo tanto βa_i es el i–ésimo coeficiente de βP y por lo tanto $\alpha(\beta a_i)$ el i–ésimo coeficiente de $\alpha(\beta P)$. Como $\alpha(\beta a_i) = (\alpha\beta)a_i \;\; \forall\, i = 1, 2, \dots, n$ por la propiedad asociativa en $\mathbb{K}$, las componentes respectivas de $\alpha(\beta\, P)$ y de $(\alpha\beta)P$ son iguales. Concluimos entonces que $\alpha(\beta\, P) = (\alpha\beta)\, P$.

P8) El producto por un escalar es distributivo respecto a la suma en $\mathbb{K}$.
$\forall \alpha, \beta \in \mathbb{K}, \forall\, P \in \mathbb{K}[x]: (\alpha + \beta)P = \alpha P + \beta P$.
D) Sea a_i el i–ésimo coeficiente del polinomio P. Por lo tanto $(\alpha + \beta)a_i$ es el i–ésimo coeficiente de $(\alpha + \beta)P$. Como $(\alpha + \beta)a_i = \alpha a_i + \beta a_i \;\; \forall\, i = 1, 2, \dots, n$ por la propiedad distributiva del producto en $\mathbb{K}$, respecto a la suma, las componentes respectivas de: $(\alpha + \beta)P$ y de $\alpha P + \beta P$ son iguales. Entonces : $(\alpha + \beta)P = \alpha P + \beta P$.

P9) El producto por un escalar es distributivo respecto a la suma en $\mathbb{K}[x]$.

$\forall \alpha \in \mathbb{K}, \forall\, P, Q \in \mathbb{K}[x] : \alpha(\, P + Q) = \alpha P + \alpha Q.$

D) Sean a_i, b_i los i–ésimos coeficientes de los polinomios $P\, y\, Q$ respectivamente. Por lo tanto $a_i + b_i$ es el i–ésimo coeficiente de $P + Q$ como también $\alpha(a_i + b_i)$ es el i–ésimo coeficiente de $\alpha(\, P + Q\,)$. Como $\alpha(a_i + \, b_i) = \alpha\, a_i + \alpha\, b_i \; \forall\, i = 1, 2, \ldots, n$ por la propiedad distributiva del producto en $\mathbb{K}$, respecto a la suma, los coeficientes respectivos de $\alpha(\, P + Q)$ y de $\alpha P + \alpha Q$ son iguales. Concluimos entonces que $\alpha(\, P + Q) = \alpha P + \alpha Q$.

P10) La unidad en $\mathbb{K}$ es el neutro para el producto por un escalar.

$\forall\, P \in \mathbb{K}[x]\colon 1\, P = P.$

Como el conjunto de los polinomios con las operaciones definidas: suma y producto por un escalar cumple con las propiedades enunciadas, se dice que este conjunto tiene estructura de espacio vectorial.

2.3. Producto de polinomios

Para multiplicar dos polinomios se multiplica cada término de uno de ellos por cada uno de los términos del otro polinomio y luego se suman los términos semejantes.

Dados los polinomios $P(x) = \sum_{i=0}^{n} a_i x^i \;\; y \;\; Q(x) = \sum_{j=0}^{m} b_j x^j$ *entonces:*

$P(x)\; Q(x) = \left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{j=0}^{m} b_j x^j\right)$

$P(x)\; Q(x) = \sum_{i=0}^{n} \sum_{j=0}^{m} \left(a_i x^i\right)\left(b_j x^j\right)$

$P(x)\; Q(x) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j x^i x^j$

$P(x)\; Q(x) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j x^{i+j}$

Ejemplo: sean los polinomios $P(x) = -2x^5 + 5x^3 - x^2 + 5x - 4$ y $Q(x) = 5x^5 - x^2 - 12x + 1$ para multiplicarlos se suele realizar la siguiente disposición práctica:

					$-2x^5$	$+5x^3$	$-x^2$	$+5x$	-4
$\times$						$5x^5$	$-x^2$	$-12x$	$+1$
				$-2x^5$		$+5x^3$	$-x^2$	$+5x$	-4
			$+24x^6$		$-60x^4$	$+12x^3$	$-60x^2$	$+48x$	
		$+2x^7$		$-5x^5$	$+x^4$	$-5x^3$	$+4x^2$		
$-10x^{10}$	$+25x^8$	$-5x^7$	$+25x^6$	$-20x^5$					
$-10x^{10}$	$+25x^8$	$-3x^7$	$+49x^6$	$-27x^5$	$-59x^4$	$+12x^3$	$-57x^2$	$+53x$	-4

$P(x).Q(x) = -10x^{10} + 25x^8 - 3x^7 + 49x^6 - 27x^5 - 59x^4 + 12x^3 - 57x^2 + 53x - 4$

Al igual que en la suma de polinomios, la definición de producto de polinomios es consistente con la evaluación de polinomios, es decir, para se verifique que:

$(P \,.\, Q)(b) = P(b).Q(b) \;\forall\, P, Q \in \mathbb{K}[x] \;\forall\, b \in \mathbb{K}$

Una consecuencia de la definición del producto $P \,.\, Q$ es que el coeficiente principal de $P \,.\, Q$ es el producto del coeficiente principal de P por el de Q.

En particular, al estar trabajando en $\mathbb{K}$ se puede afirmar que:

$P \,.\, Q = N \iff P = N \vee Q = N$ y que $gr(P \,.\, Q) = gr(P) + gr(Q)$

Propiedades del producto de polinomios

P_1) El producto es ley de composición interna en $\mathbb{C}$.

Si $P \in \mathbb{K}[x]$ y $Q \in \mathbb{K}[x]$ entonces (P . Q) $\in \mathbb{K}[x]$

Esto significa que el producto de polinomios, como está definido, es una operación que a cada par de polinomios le hace corresponder un único elemento de $\mathbb{K}[x]$.

D) Es inmediata por la definición de producto de polinomios.

P_2) El producto es asociativo en $\mathbb{K}[x]$.

$\forall\, P, Q, R \in \mathbb{K}[x] \Longrightarrow (P.Q).R = P\,.(Q\,.R)$

D) Sean a_i, b_i, c_i los i–ésimos coeficientes de los polinomios P, Q, R respectivamente. Por lo tanto $a_i\ b_i$ es el i–ésimo coeficiente de $P\ Q$ así que $(a_i\ b_i)\ c_i$ es el i–ésimo coeficiente de $(P\ Q)\,R$. Como $(a_i\ b_i)\ c_i = a_i\ (\,b_i\ c_i)\ \forall\ i = 1, 2, \dots, n$ por la propiedad asociativa en $\mathbb{K}$, los coeficientes respectivos de $(P\ Q)R$ y de $P\ (Q\ R)$ son iguales. Concluimos entonces que $(P\ Q)\ R = P\ (Q + R)$.

P_3) El producto es conmutativo en $\mathbb{K}[x]$.

$\forall\, P, Q \in \mathbb{K}[x] \Longrightarrow P\ Q = Q\ P$.

D) Sean a_i, b_i los i–ésimos coeficientes de los polinomios $P\ y\ Q$ respectivamente. Por lo tanto $a_i\ b_i$ es el i–ésimo coeficiente de $P\ Q$. Como $a_i\ \ b_i = b_i\ a_i\ \forall\ i = 1, 2, \dots, n$ por la propiedad conmutativa en $\mathbb{K}$, los coeficientes respectivos de $P\ Q$ y de $Q\ P$ son iguales. Concluimos entonces que $P\ Q = Q\ P$.

P_4) Existe un único elemento neutro para el producto en $\mathbb{K}[x]$.

El elemento neutro para el producto es el polinomio $P(x) = 1x^0 = 1$

$\forall\, P \in \mathbb{K}[x],\ P \neq N : \exists\, N_e \in \mathbb{K}[x]$ *único* $/\ P\, N_e = N_e\, P = P$

D) La existencia de este polinomio es evidente ya que en la definición de polinomios se vieron los polinomios constantes.

Para demostrar la unicidad del neutro vamos a suponer que existen dos neutros para el producto: $N_{e1}\ y\ N_{e2}$, por lo tanto, si:

N_{e1} es neutro, entonces $N_{e1}\,.\,N_{e2} = N_{e2}$

N_{e2} es neutro, entonces $N_{e2}\,.\,N_{e1} = N_{e1}$

De donde se deduce, por ser los primeros miembros de ambas igualdades iguales, que $N_{e1} = N_{e2}$ y, consecuentemente, el neutro para el producto es único.

P_5) Inverso multiplicativo.

Solo los polinomios de grado cero $(a_0x^0\ con\ a_0 \neq 0)$ tienen inverso multiplicativo, pues si P es un polinomio y Q su inverso, su producto debe ser igual al elemento neutro, es decir $P\,.\ Q = Q\,.\ P = N_e$, es decir $P\,.\ Q = 1$. Esto trae como consecuencia que la suma de los grados de $P\ y\ Q$ debe ser cero, esto exige a su vez que tanto el grado de P como el de Q sean iguales a cero.

P_6) El producto es distributivo respecto a la suma en $\mathbb{K}[x]$.

$\forall\, P, Q, R \in \mathbb{K}[x] : P(Q + R) = P\,Q + P\,R$

D) Sean a_i, b_i, c_i los i–ésimos coeficientes de los polinomios $P, Q\ y\ R$ respectivamente. Por lo tanto $b_i + c_i$ es el i–ésimo coeficiente de $Q + R$ como también $a_i(b_i + c_i)$ es el i–ésimo coeficiente de $P(\,Q + R)$. Como

$a_i(b_i + c_i) = a_i b_i + a_i c_i \; \forall \; i = 1, 2, \ldots, n$ por la propiedad distributiva del producto en $\mathbb{K}$, respecto a la suma, los coeficientes respectivos de $P(Q + R)$ y de $P\,Q + P\,R$ son iguales. Concluimos entonces que $P(Q + R) = P\,Q + P\,R$.

2.4. Raíces de un polinomio

Estudiaremos en primer lugar qué es raíz o cero de un polinomio, es decir aquellos números que, al evaluarlos en el polinomio, anulan a este.

Definición: *sean el polinomio* $P \in \mathbb{K}[x]$ *y el número* $b \in \mathbb{K}$*, se dice que* b *es un cero o una raíz de* P *si* $P(b) = 0$.

Ejemplo 1: los polinomios de grado 1, $a_1 x + a_0$ siempre tienen una única raíz $-\frac{a_0}{a_1}$. Sea el polinomio $P(x) = 2x + 4$ se anula para $x = -2$ ya que $P(-2) = 0$.

Ejemplo 2: sea el polinomio de segundo grado $P(x) = a_2 x^2 + a_1 x + a_0$ cuya notación más popular es $P(x) = ax^2 + bx + c$ con $a \neq 0 \;\; y \;\; a, b, c \in \mathbb{K}$, ya conocemos una fórmula para calcular las raíces (resolvente para la ecuación de segundo grado). Esta fórmula puede demostrarse utilizando el procedimiento conocido como "completar cuadrados":

$$P(x) = ax^2 + bx + c$$

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

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

$$P(x) = a\left[\left(x + \frac{b}{2a}\right)^2 - \left(\frac{b^2 - 4ac}{4a^2}\right)\right]$$

$$P(x) = a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right]$$

El número $\Delta = b^2 - 4ac$ se denomina *discriminante* del polinomio cuadrático P. Si Δ tiene una raíz cuadrada en $\mathbb{K}$, es decir si existe un elemento $\sqrt{\Delta} \in \mathbb{K}$ que resuelva la ecuación $x^2 = \Delta$ (lo que es posible en $\mathbb{R}$ si $\Delta \geq 0$ y siempre en $\mathbb{C}$) entonces podemos escribir P como diferencia de cuadrados:

$$P(x) = a\left[\left(x + \frac{b}{2a}\right)^2 - \left(\frac{\sqrt{\Delta}}{2a}\right)^2\right]$$

$$P(x) = a\left[\left(x + \frac{b}{2a} + \frac{\sqrt{\Delta}}{2a}\right)\left(x + \frac{b}{2a} - \frac{\sqrt{\Delta}}{2a}\right)\right]$$

$$P(x) = a\left[\left(x + \frac{b+\sqrt{\Delta}}{2a}\right)\left(x + \frac{b-\sqrt{\Delta}}{2a}\right)\right]$$

$$P(x) = a(x - \alpha_1)(x - \alpha_2) \text{ con } \alpha_1 = \frac{-b-\sqrt{\Delta}}{2a} \;\; y \;\; \alpha_2 = \frac{-b+\sqrt{\Delta}}{2a}$$

Por lo tanto P se anula cuando $x = \alpha_1$ o $x = \alpha_2$, es decir que $\alpha_1 \;\; y \;\; \alpha_2$ son exactamente las raíces de P.

En resumen, un polinomio de segundo grado tiene exactamente dos raíces, siempre que sea posible *extraer la raíz cuadrada* de su discriminante $\Delta = b^2 - 4ac$ en $\mathbb{K}$, particularmente esto es posible siempre en $\mathbb{K} = \mathbb{C}$, y en $\mathbb{K} = \mathbb{R}$ solo si $\Delta \geq 0$.

La denominación de "raíz" ha quedado por razones históricas, ya que los matemáticos pensaban, inicialmente, que los ceros de un polinomio podrían determinarse mediante fórmulas involucrando la extracción de raíces análogas a la que hemos estudiado para los polinomios de segundo grado. De hecho, esto es posible si el grado de P es tres o cuatro (las fórmulas correspondientes se complican un poco), sin embargo, luego se

demostró que esto no es en general posible para polinomios de grado mayor o igual a cinco.

Gráfica de polinomios

La gráfica de todos los polinomios son curvas suaves (no tienen elevaciones ni depresiones abruptas) y continuas (sin interrupciones en su trazo); distinguiéndose dos elementos: las ramas y la parte central. En la parte central, la gráfica se puede plegar varias veces.

a) Las ramas: no son nunca rectas, aunque pueden parecerlo si el dominio representado es muy extenso. Pueden dirigirse ambas hacia arriba, ambas hacia abajo o bien una en cada dirección. Si se representa la gráfica de un polinomio en un intervalo mayor, la forma de los extremos prácticamente no varía, vale decir, los extremos de una gráfica nos dan una idea de cómo continúa la gráfica de un polinomio.

Por ejemplo, las ramas de la gráfica del polinomio $P(x) = 4x^4 - 3x^3 - 5x^2 - x - 2$ se dirigen ambas hacia arriba y tiene tres pliegues en su parte central.

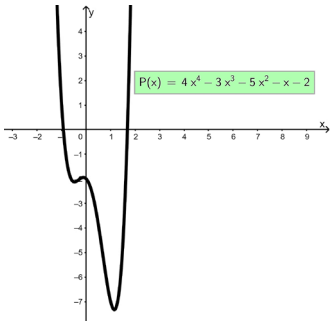


Figura 3.1. Representación gráfica de ramas y pliegues del polinomio P(x).

Pero, las ramas de la gráfica de $Q(x) = -3x^6 - 4x^5 + 3x^4 + 3x^3 + 4x^2 - x - 4$ se dirigen ambas hacia abajo y también tiene tres pliegues.

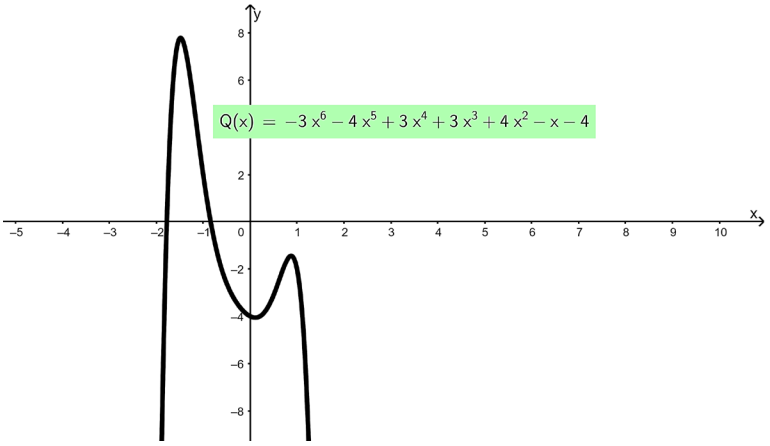


Figura 3.2. Representación gráfica de ramas y pliegues del polinomio Q(x).

Finalmente, las ramas de la gráfica de $R(x) = 5x^5 - x^4 - 3x^3 + 5x^2 - x - 3$ se dirigen una hacia abajo y la otra hacia arriba y tiene dos pliegues.

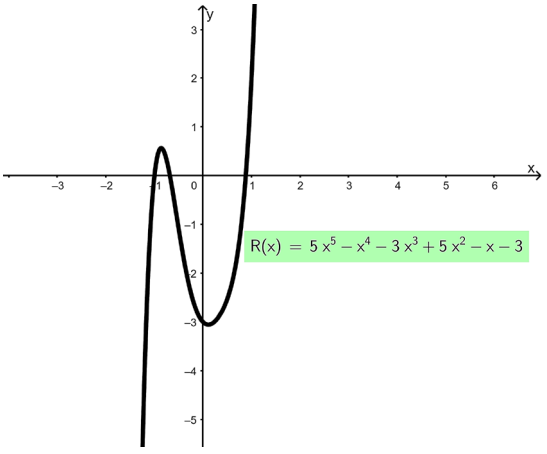


Figura 3.3. Representación gráfica de ramas y pliegues del polinomio R(x).

Estas observaciones pueden generalizarse de la siguiente manera:

Sea el polinomio $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ ... + a_1 x + a_0$ con $a_n \neq 0$, entonces:

Si n es par y $a_n < 0$, la gráfica de $P(x)$ tendrá las dos ramas hacia abajo; pero si $a_n > 0$, tendrá ambas ramas hacia arriba.

Si n es impar, una rama va hacia abajo y la otra hacia arriba.

Este comportamiento obedece a que el coeficiente principal de un polinomio determina si la gráfica del mismo crecerá o decrecerá hacia la derecha de algún valor de x. Conforme el valor de x aumenta, este término terminará por dominar a todos los demás del polinomio. Por lo tanto, si el coeficiente de este término es positivo, *en algún momento* la gráfica comenzará a crecer a medida que el valor de x aumente. Si el coeficiente principal es negativo, *en algún momento* la gráfica comenzará a decrecer a medida que el valor de x aumenta.

b) La parte central: en esta parte, la gráfica se pliega varias veces, el número de pliegues depende del grado del polinomio. El número máximo de pliegues de la gráfica de un polinomio es su grado menos 1; de este modo sabemos que, la gráfica de un polinomio lineal no puede tener pliegues, en cambio la gráfica de un polinomio cuadrático tiene exactamente un pliegue, y la gráfica de un polinomio de grado tres tiene, como máximo, dos pliegues.

La gráfica de un polinomio debe observarse de izquierda a derecha; así, por ejemplo, en la gráfica correspondiente a $P(x) = 2x^3 + 3x^2 - 12x + 3$

comprobamos que al principio se dirige hacia arriba, después hacia abajo y, finalmente, hacia arriba.

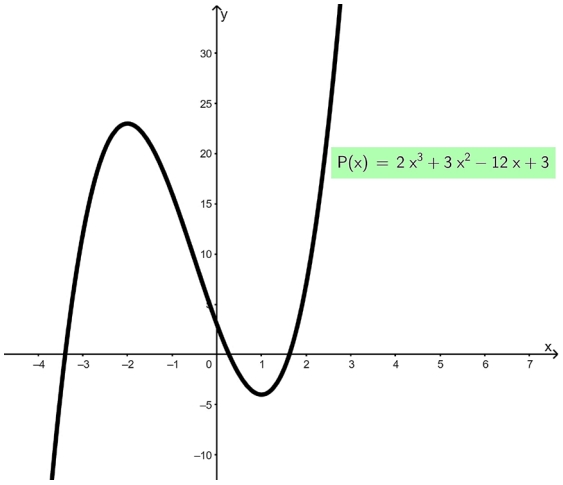


Figura 3.4. Representación gráfica del polinomio P(x).

De manera más rigurosa podemos decir que:

i) La gráfica de un polinomio será creciente cuando, a medida que aumenta el valor de x, al evaluar el polinomio también obtenemos un valor cada vez mayor.

ii) La gráfica de un polinomio será decreciente cuando, a medida que aumenta el valor de x, al evaluar el polinomio –por el contrario– obtenemos un valor cada vez menor.

De este modo, en el ejemplo anterior, la gráfica es creciente cuando $x < -2$, es decreciente entre –2 y 1, y vuelve a ser creciente a partir de 1, tal como lo muestra el siguiente gráfico.

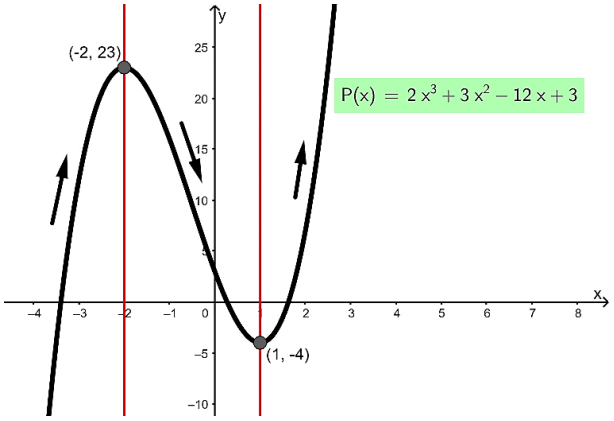


Figura 3.5. Crecimiento y decrecimiento en la gráfica del polinomio P(x).

Otros puntos importantes que podemos tener en cuenta, a la hora de observar y/o esquematizar la gráfica de un polinomio, son los siguientes:
i) Los máximos y los mínimos: se denomina máximo relativo al punto en el que la gráfica pasa de ser creciente a decreciente; el valor del polinomio en este punto es mayor que el de cualquier otro punto de la gráfica que se encuentre cercano. En cambio, un mínimo relativo es aquel punto en el que la función pasa de ser decreciente a creciente; el valor del polinomio en este punto es menor que el de cualquier otro punto de la gráfica que se encuentre cercano.
En la gráfica del ejemplo anterior, $P(x) = 2x^3 + 3x^2 - 12x + 3$ podemos observar que un máximo relativo se encuentra en $\big(-2, P(-2)\big)$ es decir en $(-2, 23)$ mientras que encontramos un mínimo relativo en $(1, P(1))$ es decir en $(1, -4)$.

ii) La intersección con el eje $\overrightarrow{oy}$: evidentemente solo existe un único punto de intersección con el eje $\overrightarrow{oy}$. Este punto es el que tiene abscisa $x = 0$.

Por ejemplo, si $P(x) = 2x^3 - 5x^2 - 4x + 10$, el único punto de intersección de la gráfica de este polinomio con el eje $\overrightarrow{oy}$ es $\big(0, P(0)\big)$ es decir $(0, 10)$ tal como se puede observar en la gráfica correspondiente.

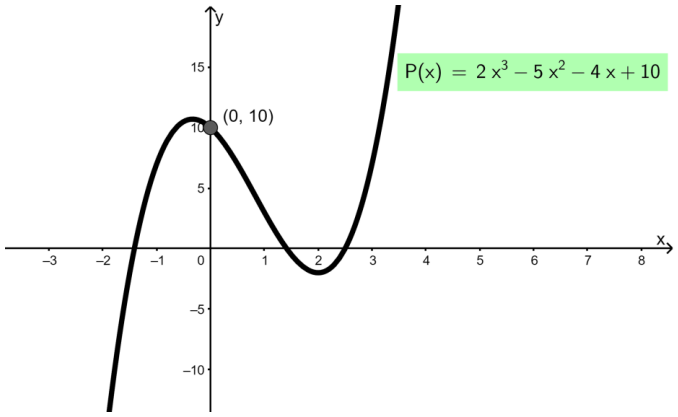


Figura 3.6. Intersección de la gráfica del polinomio P(x) con el eje $\overrightarrow{oy}$.

iii) La intersección con el eje $\overrightarrow{ox}$: en este caso puede haber un número de intersecciones igual al grado del polinomio. Para esto se puede resolver la ecuación $P(x) = 0$ o bien buscar un método –que luego se estudiará– para determinar las raíces reales del polinomio. Es decir que, gráficamente, las raíces reales del polinomio son los puntos en donde este interseca al eje $\overrightarrow{ox}$.

Si se conocen las raíces de un polinomio, se puede expresar este en función de ellas.

Por ejemplo, el polinomio $P(x) = 4x^6 - 10x^5 - 10x^4 + 14x^3 - 26x^2 + 76x - 48$ tiene tres raíces reales $x_1 = 1\ (doble), x_2 = 3\ y\ x_3 = -2$, por lo tanto el polinomio se puede expresar como: $P(x) = 2(x-1)^2(x-3)(x+2)(2x^2 + x + 4)$. Se puede comprobar que $x_1 = 1$ es una raíz doble en la gráfica correspondiente.

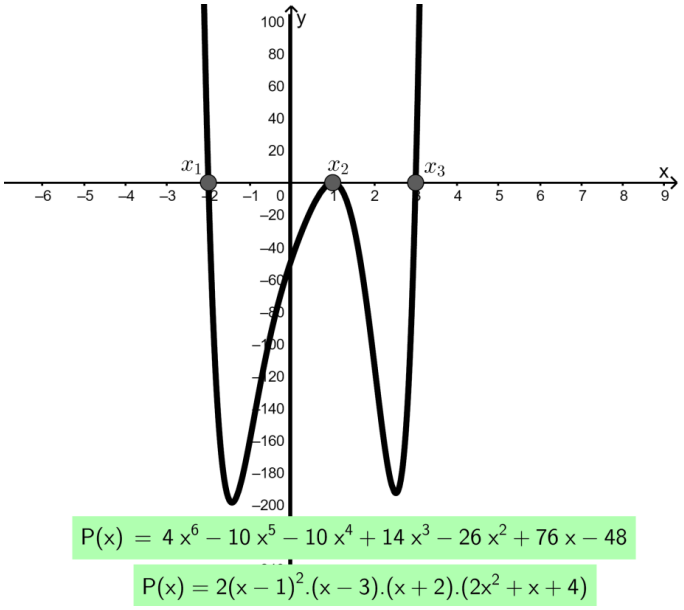


Figura 3.7. Raíces en la representación gráfica del polinomio P(x).

3. Divisibilidad de polinomios

En el conjunto de polinomios $\mathbb{K}[x]$ se pueden estudiar cuestiones de divisibilidad o factorización, ya que se definieron previamente suma, resta y producto de polinomios. Este estudio está en completa analogía con la aritmética de números enteros, cuestión que nos ayudará a entender conceptos similares, pero con polinomios.

Cabe destacar que la factorización de polinomios guarda estrecha relación con el problema de encontrar las raíces o ceros de un polinomio.

Definición: *sean los polinomios* $P, Q \in \mathbb{K}[x]$ *se dice que* P *divide a* Q *y lo denotamos como* $P|Q$ *si existe un polinomio* $R \in \mathbb{K}[x]$ *tal que* $Q = P.R$.

Sean $P, Q \in \mathbb{K}[x]$, $P|Q$ *si* $\exists\, R \in \mathbb{K}[x] \,/\, Q = P.R$

Ejemplo: el polinomio $P(x) = x - 1$ divide al polinomio $Q(x) = x^3 - 1$ en $\mathbb{Q}[x]$ ya que $Q(x)$ admite la factorización: $x^3 - 1 = (x - 1)(x^2 - x + 1)$.

La divisibilidad de polinomios tiene dos importantes propiedades:

P_1) Si $P(x)$ divide a $Q(x)$ entonces divide al producto $Q(x).S(x)$, cualquiera sea el polinomio $S(x)$.

Si $P|Q \implies P|Q.S \;\; \forall\, S \in \mathbb{K}[x]$

D) Si $P|Q$ entonces existe el polinomio $R \in \mathbb{K}[x]$ tal que:

$Q = P.R$

$Q.S = P.R.S$ si llamamos $T = Q.S \;\; y \;\; U = R.S$

$T = P.U$ lo que significa que

$P|T$ es decir que $P|Q.S$

P_2) Si $P(x)$ divide a $Q(x)$ y a $S(x)$, entonces divide a su suma y a su diferencia.

Si $P|Q \;\wedge P|S \implies P|Q + S \;\wedge P|Q - S \;\; \forall\, P, Q, S \in \mathbb{K}[x]$

D)

Si $P|Q \implies Q = P.R$ (1) con $R \in \mathbb{K}[x]$

Si $P|S \implies S = P.T$ (2) con $T \in \mathbb{K}[x]$

Sumando miembro a miembro (1) y (2)

$Q + S = P.R + P.T$

$Q + S = P\ (R + T)$

$Q + S = P\ .U$ con $U = R + T$ por lo tanto

$P|Q + S$

Para el caso de que $P|Q - S$ se procede de manera similar.

Debemos destacar que todo $P \in \mathbb{K}[x]$ es divisible por cualquier constante no nula (pensadas como polinomios de grado cero), es decir que estas constantes juegan el mismo papel en la aritmética de polinomios que los números 1 y –1 juegan en la aritmética de $\mathbb{Z}$.

A partir de la definición de divisibilidad, se puede introducir la noción de polinomio irreducible que sería la noción análoga a la de número primo en la aritmética de $\mathbb{Z}$.

Definición: *sea el polinomio no constante* $P \in \mathbb{K}[x]$ *se dice que* P *es irreducible en* $\mathbb{K}[x]$ *si no es posible factorizarlo en la forma* $P = Q.S$ *donde* Q *y* S *son polinomios no constantes de* $\mathbb{K}[x]$.

Ejemplo 1: todo polinomio de grado 1 es siempre irreducible.

Ejemplo 2: un polinomio de grado 2 será irreducible, según tenga o no raíces en $\mathbb{K}$. Sea el polinomio $x^2 + 1$ si lo pensamos en $\mathbb{R}[x]$ es irreducible pues no admite una factorización como producto de dos factores de grado 1. En cambio, si lo pensamos en $\mathbb{C}[x]$ se factoriza de la forma $x^2 + 1 = (x + i)(x - i)$ y por lo tanto no es irreducible.

En general, un polinomio que admite raíces en $\mathbb{K}$ no puede ser irreducible en $\mathbb{K}[x]$.

Un concepto importante en la aritmética de $\mathbb{Z}$ es el algoritmo de división, por ello tiene sentido preguntarse si habrá un concepto similar para polinomios.

3.1. El algoritmo de división para polinomios

La división de polinomios se define de manera análoga a la división de números enteros. Recordemos que dados $p, d \in \mathbb{Z} \wedge d \neq 0$, dividir p por d consistía en encontrar dos enteros q y r tales que $p = q\ .\ d + r$ para $0 \leq r < d$. El número q recibe el nombre de cociente de la división y r el resto, también p se llamaba dividendo y d divisor. El objetivo de dividir p por q es ver cuántas veces está contenido d en el número p (dado por el cociente q), y si no está contenido un número exacto de veces ($r = 0$), cuánto sobra hasta poder completar una vez más d (dado por el resto r). Por ejemplo, la división de $p = 386$ por $d = 12$ tiene por cociente $q = 32$ y resto $r = 2$, en efecto podemos verificar la veracidad de esta división ya que $386 = 32\ .12 + 2$ con $0 < 2 < 12$.

El algoritmo de la división de enteros es conocido, de manera similar es la división de polinomios, pero en primer lugar definamos exactamente división de polinomios.

Teorema: *sean los polinomios* $P, D \in \mathbb{K}[x],\ D \neq 0 \wedge gr(D) \leq gr(P)$, *existen los dos únicos polinomios* $Q(cociente)$ *y* $R(resto)$ *tal que* $P = Q\ .D + R$ *con* $gr(R) < gr(D)$.

Si R es el polinomio nulo, entonces $P = Q\ .D$, se dice que la división es exacta y que P es divisible por D o que D es divisor de P.

DE) Para demostrar la existencia de los polinomios $Q\ y\ R$ haremos inducción en el grado del dividendo P.

1°) Debemos verificar que lo que queremos demostrar es válido para el primer elemento del conjunto con el que se trabaja. En nuestro caso 0, es decir $gr(P) = 0$.

Si $gr(P) = 0$ (polinomios constantes) claramente podemos tomar $Q = 0\ \ y\ \ R = P$.

2°) Debemos formular la *Hipótesis Inductiva*, esta será que $gr(P) = n$ y que el teorema es válido cuando el grado del dividendo es menor que n.

3°) Debemos demostrar que el teorema es válido para el caso en el que el grado del dividendo sea mayor o igual al grado del divisor.

Sean los polinomios:

$P = \sum_{i=0}^{n} a_i x^i\ con\ a_n \neq 0\ \ (gr(P) = n)\ \ y\ \ D = \sum_{j=0}^{m} b_j x^j\ \ con\ b_m \neq 0\ \ (gr(D) = m)$

Nuevamente si $n < m$ (grado del dividendo menor que el grado del divisor) podemos tomar $Q = 0\ \ y\ \ R = P$.

Si $n \geq m$ podemos determinar un primer cociente Q_0 dividiendo el término principal de P con el término principal de D, es decir, $Q_0 = \frac{a_n}{b_m} x^{n-m}$ y en consecuencia definir un primer resto $R_0 = P - Q_0\ D$.

Si fuera $R_0 = 0$ ó $gr(R_0) < gr(D)$, hemos finalizado, tomando $Q = Q_0\ \ y\ \ R = R_0$ se demuestra la existencia de $Q\ \ y\ \ R$.

De no cumplirse el caso anterior, se repite el proceso. Debe tenerse en cuenta que $gr(R_0) < gr(P)$ ya que en la forma que hemos elegido Q_0 el

término correspondiente a la potencia x^n se cancela. Entonces, en virtud de la Hipótesis Inductiva, existirán el cociente Q_1 y el resto R_1 de la división de R_0 por D, de modo que: $R_0 = Q_1\, D + R_1$ donde $R_1 = 0$ ó $gr(R_1) < gr(D)$. Entonces

$P = Q_0\, D + R_0$

$P = Q_0\, D + Q_1\, D + R_1$

$P = D\, (Q_0 + Q_1) + R_1$

Si tomamos $Q = Q_0 + Q_1 \;\; y \;\; R = R_1$ se demuestra la existencia de $Q \;\; y \;\; R$. Esto demuestra la parte de existencia.

DU) Para demostrar la unicidad vamos a suponer que hay dos cocientes $Q \;\; y \;\; Q_1$ como también dos restos $R \;\; y \;\; R_1$. De modo que:

$P = Q\, D + R \;\; y \;\; R = 0$ ó $gr(R) < gr(D)$

$P = Q_1\, D + R_1 \;\; y \;\; R_1 = 0$ ó $gr(R_1) < gr(D)$ por lo tanto

$Q\, D + R = Q_1\, D + R_1$

$Q\, D - Q_1\, D = R_1 - R$

$(Q - Q_1)\, D = R_1 - R$

Si $R = R_1$ tendríamos que

$(Q - Q_1)\, D = 0$, como $D \neq 0$ es

$Q - Q_1 = 0 \implies Q = Q_1$

Se debe probar que no puede ocurrir $R \neq R_1$. Pero si esto ocurriera, sería $R_1 - R \neq 0$ entonces $Q - Q_1 \neq 0$ y comparando los grados tenemos:

$gr[(Q - Q_1)\, D] = gr(Q - Q_1) + gr(D) > gr(D)$ por otra parte

$gr(R_1 - R) \leq máx[gr(R), gr(R_1)] < gr(D)$

Esto significa que en la igualdad $(Q - Q_1)\, D = R_1 - R$ se tiene un polinomio que tiene grado mayor a $gr(D)$ y grado menor que $gr(D)$ al mismo tiempo, lo cual es un absurdo.

Esta contradicción proviene de suponer que $R \neq R_1$. Por lo tanto debe ser $R = R_1$ y consecuentemente $Q = Q_1$. Probando de esta manera la unicidad del cociente y del resto.

De esta manera queda demostrado el teorema.

Ejemplo: dividir el polinomio $P(x) = 2x^5 - x^2 + 4$ por $D(x) = x^3 - x + 2$

Antes de proceder a dividir, hay dos puntos importantes a considerar:

a) El dividendo y el divisor tienen que estar ordenados en forma decreciente de acuerdo al grado mayor de la indeterminada.

b) Si faltara alguna potencia de la indeterminada en el dividendo, hay que completar este polinomio poniendo los términos que faltan con coeficiente 0.

Disposición práctica para el cálculo de la división entre polinomios:

$2x^5$	$+0x^4$	$+0x^3$	$-x^2$	$+0x$	$+4$	x^3	$-x$	$+2$
$-2x^5$		$+2x^3$	$-4x^2$			$2x^2$	$+2$	$Q(x)$
		$+2x^3$	$-5x^2$	$+0x$	$+4$			
		$-2x^3$		$+2x$	-4			
			$-5x^2$	$+2x$	$R(x)$			

Entonces: $Q(x) = 2x^2 + 2$ y $R(x) = -5x^2 + 2x$

Se puede verificar (lo mismo que en la aritmética de $\mathbb{Z}$) que los cálculos son correctos haciendo el producto de $D(x)$ por $Q(x)$ y luego sumando $R(x)$ a este resultado, debiendo obtenerse el polinomio $P(x)$.

En algunas situaciones se desea saber si la división de dos polinomios es exacta o no, es decir basta con conocer si el resto es cero o no. Solo interesa conocer el resto de la división, sin efectuar esta.

Analizaremos a continuación un caso particular en que se realice una división por un polinomio de la forma $(x - a)$ con $a \in \mathbb{K}$, es decir el caso particular en donde el divisor es un binomio de grado 1 y con coeficiente principal 1.

3.2. Teorema del resto

El resto de la división de un polinomio $P \in \mathbb{K}[x]$ *por otro de la forma* $(x - a)$ *con* $a \in \mathbb{K}$ *es el valor del polinomio* P *evaluado cuando* $x = a$*, es decir* $P(a)$.

D) Dividiendo P por $D(x) = (x - a)$ podemos escribirlo como:

$P(x) = Q(x)(x - a) + R$ donde el resto R debe ser un polinomio constante ya que si $gr(P) \geq 1$ y $gr(D) = 1$ entonces $gr(Q) = gr(P) - 1$ y $gr(R) < gr(D) = 1$, por lo tanto R es un polinomio constante.

Evaluando $P(x) = Q(x)(x - a) + R$ cuando $x = a$ tenemos:

$P(a) = Q(a)(a - a) + R$

$P(a) = R$ quedando de esta manera demostrado el teorema.

Ejemplo: hallar el resto de dividir $P(x) = x^3 - 3x^2 - 1$ por $D(x) = (x + 4)$.

Como el divisor es un binomio de la forma $(x-a)$ ya que $\big(x-(-4)\big)$, podemos aplicar directamente el teorema del resto, teniendo en cuenta que $a=-4$.

$R=(-4)^3-3(-4)^2-1$

$R=-64-48-1=-113$

Una consecuencia importante del teorema anterior se enuncia en el siguiente corolario.

Corolario: *sea el polinomio* $P\in\mathbb{K}[x]$, *P es divisible por* $(x-a)$ *con* $a\in\mathbb{K}$ *si y solo si* a *es raíz de* P.

Ejemplo: sea el polinomio $P(x)=x^3-2x^2-5x+6$.

Se puede observar que $x=1$ es raíz de $P(x)$ ya que $P(1)=0$. Entonces $P(x)$ es divisible por $(x-1)$. Hacemos la división (por Regla de Ruffini) y expresamos el polinomio $P(x)$ como:

$P(x)=(x-1)(x^2-x-6)=(x-1)Q(x)$

Para que $P(x)$ se anule, debe ser $(x-1)=0$ o bien $Q(x)=0$

$Q(x)=0$ es resolver la ecuación de segundo grado $x^2-x-6=0$, resolviéndola se obtienen como raíces $x=-2\ y\ x=3$. Con lo que $Q(x)$ se factoriza como:

$Q(x)=(x+2)(x-3)$.

En consecuencia, las raíces de P son $x_1=1$, $x_2=-2\ y\ x_3=3$ y su factorización será:

$P(x)=(x-1)(x+2)(x-3)$

Utilidad: en general, este corolario se utiliza cuando se desea resolver ecuaciones, ya que una vez encontrada alguna raíz a del polinomio P, el

problema se reduce a resolver una ecuación de un grado menor, efectuando la división de P por $(x - a)$.

Notemos que la estrategia anterior pide conocer o determinar, de alguna manera, una raíz del polinomio P. Algunas veces las raíces se pueden identificar rápidamente, por ejemplo, si el polinomio no tiene término independiente, entonces $x = 0$ es raíz. Otras veces son datos del problema.

El siguiente teorema, debido al matemático Carl F. Gauss, nos permite anticipar las posibles raíces *racionales* de un polinomio con coeficientes *enteros*.

4. Raíces Racionales

4.1. Criterio de Gauss

Este criterio es de gran ayuda para determinar, en caso de que existan, la forma de las raíces racionales de un polinomio.

Criterio: *sea* $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ ... + a_1 x + a_0 \ \ con\ a_i \in \mathbb{Z} \ \ y \ \ a = \frac{p}{q} \in \mathbb{Q}$ *es una raíz racional de* $P(x)$*, escrita como fracción irreducible (o sea p y q coprimos*[18]*) se tiene necesariamente que* $p|a_0 \ y \ \ q|a_n$.

En particular si $P(x)$ *es mónico, las posibles raíces racionales de* $P(x)$ *son los divisores del término independiente* a_0.

D) Como a es raíz de $P(x)$ entonces $P(a) = 0$

[18] Dos enteros son coprimos (primos relativos) si y solo si su máximo común divisor es igual a 1.

$a_n \frac{p^n}{q^n} + a_{n-1} \frac{p^{n-1}}{q^{n-1}} + \ ... + a_1 \frac{p}{q} + a_0 = 0$ multiplicando miembro a miembro por q^n

$a_n p^n + a_{n-1} p^{n-1} q + \ ... + a_1 p\, q^{n-1} + a_0 q^n = 0$ Luego

$p(a_n p^{n-1} + a_{n-1} p^{n-2} q + \ ... + a_1 q^{n-1}) = -a_0 q^n$

En particular $p | a_0 q^n$

Pero como p es coprimo con q, p es coprimo con q^n (como consecuencia del teorema fundamental de la aritmética[19]). Por lo tanto p debe dividir a_0.

Similarmente

$q(a_{n-1} p^{n-1} + a_{n-2} p^{n-2} q + \ ... + a_1 q^{n-2} + a_0 q^{n-1}) = -a_n p^n$

Por lo tanto $q | a_n p^n$

Pero como q es coprimo con p, q es coprimo con p^n y, en consecuencia, $q | a_n$

De esta forma queda demostrado el criterio.

Debemos destacar que el criterio *no afirma* que un polinomio con coeficientes enteros tiene raíces racionales. Lo que indica es que, *de existir raíces racionales*, estas deben ser forzosamente de la forma que indica el criterio.

Debido a esto es que para determinar si el polinomio P tiene raíces racionales, debemos probar con los números racionales que tienen la forma anticipada por el criterio. Esto implica que se puede dar el caso de

[19] Todo entero positivo mayor que uno es un número primo o bien se puede escribir de manera única, salvo en el orden, como un producto de números primos.

que ninguna de las posibilidades sea raíz del polinomio y este solo tenga raíces reales y complejas.

Ejemplo 1: hallar las raíces de $P(x) = 2x^3 - 4x^2 + x + 1$.

$P(x)$ tiene coeficientes enteros y término independiente no nulo, podemos aplicar directamente el Criterio de Gauss, una raíz racional a de $P(x)$ debe ser de la forma $a = \frac{p}{q}$, donde p divide a 1 y q divide a 2, luego p puede ser ± 1 y $q \ \pm 1, \pm 2$; por lo tanto, las posibles raíces del polinomio son ± 1 ó $\pm \frac{1}{2}$. Evaluando $P(x)$ en estos valores, observamos que 1 es raíz ($P(1) = 0$), mientras que los otros valores no lo son. Por lo tanto $P(x)$ será divisible por $(x - 1)$. Aplicamos Ruffini para hallar los coeficientes del cociente.

	2	–4	1	1
1	↓	2	–2	–1
	2	–2	–1	0

Por lo tanto $P(x) = (x - 1)(2x^2 - 2x - 1)$.

Para terminar de factorizar $P(x)$ buscamos las raíces de $Q(x) = 2x^2 - 2x - 1$ en el cual también se podría aplicar el Criterio de Gauss, pero como es de segundo grado directamente aplicamos la fórmula resolvente:

$$x_{1,2} = \frac{2 \pm \sqrt{(-2)^2 - 4 \,.2(-1)}}{2 \,.2} = \frac{2 \pm \sqrt{12}}{4} = \frac{2 \pm 2\sqrt{3}}{4} = \frac{1}{2} \pm \frac{\sqrt{3}}{2}$$

En consecuencia $Q(x) = 2\left(x - \left(\frac{1}{2} + \frac{\sqrt{3}}{2}\right)\right)\left(x - \left(\frac{1}{2} - \frac{\sqrt{3}}{2}\right)\right)$ y, por lo tanto, $P(x)$ queda factorizado como:

$P(x) = 2(x-1)\left(x-\left(\frac{1}{2}+\frac{\sqrt{3}}{2}\right)\right)\left(x-\left(\frac{1}{2}-\frac{\sqrt{3}}{2}\right)\right)$ y sus raíces son:

$x_1 = 1, x_2 = \frac{1}{2}+\frac{\sqrt{3}}{2}\ \ y\ \ x_3 = \frac{1}{2}-\frac{\sqrt{3}}{2}$

Ejemplo 2: hallar las raíces de $P(x) = x^3 - 3x^2 - 2x + 6$. El Criterio de Gauss establece que si en particular, $P(x)$ es mónico, las posibles raíces racionales son los divisores del término independiente a_0. Esto se debe a que las posibles raíces tienen la forma $\frac{p}{q}$ donde p divide al término independiente y q divide al término principal, que es 1, ya que el polinomio es mónico, por lo tanto $q = \pm 1$, lo que significa que solo podemos tener en cuenta p o sea los divisores del término independiente. Entonces las posibles raíces son: $\pm 1, \pm 2\ \ y\ \ \pm 3$.

Evaluando $P(x)$ en todos estos valores observamos que $P(3) = 0$ lo que significa que 3 es raíz del polinomio. De manera similar al ejemplo anterior dividimos $P(x)$ en $(x-3)$ y obtenemos (x^2-2). Por lo tanto $P(x)$ se factoriza como: $P(x) = (x-3)(x^2-2)$ de donde observamos que las raíces son: $x_1 = 3,\ x_2 = \sqrt{2}\ \ y\ \ x_3 = -\sqrt{2}$.

Si bien el enunciado del Criterio de Gauss está hecho para polinomios a coeficientes enteros y con término independiente no nulo, este criterio puede aplicarse a polinomios a coeficientes enteros con término independiente cero y también a polinomios cualesquiera a coeficientes racionales. Veamos a continuación estas dos posibilidades:

a) Sea $P(x) \in \mathbb{Z}[x]$ con término independiente $a_0 = 0$. En este caso basta sacar factor común x^m para m el menor grado de los términos de P; de

este modo podemos escribir $P(x) = x^m Q(x)$ con Q un polinomio a coeficientes enteros con término independiente no nulo. Por lo tanto, para hallar las raíces de P aplicamos el Criterio de Gauss.

Ejemplo: hallar las raíces de $P(x) = 2x^6 + x^5 - 6x^4 + x^3 + 2x^2$.

Observamos que es un polinomio a coeficientes enteros y término independiente nulo. Comenzamos por sacar factor común x^2 ($m = 2$), luego expresamos al polinomio P como $P(x) = x^2(2x^4 + x^3 - 6x^2 + x + 2)$. Procedemos a encontrar las raíces del polinomio $Q(x) = 2x^4 + x^3 - 6x^2 + x + 2$, que tiene coeficientes enteros y término independiente no nulo; aplicamos el Criterio de Gauss.

Las posibles raíces tienen la forma $\frac{p}{q}$ en donde p divide a 2 y q divide a 2 ($a_n = a_0$), es decir que los posibles valores son: $\pm 1, \pm 2$ y $\pm \frac{1}{2}$.

Evaluando $Q(x)$ en los posibles valores, concluimos que son raíces $1, -2$ y $-\frac{1}{2}$. Aplicamos tres veces la regla de Ruffini y obtenemos:

	2	1	–6	1	2
1		2	3	–3	–2
	2	3	–3	–2	0
–2		–4	2	2	
	2	–1	–1	0	
$-1/2$		–1	1		
	2	–2	0		

Por lo tanto, podemos escribir

$Q(x) = (x-1)(x+2)\left(x+\frac{1}{2}\right)(2x-2) = 2(x-1)(x+2)\left(x+\frac{1}{2}\right)(x-1).$

Observemos cómo fuimos factorizando el polinomio P.

$P(x) = 2x^6 + x^5 - 6x^4 + x^3 + 2x^2$

$P(x) = x^2(2x^4 + x^3 - 6x^2 + x + 2)$.

$P(x) = 2x^2(x-1)^2(x+2)\left(x + \frac{1}{2}\right)$

Finalmente, las raíces de $P(x)$ son: $x_1 = x_2 = 0$, $x_3 = x_4 = 1$, $x_5 = -2$ y $x_6 = -\frac{1}{2}$.

b) Sea $P(x) \in \mathbb{Q}[x]$, es decir, $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ ... + a_1 x + a_0$ $con\ a_i \in \mathbb{Q}$. Como cada a_i tendrá la forma $\frac{b_i}{c_i}$ $con\ b_i, c_i \in \mathbb{Z}$ y $c_i \neq 0$ se puede determinar el mínimo común múltiplo de los c_i como $t_i = mcm\ (c_i)$. Multiplicando a $P(x)$ por el mínimo común múltiplo de los denominadores, se puede suponer que todos los coeficientes son enteros. Luego, aplicamos el Criterio de Gauss y dividimos por t_i para obtener P.

Ejemplo: hallar las raíces de $P(x) = \frac{1}{2}x^3 - \frac{7}{12}x^2 + \frac{1}{12}$.

Es un polinomio a coeficientes racionales, podemos multiplicarlo por $t_i = 12$. Por lo tanto: $12\,.P(x) = 12\left(\frac{1}{2}x^3 - \frac{7}{12}x^2 + \frac{1}{12}\right) = 6x^3 - 7x^2 + 1 = Q(x)$.

Aplicamos el Criterio de Gauss a $Q(x) = 6x^3 - 7x^2 + 1$. Las posibles raíces de este polinomio son: ± 1, $\pm\frac{1}{2}$, $\pm\frac{1}{3}$ y $\pm\frac{1}{6}$. Evaluando a Q en esos valores se puede observar que son raíces $1, \frac{1}{2}$ y $-\frac{1}{3}$. En consecuencia $Q(x) = 6(x-1)\left(x - \frac{1}{2}\right)\left(x + \frac{1}{3}\right)$.

Como hemos multiplicado a P por 12 y determinamos Q, ahora a la expresión factorizada de Q la dividimos en 12 para obtener P.

$P(x) = \frac{1}{12}\left(6(x-1)\left(x-\frac{1}{2}\right)\left(x+\frac{1}{3}\right)\right) = \frac{1}{2}(x-1)\left(x-\frac{1}{2}\right)\left(x+\frac{1}{3}\right)$.

Las raíces de $P(x)$son: $x_1 = 1,\ x_2 = \frac{1}{2}\ y\ x_3 = -\frac{1}{3}$.

4.2. Regla de Descartes de los signos

En algunos casos, la siguiente regla[20] es útil para eliminar candidatas de listas largas de posibles raíces racionales. Antes de describir esta regla es necesario el concepto de *variación de signo*. Si $P(x)$ es un polinomio con coeficientes reales, ordenado en forma decreciente (no siendo necesario completarlo), entonces una *variación en signo* se presenta siempre que coeficientes adyacentes tengan signos contrarios. Por ejemplo $P(x) = 5x^7 - 3x^5 + x^4 + 2x^2 + x - 3$ tiene tres variaciones en signos.

Otros ejemplos:

Polinomio	Variaciones en signo
$x^2 + 4x + 1$	0 (ninguna)
$2x^3 + x - 6$	1 (una)
$x^4 - 3x^2 - x + 2$	2 (dos)

Regla de Descartes de los signos

Sea $P(x) \in \mathbb{R}[x]$, entonces:

i) El número de raíces reales positivas de $P(x)$ es igual al número de variaciones en signo en $P(x)$ o es menor a este último número, en un número entero par.

[20] Descubierta por el filósofo y matemático francés René Descartes en el año 1637.

ii) El número de raíces reales negativas de $P(x)$ es igual al número de variaciones en signo en $P(-x)$ o es menor a este último número, en un número entero par.

Ejemplo: usar la Regla de Descartes de los signos para determinar el número posible de raíces reales positivas y negativas de $P(x) = 3x^6 + 4x^5 + 3x^3 - x - 3$.

En primer lugar podemos observar que $P(x)$ tiene una variación de signo, por lo tanto tiene, a lo sumo, una raíz positiva.

Ahora hacemos:

$P(-x) = 3(-x)^6 + 4(-x)^5 + 3(-x)^3 - (-x) - 3$

$P(-x) = 3x^6 - 4x^5 - 3x^3 + x - 3$

Es decir que $P(-x)$ tiene tres variaciones en signo, entonces $P(x)$ podría tener tres raíces reales negativas o bien una. Esto hace un total de dos o cuatro raíces reales.

5. Multiplicidad de las raíces

Sea el polinomio $P(x) = x^2 - 2x + 1$ cuya factorización es $P(x) = (x-1)(x-1)$, por una parte podemos afirmar que P tiene dos raíces que son iguales y, por otra parte, podemos escribirlo como $P(x) = (x-1)^2$. En este caso decimos que P tiene una raíz $x = 1$ doble. Generalizando esta situación, introducimos la siguiente definición.

Definición: *sean el polinomio* $P \in \mathbb{K}[x]$ *y* $a \in \mathbb{K}$ *una raíz de* P. *Se dice que a es una raíz de P de multiplicidad* m ($m \in \mathbb{N}$) si P *admite la factorización* $P(x) = (x-a)^m Q(x)$ *donde el polinomio* Q *no se anula en* $x = a$ *o sea* $Q(a) \neq 0$.

Si $m = 1$ se dice que a es raíz *simple*, si $m = 2$ que es *doble*, y así sucesivamente.

Una consecuencia de esta definición es la siguiente proposición:

Si P es un polinomio que tiene en $\mathbb{K}$ *las raíces* $a_1, a_2, \dots, a_r$ *con multiplicidades* $m_1, m_2, \dots, m_r$, *entonces P admite la factorización:*

$P(x) = (x - a_1)^{m_1} (x - a_2)^{m_2} \dots (x - a_r)^{m_r} Q(x)$ *donde* $Q(a_i) \neq 0$ para $0 < i \leq r$.

Veremos a continuación, en primer lugar, una definición alternativa de multiplicidad de una raíz y, en segundo lugar, otra caracterización de la multiplicidad de las raíces; para ambas situaciones necesitamos introducir el concepto de derivada de un polinomio (visto como función). Al derivar obtenemos nuevamente un polinomio, pero de un grado menor que el dado.

Entonces, una definición alternativa de multiplicidad de una raíz es la siguiente: *una raíz de* $P(x)$ *tiene multiplicidad* k *si también es raíz de los primeros* $(k - 1)$ *polinomios derivados* $P'(x), P''(x), \dots, P^{(k-1)}(x)$ *de* $P(x)$, *pero no de la derivada* $k - \text{ésima}$ *de* $P(x)$.

Para el segundo punto si $\mathbb{K} = \mathbb{R}$, el concepto de derivada de un polinomio coincidirá con el concepto estudiado en los cursos de análisis matemático, pero si tomamos un $\mathbb{K}$ cualquiera, este concepto se puede introducir de una manera totalmente algebraica, sin referencia alguna a otros conceptos analíticos como el de límite, por ejemplo.

Definición: *sea el polinomio* $P \in \mathbb{K}[x],\ P = \sum_{k=0}^{n} a_k x^k$, *definimos polinomio derivado* $P' = \sum_{k=1}^{n} k a_k x^{k-1}$.

Por ejemplo, sea $P(x) = 2x^4 - x^3 + 2x - 5$ el polinomio derivado de P es:

$P'(x) = 4.2x^{4-1} - 3.1x^{3-1} + 1.2x^{1-1}$

$P'(x) = 8x^3 - 3x^2 + 2$

El polinomio derivado de P' (derivado segundo) es:

$P''(x) = 24x^2 - 6x$. Así sucesivamente.

Si bien no lo demostraremos, es relativamente sencillo comprobar que esta noción de la derivada totalmente formal conserva las propiedades usuales de la derivada, como la regla para derivar una suma o un producto.

Sean $P\ y\ Q$ dos polinomios, entonces:

$(P + Q)' = P' + Q'$

$(P\ Q)' = P'Q + PQ'$

$(-P)' = -P'$

También podemos definir, recursivamente, la derivada $n - ésima\ P^{(n)}$ del polinomio P, de la siguiente manera:

$$\begin{cases} P^{(0)} = P \\ P^{(n+1)} = (P^n)' \end{cases}$$

Teorema: *sea el polinomio* $P \in \mathbb{K}[x]\ y\ a \in \mathbb{K}$, *la condición necesaria y suficiente para que* a *sea raíz múltiple de orden* h *de* P, *es que anule al polinomio y a sus* $(h-1)$ *primeras derivadas, siendo distinta de cero la derivada de orden* h. Es decir $P(a) = P'(a) = P''(a) = \ ... = P^{(h-1)}(a) = 0$ pero $P^h(a) \neq 0$.

Este teorema afirma que la multiplicidad de a como raíz de P viene dada por el orden de la primera derivada de P que no se anula cuando lo evaluamos para $x = a$.

D)

$\Rightarrow$) Suponemos que a es una raíz de orden h de P, en consecuencia P es divisible por $(x-a)^h$ y se lo puede expresar como $P = (x-a)^h\, Q$, donde Q es un polinomio tal que $Q(a) \neq 0$, ya que en caso contrario, a sería una raíz de Q y la factorización de Q contendría el factor $(x-a)$ y a sería una raíz de P con un orden de multiplicidad mayor que h.

Si derivamos P:

$P' = h(x-a)^{h-1}\, Q + (x-a)^h\, Q'$

$P' = (x-a)^{h-1}[h\, Q + (x-a)Q']$

El polinomio entre corchetes no se anula para $x = a$ ya que en este caso se reduciría a $h\, Q$, en consecuencia a es una raíz de orden $(h-1)$ de P' porque P' contiene el factor $(x-a)^{h-1}$ en su factorización.

$\Leftarrow$) Recíprocamente: si a es una raíz de un polinomio P y de su derivada, a es una raíz múltiple de P.

En efecto, pues si a fuera una raíz simple de P se podría expresar este de la manera $P = (x-a)Q$ con $Q(a) \neq 0$. Derivando P:

$P' = Q + (x-a)Q'$ que no se anula para $x = a$ contradiciendo la hipótesis.

De esta manera queda demostrado el teorema.

Ejemplo 1: hallar las raíces de $P(x) = x^3 - 5x^2 + 7x - 3$.

$P(x) = x^3 - 5x^2 + 7x - 3$

$P'(x) = 3x^2 - 10x + 7$ en consecuencia $x_1 = 1 \ y \ x_2 = \frac{7}{3}$

$P^2(x) = 6x - 10$ en consecuencia $x = \frac{5}{3}$

Por lo tanto

$P(1) = 0$

$P'(1) = 0$

$P^2(1) \neq 0$

Esto significa que $x = 1$ es raíz doble, anula a P y a su derivada primera, siendo distinta de cero la derivada segunda.

$x = \frac{7}{3} \ y \ x = \frac{5}{3}$ no anulan a P, se descartan como posibles raíces. Además si $x = 1$ es raíz doble, solo resta una raíz simple de P.

Como P es divisible por $(x - 1)$, hacemos la división por regla de Ruffini dos veces y observamos que la raíz faltante es $x = 3$. Podemos expresar a P, como:

$P(x) = (x - 1)^2(x - 3)$ entonces sus raíces son: $x_1 = 1\ (doble)\ y\ x_2 = 3\ (simple)$

6. Teorema Fundamental del Álgebra

Uno de los objetivos en el estudio de polinomios es la determinación de raíces, con sus multiplicidades. Evidentemente, este problema es más sencillo de resolver cuando el grado del polinomio es menor. Supongamos que tenemos un polinomio P y que conocemos ya una raíz a de P, y sea $Q \in \mathbb{C}[x]$ tal que $P(x) = Q(x)(x - a)$. Sabemos que las raíces de Q son también raíces de P, por lo tanto, para encontrar las otras raíces

de P, buscamos las raíces de Q pues es más sencillo buscar estas raíces ya que Q es de grado menor a P. Repitiendo este proceso, será posible –y así lo hicimos en varios ejemplos– encontrar todas las raíces de P.

Observemos que el razonamiento anterior se basa en el hecho de que conocemos una raíz de P; en consecuencia es importante garantizar la existencia (y encontrar) de al menos una raíz de P. Este hecho está garantizado por el Teorema Fundamental del Álgebra.

Teorema Fundamental del Álgebra (TFA): *todo polinomio no constante con coeficientes complejos,* $P \in \mathbb{C}[x]$, *admite al menos una raíz compleja, es decir, podemos afirmar que existe* $a \in \mathbb{C}$ *tal que* $P(a) = 0$.

Si bien existen varias demostraciones de este teorema, no es posible hacerlo sin utilizar de alguna manera conceptos analíticos como la continuidad del polinomio $P(z)$ como función de variable compleja, $z \in \mathbb{C}$. Otra cuestión a destacar es que ninguna de las demostraciones es constructiva, es decir no proporcionan un procedimiento efectivo para encontrar la raíz a, sino que utilizan una demostración por reducción al absurdo. Consecuentemente no daremos ninguna demostración de este teorema[21].

[21] Una demostración elemental, aunque no sencilla, puede verse en: Rey Pastor, J., Pi Calleja, H. y Trejo, C. A. (1969). *Análisis Matemático*. Volumen I. Octava edición. Otra demostración basada en ideas geométricas puede verse en: Courant, R. y Robbins, H. (1964) *¿Qué es la matemática?* Editorial Aguilar.

Ejemplo: hallar las raíces de $P(x) = x^3 - (1+i)^2 + (-6+i)x + 6i$ sabiendo que admite a 3 como raíz.

Acorde al razonamiento explicado anteriormente, conocemos una raíz de P, $x = 3$, por lo tanto P es divisible por $(x-3)$. Aplicamos la regla de Ruffini.

	1	–1–i	–6+i	6i
3		3	6 –3i	–6i
	1	2 –i	–2i	0

Por lo tanto $P(x) = (x-3)[x^2 + (2-i)x - 2i]$

Factorizamos $Q(x) = x^2 + (2-i)x - 2i$, como es de segundo grado aplicamos la fórmula resolvente:

$x_{1,2} = \frac{(-2+i) \pm \sqrt{(2-i)^2 - 4 \,.1\, (-2i)}}{2} = \frac{(-2+i) \pm \sqrt{3+4i}}{2}$

$x_{1,2} = \frac{(-2+i) \pm \sqrt{5_{53°\,7'\,48''}}}{2} = \frac{(-2+i) \pm \sqrt{5}_{\frac{53°\,7'\,48'' + 360°\,k}{2}} \; con\; k=0,1}{2}$

$x_{1,2} = \frac{(-2+i) \pm (2+i)}{2} = \begin{cases} x_1 = i \\ x_2 = -2 \end{cases}$. Finalmente

$P(x) = (x-3)(x+2)(x-i)$ y las raíces de P son: $x_1 = 3$, $x_2 = -2$ y $x_3 = i$

Una consecuencia importante del TFA se enuncia en el siguiente,

Corolario: *todo polinomio, no constante con coeficientes complejos* $P = \sum_{i=0}^{n} b_i x^i \in \mathbb{C}[x]$ $(b_i \in \mathbb{C},\ b_n \neq 0)$ *se factoriza como producto de polinomios lineales, en la forma:* $P(x) = b_n(x-a_1)^{m_1}(x-a_2)^{m_2} \ldots (x-a_r)^{m_r}$ *donde*

$a_1, a_2, \dots, a_r$ *son las distintas raíces complejas de* P, $m_1, m_2, \dots, m_r$ *son las correspondientes multiplicidades y* b_n *el coeficiente principal del polinomio.*

D) Por la consecuencia de la definición de multiplicidad de las raíces, sabemos que:

$P(x) = [(x - a_1)^{m_1}(x - a_2)^{m_2} \dots (x - a_r)^{m_r}]\, Q(x)$ donde $Q(a_i) \neq 0$ para $0 < i \leq r$.

Q debe ser constante. Si suponemos que no es así, por el TFA Q debe tener alguna raíz $a \in \mathbb{C}$, pero toda raíz de Q es raíz de P, luego $a = a_i$ para algún i, lo que es una contradicción pues $Q(a_i) \neq 0$.

Como Q debe ser constante al igualar los coeficientes principales de ambos miembros, deducimos que Q debe coincidir con el coeficiente principal de P, quedando demostrado el corolario.

Comparando los grados de ambos miembros, en la descomposición del corolario anterior deducimos que $n = gr(P) = m_1 + m_2 + \dots + m_r$.

Observemos que esta suma representa la cantidad de raíces de P, si contamos las raíces múltiples de acuerdo con su multiplicidad. Esto nos proporciona la siguiente consecuencia:

Un polinomio $P \in \mathbb{C}[x]$ *de grado n tiene exactamente n raíces complejas si las contamos de acuerdo con su multiplicidad.*

7. Raíces complejas de polinomios reales

Recordemos brevemente, si $z = a + bi \in \mathbb{C}$ entonces su conjugado es $\overline{z} = a - bi \in \mathbb{C}$. La conjugación tiene propiedades de las cuales nos interesa particularmente:

Si $z, w \in \mathbb{C}$, entonces

a) $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$

b) $\overline{z+w} = \overline{z} + \overline{w}$ y $\overline{z.w} = \overline{z}.\overline{w}$

c) $z + \overline{z} = 2Re(z)$ y $z.\overline{z} = \|z\|^2$ son números reales.

Recordemos que un polinomio cuadrático con coeficientes reales $P(x) = ax^2 + bx + c$ y discriminante $\Delta = b^2 - 4ac < 0$ tiene dos raíces complejas que podemos escribir como $x_1 = \frac{-b+\sqrt{-\Delta}\, i}{2a}$ y $x_2 = \frac{-b-\sqrt{-\Delta}\, i}{2a}$. De esta forma, las raíces complejas de un polinomio cuadrático forman un par de raíces conjugadas.

Se puede generalizar este hecho a polinomios con coeficientes reales de mayor grado. Pero previamente consideremos un polinomio P, cuyos coeficientes son complejos $P(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{C}[x]$ y definamos el polinomio conjugado de P, del siguiente modo $\overline{P}(x) = \sum_{i=0}^{n} \overline{a_i} x^i$.

Como consecuencia de las propiedades antes mencionadas de la conjugación, si $z \in \mathbb{C}$ tenemos que si los polinomios $P, Q \in \mathbb{C}[x]$ y $z \in \mathbb{C}$, entonces:

a) $\overline{P}(z) = \overline{P}(\overline{z})$

b) $\overline{P+Q} = \overline{P} + \overline{Q}$ y $\overline{P.Q} = \overline{P}.\overline{Q}$

Particularmente si los coeficientes de P son reales, $P \in \mathbb{R}[x]$, tendremos que $\overline{P} = P$ y resulta que $\overline{\overline{P(z)}} = P(\overline{z})$. Más específicamente si $P(z) = 0$, tenemos que $P(\overline{z}) = 0$, esto significa que: las raíces complejas de un polinomio con coeficientes reales se presentan de a pares de raíces conjugadas.

Ahora estamos en condición de demostrar el siguiente teorema:

Teorema: *sea un polinomio con coeficientes reales* $P \in \mathbb{R}[x]$. *Si* $z = a + bi$ *es una raíz de P con multiplicidad* m, *entonces su complejo conjugado* $\overline{z} = a - bi$ *también es raíz de P con multiplicidad* m.

D) Como z es raíz de P con multiplicidad m, P admite la factorización: $P(x) = (x - z)^m \, Q(x)$ donde $Q(z) \neq 0$.

Tomando el polinomio conjugado, tenemos:

$\overline{P}\,(x) = \overline{(x - z)^m \, Q(x)} = \overline{(x - z)^m} \; . \overline{\mathrm{Q}}\,(x) = (x - \overline{z})^m \; \overline{Q}\,(x)$

Pero P es a coeficientes reales, entonces $\overline{P} = P$ y como $P(\overline{z}) = 0, \overline{Q}\,(\overline{z}) = \overline{Q}\,(x) \neq 0$, esto indica que $\overline{z}$ es una raíz de P de multiplicidad m. Quedando demostrado de esta manera el teorema.

Ejemplo: hallar las raíces de $P(x) = x^3 - 3x + 52$ sabiendo que admite como raíz al número complejo $2 - 3i$

Si P admite a $2 - 3i$ como raíz también admite a su conjugado $2 + 3i$, ambas raíces son simples. Por lo tanto P es divisible por $\big(x - (2 - 3i)\big)$ y $\big(x - (2 + 3i)\big)$, recurrimos a la regla de Ruffini:

	1	0	–3	52
2 – 3i		2 – 3i	–5 – 12i	–52
	1	2 – 3i	–8 – 12i	0
2 + 3i		2 + 3i	8 + 12i	
	1	4	0	

Por lo tanto $P(x) = [x - (2 - 3i)][x - (2 + 3i)](x + 4)$ y sus raíces son: $x_1 = 2 - 3i, \; x_2 = 2 + 3i \;\; y \;\; x_3 = -4$.

Se demostró que las raíces complejas de un polinomio con coeficientes reales se presentan de a pares de raíces conjugadas, este hecho tiene una consecuencia inmediata que se enuncia de la siguiente manera:

Todo polinomio de coeficientes reales y grado impar, tiene alguna raíz real.

En el ejemplo anterior podemos observar que $P \in \mathbb{R}[x]$, $gr(P) = 3$ y también que $x_1, x_2 \in \mathbb{C}$ mientras que $x_3 \in \mathbb{R}$.

8. Algoritmo de Euclides

Como en $\mathbb{K}[x]$ hemos definido el algoritmo de la división, podemos extender a los polinomios el algoritmo de Euclides para el cálculo del máximo común divisor. Estos dos conceptos con polinomios van en perfecta sincronía con sus idénticos en la aritmética de enteros, por lo tanto, recordaremos los conceptos para números enteros y luego los formalizaremos para polinomios.

Sean $a\ y\ b$ dos números enteros no nulos, el mayor entero d tal que $d|a\ \ y\ \ d|b$ se denomina *máximo común divisor* de $a\ y\ b$, denotándose como $mcd(a, b)$.

Para calcular el mcd, todos conocemos el método de descomponer en factores primos los dos números y tomar los factores comunes con el menor exponente con el que aparezcan.

Si bien es un método útil y sencillo para conseguir lo que deseamos, tiene un problema: si los números son muy grandes, o sus factores primos lo son, el cálculo se complica y se torna tedioso. Otro método que también

persigue el mismo objetivo, pero mediante divisiones sucesivas, es el algoritmo de Euclides.

"Para calcular el mcd de dos enteros no nulos a y b, dividimos el mayor entre el menor (supongamos que $a > b$), esta división nos proporcionará un cociente Q_1 y un resto R_1. Si $R_1 = 0$ entonces $mcd(a, b) = b$, sino dividimos el b entre el R_1 obteniendo un cociente Q_2 y un resto R_2. Si $R_2 = 0$ entonces $mcd(a, b) = R_1$; sino volvemos a dividir el divisor entre el resto, así sucesivamente. Esto significa que el $mcd(a, b)$ es el último resto no nulo que obtengamos en el procedimiento anterior".

Ejemplo: calcular el $mcd\ (848, 656)$

Aplicando descomposición en factores primos:

$848 = 2^4 . 53$

$656 = 2^4 . 41$

$mcd\ (848, 656) = 2^4 = 16$

Aplicando el algoritmo de Euclides, obtenemos:

	1	3	2	2	2
848	656	192	80	32	16
192	80	32	16	0	

Como el último resto no nulo es 16, $mcd\ (848, 656) = 16$.

Estamos en condiciones de formalizar el máximo común divisor de dos polinomios y su cálculo por el algoritmo de Euclides.

Definición: *sean* $A, B \in \mathbb{K}[x]$ *dos polinomios no nulos, se dice que el polinomio* $D \in \mathbb{K}[x]$ *es el máximo común divisor* (mcd) *entre* A *y* B*, si es el único polinomio mónico que cumple simultáneamente las siguientes condiciones:*

i) $D|A \ y \ D|B$ *es decir* D *es un divisor común de* A *y* B.

ii) Si $\widetilde{D} \in \mathbb{K}[x]$ *verifica que* $\widetilde{D}|A \ y \ \widetilde{D}|B$ *entonces* $\widetilde{D}|D$, es *decir si* $\widetilde{D}$ *es otro divisor común de* A *y* B*, entonces* $\widetilde{D}$ *divide a* D.

Ejemplo: determinar el mcd entre $A(x) = x^2 - 3x + 2$ y $B(x) = x^2 - 4x + 3$.

Es mucho más fácil determinar el mcd cuando los polinomios están expresados como producto de sus factores primos, ya que en ese caso es simplemente el producto de los factores comunes a ambos con el menor exponente.

$A(x) = (x-1)(x-2)$

$B(x) = (x-1)(x-3)$ Por lo tanto

$mcd(A, B) = (x-1)$

Como en general, el problema de descomponer polinomios en sus factores primos es complicado, otro procedimiento para el cálculo del mcd de dos polinomios, sin recurrir a la descomposición en factores primos, es el algoritmo de Euclides.

Por las propiedades de la divisibilidad de dos polinomios enunciadas en el apartado "Divisibilidad", todo divisor común a dos polinomios A *y* B, divide también al polinomio $R = A - B\,Q$ cualquiera sea el polinomio Q.

Recíprocamente, todo divisor común a $B\ y\ R$ divide también a A ya que $A = B\ Q + R$.

Entonces, el cálculo del mcd entre $A\ y\ B$ lleva a calcular el mcd de $B\ y\ R$, donde R es el resto de la división de A por B. Ahora podemos enunciar el algoritmo de Euclides en completa analogía con la aritmética de $\mathbb{Z}$.

Algoritmo de Euclides

Dados dos polinomios $A, B \in \mathbb{K}[x]$ y suponiendo que $gr(A) \geq gr(B)$, para hallar su mcd usamos el Algoritmo de Euclides, cuya disposición práctica es la siguiente:

	Q_1	Q_2	Q_3	. . .	Q_{n-1}	Q_n	Q_{n+1}
A	B	R_1	R_2	. . .	R_{n-2}	R_{n-1}	R_n
R_1	R_2	R_3	R_4	. . .	R_n	0	

Procedimiento:

Dividimos A en B, obteniendo un primer cociente Q_1 y un primer resto R_1 de modo que:

$A = Q_1 B + R_1$ donde $R_1 = 0$ o bien $0 < gr(R_1) < gr(B)$. Si $R_1 = 0 \implies mcd(A, B) = B$ sino volvemos a dividir B en R_1 obteniendo un nuevo cociente Q_2 y un nuevo resto R_2 de modo que se verifica:

$B = Q_2 R_1 + R_2$ donde $R_2 = 0$ ó bien $0 < gr(R_2) < gr(R_1)$

Mientras $R_n \neq 0$ podemos seguir dividiendo R_{n-1} por R_n obteniendo un nuevo cociente Q_{n+1} y un nuevo resto R_{n+1}. Es decir

$R_1 = Q_3 R_2 + R_3$ donde $gr(R_3) < gr(R_2)$

$R_2 = Q_4 R_3 + R_4$ con $gr(R_4) < gr(R_3)$ hasta que

$R_{n-1} = Q_{n+1}R_n + 0$

Esto ocurre $(R_{n+1} = 0)$ dado que la sucesión de restos tiene grado estrictamente decreciente: $gr(A) \geq gr(B) > gr(R_1) > gr(R_2) > \ ... > gr(R_n) > gr(R_{n+1}) > \ ...$ y como los grados son enteros y en virtud del principio del buen orden[22], en algún momento debemos tener $R_n = 0$, esto significa que el algoritmo de Euclides termina luego de un número finito de pasos. Cuando esto ocurre podemos demostrar que $R_{n-1} = mcd(A, B)$.

D) Sea D un divisor de $A\ y\ B$, entonces $A\ y\ B$ se pueden escribir de la siguiente manera: $A = D\ A_1\ \ y\ \ B = D\ B_1$

Como $A = Q_1B + R_1$ tenemos que

$R_1 = A - Q_1B$

$R_1 = D\ A_1 - Q_1D\ B_1$

$R_1 = D(A_1 - Q_1B_1)$ por lo tanto D divide a R_1.

Luego, cualquier divisor de $A\ y\ B$ lo es también de $B\ y\ R_1$.

Análogamente $B = Q_2R_1 + R_2$ tenemos que

$R_2 = B - Q_2R_1$

$R_2 = D\ B_1 - Q_2\ D\widetilde{R_1}$

$R_2 = D(B_1 - Q_2\widetilde{R_1})$ por lo tanto D divide a R_2

Luego cualquier divisor de $B\ y\ R_1$ lo es también de $R_1\ y\ R_2$.

Inductivamente el proceso es:

[22] Principio del buen orden o principio del mínimo: "Todo conjunto no vacío de números naturales tiene un elemento mínimo".

$mcd(A,B) = mcd(B,R_1) = mcd(R_1,R_2) = \ ... = mcd(R_{n-1},\ R_n) = R_n$ ya que $R_n|R_{n-1}\ y\ R_n|R_n$.

El siguiente teorema resume lo que demostramos hasta aquí:

Teorema: *sean* $A,B \in \mathbb{K}[x]$ *dos polinomios. Su máximo común divisor* $D = mcd(A,B)$ *siempre existe y se puede calcular utilizando el algoritmo de Euclides.*

Particularmente si $gr(R_n) \neq 0$, es decir R_n es un número real no nulo, los polinomios $A\ y\ B$ se dicen primos entre sí.

Ejemplo: calcular el mcd entre $A = x^5 - 1$ y $B = x^3 - 1$.

	x^2	x	$x+1$
$\boldsymbol{x^5 + 0x^4 + 0x^3 + 0x^2 + 0x - 1}$ $-x^5 \qquad x^2$	$\boldsymbol{x^3 + 0x^2 + 0x - 1}$ $-x^3 \qquad x$	$x^2 + 0x - 1$ $-x^2 + x$	$x - 1$
$x^2 - 1$	$x - 1$	$x - 1$ $-x + 1$ 0	

Por lo tanto $mcd(A,B) = x - 1$

Como consecuencia de la existencia del mcd entre dos polinomios, se puede relacionar a este con las raíces en común que pudieren tener dichos polinomios. Esto se enuncia de la siguiente manera:

Corolario: *sean* $A,B \in \mathbb{K}[x]$ *dos polinomios y* $D = mcd(A,B)$ *su máximo común divisor, entonces* $a \in \mathbb{K}$ *es una raíz en común de* $A\ y\ B$*, si y solo si* a *es raíz de* D.

D) Por el Corolario del Teorema del Resto, a es raíz de D si y solo si $(x-a)$ divide a D, lo cual ocurre si y solo si $(x-a)$ divide simultáneamente a $A\ y\ B$, lo cual a su vez sucede si y solo si a es raíz de ambos polinomios.

Continuando con el ejemplo anterior $A = x^5 - 1,\ B = x^3 - 1$ y $mcd(A, B) = x - 1$. Esto significa que la única raíz que comparten los polinomios $A\ y\ B$ es $x = 1$.

Para hallar las raíces de polinomios, en muchos casos es necesario saber si este admite o no raíces múltiples, habiendo ya establecido los conceptos de polinomio derivado y cálculo del mcd de dos polinomios, la cuestión planteada es mucho más fácil de resolver, mediante el siguiente teorema.

8.1. Admisión de raíces múltiples

Teorema: *sea el polinomio* $P \in \mathbb{K}[x]$, *las raíces múltiples de P son todas las raíces de* $D \in \mathbb{K}[x]$ *siendo* $D = mcd(P, P')$.

D) Consideremos un polinomio $P(x)$, mónico de grado n y cuyas raíces múltiples son x_1 de multiplicidad h, x_2 de multiplicidad k, . . . , x_n de multiplicidad p, entonces:

$P(x) = (x - x_1)^h\ (x - x_2)^k\ \ldots (x - x_n)^p$ donde $h + k + \ldots + p = n$,

derivando obtenemos:

$$P'(x) = h(x - x_1)^{h-1}\ (x - x_2)^k\ \ldots (x - x_n)^p + (x - x_1)^h\ k(x - x_2)^{k-1}\ \ldots (x - x_n)^p + \ldots + (x - x_1)^h\ (x - x_2)^k\ \ldots p(x - x_n)^{p-1}$$

$$P'(x) = (x - x_1)^{h-1}(x - x_2)^{k-1}\ \ldots (x - x_n)^{p-1}[h(x - x_2)^k\ \ldots (x - x_n)^p + k(x - x_1)^h + \ldots + (x - x_n)^p + \ldots + p(x - x_1)^h(x - x_2)^k]$$

Por lo tanto $D = mcd(P, P')$ es:

$D(x) = (x - x_1)^{h-1}(x - x_2)^{k-1}\ \ldots (x - x_n)^{p-1}$

Esto quiere decir que las raíces de $D(x)$ son las mismas que las de $P(x)$ pero con multiplicidad disminuida en una unidad, por lo tanto, las raíces simples de $P(x)$ no figurarán en D(x).

Quedando así demostrado el teorema.

Si el mcd entre $P\ y\ P'$ es independiente de x, o sea, si $P\ y\ P'$ son primos entre sí, el polinomio P carece de raíces múltiples.

Ejemplo: determinar las raíces múltiples de $P(x) = x^5 - x^4 - 5x^3 + x^2 + 8x + 4$.

Como es $P(x) = x^5 - x^4 - 5x^3 + x^2 + 8x + 4$ por lo tanto su polinomio derivado es $P'(x) = 5x^4 - 4x^3 - 15x^2 + 2x + 8$

Aplicando el algoritmo de Euclides, se obtiene $D(x) = mcd[P, P'\,] = x^3 - 3x - 2$

Factorizando $D(x) = x^3 - 3x - 2$ obtenemos $D(x) = (x+1)^2(x-2)$

En consecuencia, $P(x)$ tiene por raíces a $x_1 = -1$ con multiplicidad 3 y $x_2 = 2$ con multiplicidad 2.

Verificamos en $P(x)$, aplicando la regla de Ruffini.

	1	–1	–5	1	8	4
–1		–1	2	3	–4	–4
	1	–2	–3	4	4	0
–1		–1	3	0	–4	
	1	–3	0	4	0	
–1		–1	4	–4		
	1	–4	4	0		
2		2	–4			
	1	–2	0			
2		2				
	1	0				

9. Relaciones entre coeficientes y raíces

Consideremos un polinomio $P(x) \in \mathbb{K}[x]$, de segundo grado $P(x) = ax^2 + bx + c$ (1) y llamemos $x_1\ y\ x_2$ a sus raíces, entonces sabemos que P admite la siguiente factorización:

$P(x) = a(x - x_1)(x - x_2)$, por lo tanto

$P(x) = ax^2 - a\,x(x_1 + x_2) + a\,x_1x_2$ (2)

Igualando los segundos miembros de (1) y (2) obtenemos

$ax^2 + bx + c = ax^2 - a\,(x_1 + x_2)x + a\,x_1x_2$ de donde

$b = -a(x_1 + x_2) \implies -\frac{b}{a} = x_1 + x_2$

$c = a\,x_1x_2 \implies \frac{c}{a} = x_1x_2$

Es decir que la suma de las raíces es igual al coeficiente lineal sobre el coeficiente principal cambiado de signo, mientras que el producto de las raíces es igual al término independiente sobre el principal. Si bien esto es ya conocido, nos permite dos cuestiones: primero, hacer idéntico procedimiento para un polinomio de tercer grado y observar qué relación hay entre sus coeficientes y raíces; en segundo lugar, preguntarnos si estas relaciones que observamos particularmente para estos dos polinomios se pueden generalizar para un polinomio cualquiera.

Ahora P es $P(x) = ax^3 + bx^2 + cx + d$ (1) y sus raíces serán $x_1, x_2\ y\ x_3$, admitiendo la siguiente factorización:

$P(x) = a(x - x_1)(x - x_2)(x - x_3)$ por lo tanto

$P(x) = a[(x^2 - (x_1 + x_2)x + x_1x_2)(x - x_3)\,]$

$P(x) = a\,x^3 - a\,(x_1 + x_2 + x_3)x^2 + a\,(x_1x_2 + x_1x_3 + x_2x_3)x -$

$a\,x_1x_2x_3$ (2)

Igualando los segundos miembros de (1) y (2) , obtenemos las siguientes relaciones

$b = -a(x_1 + x_2 + x_3) \implies -\frac{b}{a} = x_1 + x_2 + x_3$

$c = a\,(x_1x_2 + x_1x_3 + x_2x_3) \implies \frac{c}{a} = x_1x_2 + x_1x_3 + x_2x_3$

$d = -a\,x_1x_2x_3 \implies -\frac{d}{a} = x_1x_2x_3$

Este hecho se puede generalizar para cualquier polinomio de grado arbitrario, del siguiente modo.

Sea $P = \sum_{i=0}^{n} a_i x^i$ (1) un polinomio de grado n en $\mathbb{K}[x]$. Su descomposición factorial es, por lo tanto:

$P = a_n(x - x_1)(x - x_2) \dots (x - x_n)$ donde x_i con $i = 1, 2, \dots, n$ son todas sus raíces complejas, simples o múltiples. Aplicando la propiedad distributiva, obtenemos:

$P = a_n x^n - a_n(x_1 + x_2 + \cdots + x_n)x^{n-1} + a_n(x_1x_2 + x_1x_3 + \dots + x_{n-1}x_n)x^{n-2} - a_n(x_1x_2x_3 + x_1x_2x_4 + \dots + x_{n-2}x_{n-1}x_n)x^{n-3} + \cdots + (-1)^n a_n\, x_1x_2 \dots x_n$ (2)

De (1) y (2) resulta

$a_{n-1} = -a_n(x_1 + x_2 + \cdots + x_n) \implies -\frac{a_{n-1}}{a_n} = x_1 + x_2 + \cdots + x_n$

$a_{n-2} = a_n(x_1x_2 + x_1x_3 + \dots + x_{n-1}x_n) \implies \frac{a_{n-2}}{a_n} = x_1x_2 + x_1x_3 + \dots + x_{n-1}x_n$

$a_{n-3} = -a_n(x_1x_2x_3 + x_1x_2x_4 + \dots + x_{n-2}x_{n-1}x_n) \implies -\frac{a_{n-3}}{a_n} = x_1x_2x_3 + x_1x_2x_4 + \dots + x_{n-2}x_{n-1}x_n$

. . .

$a_0 = (-1)^n a_n\, x_1x_2 \dots x_n \implies (-1)^n \frac{a_0}{a_n} = x_1x_2 \dots x_n$

Por lo tanto si $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ ... + a_1 x + a_0$, las relaciones anteriores se traducen como:

a) La suma de las raíces es igual al segundo coeficiente cambiado de signo, dividido por el coeficiente principal.

b) La suma de los productos binarios de las raíces es igual al tercer coeficiente dividido por el coeficiente principal.

Las mismas reglas valen para las sumas de los productos ternarios, cuaternarios, etc. con signos $-$ o $+$ alternativamente.

El producto de las n raíces es igual al término independiente dividido por el coeficiente principal, con signo $+$ o $-$, según que n sea par o impar respectivamente.

Ejemplo: determinar el polinomio mónico de grado tres cuyas raíces son: $x_1 = 1,\ x_2 = -1\ y\ x_3 = 2$

El polinomio buscado tiene la forma:

$P(x) = x^3 + a_2 x^2 + a_1 x + a_0$ donde debemos determinar: $a_2, a_1\ y\ a_0$. Por lo tanto

$$x_1 + x_2 + x_3 = -\frac{a_2}{a_3} \Longrightarrow a_2 = -(1 + (-1) + 2) \Longrightarrow a_2 = -2$$

$$x_1x_2 + x_1x_3 + x_2x_3 = \frac{a_1}{a_3} \Longrightarrow a_1 = 1(-1) + 1(+2) + (-1)2 \Longrightarrow a_1 = -1$$

$$x_1x_2x_3 = -\frac{a_0}{a_3} \Longrightarrow a_0 = -1(-1)2 \Longrightarrow a_0 = 2$$

Por lo tanto, el polinomio solicitado es: $P(x) = x^3 - 2x^2 - x + 2$

Comprobamos que las raíces dadas anulan al $P(x)$ propuesto.

$P(1) = (+1)^3 - 2\,(+1)^2 - (+1) + 2 = 0$

$P(-1) = (-1)^3 - 2(-1)^2 - (-1) + 2 = 0$

$P(2) = (+2)^3 - 2(+2)^2 - (+2) + 2 = 0$

CAPÍTULO 4

MATRICES

1. Introducción

En el Capítulo 3: "Polinomios", se hizo mención al libro chino *Chu Chang Suan Shu* o *Arte Matemático en Nueve Secciones* en referencia a problemas sobre la resolución de ecuaciones algebraicas. En este mismo libro se puede encontrar uno de los primeros ejemplos conocidos del uso de lo que con el correr de los años se conocería como matrices: "los cuadrados mágicos". Estos cuadros también fueron conocidos por los matemáticos árabes, posiblemente desde el siglo VII, quienes pudieron haberlos conocido de los matemáticos y astrónomos de la India. Otros resultados más amplios sobre los "cuadrados mágicos" aparecieron en Bagdad en el año 983, en la *Enciclopedia de la Hermandad de Pureza*. Entonces, se podría considerar que el estudio de los "cuadrados mágicos" fue un estudio empírico y desordenado acerca de lo que se estudiará en este capítulo: Matrices.

El primero en usar el término "matriz" fue el matemático inglés James Joseph Sylvester (1814-1897) en 1850, quien se refirió a ellas como un arreglo cuadrado de términos. En 1853, el matemático, físico y astrónomo irlandés sir William Rowan Hamilton (1805-1865) realizó importantes aportes a la teoría de matrices. Sin embargo, algunos historiadores de las

matemáticas consideran a Arthur Cayley como el fundador de la teoría de matrices, ya que fue quien formalizó las operaciones básicas de suma y multiplicación con estos nuevos elementos matemáticos, a pesar de que aparecen algunos indicios de estas dos operaciones en trabajos anteriores de Euler, Lagrange y Gauss.

Otros importantes aportes de Cayley a la teoría de matrices fueron: la introducción de la notación matricial como forma abreviada de escribir un sistema de m ecuaciones lineales con n incógnitas; la publicación de un escrito en donde se hace referencia, por primera vez, a la inversa de una matriz; la definición de otra operación básica con matrices: el producto por escalares, demostración de la asociatividad del producto de matrices, introducción de las potencias de una matriz, así como las matrices simétricas y antisimétricas.

Por último, uno de los primeros libros publicados en el siglo XX donde se trata a las matrices por su propio interés es *Introducción al álgebra superior*, escrito en 1907 por el matemático estadounidense Maxime Bôcher (1867-1918).

La importancia de las matrices radica en el hecho de haberse convertido en una poderosa herramienta tanto intramatemática –debido a sus numerosas aplicaciones en las matemáticas puras– como extramatemática, con aplicaciones en diversos campos científicos como estadística, economía, física, biología, gráficas por computadora, análisis

numérico y lenguajes de programación, ya que son muchos los conceptos que se pueden representar por medio de una matriz.

Finalmente, las matrices constituyen la forma más adecuada para ordenar, analizar y operar, de un modo sistemático y relativamente sencillo, complicados cálculos numéricos y algebraicos que involucran un gran número de datos.

De esta manera en este cuarto capítulo estudiaremos, luego de los conceptos básicos con matrices, aquellas dos operaciones y sus propiedades, que permiten afirmar –de manera similar a los tres capítulos anteriores– que el conjunto de matrices constituye un espacio vectorial: la suma de matrices y producto por un escalar. Continuando con la aritmética de matrices, estudiaremos el producto, sus propiedades para luego hacer una breve introducción a la potenciación de matrices.

Respecto de las operaciones que particularmente se hacen con estos nuevos entes matemáticos, estudiaremos la trasposición e inversión de matrices con sus respectivas propiedades.

De aquellas matrices que guardan una relación especial respecto de ellas mismas o de otras matrices, se analizarán las matrices simétricas, antisimétricas, sus propiedades, matrices triangulares, diagonales, matrices idempotentes e involutivas, terminando con las matrices ortogonales.

Las transformaciones elementales entre filas o columnas de una matriz permiten definir una herramienta muy importante: la equivalencia entre

matrices y las matrices escalonadas, estas últimas se obtendrán mediante dos algoritmos: eliminación de Gauss y Gauss–Jordan. Todo esto permitirá definir y calcular el rango de una matriz, sus propiedades, la condición de existencia y el cálculo de la matriz inversa.

El capítulo finaliza con los temas submatrices y ecuaciones matriciales.

1.1. Conceptos básicos

Para poder comprender la definición de matriz –desde un punto de vista netamente matemático– es necesario recordar brevemente los conceptos de Intervalo Natural Inicial y Producto Cartesiano.

Se llama Intervalo Natural Inicial, I_n, al conjunto de los n primeros números naturales, es decir que $I_n = \{1, 2, \dots, n\}$. Por ejemplo $I_5 = \{1, 2, 3, 4, 5\}$ ó $I_8 = \{1, 2, 3, 4, 5, 6, 7, 8\}$.

Dados dos intervalos naturales $I_m = \{1, 2, \dots, m\}$ e $I_n = \{1, 2, \dots, n\}$ consideremos el producto cartesiano entre ellos, $I_m \, X \, I_n$, es decir, dados estos dos conjuntos debemos formar un tercer conjunto pero de pares ordenados, en donde el primer componente pertenece al primer conjunto dado (I_m) mientras que el segundo componente pertenece al segundo conjunto (I_n).

$I_m \, X \, I_n =$

$\{(1,1), (1,2), \dots, (1,n), (2,1), (2,2), \dots, (2,n), (m,1), (m,2), \dots, (m,n)\}$

Con los elementos del producto cartesiano, definamos una función, f, cuyo dominio sea este conjunto y cuyo codominio sean los números reales o complejos. En este último caso, para generalizar se utilizará la letra $\mathbb{K}$.

Es decir que, mediante esta letra estamos haciendo referencia al conjunto de los números reales, $\mathbb{R}$, o bien al conjunto de números complejos, $\mathbb{C}$. Entonces, $f: I_m \, X \;\; I_n \rightarrow K$.

Esto significa que la función, f, hace corresponder a cada par ordenado de elementos del conjunto $I_m \, X \, I_n$ un escalar correspondiente a $\mathbb{K}$.

Definición: *se llama matriz de orden* m x n *(o dimensión* m x n*) con elementos en el cuerpo* $\mathbb{K}$[23] *a la función* $f: I_m \, X \;\; I_n \rightarrow K$*, donde* I_i *representa un Intervalo Natural Inicial.*

Indicando con a_{ij} la imagen de un elemento cualquiera (i, j) perteneciente al dominio de la función, tenemos para el elemento:

$(1,1) \rightarrow f(1,1) = a_{11}$

$(1,2) \rightarrow f(1,2) = a_{12}$

Así sucesivamente para cualquier elemento genérico de la función:

$(i,j) \rightarrow f(i,j) = a_{ij}$

Finalmente, para el último elemento de la función:

$(m,n) \rightarrow f(m,n) = a_{mn}$

El conjunto de imágenes caracteriza la matriz que denominaremos con una letra mayúscula, en este caso A por ser un elemento cualquiera de esta

[23] Un cuerpo o campo es un conjunto K con dos operaciones binarias, usualmente llamadas suma "+" y producto ".". La suma es asociativa, conmutativa y su elemento neutro es 0 y para cada elemento existe un inverso. El producto es asociativo, conmutativo, su elemento neutro es 1 y todo elemento distinto de cero tiene inverso, además el producto se distribuye respecto de la suma.

matriz: a_{ij}. Esto permite escribir todas las imágenes en un cuadro de m filas por n columnas de la siguiente manera:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

En cada fila se escriben las imágenes de todos los pares que tienen la misma primera componente y en cada columna se anotan las imágenes de los pares que tienen la misma segunda componente.

El elemento de la matriz A, que figura en la fila i y columna j se denota por a_{ij}. Una matriz puede denotarse sintéticamente como $A = \|a_{ij}\|_{m\,x\,n}$ donde $i = 1, \dots, m$ y $j = 1, \dots, n$. Por ejemplo, el elemento de la matriz A que figura en la quinta fila y tercera columna se identifica como a_{53}.

En general, el conjunto de todas las matrices de orden $m \times n$, con elementos en $\mathbb{K}$, se simboliza por $\mathbb{K}^{m \times n}$. En muchas ocasiones para referirse tanto a columnas como filas, indistintamente, suele usarse el término línea.

Particularmente, el conjunto de todas las matrices de orden $m \times n$ con elementos en $\mathbb{R}$ se indica con $\mathbb{R}^{m \times n}$ y si tiene elementos en $\mathbb{R}$ o $\mathbb{C}$ se indica mediante $\mathbb{C}^{m \times n}$

Ejemplo 1:

$A = \begin{pmatrix} 4 & -1 & 2 \\ -3 & 5 & 6 \end{pmatrix}$ es una matriz real de orden 2 x 3, que en muchas ocasiones se suele denominar como matriz rectangular 2 x 3. En esta matriz distinguimos los elementos estudiados anteriormente:

$I_m = \{1, 2\}\ e\ I_n = \{1, 2, 3\} \Longrightarrow I_m\ X\ I_n =$

$\{\ (1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3)\}$

$f(1, 1) = 4,\ f(1, 2) = -1,\ f(1, 3) = 2,\ f(2, 1) = -3, f(2,2) =$

$5\ y\ f(2, 3) = 6$

Cuando el número de filas es igual al número de columnas ($m = n$), la matriz se dice cuadrada y simplemente se dice que es una matriz cuadrada de orden n.

$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$ donde el conjunto $\{a_{ij}\}\, con\ i = j$ o sea $\{a_{11}, a_{22}, \dots, a_{nn}\}$ constituye la diagonal principal de la matriz A.

Traza de una matriz: *sea la matriz $A \in K^{n\,x\,n}$, se llama traza de la matriz A y se denota Tr(A), al escalar $Tr(A) = \sum_{i=1}^{n} a_{ii}$.*

Esto significa que la traza de una matriz es el escalar que se obtiene al sumar los elementos de su diagonal principal.

Ejemplo 2:

$B = \begin{pmatrix} 1 & -i & 2+i \\ 0 & 4 & \sqrt{3} \\ -4i & 12 & -6 \end{pmatrix}$ es una matriz compleja de orden 3 cuya diagonal principal está formada por los elementos: $1, 4\ y\ -6$, y cuya traza es Tr(A) = 1 + 4 + (–6) = –1

Ejemplo 3: construir una matriz cuadrada de orden 2 donde $a_{ij} = 3i - j^2$.

Debemos construir una matriz C del tipo:

$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ en donde cada elemento c_{ij} se obtiene aplicando la relación $3i - j^2$. Entonces:

$c_{11} = 3(1) - 1^2 = 2$ $\quad$ $c_{12} = 3(1) - 2^2 = -1$

$c_{21} = 3(2) - 1^2 = 5$ $\quad$ $c_{22} = 3(2) - 2^2 = 2$

Luego, la matriz pedida es:

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

Un caso particular en la definición de matrices es aquella en la cual todos sus elementos son cero, se conoce como matriz nula y suele designarse con la letra N. Es decir que:

$N = \|n_{ij}\|_{m\,x\,n}$ tal que $n_{ij} = 0 \ \forall i; \ \forall j.$

Por ejemplo $N = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ es la matriz nula de orden 3 x 2.

Igualdad de matrices: *dos matrices son iguales si y solo si tienen el mismo orden y sus elementos correspondientes (aquellos que pertenecen a igual fila e igual columna) son iguales.*

Dadas $A = \|a_{ij}\|_{m\,x\,n}$ y $B = \|b_{ij}\|_{m\,x\,n}$ resulta $A = B \iff a_{ij} = b_{ij} \forall i \, ; \forall j$

Ejemplo:

Las matrices $\mathrm{A} = \begin{pmatrix} 4 & -1 \\ 2 & 0 \end{pmatrix}$ y $\mathrm{B} = \begin{pmatrix} \mathrm{x}+2 & -1 \\ 2 & \mathrm{y} \end{pmatrix}$ serán iguales solo si $x = 2 \ e \ y = 0$, mientras que las matrices $\mathrm{C} = \begin{pmatrix} -3 & 2 \\ 2 & 0 \end{pmatrix}$ y $\mathrm{D} = \begin{pmatrix} -3 & 2 \\ -2 & 0 \end{pmatrix}$ no son iguales ya que, a pesar del mismo orden, no tienen todos sus elementos iguales ($c_{21} \neq d_{21}$). Por último, las matrices $\mathrm{E} = \begin{pmatrix} 1 & -4 \\ -2 & 3 \end{pmatrix}$ y $\mathrm{F} = \begin{pmatrix} 1 & -4 & 4 \\ -2 & 3 & 0 \\ 10 & 6 & -2 \end{pmatrix}$ no son iguales porque no tienen el mismo orden.

2. Operaciones con matrices

Si bien se pueden realizar varias operaciones algebraicas con matrices, se verá en primera instancia la suma y el producto por escalares, y acto seguido las propiedades de estas dos operaciones. Esto permitirá, como en los tres capítulos anteriores, referenciar el caso particular de las matrices como espacio vectorial.

2.1. Suma de matrices

Vamos a definir esta operación de la siguiente manera:

Se llama suma de dos matrices de orden $m \times n$ *a otra matriz del mismo orden cuyos elementos son la suma de los elementos correspondientes de las matrices dadas.*

Sean $A = \|a_{ij}\|_{m \times n}$ y $B = \|b_{ij}\|_{m \times n} \in K^{m \times n}$ entonces

$$A + B = C \;/\; C = \|c_{ij}\|_{m \times n} \; con \; c_{ij} = a_{ij} + b_{ij} \; \begin{cases} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n \end{cases}$$

Ejemplo:

$$A = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 4 \end{pmatrix} \; y \; B = \begin{pmatrix} -1 & 0 & 1 \\ 2 & -5 & 3 \end{pmatrix} \implies A + B = \begin{pmatrix} 1 & 1 & 4 \\ 1 & -3 & 7 \end{pmatrix}$$

Solamente se pueden sumar matrices del mismo orden, en cuyo caso se dicen sumables, es decir, que la suma de matrices de diferentes órdenes no está definida.

2.2. Producto por una matriz por un escalar

Esta operación se define de la siguiente manera:

El producto de un escalar λ por una matriz A o de una matriz A por un escalar λ, es la matriz cuyos elementos se obtienen al multiplicar el escalar por cada elemento de A.

Si $A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$ y $\lambda \in K \implies \lambda\, A = A\, \lambda = \|\lambda\, a_{ij}\|$

Ejemplo:

Sea $A = \begin{pmatrix} 3 & -2 & 0 \\ 2 & -1 & 1 \\ 1 & 6 & -4 \end{pmatrix}$ y $\lambda = -2 \implies \lambda\, A = -2A = \begin{pmatrix} -6 & 4 & 0 \\ -4 & 2 & -2 \\ -2 & -12 & 8 \end{pmatrix}$

De acuerdo a la definición del producto de una matriz por un escalar, si la matriz que se tiene es $A = \|a_{ij}\|_{m\,x\,n}$ y el escalar es:

a) $\lambda = -1$ entonces $\lambda\, A = -A$ matriz opuesta a A.

b) $\lambda = 0$ entonces $\lambda\, A = N$ matriz nula. En general si $\lambda\, A = N$ entonces deberá cumplirse que $\lambda = 0$ o $A = N$.

Propiedades de la suma de matrices y producto por un escalar

Veremos a continuación cuáles son las propiedades algebraicas que poseen las dos operaciones básicas definidas hasta ahora. Estas propiedades son similares a las que poseen las operaciones de suma y producto de números reales.

P_1) La suma es ley de composición interna en $\mathbb{K}^{m\,x\,n}$.

$\forall\, A = \|a_{ij}\|_{m\,x\,n}$ y $B = \|b_{ij}\|_{m\,x\,n} \in K^{m\,x\,n} \implies (A + B) \in K^{m\,x\,n}$

Esto significa que la suma, como está definida, es una operación que a cada par de elementos de $K^{m\,x\,n}$ le hace corresponder un único elemento de $K^{m\,x\,n}$.

D) Es inmediata por la definición de suma en $K^{m\,x\,n}$.

P2) La suma es asociativa en $K^{m\,x\,n}$.

$\forall\, A = \|a_{ij}\|_{m\,x\,n}, B = \|b_{ij}\|_{m\,x\,n}\; y\; C = \|c_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$ se cumple que

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

D) $(A + B) + C = \|a_{ij} + b_{ij}\| + \|c_{ij}\|$

$= \|(a_{ij} + b_{ij}) + c_{ij}\|$

$= \|a_{ij} + (b_{ij} + c_{ij})\|$ Porque la suma en K es asociativa.

$= \|a_{ij}\| + \|b_{ij} + c_{ij}\|$

$= A + (B + C)$ quedando demostrada la propiedad.

P3) La suma es conmutativa en $K^{m\,x\,n}$.

$\forall\, A = \|a_{ij}\|_{m\,x\,n}\; y\; B = \|b_{ij}\|_{m\,x\,n} \in K^{m\,x\,n} \implies A + B = B + A$

D) $A + B = \|a_{ij} + b_{ij}\|$

$= \|b_{ij} + a_{ij}\|$ Porque la suma en K es conmutativa.

$= \|b_{ij}\| + \|a_{ij}\|$

$= B + A$ quedando demostrada la propiedad.

P4) Existe un único elemento neutro para la suma en $K^{m\,x\,n}$.

El elemento neutro se denota con E.

$\exists E = \|e_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$ $\acute{u}nico$ / $\forall\, A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$: $A + E = E + A = A$.

D) Para demostrar la existencia del elemento neutro, proponemos la siguiente matriz: $E = \|e_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$, por lo tanto debe cumplirse que:

$A + E = A$

$\|a_{ij}\| + \|e_{ij}\| = \|a_{ij}\|$

$\|a_{ij} + e_{ij}\| = \|a_{ij}\|$

$a_{11} + e_{11} = a_{11} \implies e_{11} = 0$

... $\implies$ ...

$a_{m1} + e_{m1} = a_{m1} \implies e_{m1} = 0$

... $\implies$ ...

$a_{mn} + e_{mn} = a_{mn} \implies e_{mn} = 0$

$\therefore\ E = \|e_{ij}\|$ con $e_{ij} = 0\ \forall i\ \forall j$

La otra parte de la igualdad: $E + P = P$ se demuestra de manera similar.

Todo esto significa que el elemento neutro de la suma de matrices existe y es la matriz nula, ya que se definió a esta como aquella matriz cuyos elementos son 0, es decir que $E = N$.

Para demostrar la unicidad del neutro vamos a suponer que existen dos neutros para la suma: $N_1\ y\ N_2$, por lo tanto, si:

N_1 es neutro, entonces $N_1 + N_2 = N_2$

N_2 es neutro, entonces $N_2 + N_1 = N_1$

De donde se deduce, por ser los primeros miembros de ambas igualdades iguales, que $N_1 = N_2$ y, consecuentemente, el neutro es único. Quedando de esta manera demostrada la propiedad.

P5) Todo elemento en $K^{m\,x\,n}$ admite un único inverso aditivo u opuesto en $K^{m\,x\,n}$.

El elemento opuesto a uno cualquiera $A = \|a_{ij}\|_{m\,x\,n}$, se nota por O.

$\forall A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}: \exists\ O \in K^{m\,x\,n}\ único\ /\ A + O = O + A = E$

D) Para demostrar la existencia del opuesto se propone $O = \|o_{ij}\|_{m\,x\,n}$, por lo tanto debe cumplirse que:

$A + O = E$

$\|a_{ij}\| + \|o_{ij}\| = \|e_{ij}\|$

$\|a_{ij} + o_{ij}\| = \|e_{ij}\|$

$a_{11} + o_{11} = e_{11} \implies o_{11} = -a_{11}$

... $\implies$ ...

$a_{m1} + o_{m1} = e_{m1} \implies o_{m1} = -a_{m1}$

... $\implies$ ...

$a_{mn} + o_{mn} = e_{mn} \implies o_{mn} = -a_{mn}$

$\therefore\ O = \|o_{ij}\|_{m\,x\,n}$ con $o_{ij} = -a_{ij}\ \forall i\ \forall j$

La otra parte de la igualdad: $O + A = E$ se demuestra de manera similar. Todo esto significa que el inverso aditivo u opuesto a un elemento cualquiera sea este, $A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$, existe y es la matriz opuesta a la dada ya que cuando se definió producto por un escalar, si el escalar es –1 al multiplicarse por cualquier matriz, A, se obtiene la opuesta de A; es decir, $O(x) = -A = \|-a_{ij}\|_{m\,x\,n} \forall\ A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$.

Para demostrar la unicidad del elemento opuesto vamos a suponer que para un elemento $A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$ existen dos opuestos $A_1\ y\ A_2$, por lo tanto se debe cumplir que $A + A_1 = N$ y también que $A + A_2 = N$ de donde se deduce que $A_1 = A_2$. Consecuentemente, el inverso aditivo u opuesto de cada elemento de $K^{m\,x\,n}$ es único, quedando demostrada, así, la propiedad.

P6) El producto es ley de composición externa en $K^{m\,x\,n}$ con escalares en $\mathbb{K}$.

$\forall \alpha \in \mathbb{K},\ \forall\ A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n} \implies (\alpha\,.\,A) \in K^{m\,x\,n}$.

Esto significa que el producto de una matriz de $K^{m\,x\,n}$ por un escalar, como está definido, es una operación que, a cada par matriz–escalar le hace corresponder un único elemento de $K^{m\,x\,n}$.

D) La existencia de tal matriz es inmediata, por la definición de producto de una matriz de $K^{m\,x\,n}$ por un escalar.

P_7) El producto por un escalar satisface la asociatividad mixta.

$\forall \alpha, \beta \in \mathbb{K},\ \forall\, A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n} : \alpha(\beta\, A) = (\alpha\beta)\, A.$

D) $\alpha(\beta\, A) = \alpha\big(\beta\|a_{ij}\|\big)$

$= \alpha\|\beta\ a_{ij}\|$

$= \|\alpha\ (\beta\ a_{ij})\ \|$

$= \|\alpha\ \beta\ a_{ij}\ \|$

$= \|(\alpha\ \beta\ )\ a_{ij}\ \|$

$= (\alpha\ \beta\ )\|\ a_{ij}\ \|$

$= (\alpha\ \beta\ )\ A$ quedando demostrada la propiedad.

P_8) El producto por un escalar es distributivo respecto a la suma en $\mathbb{K}$.

$\forall \alpha, \beta \in \mathbb{K},\ \forall\, A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}: (\alpha + \beta)\, A = \alpha\,.A + \beta\,.A.$

D) $(\alpha + \beta)\, A = (\alpha + \beta)\, \|a_{ij}\|$

$= \|(\alpha + \beta)\ a_{ij}\|$

$= \|\alpha\ a_{ij} + \beta\ a_{ij}\|$

$= \|\ \alpha\ a_{ij}\ \| + \|\ \beta\ a_{ij}\ \|$

$= \alpha\ \|\ \ a_{ij}\ \| + \beta\ \|\ \ a_{ij}\ \|$

$= \alpha\,.A + \beta\,.A$ quedando demostrada la propiedad.

P9) El producto por un escalar es distributivo respecto a la suma en $K^{m\,x\,n}$.

$\forall \alpha \in \mathbb{K},\ \forall\, A = \left\|a_{ij}\right\|_{m\,x\,n}\ y\ B = \left\|b_{ij}\right\|_{m\,x\,n} \in K^{m\,x\,n} : \alpha(A+B) = \alpha\,.A + \alpha\,.B.$

D) $\alpha(A+B) = \alpha\,\left\|a_{ij} + b_{ij}\right\|$

$= \left\|\alpha\,(a_{ij} + b_{ij})\,\right\|$

$= \left\|\alpha\,a_{ij} + \alpha\,b_{ij}\,\right\|$

$= \left\|\alpha\,a_{ij}\,\right\| + \left\|\alpha\,b_{ij}\,\right\|$

$= \alpha\,\left\|a_{ij}\right\| + \alpha\,\left\|b_{ij}\right\|$

$= \alpha\,.A + \alpha.B$ quedando demostrada la propiedad.

P10) La unidad para los $\mathbb{K}$ es el neutro para el producto por un escalar.

$\forall\, A = \left\|a_{ij}\right\|_{m\,x\,n} \in K^{m\,x\,n}: 1\,A = A$

El conjunto de las matrices de $K^{m\,x\,n}$, para las cuales se definieron las operaciones de suma y producto por un escalar con las respectivas propiedades, determinan un espacio vectorial y cada matriz es un vector.

3. Otras operaciones con matrices

Después de haber estudiado la suma de matrices, podemos dar inicio al estudio y análisis de la otra operación de gran interés entre los elementos del conjunto de matrices: el producto. Esta operación clave posibilita, entre otras tantas cosas, representar algebraicamente sistemas de ecuaciones lineales, transformaciones lineales entre vectores y operaciones geométricas tales como rotaciones.

3.1. Producto de matrices

Esta operación la definimos de la siguiente manera:

Sean las matrices $A = \|a_{ij}\|_{m\,x\,p}$ y $B = \|b_{ij}\|_{p\,x\,n}$ *llamaremos producto de las matrices A y B, en ese orden, a la matriz C de orden* $m\ x\ n$ *cuyo elemento genérico* c_{ij} *es la suma de los productos de los elementos de la fila* i *de A, por los correspondientes elementos de la columna* j *de B.*

$A = \|a_{ij}\|_{m\,x\,p} \in K^{m\,x\,p}$ y $B = \|b_{ij}\|_{p\,x\,n} \in K^{p\,x\,n} \Longrightarrow A\,.\,B = C$ *de tal manera que*

$C = \|c_{ij}\|_{m\,x\,n}$ / $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ \ldots + a_{ip}b_{pj} =$

$\sum_{k=1}^{p} a_{ik}b_{kj}$ $\begin{cases} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{cases}$

Observemos que para poder multiplicar dos matrices, en el orden dado, el número de columnas de la primera debe ser igual al número de fila de la segunda, obteniéndose una tercera matriz que tiene tantas filas como la primera matriz y tantas columnas como la segunda. Esto quiere decir que, si tenemos una matriz de orden 2x3 y la multiplicamos por otra de orden 3x5, la matriz resultante será de orden 2x5.

Un elemento cualquiera, c_{ij}, de la matriz producto se obtiene multiplicando cada elemento de la fila i de la primera matriz por el correspondiente elemento de la columna j de la segunda matriz y luego sumándolos.

También se puede observar que el producto de matrices no cumple la propiedad conmutativa, ya que en el ejemplo anterior, si multiplicamos

la segunda (3x5) por la primera (2x3), no podríamos efectuar la operación puesto que, ahora, el número de columnas de la primera matriz no coincide con el número de filas de la segunda matriz.

Ejemplo:

Sean las matrices $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & -2 \\ 0 & -1 & -1 & 1 \end{pmatrix}$ y $C = \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix}$

Calcular, de ser posible, a) A.B, b) B.A y c) A.C

a) Cálculo de A.B

El producto de estas dos matrices es posible porque A es de orden 2x3 y B de orden 3x4, por lo tanto, la matriz producto, C, será de orden 2x4.

Para realizar el producto de matrices de forma práctica y manual, dividiremos una parte de la hoja en cuatro cuadrantes, en el primero colocaremos la segunda matriz, B, mientras que en el tercero ubicaremos la primera matriz, A. Esto nos facilitará sumar los productos de cada elemento de las filas de A por los correspondientes elementos de las columnas de B y obtener la matriz producto, C, en el cuarto cuadrante. Todo esto se muestra a continuación.

			1	2	**3**	4
		B ↓	1	0	**1**	–2
			0	–1	**–1**	1
A →						
1	**0**	**–1**	1	3	**4**	3
1	2	–2	3	4	7	–2
			Matriz Producto: C			

Aquí, el elemento $c_{13}(4)$ se obtiene de la siguiente manera: 1.3 + 0.1 + (–1).(–1) = 4.

De manera similar se obtienen los otros elementos de la matriz producto.

b) El producto B.A no es posible, ya que B es de orden 3x4, mientras que A es de orden 2x3, es decir que el número de columnas de B no coincide con el número de filas de A.

c) Cálculo de A.C

El producto de estas dos matrices es posible y se procede de manera similar a lo realizado en a).

Hacemos la disposición práctica y colocamos en la parte inferior izquierda la matriz A y en la parte superior derecha la matriz C.

		C	–2 2
	A		3
1	0	–1	–5
1	2	–2	–4

En el producto A.B, de dos matrices, se dice que A multiplica a B por la izquierda (o que A premultiplica a B) y que B multiplica a A por derecha (o que B posmultiplica a A).

Propiedades del producto del producto de matrices

P_1) Conmutatividad

El producto de matrices, en general, no es conmutativo.

Si $A \in K^{m \times n}$ y $B \in K^{n \times m}$, coexisten los productos $A.B \in K^{m \times m}$ y $B.A \in K^{n \times n}$ pero son distintos si $m \neq n$.

Si $m = n$ tanto $A.B$ como $B.A$ son matrices del tipo $n \times n$, es decir cuadradas y pertenecientes a $K^{n\,x\,n}$ pero en general se verifica que $A.B \neq B.A$. Si dadas dos matrices $A\ y\ B$ se verifica que $A.B = B.A$ las matrices se dicen conmutables.

P2) Asociatividad

El producto de matrices es asociativo

Si $A \in K^{n \times m}$, $B \in K^{m \times r}$ y $C \in K^{r \times s}$ entonces $D \in K^{n \times s}$ siendo la matriz $D = (A.B).C = A.(B.C)$

D) Observemos, en primer lugar, que tanto $A.B.C \in K^{n \times s}$ como $(A.B).C \in K^{n \times s}$ y también $A.(B.C) \in K^{n \times s}$.

Tomemos particularmente los productos $(A.B).C = A.(B.C)$

Para cada $1 \leq i \leq n$, $1 \leq j \leq s$ se tiene:

$\big((A.B).C\big)_{ij} = \sum_{k=1}^{r}\big(\sum_{p=1}^{m} a_{ip}b_{pk}\big)c_{kj} = \sum_{k=1}^{r}\big(\sum_{p=1}^{m} a_{ip}b_{pk}c_{kj}\big) =$

$\sum_{p=1}^{m}\big(\sum_{k=1}^{r} a_{ip}b_{pk}c_{kj}\big) =$

$= \sum_{p=1}^{m} a_{ip}\,\big(\sum_{k=1}^{r} b_{pk}c_{kj}\big) = \sum_{p=1}^{m} a_{ip}\,(B.C)_{pj} = \big(A.(B.C)\big)_{ij}$. Quedando demostrada de esta manera la propiedad.

P3) Distributividad

El producto de matrices es distributivo con respecto a la suma de matrices.

Si $A, B \in K^{m \times n}$ y $C \in K^{n \times p}$ entonces $(A + B).C = A.C + B.C$ que indica que el producto de matrices es distributivo a derecha con respecto a la suma de matrices.

Si $A, B \in K^{m \times n}$ y $C \in K^{p \times m}$ entonces $C\,.(A + B) = C.A + C.B$ que indica que el producto de matrices es distributivo a izquierda con respecto a la suma de matrices.

D) Demostraremos la primera parte de esta propiedad: $(A + B).C = A.C + B.C$.

Como $A, B \in K^{m \times n}$ entonces la suma $(A + B) \in K^{m \times n}$ y como $C \in K^{n \times p}$, en consecuencia la matriz producto $[(A + B).C] \in K^{m \times p}$. Llamemos a esta última matriz $H \in K^{m \times p}$. En el segundo miembro de la igualdad, la matriz $(A.C) \in K^{m \times p}$ y la matriz $(B.C) \in K^{m \times p}$, en consecuencia la matriz suma: $(A.C + B.C) \in K^{m \times p}$. Llamemos a esta última matriz $Q \in K^{m \times p}$.

En resumen, las matrices $H \in K^{m \times p}$ y $Q \in K^{m \times p}$ deben ser iguales, habiéndose demostrado ya, que tienen la misma dimensión. Solo resta demostrar que los elementos correspondientes (ocupan la misma posición) son iguales.

Un elemento cualquiera de la matriz H, h_{ij}, se obtiene multiplicando la fila i de $(A + B)$ por la columna j de C, es decir:

$h_{ij} = (a_{i1} + b_{i1})c_{1j} + (a_{i2} + b_{i2})c_{2j} + \ ... + (a_{in} + b_{in})c_{nj}$ escribiendo esto como sumatoria

$h_{ij} = \sum_{k=1}^{n}(a_{ik} + b_{ik})\ c_{kj}$ por propiedades de la sumatoria

$h_{ij} = \sum_{k=1}^{n} a_{ik}\, c_{kj} + \sum_{k=1}^{n} b_{ik}\, c_{kj}$

$h_{ij} = (ac)_{ij} + (bc)_{ij}$

$h_{ij} = q_{ij}$ lo que significa que las matrices H y Q tienen los elementos correspondientes, iguales. Por lo tanto se cumple que $(A + B).C = A.C + B.C$.

La segunda parte de la propiedad se demuestra de manera similar. De este modo queda demostrada la propiedad distributiva del producto de matrices respecto a la suma de matrices.

P4) Asociatividad respecto al producto por un escalar

El producto de matrices es asociativo respecto al producto por un escalar

Si $A \in K^{m \times n}\ y\ B \in K^{n \times p}\ y\ \alpha \in K \implies (\alpha\, A)B = A\,(\alpha\, B)$

D) La matriz $(\alpha A) \in K^{m \times n}$ y como $B \in K^{n \times p}$ significa que $(\alpha\, A)B \in K^{m \times p}$, llamaremos a esta matriz $H \in K^{m \times p}$. Por otro lado $(\alpha\, B) \in K^{n \times p}$ y como $A \in K^{m \times n}$, significa que la matriz $A\ (\alpha\, B) \in K^{m\, x\, p}$, llamaremos a esta matriz $Q \in K^{m\, x\, p}$.

Dos matrices H y Q son iguales si tienen la misma dimensión y sus respectivos elementos correspondientes iguales. La primera parte de esta proposición ya está demostrada pues tanto H como Q son de dimensión mxp. Para la segunda parte, tomamos un elemento cualquiera de H, h_{ij}, que se obtiene multiplicando la fila i de (αA) por la columna j de B, es decir:

$h_{ij} = (\alpha a)_{i1}\, b_{1j} + (\alpha a)_{i2}\, b_{2j} + \ ... + (\alpha a)_{in}\, b_{nj}$ expresado como sumatoria

$h_{ij} = \sum_{k=1}^{n} (\alpha a)_{ik}\, b_{kj}$

$h_{ij} = \sum_{k=1}^{n} a_{ik}\ (\alpha b)_{kj}$

$h_{ij} = q_{ij}$ lo que significa que las matrices H y Q tienen los elementos correspondientes, iguales. Por lo tanto se cumple que $(\alpha\ A)B = A\ (\alpha\ B)$ quedando demostrada la propiedad.

P5) El hecho de que el producto de dos matrices sea la matriz nula, no implica necesariamente que alguna de las matrices sea la matriz nula.

Sean $A \in K^{m \times n}\ y\ B \in K^{n \times p}$. Si $A\ .\ B = N \nRightarrow A = N \vee B = N$

Sean las matrices no nulas, A = $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ y B = $\begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix}$ su producto es

A. B = $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

P6) Sean las matrices $A \in K^{n \times m},\ B \in K^{m \times p}\ y\ C \in K^{m \times q}$ Si $AB = AC$ no implica necesariamente que $B = C$

Sean las matrices A = $\begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix}$, B = $\begin{pmatrix} 3 & 1 \\ 4 & -7 \end{pmatrix}$ y C = $\begin{pmatrix} 1 & -1 \\ 2 & -9 \end{pmatrix}$

Los productos A. B = $\begin{pmatrix} 2 & -16 \\ 1 & -8 \end{pmatrix}$ y A. C = $\begin{pmatrix} 2 & -16 \\ 1 & -8 \end{pmatrix}$ son iguales y $B \neq C$.

En el álgebra de los números reales, si $a, b, c \in \mathbb{R}\ y\ ab = 0$, es necesariamente $a = 0$, o bien $b = 0$. También $ab = ac$ necesariamente es $b = c$. En cambio, en el álgebra matricial, como quedó establecido en P5) y P6), esto puede no ocurrir.

P7) Propiedad de la identidad multiplicativa

Para cada $n \in \mathbb{N}$, sea $I_n \in K^{n \times n}$ definida por $(I_n)_{ij} = \begin{cases} 1\ si\ \ i = j \\ 0\ si\ i \neq j \end{cases}$.

Entonces, si $A \in K^{n \times m}$ se verifica: $I_n\ .\ A = A\ .\ I_m = A$.

La matriz I_n se denomina matriz identidad de $K^{n \times n}$

D) Sean $1 \leq i \leq n,\ 1 \leq j \leq m$. Se tiene:

$(I_n \,.\, A)_{ij} = \sum_{k=1}^{n} (I_n)_{ik} \,.\, a_{kj} = 1 \,.\, a_{ij} = a_{ij} = A$

De manera similar se demuestra que $A \,.\, I_m = A$. Quedando demostrada la propiedad.

3.2. Potencia de matrices

Definida la multiplicación de matrices, se puede definir la potenciación de estas de la siguiente manera.

Sea la matriz $A \in K^{n \times n}$ *definimos las potencias de la matriz A como* $A^0 = I_n$, $A^1 = A$ *y* $A^{k+1} = A^k \,.\, A$ *para* $k = 1, 2, 3, \ \ldots$.

Esta definición permite obtener las potencias de una matriz cuadrada para exponentes enteros positivos.

Ejemplo 1:

Sea $A = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} \implies A^2 = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 15 & 16 \end{pmatrix}$

$A^3 = A^2 . A = \begin{pmatrix} 1 & 0 \\ 15 & 16 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 63 & 64 \end{pmatrix}$

Ejemplo 2:

Sea la matriz $A = \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}$ se puede verificar que, $A^n = N_{3\,x\,3}$ para $n \geq 3$.

$A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 3 & 3 & 9 \\ -1 & -1 & -3 \end{pmatrix};\ A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Ejemplo 3:

Demostrar que, en general, $(A+B)^2 \neq A^2 + 2AB + B^2$

D) $(A+B)^2 = (A+B)(A+B)$

$(A+B)^2 = (A+B)\,A + (A+B)B$

$(A+B)^2 = A^2 + BA + AB + B^2$

En general $AB \neq BA$, por lo tanto, salvo que ocurra lo contrario,

$A^2 + BA + AB + B^2 \neq A^2 + 2AB + B^2$ quedando demostrado lo pedido.

4. Trasposición de matrices

Dadas las matrices $A = \|a_{ij}\|_{m\times n} \in K^{m\times n}$ y $B = \|b_{ij}\|_{n\times m} \in K^{n\times m}$ se denomina B traspuesta de A y se indica $B = A^t$ si y solo si B tiene como filas y columnas, las columnas y filas de A (respetando el orden).

Si $A \in K^{m\times n} \Longrightarrow B = A^t \Leftrightarrow B = \|b_{ij}\|_{n\times m} con\ b_{ij} = a_{ji}\ \forall i\ \forall j$

Ejemplo:

Si $A = \begin{pmatrix} -1 & 2 \\ 3 & 5 \\ 4 & -2 \end{pmatrix}$ por lo tanto $B = A^t = \begin{pmatrix} -1 & 3 & 4 \\ 2 & 5 & -2 \end{pmatrix}$

Observemos que $A \in \mathbb{R}^{3\times 2}$ entonces $A^t \in \mathbb{R}^{2\times 3}$, además se cumple que $b_{ij} = a_{ji}\ \forall i, \forall j$ por ejemplo $a_{21} = b_{12}$o bien $a_{32} = b_{23}$.

Propiedades de la trasposición de matrices

P_1) Involución

La traspuesta de la traspuesta de una matriz A, es la misma matriz A.

Sea $A = \|a_{ij}\|_{m\times n} \in K^{m\times n} \Longrightarrow (A^t)^t = A$

D) Según la definición de trasposición de matrices, $A^t \in \mathrm{K}^{n\,x\,m}$, luego $(A^t)^t \in \mathrm{K}^{m\,x\,n}$. Por lo tanto, las matrices $A\,y\,(A^t)^t$ tienen la misma dimensión. Veamos ahora un elemento cualquiera de la matriz $(A^t)^t$:

$\left((a_{ij})^{\,t}\right)^t = (a_{ji})^{\,t} = a_{ij} = A$ esto significa que los elementos correspondientes de las dos matrices, son iguales.

Si ambas matrices tienen la misma dimensión y los elementos correspondientes iguales, quiere decir que dichas matrices son iguales. Quedando demostrada de esta manera la propiedad.

P2) Traspuesta de la suma de dos matrices

La traspuesta de una suma de matrices es igual a la suma de las traspuestas de cada uno de los sumandos

Sean $A = \|a_{ij}\|_{m\times n}$ y $B = \|b_{ij}\|_{m\times n} \in K^{m\times n} \implies (A+B)^t = A^t + B^t$

D) La matriz $(A+B)$ es de dimensión mxn, por lo tanto $(A+B)^t$ es de dimensión nxm. Las matrices $A^t y\, B^t$ son de dimensión nxm y, por lo tanto, la matriz $(A^t + B^t)$ es de dimensión nxm. Lo anterior demuestra que las matrices $(A+B)^t$ y $A^t + B^t$ tienen la misma dimensión.

Tomemos un elemento cualquiera de la matriz $(A+B)^t$:

$\left(a_{ij} + b_{ij}\right)^t = \left(a_{ji} + b_{ji}\right) = (a_{ji}) + (b_{ji}) = A^t + B^t$ esto significa que los elementos correspondientes de las matrices $(A+B)^t$ y $A^t + B^t$ son iguales.

Todo esto demuestra que las matrices $(A+B)^t$ y $A^t + B^t$ son iguales, con lo que queda demostrada la propiedad.

P3) Traspuesta del producto de un escalar por una matriz

La traspuesta del producto de un escalar por una matriz es igual al producto del escalar por la traspuesta de la matriz.

Sea $A = \|a_{ij}\|_{m \times n} \in K^{m \times n}\ y\ \alpha \in K \implies (\alpha\ A)^t = \alpha\ A^t$

D) La matriz $(\alpha\ A)^t$ es de dimensión nxm y la matriz $\alpha\ A^t$ es de dimensión n x m, por lo tanto ambas matrices tienen la misma dimensión.

Tomemos un elemento cualquiera de la matriz $(\alpha\ A)^t$:

$(\alpha\ a_{ij})^t = \alpha\ (a_{ji}) = \alpha\ A^t$ lo que significa que ambas matrices, $(\alpha\ A)^t\ y\ \alpha\ A^t$ tienen los elementos correspondientes, iguales.

Todo esto quiere decir que las matrices, $(\alpha\ A)^t\ y\ \alpha\ A^t$ son iguales, quedando de esta forma demostrada la propiedad.

P_4) Traspuesta del producto de dos matrices

La traspuesta del producto de dos matrices es igual al producto de las traspuestas de dichas matrices, pero con el orden conmutado.

Sea $A = \|a_{ij}\|_{m \times p} \in K^{m \times p}\ y\ \ B = \|b_{ij}\|_{p \times n} \in K^{p \times n} \implies (A\ B)^t = B^t\ A^t$

D) La matriz $A\ B$ es de dimensión m x n, por lo tanto la matriz $(A\ B)^t$ es de dimensión n x m. La matriz B^t es de dimensión n x p mientras que A^t es de dimensión pxm, por lo tanto la matriz $B^t\ A^t$ es de dimensión n x m. Lo anterior demuestra que las matrices $(A\ B)^t\ y\ B^t\ A^t$ tienen la misma dimensión.

Tomemos un elemento cualquiera de la matriz $(A\ B)^t$:

$((a\ b)_{ij})^t = (ab)_{ji} = a_{j1}b_{1i} + a_{j2}b_{2i} + \ ... + a_{jp}b_{pi}$ que se puede expresar como sumatoria

$\left((a\ b)_{ij}\right)^t = (ab)_{ji} = \sum_{k=1}^{p} a_{jk} b_{ki}$ (1)

Ahora, tomemos un elemento cualquiera de la matriz $B^t\ A^t$

$(b^t a^t)_{ij} = b_{i1}^t\ a_{1j}^t + b_{i2}^t\ a_{2j}^t\ +\ ...+\ b_{ip}^t\ a_{pj}^t$ que se puede expresar como sumatoria

$(b^t a^t)_{ij} = \sum_{k=1}^{p} b_{ik}^t\ a_{kj}^t = \sum_{k=1}^{p} b_{ki}\ a_{jk} = \sum_{k=1}^{p} a_{jk}\ b_{ki}$ (2)

Las expresiones (1) y (2) son iguales, lo que significa que los elementos correspondientes de las matrices $(A\ B)^t\ y\ B^t\ A^t$ son iguales.

Todo lo anterior muestra que las matrices $(A\ B)^t\ y\ B^t\ A^t$ son iguales, quedando demostrada la propiedad.

5. Matrices simétricas y antisimétricas

Matriz simétrica: *una matriz A, cuadrada de orden n, es simétrica si y solo si es igual a su traspuesta*

$A = \left\|a_{ij}\right\|_{n\times n} \in K^{n\times n}\ es\ Simétrica\ \Leftrightarrow A = A^t\ \Leftrightarrow a_{ij} = a_{ji}\ \forall i, \forall\ j$

Ejemplo:

La matriz A $= \begin{pmatrix} 3 & -7 & 0 \\ -7 & 1 & 5 \\ 0 & 5 & -2 \end{pmatrix}$ es simétrica

Matriz antisimétrica: *una matriz A, cuadrada de orden n, es antisimétrica si y solo si es igual a la opuesta de su traspuesta.*

$A = \left\|a_{ij}\right\|_{n\times n} \in K^{n\times n}\ es\ Antisimétrica \Leftrightarrow A = -A^t\ \Leftrightarrow a_{ij} = -a_{ji}\ \forall i, \forall\ j$

En una matriz antisimétrica, los elementos de la diagonal principal son todos ceros, porque en la diagonal principal $i = j$ o sea que $a_{ii} = -a_{ii} \Leftrightarrow 2a_{ii} = 0 \Leftrightarrow a_{ii} = 0$

Ejemplo:

La matriz $A = \begin{pmatrix} 0 & -1 & 5 \\ 1 & 0 & 3 \\ -5 & -3 & 0 \end{pmatrix}$ es antisimétrica.

Propiedades de las matrices simétricas y antisimétricas

P_1) El producto de una matriz por su traspuesta es igual a una matriz simétrica (no es necesario que A sea cuadrada)

Si $A = \|a_{ij}\|_{m \times n} \in K^{m \times n} \implies A \,.\, A^t = (A \,.\, A^t)^t$

D) $A \,.\, A^t = (A \,.\, A^t)^t$

$A \,.\, A^t = (A^t)^t \,.\, A^t$

$A \,.\, A^t = A \,.\, A^t$ quedando, de esta forma, demostrada la propiedad.

P_2) La suma de toda matriz cuadrada y de su traspuesta es una matriz simétrica.

Si $A = \|a_{ij}\|_{n \times n} \in K^{n \times n} \implies (A + A^t) = (A + A^t)^t$

D) $A + A^t = (A + A^t)^t$

$A + A^t = A^t + (A^t)^t$

$A + A^t = A^t + A$

$A + A^t = A + A^t$ quedando, de esta forma, demostrada la propiedad.

P_3) La diferencia de toda matriz cuadrada con su traspuesta es antisimétrica.

Si $A = \|a_{ij}\|_{n \times n} \in K^{n \times n} \implies (A - A^t) = -(A - A^t)^t$

D) $A - A^t = -(A - A^t)^t$

$A - A^t = -(A^t - (A^t)^t)$

$A - A^t = -(A^t - A)$

$A - A^t = A - A^t$ quedando, de esta forma, demostrada la propiedad.

P4) Toda matriz cuadrada es la suma de una matriz simétrica y de una matriz antisimétrica.

D) Si $A = \|a_{ij}\|_{n \times n} \in K^{n \times n}$ de las propiedades P2) y P3) resulta que $\frac{1}{2}(A + A^t)$ es simétrica y $\frac{1}{2}(A - A^t)$ es antisimétrica, por lo tanto

$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$

$A = \frac{1}{2}A + \frac{1}{2}A^t + \frac{1}{2}A - \frac{1}{2}A^t$

$A = \frac{1}{2}A + \frac{1}{2}A$

A = A quedando, de esta forma, demostrada la propiedad.

Ejemplo:

Sea la matriz $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$ entonces

$\frac{1}{2}\left[\begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}\right] = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

$\frac{1}{2}\left[\begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}\right] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ Por lo tanto

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

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

6. Matrices triangulares

Matriz triangular superior: *una matriz A cuadrada de orden n cuyos elementos a_{ij} son iguales a 0 para todo $i > j$ se llama matriz triangular superior.*

$A = \|a_{ij}\|_{n \times n} \in K^{n \times n}$ *es Triangular superior* $\Leftrightarrow a_{ij} = 0 \; \forall \; i > j$

Ejemplo:

Para que una matriz cualquiera $A \in \mathbb{R}^{3\times3}$ sea triangular superior debe tener la siguiente forma, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$ es decir, los elementos que están debajo de la diagonal principal son ceros.

Matriz triangular inferior: *una matriz A cuadrada de orden n cuyos elementos* a_{ij} *son iguales a 0 para todo* $i < j$ *se llama matriz triangular inferior.*

$A = \|a_{ij}\|_{n\times n} \in K^{n\times n}\ es\ Triangular\ inferior \iff a_{ij} = 0\ \forall\ i < j$

Ejemplo:

Para que una matriz cualquiera $A \in \mathbb{R}^{3x3}$ sea triangular inferior debe tener la siguiente forma $A = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ es decir, los elementos que están encima de la diagonal principal son ceros.

Puede demostrarse que el producto de dos matrices triangulares superiores (inferiores) es una matriz triangular superior (inferior).

Por ejemplo, sean las matrices $A\ y\ B\ \in \mathbb{R}^{3\times3}$:

$A = \begin{pmatrix} 2 & 3 & 4 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}\ y\ B = \begin{pmatrix} -1 & -3 & -5 \\ 0 & -2 & -4 \\ 0 & 0 & -3 \end{pmatrix}$ el producto de ella es:

$A.B = \begin{pmatrix} -2 & -12 & -34 \\ 0 & -8 & -31 \\ 0 & 0 & -18 \end{pmatrix}$

7. Matrices diagonales

Matriz diagonal: *una matriz A cuadrada de orden n en la cual los elementos que no están en la diagonal principal son todos ceros, se denomina matriz diagonal.*

$A = \|a_{ij}\|_{n \times n} \in K^{n \times n}\ es\ Diagonal \Leftrightarrow a_{ij} = 0\ \forall\ i \neq j$

Ejemplo:

Para que una matriz cualquiera, por ejemplo $A \in \mathbb{R}^{33}$ sea diagonal, debe cumplirse que $A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$ es decir, los elementos que no figuran en la diagonal principal son ceros.

Matriz escalar: *una matriz A cuadrada de orden n diagonal, en la cual los elementos de la diagonal principal son iguales a un escalar, se denomina matriz escalar.*

$A = \|a_{ij}\|_{n \times n} \in K^{n \times n}\ es\ Escalar \Leftrightarrow \begin{cases} a_{ij} = 0\ \forall\ i \neq j \\ a_{ij} = \alpha\ \forall\ i = j \end{cases}$

Ejemplo:

Para que una matriz $A \in \mathbb{R}^{3\times3}$ sea escalar debe cumplirse que $A = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}$ es decir, los elementos que figuran en la diagonal principal son iguales a un escalar.

Observemos que para que una matriz cuadrada de orden n sea escalar, debe ser, en primer lugar, una matriz diagonal, esto significa que toda matriz escalar es diagonal pero no toda matriz diagonal es escalar.

Matriz unidad o identidad: *una matriz A cuadrada de orden n escalar, en la cual los elementos de la diagonal principal son iguales a uno, se denomina matriz unidad o identidad.*

A la matriz identidad de orden n se la denomina I_n

$$I = \|a_{ij}\|_{n\times n} \in K^{n\times n} \; es\; Identidad \iff \begin{cases} a_{ij} = 0 \; \forall \; i \neq j \\ a_{ij} = 1 \; \forall \; i = j \end{cases}$$

Esta definición es más amplia y coincide con la dada en P7) (identidad multiplicativa) propiedades del producto de matrices.

El producto de una matriz A por la matriz identidad del orden adecuado da como resultado la misma matriz A.

Si $A = \|a_{ij}\|_{m\times n} \in K^{m\times n} \implies A\, I_n = A\, I_m = A$

Particularmente, si A es una matriz cuadrada de orden n e I, la matriz identidad del mismo orden se verifica $A\, I = I\, A = A$

Ejemplo:

Para que una matriz cualquiera $A \in \mathbb{R}^{3\times 3}$ sea unidad o identidad debe cumplirse que $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ es decir, los elementos que figuran en la diagonal principal son todos iguales a 1.

Observemos que, para que una matriz cuadrada de orden n sea unidad o identidad, debe ser en primer lugar una matriz escalar, esto significa que toda matriz identidad o unidad es escalar pero no toda matriz escalar es unidad o identidad. Ocurre lo mismo para el caso de relacionarse la matriz identidad con las matrices triangulares y diagonal; toda matriz identidad es diagonal y triangular, pero el recíproco no se cumple.

8. Matrices idempotentes e involutivas

Matriz idempotente: *una matriz A, cuadrada de orden n, es idempotente si y solo si es igual a su cuadrado.*

$A = \|a_{ij}\|_{n \times n} \in K^{n \times n}$ *es Idempotente* $\Leftrightarrow A = A^2$

La matriz $A \in \mathbb{R}^{2\times 2}$, $A = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ es idempotente porque verifica $A = A^2$.

Matriz involutiva: *una matriz A, cuadrada de orden n, es involutiva si y solo si su cuadrado es la identidad.*

$A = \|a_{ij}\|_{n \times n} \in K^{n \times n}$ *es Involutiva* $\Leftrightarrow A^2 = I_n$

La matriz $A \in \mathbb{R}^{2\times 2}$, $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ es involutiva pues $A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$.

Ejemplo 1:

Sean la matriz <u>idempotente</u> $A \in K^{n \times n}$ y la matriz identidad I_n. Desarrollar $(A - I)^2$

$(A - I)^2 = (A - I)(A - I)$

$(A - I)^2 = A.A - A.I - I.A + I.I$

$(A - I)^2 = A^2 - A - A + I$

$(A - I)^2 = A - 2A + I$

$(A - I)^2 = I - A$

Ejemplo 2:

Sean la matriz $A \in K^{n \times n}$ y la matriz identidad I_n. Desarrollar $(A - I)^2$

$(A - I)^2 = (A - I)(A - I)$

$(A - I)^2 = A.A - A.I - I.A + I.I$

$(A-I)^2 = A^2 - A - A + I^2$

$(A-I)^2 = A^2 - 2A + I^2$

$(A-I)^2 = A^2 - 2A + I$

Observemos que este desarrollo coincide con el desarrollo que, habitualmente, conocemos con los números reales, $(a+b)^2 = a^2 + 2ab + b^2$ con $a, b \in \mathbb{R}$.

En el apartado Potencia de matrices, en el Ejemplo 3, se demostró que, en general, dadas las matrices A y B, $(A+B)^2 \neq A^2 + 2AB + B^2$ y que la igualdad solo podría establecerse si $AB = BA$.

Por lo tanto, para que el cuadrado de la suma de dos matrices coincida con el desarrollo del cuadrado de la suma de dos números reales, debe ocurrir que el producto de las matrices sea conmutable o bien que una de las dos matrices sea la matriz identidad.

Ejemplo 3:

Sea la matriz involutiva $A \in K^{n \times n}$, demostrar que $\frac{1}{2}(I-A)$ es idempotente.

Debemos demostrar que $\left[\frac{1}{2}(I-A)\right]^2 = \frac{1}{2}(I-A)$ sabiendo que A es involutiva. En efecto:

$\left[\frac{1}{2}(I-A)\right]^2 = \frac{1}{4}(I-A)^2$

$\left[\frac{1}{2}(I-A)\right]^2 = \frac{1}{4}(I^2 - 2\,I\,A + A^2)$

$\left[\frac{1}{2}(I-A)\right]^2 = \frac{1}{4}(I - 2\,A + I)$

$\left[\frac{1}{2}(I-A)\right]^2 = \frac{1}{4}(2\,I - 2\,A)$

$\left[\frac{1}{2}(I-A)\right]^2 = \frac{1}{4}2\,(I-A)$

$\left[\frac{1}{2}(I-A)\right]^2 = \frac{1}{2}(I-A)$ lo que demuestra que $\frac{1}{2}(I-A)$ es idempotente.

9. Inversión de matrices

Matriz inversa: *la matriz, cuadrada de orden n,* $A = \|a_{ij}\|_{n\times n} \in K^{n\times n}$ *es inversible, regular o no singular si y solo si existe una matriz* $B = \|b_{ij}\|_{n\times n} \in K^{n\times n}$ *tal que su producto por A a izquierda y a derecha es la matriz unidad.*

$A = \|a_{ij}\|_{n\times n} \in K^{n\times n}$ *es regular* $\Leftrightarrow \exists B = \|b_{ij}\|_{n\times n} \in K^{n\ n} / A\,B = B\,A = I$

La inversa de A, si existe, se denota por A^{-1}

Propiedades de la inversión de matrices

P_1) Si existe la inversa de una matriz cuadrada A, esta es única.

D) Suponemos que la matriz A tiene dos matrices inversas: $B\ y\ C$.

Si B es inversa de A, entonces $A\,B = B\,A = I$

Si C es inversa de A, entonces $A\,C = C\,A = I$

Sabemos que:

$B = B\,I$

$B = B\,(A\,C)$

$B = (B\,A)\,C$

$B = I\,C$

$B = C$ esto significa que, de existir la matriz inversa, esta es única; quedando, así, demostrada la propiedad.

P2) Involución: la inversa de la inversa de una matriz cuadrada, regular A, es la misma matriz A.

Sea $A = \|a_{ij}\|_{n\times n} \in K^{n\times n}\ invertible \implies (A^{-1})^{-1} = A$

D) Por definición es $A.A^{-1} = A^{-1}.A = I$

Se puede observar que A y A^{-1} son inversas entre sí, por lo tanto $(A^{-1})^{-1} = A$. Con esto queda demostrada la propiedad.

P3) Si dos matrices cuadradas $A\ y\ B$ son invertibles, entonces la inversa del producto es igual al producto de las inversas con el orden permutado.

Sean

$A = \|a_{ij}\|_{n\times n} \in K^{n\,x\,n}\ y\ B = \|b_{ij}\|_{n\times n} \in K^{n\times n}\ invertibles \implies (A\ B)^{-1} = B^{-1}A^{-1}$

D) $(A\ B)^{-1} = B^{-1}A^{-1}$ esto significa que la inversa de $A\ B$ es $B^{-1}A^{-1}$, por lo tanto, el producto a izquierda y a derecha de $A\ B$ con $B^{-1}A^{-1}$ debe ser la matriz unidad.

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

Por lo tanto, la inversa de $A\ B$ es $B^{-1}A^{-1}$ es decir $(A\ B)^{-1} = B^{-1}A^{-1}$, demostrándose así la propiedad.

Este resultado se puede generalizar: $(A_1\ A_2\ \ldots A_n)^{-1} = A_n^{-1}\ A_{n-1}^{-1}\ \ldots\ A_1^{-1}$

P4) Si una matriz es invertible, también lo es su traspuesta, y la inversa de la traspuesta de una matriz cuadrada A regular es igual a la traspuesta de la inversa de A.

Sea $A = \|a_{ij}\|_{n \times n} \in K^{n \times n}\ invertible \implies (A^t)^{-1} = (A^{-1})^t$

D) $A\,A^{-1} = I$ $\qquad$ $A^{-1}\,A = I$

$(A\,A^{-1})^t = I^t$ $\qquad$ $(A^{-1}A)^t = I^t$

$(A^{-1})^t\,A^t = I$ $\qquad$ $A^t\,(A^{-1})^t = I$

Hemos multiplicado a izquierda y a derecha respectivamente, $A^t\ y\ (A^{-1})^t$ y se obtuvo la matriz identidad, por lo tanto, la inversa de A^t es $(A^{-1})^t$, es decir que $(A^t)^{-1} = (A^{-1})^t$.

Queda así demostrada la propiedad.

P5) Sea A una matriz cuadrada de orden n, invertible, si $B\ y\ C$ son dos matrices de orden $m \times n$, entonces $i)\ B\,A = N \implies B = N\ y\ ii)\ BA = CA \implies B = C$

Sean la matriz $A = \|a_{ij}\|_{n\,x\,n} \in K^{n\,x\,n}$ invertible y las matrices

$B = \|b_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}\ y\ C = \|c_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$, entonces

$i)\ B\,A = N \implies B = N$ $\qquad$ $ii)\ BA = CA \implies B = C$

D)

i) $B\,A = N \implies B = N$ $\qquad$ *ii)* $BA = CA \implies B = C$

$B\,A = N$ $\qquad$ $BA = CA$

$B\,A\,A^{-1} = N\,A^{-1}$ $\qquad$ $B\,A\,A^{-1} = C\,A\,A^{-1}$

$B\,I = N$ $\qquad$ $B\,I = C\,I$

$B = N$ $\qquad$ $B = C$

Quedando, de esta manera, demostrada la propiedad.

P6) Sea A una matriz cuadrada de orden n, invertible, si k es un escalar no nulo, entonces la matriz (k A) también es invertible.

Sea $A = \|a_{ij}\|_{n\times n} \in K^{n\times n}\ invertible\ y\ k \in \mathbb{K}\ con\ k \neq 0 \implies (kA)^{-1} = \frac{1}{k}A^{-1}$

D) $(kA)^{-1}k\ A = I$

$(kA)^{-1}k\ A\ A^{-1} = I\ A^{-1}$

$(kA)^{-1}k\ \ I = A^{-1}$

$k\ (kA)^{-1}\ I = A^{-1}$

$k\ (kA)^{-1} = A^{-1}$

$(kA)^{-1} = \frac{1}{k}\ A^{-1}$ quedando, de esta manera, demostrada la propiedad.

P7) Sea A una matriz cuadrada de orden n, invertible, si k es un entero positivo, entonces la matriz A^k también es invertible y se tiene que $\left(A^k\right)^{-1} = (A^{-1})^k$

Sea $A = \|a_{ij}\|_{n\times n} \in K^{n\times n}\ invertible\ y\ k \in \mathbb{Z}^+$ entonces

A^k es invertible y $\left(A^k\right)^{-1} = (A^{-1})^k$

D1) Si A^k es invertible debe cumplirse que

$(A^{-1})^k A^k = I$

$(A\ A^{-1})^k = I$

$I^k = I$

$I = I$

El producto por derecha es análogo; con lo que queda demostrado que A^k es invertible.

D2) $\left(A^k\right)^{-1} A^k = I$

$\left(A^k\right)^{-1} \underbrace{A \,.\, A \,.\, A \;\ldots\; A}_{k-veces} = I$

$\left(A^k\right)^{-1} A \,.\, A \,.\, A \;\ldots\; A \,.\, A \,.\, A^{-1} = I \,.\, A^{-1}$

$\left(A^k\right)^{-1} A \,.\, A \,.\, A \;\ldots\; A \,.\, I = A^{-1}$

$\left(A^k\right)^{-1} A \,.\, A \,.\, A \;\ldots\; A = A^{-1}$

Nuevamente multiplicamos a derecha por A^{-1}, ambos miembros de la igualdad, y obtendremos:

$\left(A^k\right)^{-1} A \,.\, A \,.\, A \;\ldots\; A = (A^{-1})^2$

Si repetimos el proceso k–veces:

$\left(A^k\right)^{-1} = (A^{-1})^k$ de esta manera, queda demostrada la propiedad.

P8) Si A tiene una fila o una columna de ceros, entonces no es invertible.

D) Si la fila i de A es de ceros, al multiplicarla a la derecha por cualquier matriz, ésta tendrá la fila i de ceros. Lo mismo ocurre con las columnas, multiplicando a la izquierda. Por lo tanto, no se cumplirá la existencia de A^{-1} tal que, si se multiplica a izquierda y a derecha por A se obtenga la matriz unidad. Quedando así demostrada la propiedad.

Hasta el momento solo se definió y se demostraron las propiedades de la matriz inversa, tanto la condición de existencia como el cálculo de esta matriz se estudiará en el apartado "Condición de existencia y cálculo de la matriz inversa".

10. Matrices ortogonales

Matriz ortogonal 1: *una matriz cuadrada de orden n, invertible se dice ortogonal si y solo si su inversa es igual a su traspuesta.*

Sea $A = \|a_{ij}\|_{n\times n} \in K^{n\times n}, A\ es\ ortogonal \iff A^{-1} = A^t$

Matriz ortogonal 2: *una matriz cuadrada de orden n, invertible se dice ortogonal si y solo si el producto de dicha matriz por su traspuesta es la unidad.*

Sea $A = \|a_{ij}\|_{n\times n} \in K^{n\times n}, A\ es\ ortogonal \iff A.A^t = A^t.A = I_n$

Ejemplo:

La matriz $A = \begin{pmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix}$ es ortogonal, pues verifica $A.A^t = I_n$

$$\begin{pmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix}.\begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Propiedad

El producto de dos matrices ortogonales, es ortogonal.

Sean $A = \|a_{ij}\|_{n\times n} \in K^{n\times n},\ B = \|b_{ij}\|_{m\times n} \in K^{m\times n}\ ortogonales \Longrightarrow$

$(A.B)(A.B)^t = (A.B)^t(A.B) = I_n$

D) $(A.B)(A.B)^t = A.B.B^t.A^t$

$(A.B)(A.B)^t = A.I.A^t.$

$(A.B)(A.B)^t = A.A^t$

$(A.B)(A.B)^t = I$

De manera similar se demuestra $(A.B)^t(A.B) = I_n$, quedando, de esta manera, demostrada la propiedad.

11. Transformaciones elementales: matrices equivalentes, escalonadas y reducidas

Para abordar, con más facilidad, el análisis de los sistemas de ecuaciones lineales y el estudio de las propiedades de los determinantes, las matrices proveen de una importante herramienta, las llamadas transformaciones elementales de matrices. Estas se definen, tanto para las filas como para las columnas, de la siguiente manera:

Transformaciones elementales u operaciones elementales: *entre las filas (columnas) de una matriz* $A = \|a_{ij}\|_{m \times n} \in K^{m \times n}$ *son las siguientes:*

a) Permutación de filas (columnas) entre sí.

b) Adición de una fila (columna) a otra.

c) Multiplicación de una fila (columna) por un escalar no nulo.

A una matriz se la puede considerar como un conjunto de vectores filas (columnas), por lo tanto, las operaciones elementales se expresan como:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ \dots \\ F_m \end{pmatrix} \Rightarrow \begin{cases} F_i \leftrightarrow F_j \\ F_i \rightarrow F_i + F_j \\ F_i \rightarrow \alpha F_i \; con \; \alpha \in K \;\; \wedge \alpha \neq 0 \end{cases}$$

A partir de estas transformaciones de una matriz, podemos dar la definición de:

Matrices equivalentes: *dadas dos matrices A y B decimos que B es equivalente a A si y solo si B se puede obtener efectuando un número finito de operaciones elementales sobre A. Para indicar que A es equivalente a B escribimos* $A \sim B$.

Ejemplo:

Sea la matriz $A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 5 & 0 \\ 1 & -1 & 1 \\ 7 & 9 & 1 \end{pmatrix}$ a partir de ella se pueden obtener, operando con filas, las siguientes matrices equivalentes:

$$\begin{pmatrix} 0 & 1 & 2 \\ 3 & 5 & 0 \\ 1 & -1 & 1 \\ 7 & 9 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 3 & 5 & 0 \\ 0 & 1 & 2 \\ 7 & 9 & 1 \end{pmatrix} \sim \begin{matrix} -3F_1 + F_2 \\ \\ \\ -7F_1 + F_4 \end{matrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 8 & -3 \\ 0 & 1 & 2 \\ 0 & 16 & -6 \end{pmatrix} \sim \ \ldots$$

Se puede observar que la primera matriz equivalente se obtuvo intercambiando la fila 1 *(F_1)* por la fila 3 *(F_3)*, en este caso no se utilizó ninguna notación debido a la simplicidad de la operación. Para obtener la segunda matriz equivalente se procedió así: a la fila 2 *(F_2)* se le sumó la fila 1 *(F_1)* multiplicada por el escalar –3, indicándose esto mediante la notación $-3F_1 + F_2$, luego a la fila 4 *(F_4)* se le sumó la fila 1 *(F_1)* multiplicada por el escalar –7, indicándose esto mediante la notación $-7F_1 + F_4$, y así sucesivamente se pueden obtener infinitas matrices equivalentes a una dada.

Matrices equivalentes por fila: *dadas dos matrices A y B decimos que B es equivalente por fila a A, si B se puede obtener efectuando un número finito de operaciones elementales sobre las filas de A. Para indicar que A es equivalente por fila a B, escribimos* $A \sim_f B$.

Si $A \sim_f B \implies A \sim B$ el recíproco no vale, o sea que $A \sim B \nRightarrow A \sim_f B$.

Matrices equivalentes por columna: *dadas dos matrices A y B decimos que B es equivalente por columna a A, si B se puede obtener efectuando un número*

finito de operaciones elementales sobre las columnas de A. Para indicar que A es equivalente por columna a B escribimos $A \sim_c B$.

Si $A \sim_c B \implies A \sim B$, el recíproco no vale, o sea que $A \sim B \not\Rightarrow A \sim_c B$

Ejemplo:

Sea la matriz $A = \begin{pmatrix} 1 & -1 & 1 \\ 3 & 5 & 0 \\ 0 & 1 & 2 \\ 7 & 9 & 1 \end{pmatrix}$ a partir de ella se pueden obtener, operando con columnas, las siguientes matrices equivalentes

$$\begin{pmatrix} 1 & -1 & 1 \\ 3 & 5 & 0 \\ 0 & 1 & 2 \\ 7 & 9 & 1 \end{pmatrix} \sim \begin{matrix} 2C_1 + C_2 \\ -2C_1 + C_3 \end{matrix} \begin{pmatrix} 1 & 1 & -1 \\ 3 & 11 & -6 \\ 0 & 1 & 2 \\ 7 & 23 & -13 \end{pmatrix}$$

Sea A una matriz y Fi una fila cualquiera de A, si Fi no es una fila totalmente de ceros (nula), se dice que el primer número distinto de cero, de izquierda a derecha, es el *pivot* (p_i) de la fila y a esta se la suele llamar fila pivotante.

Matriz escalonada por filas: *es aquella que verifica las siguientes propiedades:*

a) Todas las filas nulas (si existiesen) se encuentran en la parte inferior de la matriz.

b) El pivot de cada fila no nula se encuentra estrictamente más a la derecha que el pivot de la fila de encima.

Ejemplos:

Sean las matrices:

$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 7 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 5 & 8 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 7 & 9 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 5 & 7 \\ 0 & 1 & 2 \end{pmatrix}$, $D = \begin{pmatrix} 2 & 0 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$

Las matrices A y B no son escalonadas, mientras que C y D sí lo son.

Propiedades de las matrices escalonadas por filas

P_1) Todas las formas escalonadas por filas de una matriz tienen el mismo número de filas no nulas. A dicho número se lo suele denominar ρ.

P_2) Todas las formas escalonadas por filas de una matriz tienen las columnas pivotales en las mismas posiciones y el número de ellas coincide con el número de filas no nulas, ρ.

Matriz escalonada reducida por filas: *es aquella que verifica las siguientes propiedades:*

a) En cada fila el pivot es igual a 1.

b) Todos los elementos, por encima de los pivotes, son nulos.

Ejemplos:

Sean las matrices $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 2 \end{pmatrix}$

Las tres matrices son escalonadas reducidas por filas.

Debemos destacar que, para que una matriz sea escalonada reducida debe ser, en primer lugar, escalonada, es decir que, toda matriz escalonada

reducida es escalonada pero no toda matriz escalonada es reducida. Dicho de otra forma, para que una matriz se pueda clasificar como escalonada debe cumplir dos propiedades mientras que, para ser escalonada reducida debe cumplir cuatro propiedades, las dos de escalonada más otras dos.

A la forma escalonada reducida por filas también se le suele llamar "forma normal de Hermite". Se puede demostrar que "*Dada una matriz A, la forma escalonada reducida por filas es única*". A pesar de que la forma escalonada reducida por filas es única, se puede llegar a ella por distintas secuencias de operaciones elementales por filas.

El proceso de obtener una forma escalonada por filas de una matriz se denomina "eliminación gaussiana".

11.1. Obtención de la forma escalonada reducida por filas: algoritmo de eliminación gaussiana

Describiremos, a continuación, los pasos de un algoritmo para obtener, a partir de una matriz $A = \|a_{ij}\|_{m \times n} \in K^{m \times n}$, una determinada forma escalonada reducida por fila. Este algoritmo se puede resumir de la siguiente manera:

1.- Determine, de izquierda a derecha, la primera columna no cero (pivote).

2.- Si el primer elemento de la columna es cero, intercambie la fila por otra que no tenga cero, esto se ha de realizar con la primera fila siguiente que sí posea pivote.

3.- Obtenga ceros debajo del pivote sumando múltiplos adecuados a las filas debajo de él.

4.- Sin considerar la fila ni la columna de trabajo, repita el proceso comenzado en el paso 1. Al término del ciclo desde el paso 1 al 4, es decir, cuando se hayan utilizado todas las filas, la matriz deberá tener forma escalonada.

5.- Comenzando con la última fila no nula, avance hacia arriba para que en cada fila tenga un 1 como pivote y arriba de él solo ceros. Para esto deberá sumar múltiplos adecuados de la fila pivotante a las filas correspondientes.

Es importante observar que, en este algoritmo, los pasos del 1 al 4 aplicados repetidamente permiten obtener una matriz escalonada, mientras que el paso 5 permite obtener una matriz escalonada reducida. En el paso 2, si el elemento no es cero no se debe realizar intercambio alguno y, en el paso 3, los elementos que se hacen cero son solo los elementos debajo del pivote.

Ejemplo:

Dada la matriz $A = \begin{pmatrix} 3 & 6 & -9 & 3 \\ 2 & 4 & -8 & 0 \\ -2 & -3 & 4 & -1 \end{pmatrix}$ aplicar el algoritmo de eliminación gaussiana.

En este caso, la primera columna tiene elementos distintos de cero ($a_{11} = 3$), continúa entonces el paso 2, pero como el elemento $a_{11} \neq 0$, se continúa con el paso 3. El elemento a_{11} es el pivote y será usado para hacer

ceros los elementos debajo de él; para esto debemos sumar múltiplos adecuados de la fila pivotante a las filas inferiores.

$$\begin{pmatrix} 3 & 6 & -9 & 3 \\ 2 & 4 & -8 & 0 \\ -2 & -3 & 4 & -1 \end{pmatrix} \sim \begin{matrix} -\frac{2}{3}F_1 + F_2 \\ \frac{2}{3}F_1 + F_3 \end{matrix} \begin{pmatrix} 3 & 6 & -9 & 3 \\ 0 & 0 & -2 & -2 \\ 0 & 1 & -2 & 1 \end{pmatrix}$$

No se debe considerar, a partir de ahora, ni la primera fila ni la primera columna, pues con ellas ya se trabajó. Al repetir el paso 1, el algoritmo busca la primera columna distinta de cero; en este caso, a partir de la segunda columna, y continúa con el paso 2. Como el elemento $a_{22} = 0$ debemos buscar en la parte inferior de la columna dos el primer elemento distinto de cero ($a_{32} = 1$) y realizar el intercambio de filas.

$$\begin{pmatrix} 3 & 6 & -9 & 3 \\ 0 & 0 & -2 & -2 \\ 0 & 1 & -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 3 & 6 & -9 & 3 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix}$$

Ahora, el pivote es el elemento $a_{22} = 1$.

Continuamos con el paso 3. En este caso los elementos debajo del pivote son cero y el algoritmo procede a la siguiente columna siendo el pivote el elemento $a_{33} = -2$ que no tiene elementos debajo de él.

Hasta aquí la matriz es escalonada, ahora se procede con el paso 5; hacer uno cada pivote y, posteriormente, hacer ceros los elementos arriba de cada uno de ellos.

Hagamos 1 el p_3 ($a_{33} = -2$).

$$\begin{pmatrix} 3 & 6 & -9 & 3 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix} \sim \begin{matrix} \\ \\ -\frac{1}{2}F_3 \end{matrix} \begin{pmatrix} 3 & 6 & -9 & 3 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Debemos hacer cero los elementos por arriba de p_3.

$$\begin{pmatrix} 3 & 6 & -9 & 3 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{matrix} 9F_3 + F_1 \\ \sim 2F_3 + F_2 \end{matrix} \begin{pmatrix} 3 & 6 & 0 & 12 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Procedemos de igual manera con p_2 ($a_{22} = 1$) que ya es 1, entonces solo resta hacer cero los elementos por encima de él.

$$\begin{pmatrix} 3 & 6 & 0 & 12 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{matrix} -6F_2 + F_1 \\ \\ \end{matrix} \begin{pmatrix} 3 & 0 & 0 & -6 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Finalmente, debemos hacer 1 el p_1 ($a_{11} = 3$).

$$\begin{pmatrix} 3 & 0 & 0 & -6 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{matrix} \frac{1}{3}F_1 \\ \\ \end{matrix} \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Rango de una matriz: *sea la matriz* $A = \|a_{ij}\|_{m \times n} \in K^{m \times n}$ *se denomina rango de A y se denota como* $\rho(A)$*, al número de columnas pivotales o número de filas no nulas de cualquier forma escalonada por filas de dicha matriz.*

Observemos que esta definición de rango se basa en cualquier forma escalonada por fila de la matriz, es decir que no es necesario que sea escalonada reducida por fila. Por lo tanto, para obtener el rango de una matriz se puede aplicar el algoritmo de eliminación gaussiana del paso 1 al 4, únicamente.

Ejemplo:

Calcular el rango de $A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 5 & 0 \\ 1 & -1 & 1 \\ 7 & 9 & 1 \end{pmatrix}$

Como $a_{11} = 0$ se intercambian la fila 1 con la fila 3.

$$\begin{pmatrix} 0 & 1 & 2 \\ 3 & 5 & 0 \\ 1 & -1 & 1 \\ 7 & 9 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 3 & 5 & 0 \\ 0 & 1 & 2 \\ 7 & 9 & 1 \end{pmatrix}$$

Ahora $a_{11} = 1 = p_1$. Debemos hacer cero todos los elementos de la primera columna y ubicados debajo de p_1.

$$\begin{pmatrix} 1 & -1 & 1 \\ 3 & 5 & 0 \\ 0 & 1 & 2 \\ 7 & 9 & 1 \end{pmatrix} \sim \begin{matrix} -3\mathrm{F}_1 + \mathrm{F}_2 \\ \\ \\ -7\mathrm{F}_1 + \mathrm{F}_4 \end{matrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 8 & -3 \\ 0 & 1 & 2 \\ 0 & 16 & -6 \end{pmatrix}$$

Corresponde que $\mathrm{a}_{22} = 8$ sea el segundo pivote, p_2. Debemos hacer cero todos los elementos debajo de p_2.

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 8 & -3 \\ 0 & 1 & 2 \\ 0 & 16 & -6 \end{pmatrix} \sim \begin{matrix} \\ -8\mathrm{F}_3 + \mathrm{F}_2 \\ \\ -2\mathrm{F}_2 + \mathrm{F}_4 \end{matrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 8 & -3 \\ 0 & 0 & -19 \\ 0 & 0 & 0 \end{pmatrix}$$

Finalmente $p_3 = -19$ y ya son cero los elementos debajo de él.

Por lo tanto $\rho(\mathrm{A}) = 3$ ya que esta forma escalonada de A tiene 3 filas pivotales o tres filas no nulas.

Propiedades del rango de una matriz

P1) Dada una matriz $A = \|a_{ij}\|_{m \times n} \in K^{m \times n}$ el rango es menor o igual que m y menor o igual que n.

P2) Todas las matrices equivalentes entre sí (por filas) tienen el mismo rango.

Nótese que el recíproco de esta propiedad no se cumple.

P3) Dada una matriz $A = \|a_{ij}\|_{m \times n} \in K^{m \times n}$, el rango de A es igual al rango de su traspuesta.

P4) El rango de la matriz I_n es n.

D) La demostración de esta propiedad es evidente ya que la matriz identidad ya es escalonada reducida por filas y, por lo tanto, es escalonada por filas y tiene n filas no nulas o n columnas pivotales; en consecuencia, el rango de la matriz I_n es n.

P5) Toda matriz equivalente por filas a la identidad tiene rango n.

D) La demostración de esta propiedad se pone en evidencia si consideramos la segunda propiedad enunciada: Todas las matrices equivalentes entre sí (por filas) tienen el mismo rango, por lo tanto, si una matriz es equivalente por fila a la matriz identidad, I_n, y esta tiene rango n, la matriz equivalente también tendrá rango n.

Ejemplos:

Sean las matrices:

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 5 \end{pmatrix},\ B = \begin{pmatrix} 1 & 2 & 4 & -1 \\ 0 & 2 & -4 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ C = \begin{pmatrix} 1 & -2 & 4 & 1 \\ 0 & 2 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix},\ I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\rho(A) = 2,\ \rho(B) = 2,\ \rho(C) = 3 \quad y \quad \rho(I_4) = 4$$

11.2. Obtención de la forma escalonada reducida por filas: algoritmo de Gauss–Jordan

Describiremos a continuación otro algoritmo para obtener, a partir de una matriz dada $A = \|a_{ij}\|_{m\,x\,n} \in K^{m\,x\,n}$, una determinada forma escalonada reducida por fila. Este algoritmo consta de los siguientes pasos:

1.- Determine, de izquierda a derecha, la primera columna no cero (pivote).

2.- Si el primer elemento de la columna es cero, intercambie la fila por otra que no tenga cero, esto se ha de realizar con la primera fila siguiente que sí posea pivote. Multiplicando apropiadamente, haga 1 el pivote, p_1.

3.- Obtenga ceros arriba y abajo del pivote sumando múltiplos adecuados a las filas debajo (arriba) de él.

4.- Sin considerar la fila ni la columna de trabajo, repita el proceso comenzado en el paso 1 en la columna siguiente.

Es importante observar que la idea de este algoritmo es ir escalonando la matriz y volviéndola reducida por filas, simultáneamente.

El paso 1 de este algoritmo es idéntico al de la eliminación gaussiana, en el paso 2 se debe hacer el intercambio de filas solo si el primer elemento de la columna es cero, igual que en el algoritmo de eliminación gaussiana, pero en este caso al pivote se lo transforma en 1. La diferencia en el paso 3 de los algoritmos es que en la eliminación gaussiana se hacen ceros solamente los elementos por debajo del pivot, mientras que en Gauss–Jordan se vuelven cero los elementos por debajo y encima del pivot. El paso 4 es similar, obteniéndose en la eliminación gaussiana una matriz escalonada.

Ejemplo:

Dada la siguiente matriz, aplicar el algoritmo de Gauss–Jordan para obtener una matriz escalonada reducida por filas.

$$A = \begin{pmatrix} 3 & 6 & -9 & 3 \\ 2 & 4 & -8 & 0 \\ -2 & -3 & 4 & -1 \end{pmatrix}$$

Observamos que el elemento $a_{11} = 3$ es distinto de cero, por lo tanto es el p_1. Corresponde volver 1 a dicho pivot y volver cero los elementos, en este caso, debajo de él.

$$\begin{pmatrix} 3 & 6 & -9 & 3 \\ 2 & 4 & -8 & 0 \\ -2 & -3 & 4 & -1 \end{pmatrix} \sim^{\frac{1}{3}F_1} \begin{pmatrix} 1 & 2 & -3 & 1 \\ 2 & 4 & -8 & 0 \\ -2 & -3 & 4 & -1 \end{pmatrix} \sim \begin{matrix} -2F_1 + F_2 \\ 2F_1 + F_3 \end{matrix} \begin{pmatrix} 1 & 2 & -3 & 1 \\ 0 & 0 & -2 & -2 \\ 0 & 1 & -2 & 1 \end{pmatrix}$$

Sin considerar ni la primera fila ni la primera columna, procedemos a identificar el segundo pivot, p_2, observamos que $a_{22} = 0$, por lo tanto procedemos a intercambiar la fila 2 por la inmediata inferior que no tenga cero en la posición deseada, en este caso la fila 3.

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

Como p_2 ya es 1, procedemos a hacer cero los elementos por debajo y por encima de p_2, en este caso el elemento por debajo de p_2 ya es cero, solo corresponde trabajar con el elemento a_{12}.

$$\begin{pmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix} \sim \begin{matrix} -2F_2 + F_1 \\ \\ \\ \end{matrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix}$$

Sin considerar ni la segunda fila ni la segunda columna, procedemos a identificar el tercer pivot, p_3, observamos que $a_{33} = -2$, corresponde volver 1 a p_3 y cero los elementos, en este caso, por encima de él.

$$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix} \sim \begin{matrix} \\ \\ -\frac{1}{2}F_3 \end{matrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{matrix} -F_3 + F_1 \\ 2F_3 + F_2 \\ \\ \end{matrix} \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

De esta forma finaliza el algoritmo y se obtuvo a partir de la matriz dada otra matriz, pero escalonada reducida por filas. Además, podemos decir en este caso que $\rho(A) = 3$.

12. Condición de existencia y cálculo de la matriz inversa

Tal como se dijo anteriormente, en este apartado se demostrará una condición necesaria y suficiente para que una matriz cuadrada posea inversa y se ejemplificará su cálculo. Previo a esto será necesario demostrar el siguiente teorema.

Teorema: *Dada una matriz cualquiera* $A = \|a_{ij}\|_{m \times n} \in \mathbb{K}^{m \times n}$ *existen dos matrices P y U tales que P.A = U, siendo U una matriz escalonada.*

D) La demostración de este teorema se realizará de forma constructiva.

Paso 1: Anular todos los elementos a_{i1} con $1 < i \leq n$.

Si $a_{11} \neq 0$ mediante transformaciones elementales por filas podemos anular todos los elementos de la primera columna situados por debajo de él.

Si $a_{11} = 0$ y algún elemento de la primera columna es no nulo, entonces podemos llevarlo –mediante una transformación elemental– a la posición $(1, 1)$ y proceder como en el caso anterior.

Si $a_{i1} = 0$ para $i = 1, \dots, m$ significa que la primera columna es de ceros, por lo tanto $a_{i1} = 0 \; \forall \; i > 1$ lo que significa que se trata de una columna del tipo de las matrices escalonadas.

Paso 2: Procedemos con el elemento a_{22} (resultante de las transformaciones anteriores) de la misma manera que se procedió con

el a_{11}, es decir, si $a_{22} \neq 0$ lo utilizamos para hacer cero los elementos debajo de él en la segunda columna. Si $a_{22} = 0$ buscamos si existe por debajo de él, el primer elemento $a_{i2} \neq 0$ y, en caso de existir, realizamos la transformación correspondiente, etc.

Paso 3: Reiterando los pasos 1 y 2, llegamos a una matriz escalonada.

De esta manera, queda demostrado –de manera constructiva– el teorema.

Observemos que la matriz P es el producto de las matrices de las transformaciones elementales por filas[24] realizadas para pasar de A a U.

Ejemplo:

Sea la matriz $A = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix}$ encontrar las matrices P y U (escalonada) tal que $P.A = U$

$$\begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix} \sim \underbrace{-2F_1 + F_2 \begin{pmatrix} 2 & 1 & 3 & 4 \\ 0 & 0 & -5 & -3 \\ 1 & 0 & 2 & 3 \end{pmatrix}}_{P_1} \sim \underbrace{-\frac{1}{2}F_1 + F_3 \begin{pmatrix} 2 & 1 & 3 & 4 \\ 0 & 0 & -5 & -3 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}}_{P_2} \sim$$

$$\sim \underbrace{F_2 \leftrightarrow F_3 \begin{pmatrix} 2 & 1 & 3 & 4 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -5 & -3 \end{pmatrix}}_{P_3} = U.$$

Como $P_3.P_2.P_1.A = U$ entonces $P.A = U$ siendo

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}}_{P_3} . \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}}_{P_2} . \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P_1} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}}_{P}$$

[24] Para ampliar este tema, ver el apéndice al final de este capítulo.

Se verifica que $P.A = U$

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix} . \begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -5 & -3 \end{pmatrix}$$

Finalmente $P = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$ y $U = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -5 & -3 \end{pmatrix}$

Este teorema nos permite demostrar una condición necesaria y suficiente para que una matriz cuadrada posea inversa.

Teorema: *una condición necesaria y suficiente para que una matriz cuadrada de orden n posea inversa es que su forma escalonada reducida por filas sea la matriz unidad.*

D) $\Longrightarrow$) Si su forma escalonada reducida por fila es I_n existe $P \in \mathbb{K}^{n \times n}$ tal que $P.A = I_n$ por lo tanto: $P.A = I_n \iff P.A.A^{-1} = I_n.A^{-1} \iff P = A^{-1}$.

$\Longleftarrow$) Si existe A^{-1} tal que $A^{-1}.A = I_n \implies \exists P = A^{-1}$ tal que $P.A = I_n$ y, por lo tanto, $I_n$es la forma escalonada reducida por fila de A.

De esta manera queda demostrado el teorema.

Se puede establecer una relación entre el rango de una matriz cuadrada de orden n y su matriz inversa, mediante el siguiente corolario.

Corolario: *la matriz $A \in \mathbb{K}^{n \times n}$ es invertible si y solo si $\rho(A) = n$.*

D) $\Longrightarrow$) Si A es invertible, por el teorema anterior podemos afirmar que su forma escalonada reducida es I_n, que tienen n filas no nulas y por lo tanto $\rho(A) = n$.

$\Leftarrow$) Si $\rho(A) < n$ entonces U, la forma escalonada reducida de A, tiene una fila de ceros, luego no es invertible. Pero sabemos que $U = P.A$, por lo que, si A fuera invertible, U también lo sería.

Quedando de esta manera demostrado el corolario y estableciendo una segunda condición necesaria y suficiente para que una matriz posea inversa.

El teorema nos permite calcular la matriz inversa de una matriz A cuadrada de orden n, mediante transformaciones elementales entre sus filas, puesto que "*efectuando la misma secuencia de operaciones elementales por filas y en el mismo orden que llevan de A a I_n, la matriz I_n se transforma en A^{-1}*".

Esto determina un procedimiento para obtener la inversa de una matriz denominado método de Gauss–Jordan, cuyo esquema es el siguiente:

$$(A \mid I) \quad \underbrace{\rightarrow \rightarrow \rightarrow \rightarrow \ \ldots \ \rightarrow \rightarrow \rightarrow \rightarrow}_{operaciones\ elementales\ por\ filas} \quad (I \mid A^{-1})$$

Ejemplo 1:

Determinar por el método de Gauss–Jordan la matriz inversa de A, si dicha matriz existe.

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

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \sim \begin{matrix} \\ -F_{1+F_3} \end{matrix} \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{array}\right) \sim$$

$$\begin{matrix} -3F_2 + F_1 \\ \sim \\ F_2 + F_3 \end{matrix} \left(\begin{array}{ccc|ccc} 1 & 0 & -3 & 1 & -3 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right) \sim \begin{matrix} 3F_3 + F_1 \\ -F_3 + F_2 \\ \\ \end{matrix} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 3 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right)$$

Para verificar este resultado, al multiplicar la matriz A por su inversa se debe obtener la matriz identidad.

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

Por lo tanto, existe la inversa de A y es: $A^{-1}=\begin{pmatrix}-2 & 0 & 3\\ 1 & 0 & -1\\ -1 & 1 & 1\end{pmatrix}$

Ejemplo 2:

Determinar por el método de Gauss–Jordan la matriz inversa de A, si dicha matriz existe.

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

$$\left(\begin{array}{ccc|ccc}1 & 1 & 3 & 1 & 0 & 0\\ -1 & -1 & 2 & 0 & 1 & 0\\ 1 & 1 & 0 & 0 & 0 & 1\end{array}\right)\sim\begin{matrix}\\ F_1+F_2\\ -F_1+F_3\end{matrix}\left(\begin{array}{ccc|ccc}1 & 1 & 3 & 1 & 0 & 0\\ 0 & 0 & 5 & 1 & 1 & 0\\ 0 & 0 & -3 & -1 & 0 & 1\end{array}\right)$$

A partir de aquí podemos observar que no debemos considerar ni la primera fila ni la primera columna, además el elemento $a_{22}=0$ y el único elemento inmediatamente debajo de él también es cero, $a_{32}=0$. Por lo tanto, no podemos seguir transformando la matriz A en la matriz identidad, esto significa que la matriz A no posee inversa.

Otra condición necesaria y suficiente para determinar, más rápidamente, si una matriz cuadrada de orden n admite o no matriz inversa, es calculando su determinante. Este y otros temas relacionados se estudiarán en el capítulo correspondiente al tema Determinantes.

13. Submatrices

Submatriz: *dada la matriz* $A = \|a_{ij}\|_{m \times n}$ *si se eligen I filas de las* m *posibles y J columnas de las* n *posibles, se define la submatriz de A de orden* $I \times J$ *como la matriz* $[A]_{I \times J}$ *cuyos elementos* a_{ij} *de A tales que* $i \in I \ \ y \ \ j \in J$.

Una submatriz de la matriz A es la matriz que se obtiene de A al eliminar las filas que no están en I y las columnas que no están en J; es decir, es la matriz resultante de eliminar de A determinadas filas y/o columnas.

Sea $A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 1 \\ 4 & 2 & -1 \\ 2 & 1 & 0 \end{pmatrix}$ entonces, algunas de las submatrices son:

$[A]_{2\,x\,2} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$ $\qquad$ $[A]_{3\,x\,2} = \begin{pmatrix} 0 & 1 \\ 4 & 2 \\ 2 & 1 \end{pmatrix}$ $\qquad$ $[A]_{1\,x\,3} =$

$(4 \quad 2 \quad -1)$

14. Ecuaciones matriciales

Para poner en práctica el álgebra con matrices estudiadas en los apartados anteriores, se propone a continuación la resolución –en forma genérica– de ecuaciones donde se involucran incógnitas que representan matrices. La resolución genérica se realizará por medio de ejemplos.

Ejemplo 1:

Siendo A, B, X matrices y k un escalar no nulo, determinar X en: $k.X + A = B$

Los procedimientos que se emplean para encontrar la solución en esta ecuación matricial son similares a los seguidos en las ecuaciones algebraicas.

$k.X + A = B$

$k.X + A - A = B - A$ sumando en ambos miembros la opuesta de A

$k.X + (A - A) = B - A$ por asociatividad de la suma (resta) de matrices

$k.X = B - A$ por suma (resta) de matrices

$\frac{1}{k}(k.X) = \frac{1}{k}(B - A)$ multiplicando ambos miembros por el escalar $\frac{1}{k}$

$X = \frac{1}{k}(B - A)$ por la asociatividad del producto de una matriz por un escalar.

Este procedimiento se realizó con detalles para evidenciar la similitud enunciada en el primer párrafo. Evidentemente, este procedimiento se puede realizar –sobre todo en la práctica– de manera más sintética si se eliminan los pasos obvios:

$k.X + A = B$

$k.X = B - A$

$X = \frac{1}{k}(B - A)$

Ejemplo 2:

Siendo A y X matrices y A invertible, determinar X en: $A.X = B$

Como la matriz A está multiplicando por izquierda a la matriz X, de igual manera se multiplicará ambos miembros por la inversa de A.

$A^{-1}.A.X = A^{-1}.B$

$(A^{-1}.A).X = A^{-1}.B$ por la asociatividad del producto de matrices

$I.X = A^{-1}.B$ por definición de inversa de una matriz

$X = A^{-1}.B$ por ser la matriz identidad el elemento neutro del producto de matrices

$X = A^{-1}.B$

Observemos, en primer lugar que, como en el caso anterior, aquí también se puede hacer un procedimiento más sintético que el expuesto. En segundo lugar que, en el primer paso del procedimiento no se puede multiplicar por A^{-1} ambos miembros de la ecuación por derecha. Es decir, hay que distinguir y recordar las dos siguientes situaciones que no son iguales.

$A.X = B \Longrightarrow X = A^{-1}.B$ o bien $X.A = B \Longrightarrow X = B.A^{-1}$

Finalmente, esto último también nos indica que, para determinar la matriz X, no se deben dividir ambos miembros de la ecuación, por una misma matriz. No está definida la división de matrices.

Ejemplo 3:

Sabiendo que A, B, X son matrices y que además A y B son invertibles, determinar X en: $A.B.X = C$

Solución 1

$A.B.X = C$

$(A.B).X = C$

$A^{-1}.C$

$(A.B)^{-1}.(A.B).X = (A.B)^{-1}.C$

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

$B^{-1}.A^{-1}.C$

$X = (A.B)^{-1}.C$

$B^{-1}.A^{-1}.C$

Solución 2

$A.B.X = C$

$A^{-1}.A.B.X =$

$B.X = A^{-1}.C$

$B^{-1}.B.X =$

$X =$

Debido a que $(A.B)^{-1} = B^{-1}.A^{-1}$, ambos resultados son igualmente correctos.

Ejemplo 4:

Sabiendo que A, X son matrices, determinar X en los siguientes dos casos:

$X^t = A$ y $X^{-1} = A$

Ambos casos se basan en propiedades similares de la trasposición e inversión de matrices. Se los resolverá sintéticamente y de manera simultánea.

$X^t = A$ $\qquad\qquad$ $X^{-1} = A$

$(X^t)^t = A^t$ $\qquad\qquad$ $(X^{-1})^{-1} = A^{-1}$

$X = A^t$ $\qquad\qquad$ $X = A^{-1}$

Los siguientes ejemplos son una combinación de los ejemplos anteriores, es importante ser cuidadoso y riguroso en cuanto a los pasos necesarios para determinar la incógnita.

Ejemplo 5:

Sabiendo que A, B, C, D, X son matrices, X invertible y k un escalar no nulo, determinar X en: $A.(kX + B) + C = D$

$A.(kX + B) + C = D$

$A.(kX + B) = D - C$

$kX + B = A^{-1}.(D - C)$

$kX = A^{-1}.(D - C) - B$

$X = \frac{1}{k}(A^{-1}.(D - C) - B)$

Ejemplo 6:

Sabiendo que A, B, C, D, E, X son matrices invertibles, determinar X en:

$A.((B.X)^{-1}+C)^t+D=E$

$A.((B.X)^{-1}+C)^t=E-D$

$((B.X)^{-1}+C)^t=A^{-1}.(E-D)$

$(B.X)^{-1}+C=(A^{-1}.(E-D))^t$

$(B.X)^{-1}=(A^{-1}.(E-D))^t-C$

$B.X=((A^{-1}.(E-D))^t-C)^{-1}$

$X=B^{-1}.((A^{-1}.(E-D))^t-C)^{-1}$

Apéndice

Este apartado pretende profundizar el tema: Transformaciones elementales de matrices, también llamadas operaciones elementales entre filas o columnas de una matriz.

Recordemos, en primer lugar, que se definieron tres tipos de transformaciones: permutación de líneas, adición de una línea a otra y multiplicación de una línea por un escalar no nulo; y en segundo, es válido realizar simultáneamente más de una transformación.

La primera transformación elemental sobre las filas de una matriz $A \in \mathbb{K}^{m \times n}$ es el intercambio de la fila i y j de ella. Este efecto se logra al multiplicar por izquierda la matriz A por la matriz P_{ij}, siendo esta el resultado de intercambiar las filas i y j de la matriz I_m.

Por ejemplo, consideremos la matriz $A = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix}$, para intercambiar $F_2\ y\ F_3$ multiplicamos por izquierda la matriz A por la matriz $P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ (en $I_3\ (m = 3)$ se intercambiaron las filas segunda y tercera).

$$P_{23}.A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.\begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 1 & 0 & 2 & 3 \\ 4 & 2 & 1 & 5 \end{pmatrix}$$

Observemos que se obtuvo una matriz con las mismas filas que A, pero con la segunda y tercera fila intercambiada.

En la práctica, esto no suele aplicarse ya que es simplemente volver a "escribir" la matriz A con sus filas intercambiadas, pero en el ámbito computacional el intercambio de filas de una matriz se realiza con este procedimiento.

La segunda transformación elemental sobre las filas de una matriz $A \in \mathbb{K}^{m\,x\,n}$ es multiplicar la fila i por un número $\alpha \neq 0$. Este efecto se logra al multiplicar por izquierda la matriz A por la matriz $P_i(\alpha)$, siendo esta el resultado de multiplicar por α la fila i de la matriz I_m.

Por ejemplo, consideremos la matriz $A = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix}$ para multiplicar por 3 ($\alpha = 3$) la segunda fila de A, multiplicamos por izquierda A y la matriz $P_2(3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (en I_3 ($m = 3$) se multiplicó la segunda fila por 3).

$$P_2(3).A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.\begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 12 & 6 & 3 & 15 \\ 1 & 0 & 2 & 3 \end{pmatrix}$$

Observemos que se obtuvo una matriz cuyas primera y tercera fila son idénticas a las de la matriz A, y cuya segunda fila es igual a la segunda fila de A multiplicada por 3.

La tercera transformación elemental sobre las filas de una matriz $A \in \mathbb{K}^{m \times n}$ es sumar a la fila i la fila j multiplicada por un número $\alpha \neq 0$. Este efecto se logra al multiplicar por izquierda la matriz A por la matriz $P_{ij}(\alpha)$, siendo esta el resultado de sumar a la fila i de la matriz I_m su fila j

multiplicada por α, es decir, la matriz resultante de sustituir el elemento $i_{ij} = 0$ por α.

Por ejemplo, consideremos la matriz $A = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix}$, para restar a la segunda fila de A el doble de la primera utilizamos la matriz $P_{21}(-2) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (en I_3 $(m = 3)$ se sustituyó el elemento $i_{21} = 0$ por –2).

$$P_{21}(-2).A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.\begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 1 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 0 & 0 & -5 & -3 \\ 1 & 0 & 2 & 3 \end{pmatrix}$$

Observemos que en la matriz A se ha producido el efecto que deseamos.

Si bien se detallaron las transformaciones para las filas de una matriz A, se puede trabajar de manera similar, pero con las columnas de A; es decir, las transformaciones elementales para las columnas de A son las mismas que para las filas, nada más que operando por columna.

CAPÍTULO 5

DETERMINANTES

1. Introducción

Los determinantes se definen como una función cuyo dominio son las matrices cuadradas de orden n, es decir que los determinantes son una función de funciones y en los mismos se relacionan las filas y las columnas de la matriz mediante sumatorias.

Una de las primeras culturas que abordó este tema fue la japonesa. Al matemático Seki Kowa (1642–1708) se le adjudica la invención de los determinantes ya que fue quien inició estudios sobre la teoría de las ecuaciones según los métodos de los matemáticos chinos. En el año 1683 introdujo el tema de los determinantes en el Kai Fukundai no Ho (Método de solución de problemas Fukundai). Con un procedimiento denominado *tatmu*, derivó las ecuaciones lineales de dos ecuaciones cuadráticas y encontró su determinante. Otros aportes de este mismo autor fueron la simplificación de los determinantes de grado mayor a dos al identificar el factor común de una determinada fila, la formulación del concepto de determinante y el establecimiento de muchas de sus propiedades.

En Europa, los determinantes aparecieron en la literatura matemática más de un siglo antes que las matrices. En 1693, el filósofo alemán Gottfried W. Leibniz (1646–1716) escribió una carta a Guillaume F. Antoine (1661–

1704), marqués de l'Hôpital, en la que resolvía un sistema de ecuaciones, proporcionando de esta manera la descripción del determinante por permutaciones que actualmente se utiliza.

En 1748 se publicó un tratado de álgebra del matemático británico Colin Maclaurin (1698–1746), donde se exponía la resolución de sistemas de dos y tres ecuaciones lineales con el mismo número de incógnitas, llegándose de este modo a lo que hoy se conoce como el determinante de segundo y tercer orden.

Algo similar ocurrió con el matemático suizo Gabriel Cramer (1704–1752) quien publicó, en 1750, el tratado de geometría *Introducción al análisis de las líneas curvas algebraicas*, donde proponía la solución de sistemas de ecuaciones lineales con el mismo número de incógnitas y ecuaciones. El método es similar al publicado en el tratado de Maclaurin, pero es más conocido por su claridad de notación y generalización.

En 1771, Alexandre T. Vandermonde (1735–1796) realizó la primera exposición lógica y formal de la teoría de los determinantes, reconociéndolos como funciones independientes; mientras que, en 1772, Pierre Laplace (1749–1827) generalizó el método de Vandermonde para el desarrollo de los determinantes en productos de menores, y presentó el método general de la expansión de un determinante en términos de sus menores complementarios. Al año siguiente, Joseph L. de Lagrange (1736–1813) expuso la solución de determinantes de segundo y tercer

orden y los utilizó para fines distintos al de la solución de ecuaciones lineales.

En 1801, Carl F. Gauss (1777–1855) utilizó la palabra determinante en su teoría de números. Para 1812, el matemático francés Agustin L. Cauchy (1789–1857) presentó en la Academia de Ciencias de Francia un artículo con el teorema de la multiplicación de determinantes, utilizando la palabra determinante de Gauss en su sentido actual, con lo que mejoró la notación. En tanto que, en 1815, este mismo autor publicó una memoria en la que mejoraba el desarrollo de Laplace, proporcionando así la primera exposición sistemática de los determinantes mediante la disposición de los elementos en filas y columnas y la notación de los índices dobles.

En 1829, el matemático alemán Carl G. Jakob Jacobi (1804–1851) utilizó por primera vez los determinantes funcionales, que más tarde fueron llamados jacobianos por Sylvester en sus memorias del *Journal Crelle*, en 1841.

Durante el último cuarto del siglo XIX, el matemático británico Charles Lutwidge Dogson (1832–1898) –mejor conocido por el seudónimo de Lewis Carroll– y algunos otros estudiosos, también enriquecieron la teoría de los determinantes con operaciones numéricas, así como con resultados nuevos y complementarios a los ya desarrollados por sus antecesores.

Finalmente, las aplicaciones de los determinantes son diversas y variadas. A modo de ejemplo, en economía se los utiliza para encontrar los puntos de equilibrio de un sistema financiero descrito matricialmente; en ingeniería para optimizar procesos; y, en general, en matemáticas se los utiliza –en el caso ya estudiado de las matrices– para saber si una matriz cuadrada es invertible o no o bien para calcular el rango de las mismas. Otros temas que se verán más adelante y en los cuales se hace uso de los determinantes son: resolución de sistemas de ecuaciones lineales y la independencia lineal de vectores.

Otras aplicaciones, en geometría, son el cálculo de áreas de polígonos en el plano de los que se conocen las coordenadas de los vértices o bien las de ciertos vectores, y el cálculo de volúmenes de paralelepípedos.

De esta manera, en este quinto capítulo se estudiará, en primer lugar, la definición de determinantes, para lo cual se hace necesario hacer un pequeño análisis de permutaciones simples, cómo calcularlas y formarlas; como así también, de algunos conceptos relacionados con éstas. Para poder definir matriz adjunta y de cofactores, primero se estudiará menor complementario y adjunto de un elemento de una matriz cuadrada. Todo esto ayudará a que el desarrollo de un determinante por los elementos de una línea sea mucho más comprensible y, de esta forma, se podrá arribar, de una manera más sencilla, a la definición inductiva del determinante de una matriz cuadrada de orden n.

Las propiedades de los determinantes, que se estudian en el apartado 4, permitirán, luego de analizar la relación entre determinantes y matrices equivalentes, proponer el cálculo de determinantes de orden superior a tres mediante dos métodos: el de triangulación y el de Chío.

En el apartado 7 se propone un segundo método para calcular la matriz inversa: mediante la matriz adjunta; mientras que, en el apartado 8, simplemente se pone en evidencia cómo se vinculan los determinantes con la matriz inversa, el rango de una matriz y la equivalencia con la matriz identidad.

Finalmente, se utilizan los determinantes para el cálculo del rango de una matriz cualquiera por el método del orlado.

1.1. Conceptos básicos

Una parte importante de la matemática discreta es la "*combinatoria*", que consiste en estudiar las posibles agrupaciones de objetos. Por ejemplo, contar el número de objetos que cumplen cierta propiedad, es un problema que se resuelve con técnicas combinatorias. Se pueden interpretar muchos problemas combinatorios en términos de las agrupaciones ordenadas o desordenadas de los elementos de un conjunto; a estas agrupaciones se las suelen llamar permutaciones y combinaciones.

Permutaciones simples: *Dado un conjunto finito de* n *elementos, se llama permutaciones simples de orden* n, (P_n), *a todo conjunto ordenado y formado con los* n *elementos.*

P_n: "se lee permutaciones de orden n"

Dos permutaciones simples cualesquiera tienen exactamente los mismos elementos y difieren solamente en el orden de colocación de estos.

Ejemplo:

Sea el conjunto $\{a, b, c, d\}$, es decir $n = 4$, algunas de las P_4 son: abcd, bacd, cdba, dcab, cabd, etc.

¿Cuántas permutaciones de orden n pueden obtenerse?

Supóngase que a los n elementos se los debe ubicar en n casillas, pues se deben utilizar todos ellos. Esto se muestra en el siguiente diagrama.

n–c a s i l l a s

C_1	C_2	C_3	C_4	. . .	C_5
n elecciones	n – 1 elecciones	n – 2 elecciones	n – 3 elecciones	. . .	n – (n – 1) elecciones

La cantidad de elecciones distintas que se pueden hacer, al tratar de llenar la primera casilla, es n. Elegido un elemento para llenar la primera casilla (C_1), para ocupar la segunda casilla, (C_2), podrán hacerse $(n-1)$ elecciones diferentes. Es decir que habrá $n(n-1)$ elecciones posibles para las dos primeras casillas. Siguiendo este razonamiento, para las tres primeras casillas $(C_1, C_2\ y\ C_3)$, habrá $n(n-1)(n-2)$ posibles elecciones distintas. Finalmente, solo restará una elección para llenar la última casilla.

Luego, el número de permutaciones simples de n elementos se puede calcular como:

$P_n = n(n-1)(n-2) \ldots 3 .2 .1$

$P_n = n!$ [25] en donde $n!$ se lee "n factorial o bien factorial de n".

Por ejemplo, las $P_4 = 4! = 4 .3 .2 .1 = 24$

Es decir que con 4 elementos distintos {a, b, c, d}, podemos formar 24 permutaciones simples. ¿Cuáles son esas permutaciones?

Debemos colocar cuatro elementos {a, b, c, d} en cuatro casillas C_1, C_2, C_3, C_4 . Supongamos que colocamos en C_1 el elemento a, por lo tanto, en C_2 podemos colocar b, c o d, y así sucesivamente con las demás casillas y elementos. Todo esto se muestra en el siguiente diagrama de árbol.

$$\overset{C_1}{\widetilde{a}}\begin{cases}\overset{C_2}{\widetilde{b}}\begin{cases}\overset{C_3}{\widetilde{c}}\begin{cases}\overset{C_4}{\widetilde{d}} \rightarrow \overbrace{abcd}^{PERM.}\\ d\{c \rightarrow abdc\end{cases}\\ c\begin{cases}b\{d \rightarrow acbd\\ d\{b \rightarrow acdb\end{cases}\\ d\begin{cases}b\{c \rightarrow adbc\\ c\{b \rightarrow adcb\end{cases}\end{cases}$$

Estas son las seis permutaciones simples cuyo primer elemento es a. Idéntico razonamiento se hace colocando en C_1 el elemento b y en C_2 los restantes elementos, y de igual manera para los elementos $c\ y\ d$, obteniéndose las 24 permutaciones simples.

[25] La función factorial se representa con un signo de exclamación "!" detrás de un número (n). Esto significa que se debe multiplicar todos los números enteros positivos que hay entre ese número y el 1. Por convención 1! = 0! = 1.

Otro ejemplo que nos será de utilidad en un desarrollo posterior es considerar que los objetos con los cuales formamos las permutaciones son números (subíndices).

¿Cuántas y cuáles son las permutaciones que se pueden hacer con los elementos del conjunto {1, 2, 3}?

Cantidad de permutaciones: $P_3 = 3! = 3 .2 .1 = 6$

$$Permutaciones \begin{cases} 1 \begin{cases} 2\,\{3 \longrightarrow 123 \\ 3\,\{2 \longrightarrow 132 \end{cases} \\ 2 \begin{cases} 1\,\{3 \longrightarrow 213 \\ 3\,\{1 \longrightarrow 231 \end{cases} \\ 3 \begin{cases} 1\,\{2 \longrightarrow 312 \\ 2\,\{1 \longrightarrow 321 \end{cases} \end{cases}$$

Existen otras cuestiones, respecto de las permutaciones, que nos serán de gran ayuda a la hora de entender la definición de determinantes.

Llamamos *permutación principal* a la permutación en la cual los n elementos están en un orden natural según los elementos con los que se trabaje. En el primer ejemplo la permutación principal es $abcd$, mientras que en el último ejemplo es 123.

En otra permutación cualquiera, decimos que dos elementos forman una *permanencia* cuando, sin tener en cuenta los restantes, están en el mismo orden que en la permutación principal, y que forman *inversión* cuando están en orden contrario.

Por ejemplo, la permutación:

a) 132 presenta una inversión, pues 3 precede a 2

b) 312 presenta dos inversiones, ya que 3 precede a 2, y 3 precede a 1

Decimos que una permutación es de *clase par* si, como en el ejemplo b), es par el número de sus inversiones y, que es de *clase impar* si, como en el ejemplo a), es impar el número de inversiones. A toda permutación de clase par o simplemente permutación par, se le asigna un signo positivo, mientras que, a una impar, se le asigna un signo negativo.

Por ejemplo, la permutación:

a) 132, de clase impar, se le asigna un signo negativo (–)

b) 312, de clase par, se le asigna un signo positivo (+)

Con todos estos conceptos vistos hasta el momento, estamos en condiciones de comprender la siguiente:

Definición: *Llamaremos determinante a la función cuyo dominio es el conjunto de las matrices cuadradas de orden* n *y cuyo codominio es el conjunto* $\mathbb{K}$*, de modo que la imagen de una matriz es un número que se obtiene formando todos los productos posibles de* n *elementos elegidos entre los* n^2 *de la matriz dada, siempre que en cada producto haya un factor de cada fila y un factor de cada columna, anteponiendo a cada producto el signo más o menos, según que las permutaciones que indican las filas y las columnas sean de clase par o de clase impar, respectivamente.*

Dada la matriz cuadrada de orden 2,

$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ el determinante de A, que indicaremos como $det(A)$, $\Delta(A)$ o $| A |$, es:

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

Observemos los siguientes detalles:

Dada la matriz A, su correspondiente determinante es $det(A)\ o\ \Delta(A)\ o\ |A|$

$|A|$ es un número que se obtiene sumando los productos de 2 elementos $(n = 2)$ cada uno, que se pueden formar con los cuatro $(n^2 = 4)$ elementos de la matriz.

El producto $a_{11}\ a_{22}$ es precedido por un signo positivo porque, si dejamos en el orden natural los subíndices que indican las filas, los subíndices que indican las columnas no presentan inversión alguna, por lo tanto, se le asigna un signo positivo. Por el contrario, en el producto $a_{12}\ a_{21}$, los subíndices que indican las columnas presentan una inversión y esta permutación es de clase impar, en consecuencia, se le asigna un signo negativo.

Ejemplo: hallar el determinante de la matriz $A = \begin{pmatrix} -1 & -3 \\ 5 & -2 \end{pmatrix}$

$$|A| = \begin{vmatrix} -1 & -3 \\ 5 & -2 \end{vmatrix} = -1(-2) - (-3)5 = 17$$

En general, como los factores que forman los productos del determinante pertenecen a filas sucesivas, los primeros subíndices están en orden natural $1, 2, \ldots n$, pero estos factores también pertenecen a columnas sucesivas, entonces la sucesión de los segundos subíndices forma una permutación P tal que $P = p_1\ p_2\ p_3\ \ldots p_n$.

Por lo tanto, a partir de una matriz cuadrada de orden n, se pueden formar $n!$ productos distintos. Por ejemplo, para la matriz $A =$

$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, cuadrada de orden 2, podemos formar $2! = 2$ productos distintos $a_{11}\, a_{22}$ y $a_{12}\, a_{21}$.

Veamos, ahora, el determinante de una matriz de orden tres.

Dada la matriz A:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \Longrightarrow |\, A \,| = \sum_P sg(P) a_{1p_1} a_{2p_2} a_{3p_3}$$

donde $sg(P)$ es el signo que depende de la paridad de la permutación.

$$|\, A \,| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum_P sg(P) a_{1p_1} a_{2p_2} a_{3p_3}$$

$$1\begin{cases} 2\begin{cases} 3 \rightarrow \overbrace{1\,2\,3}^{p_1\, p_2\, p_3} + \\ \\ 3\{\ 2 \rightarrow 132 \quad - \end{cases} \\ 2\begin{cases} 1\{3 \rightarrow 213 - \\ 3\{1 \rightarrow 231 + \end{cases} \\ 3\begin{cases} 1\{2 \rightarrow 312 + \\ 2\{1 \rightarrow 321 - \end{cases} \end{cases}$$

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

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

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

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

Esta última expresión es el desarrollo de un determinante de orden tres por la Regla de Sarrus, la que solo se aplica a determinantes de hasta orden tres.

Obsérvese que, por ser la matriz de orden 3 ($n = 3$), se pudo hacer $3! = 6$ productos distintos. Si la matriz fuese de orden cuatro, el desarrollo tendría $4! = 24$ términos y, si fuese de orden 5, el desarrollo tendría $5! = 120$ términos; lo que hace que el cálculo del determinante se haga muy trabajoso aplicando la definición dada.

En general, dada una matriz A cuadrada de orden n, su determinante será:

$$\sum_P sg(P) a_{1p_1} a_{2p_2} a_{3p_3} \ \dots \ a_{np_n} \ \ donde\ P = P_n \ \ y\ sg(P) = \begin{cases} + si\ P\ es\ par \\ - si\ P\ es\ impar \end{cases}$$

Ejemplo: hallar el determinante asociado a cada una de las siguientes matrices:

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

$$\Longrightarrow |A| = \begin{vmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{vmatrix} = 0 + 2 - 9 + 0 - 12 + 15 = -4$$

$$A = \begin{pmatrix} 2+i & 4-i \\ 6 & 5i \end{pmatrix}$$

$$\Longrightarrow |A| = \begin{vmatrix} 2+i & 4-i \\ 6 & 5i \end{vmatrix} = (2+i)5i - 6(4-i) = 16i - 29$$

2. Matriz de cofactores y Matriz adjunta

En todas las ramas de las matemáticas, las matrices son una herramienta muy valiosa, sobre todo, aquellas matrices que son invertibles, es decir, aquellas que admiten inversas. Existen diversos métodos para la obtención de la matriz inversa, por ejemplo, el de Gauss–Jordan o aquel que utiliza la matriz adjunta. El primero se estudió en el capítulo

correspondiente a matrices y utiliza transformaciones elementales entre las filas de la matriz dada, mientras que, en el segundo, se calculan determinantes. Cada método tiene sus ventajas y desventajas.

Antes de definir matriz adjunta, necesitamos desarrollar otros conceptos.

Menor complementario: *Dada la matriz A cuadrada de orden n, se define menor complementario de un elemento a_{ij}, denotado como m_{ij}, como el determinante de la matriz de orden $(n-1)$ que resulta de suprimir en A, la fila i y la columna j.*

Cofactor[26]**:** *de un elemento a_{ij}, denotado como A_{ij}, es el producto de su menor complementario, m_{ij}, por el signo que resulta de calcular $(-1)^{i+j}$.*

Dada $A \in K^{n\,x\,n}$, $A_{ij} = (-1)^{i+j} \,.\, m_{ij}$.

Ejemplo: Sea la matriz $A = \begin{pmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{pmatrix}$:

$m_{23} = \begin{vmatrix} 2 & -1 \\ 1 & -3 \end{vmatrix} = -5$, mientras que $m_{33} = \begin{vmatrix} 2 & -1 \\ 3 & 0 \end{vmatrix} = 3$. Por lo tanto:

$A_{23} = (-1)^{2+3} \,.\, m_{23} = (-1)^5 \,.\, (-5) = 5$, mientras que

$A_{33} = (-1)^{3+3} \,.\, m_{33} = (-1)^6 .\ 3 = 3$

Debemos destacar que el menor complementario m_{ij}, de un elemento de A, es igual a m_{ji} en A^t, ya que eliminar las líneas (i, j) en A es lo mismo que eliminar las líneas (j, i) en A^t. De idéntica manera, A_{ij} en A es igual a A_{ji} en A^t, es decir, $A_{ij} = A^t_{ij}$.

[26] También suele denominarse Adjunto.

Matriz de cofactores: *de A, es la matriz que resulta de sustituir cada elemento de A por su cofactor A_{ij}. Suele denotarse como $Cof(A)$.*

La importancia de esta última matriz es que permite definir matriz adjunta.

Ejemplo: Sea la matriz $A = \begin{pmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{pmatrix}$

$A_{11} = -6$ $\quad A_{21} = 2$ $\quad A_{31} = 2$

$A_{12} = -17$ $\quad A_{22} = 9$ $\quad A_{32} = 7$

$A_{13} = -9$ $\quad A_{23} = 5$ $\quad A_{33} = 3$

Por lo tanto

$$Cof(A) = \begin{pmatrix} -6 & -17 & -9 \\ 2 & 9 & 5 \\ 2 & 7 & 3 \end{pmatrix}$$

Se puede demostrar que $\big(Cof\ (A)\big)^t = Cof(A^t)$.

Analicemos los cofactores de A^t: m_{ij} de A^t es igual a m_{ji} de A, porque eliminando las líneas (i, j) de A^t y las líneas (j, i) de A, nos siguen quedando dos matrices que son una la traspuesta de la otra (con esas líneas eliminadas y de orden $(n-1)$). Por ser esas matrices traspuesta una de otra, tienen el mismo determinante[27].

Multiplicando los menores complementarios anteriores por $(-1)^{i+j}$ tenemos:

$(-1)^{i+j}\ m_{ij} = (-1)^{i+j}\ m_{ji}$

$A_{ij}\ de\ A^t = A_{ji}\ de\ A.$

[27] Esta propiedad se demostrará en el apartado Propiedades de los determinantes.

Si denotamos la matriz de cofactores de A^t como B y a la matriz de cofactores de A como C, la igualdad anterior nos indica $b_{ij} = c_{ji}$ es decir $B = C^t$. Por lo tanto:

$\big(Cof\ (A)\big)^t = Cof(A^t)$

Esto trae como consecuencia que la matriz adjunta se pueda calcular como la traspuesta de la matriz de cofactores de la matriz dada o bien como la matriz de cofactores de la traspuesta de la matriz dada.

Matriz adjunta: *Dada una matriz B cuadrada de orden n, se llama matriz adjunta de B, denotada como $Adj\ B$, a la matriz traspuesta de los cofactores de los elementos b_{ij} de B.*

Sea $B = \|b_{ij}\| \in K^{n\,x\,n} \Longrightarrow Adj\ B = \|A_{ji}\|$

$$Adj\ B = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \cdots & A_{n1} \\ A_{12} & A_{22} & A_{32} & \cdots & A_{n2} \\ A_{13} & A_{23} & A_{33} & \cdots & A_{n3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & A_{3n} & \cdots & A_{nn} \end{pmatrix}$$

Ejemplo: sea la matriz $B = \begin{pmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{pmatrix}$, sabemos, por el ejemplo anterior, que:

$Cof(B) = \begin{pmatrix} -6 & -17 & -9 \\ 2 & 9 & 5 \\ 2 & 7 & 3 \end{pmatrix}$, por lo tanto $Adj\ B = \begin{pmatrix} -6 & 2 & 2 \\ -17 & 9 & 7 \\ -9 & 5 & 3 \end{pmatrix}$

3. Desarrollo de un determinante por los elementos de una línea

A partir de las definiciones de menor complementario y de cofactor, se puede dar una nueva definición de determinante para poder hacer su

cálculo más fácil, sobre todo, cuando éste es de orden mayor a tres. A diferencia del cálculo por definición, cuya cantidad de términos depende del orden del determinante –orden 2, $2! = 4$ términos, orden 3, $3! = 6$ términos, orden 4, $4! = 24$ términos, así sucesivamente el de orden n, tendrá $n!$ términos–, en esta nueva definición inductiva[28], se recurre a disminuir el orden del determinante que se calcula.

Consideremos, en primer lugar, el caso trivial, sea la matriz de orden 1, $A = (a)$, cuyo determinante será $| A | = a$.

En segundo lugar, sabemos calcular el determinante de una matriz de orden dos, es decir: sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, cuyo determinante será $| A | = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$.

Reescribiendo este determinante: $| A | = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12}$, es decir, se utilizó los cofactores de los elementos de la primera fila. Esta es la expansión del $det(A)$ a lo largo de la fila 1.

También podemos escribir el mismo determinante como:

$$| A | = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{21}(-a_{12}) + a_{22}\, a_{11}$$

$$| A | = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{21}\, A_{21} + a_{22}\, A_{22}$$

Esta es la expansión por cofactores del $det(A)$ a lo largo de la fila 2.

[28] Definir inductivamente objetos de un conjunto, consiste en indicar i) los objetos iniciales o elementales del conjunto, ii) las reglas u operaciones que permiten formar con los objetos que se dispone, nuevos objetos del conjunto.

Usando ahora las columnas en lugar de las filas, podemos realizar una expansión del $det(A)$ a lo largo de cualquiera de las columnas.

$de\,t(A) = a_{11}a_{22+}a_{21}(-a_{12}) = a_{11}A_{11} + a_{21}A_{21}$ (primera columna)

$de\,t(A) = a_{12}(-a_{21}) + a_{22}\,a_{11} = a_{12}\,A_{12} + a_{22}\,A_{22}$ (segunda columna)

Comprobemos estos resultados mediante un ejemplo.

Calcular el determinante de la matriz $A = \begin{pmatrix} 2 & -1 \\ -2 & 3 \end{pmatrix}$

$|\,A\,| = \begin{vmatrix} 2 & -1 \\ -2 & 3 \end{vmatrix} = 2(-1)^2 3 + (-1)(-1)^3(-2) = 6 - 2 = 4$ expandido por la F_1

$|\,A\,| = \begin{vmatrix} 2 & -1 \\ -2 & 3 \end{vmatrix} = -2(-1)^3(-1) + 3(-1)^4 2 = -2 + 6 = 4$ expandido por la F_2

$|\,A\,| = \begin{vmatrix} 2 & -1 \\ -2 & 3 \end{vmatrix} = 2(-1)^2\,3 + (-2)(-1)^3(-1) = 6 - 2 = 4$ expandido por la C_1

$|\,A\,| = \begin{vmatrix} 2 & -1 \\ -2 & 3 \end{vmatrix} = -1(-1)^3(-2) + 3(-1)^4 2 = -2 + 6 = 4$ expandido por la C_2

En tercer lugar, consideremos una matriz de orden 3, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, entonces:

$|\,A\,| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$ si expandimos por la F_1

Y de la misma manera que en el caso anterior, el mismo determinante se puede expandir a lo largo de cualquier fila o cualquier columna.

$det(A) = a_{i1}\,A_{i1} + a_{i2}\,A_{i2} + a_{i3}\,A_{i3}$ con $i = 1, 2, 3$ expansión por fila i

$det(A) = a_{1j}\,A_{1j} + a_{2j}\,A_{2j} + a_{3j}\,A_{3j}$ con $j = 1, 2, 3$ expansión por columna j

Hagamos uso de la matriz $\mathrm{A} = \begin{pmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{pmatrix}$ cuyo determinante ya fue calculado por definición ($det(A) = -4$).

$$|\mathrm{A}| = \begin{vmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{vmatrix} = 1(-1)^4 \begin{vmatrix} -1 & 1 \\ 0 & -2 \end{vmatrix} + (-3)(-1)^5 \begin{vmatrix} 2 & 1 \\ 3 & -2 \end{vmatrix} +$$

$5(-1)^6 \begin{vmatrix} 2 & -1 \\ 3 & 0 \end{vmatrix}$ si expandimos por la F_3.

$$det(A) = 2 + 3(-7) + 5\,.3 = 2 - 21 + 15 = -4$$

$$|\mathrm{A}| = \begin{vmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{vmatrix} = 2(-1)^2 \begin{vmatrix} 0 & -2 \\ -3 & 5 \end{vmatrix} + 3(-1)^3 \begin{vmatrix} -1 & 1 \\ -3 & 5 \end{vmatrix} +$$

$1(-1)^4 \begin{vmatrix} -1 & 1 \\ 0 & -2 \end{vmatrix}$ si expandimos por la C_1.

$$det(A) = 2(-6) + (-3)(-2) + 2 = -12 + 6 + 2 = -4$$

Estos ejemplos nos permiten ir haciendo algunas conjeturas; en primer lugar, que no importa cómo se haga la expansión del determinante, siempre se obtiene el mismo resultado y, en segundo lugar, que el resultado del determinante es único.

Finalmente, consideremos el caso general de una matriz de orden n y formalicemos la siguiente:

Definición inductiva: *sea A una matriz de orden n, entonces, si:*

$n = 1$, *caso trivial*, $det(A) = a_{11}$

$n > 1$ *podemos definir el* $det(A)$ *como:*

$$det(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \quad \textit{expansión por fila } i$$

$$det(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \quad \textit{expansión por columna } j$$

Donde A_{ij} *es el cofactor del elementos* (i, j) *de la matriz A.*

De esta forma, el determinante de una matriz de orden n queda expresado como una combinación de n determinantes de orden $(n-1)$.

En la práctica, los determinantes se expanden a lo largo de la fila o columna que contenga la mayor cantidad de ceros, para evitar el cálculo del cofactor de dichos elementos.

Como caso particular del desarrollo de un determinante por los elementos de una línea, si en un determinante todos los elementos de una línea son nulos (excepto uno), el determinante es igual al producto de ese elemento no nulo por el adjunto del mismo.

Ejemplo:

$$|\,A\,| = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 5 & 2 & -1 & 1 \\ -8 & 3 & 0 & -2 \\ 7 & 1 & -3 & 5 \end{vmatrix} = 1(-1)^2 \begin{vmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{vmatrix} = -4$$

Con esta definición, la obtención de un determinante de orden n se reduce a calcular n determinantes de orden $(n-1)$, y así hasta que los determinantes de orden inferior a resolver sean de orden 2 o 3. De esta

forma, un determinante de orden 5, en general se resolvería con 5 determinantes de orden 4, uno de orden 4 con 4 determinantes de orden 3, uno de orden 3 con 3 de orden 2. Así, un determinante de orden 5 se resolvería con $5 \cdot 4 = 20$ determinantes de orden 3, o 60 de orden 2, $(5 \cdot 4 \cdot 3 = 60)$. Se reduce el problema de calcular un determinante de orden superior a calcular más determinantes, pero de orden menor y obviamente más sencillos. Más adelante se explicará un método, basado en propiedades de los determinantes, que reducirá en la práctica el número de menores (determinantes de orden inferior) a resolver, pues se omitirá el cálculo de muchos de los adjuntos, al conseguir que el elemento correspondiente sea cero, lo que anula el sumando resultante de multiplicar el elemento nulo por su adjunto.

4. Propiedades de los determinantes

Las propiedades de los determinantes permiten, conjuntamente con la definición anterior, hacer mucho más sencillos los cálculos cuando se trata de matrices de orden superior o igual a tres. A continuación, se enunciarán y demostrarán estas propiedades, haciendo uso de un determinante genérico de orden tres, ya que esto no hace que se pierda generalidad.

Sea $A \in K^{n \times n}$ una matriz cuadrada de orden n, se cumple que:

P_1) El determinante de una matriz A y el de su traspuesta son iguales.

D) $| A | = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ por lo tanto $| A^t | = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$

Todo término del $|A|$ está formado por n elementos, uno de cada fila y uno de cada columna, luego pertenece también a $|A^t|$.

Las dos permutaciones que indican filas (columnas) en $|A^t|$ son las mismas que indican columnas (filas) en $|A|$, luego el signo de dicho término en ambos determinantes es $(+)\, o\, (-)$, según que ambas permutaciones sean de clase par o impar, respectivamente.

Así, por ejemplo, en el caso considerado:

En $|A|$

$$a_{12}\, a_{23}\, a_{31} \begin{cases} 1\ 2\ 3\ par\ (+) \\ 2\ 3\ 1\ par\ (+) \end{cases} \Longrightarrow (+)(+) = +\, par$$

En $|A^t|$

$$a_{21}\, a_{32}\, a_{13} \begin{cases} 2\ 3\ 1\ par\ (+) \\ 1\ 2\ 3\ par\ (+) \end{cases} \Longrightarrow (+)(+) = +\, par$$

Luego $|A| = |A^t|$

Esta propiedad permite afirmar que toda otra propiedad de los determinantes referida a sus filas es también válida para sus columnas.

Ejemplo: sea la matriz $A = \begin{pmatrix} 3 & 4 & -1 \\ 2 & -2 & 5 \\ 1 & 0 & 3 \end{pmatrix}$, entonces $A^t = \begin{pmatrix} 3 & 2 & 1 \\ 4 & -2 & 0 \\ -1 & 5 & 3 \end{pmatrix}$,

por lo tanto

$$|A| = \begin{vmatrix} 3 & 4 & -1 \\ 2 & -2 & 5 \\ 1 & 0 & 3 \end{vmatrix} = -18 + 0 + 20 - 2 - 0 - 24 = -24$$

$$|A^t| = \begin{vmatrix} 3 & 2 & 1 \\ 4 & -2 & 0 \\ -1 & 5 & 3 \end{vmatrix} = -18 + 20 + 0 - 2 - 24 - 0 = -24$$

P2) Si se permutan dos líneas cualesquiera de una matriz, los correspondientes determinantes son opuestos.

D)

$$|A_1| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \text{ por lo tanto } |A_2| = \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Intercambiar entre sí dos filas significa una trasposición en los primeros índices, lo cual cambia la clase de la permutación y, como la de los segundos índices correspondientes a las columnas no ha variado, habrá un cambio de signo en cada término del determinante.

En $|A_1|$

$$a_{12}\, a_{23}\, a_{31} \begin{cases} 1\ 2\ 3\ par\ (+) \\ 2\ 3\ 1\ par\ (+) \end{cases} \Longrightarrow (+)(+) = +\, par$$

En $|A_2|$

$$a_{22}\, a_{13}\, a_{31} \begin{cases} 2\ 1\ 3\ impar\ (-) \\ 2\ 3\ 1\ par\ (+) \end{cases} \Longrightarrow (-)(+) = -\, impar$$

Luego $|A_1| = -|A_2|$

Ejemplo: sea la matriz $A_1 = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 1 & 5 \\ 6 & 3 & 3 \end{pmatrix}$, entonces $A_2 = \begin{pmatrix} 1 & 3 & 2 \\ 6 & 3 & 3 \\ 3 & 1 & 5 \end{pmatrix}$, por lo tanto

$$|A_1| = \begin{vmatrix} 1 & 3 & 2 \\ 3 & 1 & 5 \\ 6 & 3 & 3 \end{vmatrix} = 3 + 18 + 90 - 12 - 27 - 15 = 57$$

$$|A_2| = \begin{vmatrix} 1 & 3 & 2 \\ 6 & 3 & 3 \\ 3 & 1 & 5 \end{vmatrix} = 15 + 12 + 27 - 18 - 3 - 90 = -57$$

Consecuencia de la segunda propiedad:

CP_2) El determinante de toda matriz que tenga dos líneas idénticas es nulo.

Si $|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$

Al permutar dos filas consecutivas (la primera y la segunda, por ejemplo), por la propiedad anterior, el determinante será $-|A|$, pero siendo idénticas las filas, el nuevo determinante es igual al anterior. Es decir $|A| = -|A|$

Luego: $|A| = -|A| \Leftrightarrow 2|A| = 0 \Leftrightarrow |A| = 0$

Ejemplo: sea la matriz $A = \begin{pmatrix} 1 & 1 & 3 \\ -1 & -1 & 2 \\ 1 & 1 & 0 \end{pmatrix}$, por lo tanto:

$$|A| = \begin{vmatrix} 1 & 1 & 3 \\ -1 & -1 & 2 \\ 1 & 1 & 0 \end{vmatrix} = 0 - 3 + 2 + 3 - 2 + 0 = 0$$

P_3) Si se multiplican todos los elementos de una línea de una matriz por un mismo escalar λ no nulo, el determinante correspondiente queda multiplicado por dicho escalar.

Si $|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ y multiplicamos por λ los elementos de la primera fila, resulta

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

Según el desarrollo de los determinantes por elementos de una línea, tenemos:

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

$$= \lambda [a_{11}A_{11} + a_{12} A_{12} + a_{13} A_{13}] = \lambda |A|$$

Esta propiedad permite separar, como factor del determinante, un factor común a todos los elementos de una línea cualquiera, simplificando esta.

Ejemplo: sea la matriz $A = \begin{pmatrix} 10 & -5 & 5 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{pmatrix}$, por lo tanto

$$|A| = \begin{vmatrix} 10 & -5 & 5 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{vmatrix} = 0 + 10 - 45 + 0 - 60 + 75 = -20$$

Extrayendo 5 como factor común de la primera fila:

$$|A| = 5\begin{vmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \\ 1 & -3 & 5 \end{vmatrix} = 5(0 - 9 + 2 - 0 + 15 - 12) = 5(-4) = -20$$

Consecuencia de la tercera propiedad:

CP_3) Si en el determinante de una matriz, los elementos de una línea son proporcionales a los correspondientes de otra paralela a ella, el determinante es nulo.

En efecto, si separamos el coeficiente de proporcionalidad como factor del determinante, queda otro con dos líneas iguales y por CP_2) es nulo.

Ejemplo: sea la matriz $A = \begin{pmatrix} 2 & -1 & 8 \\ 3 & 0 & 12 \\ 1 & -3 & 4 \end{pmatrix}$, por lo tanto $|A| = \begin{vmatrix} 2 & -1 & 8 \\ 3 & 0 & 12 \\ 1 & -3 & 4 \end{vmatrix}$

Se puede observar que la tercera columna es proporcional a la primera, $3°C = 4\,.\,1°C$

$$|A\,| = 4\begin{vmatrix} 2 & -1 & 2 \\ 3 & 0 & 3 \\ 1 & -3 & 1 \end{vmatrix} = 4\,.0 = 0$$

Esta consecuencia puede generalizarse de la siguiente manera:

CGP3) Si la línea de una matriz es combinación lineal[29] de las demás, el determinante correspondiente es nulo.

Ejemplo: sea la matriz $A = \begin{pmatrix} 2 & -1 & 8 \\ 3 & 2 & -2 \\ 8 & 3 & 4 \end{pmatrix}$, se puede observar que la fila 3 se obtiene sumando la fila 1 con el doble de la fila 2, es decir, $F_3 = F_1 + 2\,F_2$. Por lo tanto:

$$|A\,| = \begin{vmatrix} 2 & -1 & 8 \\ 3 & 2 & -2 \\ 8 & 3 & 4 \end{vmatrix} = 16 + 72 + 16 - 128 + 12 + 12 = 0$$

Segunda consecuencia de la tercera propiedad:

C2P3) Si se multiplican todos los elementos de una matriz por un mismo escalar λ no nulo, el determinante correspondiente queda multiplicado por el escalar elevado a una potencia igual al orden de la matriz.

La demostración de esta propiedad es evidente si se considera que toda matriz de orden n tiene n líneas, si se aplica n veces la propiedad 3 (una

[29] Para el caso de las líneas de una matriz, combinación lineal significa que una línea se obtuvo a partir de la suma de otras pudiendo estas estar multiplicadas por un escalar.

vez por cada línea) el determinante de esta matriz quedará multiplicado n veces por el escalar λ, o sea λ^n.

Ejemplo: sea la matriz $A = \begin{pmatrix} 1 & 3 & -2 \\ 6 & 3 & 3 \\ 3 & 1 & 5 \end{pmatrix}$ tal que $|A| = -45$, entonces

$-2A = \begin{pmatrix} -2 & -6 & 4 \\ -12 & -6 & -6 \\ -6 & -2 & -10 \end{pmatrix}$ por lo tanto $|-2A| = 360 = (-2)^3(-45)$

P4) Si todos los elementos de una línea de una matriz son ceros, su determinante es nulo.

D1) Esto es verdad, ya que, por la tercera propiedad, separando como factor común del determinante, el elemento nulo, resultara: $0 \,.\, |A| = 0$

D2) La veracidad de esta propiedad es evidente si aplicamos el desarrollo del determinante por los elementos de la línea formada por ceros.

Ejemplo: sea la matriz $A = \begin{pmatrix} 2 & 0 & 8 \\ 3 & 0 & -2 \\ 8 & 0 & 4 \end{pmatrix}$, entonces:

$|A| = \begin{vmatrix} 2 & 0 & 8 \\ 3 & 0 & -2 \\ 8 & 0 & 4 \end{vmatrix} = 0 \,.\, \begin{vmatrix} 2 & 0 & 8 \\ 3 & 0 & -2 \\ 8 & 0 & 4 \end{vmatrix} = 0$, o bien:

$$|A| = \begin{vmatrix} 2 & 0 & 8 \\ 3 & 0 & -2 \\ 8 & 0 & 4 \end{vmatrix} = 0.(-1)^3 28 + 0.(-1)^4(-56) + 0.(-1)^5(-28) = 0$$

P5) Si los elementos de una línea de una matriz se pueden escribir como suma de n términos, puede descomponerse el determinante correspondiente en suma de n determinantes que tienen las mismas restantes líneas y, en lugar de aquella, la formada por los primeros

sumandos, por los segundos, . . . , y por los n–ésimos sumandos, respectivamente.

D) Sea $A = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ entonces,

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

Desarrollando el determinante por los elementos de la primera fila, tenemos:

$$|A| = (a_{11}+b_{11})\,A_{11} + (a_{12}+b_{12})\,A_{12} + (a_{13}+b_{13})\,A_{13}$$

$$|A| = a_{11}A_{11} + b_{11}\,A_{11} + a_{12}A_{12} + b_{12}\,A_{12} + a_{13}A_{13} + b_{13}\,A_{13}$$

$$|A| = [\,a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}] + [\,b_{11}\,A_{11} + b_{12}\,A_{12} + b_{13}\,A_{13}]$$

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

Consecuencia de la quinta propiedad:

CP5) El determinante de una matriz no varía si a una línea se le suma una combinación lineal de otras.

Sea la matriz $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, entonces $|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$

Si a la primera fila le sumamos la segunda multiplicada por cualquier escalar λ, no nulo, resulta:

$\begin{vmatrix} a_{11}+\lambda a_{21} & a_{12}+\lambda a_{22} & a_{13}+\lambda a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ por la quinta propiedad

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} \lambda\, a_{21} & \lambda\, a_{22} & \lambda\, a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \text{ por la tercera propiedad}$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \lambda \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \text{ por la segunda propiedad}$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \lambda\, .0$$

$$= | A |$$

Esta consecuencia permite simplificar los determinantes, reduciendo a cero varios elementos de una misma línea mediante operaciones elementales entre líneas; cada elemento que se logre anular así, evita el cálculo de un cofactor al desarrollar el determinante por los elementos de esa línea.

Ejemplo: sea la matriz $A = \begin{pmatrix} 1 & 1 & -2 & 4 \\ 0 & 1 & 1 & 3 \\ 2 & -1 & 1 & 0 \\ 3 & 4 & 2 & -1 \end{pmatrix}$, entonces:

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

Efectuando operaciones elementales entre columnas, reducir a cero todos los elementos de una línea, excepto uno, y luego desarrollar el determinante por los elementos de esa línea.

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

$$|A| = \overset{C_2 + C_3}{} \begin{vmatrix} 1 & -1 & -2 & 4 \\ 0 & 2 & 1 & 3 \\ 2 & 0 & 1 & 0 \\ 3 & 6 & 2 & -1 \end{vmatrix}$$

$$|A| = \overset{C_1 + (-2)C_3}{} \begin{vmatrix} 5 & -1 & -2 & 4 \\ -2 & 2 & 1 & 3 \\ 0 & 0 & 1 & 0 \\ -1 & 6 & 2 & -1 \end{vmatrix}$$

Desarrollamos el determinante obtenido por los elementos de la tercera fila.

$$|A| = 1.(-1)^6 \begin{vmatrix} 5 & -1 & 4 \\ -2 & 2 & 3 \\ -1 & 6 & -1 \end{vmatrix}$$

$$|A| = -135$$

P6) El determinante de las matrices triangulares y diagonales es igual al producto de los elementos de la diagonal principal.

D1) Sea la matriz triangular superior $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$, por lo tanto:

$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix}$. Si desarrollamos el $det(A)$ por los elementos de la tercera fila:

$$|A| = (-1)^6 a_{33} \begin{vmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{vmatrix} = a_{11}\, a_{22}\, a_{33}$$

Similar razonamiento se sigue para el caso de una matriz triangular inferior.

D2) Sea la matriz diagonal $A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$, por lo tanto:

$$| A | = \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix}$$

Si desarrollamos el $det(A)$ por los elementos de la tercera fila o columna, tenemos

$$| A | = (-1)^6 a_{33} \begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix} = a_{11}\, a_{22}\, a_{33}$$

Similar razonamiento se sigue para el caso de una matriz escalar.

De esta propiedad se deduce en forma inmediata que el determinante de la matriz identidad es 1. Es decir que $| I_n | = 1$.

P7) Dadas dos matrices $A \in K^{n \times n}$ y $B \in K^{n \times n}$, el determinante del producto de $A.B$ es igual al producto de sus respectivos determinantes.

Sean $A, B \in K^{n \times n}$, entonces: $| A \,.\, B | = | A | \,.\, | B |$

Antes de demostrar esta propiedad, conocida también como Teorema de Cauchy–Binet, necesitaremos una proposición y un corolario de la misma, los que admitiremos sin demostración ya que estas exceden los objetivos del presente capítulo.

Proposición: Si $A \in K^{n \times n}$ es una matriz cualquiera y $P_1, \ldots, P_r$ son matrices elementales[30], entonces $| A\, P_1 \ \ldots P_r | = | A | \, | P_1 | \ \ldots \ | P_r |$.

Corolario: Si $P \in K^{n \times n}$ es producto de matrices elementales: $P = P_1 \ \ldots \ P_r$, entonces

[30] Matrices elementales son aquellas que se obtienen como el resultado de aplicar a la matriz identidad uno de los tres tipos de transformaciones elementales.

$|P| = |P_1| \ \ldots \ |P_r|$.

La demostración de este corolario es muy simple pues se toma como un caso particular del resultado anterior, tomando $A = I$ y recordando que $|I| = 1$.

Ahora estamos en condiciones de demostrar el enunciado de la séptima propiedad.

D) Supóngase, en primer lugar, que B es singular, es decir no admite inversa. En ese caso $|B| = 0$ y B' la forma reducida por columnas de B, tiene una columna de ceros. Pero $B' = B\,P$, donde P es el producto de matrices elementales, luego $B = B'P^{-1}$ donde P^{-1} también es producto de matrices elementales (la inversa de una matriz elemental también es una matriz elemental). Por lo tanto $A\,B = A\,B'P^{-1}$. Como B' tiene una columna de ceros, $A\,B'$ también la tiene, por lo tanto $|A\,B'| = 0$. Pero sabemos que, al ser P^{-1} producto de matrices elementales, $|A\ B| = |A\,B'\,P^{-1}| = |A\,B'|\,|P^{-1}| = 0$. Por lo tanto $|A\ B| = 0$ y el resultado es cierto en este caso.

En segundo lugar, corresponde suponer que B es no singular. Entonces tiene rango n, luego es producto de matrices elementales: $B = P_1 \ \ldots \ P_r$, pero en este caso la proposición anterior y el corolario nos indican que $|A\ B| = |A|\,|P_1| \ \ldots \ |P_r| = |A|\,|B|$. Con lo que queda demostrada la propiedad.

Este resultado se puede generalizar para el caso de m matrices cuadradas de orden n.

P8) Sea la matriz $A \in K^{n \times n}$ invertible, entonces su determinante es no nulo y además se cumple que $|A^{-1}| = \frac{1}{|A|}$

D) Sabemos que $A\,A^{-1} = A^{-1}\,A = I$, obteniendo los determinantes:

$|A\,A^{-1}| = |A^{-1}\,A| = |I|$ por la P 7

$|A||A^{-1}| = |A^{-1}||A| = 1$. Esto implica tres resultados:

$a)\ |A| \neq 0,\ \ b)\ |A^{-1}| \neq 0\ \ y\ \ c)\ \ |A^{-1}| = \frac{1}{|A|}$

P9) Sea la matriz $A \in K^{n \times n}$ ortogonal, entonces $|A| = 1\ \vee\ |A| = -1$

D) Teniendo en cuenta que una matriz cuadrada A invertible se dice ortogonal si y solo si el producto de dicha matriz por su traspuesta es la unidad[31].

Por lo tanto, si A es ortogonal se debe cumplir que:

$A.A^t = A^t.A = I_n$. Tomando determinantes y teniendo en cuenta la P1, $|A| = |A^t|$:

$|A||\ A^t| = |A^t||\ A| = |\ I_n| \Longrightarrow |A||\ A| = 1$, esto implica dos posibilidades a saber: $|A| = 1\ \vee\ |A| = -1$

P10) La suma de los productos de los elementos de una línea de una matriz por los adjuntos de los elementos correspondientes a una línea paralela a ella, es cero.

Sea la matriz $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, entonces: $|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$

[31] Definición 2 del Capítulo 4 "Matrices".

Sumando los productos de los elementos de la segunda fila por los adjuntos de la primera, obtenemos:

$a_{21}\,A_{11} + a_{22}\,A_{12} + a_{23}\,A_{13} = \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0$ por tener dos filas iguales.

5. Relación entre los determinantes de matrices equivalentes

Una cuestión importante, a la hora de calcular determinantes, es analizar cómo varía el determinante de una matriz al efectuar sobre ella operaciones elementales. La importancia radica en analizar qué relación tiene el determinante de la matriz equivalente obtenida con el de la matriz dada.

Para esto, debemos tener presente tres propiedades de los determinantes. En primer lugar, la P_2) "Si se permutan dos líneas cualesquiera de una matriz, los correspondientes determinantes son opuestos"; en segundo lugar, la P_3) "Si se multiplican todos los elementos de una línea de una matriz por un mismo escalar λ no nulo, el determinante correspondiente queda multiplicado por dicho escalar"; y, finalmente, la consecuencia de la quinta propiedad, CP_5) "El determinante de una matriz no varía si a una línea se le suma una combinación lineal de otras".

Por lo tanto, si $A \sim B$, entonces $|\,B\,| = |\,A\,|\,(-1)^k\,\lambda_1\,\lambda_2\,\ldots\,\lambda_p$ siendo k el número de intercambios de líneas y λ_i, todos no nulos, los factores utilizados al transformar A en B por medio de operaciones elementales.

La expresión anterior está indicando $A \sim B$, $|\,A\,| = 0 \iff |\,B\,| = 0$

Ahora, podemos analizar el caso particular, en que partiendo de $A \in K^{n \times n}$ realizamos operaciones elementales por filas hasta llegar a una forma escalonada que denotamos como U (eliminación guassiana) y de allí hasta llegar a la forma escalonada reducida.

Por ser U cuadrada y escalonada, es triangular superior, por lo tanto, $|\ U\ |$ es igual al producto de los elementos de la diagonal principal. Es decir:

$|\ U\ | = u_{11}\,.\,u_{22}\,.\,\ldots\ .\,u_{nn}$ y de aquí podemos deducir que:

a) $|A\ | \neq 0 \Leftrightarrow |\ U\ | \neq 0 \Leftrightarrow u_{ii} \neq 0\ \ \forall i = 1, \ldots, n \Leftrightarrow n°\ de\ pivotes = n \Leftrightarrow \rho(A) = n$

Si continuamos con las operaciones elementales por filas hasta llegar a la forma escalonada reducida, esta sería I_n.

b) $|\ A\ | = 0 \ \Leftrightarrow\ |\ U\ | = 0 \ \Leftrightarrow \exists\ u_{ii} = 0 \ \Leftrightarrow n°\ de\ pivotes\ menor\ que\ n,$ es decir, $\rho(A) < n$

Si continuamos con las operaciones elementales por filas hasta llegar a la forma escalonada reducida, esta NO sería I_n, ya que el número de pivotes es menor que n.

6. Cálculo de determinantes

Hasta el momento sabemos calcular determinantes de orden mayor a tres mediante el desarrollo por los elementos de una de sus líneas, sin embargo, la eficacia de este método no es muy buena, ya que para calcular el determinante de una matriz de orden n hay que calcular n determinantes de orden $(n-1)$ y, para cada uno de estos hay que calcular

$(n-1)$ determinantes de orden $(n-2)$, y así sucesivamente. Por lo tanto, el número de operaciones que hay que efectuar es del orden $n!$.

A continuación, se estudiarán dos métodos mucho más simples que permiten calcular determinantes de una manera más rápida y que se basan, fundamentalmente, en la relación que hay entre los determinantes de matrices equivalentes.

6.1. Método por triangulación o eliminación Gaussiana

El método se basa en transformar la matriz dada, mediante operaciones elementales entre sus líneas, en una matriz triangular. A partir de esto, se aplica la P_6) de los determinantes: "El determinante de las matrices triangulares y diagonales es igual al producto de los elementos de la diagonal principal".

Por otra parte, hay que tener en cuenta que toda matriz cuadrada de orden n escalonada por filas (columnas) es triangular superior (inferior) y que, la traspuesta de una matriz triangular superior (inferior) es triangular inferior (superior).

Por lo tanto, dada una matriz $A \in K^{n \times n}$, usamos el método de eliminación de Gauss –pasos 1 a 4– para hallar la forma escalonada de A. Luego calculamos el determinante de esta última matriz.

Ejemplo: calcular el determinante de la matriz $A = \begin{pmatrix} -3 & 2 & -1 & -1 \\ 2 & -2 & 4 & 6 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{pmatrix}$ escalonando la matriz.

$$|A| = \begin{vmatrix} -3 & 2 & -1 & -1 \\ 2 & -2 & 4 & 6 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{vmatrix} = (-1)^{P_2)} \begin{vmatrix} 2 & -2 & 4 & 6 \\ -3 & 2 & -1 & -1 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{vmatrix} =$$

$$(-1)\,2 \overset{P_3)}{} \begin{vmatrix} 1 & -1 & 2 & 3 \\ -3 & 2 & -1 & -1 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{vmatrix}$$

$$|A| = (-2) \begin{matrix} 3F_1 + F_2 \\ -5F_1 + F_3 \\ 2F_1 + F_4 \end{matrix} \begin{vmatrix} 1 & -1 & 2 & 3 \\ 0 & -1 & 5 & 8 \\ 0 & 9 & -15 & -12 \\ 0 & 1 & 10 & 12 \end{vmatrix} =$$

$$(-2) \begin{matrix} \\ 9F_2 + F_3 \\ F_2 + F_4 \end{matrix} \begin{vmatrix} 1 & -1 & 2 & 3 \\ 0 & -1 & 5 & 8 \\ 0 & 0 & 30 & 60 \\ 0 & 0 & 15 & 20 \end{vmatrix}$$

$$|A| = (-2) \begin{matrix} \\ \\ -\frac{1}{2}F_3 + F_4 \end{matrix} \begin{vmatrix} 1 & -1 & 2 & 3 \\ 0 & -1 & 5 & 8 \\ 0 & 0 & 30 & 60 \\ 0 & 0 & 0 & -10 \end{vmatrix}$$

$$|A| = (-2)\ 1\ (-1)\ 30\ (-10)$$

$$|A| = -600$$

6.2. Método de Chío

Este método se basa en la consecuencia de la quinta propiedad de los determinantes: "El determinante de una matriz no varía si a una línea se le suma una combinación lineal de otras"; es decir, esta propiedad permite a los elementos de una línea sumar los correspondientes de otras paralelas multiplicadas por un escalar, de este modo el determinante de la nueva matriz es igual al de la matriz dada.

Basándonos en esta propiedad, se puede obtener un determinante igual al de la matriz dada, pero con los elementos de una fila o columna todos ceros, excepto uno. Se aplica el desarrollo de un determinante por los elementos de una línea y se reduce el orden del determinante en una unidad.

Este proceso se puede repetir tantas veces como se desee hasta obtener un determinante de orden tres o dos, es decir, la idea principal del método es ir obteniendo determinantes de un orden menor al dado.

Se parte de una matriz cuadrada de orden n que tenga un elemento $a_{ij} = 1$, pivote, (si no lo tuviere, se elige un elemento cualquiera y se extrae dicho elemento como factor común de toda la línea).

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & 1 & a_{in} \\ a_{n1} & a_{n2} & \dots & a_{nj} & a_{nn} \end{pmatrix}$$

Reducir a cero los elementos restantes de la fila del pivote. Es decir, transformar la matriz A en otra B mediante las siguientes operaciones elementales entre columnas:

A la 1° columna se le resta la j-ésima columna multiplicada por a_{i1}

A la 2° columna se le resta la j-ésima columna multiplicada por a_{i2}

A la 3° columna se le resta la j-ésima columna multiplicada por a_{i3}

Se continúa así hasta que:

A la n-ésima columna se le resta la j-ésima columna multiplicada por a_{in} con $n \neq j$

$$B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{ij} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & a_{2n} \\ \dots & \dots & \cdots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \\ a_{n1} & a_{n2} & \dots & a_{nj} & a_{nn} \end{pmatrix}$$

Luego $| B | = a_{ij}A_{ij}$, pero $a_{ij} = 1 \implies | B | = A_{ij}$

$| B | = (-1)^{i+j}M_{ij}$ pero por propiedades de los determinantes $| B | = |A |$, por lo tanto:

$| A | = (-1)^{i+j}M_{ij}$

Ejemplo: calcular el determinante de la matriz $A = \begin{pmatrix} -3 & 2 & -1 & -1 \\ 2 & -2 & 4 & 6 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{pmatrix}$ por el método de Chío.

$| A | = \begin{vmatrix} -3 & 2 & -1 & -1 \\ 2 & -2 & 4 & 6 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{vmatrix} = 2^{P_3)} \begin{vmatrix} -3 & 2 & -1 & -1 \\ 1 & -1 & 2 & 3 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{vmatrix}$, elegimos como pivote $a_{21} = 1$ es decir que $i = 2$ y $j = 1$

$$| A | = 2 \begin{vmatrix} -3 & 2-(-3)(-1) & -1-(-3)2 & -1-(-3)3 \\ 1 & -1-1(-1) & 2-1.2 & 3-1.3 \\ 5 & 4-5(-1) & -5-5.2 & 3-5.3 \\ -2 & 3-(-2)(-1) & 6-(-2)2 & 6-(-2).3 \end{vmatrix}$$

$$| A | = 2 \begin{vmatrix} -3 & -1 & 5 & 8 \\ 1 & 0 & 0 & 0 \\ 5 & 9 & -15 & -12 \\ -2 & 1 & 10 & 12 \end{vmatrix}$$

$$| A | = 2\,(-1)^3\,1 \begin{vmatrix} -1 & 5 & 8 \\ 9 & -15 & -12 \\ 1 & 10 & 12 \end{vmatrix}$$

$$| A | = -2\,(180 - 60 + 720 + 120 - 120 - 540)$$

$$|A| = -2 .300 = -600$$

Para hacer el proceso más directo, se suele aplicar esta regla de manera directa y de la siguiente manera.

Una vez logrado el pivote, hacemos:

$$|A| = \begin{vmatrix} -3 & 2 & -1 & -1 \\ 2 & -2 & 4 & 6 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{vmatrix} = 2^{P_3)} \begin{vmatrix} -3 & 2 & -1 & -1 \\ 1 & -1 & 2 & 3 \\ 5 & 4 & -5 & 3 \\ -2 & 3 & 6 & 6 \end{vmatrix}$$

A cada elemento que no esté ni en la fila ni en la columna del pivote (por ej. $a_{12} = 2$) los multiplicamos por el pivote ($a_{12} = 2 .1$) y le restamos el producto de los elementos que están en la otra diagonal del rectángulo que forman con un elemento de la fila y de la columna del pivote ($a_{22} = -1 \ y \ a_{11} = -3$), es decir $1.2 - (-3)(-1) = -1$.

Ordenadamente iremos obteniendo los elementos de un determinante de un orden menor al dado.

$$|A| = 2\ (-1)^3\ 1 \begin{vmatrix} -1 & 5 & 8 \\ 9 & -15 & -12 \\ 1 & 10 & 12 \end{vmatrix}$$

$$|A| = -600$$

7. Cálculo de la inversa a partir de la matriz adjunta

En este apartado se evidenciará, en primer lugar, la relación que existe entre el cálculo de los determinantes y la matriz inversa de una matriz cuadrada; en segundo lugar, se brindará y ejemplificará un segundo método para el cálculo de la matriz inversa.

Sea la matriz $A \in K^{n \times n}$ y su matriz adjunta, $Adj\ (A)$.

Consideremos, en primer lugar, el producto de estas dos matrices.

$$A\ \ Adj\ (A) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{pmatrix}$$

Para hacer este producto debemos tener presente:

a) La suma del producto de los elementos de una línea de A por sus respectivos cofactores es el determinante de A.

b) La suma del producto de los elementos de una línea de A por los cofactores de otra paralela a ella es nulo.

$$A\ \ Adj\ (A) = \begin{pmatrix} |\ A\ | & 0 & \cdots & 0 \\ 0 & |\ A\ | & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & |\ A\ | \end{pmatrix}$$

$A\ \ Adj\ (A) = |\ A\ |\ I_n$

Si $|\ A\ | \neq 0$ podemos trasponer términos y obtener

$A\ \ \frac{Adj\ (A)}{|\ A\ |}\ =\ I_n\ (1)$

Consideremos, en segundo lugar, el producto conmutado de A y su adjunta.

$Adj\ (A)\ A$ haciendo un desarrollo similar al anterior, obtenemos

$\frac{Adj\ (A)}{|\ A\ |}\ A\ =\ I_n\ (2)$

De (1) y (2) por la definición de matriz inversa, obtenemos

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

En la P8 se demostró que si la matriz A es invertible, implica que $|A| \neq 0$ de lo que se concluye que si $|A| = 0$ la matriz no es invertible.

En este apartado se demostró que si $|A| \neq 0$, existe la matriz inversa de A, dada por $A^{-1} = \frac{Adj\,(A)}{|A|}$.

Por lo tanto, podemos concluir que una matriz invertible es equivalente a decir que esa matriz tiene un determinante no nulo. Es decir:

"La condición necesaria y suficiente para que una matriz cuadrada de orden n admita inversa es que su determinante sea no nulo".

Luego, dada la matriz $A \in K^{n \times n}$, para hallar su inversa utilizando su adjunta, conviene:

a) Hallar el determinante asociado a la matriz A; si este es distinto de cero, A admite inversa.

b) Hallar la matriz de los cofactores.

c) Hallar la traspuesta de la matriz de los cofactores.

d) Multiplicar cada elemento de esta última matriz por $\frac{1}{|A|}$

e) Verificar el resultado obtenido.

Ejemplo: calcular la matriz inversa de la matriz A, usando su adjunta.

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

a) $|A| = \begin{vmatrix} 1 & 0 & -1 \\ 2 & -1 & 1 \\ 1 & 2 & -2 \end{vmatrix} = -5$ por lo tanto A admite inversa.

b) $A_{11} = (-1)^2 \begin{vmatrix} -1 & 1 \\ 2 & -2 \end{vmatrix} = 0$; $A_{12} = (-1)^3 \begin{vmatrix} 2 & 1 \\ 1 & -2 \end{vmatrix} = 5$; y finalmente

$A_{13} = (-1)^4 \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = 5$

$A_{21} = (-1)^3 \begin{vmatrix} 0 & -1 \\ 2 & -2 \end{vmatrix} = -2;\ A_{22} = (-1)^4 \begin{vmatrix} 1 & -1 \\ 1 & -2 \end{vmatrix} = -1;\ A_{23} =$

$(-1)^5 \begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = -2$

$A_{31} = (-1)^4 \begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix} = -1;\ \ A_{32} = (-1)^5 \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = -3$; y, por último

$A_{33} = (-1)^6 \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} = -1$

Por lo tanto, la matriz de los cofactores es:

$Cof\ (A) = \begin{pmatrix} 0 & 5 & 5 \\ -2 & -1 & -2 \\ -1 & -3 & -1 \end{pmatrix}$ y, en consecuencia, la matriz adjunta es:

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

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

$A^{-1} = -\frac{1}{5} \begin{pmatrix} 0 & -2 & -1 \\ 5 & -1 & -3 \\ 5 & -2 & -1 \end{pmatrix}$

$A^{-1} = \begin{pmatrix} 0 & \frac{2}{5} & \frac{1}{5} \\ -1 & \frac{1}{5} & \frac{3}{5} \\ -1 & \frac{2}{5} & \frac{1}{5} \end{pmatrix}$

e) Para verificar hacemos: $A\,.\,A^{-1} = I$

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

8. Determinante, inversa, rango y equivalencia a la identidad

Simplemente evidenciaremos y ejemplificaremos la relación que existe entre los cuatro temas ya estudiados y que se enuncian en el título del presente apartado.

Partiendo de $A \in K^{n \times n}$ podemos afirmar que:

$|A| \neq 0 \iff A \text{ es invertible} \iff \rho(A) = n \iff A \sim I_n \iff A$ es producto de matrices elementales.

Ejemplo 1: determinar si las siguientes matrices son invertibles.

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

$$|A| = \begin{vmatrix} 1 & 0 & 2 & 0 \\ 1 & 2 & 0 & 1 \\ -1 & 1 & 4 & -1 \\ 3 & -1 & -3 & -2 \end{vmatrix} = \begin{matrix} -F_1 + F_2 \\ F_1 + F_3 \\ -3F_1 + F_4 \end{matrix} \begin{vmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -2 & 1 \\ 0 & 1 & 6 & -1 \\ 0 & -1 & -9 & -2 \end{vmatrix} =$$

$$\begin{matrix} \\ -\frac{1}{2}F_2 + F_3 \\ \frac{1}{2}F_2 + F_4 \end{matrix} \begin{vmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -2 & 1 \\ 0 & 0 & 7 & -\frac{3}{2} \\ 0 & 0 & -10 & -\frac{3}{2} \end{vmatrix}$$

$$|A| = \begin{matrix} \\ \\ \frac{10}{7}F_3 + F_4 \end{matrix} \begin{vmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -2 & 1 \\ 0 & 0 & 7 & -\frac{3}{2} \\ 0 & 0 & 0 & -\frac{51}{14} \end{vmatrix} = -51$$

Rta1: A es invertible pues $|A| \neq 0$

Rta2: A es invertible pues $\rho(A) = 4$, ya que tiene 4 columnas pivotales.

Rta3: A es invertible, por tener cuatro columnas pivotales, A es equivalente por filas a I_4 (solo hace falta "convertir" los pivotes y realizar la

eliminación de Gauss – Jordan), lo cual indicaría también que la matriz admite inversa.

b) $A = \begin{pmatrix} 1 & 5 & 4 \\ 2 & 4 & 2 \\ 0 & -2 & -2 \end{pmatrix}$

$$|A| = \begin{vmatrix} 1 & 5 & 4 \\ 2 & 4 & 2 \\ 0 & -2 & -2 \end{vmatrix} \underset{}{=} \underset{-2F_1 + F_2}{} \begin{vmatrix} 1 & 5 & 4 \\ 0 & -6 & -6 \\ 0 & -2 & -2 \end{vmatrix} = -6 \begin{vmatrix} 1 & 5 & 4 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{vmatrix} =$$

$$\underset{2F_2 + F_3}{-6} \begin{vmatrix} 1 & 5 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0$$

Rta1: A es no invertible, pues $|A| = 0$

Rta2: A no es invertible, pues $\rho(A) = 2$ ya que tienen 2 columnas pivotales.

Rta3: A no es invertible, por tener 2 columnas pivotales, A no es equivalente por filas a la matriz identidad de orden 3, I_3.

Ejemplo 2: determinar el valor de k para que $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & k & 2 \\ 2 & 3 & 0 \end{pmatrix}$ admita inversa.

$$|A| = \begin{vmatrix} 1 & 2 & 1 \\ 2 & k & 2 \\ 2 & 3 & 0 \end{vmatrix} \underset{-2F_1 + F_3}{\overset{-2F_1 + F_2}{=}} \begin{vmatrix} 1 & 2 & 1 \\ 0 & -4+k & 0 \\ 0 & -1 & -2 \end{vmatrix} = -1 \begin{vmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & -4+k & 0 \end{vmatrix} =$$

$$-1(-1) \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & -4+k & 0 \end{vmatrix}$$

$$|A| = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & -4+k & 0 \end{vmatrix} \underset{(4-k)F_2 + F_3}{=} \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2(4-k) \end{vmatrix} = 2(k-4)$$

Nótese que la operación $(4-k)F_2 + F_3$ puede hacerse cualquiera sea el valor de k.

Si $2(k-4) = 0 \implies k = 4$, por lo tanto, si $k = 4$, obtenemos:

$|A| = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{vmatrix} = 0$; por lo tanto, si $k = 4$, A no admite inversa y podemos afirmar, además que su rango es 2, menor que n, pues solamente tiene dos filas no nulas o dos columnas pivotales.

Si $2(k-4) \neq 0 \implies k \neq 4$, por lo tanto, si $k \neq 4$, obtenemos:

$$|A| = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2(4-k) \end{vmatrix} = \frac{1}{2(4-k)} F_3 \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix}$$

Nótese que la operación $\frac{1}{2(4-k)} F_3$ puede realizarse pues $k \neq 4$.

Por lo tanto, $|A| \neq 0$ y A admite inversa. También podemos afirmar que $\rho(A) = 3$ y A es equivalente a la matriz I_3.

9. Determinantes y rango

Veremos cómo se pueden utilizar los determinantes para calcular el rango de una matriz de cualquier orden.

Dada una matriz $A \in K^{m \times n}$ $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$, podemos definir:

Menor de orden p de A, con $p \leq m$ y $p \leq n$, al determinante de la submatriz cuadrada de orden p cuyos elementos están situados en la intersección de las p filas, $i_1, \dots, i_p$ y de las p columnas, $j_1, \dots, j_p$ de A. Cabe

destacar que la elección de distintos conjuntos de filas y columnas –siempre en orden natural– dará lugar a los distintos menores. Luego, los menores no son submatrices sino los determinantes de esas submatrices cuadradas.

Por ejemplo: sea $A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$, algunos de sus menores son:

M_1 de orden 2 que toma las filas 1 y 2 y las columnas 1 y 4, $M_1 = \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = 2$

M_2 de orden 2 que toma las filas 1 y 3 y las columnas 2 y 3, $M_2 = \begin{vmatrix} 2 & 0 \\ 0 & 0 \end{vmatrix} = 0$

M_3 de orden 3 que toma las columnas 2, 3 y 4; $M_3 = \begin{vmatrix} 2 & 0 & 1 \\ -1 & 1 & 2 \\ 0 & 0 & 3 \end{vmatrix} = 6$

El número de menores de orden 2 para la matriz A es 18 pues existen tres elecciones para el par de filas ($C_{3,2} = 3 \implies F_1\,F_2;\ F_1F_3\ y\ F_2F_3$) y 6 para el par de columnas ($C_{4,2} = 6 \implies C_1C_2;\ C_1C_3;\ C_1C_4;\ C_2C_3;\ C_2C_4\ y\ C_3C_4$)

El número de menores de orden 3 para la matriz A es 4 pues existe una sola elección para las filas ($C_{3,3} = 1 \implies F_1F_2F_3$) y 4 elecciones para la terna de columnas ($C_{4,3} = 4 \implies C_1C_2C_3;\ C_1C_2C_4;\ C_1C_3C_4;\ C_2C_3C_4$).

Respecto al rango de una matriz y submatrices se establece la siguiente relación:

Si $[A]_{I \times J}$ es una submatriz de A, entonces $\rho\big([A]_{I \times J}\big) \leq \rho(A)$

A partir de la definición de menor y la relación establecida anteriormente, podemos evidenciar la utilidad de los menores para calcular el rango de una matriz cualquiera.

Definición: *El rango de una matriz A es igual al mayor orden r alcanzado por los menores no nulos de A, es decir, $\rho(A) = r$ si y solo si:*

a) Existe un menor $m_{I,J} \neq 0$ de orden r

b) Todos los menores de A de orden mayor a r son nulos.

En un principio, para ver que una matriz A tiene rango r habría que encontrar un menor, M, no nulo de orden r y calcular todos los menores de orden $r + 1$ que corresponden a submatrices de A que contienen a M, si vemos que se anulan todos esos menores, entonces $\rho(A) = r$

Teorema del método del orlado

Sea la matriz $A = \|a_{ij}\|_{m \times n}$ con una submatriz cuadrada M de orden r, de tal modo que $|M| \neq 0$. Supongamos que todas las submatrices cuadradas de orden $r + 1$ que contienen a M tienen determinante nulo, entonces $\rho(A) = r$

9.1. Método del orlado[32] para calcular el rango de una matriz

Sea la matriz $A = \|a_{ij}\|_{m \times n}$, el método del orlado para calcular el rango r, es:

[32] Orlar: si consideramos un menor de orden p de una matriz A y el menor de orden (p+1) resultante de añadirle la fila i y la columna j no pertenecientes a él, entonces se dice que se ha construido "orlando" dicho menor con la fila i y la columna j.

Si $A = N$, entonces $\rho(A) = 0$, caso contrario tomamos un menor M de orden 1 no nulo. Comenzando por $r = 1$, se realizan los siguientes pasos:

a) Hallar, si existe, un menor no nulo de orden $r + 1$ que contenga a M

b) Si no existe, entonces $\rho(A) = r$

c) Si existe repetir el paso a) incrementando a r en una unidad y reemplazando M por el menor encontrado.

Obviamente, no es necesario comenzar con $r = 1$. Si se encuentra un menor no nulo de orden r, se puede empezar con dicho r en el paso a).

Ejemplo: calcular el rango de las siguientes matrices:

a) $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 0 \end{pmatrix}$

$M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \Rightarrow |M| = 1 \neq 0$

$\rho(A) \leq 2$

$\therefore\ \rho(A) = 2$

b) $B = \begin{pmatrix} 3 & -1 & 0 & 2 \\ 4 & 0 & 3 & -1 \\ -1 & -1 & -3 & 3 \end{pmatrix}$

$M = \begin{pmatrix} 3 & -1 \\ 4 & 0 \end{pmatrix} \Rightarrow |M| = 4 \neq 0$

$\rho(B) \leq 3$

$M_1 = \begin{pmatrix} 3 & -1 & 0 \\ 4 & 0 & 3 \\ -1 & -1 & -3 \end{pmatrix} \Rightarrow |M_1| = 0$

$M_2 = \begin{pmatrix} 3 & -1 & 2 \\ 4 & 0 & -1 \\ -1 & -1 & 3 \end{pmatrix} \Rightarrow |M_2| = 0$

$M_3 = \begin{pmatrix} 3 & 0 & 2 \\ 4 & 3 & -1 \\ -1 & -3 & 3 \end{pmatrix} \Rightarrow |M_3| = 0$

$M_4 = \begin{pmatrix} -1 & 0 & 2 \\ 0 & 3 & -1 \\ -1 & -3 & 3 \end{pmatrix} \Rightarrow |M_4| = 0$

Si bien $M_3\ y\ M_4$ no contienen a M, se calculó sus menores para evidenciar la cantidad de submatrices de orden tres que se pueden obtener a partir

de B. Es decir, el cálculo para determinar el rango de la matriz puede terminar al calcular M_2.

$\therefore\ \rho(B) = 2$

c) $C = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix}$

$\rho(C) \leq 4$

$M = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \Longrightarrow |M| = 0$

$M = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix} \Longrightarrow |M| = 2 \neq 0$

$M_1 = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Longrightarrow |M_1| = 0$

$M_2 = \begin{pmatrix} 2 & -1 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \Longrightarrow |M_2| = 4 \neq 0$

$M_3 = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 \end{pmatrix} = 0$

$M_4 = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} = 0$

$M_5 = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 2 & 1 & 0 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} = 0$

$M_6 = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} = 0$

$$M_7 = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} = 0$$

Vale la misma observación respecto de $M_2 \ y \ M_5$, M_6, M_7

$\therefore \ \rho(C) = 3$

Finalmente, con esta nueva definición de rango de una matriz y las propiedades de los determinantes se puede demostrar el siguiente:

Teorema: *El rango de una matriz no se altera si:*

i) intercambiamos entre sí dos líneas,

ii) trasponemos una fila (F_i) *con una columna* (C_i)*,*

iii) multiplicamos los elementos de una línea por un escalar no nulo,

iv) sumamos a una línea los elementos de otra línea multiplicados por un escalar no nulo.

CAPÍTULO 6

SISTEMAS DE ECUACIONES LINEALES

1. Introducción

Los sistemas de ecuaciones lineales son una poderosa herramienta matemática para modelar problemas como la distribución de cosechas o el presupuesto de un país, el cálculo de la órbita de un planeta, el cálculo de la estabilidad estructural de un edificio en ingeniería civil, entre otros muchos. De las muy variadas y diversas aplicaciones se puede mencionar uno de los paradigmas económicos más empleados: el modelo input–output de Wassily Leontief, quien, en 1973, recibió el Premio Nobel en Ciencias Económicas. Si bien la teoría se le atribuye a Leontief, algunos historiadores afirman que el procedimiento fue descrito en 1758 por el economista francés François Quesnay, quien desarrolló una versión más rudimentaria. Karl Marx fue el primero en traducir la obra de Quesnay a un sistema matricial de ecuaciones, en los llamados modelos de reproducción simple y reproducción ampliada (*El Capital,* vol. II).

Otra de las obras célebres de la teoría del álgebra lineal fue propuesta por George Bernard Dantzig (1914–2005), quien, como consejero matemático de las Fuerzas Áreas estadounidenses, presentó por primera vez, en el año 1947, un problema de programación lineal y propuso el Método Simplex para resolverlo haciendo uso de los sistemas de ecuaciones lineales.

La búsqueda de soluciones tanto a las ecuaciones algebraicas, como a los sistemas de ecuaciones lineales, es tan importante como antigua. En el segundo caso, aparecieron incluso antes del estudio de las matrices y determinantes.

El médico y matemático italiano Girolamo (Gerolamo) Cardano (1501–1576) se destacó por sus trabajos en álgebra. En 1945 publicó su obra *Ars Magna*, donde expuso las soluciones de las ecuaciones de tercer y cuarto grado y muestra una regla para resolver sistema de dos ecuaciones lineales con dos incógnitas que llama "regula de modo" y que en esencia es la conocida regla de Cramer.

Francois Viète (1540–1603) fue el primero en representar los parámetros de una ecuación mediante letras. Esto marcó el inicio de una nueva etapa, en la cual René Descartes (1596–1650) aportó de forma destacada al desarrollo de dicha notación. A partir de allí, el álgebra se convierte en la ciencia de los cálculos simbólicos y de las ecuaciones.

Tiempo después, Euler la define como la teoría de los "cálculos con cantidades de distintas clases", mientras que Leibniz, al estudiar los sistemas de ecuaciones lineales y no disponer de una notación matricial, representaba los coeficientes de las incógnitas mediante una pareja de subíndices. Dos años luego de su muerte, se publicó la obra de Colin Maclaurin (1698–1746), *Tratado de Álgebra*, en donde aparecieron los primeros resultados sobre determinantes y se probó la Regla de Cramer para sistemas hasta tres ecuaciones con tres incógnitas; mientras que el

propio Gabriel Cramer (1704–1752) anunció la regla general para sistemas de n ecuaciones con n incógnitas en su obra *Introducción al análisis de líneas algebraicas* en 1750.

Más adelante, en 1764, Etienne Bézout (1730–1783) expuso nuevos métodos para calcular determinantes al igual que Alexandre Théophile Vandermonde (1735–1796), pero en el año 1771. Un año más tarde, Pierre Simon Laplace (1749–1827) criticó tanto a Cramer como a Bézout tildando sus métodos como imprácticos y describió un método para resolver sistema de ecuaciones lineales sin tener que calcular, de manera explícita, determinantes.

De esta forma, el álgebra lineal tuvo un fuerte empuje gracias al estudio de los sistemas de ecuaciones lineales, pero hasta el siglo XVIII era una disciplina dedicada a resolver ecuaciones de grado arbitrario. Hasta que el matemático y filósofo francés Jean le Rond D'Alembert (1717–1783) descubrió que las soluciones de un sistema matricial de la forma $A\,X = B$, cumplen ciertas propiedades, dando inicio así al desarrollo de un álgebra más teórica. Este mismo autor, junto con Euler y Lagrange, se dieron cuenta de que la solución del sistema homogéneo $A.X = N$, donde N es la matriz nula, es una combinación lineal de algunas soluciones particulares.

En este capítulo, luego de ver los conceptos básicos para poder manipular y entender sistemas de ecuaciones lineales, se estudian dos aspectos básicos en este tema: análisis del sistema para determinar su

compatibilidad y la resolución por medio de distintos métodos. Dedicamos un apartado para los sistemas homogéneos asociados a los no homogéneos y a aquellos en los cuales el número de incógnitas es igual al número de ecuaciones. Finalmente se ejemplifica, para poder afianzar la clasificación de los sistemas, aquellos en donde intervienen parámetros y se establece la relación que existe entre la solución o no de un sistema con las características de las matrices del sistema y la correspondiente matriz ampliada.

1.1. Conceptos básicos

El hombre siempre recurre a la matemática para poder analizar y entender la realidad, ya que esta disciplina le permite modelar, con distintas herramientas, fenómenos químicos, físicos o biológicos, entre otros tantos.

Una de las herramientas que permiten resolver problemas concretos o que ayudan a encontrar una solución a problemas específicos, son las ecuaciones algebraicas. Recordemos que la palabra ecuación proviene del latín "*aequatio, aequationis*" que significa "*nivelación, igualación o repartición igual de algo*", entonces una ecuación es una igualdad que contiene algunas cantidades desconocidas (variables o incógnitas) y que se distinguen de las identidades que también son igualdades pero que se verifican para cualquier valor que se asignen a sus variables. Ejemplos de identidades ya conocidas son: $a^2 - b^2 = (a + b)(a - b)$ o $(a + b)^2 = a^2 + 2ab + b^2$.

Las siguientes expresiones representan ecuaciones:

$a)\ 2x - 5 = 3$ $\qquad$ $e)\ x - 3xy = 7$

$b)\ -2x + y = 5$ $\qquad$ $f)\ (x + y)(3x - y)^{-1} = x$

$c)\ x - 2y + 3z = 2$ $\qquad$ $g)\ log\ (2x - 1) = 4$

$d)\ 2x^2 + 3x = 0$ $\qquad$ $h)\ 2\ sen\ x = 1$

En general, el estudio de las ecuaciones dependerá del tipo que sean estas; en nuestro caso, restringiremos el análisis y estudio al caso de ecuaciones lineales, de allí la siguiente:

Definición: *Se llama ecuación lineal a una igualdad donde las variables que intervienen solo tienen constantes y potencia 1. Si en la ecuación intervienen n variables, entonces se dice que la ecuación es lineal con n variables.*

En los ejemplos anteriores, los tres primeros representan ecuaciones lineales, el inciso $a)$ de una sola variable, mientras que $b)\ y\ c)$ son de dos y tres variables, respectivamente.

Los incisos $d), e)\ y\ f)$ no son ecuaciones lineales pues en ellos intervienen potencias mayores a 1 o bien producto o cocientes entre sus variables. En los restantes incisos interviene una función logarítmica o trigonométrica.

El estudio de ecuaciones lineales obedece al hecho de que muchos problemas en múltiples disciplinas tienen un carácter lineal y puede resolverse este tipo de ecuaciones mediante algoritmos que siempre dan una respuesta, más allá de la cantidad de ecuaciones con las que se trabaje y de las incógnitas que estas tengan.

El caso más sencillo de un sistema de ecuaciones lineales es aquel que posee una sola ecuación con una sola incógnita "x". Por ejemplo:

$ax = b$ con $a\ y\ b\ \in \mathbb{R}$. La solución de esta ecuación –de existir– dependerá de a.

i) Si a posee inverso multiplicativo, es decir $a \neq 0$, entonces la ecuación tiene solución única. $x = a^{-1}b$ para cualquier valor de b.

ii) Si a no posee inverso multiplicativo, es decir $a = 0$, entonces la solución de la ecuación dependerá de b.

ii$_1$) Si $b = 0$, entonces la ecuación tiene infinitas soluciones ya que cualquier $x \in \mathbb{R}$ satisface la ecuación.

ii$_2$) Si $b \neq 0$, entonces la ecuación no tiene solución.

El caso más general está dado por la siguiente:

Definición: *Una ecuación lineal con n incógnitas* $x_1, x_2, \dots, x_n$*, es una ecuación que tiene la forma* $a_1x_1 + a_2x_2 + \dots + a_nx_n = b$ *o de manera reducida* $\sum_{i=1}^{n} a_ix_i = b$ *donde* $n \in \mathbb{Z}^+$*, los coeficientes* a_i *y el término* b *son elementos de un cuerpo* $\mathbb{K}$[33].

Los incisos $b)\ y\ c)$ del ejemplo anterior pueden ser reescritos de la siguiente manera.

[33] Un cuerpo o campo es un conjunto K con dos operaciones binarias, usualmente llamadas suma "+" y producto ".". La suma es asociativa, conmutativa y su elemento neutro es 0, además para cada elemento existe un opuesto. El producto es asociativo, conmutativo, su elemento neutro es 1 y todo elemento distinto de cero tiene inverso, además el producto se distribuye respecto de la suma.

$b) -2x_1 + x_2 = 5$ y $c) x_1 - 2x_2 + 3x_3 = 2$

Solución: *de la ecuación lineal de n incógnitas, $\sum_{i=1}^{n} a_i x_i = b$, es la n-tupla $\alpha_1, \alpha_2, \dots, \alpha_n$ que la satisfacen, es decir verifican $\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n = b$.*

Otro sistema, seguramente ya estudiado, es por ejemplo $\begin{cases} 2x - y = 4 \\ x - y = 1 \end{cases}$. Se trata de un sistema lineal de dos ecuaciones con dos incógnitas. Particularmente, este sistema puede verse desde dos perspectivas.

a) Desde el punto de vista *geométrico:* cada una de las dos ecuaciones representa una recta en el plano cartesiano. Resolver este sistema consiste en encontrar, si existen, los puntos de intersección de ambas rectas. Es por esta razón que estos sistemas se llaman lineales.

b) Desde el punto de vista *algebraico:* el problema consiste en hallar las coordenadas cartesianas de los puntos que satisfagan simultáneamente ambas ecuaciones. Como ya se dijo, este tipo de ecuaciones son lineales porque cada término (excepto los independientes) es de grado 1.

Si nos quedamos en el marco algebraico, nada nos impide generalizar el concepto de sistema lineal a más de dos ecuaciones con más de dos incógnitas. Así, tenemos la siguiente:

Definición: *Un sistema de m ecuaciones lineales con n incógnitas $x_1, x_2, \dots, x_n$, es un conjunto de ecuaciones lineales, de la forma:*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ \ldots + a_{2n}x_n = b_2 \\ \quad \vdots \qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ \ldots + a_{mn}x_n = b_m \end{cases}$$ *o de manera resumida:* $\sum_{j=1}^{n} a_{ij}\, x_j = b_i$,

$con\ i = 1, 2, \ \ldots, m.$

Donde los a_{ij} *y* b_i *son elementos de un cuerpo* $\mathbb{K}$.

Se han asignado dos subíndices a los coeficientes de cada una de las incógnitas para que queden unívocamente localizados. El primer subíndice remite a la ecuación a la que pertenece (fila), mientras que el segundo hace referencia a la incógnita de la cual es coeficiente (columna). Por otro lado, esta doble indexación será de gran utilidad a la hora de construir matrices para el análisis y resolución de los sistemas de ecuaciones lineales.

Según sea la cantidad de ecuaciones (m) y de incógnitas (n), suele llamarse al sistema cuadrado si $m = n$ o bien rectangular si $m \neq n$. En este último caso, si el sistema tiene más ecuaciones que incógnitas, $m > n$, algunos autores suelen denominarlos sobredeterminados, mientras que, si tiene menos ecuaciones que incógnitas, $m < n$, se habla de sistemas subdeterminados.

Solución: *de un sistema lineal de m ecuaciones con* n *incógnitas, expresado como* $\sum_{j=1}^{n} a_{ij}\, x_j = b_i \ \ con\ i = 1, 2, \ \ldots, m$, *es la n-tupla* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *que satisface simultáneamente las* m *ecuaciones.*

Puede ocurrir que uno o más sistemas lineales tengan el mismo conjunto solución. En ese caso podemos establecer la siguiente:

Definición: *Dos o más sistemas lineales con el mismo conjunto de incógnitas, se dicen equivalentes si y solo si tiene el mismo conjunto solución.*

Ejemplo: sean los sistemas $a)\begin{cases}2x_1+2x_2=6\\2x_1+3x_2=8\end{cases}$; $b)\begin{cases}x_1-x_2+x_3=3\\2x_1+x_2-x_3=6\end{cases}$;

$c)\begin{cases}x_1+x_2=2\\x_1-x_2=4\\x_1+2x_2=0\end{cases}$

En $a)$, sistema cuadrado, puede verificarse que el par ordenado $(x_1,x_2)=(1,2)$ satisface ambas ecuaciones simultáneamente, por lo tanto, es solución. También puede verificarse que es la *única solución*.

En $b)$, sistema rectangular con menos ecuaciones que incógnitas, la terna ordenada $(x_1,x_2,x_3)=(3,0,0)$ satisface ambas ecuaciones, pero también lo hace la terna $(x_1,x_2,x_3)=(3,\alpha,\alpha)$ con $\alpha\in\mathbb{R}$. En este caso existen *infinitas soluciones*, ya que hay infinitas ternas (3–tuplas) ordenadas que satisfacen las dos ecuaciones.

En $c)$, sistema rectangular con más ecuaciones que incógnitas, *no existe solución,* ya que al sumar miembro a miembro la primera ecuación con la segunda obtenemos $2x_1=6$, lo que implica que $x_1=3$. Reemplazando este valor en la primera o segunda ecuación, obtenemos $x_2=-1$. Pero estos dos valores, $x_1=3\ \ y\ \ x_2=-1$, no satisfacen la tercera ecuación, por lo tanto, no existe un par ordenado que satisfaga simultáneamente estas tres ecuaciones.

Debemos recalcar que los tres sistemas anteriores son simplemente ejemplos esclarecedores de lo que estamos exponiendo, no debe inferirse

de ellos que todo sistema cuadrado siempre posee una única solución (compatible determinado), que todo sistema con menos ecuaciones que incógnitas siempre tiene infinitas soluciones (compatible indeterminado) y que, finalmente, todo sistema con más ecuaciones que incógnitas nunca tiene solución (incompatible).

Se demostrará, en los siguientes apartados, que todo sistema de ecuaciones lineales puede tener o no tener solución. En caso de tenerla (compatible), esta a su vez puede ser única (compatible determinado) o infinitas (compatible indeterminado). Si el sistema no posee solución, se dice incompatible. Esto nos lleva a la siguiente:

Definición: *Un sistema de m ecuaciones lineales con n incógnitas con al menos una solución, se denomina compatible; si es única se denomina sistema compatible determinado, y si existen infinitas soluciones se lo denomina sistema compatible indeterminado. Un sistema sin solución se denomina sistema incompatible. El conjunto de todas las soluciones de un sistema suele llamársele conjunto solución.*

En resumidas palabras:

$$Sistema\ de\ ecuaciones\ lineales \begin{cases} Compatible \begin{cases} Determinado \\ Indeterminado \end{cases} \\ Incompatible \end{cases}$$

Recurramos ahora a la perspectiva geométrica y analicemos algunos ejemplos de sistemas que ya sabemos resolver.

Una ecuación lineal $ax + by = c$ con dos incógnitas $x\ e\ y$, se interpreta geométricamente como ecuación cartesiana de una *recta* en $\mathbb{R}^2$. Por lo

tanto, resolver un sistema de m ecuaciones, con $m = 2 \vee m = 3$, y con dos incógnitas, $n = 2$, significa encontrar pares ordenados, (x, y), que satisfagan todas las ecuaciones; lo que es equivalente a analizar si dos o tres rectas en el plano se intersecan en un punto (compatible determinado), si son coincidentes (compatible indeterminado) o bien si no tienen ningún punto en común (incompatible). Por ejemplo:

a) $\begin{cases} 2x + y = 5 \\ x + 2y = 1 \end{cases}$

Para resolver este sistema usaremos el método de reducción por sumas y restas, ya que, en primer lugar, dicho procedimiento es ya conocido y, en segundo lugar, se generalizará en los apartados correspondientes a la resolución de sistemas lineales.

Si a la segunda ecuación la multiplicamos por el escalar $\lambda = -2$ y luego sumamos ambas ecuaciones, obtenemos $-3y = 3$, por lo tanto $y = -1$, con lo que $x = 3$.

Solución única: $(x, y) = (3, -1)$. El sistema compatible determinado. Geométricamente, las dos ecuaciones ($m = 2$) corresponden a dos rectas que se intersecan en el punto $(3, -1)$.

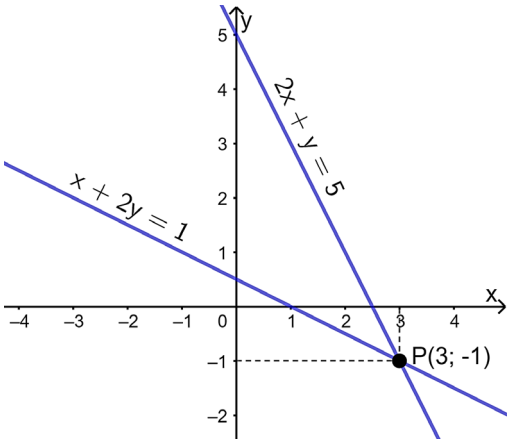


Figura 6.1. Sistema compatible determinado.

b) $\begin{cases} x + y = 2 \\ -x - y = 0 \end{cases}$

Como observamos que los coeficientes de las incógnitas son opuestos, procedemos a sumar ambas ecuaciones, obteniendo $0 = 2$, lo que es algebraicamente una incoherencia. Pero podemos colocar cada ecuación del sistema en su forma explícita.

$\begin{cases} y = -x + 2 \\ y = -x \end{cases}$ y podemos observar que ambas ecuaciones tienen la misma pendiente ($p = -1$) y distintas ordenadas al origen, lo que significa que son paralelas. Esto se muestra en la Figura 6.2.

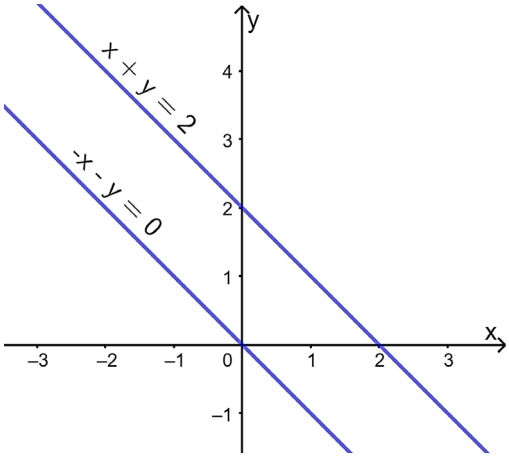


Figura 6.2. Sistema incompatible.

c) $\begin{cases} x + y = 3 \\ 2x + 2y = 6 \end{cases}$

Podemos observar, o bien que la segunda ecuación se obtuvo multiplicando la primera por el escalar $\lambda = 2$, con lo que concluimos que se trata de la misma ecuación, o bien aplicar el procedimiento indicado anteriormente y multiplicar la primera ecuación por el escalar $\lambda = -2$, sumar ambas ecuaciones y obtener $0 = 0$, expresión tautológica (siempre es verdadera). En ambos casos se llega a la misma conclusión, que se trata en realidad de la misma ecuación $x + y = 3$, o sea que $y = 3 - x$. Por lo tanto, cualquier par ordenado de la forma $(x, 3 - x)$ es solución del sistema. El sistema es compatible indeterminado.

Geométricamente, ambas ecuaciones representan la misma recta y las infinitas soluciones son todos los puntos de la forma $(x, 3 - x)$, que son los puntos de la recta.

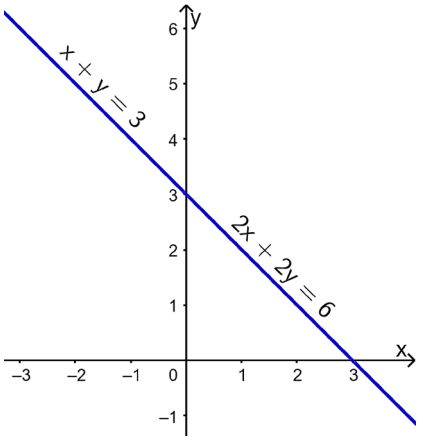


Figura 6.3. Sistema compatible indeterminado.

Similar análisis puede hacerse cuando $m = 3$, es decir tres ecuaciones con dos incógnitas.

d) $\begin{cases} 2x + y = 5 \\ x - y = 4 \\ x + 3y = 0 \end{cases}$

Sistema compatible determinado cuya solución es $(x, y) = (3, -1)$. Geométricamente, son tres rectas que tienen un punto en común, tal como se muestra en la Figura 6.4.

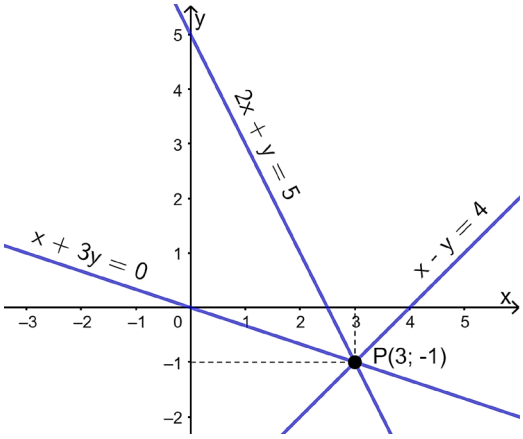


Figura 6.4. Sistema compatible determinado.

e) $\begin{cases} x + y = 3 \\ x - y = 2 \\ x + 3y = 0 \end{cases}$

El sistema es incompatible. Geométricamente, son tres rectas que no tienen ningún punto en común, tal como se muestra en la Figura 6.5. Obsérvese que, si las rectas se toman de a pares, estas tienen un punto en común, pero las tres juntas, no.

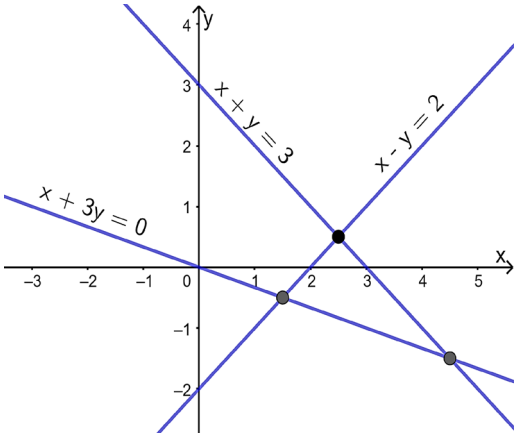


Figura 6.5. Sistema incompatible.

2. Distintas formas de expresar un sistema de ecuaciones lineales

Veremos, a continuación, distintas formas en las que se puede expresar un mismo sistema de ecuaciones lineales, destacando que no son las únicas, sino aquellas a las que recurriremos en el desarrollo del presente capítulo.

Un sistema de m ecuaciones lineales con n incógnitas puede expresarse como una única ecuación matricial $A.X = B$.

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

Dado el sistema de ecuaciones lineales

$$\begin{cases} a_{11}x_1 + & a_{12}x_2 + & \dots & +a_{1n}x_n = & b_1 \\ a_{21}x_1 + & a_{22}x_2 + & \dots & +a_{2n}x_n = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \dots & +a_{mn}x_n = & b_m \end{cases}$$

Se pueden construir tres matrices:

a) La matriz de los coeficientes, A, matriz de orden $(m \times n)$ y cuyos elementos son los coeficientes de las incógnitas.

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

b) La matriz de las incógnitas, X, matriz de orden $(n \times 1)$ y cuyos elementos son las n incógnitas.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

c) La matriz de los términos independientes, B, matriz de orden $(m \times 1)$ y cuyos elementos son los términos independientes.

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Como A es de orden $(m \times n)$ y X es de orden $(n \times 1)$, el producto, $A.X$, es posible y da por resultado una matriz de orden $(m \times 1)$. Entonces es posible escribir $A.X = B$.

Además de estas tres matrices es posible construir una cuarta matriz, llamada matriz ampliada, A' o $(A|B)$, que es una matriz de orden $[m \times (n+1)]$ y sus columnas son las columnas de la matriz A más la columna de los términos independientes.

$$A' = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

Otra forma de expresar un sistema de ecuaciones lineal es de manera vectorial.

Con los coeficientes de las incógnitas y los términos independientes, se pueden construir vectores columnas.

Por lo tanto, el sistema de ecuaciones lineales se puede expresar de la siguiente manera:

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \quad \ldots \quad \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Llamada forma vectorial del sistema de ecuaciones lineales; es decir:

$x_1 \overrightarrow{c_1} + x_2 \overrightarrow{c_2} + \ \ldots + x_n \overrightarrow{c_n} = \overrightarrow{b}$, en donde $\overrightarrow{c_i}$ es el vector columna de posición i de la matriz A, y $\overrightarrow{b}$ es el vector de $\mathbb{R}^m$ cuyas componentes son los términos independientes del sistema.

De esta manera, las soluciones del sistema son simplemente los escalares que lleven a que el vector $\overrightarrow{b}$ sea una combinación lineal de los vectores columnas $\overrightarrow{c_1}, \ \overrightarrow{c_2}, \ \ldots, \overrightarrow{c_n}$.

Finalmente, un sistema puede ser visto como una transformación lineal. Las transformaciones lineales son funciones especiales y se estudiarán en capítulos posteriores.

Dos son las cuestiones que se presentan al estudiar sistemas de ecuaciones lineales, en primer lugar, analizarlos, o sea, determinar si tienen solución (compatible) o no (incompatible), y, en caso de tener solución –como segunda instancia–, determinar cuántas y cuáles son las soluciones.

Para la primera instancia utilizaremos, previa demostración, un valioso teorema, el teorema de Rouché–Frobenius. Destacamos que este teorema permite analizar un sistema; no resolverlo. Para ese último caso se analizarán y ejemplificarán distintos métodos de resolución como el de Gauss, Gauss–Jordan, Cramer e inversión de matrices.

3. Análisis de sistemas de ecuaciones lineales

El nombre del teorema se debe al matemático francés Eugène Rouché, que lo enunció en 1875 y a Frobenius. Poco después de su publicación, el francés Georges Fontené se atribuyó la primera demostración del mismo, y Frobenius, en 1905, acreditó la demostración a Rouché y Fontené.

Sin embargo, el nombre por el que se conoce el resultado (especialmente en los países de habla hispana) se debe al matemático hispano-argentino Julio Rey Pastor, que lo refirió de tal forma. En otros países es conocido como Teorema de Rouché–Fontené o Teorema de Kronecker–Capelli[34].

Teorema de Rouché–Frobenius: *La condición necesaria y suficiente para que un sistema de m ecuaciones lineales con n incógnitas, $A.X = B$, sea compatible es que el rango de la matriz de los coeficientes sea igual al rango de la matriz ampliada.*

Sea $A.X = B$ un sistema de m ecuaciones lineales con n incógnitas, $A.X = B$ es compatible $\Leftrightarrow$ $\rho(A) = \rho(A')$.

$D \Longrightarrow$) Si el sistema es compatible, entonces $\rho(A) = \rho(A')$

Sea el sistema de ecuaciones lineales

$$\begin{cases} a_{11}x_1 + & a_{12}x_2 + & \dots & +a_{1n}x_n = & b_1 \\ a_{21}x_1 + & a_{22}x_2 + & \dots & +a_{2n}x_n = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \dots & +a_{mn}x_n = & b_m \end{cases} \quad \text{y}$$

[34] Ernesto Aranda Ortega (2013). *Álgebra lineal con aplicaciones y Python*. Impreso por Lulu.com. Ciudad Real. España.

$$A' = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$ su matriz ampliada.

Por la hipótesis, el sistema es compatible, o sea que tiene solución, es decir que existen n valores, $s_1, s_2, \dots, s_n$, para las incógnitas que verifican todas las ecuaciones del sistema.

$$\begin{cases} a_{11}s_1 + & a_{12}s_2 + & \dots & +a_{1n}s_n = & b_1 \\ a_{21}s_1 + & a_{22}s_2 + & \dots & +a_{2n}s_n = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}s_1 + & a_{m2}s_2 + & \dots & +a_{mn}s_n = & b_m \end{cases}$$

Si en A' sustituimos $b_1, b_2, \dots, b_m$, obtenemos:

$$A' = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & a_{11}s_1 + a_{12}s_2 + \cdots + a_{1n}s_n \\ a_{21} & a_{22} & \dots & a_{2n} & a_{21}s_1 + a_{22}s_2 + \cdots + a_{2n}s_n \\ \vdots & \vdots & \cdots & & \vdots \quad \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & a_{m1}s_1 + a_{m2}s_2 + \cdots + a_{mn}s_n \end{array}\right)$$

La última columna es combinación lineal de las demás, por lo tanto, a efectos del cálculo del rango puede suprimirse. Por lo tanto $\rho(A) = \rho(A')$.

Recíprocamente,

$D \Longleftarrow) \ \rho(A) = \rho(A')$, entonces el sistema es compatible.

Si $\rho(A) = \rho(A') = r$ significa que existe al menos un menor de orden r distinto de cero.

Sea $| C_r | \neq 0$ el determinante de orden r, no nulo.

Si eliminamos las ecuaciones que no intervienen en este menor por ser combinación lineal de las r ecuaciones que sí intervienen, el sistema que obtenemos es equivalente al inicial.

Consideramos las r ecuaciones que intervienen en C_r como principales y las otras como parámetros, transponiéndolas a que formen parte del término independiente. De este modo, obtenemos un sistema equivalente al dado y de Cramer[35] ya que C_r sería la matriz de los coeficientes, luego el sistema sería compatible, con lo que queda demostrado el teorema.

Una consecuencia inmediata del teorema demostrado es que, si el rango de la matriz del sistema y el de la ampliada no coinciden, el sistema es incompatible.

Si $\rho(A) \neq \rho(A')$*, el sistema* $A.X = B$ *es incompatible.*

Esto significa que el sistema tiene al menos una ecuación inconsistente. O sea, ecuaciones de la forma $0x_1 + 0x_2 + \ ... + 0x_n = k$ con la constante $k \neq 0$. Este tipo de ecuaciones son fácilmente reconocibles cuando se escalona la matriz ampliada del sistema. Más adelante se retomará este tema.

Los siguientes teoremas completan lo enunciado en el teorema de Rouché–Frobenius respecto de los sistemas compatibles.

Teorema: *Un sistema de* m *ecuaciones lineales con* n *incógnitas,* $A.X = B$*, tiene solución única, si y solo si* $\rho(A) = \rho(A') = n°\ de\ incógnitas$*.*

$A.X = B$ *tiene única solución* $\Leftrightarrow\ \rho(A) = \rho(A') = n°\ de\ incógnitas$*.*

D) La primera igualdad $\rho(A) = \rho(A')$ está demostrada por el teorema de Rouché–Frobenius con lo que el sistema es compatible.

[35] En el próximo apartado se estudiará la Regla de Cramer.

Ahora bien, si $\rho(A) = \rho(A') = n°\ de\ incógnitas$ existirán n variables principales y el sistema será compatible determinado si y solo si todas sus variables son principales y esto solamente ocurre cuando $\rho(A) = \rho(A') = n°\ de\ incógnitas$.

El siguiente teorema demuestra la condición necesaria y suficiente para que un sistema tenga infinitas soluciones, o soluciones múltiples. Pero previamente debe tenerse en cuenta la siguiente:

Definición: *La solución de un sistema de ecuaciones lineales es múltiple si está en función de al menos una variable (incógnita) que puede tomar cualquier valor en* $\mathbb{K}$*. Este tipo de variables se llaman variables libres.*

Con esta definición se puede enunciar el correspondiente teorema:

Teorema: *Un sistema de* m *ecuaciones lineales con* n *incógnitas,* $A.X = B$*, tiene solución múltiple, si y solo si* $\rho(A) = \rho(A') < n°\ de\ incógnitas$*. Además, el sistema tiene solución múltiple con* $s = n - \rho(A)$ *variables libres.*

$A.X = B$ *tiene infinitas soluciones* $\Leftrightarrow\ \rho(A) = \rho(A') < n°\ de\ incógnitas.$

$D \Longrightarrow$) Supóngase que el sistema tiene infinitas soluciones, entonces al formar la matriz escalonada o escalonada reducida, no pueden aparecer ecuaciones que posean la forma $(0\ \ 0\ \ ...0\ \ \ k)$ con $k \neq 0$ (caso contrario el sistema no tendría solución), esto quiere decir que $\rho(A) = \rho(A')$. Por otro lado, de la definición de solución múltiple, resulta que existe al menos una variable libre, esto quiere decir que tenemos más variables que pivotes, pero la cantidad de pivotes es el rango de la matriz. De lo anterior se puede concluir que $\rho(A) = \rho(A') < n°\ de\ incógnitas$.

$D \Longleftarrow$) Supóngase que $\rho(A) = \rho(A') < n° \, de \, incógnitas$. Por cumplirse la igualdad de rangos, $\rho(A) = \rho(A')$, esto implica que en la matriz escalonada o escalonada reducida del sistema no existen ecuaciones de la forma $(0 \ 0 \ ... 0 \ \ k)$ con $k \neq 0$, esto quiere decir que el sistema tiene solución. La solución es múltiple puesto que existen más variables que pivot, es decir, $\rho(A) = \rho(A') < n° \, de \, incógnitas$.

La cantidad de variables libres es igual a la cantidad de variables, n, menos la cantidad de pivotes en la matriz escalonada o escalonada reducida. Entonces $s = n - \rho(A)$.

En pocas palabras, luego de estos tres teoremas, podemos afirmar que para analizar un sistema de m ecuaciones lineales con n incógnitas, $A.X = B$, debemos calcular el rango de la matriz de los coeficientes, $\rho(A)$, y el de la matriz ampliada, $\rho(A')$. Entonces:

$$Si \begin{cases} \rho(A) = \rho(A'); A.X = B \ es \ compatible, y \begin{cases} \rho(A) = \rho(A') = n; A.X = B \ es \ compatible \ determinado \\ \rho(A) = \rho(A') < n; A.X = B \ es \ compatible \ indeterminado \end{cases} \\ \rho(A) \neq \rho(A'); A.X = B \ es \ incompatible \end{cases}$$

4. Resolución de sistemas de ecuaciones lineales

Resolver un sistema de ecuaciones lineales significa determinar los valores de las incógnitas que satisfacen simultáneamente todas las ecuaciones del sistema. Los sistemas compatibles determinados tendrán un solo conjunto de valores, mientras que los compatibles indeterminados tendrán infinitos conjuntos. En este último caso, suele hablarse de solución general –aquella que expresa todas las soluciones del sistema y en la cual intervienen determinados parámetros– y de soluciones

particulares, aquellas que se obtienen luego de dar valores a los parámetros que intervienen en la solución general.

4.1. Método de eliminación de Gauss

Es un método de resolución de sistemas de ecuaciones lineales que se puede programar muy fácilmente en diversos lenguajes, esto permite resolver un sistema con la ayuda de una computadora.

La idea del método consiste en aplicar, a la matriz ampliada del sistema, transformaciones elementales sobre filas (no pueden realizarse estas transformaciones sobre columnas), obteniendo, de esta forma, sistemas equivalentes al dado, pero, al final, con una matriz escalonada por filas. El trabajo es más sencillo si se emplean matrices ampliadas para representar en todo momento a los sistemas lineales equivalentes que resultan tras las transformaciones.

Recordemos que cuando una matriz está en forma escalonada, los primeros elementos distintos de cero de cada fila reciben el nombre de *pivotes*. Por ser el pivote el primer elemento no nulo de la fila, no hay manera de que una fila tenga más de un pivote, puede no tener pivote, en el caso de ser una fila con todos sus elementos nulos (fila nula), pero no puede tener más de un pivote. Por otro lado, notemos que, por estar la matriz ampliada escalonada por filas, no hay forma que dos pivotes queden en la misma columna, puede una columna no tener pivote, pero, si lo tiene, no puede tener más de un pivote.

De todo esto se puede concluir que una matriz de orden $(m \times n)$ no puede tener más de m pivotes porque tiene a lo sumo uno por cada fila.

Algunos autores piden que el pivote de cada fila sea 1, lo que requiere que se multiplique cada fila no nula por una constante adecuada. Si bien este paso es conveniente para despejar directamente las variables (incógnitas) dependientes, no es imprescindible.

A partir de haber obtenido la forma escalonada de la matriz ampliada de un sistema de ecuaciones lineales, se pueden realizar las siguientes observaciones válidas para cualquier sistema $(m \times n)$.

1.- Si la forma escalonada por fila de la matriz ampliada incluye alguna fila de la forma:

$(0 \ \ 0 \ \ \ldots 0 \mid k)$ con $k \neq 0$ entonces el sistema es incompatible, es decir no tiene solución $\rho(A) \neq \rho(A')$.

En efecto, tal fila significa que $0x_1 + \ 0x_2 + \ \ldots + 0x_n = \ k$ con $k \neq 0$, lo cual no puede satisfacer ninguna n–tupla $(x_1, x_2, \ \ldots, x_n)$.

2.- Si no existe ninguna fila como la indicada anteriormente, entonces el sistema es compatible, $\rho(A) = \rho(A')$, es decir tiene solución, en cuyo caso se presentan dos posiblidades.

2a).- Si las filas no nulas de la forma escalonada por filas de la matriz ampliada del sistema forman una matriz triangular, es decir una matriz de la forma:

$\left(\begin{array}{cccc|c} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 \\ 0 & 0 & 1 & \alpha_8 & \alpha_9 \\ 0 & 0 & 0 & 1 & \alpha_{10} \end{array}\right)$ donde los α_i indican números arbitrarios, entonces el sistema es compatible determinado, $\rho(A) = \rho(A') = n$, es decir tiene única solución. Partiendo desde la última fila y reemplazando los valores obtenidos en las filas inmediatas superiores, se obtiene un valor único para cada incógnita.

Este caso no puede darse en los sistema en donde el número de ecuaciones es menor al número de incógnitas ($m < n$), ya que requiere un número total de filas al menos igual al número de columnas.

2b).- Si no se da el caso anterior, significa que existe una submatriz triangular correspondiente a un subconjunto de las variables (las variables dependientes o pivotes, que corresponden a las columnas con pivotes) y las restantes son variables (incógnitas) *libres* (independientes) que pueden tomar cualquier valor. Por lo tanto el sistema es compatible indeterminado, $\rho(A) = \rho(A') < n$, es decir tiene infinitas soluciones, y por lo general se suele dar una solución general y algunas soluciones particulares. Este último caso puede darse tanto si $m < n$ como si $m = n$ ó $m > n$.

Por ejemplo, si la forma escalonada de la matriz ampliada es la siguiente:

$$\left(\begin{array}{cccccc|c} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 \\ 0 & 0 & 1 & \alpha_7 & \alpha_8 & \alpha_9 & \alpha_{10} \\ 0 & 0 & 0 & 1 & \alpha_{11} & \alpha_{12} & \alpha_{13} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$ el sistema es compatible indeterminado, siendo $x_2, x_5\, y\, x_6$ las varibles libres y $x_1, x_3\, y\, x_4$ las variables dependientes.

De todas estas observaciones podemos concluir lo siguiente:

1.- Un sistema de ecuaciones lineales que tenga menos ecuaciones que incógnitas, es decir, $m < n$, solamente puede ser compatible indeterminado o incompatible.

2.- Los sitemas cuadrados ($m = n$) o los que tienen más ecuaciones que incógnitas, es decir, $m > n$, pueden ser compatibles determinados, compatibles indeterminados o bien incompatibles.

Ejemplos: analizar y resolver, si correspondiere, cada uno de los siguientes sistemas de ecuaciones lineales, aplicando el método de eliminación de Gauss.

a) $\begin{cases} x - y + z = 2 \\ 2x + y - 2z = 3 \\ x + y - z = 2 \\ -x + 2y + z = 1 \end{cases}$

Escribimos la matriz ampliada del sistema y, a partir de ella, la transformamos en una matriz escalonada por filas.

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 2 & 1 & -2 & 3 \\ 1 & 1 & -1 & 2 \\ -1 & 2 & 1 & 1 \end{array}\right) \sim \begin{matrix} -2F_1 + F_2 \\ -1F_1 + F_3 \\ F_1 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 3 & -4 & -1 \\ 0 & 2 & -2 & 0 \\ 0 & 1 & 2 & 3 \end{array}\right) \sim$$

$$\sim \begin{matrix} -\frac{2}{3}F_2 + F_3 \\ -\frac{1}{3}F_2 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 3 & -4 & -1 \\ 0 & 0 & \frac{2}{3} & \frac{2}{3} \\ 0 & 0 & \frac{10}{3} & \frac{10}{3} \end{array}\right) \sim \begin{matrix} \\ \frac{3}{2}F_3 \\ \frac{3}{10}F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 3 & -4 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

$$\sim \begin{matrix} \\ -1F_3 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 3 & -4 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Análisis: Tanto para A como para A' se puede observar que el número de pivotes o el número de columnas pivotales o el número de filas no nulas, es en todos los casos 3; por lo tanto, podemos afirmar que $\rho(A) = \rho(A') = 3$, es decir el sistema es compatible. Como el número de incógnitas es 3, también afirmamos que $\rho(A) = \rho(A') = n = 3$, por lo que el sistema es compatible determinado. También podemos observar que la matriz escalonada obtenida no posee ninguna fila de la forma $(0\ 0\ \ldots 0\ \ k)$ con $k \neq 0$ y las filas no nulas forman una matriz triangular.

Resolución: Reconstruimos el sistema, obteniendo

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

A partir de la última ecuación, $z = 1$, sustituimos en la segunda ecuación, obteniendo:

$3y - 4.1 = -1 \implies y = 1$. Sustituimos $z = 1$ e $y = 1$ en la primera ecuación, obteniendo:

$1x - 1.1 + 1.1 = 2 \implies x = 2$. Este procedimiento suele llamarse sustitución hacia atrás o sustitución regresiva.

Por lo tanto, la única solución de este sistema es la terna $(2,1,1)$ y el conjunto solución será: $S = \{\ (2,1,1)\ \}$.

Verificación:

$$\begin{cases} x-y+z=2 \\ 2x+y-2z=3 \\ x+y-z=2 \\ -x+2y+z=1 \end{cases} \Rightarrow \begin{cases} 2-1+1=2 \\ 2.2+1-2.1=3 \\ 2+1-1=2 \\ -2+2.1+1=1 \end{cases} \Rightarrow \begin{cases} 2=2 \\ 3=3 \\ 2=2 \\ 1=1 \end{cases}$$

b) $\begin{cases} 3x+2y+w=0 \\ x+2y+z=1 \\ 5x+6y+2z+w=2 \end{cases}$

Obsérvese que es un sistema con menos ecuaciones que incógnitas y recuérdese la primera conclusión realizada luego de haber analizado el método de eliminación de Gauss.

Escribimos la matriz ampliada del sistema y a partir de ella la transformamos en una matriz escalonada por filas.

$$\left(\begin{array}{cccc|c} 3 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 \\ 5 & 6 & 2 & 1 & 2 \end{array}\right) \sim_{-\frac{1}{3}F_1+F_2 \atop -\frac{5}{3}F_1+F_3} \left(\begin{array}{cccc|c} 3 & 2 & 0 & 1 & 0 \\ 0 & \frac{4}{3} & 1 & -\frac{1}{3} & 1 \\ 0 & \frac{8}{3} & 2 & -\frac{2}{3} & 2 \end{array}\right) \sim_{3F_2 \atop 3F_3} \left(\begin{array}{cccc|c} 3 & 2 & 0 & 1 & 0 \\ 0 & 4 & 3 & -1 & 3 \\ 0 & 8 & 6 & -2 & 6 \end{array}\right)$$

$$\sim_{-2F_2+F_3} \left(\begin{array}{cccc|c} 3 & 2 & 0 & 1 & 0 \\ 0 & 4 & 3 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Análisis: Tanto para A como para A' se puede observar que el número de pivotes o el número de columnas pivotales o el número de filas no nulas es, en todos los casos, 2; por lo tanto, podemos afirmar que $\rho(A) = \rho(A') = 2$, es decir, el sistema es compatible. Como el número de incógnitas es 4, también afirmamos que $\rho(A) = \rho(A') < n = 4$, por lo que el sistema es

compatible indeterminado. También podemos observar que la matriz escalonada obtenida no posee ninguna fila de la forma $(0\ \ 0\ \ ...0\ \ \ k)$ con $k \neq 0$ y las filas no nulas forman una submatriz triangular. El sistema tiene dos variables dependientes, $x\ e\ y$, y dos variables (incógnitas) libres, z junto con w. Por lo tanto, a z y w las trasponemos junto con los términos independientes, pues tendrán el papel de ser parámetros.

Resolución: Reconstruimos el sistema, obteniendo

$\begin{cases} 3x + 2y + w = 0 \\ 4y + 3z - w = 3 \end{cases}$ por ser compatible indeterminado, tenemos

$\begin{cases} 3x + 2y = -w \\ 4y = 3 + w - 3z \end{cases}$

De la segunda ecuación obtenemos:

$y = \frac{3}{4} + \frac{1}{4}w - \frac{3}{4}z$, reordenando $y = \frac{1}{4}w - \frac{3}{4}z + \frac{3}{4}$. Reemplazando en la primera ecuación:

$3x + 2y = -w \Leftrightarrow 3x = -w - 2\left(\frac{1}{4}w - \frac{3}{4}z + \frac{3}{4}\right) \Leftrightarrow$

$x = -\frac{1}{3}w - \frac{1}{6}w + \frac{1}{2}z - \frac{1}{2}$

$x = \frac{1}{2}z - \frac{1}{2}w - \frac{1}{2}$.

Por lo tanto, el sistema tiene infinitas soluciones cuyo conjunto solución o solución general se expresa como: $S_G = \left\{(x, y, z, w)/x = \frac{1}{2}z - \frac{1}{2}w - \frac{1}{2} \wedge y = \frac{1}{4}w - \frac{3}{4}z + \frac{3}{4} \wedge z, w \in \mathbb{R}\right\}$ mientras que algunas soluciones particulares se pueden obtener dándoles valores particulares a z y w. Por ejemplo:

Si $z = 0\ \ y\ \ w = 0 \Longrightarrow x = -\frac{1}{2}\ e\ y = \frac{3}{4}$ por lo tanto $S_{P1} = \left(-\frac{1}{2}, \frac{3}{4}, 0, 0\right)$

Si $z = 1 \ y \ w = 1 \Longrightarrow x = -\frac{1}{2} \ e \ y = \frac{1}{4}$ por lo tanto $S_{P2} = (-\frac{1}{2}, \frac{1}{4}, 1, 1)$

c) $\begin{cases} x + y - 3z = -1 \\ 2x + y - 2z = 1 \\ x + y + z = 3 \\ x + 2y - 3z = 1 \end{cases}$

De manera similar a los casos anteriores, escribimos la matriz ampliada del sistema y a partir de ella la transformamos en una matriz escalonada por filas.

$$\left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 1 & 2 & -3 & 1 \end{array}\right) \sim \begin{matrix} -2F_1 + F_2 \\ -1F_1 + F_3 \\ -1F_1 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & -1 & 4 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 1 & 0 & 2 \end{array}\right) \sim$$

$$\sim \begin{matrix} \\ -1F_2 \\ \\ \end{matrix} \left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & 4 & 4 \\ 0 & 1 & 0 & 2 \end{array}\right) \sim \begin{matrix} \\ \\ -F_2 + F_4 \\ \end{matrix} \left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 4 & 5 \end{array}\right)$$

$$\sim \begin{matrix} \\ \\ \\ -1F_3 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Análisis: La matriz A tiene tres pivotes o tres columnas pivotales, mientras que A' tiene cuatro, ya que $a_{44} = 1$ es pivote, pero al no tener elementos por debajo, no se hace evidente su uso. Es decir que $\rho(A) = 3 \ \neq \rho(A') = 4$. Por otra parte, obsérvese la cuarta fila de la matriz escalonada obtenida, es de la forma $(0 \ 0 \ ... 0 \ \ k)$ con $k \neq 0$. Es decir que el sistema tiene una ecuación inconsistente, por lo tanto es incompatible.

4.2. Método de Gauss–Jordan

Este método es una variante del método de eliminación de Gauss; también permite analizar el sistema, acorde con el Teorema de Rouché–Frobenius, y resolverlo en caso de ser compatible.

El método consiste no solo en eliminar la incógnita x_i de las ecuaciones $E_{i+1}, E_{i+2}, \dots, E_n$ como hicimos en el de Gauss, sino también de las ecuaciones $E_1, E_2, \dots, E_{i-1}$, es decir, eliminamos en cada columna los elementos que están por debajo y por encima de la diagonal principal.

Este procedimiento va transformando la matriz ampliada del sistema en una matriz escalonada reducida por filas, evitando de esta manera, en primer lugar, la necesidad de la sustitución hacia atrás como en la eliminación gaussiana. Por otro lado, la columna correspondiente a los términos independientes serán las soluciones del sistema, si este es compatible determinado. La idea general es ir escalonando y reduciendo simultáneamente.

Como ejemplo de este método volveremos a resolver los tres sistemas resueltos en el método anterior, de esta manera quedará en evidencia que sin importar el método a utilizar se debe llegar a resultados equivalentes.

Ejemplos: analizar y resolver, si correspondiere, cada uno de los siguientes sistemas de ecuaciones lineales, aplicando el método de Gauss–Jordan.

a) $\begin{cases} x - y + z = 2 \\ 2x + y - 2z = 3 \\ x + y - z = 2 \\ -x + 2y + z = 1 \end{cases}$

Escribimos la matriz ampliada del sistema y a partir de ella la transformamos en una matriz escalonada reducida por filas.

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 2 & 1 & -2 & 3 \\ 1 & 1 & -1 & 2 \\ -1 & 2 & 1 & 1 \end{array}\right) \sim \begin{matrix} \\ -2F_1 + F_2 \\ -1F_1 + F_3 \\ F_1 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 3 & -4 & -1 \\ 0 & 2 & -2 & 0 \\ 0 & 1 & 2 & 3 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & -2 & 0 \\ 0 & 3 & -4 & -1 \end{array}\right)$$

$$\begin{matrix} F_2 + F_1 \\ -2F_2 + F_3 \\ -3F_2 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & 3 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -6 & -6 \\ 0 & 0 & -10 & -10 \end{array}\right) \sim \begin{matrix} \\ \\ -\frac{1}{6}F_3 \\ -\frac{1}{10}F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & 3 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \sim$$

$$\sim \begin{matrix} -3F_3 + F_1 \\ -2F_3 + F_2 \\ \\ -1F_3 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Obsérvese que la obtención de la primera matriz equivalente es de manera idéntica a cómo se hizo en el método de eliminación de Gauss.

Análisis: Tanto para A como para A' se puede observar que el número de pivotes o el número de columnas pivotales o el número de filas no nulas es en todos los casos 3, por lo tanto, podemos afirmar que $\rho(A) = \rho(A') = 3$, es decir, el sistema es compatible. Como el número de incógnitas es 3, también afirmamos que $\rho(A) = \rho(A') = n = 3$, por lo que el sistema es compatible determinado. También podemos observar que la matriz escalonada reducida obtenida no posee ninguna fila de la forma $(0\ \ 0\ \ \ldots 0\ \ \ k)$ con $k \neq 0$.

Resolución: No hace falta reconstruir el sistema, puesto que directamente podemos leer de la matriz escalonada reducida por fila: $x = 2,\ y = 1,\ z = 1$.

Por lo tanto, la única solución de este sistema es la terna $(2, 1, 1)$ y el conjunto solución será: $S = \{\ (2, 1, 1)\ \}$.

Verificación:

$$\begin{cases} x - y + z = 2 \\ 2x + y - 2z = 3 \\ x + y - z = 2 \\ -x + 2y + z = 1 \end{cases} \Rightarrow \begin{cases} 2 - 1 + 1 = 2 \\ 2.2 + 1 - 2.1 = 3 \\ 2 + 1 - 1 = 2 \\ -2 + 2.1 + 1 = 1 \end{cases} \Rightarrow \begin{cases} 2 = 2 \\ 3 = 3 \\ 2 = 2 \\ 1 = 1 \end{cases}$$

b) $\begin{cases} 3x + 2y + w = 0 \\ x + 2y + z = 1 \\ 5x + 6y + 2z + w = 2 \end{cases}$

Escribimos la matriz ampliada del sistema y a partir de ella la transformamos en una matriz escalonada reducida por filas.

$$\left(\begin{array}{cccc|c} 3 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 \\ 5 & 6 & 2 & 1 & 2 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 3 & 2 & 0 & 1 & 0 \\ 5 & 6 & 2 & 1 & 2 \end{array}\right) \sim \begin{matrix} \\ -3F_1 + F_2 \\ -5F_1 + F_3 \end{matrix} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 0 & -4 & -3 & 1 & -3 \\ 0 & -4 & -3 & 1 & -3 \end{array}\right)$$

$$\begin{matrix} \\ -1F_2 + F_3 \end{matrix} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 0 & -4 & -3 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \sim -\frac{1}{4}F_2 \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\sim \begin{matrix} -2F_2 + F_1 \end{matrix} \left(\begin{array}{cccc|c} 1 & 0 & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Análisis: Tanto para A como para A' se puede observar que el número de pivotes o el número de columnas pivotales o el número de filas no nulas es, en todos los casos 2, por lo tanto, podemos afirmar que $\rho(A) = \rho(A') = 2$, es decir, el sistema es compatible. Como el número de incógnitas es 4, también afirmamos que $\rho(A) = \rho(A') < n = 4$, por lo que el sistema es compatible indeterminado.

Resolución: El sistema tiene dos variables dependientes, $x\ e\ y$, y dos variables (incógnitas) libres, z junto con w. Por lo tanto, a z y w las trasponemos junto con los términos independientes, pues tendrán el papel de ser parámetros.

$$\begin{cases} 2x - z + w = -1 \\ 4y + 3z - w = 3 \end{cases} \Longrightarrow \begin{cases} 2x = z - w - 1 \\ 4y = -3z + w + 3 \end{cases}$$

De la última ecuación $y = \frac{-3z+w+3}{4}$ y de la primera ecuación $x = \frac{z-w-1}{2}$

Por lo tanto, el sistema tiene infinitas soluciones, cuya solución general se expresa como: $S_G = \left\{(x, y, z, w)/x = \frac{1}{2}z - \frac{1}{2}w - \frac{1}{2} \wedge y = -\frac{3}{4}z + \frac{1}{4}w + \frac{3}{4} \wedge z, w \in \mathbb{R}\right\}$, mientras que algunas soluciones particulares se pueden obtener dándoles valores particulares a z y w. Por ejemplo:

Si $z = 0\ y\ w = 0 \Longrightarrow x = -\frac{1}{2}\ e\ y = \frac{3}{4}$ por lo tanto $S_{P1} = (-\frac{1}{2}, \frac{3}{4}, 0, 0)$

Si $z = 0\ y\ w = 1 \Longrightarrow x = -1\ e\ y = 1$ por lo tanto $S_{P2} = (-1, 1, 0, 1)$

c) $\begin{cases} x + y - 3z = -1 \\ 2x + y - 2z = 1 \\ x + y + z = 3 \\ x + 2y - 3z = 1 \end{cases}$

De manera similar a los casos anteriores, escribimos la matriz ampliada del sistema y a partir de ella la transformamos en una matriz escalonada reducida por filas.

$$\left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 1 & 2 & -3 & 1 \end{array}\right) \sim \begin{matrix} -2F_1 + F_2 \\ -1F_1 + F_3 \\ -1F_1 + F_4 \end{matrix} \left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & -1 & 4 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 1 & 0 & 2 \end{array}\right) \sim -1F_2 \left(\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & 4 & 4 \\ 0 & 1 & 0 & 2 \end{array}\right)$$

$$-1F_2\begin{pmatrix}1 & 1 & -3 & -1\\ 0 & 1 & -4 & -3\\ 0 & 0 & 4 & 4\\ 0 & 1 & 0 & 2\end{pmatrix}\sim\begin{matrix}\\ -1F_2+F_1\\ \\ -1F_2+F_4\end{matrix}\begin{pmatrix}1 & 0 & 1 & 2\\ 0 & 1 & -4 & -3\\ 0 & 0 & 4 & 4\\ 0 & 0 & 4 & 5\end{pmatrix}\sim_{\frac{1}{4}F_3}\begin{pmatrix}1 & 0 & 1 & 2\\ 0 & 1 & -4 & -3\\ 0 & 0 & 1 & 1\\ 0 & 0 & 4 & 5\end{pmatrix}$$

$$\begin{matrix}-1F_3+F_1\\ 4F_3+F_2\\ \\ -4F_3+F_4\end{matrix}\begin{pmatrix}1 & 0 & 0 & 1\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1\end{pmatrix}\sim\begin{matrix}-1F_4+F_1\\ -1F_4+F_2\\ -1F_4+F_3\\ \\ \end{matrix}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Análisis: La matriz A tiene tres pivotes o tres columnas pivotales o tres filas no nulas, pero A' tiene cuatro. Es decir que $\rho(A) = 3 \neq \rho(A') = 4$. Por otra parte, obsérvese la cuarta fila de la matriz escalonada reducida obtenida, es de la forma $(0\ 0\ \ldots 0\ \ k)$ con $k \neq 0$. Esto significa que el sistema tiene una ecuación inconsistente, por lo tanto, el sistema es incompatible.

4.3. Método de la matriz inversa

Este método consiste en utilizar la matriz inversa de la matriz de los coeficientes para hallar la matriz columna de las incógnitas. Entonces debemos tener, en primer lugar, la seguridad de que existe dicha matriz inversa. Luego se debe multiplicar por la izquierda los dos miembros de la expresión matricial del sistema de ecuaciones por la matriz inversa de la matriz de los coeficientes.

Sea un sistema de m ecuaciones lineales con n incógnitas, expresado como ecuación matricial $A.X = B$.

Analizado el sistema, resulta $\rho(A) = \rho(A') = r$ con $r \leq m$ y $r \leq n$. Entonces, se puede construir un sistema equivalente al dado $C.X = D$ donde C es una matriz cuadrada, invertible de orden r, extraída de A.

$$C = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \cdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{pmatrix}$$

Las matrices X y D serán

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} \text{ y } D = \begin{pmatrix} b_1 - \sum_{j=r+1}^{n} a_{1j}x_j \\ b_2 - \sum_{j=r+1}^{n} a_{2j}x_j \\ \vdots \\ b_r - \sum_{j=r+1}^{n} a_{rj}x_j \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \end{pmatrix}$$

A partir de este nuevo sistema, $C.X = D$, solo resta multiplicar por izquierda ambos miembros de la ecuación por la inversa de C, obteniendo $C^{-1}.C.X = C^{-1}.D$ de donde

$I.X = C^{-1}.D$ finalmente

$X = C^{-1}.D$

Puesto que los sistemas son equivalentes, las soluciones obtenidas mediante esta última expresión son también las soluciones del sistema original, $A.X = B$. Esto significa que para obtener la matriz columna de las incógnitas debemos multiplicar la matriz inversa de C por la matriz columna de los términos independientes.

Ejemplo: resolver el siguiente sistema por el método de la matriz inversa.

$$\begin{cases} x+y+z+w=2 \\ x-y-z+w=1 \\ x+y-z-w=0 \\ 3x-y-3z+w=2 \end{cases}$$

Como tenemos que averiguar tanto el rango de la matriz del sistema como el de la matriz ampliada y calcular una matriz inversa, conviene hacer uso de los determinantes.

Calculamos un menor de orden dos, por ejemplo:

$\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$, como es distinto de cero, calculamos un menor de orden tres que incluya el menor anterior

$\begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{vmatrix} = 4$, como es distinto de cero, calculamos un menor de orden cuatro que incluya el menor anterior

$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 3 & -1 & -3 & 1 \end{vmatrix} = 0$, como es el único determinante que podemos construir para A, el rango de la matriz del sistema es 3, $\rho(A) = 3$.

Para la matriz ampliada ocurrirá algo similar, ya que a partir del menor de orden 3 no nulo que calculamos, podemos construir dos menores de orden cuatro, que incluyan al de orden tres.

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

Como los menores de orden cuatro son nulos, significa que el rango de la matriz ampliada es 3, $\rho(A') = 3$. Por lo tanto:

$\rho(A) = \rho(A') = 3 < n°\ de\ incógnitas = 4$ significa que el sistema es compatible indeterminado. Como el número de incógnitas es cuatro y el rango de las matrices es tres, significa que hay una incógnita no principal que debemos hacer actuar como parámetro. Para construir el sistema equivalente al dado, tomamos como matriz del sistema, aquella que corresponde al menor de orden tres no nulo que hayamos calculado. Es decir que nuestro sistema será:

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

La matriz de este nuevo sistema admite inversa, pues tiene su determinante distinto de cero y las matrices que nos interesan son:

$$C = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{pmatrix},\ X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ D = \begin{pmatrix} 2 - w \\ 1 - w \\ w \end{pmatrix}$$

Calculamos la matriz inversa de C, C^{-1} por cualquiera de los dos métodos ya estudiados.

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

Luego de verificar el cálculo de C^{-1} ($C\,.\,C^{-1} = I_3$), procedemos a realizar la multiplicación $C^{-1}\,.\,D$ para obtener la matriz de las incógnitas.

$\mathbf{C^{-1}}$			**D**
			$2 - w$
			$1 - w$
			w
$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{3}{2} - w$
0	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2} + w$
$\frac{1}{2}$	0	$-\frac{1}{2}$	$1 - w$

Por lo tanto, la solución general será:

$S_G = \{(x, y, z, w) / x = \frac{3}{2} - w,\ y = -\frac{1}{2} + w,\ z = 1 - w, w \in \mathbb{R}\}$

Mientras que algunas de las soluciones particulares serán:

$S_P = \left\{\left(\frac{3}{2}, -\frac{1}{2}, 1, 0\right), \left(\frac{1}{2}, \frac{1}{2}, 0, 1\right), \ldots\right\}$

4.4. Regla de Cramer

Finalizaremos el estudio de los métodos, para resolver sistemas de ecuaciones lineales, haciendo hincapié en cómo las propiedades de los determinantes también pueden ayudar en este tema.

Más exactamente, veremos cómo los determinantes permiten calcular explícitamente el valor de cada incógnita, por medio de la Regla de Cramer[36]. Esta regla tiene su utilidad, pues proporciona una expresión directa de cada incógnita del sistema, pero en contrapartida, cuando el sistema es grande, obliga al cálculo de determinantes de tamaño excesivo.

Sea un sistema de m ecuaciones lineales con n incógnitas, expresado como ecuación matricial $A.X = B$.

Analizado el sistema, resulta $\rho(A) = \rho(A') = r$ con $r \leq m\ y\ r \leq n$. Entonces, se puede construir un sistema equivalente al dado $C.X = D$ donde C es una matriz cuadrada, invertible de orden r, extraída de A.

[36] Esta regla lleva el nombre del matemático suizo Gabriel Cramer que la publicó y popularizó en 1750, aunque el mismo resultado ya había sido publicado, dos años antes, por el escosés Colin Maclaurin.

$$C = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \vdots & \vdots & \cdots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{pmatrix}$$

Las matrices X y D serán

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} \text{ y } D = \begin{pmatrix} b_1 - \sum_{j=r+1}^{n} a_{1j}x_j \\ b_2 - \sum_{j=r+1}^{n} a_{2j}x_j \\ \vdots \\ b_r - \sum_{j=r+1}^{n} a_{rj}x_j \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \end{pmatrix}$$

Obsérvese que esta primera parte es idéntica al método de la matriz inversa. A partir de aquí enunciaremos y demostraremos la:

Regla de Cramer: *Sea el sistema de ecuaciones lineales $C.X = D$, donde C es una matriz cuadrada, invertible de orden r; sea D_i la matriz que se obtiene de C al sustituir su columna i por D, para $i = 1, \dots, r$; entonces, la solución del sistema viene dada por:*

$x_i = \frac{det(D_i)}{det(C)} \ \forall\, i = 1, \dots, r.$

D) Sabemos que $X = C^{-1}D$, luego la incógnita x_i será el producto de la fila i de C^{-1} por la matriz columna D. Como sabemos que $C^{-1} = \frac{1}{det(C)}\, Adj\,(C)$, la fila i de esta matriz será:

$\left(\frac{C_{1i}}{det(C)}, \frac{C_{2i}}{det(C)}, \dots, \frac{C_{ri}}{det(C)}\right)$. Por lo tanto, tendremos:

$$x_i = \left(b_1 \frac{C_{1i}}{det(C)}, b_2 \frac{C_{2i}}{det(C)}, \dots, b_r \frac{C_{ri}}{det(C)}\right)$$

$x_i = \frac{1}{det(C)}(b_1C_{1i} + b_2C_{2i} + \ ... + b_rC_{ri})$ Pero el factor entre paréntesis es el desarrollo por la columna i del determinante de la matriz D_i, por lo tanto: $x_i = \frac{det(D_i)}{det(C)} \ \forall\, i = 1, ..., r$, como queríamos demostrar.

Como ejemplo resolveremos el mismo sistema planteado en el método anterior; pues en este caso al trabajar directamente con determinantes conviene, con mucha más razón, calcular los rangos haciendo uso de ellos.

Resolver el siguiente sistema por la Regla de Cramer.

$$\begin{cases} x + y + z + w = 2 \\ x - y - z + w = 1 \\ x + y - z - w = 0 \\ 3x - y - 3z + w = 2 \end{cases}$$

$\rho(A) = \rho(A') = 3 < n°\ de\ incógnitas = 4$ significa que el sistema es compatible indeterminado. Nuevamente, la diferencia entre el número de incógnitas y el rango de las matrices nos da la cantidad de variables no principales que deberemos hacer actuar como parte del término independiente.

El sistema equivalente será:

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

Las matrices que nos interesan son:

$$C = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{pmatrix},\ X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ D = \begin{pmatrix} 2 - w \\ 1 - w \\ w \end{pmatrix}$$

$| C | = 4$

$C_x = \begin{pmatrix} 2-w & 1 & 1 \\ 1-w & -1 & -1 \\ w & 1 & -1 \end{pmatrix}$, en consecuencia:

$|\, C_x \,| = \begin{vmatrix} 2-w & 1 & 1 \\ 1-w & -1 & -1 \\ w & 1 & -1 \end{vmatrix} = 6 - 4w$. Por lo tanto:

$x = \frac{|\, C_x \,|}{|\, C \,|} = \frac{6-4w}{4} = \frac{3}{2} - w$

$C_y = \begin{pmatrix} 1 & 2-w & 1 \\ 1 & 1-w & -1 \\ 1 & w & -1 \end{pmatrix}$, en consecuencia:

$\left|\, C_y \,\right| = \begin{vmatrix} 1 & 2-w & 1 \\ 1 & 1-w & -1 \\ 1 & w & -1 \end{vmatrix} = -2 + 4w$. Por lo tanto:

$y = \frac{|\, C_y \,|}{|\, C \,|} = \frac{-2+4w}{4} = -\frac{1}{2} + w$

$C_z = \begin{pmatrix} 1 & 1 & 2-w \\ 1 & -1 & 1-w \\ 1 & 1 & w \end{pmatrix}$, en consecuencia:

$|\, C_z \,| = \begin{vmatrix} 1 & 1 & 2-w \\ 1 & -1 & 1-w \\ 1 & 1 & w \end{vmatrix} = 4 - 4w$. Por lo tanto:

$z = \frac{|\, C_z \,|}{|\, C \,|} = \frac{4-4w}{4} = 1 - w$

La solución general es:

$S_G = \{(x, y, z, w) \,/ x = \frac{3}{2} - w,\ y = -\frac{1}{2} + w,\ z = 1 - w \,, w \in \mathbb{R}\}$

Mientras que algunas de las soluciones particulares, distintas a las del método anterior, serán:

$S_P = \left\{\left(\frac{5}{2}, -\frac{3}{2}, 2, -1\right), \left(-\frac{1}{2}, \frac{3}{2}, -1, 2\right), \ldots\right\}$

5. Sistemas de ecuaciones lineales homogéneos

Un sistema de m ecuaciones lineales con n incógnitas se dice homogéneo si todos sus términos independientes son nulos, es decir, es un sistema del tipo:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ \ldots + a_{2n}x_n = 0 \\ \quad \vdots \qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ \ldots + a_{mn}x_n = 0 \end{cases}$$ que podemos escribir, de manera resumida, como $\sum_{j=1}^{n} a_{ij}x_j = 0\,;\ i = 1,2,\ldots,m$

Donde los a_{ij} son elementos de un cuerpo $\mathbb{K}$.

Es inmediato comprobar que todo sistema de ecuaciones lineales homogéneo es compatible, ya que la n–tupla $(0, 0, \ldots, 0)$, comúnmente llamada solución trivial, es solución del mismo. Por otra parte, si aplicamos, por ejemplo, el método de eliminación de Gauss o el de Gauss–Jordan, al tener la matriz ampliada la última columna nula, por más transformaciones elementales por filas que se haga, esta columna siempre se mantendrá nula. Esto significa que nunca habrá un pivote en esa columna y, por lo tanto, el rango de la matriz del sistema y el de la matriz ampliada siempre serán iguales, lo que significa que todo sistema de ecuaciones lineales homogéneo siempre es compatible.

Por supuesto, si los rangos de las matrices del sistema y ampliada coinciden con el número de incógnitas, el sistema será determinado, siendo su única solución la trivial. Si el rango es menor al número de incógnitas, el sistema será indeterminado, teniendo además de la solución trivial, otras infinitas soluciones más.

Un sistema de ecuaciones lineales homogéneo que tenga menos ecuaciones que incógnitas ($m < n$) es necesariamente compatible indeterminado (tiene infinitas soluciones). Esto es cierto, pues al estudiar el método de eliminación de Gauss, se hicieron unas observaciones válidas para cualquier sistema y acto seguido se llegó a dos conclusiones, una de las cuales establecía: "Un sistema de ecuaciones lineales que tenga menos ecuaciones que incógnitas ($m < n$) solamente puede ser compatible indeterminado o incompatible". Como ya establecimos que los sistemas homogéneos son siempre compatibles, se demuestra lo enunciado en el inicio del párrafo.

En resumen:

Sea el sistema de m ecuaciones lineales con n incógnitas, homogéneo $A \,.\, X = N$ con la matriz $A \in \mathbb{K}^{m \times n}$ y N matriz nula de orden $m \times 1$, entonces:

a) Si $n = \rho(A)$, el sistema es compatible determinado, es decir, la única solución que admite es la trivial. Esto es evidente ya que llevamos A a su forma escalonada reducida, reconstruimos el sistema y todas las variables quedarán igualadas a cero.

b) Si $n > \rho(A)$, el sistema es compatible indeterminado, es decir que tiene infinitas soluciones, con $s = n - \rho(A)$ variables libres que pueden tomar cualquier valor en $\mathbb{R}$.

Ejemplo: analizar y resolver los siguientes sistemas.

$$a)\begin{cases} -2x+6y-4z=0 \\ -x+3y-z=0 \\ 5x-15y+10z=0 \\ x-3y=0 \end{cases}$$

Como es un sistema homogéneo sabemos que es compatible y para ver si es determinado o indeterminado escalonamos la matriz del sistema.

$$\begin{pmatrix} -2 & 6 & -4 \\ -1 & 3 & -1 \\ 5 & -15 & 10 \\ 1 & -3 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -3 & 0 \\ -1 & 3 & -1 \\ 5 & -15 & 10 \\ -2 & 6 & -4 \end{pmatrix} \sim \begin{matrix} \\ F_1+F_2 \\ -5F_1+F_3 \\ 2F_1+F_4 \end{matrix} \begin{pmatrix} 1 & -3 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 10 \\ 0 & 0 & -4 \end{pmatrix} \sim$$

$$\sim \begin{matrix} \\ \\ 10F_2+F_3 \\ 4F_2+F_4 \end{matrix} \begin{pmatrix} 1 & -3 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Con lo que concluimos que $\rho(A)=\rho(A')=2<4(n°\ de\ incógnitas)$, por lo que el sistema es compatible indeterminado, con una variable libre.

Reconstruyendo el sistema tenemos

$\begin{cases} x-3y=0 \\ -z=0 \end{cases}$ por lo tanto $\begin{cases} x=3y \\ z=0 \end{cases}$

La solución general del sistema es $S_G=\{(x,y,z)/x=3y, z=0, y\in\mathbb{R}\}$. Para obtener las soluciones particulares, damos valores convenientes a la variable y para obtener los valores de x, y z siempre valdrá cero.

$$b)\begin{cases} 4x+5y-z=0 \\ 3x-2y+z=0 \\ -x+2y+2z=0 \end{cases}$$

Como observamos que es un sistema homogéneo de tres ecuaciones con tres incógnitas, recurrimos directamente al determinante del sistema para saber si es compatible determinado o indeterminado (esta idea se

específica con más precisión en el apartado *Sistema de ecuaciones lineales con igual número de ecuaciones e incógnitas*).

$\begin{vmatrix} 4 & 5 & -1 \\ 3 & -2 & 1 \\ -1 & 2 & 2 \end{vmatrix} = -63$, por lo tanto, el sistema es compatible determinado, siendo su única solución la trivial. $S_G = \{(0,0,0)\}$.

5.1. Relación entre los sistemas no homogéneos y su sistema homogéneo asociado

Analicemos el siguiente sistema de ecuaciones lineales, aplicando Gauss–Jordan

$$\begin{cases} 2x - y + 3z = 3 \\ x + z = 1 \\ 4x - y + 5z = 5 \end{cases}$$

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & 3 \\ 1 & 0 & 1 & 1 \\ 4 & -1 & 5 & 5 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 2 & -1 & 3 & 3 \\ 4 & -1 & 5 & 5 \end{array}\right) \sim \begin{matrix} \\ -2F_1 + F_2 \\ -4F_1 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \\ 0 & -1 & 1 & 1 \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & -1 & 1 & 1 \end{array}\right) \sim \begin{matrix} \\ \\ F_2 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$ Por lo tanto:

$\rho(A) = \rho(A') = 2 < 3(n° \, de \, incógnitas)$, en consecuencia, el sistema se clasifica como SCI. Cantidad de variables libres: 1(una).

Reconstruimos el sistema

$\begin{cases} x + z = 1 \\ y - z = -1 \end{cases}$ (1) por lo tanto $\begin{cases} x = 1 - z \\ y = z - 1 \end{cases}$

La solución general del sistema es $S_G = \{(x,y,z)/x = 1 - z,\ y = z - 1,\ z \in \mathbb{R}\}$.

El vector solución puede ser reescrito de la siguiente manera:

$(x,y,z) = (1 - z,\ z - 1,\ z)$

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

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

Si consideramos a la variable z como un parámetro cualquiera, λ por ejemplo, la solución general se puede escribir como $S_G = \{(x, y, z)/ (x, y, z) = \lambda(-1,\ 1,\ 1) + (1, -1,\ 0)\}$

Gráficamente, la solución del sistema es la recta que resulta de la intersección de los dos planos dados por (1).

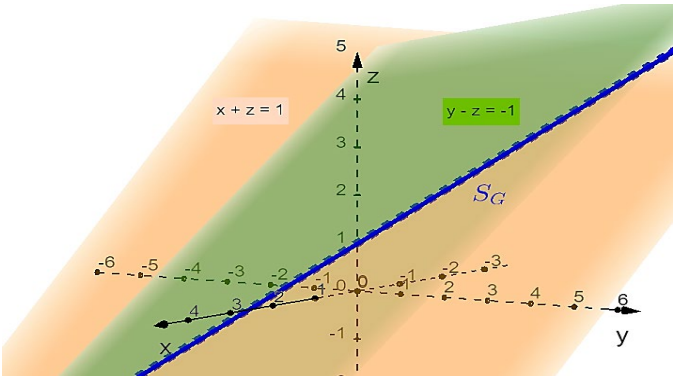


Figura 6.6. Solución del sistema.

Resolvemos ahora el sistema homogéneo asociado al dado, es decir:

$$\begin{cases} 2x - y + 3z = 0 \\ x + z = 0 \\ 4x - y + 5z = 0 \end{cases}$$

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 1 & 0 & 1 & 0 \\ 4 & -1 & 5 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & -1 & 3 & 0 \\ 4 & -1 & 5 & 0 \end{array}\right) \sim \begin{matrix} \\ -2F_1 + F_2 \\ -4F_1 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right) \sim \begin{matrix} \\ \\ F_2 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$ Por lo tanto:

$\rho(A) = \rho(A') = 2 < 3(n° \ de \ incógnitas)$, en consecuencia, el sistema se clasifica como SCI. Cantidad de variables libres: 1 (una).

Reconstruimos el sistema:

$\begin{cases} x + z = 0 \\ y - z = 0 \end{cases}$ (2) por lo tanto $\begin{cases} x = -z \\ y = z \end{cases}$

La solución general del sistema homogéneo es $S_{GH} = \{(x, y, z)/x = -z, y = z,\ z \in \mathbb{R}\}$.

El vector solución puede ser reescrito de la siguiente manera:

$(x, y, z) = (-z,\ z,\ z)$

$(x, y, z) = z(-1,\ 1,\ 1)$

Si consideramos a la variable z como un parámetro cualquiera, λ por ejemplo, la solución general se puede escribir como $S_{GH} = \{(x, y, z)/ (x, y, z) = \lambda(-1,\ 1,\ 1)\}$

Gráficamente, la solución del sistema es la recta que resulta de la intersección de los dos planos dados por (2).

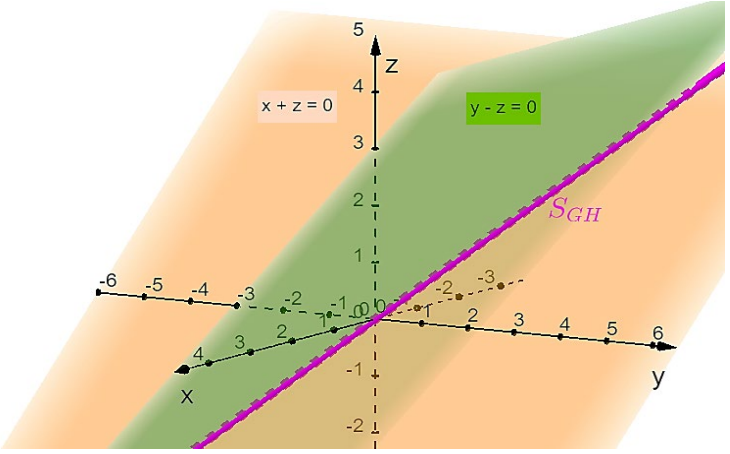


Figura 6.7. Solución del sistema.

Si graficamos ambas soluciones en un mismo sistema, tenemos:

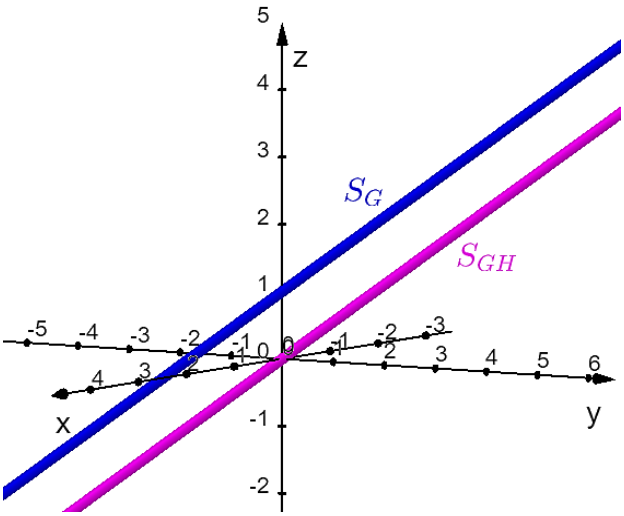


Figura 6.8. Soluciones de ambos sistemas.

Podemos observar que:

a) Tanto la solución S_G como S_{GH} representan una recta.

b) Las rectas que representan la solución de los sistemas son paralelas.

c) La recta que representa la solución S_G no pasa por el origen de coordenadas, mientras que la que representa a S_{GH}, sí lo hace.

Este ejemplo nos permite preguntarnos, ¿qué relación existe entre los conjuntos solución de un sistema no homogéneo y el de su homogéneo asociado?

Las siguientes propiedades nos permiten responder a esta pregunta.

Sea el sistema de ecuaciones lineales $A.X = B$ con $A \in \mathbb{K}^{m\times n}$ y $B \in \mathbb{K}^{m\times 1}$, y $A.X = N$ sistema homogéneo asociado con N matriz nula de orden $m \times 1$, entonces:

P1) La diferencia de dos soluciones particulares de $A.X = B$ es solución del sistema homogéneo $A.X = N$ asociado.

D) Si $X_1\ y\ X_2$ son dos soluciones particulares de $A.X = B$, entonces:

$A.X_1 = B$

$A.X_2 = B$

Restando miembro a miembro

$A.X_1 - A.X_2 = N$

$A.(X_1 - X_2) = N$

Esto significa que $(X_1 - X_2)$ es solución de $A.X = N$ que es lo que se quería demostrar.

P2) La suma de una solución particular de $A.X = B$ y una del $A.X = N$ asociado es solución del sistema $A.X = B$.

D) Sean:

X_P una solución particular de $A.X = B$, es decir que $A.X_P = B$.

X_H una solución particular de $A.X = N$, es decir que $A.X_H = N$.

Queremos probar que $(X_P + X_H)$ es solución de $A.X = B$, o sea que

$A.(X_P + X_H) = B$

$A.(X_P + X_H) = B$

$A.X_P + A.X_H = B$

$B + N = B$

$B = B$ quedando demostrada la propiedad.

El recíproco de la P2) es también válido y se enuncia en la tercera propiedad.

P3) Cualquier solución de $A.X = B$ puede expresarse como $(X_P + X_H)$ con X_P solución particular de $A.X = B$ y X_H solución del $A.X = N$ asociado.

D) Sea X una solución de $A.X = B$, entonces

$X = X - X_P + X_P$

$X = (X - X_P) + X_P$. Por la P1) $(X - X_P)$ es solución de $A.X = N$ o sea que $X = X_H + X_P$ que es lo que se quería demostrar.

Una consecuencia inmediata de P2) y P3) se puede enunciar de la siguiente manera:

"El conjunto solución del sistema de ecuaciones lineales $A.X = B$ puede expresarse como suma de una solución particular de $A.X = B$ y la solución del sistema homogéneo asociado".

Ejemplo 1: consideremos el siguiente sistema de ecuaciones lineales no homogéneo.

$$\begin{cases} x - y + 2z = -4 \\ 3x - 5y + 8z = -14 \\ x + 3y - 2z = 0 \end{cases}$$

Al resolverlo, tenemos

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & -4 \\ 3 & -5 & 8 & -14 \\ 1 & 3 & -2 & 0 \end{array}\right) \sim_{\substack{-3F_1 + F_2 \\ -F_1 + F_3}} \left(\begin{array}{ccc|c} 1 & -1 & 2 & -4 \\ 0 & -2 & 2 & -2 \\ 0 & 4 & -4 & 4 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 2 & -4 \\ 0 & 1 & -1 & 1 \\ 0 & 4 & -4 & 4 \end{array}\right)$$

$$\sim_{\substack{F_2 + F_1 \\ -4F_2 + F_3}} \left(\begin{array}{ccc|c} 1 & 0 & 1 & -3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Como $\rho(A) = \rho(A') = 2 < 3$ $(n°\ de\ incógnitas)$, el sistema se clasifica como SCI. Cantidad de variables libres: 1 (una).

Reconstruimos el sistema

$\begin{cases} x+z=-3 \\ y-z=1 \end{cases}$ por lo tanto $\begin{cases} x=-z-3 \\ y=z+1 \end{cases}$

$S_G = \{(x,y,z)/\, x=-z-3, y=z+1, z\in\mathbb{R}\}$

$(x,y,z)=(-z-3,z+1,z)$

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

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

Tomando a z como parámetro:

$(x,y,z)=\lambda(-1,1,1)+(-3,1,0)$

Por lo tanto:

$S_P=(-3,1,0)$ es solución particular del sistema no homogéneo y

$S_H=\lambda(-1,1,1)$ es la solución del sistema homogéneo asociado.

Verificamos que el vector $(-3,1,0)$ es una solución particular del sistema no homogéneo.

$\begin{cases} x-y+2z=-4 \\ 3x-5y+8z=-14, \\ x+3y-2z=0 \end{cases}$ por lo tanto: $\begin{cases} -3-1+2.0=-4 \\ 3.(-3)-5.1+8.0=-14 \\ -3+3.1-2.0=0 \end{cases} \Longrightarrow$

$\begin{cases} -4=-4 \\ -14=-14 \\ 0=0 \end{cases}$

Verificamos que el vector $(-1,1,1)$ es una solución particular ($\lambda=1$) del sistema homogéneo asociado al no homogéneo.

$\begin{cases} x-y+2z=0 \\ 3x-5y+8z=0, \\ x+3y-2z=0 \end{cases}$ por lo tanto: $\begin{cases} -1-1+2.1=0 \\ 3.(-1)-5.1+8.1=0 \\ -1+3.1-2.1=0 \end{cases} \Longrightarrow \begin{cases} 0=0 \\ 0=0 \\ 0=0 \end{cases}$

Cualquier otra solución del sistema homogéneo asociado se obtiene dándole valor al parámetro λ, por ejemplo: si $\lambda=-2$, la solución particular será $(2,-2,-2)$.

$\begin{cases} x - y + 2z = 0 \\ 3x - 5y + 8z = 0 \\ x + 3y - 2z = 0 \end{cases}$, por lo tanto: $\begin{cases} 2 - (-2) + 2.(-2) = 0 \\ 3.2 - 5(-2) + 8(-2) = 0 \\ 2 + 3(-2) - 2(-2) = 0 \end{cases} \Longrightarrow \begin{cases} 0 = 0 \\ 0 = 0 \\ 0 = 0 \end{cases}$

Ejemplo 2: consideremos el siguiente sistema de ecuaciones lineales no homogéneo.

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

Al analizarlo, tenemos

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & -1 & 2 \\ 1 & 1 & -1 & 2 & 1 \end{array}\right) \sim \underset{-F_1 + F_3}{} \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & -1 & 1 & -2 \end{array}\right) \sim$$

$$\sim \underset{F_2 + F_3}{} \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Como $\rho(A) = \rho(A') = 2 < 4$ $(n° \, de \, incógnitas)$, el sistema se clasifica como SCI. Cantidad de variables libres: 2 (dos).

Reconstruimos el sistema:

$\begin{cases} x + y + w = 3 \\ z - w = 2 \end{cases}$, por lo tanto: $\begin{cases} x = -y - w + 3 \\ z = w + 2 \end{cases}$

$S_G = \{(x, y, z, w) / \, x = -y - w + 3, \; y, \; w + 2, \; w \; con \; y, \; w \in \mathbb{R}\}$

$(x, y, z, w) = (-y - w + 3, \; y, \; w + 2, \; w)$

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

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

Tomando a y, w como parámetros:

$(x, y, z, w) = \lambda \, (-1, 1, 0, 0) + \beta \, (-1, 0, 1, 1) + (3, 0, 2, 0)$

Por lo tanto:

$S_P = (3, 0, 2, 0)$ es solución particular del sistema no homogéneo y

$S_H = \lambda\,(-1,1,0,0) + \beta\,(-1,0,1,1)$ es la solución del sistema homogéneo asociado.

Verificamos que el vector $(3, 0{,}2, 0)$ es una solución particular del sistema no homogéneo.

$$\begin{cases} x+y+w=3 \\ z-w=2 \\ x+y-z+2w=1 \end{cases}$$, por lo tanto: $$\begin{cases} 3+0+0=3 \\ 2-0=2 \\ 3+0-2+2.0=1 \end{cases} \Longrightarrow \begin{cases} 3=3 \\ 2=2 \\ 1=1 \end{cases}$$

Verificamos que el vector $\lambda\,(-1,1,0,0) + \beta\,(-1,0,1,1)$ es una solución particular del sistema homogéneo asociado al no homogéneo. Para $\lambda = \beta = 1$ tenemos $(-2,1,1,1)$

$$\begin{cases} x+y+w=0 \\ z-w=0 \\ x+y-z+2w=0 \end{cases}$$, por lo tanto: $$\begin{cases} -2+1+1=0 \\ 1-1=0 \\ -2+1-1+2=0 \end{cases} \Longrightarrow \begin{cases} 0=0 \\ 0=0 \\ 0=0 \end{cases}$$

Cualquier otra solución del sistema homogéneo asociado se obtiene dándole valor a los parámetros $\lambda\ y\ \beta$.

Terminemos esta sección viendo cómo se relaciona lo que vimos anteriormente con el rango de una matriz.

Sea el sistema de ecuaciones lineales $A.X = B$ con $A \in \mathbb{R}^{m\times n}\ y\ B \in \mathbb{R}^{m\times 1}$ y el sistema $A.X = N$ el sistema homogéneo asociado con N matriz nula de orden $m \times 1$, observemos que:

a) Cada ecuación que produzca un pivote en la forma escalonada de la matriz del sistema es una ecuación que aporta información para resolver el sistema homogéneo asociado, que no está contemplada en las otras ecuaciones.

b) Una fila nula en la forma escalonada de la matriz del sistema representa una ecuación que puede omitirse del sistema homogéneo asociado sin que se pierda información alguna. Esta fila es una ecuación redundante o superflua.

Recordando que el rango de una matriz es el número de filas no nulas o de columnas pivotales que tiene la matriz escalonada equivalente a la dada, podemos decir que el rango informa el número de ecuaciones que realmente aporta información para resolver el sistema homogéneo $A.X = N$. Esto equivale a decir que el rango de una matriz es:

a) el máximo número de ecuaciones del sistema $A.X = N$ que no incluyen ecuaciones redundantes y

b) el mínimo número de ecuaciones que se deben preservar en el sistema $A.X = N$ para obtener un sistema equivalente.

6. Sistema de ecuaciones lineales con igual número de ecuaciones e incógnitas

Sea el sistema de m ecuaciones lineales con n incógnita $A.X = B$ con $A \in \mathbb{K}^{m\times n}, B \in \mathbb{K}^{m\times 1}$ y X la matriz de las incógnitas de orden $n \times 1$; particularmente interesa estudiar el caso en donde el número de ecuaciones es igual al número de incógnitas ($m = n$), puesto que en ese caso se puede recurrir a los determinantes para determinar si el sistema tiene solución única o no. A estos sistemas se los suele llamar cuadrados.

Una forma de saber si un sistema de ecuaciones lineales con igual número de ecuaciones y de incógnitas tiene solución única, es tener presente el siguiente:

Teorema de la solución única. *Sea un sistema de ecuaciones lineales* $A.X = B$ *tal que* $A \in \mathbb{K}^{n\times n}$ *y cuya matriz ampliada es* $(A|B)$*, entonces el sistema tiene solución única si y solo si la matriz ampliada es equivalente, mediante operaciones elementales entre filas, a una matriz de la forma* $(I|B^*)$*. En ese caso la solución única está determinada por* $x_i = b_i^*$ *para* $i = 1, 2, \dots, n$.

$D \Longrightarrow$) Por hipótesis, mediante operaciones elementales entre filas obtenemos $(A|B) \sim (I|B^*)$, por lo tanto, la demostración es directa ya que los respectivos sistemas $(A|B)$ y $(I|B^*)$ también son equivalentes. En este caso, la solución única está dada por $x_i = b_i^*$ para $i = 1, 2, \dots, n$.

$D \Longleftarrow$) En este sentido, la demostración también resulta directa ya que las operaciones elementales entre filas, por construcción, son invertibles, entonces a partir del sistema $(I|B^*)$ y operaciones elementales entre filas se puede llegar al sistema $(A|B)$ que tendrá la misma solución que $(I|B^*)$, la cual es única.

Lo interesante del teorema anterior es que de él se deriva el teorema para saber cuándo un sistema con igual número de ecuaciones que de incógnitas, es compatible determinado, es decir, tiene solución única.

Teorema. *Un sistema de ecuaciones lineales* $A.X = B$ *con* $A \in \mathbb{K}^{n\times n}$ *y matriz ampliada* $(A|B)$ *tiene solución única si y solo si* $det(A) \neq 0$.

D) Por el teorema anterior se puede establecer que si el sistema tiene solución única, entonces los sistemas $(A|B)$ y $(I|B^*)$ son equivalentes. Por otro lado, se sabe que los determinantes de matrices que se obtienen por medio de operaciones elementales son iguales, multiplicados por una constante no nula, $(k \neq 0)$. Entonces se puede concluir que $|A| = k|I| = k.1 = k \neq 0$.

Por ejemplo, sea el sistema de ecuaciones de cuatro ecuaciones con cuatro incógnitas

$$\begin{cases} x - 3y + 6z + 8w = 35 \\ x - 2y + 6z + 6w = 26 \\ -x + 3y - 3z - 7w = -30 \\ 4y + 3z - 6w = -26 \end{cases}$$

Calculamos, por el método de triangulación de Gauss, el determinante de la matriz del sistema.

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

Por ser el determinante de la matriz del sistema distinto de cero, ser un sistema con igual número de ecuaciones que de incógnitas y por el teorema anterior, el sistema es compatible determinado, es decir, con solución única.

Continuamos con el cálculo de determinantes para obtener los valores de las incógnitas. De esta forma

$$\frac{\Delta_x}{\Delta} = \frac{-6}{3} = -2, \frac{\Delta_y}{\Delta} = \frac{3}{3} = 1, \frac{\Delta_z}{\Delta} = \frac{0}{3} = 0, \frac{\Delta_w}{\Delta} = \frac{15}{3} = 5$$

Previa verificación podemos decir que la solución del sistema es:

$S = \{(x, y, z, w)/x = -2, y = 1, z = 0, w = 5\}$

7. Sistema de ecuaciones lineales con parámetros

Los siguientes ejemplos ayudarán a afianzar más los distintos tipos de soluciones de un sistema de ecuaciones lineales.

Ejemplo 1: encontrar, de ser posible, los valores de k para que el siguiente sistema sea compatible determinado, compatible indeterminado e incompatible.

$$\begin{cases} x + 2y - 3z = 4 \\ 3x - y + 5z = 2 \\ 4x + y + (k^2 - 14)z = k + 2 \end{cases}$$

Primer procedimiento

Llevamos la matriz ampliada del sistema dado, a su forma escalonada.

$$\begin{pmatrix} 1 & 2 & -3 & 4 \\ 3 & -1 & 5 & 2 \\ 4 & 1 & k^2 - 14 & k + 2 \end{pmatrix} \sim_{\substack{-3F_1 + F_2 \\ -4F_1 + F_3}} \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & -7 & k^2 - 2 & k - 14 \end{pmatrix}$$

$$\sim_{-1F_2 + F_3} \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & 0 & k^2 - 16 & k - 4 \end{pmatrix}$$

En la tercera fila podemos observar que $k^2 - 16 = 0$ cuando $k = \pm 4$, entonces

Con $k = 4$ se obtiene:

$\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ lo que indica, analizando los rangos de las matrices A y el de la ampliada, que el sistema es compatible indeterminado.

Con $k = -4$ se obtiene:

$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & 0 & 0 & -8 \end{array}\right)$ lo que indica, analizando los rangos de las matrices A y el de la ampliada, que el sistema es incompatible.

Por lo tanto, si:

$k \neq \pm 4$ el sistema será compatible determinado

$k = 4$ el sistema será compatible indeterminado y

$k = -4$ el sistema será incompatible.

Segundo procedimiento

a) Calcular el determinante de la matriz del sistema

$$|A| = \begin{vmatrix} 1 & 2 & -3 \\ 3 & -1 & 5 \\ 4 & 1 & k^2 - 14 \end{vmatrix} = -7k^2 + 112$$

b) Resolver la ecuación, en términos de k, que resulta del cálculo del determinante.

$-7k^2 + 112 = 0 \implies k = \pm 4$.

c) Sustituir cada valor del parámetro en el sistema dado y determinar el tipo de solución.

Para $k = 4$

$$\begin{cases} x + 2y - 3z = 4 \\ 3x - y + 5z = 2 \\ 4x + y + (k^2 - 14)z = k + 2 \end{cases} \Longrightarrow \begin{cases} x + 2y - 3z = 4 \\ 3x - y + 5z = 2 \\ 4x + y + 2z = 6 \end{cases}$$

Entonces:

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 3 & -1 & 5 & 2 \\ 4 & 1 & 2 & 6 \end{array}\right) \sim \begin{matrix} \\ -3F_1 + F_2 \\ -4F_1 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & -7 & 14 & -10 \end{array}\right)$$

$$\underset{-1F_2+F_3}{\sim}\left(\begin{array}{ccc|c}1 & 2 & -3 & 4\\ 0 & -7 & 14 & -10\\ 0 & 0 & 0 & 0\end{array}\right)$$: Sistema compatible indeterminado.

Para $k=-4$

$$\begin{cases} x+2y-3z = 4\\ 3x-y+5z = 2\\ 4x+y+(k^2-14)z = k+2\end{cases} \Longrightarrow \begin{cases} x+2y-3z = 4\\ 3x-y+5z = 2\\ 4x+y+2z = -2\end{cases}$$

Entonces:

$$\left(\begin{array}{ccc|c}1 & 2 & -3 & 4\\ 3 & -1 & 5 & 2\\ 4 & 1 & 2 & -2\end{array}\right)\underset{-4F_1+F_3}{\sim -3F_1+F_2}\left(\begin{array}{ccc|c}1 & 2 & -3 & 4\\ 0 & -7 & 14 & -10\\ 0 & -7 & 14 & -18\end{array}\right)$$

$$\underset{-1F_2+F_3}{\sim}\left(\begin{array}{ccc|c}1 & 2 & -3 & 4\\ 0 & -7 & 14 & -10\\ 0 & 0 & 0 & -8\end{array}\right)$$: Sistema incompatible.

Finalmente se llega a las mismas conclusiones ya enunciadas respecto a los distintos valores de k y las posibles soluciones.

Ejemplo 2: encontrar, de ser posible, los valores de k para que el siguiente sistema sea compatible determinado, compatible indeterminado e incompatible.

$$\begin{cases}(5-k)x-2y-z = 1\\ -2x+(2-k)y-2z = 2\\ -x-2y+(5-k)z = 1\end{cases}$$

Primer procedimiento

Llevamos la matriz ampliada del sistema dada a su forma escalonada.

$$\left(\begin{array}{ccc|c}5-k & -2 & -1 & 1\\ -2 & 2-k & -2 & 2\\ -1 & -2 & 5-k & 1\end{array}\right)\underset{(5-k)F_1+F_3}{\sim -2F_1+F_2}\left(\begin{array}{ccc|c}-1 & -2 & 5-k & 1\\ -2 & 2-k & -2 & 2\\ 5-k & -2 & -1 & 1\end{array}\right)$$

$$\sim \underset{2F_2+F_3}{\begin{pmatrix} -1 & -2 & 5-k & 1 \\ 0 & 6-k & 2(k-6) & 0 \\ 0 & -2(6-k) & (k-4)(k-6) & 6-k \end{pmatrix}} \sim$$

$$\sim \left(\begin{array}{ccc|c} -1 & -2 & 5-k & 1 \\ 0 & 6-k & 2(k-6) & 0 \\ 0 & 0 & k(k-6) & 6-k \end{array}\right)$$

En la tercera fila podemos observar que $k(k-6)=0$ cuando $k=0 \vee k=6$, entonces:

Con $k=0$ se obtiene

$\left(\begin{array}{ccc|c} -1 & -2 & 5 & 1 \\ 0 & 6 & -12 & 0 \\ 0 & 0 & 0 & 6 \end{array}\right)$ lo que indica, analizando los rangos de las matrices A y el de la ampliada, que el sistema es incompatible.

Con $k=6$ se obtiene

$\left(\begin{array}{ccc|c} -1 & -2 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$ lo que indica, analizando los rangos de las matrices A y el de la ampliada, que el sistema es compatible indeterminado.

Por lo tanto, si:

$k \neq 0 \wedge k \neq 6$, el sistema será compatible determinado

$k=6$, el sistema será compatible indeterminado y

$k=0$, el sistema será incompatible.

Segundo procedimiento

a) Calcular el determinante de la matriz del sistema

$$|A| = \begin{vmatrix} 5-k & -2 & -1 \\ -2 & 2-k & -2 \\ -1 & -2 & 5-k \end{vmatrix} = k^3 - 12k + 36k$$

b) Resolver la ecuación, en términos de k, que resulta del cálculo del determinante.

$k^3 - 12k + 36k = 0 \Leftrightarrow k(k^2 - 12k + 36) \Leftrightarrow k(k-6)^2 = 0$. Por lo tanto, la ecuación se anulará para $k = 0 \vee k = 6$.

c) Sustituir cada valor del parámetro en el sistema dado y determinar el tipo de solución.

Para $k = 0$

$$\begin{cases} (5-k)x - 2y - z = 1 \\ -2x + (2-k)y - 2z = 2 \\ -x - 2y + (5-k)z = 1 \end{cases} \Longrightarrow \begin{cases} 5x - 2y - z = 1 \\ -2x + 2y - 2z = 2 \\ -x - 2y + 5z = 1 \end{cases}$$

Analizando la matriz escalonada de la matriz ampliada del sistema, se comprobará que para $k = 0$ el sistema es incompatible.

Para $k = 6$

$$\begin{cases} (5-k)x - 2y - z = 1 \\ -2x + (2-k)y - 2z = 2 \\ -x - 2y + (5-k)z = 1 \end{cases} \Longrightarrow \begin{cases} -x - 2y - z = 1 \\ -2x - 4y - 2z = 2 \\ -x - 2y - 1z = 1 \end{cases}$$

Analizando la matriz escalonada de la matriz ampliada del sistema, se comprobará que para $k = 6$ el sistema es compatible indeterminado.

Finalmente, se llega a las mismas conclusiones ya enunciadas respecto a los distintos valores de k y a las posibles soluciones.

8. Relación entre las soluciones de un sistema y sus matrices

A través del estudio de los sistemas de m ecuaciones lineales con n incógnitas se evidenció la relación que existe entre las soluciones, o no, del sistema con la matriz del sistema y la matriz ampliada. Las características de estas matrices, como ser: rango, filas o columnas nulas,

número de filas (ecuaciones) y columnas (incógnitas), etc. permiten determinar las soluciones, o no, del sistema.

En este apartado se pondrá en evidencia, brevemente, la relación que une las soluciones o no, de un sistema con las características de la matriz del sistema y la matriz ampliada.

8.1. Sistema con solución

Sea el sistema de m ecuaciones lineales con n incógnitas $A.X = B$ con $A \in \mathbb{K}^{m\times n}, B \in \mathbb{K}^{m\times 1}$ y X la matriz de las incógnitas de orden $n \times 1$, se dice que este sistema es compatible, tiene solución, cuando no tiene ecuaciones inconsistentes. La ausencia de estas ecuaciones se refleja en que tanto la forma escalonada de la matriz del sistema, A, como la matriz ampliada, A' o $(A|B)$, tienen el mismo número de filas no nulas o de columnas pivotales, lo que equivale a decir que $\rho(A) = \rho(A')$.

Por ejemplo, los siguientes esquemas en donde $\alpha_i \in \mathbb{R}$

$$\left(\begin{array}{cccc|c} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ 0 & 0 & 1 & \alpha_5 & \alpha_6 \\ 0 & 0 & 0 & 1 & \alpha_7 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{cccc|c} 0 & 1 & \alpha_1 & \alpha_2 & \alpha_3 \\ 0 & 0 & 1 & \alpha_4 & \alpha_5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

representan matrices ampliadas de dos sistemas de ecuaciones compatibles. Para el sistema asociado a la matriz de la izquierda $\rho(A) = \rho(A') = 3$, mientras que para el sistema asociado a la matriz ampliada de la derecha $\rho(A) = \rho(A') = 2$.

Ahora bien, en primer lugar, si además de cumplirse la igualdad de los rangos entre las matrices del sistema y la ampliada, este número coincide con el número de incógnitas que tiene el sistema, estamos en presencia,

como se estableció en el Teorema de Rouché-Frobenius y demás teoremas, de un sistema compatible determinado. O lo que es lo mismo:

a) el número de filas no nulas de la matriz del sistema, en su forma escalonada, es igual al número de incógnitas que tiene el sistema o

b) cada columna de la forma escalonada de la matriz del sistema tiene un pivote o

c) $\rho(A) = \rho(A') = n$ $(número\ de\ incógnitas\ del\ sistema)$.

Por ejemplo, los siguientes esquemas

$$\left(\begin{array}{cccc|c} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 \\ 0 & 0 & 1 & \alpha_8 & \alpha_9 \\ 0 & 0 & 0 & 1 & \alpha_{10} \end{array}\right) \qquad \left(\begin{array}{cccc|c} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ 0 & 1 & \alpha_5 & \alpha_6 & \alpha_7 \\ 0 & 0 & 1 & \alpha_8 & \alpha_9 \\ 0 & 0 & 0 & 1 & \alpha_{10} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

representan matrices ampliadas de dos sistemas de ecuaciones compatibles determinados.

En segundo lugar, si además de cumplirse la igualdad de los rangos entre las matrices del sistema y la ampliada, este número es menor que el número de incógnitas que tiene el sistema, estamos en presencia, como se estableció en el Teorema de Rouché-Frobenius y demás teoremas, de un sistema compatible indeterminado. O lo que, sintéticamente, es lo mismo $\rho(A) = \rho(A') < n$ $(número\ de\ incógnitas\ del\ sistema)$.

Esto significa que la forma escalonada de la matriz del sistema tiene columnas no pivotales, el número de estas columnas corresponde al número de variables libres (s) que posee el sistema y, como se dijo en el apartado tres, $s = n - \rho(A)$.

Por ejemplo, sean las matrices ampliadas, en la forma escalonada reducida de dos sistemas de ecuaciones lineales.

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Se tiene para la primera matriz $\rho(A) = \rho(A') = 3 < 4\ (número\ de\ incógnitas)$, por lo tanto, el sistema es compatible indeterminado. Tiene infinitas soluciones y la cantidad de variables libres es 1 (una). Reconstruimos el sistema:

$\begin{cases} x - y = -1 \\ z = 0 \\ w = 2 \end{cases}$ decidimos que la variable "y" es la variable libre, entonces

$\begin{cases} x = y - 1 \\ z = 0 \\ w = 2 \end{cases}$

La solución general del sistema es $S_G = \left\{\frac{(x,y,z,w)}{x} = y - 1, z = 0, w = 2, y \in \mathbb{R}\right\}$.

Para el sistema representado en la matriz de la derecha $\rho(A) = \rho(A') = 2 < 4\ (n)$, por lo tanto, el sistema es compatible indeterminado. Tiene infinitas soluciones y la cantidad de variables libres es 2 (dos). Reconstruimos el sistema:

$\begin{cases} y + w = 3 \\ z - 2w = 2 \end{cases}$ decidimos que las variables $x\ y\ w$ sean las variables libres, entonces $\begin{cases} y = 3 - w \\ z = 2w + 2 \end{cases}$

Por lo tanto $S_G = \{(x, y, z, w)/y = 3 - w, z = 2w + 2, x, w \in \mathbb{R}\}$.

8.2. Sistema sin solución

Sea el sistema de m ecuaciones lineales con n incógnita $A.X = B$ con $A \in \mathbb{K}^{m\times n}$, la matriz $B \in \mathbb{K}^{m\times 1}$ y X la matriz de las incógnitas de orden $n \times 1$, se dice que este sistema es incompatible, no tiene solución, cuando tiene ecuaciones inconsistentes, es decir ecuaciones de la forma $(0\ \ 0\ \ \ldots 0\ \ \ k)$ con $k \neq 0$.

Es muy sencillo reconocer ecuaciones inconsistentes en la forma escalonada de la matriz ampliada de un sistema. También se puede reconocer este tipo de ecuaciones cuando en el sistema hay dos ecuaciones con los mismos coeficientes de las respectivas incógnitas, pero con distintos términos independientes. Por ejemplo, en el siguiente sistema

$$\begin{cases} x + 5z - w = 1 \\ 2x - y + 5z + w = 3 \\ 3x + y - 2z + 3w = -1 \\ 2x - y + 5z + w = 6 \end{cases}$$

Podemos observar que la segunda y cuarta ecuación solo difieren en el segundo miembro. Si escalonamos la matriz ampliada del sistema se evidenciará la ecuación de inconsistencia que lo hace incompatible.

$$\left(\begin{array}{cccc|c} 1 & 0 & 5 & -1 & 1 \\ 2 & -1 & 5 & 1 & 3 \\ 3 & 1 & -2 & 3 & -1 \\ 2 & -1 & 5 & 1 & 6 \end{array}\right) \sim \begin{matrix} \\ -2F_1 + F_2 \\ -3F_1 + F_3 \\ -2F_1 + F_4 \end{matrix} \left(\begin{array}{cccc|c} 1 & 0 & 5 & -1 & 1 \\ 0 & -1 & -5 & 3 & 1 \\ 0 & 1 & -17 & 6 & -4 \\ 0 & -1 & -5 & 3 & 4 \end{array}\right)$$

$$\sim \begin{matrix} \\ F_2 + F_3 \\ -F_2 + F_4 \end{matrix} \left(\begin{array}{cccc|c} 1 & 0 & 5 & -1 & 1 \\ 0 & -1 & -5 & 3 & 1 \\ 0 & 0 & -22 & 9 & -3 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right)$$

No en todos los casos es tan simple reconocer ecuaciones inconsistentes, hay casos de dependencia entre las ecuaciones que las pueden hacer inconsistentes.

Sin embargo, existen otros tipos de dependencias entre las ecuaciones que las pueden hacer inconsistentes. El hecho de que un sistema tenga ecuaciones inconsistentes se relaciona con conceptos como el de rango, pero también con el de independencia lineal (que se estudiará en el próximo capítulo), pero, en todos los casos, la forma escalonada de la matriz ampliada del sistema hará evidente la existencia de, al menos, una ecuación inconsistente cuando el sistema no tenga solución.

Por ejemplo, en el siguiente esquema en donde $\alpha_i \in \mathbb{R}$ que representa la matriz ampliada de un sistema de ecuaciones, se puede observar que la tercera fila representa una ecuación inconsistente y hace que los rangos de la matriz del sistema y la matriz ampliada sean distintos. Más precisamente $2 = \rho(A) \neq 3 = \rho(A')$.

$$\left(\begin{array}{cccc|c} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ 0 & 0 & 1 & \alpha_5 & \alpha_6 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

En general $\rho(A) < \rho(A') \Leftrightarrow A.X = B$ tiene ecuaciones inconsistentes.

8.3. Sistema con menos ecuaciones que incógnitas

Si en un sistema de m ecuaciones lineales con n incógnita $A.X = B$ con $A \in \mathbb{K}^{m\times n}, B \in \mathbb{K}^{m\times 1}$ y X la matriz de las incógnitas de orden $m \times 1$, se cumple que $m < n$, el número de filas no nulas en la forma escalonada de la matriz ampliada del sistema será necesariamente menor al número de

incógnitas, luego un sistema de este tipo no puede tener solución única. Si, además, este tipo de sistema es homogéneo, necesariamente tiene un número infinito de soluciones.

CAPÍTULO 7

ESPACIOS VECTORIALES

1. Introducción

Jean Luc Dorier[37] determinó dos periodos unificadores en la evolución histórica del concepto de espacio vectorial y otros conectados con él –dependencia lineal, base, dimensión, etc.–. El primero de ellos giró en torno al concepto de determinantes e inicia a mediados del siglo XVIII con las publicaciones de Gabriel Cramer sobre este tema y culmina a principios del siglo XX con el desplazamiento de los determinantes por la formulación axiomática de la teoría de espacios vectoriales. En este periodo surgen –estudiando sistema de ecuaciones– conceptos fundamentales como la dependencia lineal y rango, aplicados a espacios de dimensión finita. De hecho, y para remarcar la importancia de los determinantes, estos eran una herramienta imprescindible para abordar cualquier problema de naturaleza lineal. La geometría jugó aquí un papel muy importante, pues no solo aportó el marco intuitivo de los descubrimientos algebraicos que desembocaron en lo que hoy se conoce como análisis vectorial, sino que, por otra parte, propició la búsqueda de un cálculo geométrico intrínseco que condujo al estudio de los vectores,

[37] Dorier, J. L. (1995). *A general outline of the genesis of vector space theory. Historia Mathematica*, 22(3), 227–261.

lo que sirvió como antecedentes a los primeros trabajos axiomáticos. Sin embargo, a inicios del siglo XX, conceptos tan importantes como el de rango de una matriz aún permanecían estrechamente vinculados al de determinantes. Prueba de esto es el breve artículo que publicó George A. Miller[38] en el cual demuestra un teorema sobre sistemas compatibles de ecuaciones lineales utilizando la definición de rango de una matriz por medio de determinantes.

El segundo periodo se extiende desde la publicación del primer conjunto de axiomas para n–tuplas dado por Darboux, en 1875, y culmina con la publicación del texto de Birkhoff y MacLane, en 1941. En este lapso, van perdiendo el papel preponderante que tenían los determinantes y se da un proceso de refinamiento y aceptación de la aproximación axiomática a la teoría de espacios vectoriales debido al impulso como consecuencia de la necesidad de unificar el gran número de nuevas teorías que se expandían rápidamente (Moore, 1995)[39].

Al final de este periodo se logró unificar bajo un conjunto de axiomas los espacios vectoriales de dimensión finita e infinita, estableciendo un precedente importante para la creación del Álgebra Lineal como campo de estudio independiente que agrupa muchos fenómenos lineales de

[38] Miller, G. A. (1910). *On the Solution of a System of Linear Equations, The American Mathematical Monthly*, 17: 6-7, 137-139.

[39] Moore, G. H. (1995). *The axiomatization of linear algebra: 1875-1940. Historia Mathematica*, 22(3), 262-303.

distintas partes de la matemática. Es debido a estas características en su gestación y su evolución que Jean Luc Dorier ha incluido el concepto de Espacios Vectoriales junto al de Grupo en una categoría de conceptos que ha denominado unificador y generalizador, puesto que realizaron estas dos acciones con diferentes métodos, herramientas y objetos matemáticos que se encontraban dispersos en contextos como la geometría, los sistemas de ecuaciones lineales, entre otros. En otras palabras, estos conceptos poseen tres características básicas: no fueron creados para resolver nuevos problemas –como la integral de Riemann–, unifican conceptos y métodos ya existentes y reorganizan en una nueva teoría resultados dispersos.

En este capítulo, antes de definir espacios vectoriales, repasamos muy brevemente los axiomas de cuerpo de los números reales y generalizamos el concepto de las operaciones: suma y producto por un escalar. A partir de esto, dedicamos un apartado para la definición de espacios vectoriales, diversos ejemplos y algunos postulados que se demuestran directamente de la definición precedente. El apartado tres está destinado a la definición de subespacio vectorial y a la demostración de una condición necesaria y suficiente para determinar si un conjunto de vectores es o no subespacio vectorial.

En el apartado cuatro, bajo la idea de que en cualquier espacio vectorial se pueden combinar sus elementos mediante la suma y el producto por un escalar, se define –de dos maneras distintas– combinación lineal de

vectores y, a partir de esto, el concepto de subespacio generado. El concepto de combinación lineal nos permite analizar la dependencia lineal de un conjunto de vectores, para lo cual estudiamos –en el apartado siguiente– vectores linealmente independientes y dependientes con su respectiva interpretación geométrica y propiedades.

Una de las ideas más importantes en espacios vectoriales –Sistema de Generadores– se estudia en el apartado seis, mientras que, el siguiente está dedicado a la idea que reúne los conceptos de independencia lineal y sistema de generadores, base de un espacio vectorial, que está directamente relacionado con coordenadas de un vector y dimensión de un espacio vectorial.

1.1. Conceptos básicos

El primer paso para iniciar el estudio de espacios vectoriales es comprender la estructura y las principales propiedades del conjunto de números reales. Presentaremos este conjunto sin dar una definición concreta de número real ya que lo importante –en esta asignatura– no es saber qué es un número real, sino qué propiedades tiene el conjunto de ellos.

Enunciaremos estas propiedades como postulados básicos elementales: axiomas[40], y solamente aquellos que están asociados a la igualdad de números reales que son conocidos como axiomas de cuerpo.

En el conjunto de números reales ($\mathbb{R}$) se dispone de dos operaciones; una llamada suma, simbolizada por " + " y que hace corresponder a cada par, (x, y), de números reales un único número real, la suma de x con y, denotada por $x + y$; la otra operación llamada producto, simbolizada por " $\cdot$ " y que hace corresponder a cada par (x, y) de números reales un único número real, el producto de x con y, denotado por $x.y$ o simplemente xy.

Axiomas de Cuerpo de los Reales

Los axiomas del $\mathbb{R}$ sobre la igualdad son, en muchos casos, llamados axiomas de cuerpo de los reales y pueden agruparse de la siguiente manera:

A_1) Conmutatividad

i) La suma de cualquier par de reales, $x\ e\ y$ es un real independiente del orden de los sumandos. $\forall\ x, y \in \mathbb{R}$: $x + y = y + x$.

ii) El producto de cualquier par de reales, $x\ e\ y$ es un real independiente del orden de los factores. $\forall\ x, y \in \mathbb{R}$: $x\ y = y\ x$.

A_2) Asociatividad

[40] Un axioma es un enunciado que sirve como punto de inicio y del cual se derivan, lógicamente, otros enunciados. Se consideran una verdad evidente que no necesita demostración alguna.

i) La suma de cualquier terna de reales, x, y, z es independiente de la forma en la que se pueden asociar los sumandos. $\forall\, x, y, z \in \mathbb{R}: x + (y + z) = (x + y) + z$.

ii) El producto de cualquier terna de reales, x, y, z es independiente de la forma en la que se pueden asociar los factores. $\forall\, x, y, z \in \mathbb{R}: x\,(y\,z) = (x\,y)\,z$.

A_3) Distributividad del producto respecto a la suma

i) El producto de un real x por la suma de otros dos, y, z, es equivalente a la suma de los productos del factor por cada sumando. $\forall\, x, y, z \in \mathbb{R}: x\,(y + z) = x\,y + x\,z$.

Muchas veces suele admitirse como una segunda parte de este axioma el siguiente enunciado ii) $\forall\, x, y, z \in \mathbb{R}: (x + y)z = x\,z + y\,z$. Si bien no hay problemas en aceptar esto y utilizar libremente ambas expresiones en cualquier demostración, es necesario destacar que ii) es una consecuencia de i) y el axioma de la conmutatividad del producto. Es decir que se puede demostrar que $\forall\, x, y, z \in \mathbb{R}: (x + y)\,z = x\,z + y\,z$.

$(x + y)\,z = z\,(x + y)$ por conmutatividad del producto.

$(x + y)\,z = z\,x + z\,y$ por distributividad.

$(x + y)\,z = x\,z + y\,z$ por conmutatividad del producto. Con lo que queda demostrada la propiedad.

Este ejemplo permite establecer que existen afirmaciones, expresiones y/o propiedades que pueden demostrarse utilizando axiomas y/u otras

propiedades ya estudiadas. A estas afirmaciones más la correspondiente demostración se las suele conocer como teoremas.

Los siguientes axiomas permiten afirmar la existencia de ciertos elementos especiales en el $\mathbb{R}$ y tienen como consecuencia directa la confirmación de que el conjunto $\mathbb{R}$ es no vacío.

A_4) Existencia de elementos neutros

i) Para la suma.

En $\mathbb{R}$ existe cierto número denotado por la letra e que no afecta el resultado de la operación suma. $\forall\, x \in \mathbb{R}: \exists\, e \in \mathbb{R} / x + e = x$.

Todo elemento e que cumpla con esta propiedad se dirá neutro para la suma.

Destaquemos que este axioma solo garantiza la existencia del elemento neutro, pero no nos indica nada respecto de cuántos hay. Sabemos que hay un único elemento neutro para la suma en $\mathbb{R}$, esa afirmación se puede demostrar usando los axiomas y/o propiedades ya estudiadas.

Se deja para el estudiante la demostración del siguiente:

Teorema: El elemento neutro para la suma en $\mathbb{R}$ es único[41].

Una vez demostrado el teorema, se suele denominar al único neutro aditivo "*cero*" y denotarlo por "*0*".

ii) Para el producto.

[41] Se recomienda repasar los procedimientos utilizados en la propiedad P_4) de la suma en $\mathbb{R}^n$ y en $\mathbb{K}^{m \times n}$.

En $\mathbb{R}$ existe cierto número denotado por la letra e que no afecta el resultado de la operación producto. $\forall\, x \in \mathbb{R}: \exists\, e \in \mathbb{R} / \, x\, e = x$.

Todo elemento e que cumpla con esta propiedad se dirá neutro para el producto.

De manera similar al caso de la suma, este axioma solo garantiza la existencia del elemento neutro, pero sabemos que hay un único elemento neutro para esta operación en $\mathbb{R}$. Nuevamente esta afirmación se puede demostrar –de manera análoga a la suma– usando los axiomas y/o propiedades ya estudiadas.

Se deja para el estudiante la demostración del siguiente:

Teorema: El elemento neutro para el producto en $\mathbb{R}$ es único.

Una vez demostrado el teorema, se suele denominar al único neutro multiplicativo "*uno*" y denotarlo por "*1*".

A_5) Existencia de elementos inversos

i) Para la suma.

Para cada número real x existe otro real asociado a él, llamado opuesto o inverso aditivo tal que al sumarlos se obtiene el elemento neutro de esta operación. Es decir $\forall\, x \in \mathbb{R}: \exists\, o \in \mathbb{R} / x + o = 0$.

Sea el siguiente:

Teorema: El inverso aditivo para cualquier número real x, es único.

Si bien este teorema se puede demostrar con un procedimiento análogo al utilizado en las propiedades P_5) de la suma en $\mathbb{R}^n$ y de la suma en $\mathbb{K}^{m \times n}$, se brindará a continuación una demostración alternativa que sigue la

misma idea (proponer dos inversos aditivos distintos para un mismo real y demostrar que son iguales), pero utiliza los axiomas que estamos estudiando.

D) Sean los inversos aditivos $o_1\ y\ o_2$ del mismo real arbitrario x, estos satisfacen las ecuaciones:

$x + o_1 = 0$ (1) y

$x + o_2 = 0$ (2)

Debemos probar que $o_1 = o_2$. En efecto haciendo uso de los axiomas estudiados –en este caso particular de la suma– y las ecuaciones anteriores, tenemos:

$o_1 = o_1 + 0$ por el elemento neutro

$o_1 = o_1 + (x + o_2)$ por la ecuación (2)

$o_1 = (o_1 + x) + o_2$ por asociatividad

$o_1 = (x + o_1) + o_2$ por conmutatividad

$o_1 = 0 + o_2$ por la ecuación (1)

$o_1 = o_2 + 0$ por la conmutatividad

$o_1 = o_2$ por el elemento neutro. Con lo que queda demostrado que $o_1 = o_2$ y la unicidad del inverso aditivo de la suma en $\mathbb{R}$. El inverso aditivo del real x suele simbolizarse generalmente como $-x$.

ii) Para el producto

Para cada número real x no nulo existe otro real asociado a él, llamado inverso multiplicativo o recíproco tal que al multiplicarlos se obtiene el elemento neutro de esta operación. Es decir $\forall\ x \in\ \mathbb{R}: \exists\ o \in \mathbb{R}/x\ o = 1$.

Se deja para el estudiante la demostración del siguiente teorema, haciendo uso de los axiomas estudiados.

Teorema: El inverso multiplicativo para cualquier número real x, es único.

El inverso multiplicativo de un real x, suele simbolizarse como x^{-1}.

Por las operaciones de suma, producto en $\mathbb{R}$ y los axiomas enunciados anteriormente se dice que este conjunto forma un Cuerpo. Es decir, como se afirmó en el capítulo de números complejos, un cuerpo o campo, es un conjunto K con dos operaciones binarias, usualmente llamadas suma "+" y producto "." que son asociativa, conmutativa y sus elementos neutros son 0 y 1, respectivamente. Además, para cada elemento –en el caso del producto distintos de cero– existe un inverso. Finalmente, el producto es distributivo respecto de la suma.

De manera similar se puede demostrar que el conjunto de números complejos con las operaciones suma y producto forman un Cuerpo que, por lo general, se denota como $\mathbb{C}$. Como existen definiciones, teoremas y propiedades en las cuales es indistinto trabajar con el cuerpo de los reales o de los complejos, se suele denotar, en estos casos, ambos cuerpos como $\mathbb{K}$.

Recordando lo enunciado en el inicio del apartado, la segunda cuestión que se debe tener presente para comprender mejor este capítulo referido a espacios vectoriales, es tener presente que al estudiar vectores en $\mathbb{R}^n$ se definieron dos operaciones: suma de vectores y producto de un vector por

un escalar, luego se enunciaron y demostraron diez propiedades que poseen estas dos operaciones. Situación similar ocurrió al estudiar matrices en $\mathbb{K}^{m\,x\,n}$, se definieron dos operaciones: suma de matrices y producto de una matriz por un escalar e inmediatamente se enunciaron y demostraron las diez propiedades que cumplen estas dos operaciones.

Existen –como se estudió en capítulos anteriores– otros conjuntos a los que se les puede observar una estructura similar (definir en ellos dos operaciones y establecer las diez propiedades de las mismas) a la de los vectores en $\mathbb{R}^n$ y las matrices en $\mathbb{K}^{m\,x\,n}$: los números complejos y los polinomios.

Cuando en varios conjuntos distintos aparecen estructuras similares, lo más conveniente es axiomatizar[42] estas, generalizar y dar un nombre al objeto que se define a través del conjunto de propiedades. La principal ventaja que se obtiene es que estudiando esta estructura quedan estudiadas todas las estructuras particulares que ella encierra. Específicamente, cuando en un conjunto (vectores en $\mathbb{R}^n$, matrices en $\mathbb{K}^{m\,x\,n}$, números complejos, polinomios, funciones, entre otros tantos) se da una estructura similar se dice que este conjunto tiene estructura de espacio vectorial.

Generalizar significa, en nuestro caso, construir una estructura o teoría que englobe hechos que ya conocemos, tiene que ver con economía de

[42] Determinar los axiomas sobre los cuales se asienta una teoría.

trabajo y con determinar cuáles son los elementos mínimos responsables de que ciertos resultados ocurran.

Por lo tanto, el objetivo que perseguimos en este capítulo es el de introducir aquella estructura abstracta (espacios vectoriales) que englobe las anteriores construcciones (vectores en $\mathbb{R}^n$, matrices en $\mathbb{K}^{m\times n}$, polinomios, números complejos) y ver qué resultados se pueden obtener en lo general sin importar a cuál de estas estructuras específicas se haga referencia.

Previo a definir espacios vectoriales, veamos –por medio de ejemplos– que el concepto de las operaciones suma y producto por un escalar podrían ser definidos de manera distinta a la que generalmente conocemos.

Ejemplo 1: supongamos que tenemos el conjunto $V = R^2$ y se define la operación:

$$(x, y) \oplus (z, w) = (2x + z, 3w + y)$$

Si $u = (1, -1),\ v = (-2, 2)\ y\ \ t = (0, 2)$, podría calcularse

$a)\ u \oplus v = (1, -1) \oplus (-2, 2) = \big(2.1 + (-2);\ 3.2 + (-1)\big) = (0, -5)$

$b)\ v\ \oplus\ \ u = (-2, 2)\ \ \oplus\ \ (1, -1) = (2(-2) + 1;\ 3(-1) + 2) = (-3, -1)$

$c)\ (u \oplus v) \oplus t = (0, -5) \oplus (0, 2) = (2.0 + 0; 3.2 + (-5)) = (0, 1)$

Ejemplo 2: supongamos que tenemos el conjunto $V = R^2$ y se definen las operaciones:

$(x, y) \oplus (z, w) = (-x + 2w;\ z + y)$ y $\lambda \odot (x, y) = (2\lambda x, -\lambda y)$

Si $u = (0, 1),\ \lambda_1 = 2\ y\ \ \lambda_2 = -3$, podría calcularse

$a)\ (\lambda_1 + \lambda_2) \odot u = (2-3) \odot u = -1 \odot (0,1) = (2(-1)0;\ -(-1)1) = (0,1)$

$b)\ (\lambda_1 \odot u)\ \oplus (\lambda_2\ \odot u) = (2\ \odot (0{,}1)) \oplus (-3 \odot (0,1)) =$

$(2.2.0;\ -2.1) \oplus (2(-3)0;\ -(-3)1)$

$= (0,-2)\ \oplus (0,3) = (-0+2.3;\ 0+(-2)) = (6,-2)$

$c)\ (\lambda_1\ \lambda_2) \odot u = -6\ \odot\ (0,1) = (2(-6)0;\ -(-6)1) = (0,6)$

Teniendo presente el concepto de cuerpo de números reales y complejos y la generalización de las operaciones, estamos en condición de estudiar lo que es un espacio vectorial, en conjunto con otros conceptos estrechamente ligados a él como: subespacio vectorial, dependencia lineal, base, dimensión, etc.

2. Espacio Vectorial

Es momento de definir esa estructura abstracta que engloba lo ya conocido y estudiado por nosotros y cuyos resultados son generales, de tal forma que al estudiar otros conjuntos que también tengan la misma estructura sean igualmente válidos.

Iniciamos con la siguiente:

Definición: Sean V un conjunto no vacío y K un cuerpo de escalares, " $+$ " y " $\cdot$ " dos operaciones que llamaremos suma y producto respectivamente; el objeto $(V,+,K,\cdot)$ es un espacio vectorial si y solo si se verifican los siguientes axiomas:

A_1) La suma es Ley de Composición interna en V

$\forall\ x,y \in V \Longrightarrow (x+y) \in V$

A_2) La suma es asociativa en V

$\forall\, x, y, z \in V \Longrightarrow (x + y) + z = x + (y + z)$

A3) La suma es conmutativa en V

$\forall\, x, y \in V \Longrightarrow x + y = y + x$

A4) Existe un único neutro para la suma en V.

El elemento neutro se denota con 0

$\exists \mathbf{0} \in V\ \acute{u}nico\ / \forall x \in V : x + \mathbf{0} = x$

A5) Todo elemento en V admite inverso aditivo u opuesto en V

El elemento opuesto a uno cualquiera x, se nota por $-x$.

$\forall x \in V, \exists (-x) \in V / x + (-x) = \mathbf{0}$

A6) El producto es Ley de Composición Externa en V con escalares en K

$\forall \alpha \in K, \forall x \in V \Longrightarrow \alpha x \in V$

A7) El producto satisface la asociatividad mixta

$\forall \alpha, \beta \in K, \forall x \in V : \alpha(\beta x) = (\alpha\beta)x$

A8) El producto es distributivo respecto a la suma en K

$\forall \alpha, \beta \in K, \forall x \in V : (\alpha + \beta)x = \alpha x + \beta x$

A9) El producto es distributivo respecto a la suma en V

$\forall \alpha \in K, \forall x, y \in V : \alpha(x + y) = \alpha x + \alpha y$

A10) La unidad para el cuerpo de escalares es el neutro para el producto

La unidad del cuerpo se denota como 1.

$\forall x \in V : 1x = x$

Se conviene en llamar genéricamente "vectores" a los elementos de los espacios vectoriales, por lo tanto, a dichos elementos se los representa con una supraflecha.

Si $K = \mathbb{R}$ el espacio vectorial se dice espacio vectorial real, mientras que si $K = \mathbb{C}$ el espacio vectorial se dice complejo.

Se deja para los estudiantes deducir, a partir de los axiomas A_4), A_5) y de la conmutatividad de la suma, que $\overrightarrow{0} + \overrightarrow{x} = \overrightarrow{x}$ y $-\overrightarrow{x} + \overrightarrow{x} = \overrightarrow{0}$.

Ejemplos:

a) La cuaterna $(\mathbb{R}^n, +, \mathbb{R}, \cdot)$ con la suma y el producto por un escalar, tal como se definió en el correspondiente capítulo (vectores), es un espacio vectorial real. En general, $(\mathbb{K}^n, +, \mathbb{K}, \cdot)$ con la suma y el producto por un escalar usual es un espacio vectorial.

b) La cuaterna $(\mathbb{K}^{m \times n}, +, \mathbb{R}, \cdot)$ con la suma y el producto por un escalar, tal como se definió en el correspondiente capítulo (matrices), es un espacio vectorial real.

c) Particularmente $(\mathbb{K}^2, +, \mathbb{R}, \cdot)$ con la suma y el producto por un escalar usual, es espacio vectorial.

d) $(\mathbb{C}, +, \mathbb{C}, \cdot)$ con la suma usual de números complejos y el producto usual de números complejos por un escalar, es un espacio vectorial complejo.

e) $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$ con la suma usual de ternas ordenadas de número reales (vectores en $\mathbb{R}^3$) y el producto usual de una terna por un número real, es un espacio vectorial real.

f) $(P(x), +, \mathbb{K}, \cdot)$ donde $P(x)$ es el conjunto de polinomios de coeficientes reales, variable x y de grado menor o igual a n, con la suma usual de polinomios y el producto usual de un polinomio por un número real, es un espacio vectorial.

g) $(\mathbb{R}^{m\times n}, +, \mathbb{K}, \cdot)$ donde $\mathbb{R}^{m\times n}$ es el conjunto de matrices, de orden $m \times n$, de números reales con la suma usual de matrices y el producto usual de una matriz por un número real, es un espacio vectorial.

h) $(\mathbb{R}^2, +, \mathbb{C}, \cdot)$ no es espacio vectorial. Si $\overrightarrow{x} = (2,3)$ y $\alpha = 2 + i$ entonces al realizar el producto $\alpha\overrightarrow{x}$, obtenemos $\alpha\overrightarrow{x} = (4 + 2i, 6 + 3i) \notin \mathbb{R}^2$

i) $(\mathbb{Z}^{2x2}, +, \mathbb{R}, \cdot)$ no es espacio vectorial. Si $\overrightarrow{x} = \begin{pmatrix} 2 & -1 \\ 4 & -6 \end{pmatrix}$ y $\alpha = 1{,}2 \Longrightarrow$ $\alpha\overrightarrow{x} \notin \mathbb{Z}^{2x2}$

j) $(\mathbb{R}^{n\times n}, +, \mathbb{K}, \cdot)$ donde $\mathbb{R}^{n\times n}$ tiene determinante nulo, con las operaciones suma de matrices y producto de una matriz por un escalar usuales, no es un espacio vectorial. Basta con proponer

$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ matrices cuadradas tales que $det(A) = det(B) = 0$. Pero se puede comprobar fácilmente que $\det(A + B) \neq 0$.

A continuación, se demostrarán algunos postulados que se deducen directamente de los axiomas que definen un espacio vectorial.

Sea $(V, +, K, \cdot)$ un espacio vectorial, se cumple:

P_1) El producto del escalar 0 por cualquier vector, es el vector nulo.

$0.\overrightarrow{x} = \overrightarrow{0} \ \ \forall\overrightarrow{x} \in V \wedge 0 \in K$

D) $\alpha\overrightarrow{x} = (\alpha + 0)\overrightarrow{x} = \alpha\overrightarrow{x} + 0\overrightarrow{x}$ (1)

$\alpha\overrightarrow{x} = \alpha\overrightarrow{x} + \overrightarrow{0}$ (2)

Por lo tanto, de (1) y (2)

$\alpha\overrightarrow{x} + 0\overrightarrow{x} = \alpha\overrightarrow{x} + \overrightarrow{0}$

$-\alpha\overrightarrow{x} + \alpha\overrightarrow{x} + 0\overrightarrow{x} = -\alpha\overrightarrow{x} + \alpha\overrightarrow{x} + \overrightarrow{0}$

$0\overrightarrow{x} = \overrightarrow{0}$

P2) El producto de cualquier escalar por el vector nulo, es el vector nulo

$\alpha.\overrightarrow{0} = \overrightarrow{0} \ \ \forall \alpha \in K$

D) $\alpha\overrightarrow{x} = \alpha\left(\overrightarrow{x} + \overrightarrow{0}\right)$

$\alpha\overrightarrow{x} = \alpha\overrightarrow{x} + \alpha\overrightarrow{0}$

$\alpha\overrightarrow{x} + \overrightarrow{0} = \alpha\overrightarrow{x} + \alpha\overrightarrow{0}$

$\alpha\overrightarrow{0} = \overrightarrow{0}$

P3) Si el producto de un escalar por un vector es el vector nulo, entonces el escalar es 0 o el vector es nulo.

$\alpha\overrightarrow{x} = \overrightarrow{0} \Longrightarrow \alpha = 0 \ \vee \ \overrightarrow{x} = \overrightarrow{0}$

Se presentan dos posibilidades: $i)$ $\alpha = 0$ o bien $ii)$ $\alpha \neq 0$.

En el primer caso, $\alpha = 0$, es verdadera la primera proposición de la disyunción que figura en la tesis y, por consiguiente, esta es verdadera.

En el segundo caso, $\alpha \neq 0$, es necesario probar que $\overrightarrow{x} = \overrightarrow{0}$.

Primeramente, dejemos establecido que como $\alpha \neq 0$ entonces existe $\alpha^{-1} \ \in K$.

$\alpha\overrightarrow{x} = \overrightarrow{0}$

$\alpha^{-1}\,\alpha\overrightarrow{x} = \alpha^{-1}\,\overrightarrow{0}$

$(\alpha^{-1}\,\alpha)\overrightarrow{x} = \alpha^{-1}\,\overrightarrow{0}$

$1\,\overrightarrow{x} = \ \overrightarrow{0}$

$\overrightarrow{x} = \ \overrightarrow{0}$

P4) El opuesto de cualquier escalar por un vector es igual al opuesto de su producto

$(-\alpha)\overrightarrow{x} = -(\alpha\overrightarrow{x})$

D) $-(\alpha\overrightarrow{x}) + \alpha\overrightarrow{x} = \overrightarrow{0} \ \ (1)$

$\overrightarrow{0} = 0.\overrightarrow{x} = (-\alpha + \alpha)\overrightarrow{x} = (-\alpha)\overrightarrow{x} + \alpha\overrightarrow{x}$

$\overrightarrow{0} = (-\alpha)\overrightarrow{x} + \alpha\overrightarrow{x}$ (2)

Por lo tanto, de (1) y (2)

$-(\alpha\overrightarrow{x}) + \alpha\overrightarrow{x} = (-\alpha)\overrightarrow{x} + \alpha\overrightarrow{x}$

$(-\alpha)\overrightarrow{x} = -(\alpha\overrightarrow{x})$

Particularmente, tenemos

$(-1)\overrightarrow{x} = -(1\overrightarrow{x}) = -\overrightarrow{x}$

P_5) Ley cancelativa para la suma en V

$\forall\, \overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z} \in V: \overrightarrow{x} + \overrightarrow{y} = \overrightarrow{z} + \overrightarrow{y} \Longrightarrow \overrightarrow{x} = \overrightarrow{z}$

D) $\overrightarrow{x} + \overrightarrow{y} = \overrightarrow{z} + \overrightarrow{y}$

$\overrightarrow{x} + \overrightarrow{y} + (-\overrightarrow{y}) = \overrightarrow{z} + \overrightarrow{y} + (-\overrightarrow{y})$

$\overrightarrow{x} + \left(\overrightarrow{y} + (-\overrightarrow{y})\right) = \overrightarrow{z} + \left(\overrightarrow{y} + (-\overrightarrow{y})\right)$

$\overrightarrow{x} + \overrightarrow{0} = \overrightarrow{z} + \overrightarrow{0}$

$\overrightarrow{x} = \overrightarrow{z}$

P_6) El elemento neutro de la suma en V es único

D) Suponemos que existen dos neutros en V, $\overrightarrow{0_1}$ y $\overrightarrow{0_2}$, entonces

Si $\overrightarrow{0_1}$ es neutro, se cumple que $\overrightarrow{0_1} + \overrightarrow{0_2} = \overrightarrow{0_2}$

Si $\overrightarrow{0_2}$ es neutro, se cumple que $\overrightarrow{0_2} + \overrightarrow{0_1} = \overrightarrow{0_1}$

De estas dos expresiones se deduce que $\overrightarrow{0_1} = \overrightarrow{0_2}$ y, en consecuencia, el neutro para la suma en V es único.

P_7) El inverso aditivo u opuesto de un elemento de V es único

D) Suponemos que existen dos opuestos para $\overrightarrow{v}$: $\overrightarrow{v_1}$ y $\overrightarrow{v_2}$ entonces

Si $\overrightarrow{v_1}$ es opuesto, se cumple que $\overrightarrow{v} + \overrightarrow{v_1} = \overrightarrow{0}$

Si $\overrightarrow{v_2}$ es opuesto, se cumple que $\overrightarrow{v} + \overrightarrow{v_2} = \overrightarrow{0}$

De estas dos expresiones $\overrightarrow{v} + \overrightarrow{v_1} = \overrightarrow{v} + \overrightarrow{v_2} \implies \overrightarrow{v_1} = \overrightarrow{v_2}$

En consecuencia, el inverso aditivo u opuesto de un elemento de V es único

3. Subespacio Vectorial

Algunos subconjuntos de un espacio vectorial V, son por sí mismos espacios vectoriales con las mismas operaciones definidas en V; en estos casos, estos subconjuntos reciben el nombre de subespacios vectoriales.

Definición: sea el espacio vectorial $(V, +, K, \cdot)$ y S un subconjunto no vacío de V, es decir $S \subset V \; y \; S \neq \emptyset$, S es subespacio de $(V, +, K, \cdot)$ si y solo si $(S, +, K, \cdot)$ es espacio vectorial en sí mismo, respecto de las mismas operaciones definidas en V.

Cualquiera que sea $(V, +, K, \cdot)$, tanto V como $\{ \overrightarrow{0} \}$ son subespacios de V, llamados triviales.

Ejemplos: consideremos el espacio vectorial $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$ con la suma y el producto usual y los subconjuntos:

$a)\; T = \{(x, y) \in \mathbb{R}^2 / y = x + 1\}$

T no es un subespacio vectorial ya que el vector nulo $(0, 0) \notin T$, puesto que si $x = y = 0$ entonces $0 = 1$, lo que no es posible. Por lo tanto, la recta $y = x + 1$ de $\mathbb{R}^2$, no es subespacio vectorial.

$b)\; S = \{(x, y) \in \mathbb{R}^2 / y = x\}$

S si es subespacio vectorial, debiendo demostrar –solamente a modo de ejemplo y para completar el apartado 2– que se cumplen los diez axiomas de la definición de espacios vectoriales.

Sean los vectores de S, $x = (x_1, x_2) \in S \implies x_1 = x_2$, $y = (y_1, y_2) \in S \implies y_1 = y_2$, $z = (z_1, z_2) \in S \implies z_1 = z_2$ y $\alpha, \beta \in \mathbb{R}$ escalares.

A1) La suma es Ley de Composición interna en S

$x = (x_1, x_2) \in S \implies x_1 = x_2$

$y = (y_1, y_2) \in S \implies y_1 = y_2$ sumando miembro a miembro

$x + y = (x_1 + y_1; x_2 + y_2) \implies x_1 + y_1 = x_2 + y_2$

Por lo tanto $(x + y) \in S$.

A2) La suma es asociativa en S

$(x + y) + z = (x_1 + y_1; x_2 + y_2) + (z_1, z_2)$

$(x + y) + z = (x_1 + y_1 + z_1; x_2 + y_2 + z_2)$

$(x + y) + z = (x_1 + (y_1 + z_1)\,; x_2 + (y_2 + z_2))$

$(x + y) + z = (x_1, x_2) + (y_1 + z_1; y_2 + z_2)$

$(x + y) + z = x + (y + z)$

A3) La suma es conmutativa en S

$x + y = (x_1 + y_1; x_2 + y_2)$

$x + y = (y_1 + x_1; y_2 + x_2)$

$x + y = y + x$

A4) Existe un neutro para la suma en S.

El neutro para la suma en S es $(0, 0)$ puesto que

i) $(0, 0) \in S$ ya que $0 = 0$ y

ii) si $x = (x_1, x_2) \in S$, entonces

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

A5) Todo elemento en S admite inverso aditivo u opuesto en S

El inverso aditivo para $x = (x_1, x_2) \in S$ es $-x = (-x_1, -x_2)$ puesto que

i) $(-x_1, -x_2) \in S$ debido a que $x_1 = x_2 \implies -x_1 = -x_2$, además

ii) $(x_1, x_2) + (-x_1, -x_2) = (0, 0)$.

A6) El producto es Ley de Composición Externa en S con escalares en $\mathbb{R}$

Como $\alpha \in \mathbb{R}$ entonces

$x = (x_1, x_2) \in S \implies x_1 = x_2$

$\alpha = \alpha$ multiplicando miembro a miembro

$\alpha\, x = (\alpha x_1, \alpha x_2) \Longrightarrow \alpha x_1 = \alpha x_2$ por que $x_1 = x_2$

A7) El producto satisface la asociatividad mixta

$\alpha(\beta x) = \alpha(\beta x_1, \beta x_2)$

$\alpha(\beta x) = (\alpha\beta x_1, \alpha\beta x_2)$

$\alpha(\beta x) = ((\alpha\beta)x_1, (\alpha\beta)x_2)$

$\alpha(\beta x) = (\alpha\beta)(x_1, x_2)$

$\alpha(\beta x) = (\alpha\beta)x$

A8) El producto es distributivo respecto a la suma en $\mathbb{R}$

$(\alpha + \beta)x = (\alpha + \beta)(x_1, x_2)$

$(\alpha + \beta)x = ((\alpha + \beta)x_1, (\alpha + \beta)x_2)$

$(\alpha + \beta)x = (\alpha x_1 + \beta x_1;\ \alpha x_2 + \beta x_2)$

$(\alpha + \beta)x = (\alpha x_1, \alpha x_2) + (\beta x_1;\ \beta x_2)$

$(\alpha + \beta)x = \alpha(x_1, x_2) + \beta(x_1;\ x_2)$

$(\alpha + \beta)x = \alpha x + \beta x$

A9) El producto es distributivo respecto a la suma en S

$\alpha(x + y) = \alpha((x_1, x_2) + (y_1, y_2))$

$\alpha(x + y) = \alpha(x_1 + y_1; x_2 + y_2)$

$\alpha(x+y) = (\alpha(x_1 + y_1); \alpha(x_2 + y_2))$

$\alpha(x+y) = (\alpha\, x_1 + \alpha\, y_1; \alpha\, x_2 + \alpha\, y_2)$

$\alpha(x+y) = (\alpha\, x_1, \alpha\, x_2) + (\alpha\, y_1, \alpha\, y_2)$

$\alpha(x+y) = \alpha(x_1, x_2) + \alpha(\, y_1,\, y_2)$

$\alpha(x+y) = \alpha x + \alpha y$

A10) La unidad para el cuerpo de escalares es el neutro para el producto

La unidad en $\mathbb{R}$ es 1.

$1x = 1(x_1, x_2)$

$1x = (1\, x_1, 1\, x_2)$

$1x = (x_1, x_2)$

$1x = x$

Por lo tanto, $(S, +, \mathbb{R}, \cdot)$ es subespacio vectorial de $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$. Es decir que la recta de $\mathbb{R}^2, y = x$ es subespacio vectorial.

A partir de estos ejemplos podemos establecer lo siguiente:

i) una forma sencilla y rápida para descartar que un conjunto es un (sub)espacio vectorial es observar, si se da el caso, que el vector nulo no pertenece al conjunto. Obviamente hay conjuntos que contienen a este vector y que no son (sub)espacios vectoriales, y

ii) para demostrar que un subconjunto S es subespacio vectorial no es necesario comprobar de nuevo que cumple con todos los axiomas de espacio vectorial, en realidad es suficiente que S sea cerrado para la suma y para el producto por escalares. Debido a esto, demostraremos una condición necesaria y suficiente que nos permitirá decidir si un subconjunto no vacío de un espacio vectorial es subespacio. Esta condición es conocida como criterio de subespacio.

3.1. Criterio de subespacio

Sea el espacio vectorial $(V,+,K,\cdot)$ y S un subconjunto no vacío de V, $(S,+,K,\cdot)$ es subespacio vectorial de V si y solo si S es cerrado para la suma y producto por escalares.

Sea el espacio vectorial $(V,+,K,\cdot)\ y\ S \subset V, S \neq \emptyset$ se dice que S es subespacio vectorial del espacio vectorial $V \Leftrightarrow \begin{cases} \forall\, \overrightarrow{x}, \overrightarrow{y} \in S: (\overrightarrow{x} + \overrightarrow{y}) \in S \\ \forall\, \overrightarrow{x} \in S\ \forall \alpha \in K: \alpha\overrightarrow{x} \in S \end{cases}$

Observemos que la segunda condición implica que el vector nulo de V está también en S, porque si $\overrightarrow{x} \in S$ entonces $0\,\overrightarrow{x} = \overrightarrow{0}\ \in S$.

Demostración

$\Longrightarrow) S\ es\ subespacio\ vectorial\ de\ V \Longrightarrow \begin{cases} \forall\, \overrightarrow{x}, \overrightarrow{y} \in S: (\overrightarrow{x} + \overrightarrow{y}) \in S \\ \forall\, \overrightarrow{x} \in S\ \forall \alpha \in K: \alpha\overrightarrow{x} \in S \end{cases}$

Si S es subespacio vectorial de V cumple, de acuerdo a la definición, con todas las propiedades de un espacio vectorial. Por lo tanto, la suma y el producto por escalares son operaciones cerradas en S, o sea $\forall\, \overrightarrow{x}, \overrightarrow{y} \in S: (\overrightarrow{x} + \overrightarrow{y}) \in S\ \wedge\ \forall\, \overrightarrow{x} \in S\ \forall \alpha \in K: \alpha\overrightarrow{x} \in S$.

Con lo que queda la primera parte del criterio.

$\Longleftarrow) \left.\begin{matrix} \forall\, \overrightarrow{x}, \overrightarrow{y} \in S: (\overrightarrow{x} + \overrightarrow{y}) \in S \\ \forall\, \overrightarrow{x} \in S\ \forall \alpha \in K: \alpha\overrightarrow{x} \in S \end{matrix}\right\} \Longrightarrow S\ es\ subespacio\ vectorial\ de\ V$

Como S es subconjunto de V, sus elementos verifican los axiomas A_1, A_2, A_3, A_6, A_7, A_8, A_9 y A_{10}; solo restaría ver que se cumplan los axiomas A_4 que el elemento neutro para la suma también pertenezca a S y A_5 todo elemento de S tiene su opuesto en S.

A_4) Sea $\overrightarrow{x} \in S\ ya\ que\ S \neq \emptyset\ \wedge 0 \in K \Longrightarrow 0.\overrightarrow{x} = \overrightarrow{0}$

Como el producto de un escalar por un elemento de S da como resultado otro elemento de S, podemos asegurar que $\overrightarrow{0} \in S$

A5) Sea $\overrightarrow{x} \in S \; ya \; que \; S \neq \emptyset \; \wedge -1 \in K \Longrightarrow \; (-1).\overrightarrow{x} = -\overrightarrow{x}$

Como el producto de un escalar por un elemento de S da como resultado otro elemento de S, podemos asegurar que $-\overrightarrow{x} \in S$

Por lo tanto, S es subespacio vectorial de V, quedando demostrado así el criterio de subespacio.

Demostrar si los siguientes subconjuntos son o no subespacios vectoriales de los espacios vectoriales indicados.

Ejemplo 1: sea el conjunto $W = \{A \in \mathbb{R}^{2\times 2} / A = A^T\} \subset (\mathbb{R}^{n\times n}, +, \mathbb{R}, \cdot)$

i) $A \in W \; \Longrightarrow A = A^T$

$B \in W \; \Longrightarrow B = B^T$

$A + B = A^T + B^T = (A + B)^T$ por lo tanto, $(A + B) \; \in W$.

ii) $A \in W \; \Longrightarrow A = A^T$ y $\alpha \in \; \mathbb{R}$

$\alpha \; A = \alpha \; A^T = (\alpha \; A)^T$ por lo tanto $\alpha \; A \in W$

Por i) y ii) W es subespacio vectorial. Es decir, todas las matrices reales de orden 2 simétricas conforman un subespacio vectorial.

Ejemplo 2: sea el conjunto $W = \{(x, y) \in \mathbb{R}^2 / x^2 - y^2 = 0\} \subset (\mathbb{R}^2, +, \mathbb{R}, \cdot)$

i) $x = (x_1, y_1) \in W \Longrightarrow x_1^2 - y_1^2 = 0$

$y = (x_2, y_2) \in W \Longrightarrow x_2^2 - y_2^2 = 0$

$x + y = (x_1 + x_2; y_1 + y_2) \not\Rightarrow (x_1 - x_2)^2 + (y_1 - y_2)^2 = 0$. Esto se puede demostrar con un contraejemplo (ejemplo en contra).

$(1,1) \in W$ y $(1,-1) \in W$, pero claramente la suma de estos dos vectores $(2,0) \notin W$.

Por lo tanto, no es necesario comprobar la condición ii), ya podemos afirmar que W no es subespacio vectorial de $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$.

Observemos que en este ejemplo, el vector nulo, $(0,0) \in W$ y, sin embargo, W no es subespacio vectorial.

4. Combinación lineal de vectores

En esta sección formalizaremos el concepto de combinación lineal de vectores; para esto recordemos que en cualquier espacio vectorial existe la posibilidad de combinar sus elementos con multiplicación por escalares y sumándolos.

Sea el espacio vectorial V y $\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}$ tres elementos de él. Por ejemplo $3\overrightarrow{x}$ es una combinación lineal de $\overrightarrow{x}$, mientras que $2\overrightarrow{x} + 3\overrightarrow{y}$ es una combinación lineal de $\overrightarrow{x}$ e $\overrightarrow{y}$, mientras que si hacemos $\frac{1}{2}\overrightarrow{x} - 5\,\overrightarrow{y} + \frac{3}{4}\,\overrightarrow{z}$ estamos haciendo una combinación lineal de $\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}$. Esto significa que si tenemos un conjunto, A, de vectores de un espacio vectorial $(V, +, K, \cdot)$ entenderemos por combinación lineal de los vectores de A a toda suma de productos de escalares arbitrarios de K por los vectores de A. Con esto en mente podemos dar la siguiente:

Definición: dado un conjunto de n vectores $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}\}$ del espacio vectorial $(V, +, K, \cdot)$ combinación lineal de los vectores de A es todo vector del tipo:

$\sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} = \alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ \ldots + \alpha_n\overrightarrow{v_n} \ \ tq\ \alpha_i \in K \ \wedge \ \overrightarrow{v_i} \in A$

Ejemplo 1: la Figura 7.1 identifica combinaciones lineales seleccionadas de los vectores de $\mathbb{R}^2$ $\overrightarrow{v_1} = (-1, 1)$ y $\overrightarrow{v_2} = (2, 1)$. Observemos que las líneas paralelas de la rejilla están trazadas mediante múltiplos de $\overrightarrow{v_1}$ y $\overrightarrow{v_2}$. Estimar las combinaciones lineales de $\overrightarrow{v_1}$ y $\overrightarrow{v_2}$ que generan los vectores $\overrightarrow{u}$ y $\overrightarrow{w}$.

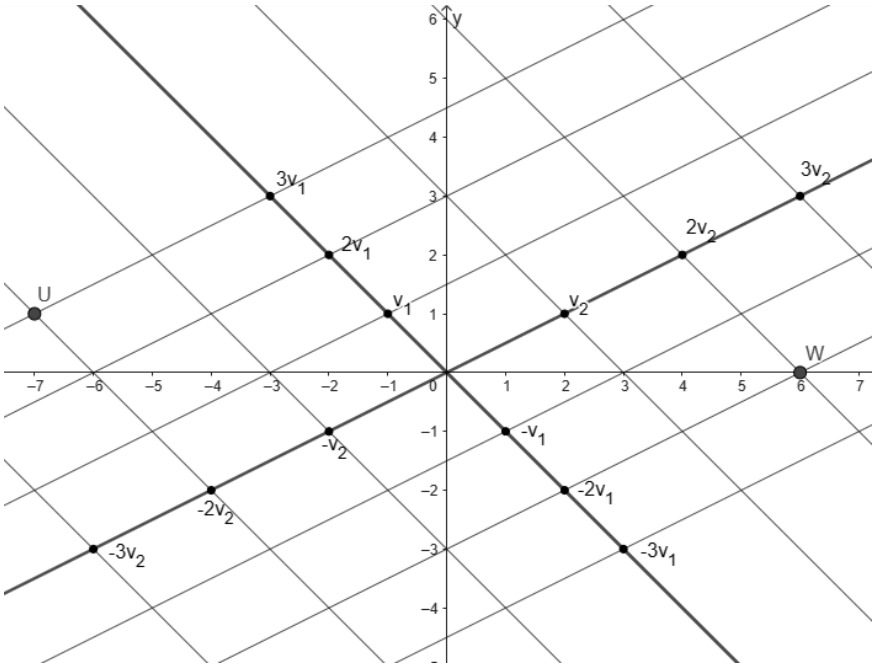


Figura 7.1. Combinación lineal de vectores.

Para encontrar la solución recurramos a la regla del paralelogramo para la suma de vectores. Esta regla nos indica que $\overrightarrow{u} = 3\overrightarrow{v_1} - 2\overrightarrow{v_2}$. Esta expresión para $\overrightarrow{u}$ puede interpretarse como dos instrucciones para desplazarse desde el origen hacia dicho vector.

i) desplácese 3 unidades en la dirección de $\overrightarrow{v_1}$, hasta 3 $\overrightarrow{v_1}$ y luego,

ii) avance –2 unidades en la dirección de $\overrightarrow{v_2}$.

Con la misma idea se establece que el vector $\vec{w}$ se puede obtener como combinación lineal de los vectores $\overrightarrow{v_1}\ y\ \overrightarrow{v_2}$ mediante la expresión $\vec{w} = -2\overrightarrow{v_1} + 2\overrightarrow{v_2}$.

Ejemplo 2: en el espacio vectorial $(\mathbb{R}^{2\times 2}, +, \mathbb{R}, \cdot)$ consideremos los siguientes vectores (matrices).

$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$

¿Se podrá expresar la matriz $D = \begin{pmatrix} 1 & 0 \\ 2 & 5 \end{pmatrix}$ como combinación lineal de $A, B\ y\ C$?

Esto significa que se debe poder expresar D como la suma de $A, B\ y\ C$ multiplicadas por escalares, es decir: $D = \alpha\ A + \beta B + \delta\ C$ con $\alpha, \beta\ y\ \delta \in \mathbb{R}$.

O sea

$$\begin{pmatrix} 1 & 0 \\ 2 & 5 \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} + \begin{pmatrix} \beta & 0 \\ \beta & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \delta & \delta \end{pmatrix}$$

$\begin{pmatrix} 1 & 0 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} \alpha + \beta & 0 \\ \beta + \delta & \alpha + \delta \end{pmatrix}$ Esto nos permite establecer, por igualdad de matrices, el siguiente sistema de ecuaciones lineales.

$$\begin{cases} \alpha + \beta = 1 \\ \beta + \delta = 2 \\ \alpha + \delta = 5 \end{cases}$$

Si bien este sistema se puede resolver con algunas manipulaciones algebraicas, vamos a hacerlo por el método de Gauss–Jordan estudiado en el capítulo 6. Sistema de Ecuaciones Lineales, pues esto nos ayudará

en un próximo análisis en esta misma sección y en el tema independencia lineal.

$$\left(\begin{array}{ccc|c}1 & 1 & 0 & 1\\ 0 & 1 & 1 & 2\\ 1 & 0 & 1 & 5\end{array}\right)\sim \begin{matrix}\\ \\ -F_1+F_3\end{matrix}\left(\begin{array}{ccc|c}1 & 1 & 0 & 1\\ 0 & 1 & 1 & 2\\ 0 & -1 & 1 & 4\end{array}\right)\sim \begin{matrix}F_2+F_1\\ \\ F_2+F_3\end{matrix}\left(\begin{array}{ccc|c}1 & 0 & -1 & -1\\ 0 & 1 & 1 & 2\\ 0 & 0 & 2 & 6\end{array}\right)\sim$$

$$\frac{1}{2}F_3\left(\begin{array}{ccc|c}1 & 0 & -1 & -1\\ 0 & 1 & 1 & 2\\ 0 & 0 & 1 & 3\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 0 & -1 & -1\\ 0 & 1 & 1 & 2\\ 0 & 0 & 1 & 3\end{array}\right)\sim \begin{matrix}F_3+F_1\\ -F_3+F_2\\ \\ \end{matrix}\left(\begin{array}{ccc|c}1 & 0 & 0 & 2\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & 3\end{array}\right)$$

Como el rango de la matriz del sistema es igual al rango de la matriz ampliada e igual al número de incógnitas, el sistema es compatible determinado y, por lo tanto, la solución del mismo es $S=\{\alpha=2,\beta=-1,\delta=3\}$. La respuesta a la pregunta que generó este desarrollo es afirmativa, es decir sí se puede expresar D como combinación lineal de $A, B\ y\ C$. Si hacemos en $\alpha\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}+\beta\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}+\delta\begin{pmatrix}0 & 0\\ 1 & 1\end{pmatrix}, \alpha=2,\beta=-1,\delta=3$ obtendremos $\begin{pmatrix}1 & 0\\ 2 & 5\end{pmatrix}$. Se sugiere que el estudiante realice esta verificación.

Ejemplo 3: determinar si el vector $\overrightarrow{v}=(-1,1,3)$ de $\mathbb{R}^3$ se puede generar (o escribir) como combinación lineal de $\overrightarrow{v_1}=(-1,0,2)\ y\ \overrightarrow{v_2}=(-1,2,4)$. Lo que nos solicitan aquí es que determinemos los coeficientes $\alpha\ y\ \beta$ de la ecuación vectorial

$\alpha\ \overrightarrow{v_1}+\beta\ \overrightarrow{v_2}=\overrightarrow{v}$ (1) o lo que es lo mismo

$\alpha\ (-1,0,2)+\beta\ (-1,2,4)=(-1,1,3)$

Aplicando las definiciones de producto por un escalar y suma de vectores podemos escribir directamente el sistema de ecuaciones lineales que deben satisfacer $\alpha \; y \; \beta$.

$\begin{cases} -\alpha - \beta = -1 \\ 2\beta = 1 \\ 2\alpha + 4\beta = 3 \end{cases}$ Resolvemos este sistema por algún método conocido, escribiendo su matriz ampliada

$\left(\begin{array}{cc|c} -1 & -1 & -1 \\ 0 & 2 & 1 \\ 2 & 4 & 3 \end{array}\right)$ (2) y a partir de allí analizamos el sistema.

$$\left(\begin{array}{cc|c} -1 & -1 & -1 \\ 0 & 2 & 1 \\ 2 & 4 & 3 \end{array}\right) \sim \overset{-F_1}{} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 2 & 4 & 3 \end{array}\right) \sim \overset{}{-2F_1 + F_3} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 2 & 1 \end{array}\right)$$

$$\sim \overset{}{-F_2 + F_3} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{array}\right) \sim \frac{1}{2}F_2 \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{array}\right) \sim \overset{-F_2 + F_1}{} \left(\begin{array}{cc|c} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{array}\right)$$

Como es un sistema compatible determinado, la solución es $\alpha = \frac{1}{2} \; y \; \beta = \frac{1}{2}$. Esto quiere decir que $\overrightarrow{v}$ es combinación lineal de $\overrightarrow{v_1} \; y \; \overrightarrow{v_2}$.

Verificación

$(-1, 1, 3) = \frac{1}{2}(-1, 0, 2) + \frac{1}{2}(-1, 2, 4)$

$(-1, 1, 3) = \left(-\frac{1}{2}, 0, 1\right) + (-\frac{1}{2}, 1, 2)$

$(-1, 1, 3) = (-1, 1, 3)$

Ejemplo 4: determinar si el vector $\overrightarrow{u} = (1, 2, 2)$ de $\mathbb{R}^3$ se puede generar (o escribir) como combinación lineal de $\overrightarrow{v_1} = (-1, 0, 2) \; y \; \overrightarrow{v_2} = (-1, 2, 4)$.

Debemos determinar los coeficientes $\alpha \; y \; \beta$ de la ecuación vectorial:

$\alpha\ \overrightarrow{v_1} + \beta\ \overrightarrow{v_2} = \overrightarrow{u}$ (1) o lo que es lo mismo,

$\alpha\ (-1, 0, 2) + \beta\ (-1, 2, 4) = (1, 2, 2)$ de donde escribimos directamente el sistema de ecuaciones lineales.

$\begin{cases} -\alpha - \beta = 1 \\ 2\beta = 2 \\ 2\alpha + 4\beta = 2 \end{cases}$ Resolvemos este sistema por algún método conocido, escribiendo su matriz ampliada

$\begin{pmatrix} -1 & -1 & 1 \\ 0 & 2 & 2 \\ 2 & 4 & 2 \end{pmatrix}$ (2) y a partir de allí analizamos el sistema.

$$\begin{pmatrix} -1 & -1 & 1 \\ 0 & 2 & 2 \\ 2 & 4 & 2 \end{pmatrix} \sim \begin{matrix} \\ \\ -F_1 \end{matrix} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 2 & 4 & 2 \end{pmatrix} \sim \begin{matrix} \\ \\ -2F_1 + F_3 \end{matrix} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 0 & 2 & 4 \end{pmatrix}$$

$$\sim \begin{matrix} \\ \frac{1}{2}F_2 \\ \frac{1}{2}F_3 \end{matrix} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \sim \begin{matrix} \\ -F_2 + F_1 \\ -F_2 + F_3 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

En este punto del análisis ya podemos afirmar que es un sistema incompatible, por lo tanto, no tiene solución. Esto significa que no se puede escribir el vector $\overrightarrow{u}$ como combinación lineal de los vectores $\overrightarrow{v_1}\ y\ \overrightarrow{v_2}$.

Podemos generalizar los resultados obtenidos en los ejemplos 2, 3 y 4 mediante la siguiente:

Definición: dado un conjunto de n vectores $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\}$ del espacio vectorial $(V, +, K, \cdot)$, el vector $\overrightarrow{v} \in V$ es combinación lineal de los vectores de A si y solo si existen escalares tales que: $\overrightarrow{v} = \sum_{i=1}^{n} \alpha_i\ \overrightarrow{v_i}\ con\ \alpha_i \in K\ \wedge\ \overrightarrow{v_i} \in A$

Otra cuestión que nos permite analizar más fácilmente los ejemplos 3 y 4 es la siguiente.

Observemos que en estos ejemplos los vectores $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v}$ y $\overrightarrow{u}$ son las columnas de las correspondientes matrices ampliadas.

Ejemplo 3: $\begin{pmatrix} -1 & -1 & -1 \\ 0 & 2 & 1 \\ 2 & 4 & 3 \\ \underbrace{\quad}_{\overrightarrow{v_1}} & \underbrace{\quad}_{\overrightarrow{v_2}} & \underbrace{\quad}_{\overrightarrow{v}} \end{pmatrix}$ y ejemplo 4: $\begin{pmatrix} -1 & -1 & 1 \\ 0 & 2 & 2 \\ 2 & 4 & 2 \\ \underbrace{\quad}_{\overrightarrow{v_1}} & \underbrace{\quad}_{\overrightarrow{v_2}} & \underbrace{\quad}_{\overrightarrow{u}} \end{pmatrix}$

Por brevedad de espacio se puede escribir estas matrices en una forma que identifique sus columnas, ($\overrightarrow{v_1}$ $\overrightarrow{v_2}$ $\overrightarrow{v}$).

En estos ejemplos de la ecuación vectorial (1), es claro cómo pasar a la matriz ampliada (2) sin realizar los pasos intermedios. Tomamos los vectores en el orden en que aparecen en (1) y los colocamos como columnas de una matriz como la (2).

Es decir, para el ejemplo 3 y 4 respectivamente.

De $\alpha\,\overrightarrow{v_1} + \beta\,\overrightarrow{v_2} = \overrightarrow{v}$ (1) podemos escribir directamente $\begin{pmatrix} -1 & -1 & -1 \\ 0 & 2 & 1 \\ 2 & 4 & 3 \end{pmatrix}$ (2).

De $\alpha\,\overrightarrow{v_1} + \beta\,\overrightarrow{v_2} = \overrightarrow{u}$ (1) podemos escribir directamente $\begin{pmatrix} -1 & -1 & 1 \\ 0 & 2 & 2 \\ 2 & 4 & 2 \end{pmatrix}$ (2)

Esta observación puede generalizarse estableciendo la siguiente:

Observación: una ecuación vectorial $\alpha_1\,\overrightarrow{v_1} + \alpha_2\,\overrightarrow{v_2} + \ldots + \alpha_n\,\overrightarrow{v_n} = \overrightarrow{v}$ tiene el mismo conjunto solución que el sistema lineal con matriz ampliada ($\overrightarrow{v_1}$ $\overrightarrow{v_2}$ … $\overrightarrow{v_n}$ $\overrightarrow{v}$) (3).

Particularmente, $\overrightarrow{v}$ se puede generar por una combinación lineal de $\overrightarrow{v_1}$, $\overrightarrow{v_2}$, … $\overrightarrow{v_n}$ si y solo si existe una solución al sistema lineal correspondiente a la matriz (3).

La última cuestión que podemos observar con los ejemplos 2, 3 y 4 es que al tener un conjunto de n vectores de un espacio vectorial V, entonces en general algunos vectores de V pueden ser combinaciones lineales de los n vectores dados –ejemplo 2 y 3– mientras que otros no –ejemplo 4–. Esta idea se profundizará en el siguiente apartado.

4.1. Subespacio generado

Una de las ideas muy importantes en álgebra lineal es el estudio del conjunto de todos los vectores que se pueden generar o escribir como combinación lineal de un conjunto fijo de vectores dados. Esto significa que es posible construir un conjunto que conste de todos los vectores que se pueden expresar como combinación lineal de n vectores dados. Esto nos lleva a la siguiente idea:

Sea A un conjunto no vacío de n vectores del espacio vectorial $(V, +, K, \cdot)$, se puede formar el subconjunto de V cuyos elementos sean todas las combinaciones lineales de los vectores de A. Este nuevo conjunto lo denotaremos con $\overline{A}$. Es decir:

Si $A = \{ \overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n} \}$ entonces $\overline{A} = \{ \sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} / \alpha_i \in K \ \wedge \ \overrightarrow{v_i} \in A \}$.

Ejemplo 1: sea $A = \{ (1,0,1), (0,1,1) \} \in \mathbb{R}^3$, calcular el conjunto de todas las combinaciones lineales que se pueden generar con A.

$\overline{A} = \{\alpha_1 (1,0,1) + \alpha_2 (0,1,1) / \alpha_1, \alpha_2 \in \mathbb{R}\}$

$\overline{A} = \{ (\alpha_1,\ \alpha_2,\ \alpha_1 + \alpha_2) / \alpha_1, \alpha_2 \in \mathbb{R}\}$

Esto significa que a $\overline{A}$ pertenecen todas las ternas de números reales cuya tercera componente es la suma de las dos primeras. Podemos escribir:

$\overline{A} = \{ (x_1, x_2, x_3) \in \mathbb{R}^3 / x_3 = x_1 + x_2 \}$

Este conjunto de todas las combinaciones posibles que se pueden construir a partir de un conjunto de n vectores tiene la propiedad de ser subespacio vectorial del espacio dado. Esto se demuestra en el siguiente:

Teorema: el conjunto de las combinaciones lineales de todo subconjunto no vacío de un espacio vectorial es un subespacio del mismo. Es decir:

Si $A = \{ \overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n} \} \subset (V, +, K, \cdot)$ entonces $(\overline{A}, +, K, \cdot)$ es subespacio de V.

D) Para demostrar este teorema utilizaremos el criterio de subespacio, o sea debemos demostrar que la suma y el producto por escalar son operaciones cerradas para $\overline{A}$.

Sean $\overrightarrow{v} = \sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} \in \overline{A}$ y $\vec{u} \sum_{i=1}^{n} \beta_i \overrightarrow{v_i} \in \overline{A}$ y $\lambda \in K$, entonces

$i) \overrightarrow{v} + \vec{u} = \sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} + \sum_{i=1}^{n} \beta_i \overrightarrow{v_i}$

$\overrightarrow{v} + \vec{u} = \sum_{i=1}^{n} (\alpha_i \overrightarrow{v_i} + \beta_i \overrightarrow{v_i})$

$\overrightarrow{v} + \vec{u} = \sum_{i=1}^{n} (\alpha_i + \beta_i) \overrightarrow{v_i}$

$\overrightarrow{v} + \vec{u} = \sum_{i=1}^{n} \delta_i \overrightarrow{v_i}$ con $\delta_i = \alpha_i + \beta_i$

$(\overrightarrow{v} + \vec{u}) \in \overline{A}$

$ii) \lambda \overrightarrow{v} = \lambda \sum_{i=1}^{n} \alpha_i \overrightarrow{v_i}$

$\lambda \overrightarrow{v} = \sum_{i=1}^{n} \lambda (\alpha_i \overrightarrow{v_i})$

$\lambda \overrightarrow{v} = \sum_{i=1}^{n} (\lambda \alpha_i) \overrightarrow{v_i}$

$\lambda \overrightarrow{v} = \sum_{i=1}^{n} \rho \overrightarrow{v_i}$ con $\rho = \lambda \alpha_i$

$\lambda \overrightarrow{v} \in \overline{A}$

Por i) y ii) $\overline{A}$ es subespacio vectorial de V, quedando demostrado el teorema.

La idea con la que se inició el apartado, sumada al teorema demostrado anteriormente, nos permite dar la siguiente:

Definición: el subespacio de las combinaciones lineales de un conjunto A no vacío de n vectores se denomina subespacio generado por A y se simboliza $\overline{A}$.

Ejemplo 2: determinar el subespacio de $\mathbb{R}^2$ generado por el vector $S = \{(1,2)\}$.

$\overline{S} = \{(x_1, x_2)/(x_1, x_2) = k(x_1, x_2)\}$. Por lo tanto

$(x_1, x_2) = k(1,2)$

$(x_1, x_2) = (k,\ 2k)$

$$\begin{cases} x_1 = k \\ x_2 = 2k \end{cases} \Longrightarrow \begin{cases} -2x_1 = -2k \\ x_2 = 2k \end{cases} \Longrightarrow -2x_1 + x_2 = 0$$

O lo que es lo mismo $2x_1 - x_2 = 0$

$\overline{S} = \{(x_1, x_2)/2x_1 - x_2 = 0\}$

Geométricamente $\overline{S}$ es la recta que pasa por el origen y el punto $(1,2)$, esto se muestra en la Figura 7.2.

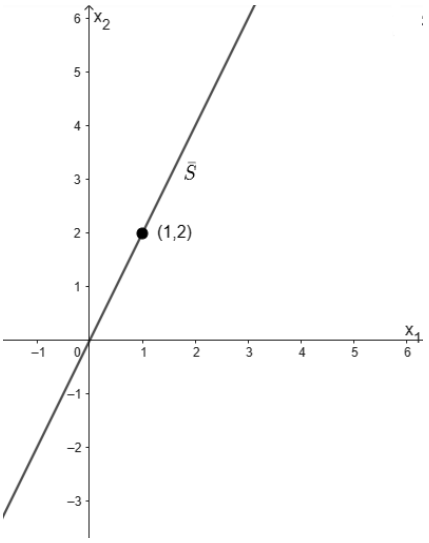


Figura 7.2. Subespacio generado.

Observación: en general, el subespacio de $(V, +, K, \cdot)$ generado por un vector $\overrightarrow{v}$ es, en particular, el conjunto de todos los múltiplos escalares de ese vector, o sea $\{k\ \overrightarrow{v} / k \in K\}$

Ejemplo 3: determinar el subespacio generado por los vectores del siguiente conjunto $P = \{2x^2 - 1, -x^2 + 2x + 1\} \subset (\mathbb{P}[x], +, \mathbb{R}\cdot)$ en donde $\mathbb{P}[x]$ son los polinomios de una indeterminada con coeficientes reales y de grado menor igual a 2.

$\overline{P} = \{\alpha_1\ (2x^2 - 1) + \alpha_2(-x^2 + 2x + 1) / \alpha_1\ y\ \alpha_2 \in \mathbb{R}\}$

$\overline{P} = \{(2\alpha_1 x^2 - \alpha_1) + (-\alpha_2 x^2 + 2\alpha_2 x + \alpha_2)\}$

$\overline{P} = \{(2\alpha_1 - \alpha_2)x^2 + 2\alpha_2 x + (-\alpha_1 + \alpha_2)\}$

$\overline{P} = \{a_2 x^2 + a_1 x + a_0\ \ tq\ \ a_2 = 2\alpha_1 - \alpha_2,\ a_1 = 2\alpha_2\ \ y\ \ a_0 = -\alpha_1 + \alpha_2\}$,

$\alpha_1, \alpha_2\ \in \mathbb{R}$

¿Se puede escribir $Q(x) = 12x + 3$ como combinación lineal de los vectores de P?

Como tenemos todas las combinaciones posibles de los vectores de P, podemos hacer

$(2\alpha_1 - \alpha_2)x^2 + 2\alpha_2 x + (-\alpha_1 + \alpha_2) = 0x^2 + 12x + 3$. Escribimos el sistema de ecuaciones lineales.

$$\begin{cases} 2\alpha_1 - \alpha_2 = 0 \\ 2\alpha_2 = 12 \\ -\alpha_1 + \alpha_2 = 3 \end{cases}$$

Resolvemos este sistema por manipulación algebraica debido a sus características.

$\alpha_2 = 6$ por la segunda ecuación, por lo tanto, por la tercera ecuación $\alpha_1 = 3$. Estos valores de α_1 y α_2 satisfacen la primera ecuación.

Esto significa que sí se puede escribir $Q(x) = 12x + 3$ como combinación lineal de los vectores de P.

$12x + 3 = 3(2x^2 - 1) + 6(-x^2 + 2x + 1)$

$12x + 3 = (6x^2 - 3) + (-6x^2 + 12x + 6)$

$12x + 3 = 12x + 3$

5. Dependencia lineal

En el apartado 4 se estableció que la ecuación $\alpha_1 \overrightarrow{v_1} + \alpha_2 \overrightarrow{v_2} + \ldots + \alpha_n \overrightarrow{v_n} = \overrightarrow{v}$ tiene el mismo conjunto solución que el sistema lineal con matriz ampliada ($\overrightarrow{v_1}$ $\overrightarrow{v_2}$ … $\overrightarrow{v_n}$ $\overrightarrow{v}$). Esta observación se torna más importante en el caso de que el vector $\overrightarrow{v}$ sea el vector nulo y por consiguiente el sistema de ecuaciones lineales se vuelve homogéneo. Esto

debido a que la atención se transfiere de las soluciones desconocidas del sistema homogéneo a los vectores que aparecen en la ecuación vectorial.

Por ejemplo, sea expresar $(0,0) \in \mathbb{R}^2$como combinación lineal de $(-1,1)\ y\ (1,0)$.

La ecuación vectorial es $\alpha_1(-1,1) + \alpha_2(1,0) = (0,0)$, desde la cual podemos escribir directamente la matriz ampliada del sistema de ecuaciones: $A = \left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right)$.

Como es un sistema homogéneo cuadrado, calculamos directamente el determinante del sistema $\det(A) = -1$, como $\det(A) \neq 0$ el sistema es compatible determinado, es decir la única solución es $\alpha_1 = 0\ y\ \alpha_2 = 0$.

Esto nos permite dar la siguiente:

Definición: dado un conjunto de n vectores, $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\}$, del espacio vectorial $(V, +, K, \cdot)$ se dice que A es linealmente independiente si y solo si la única combinación lineal, cuyo resultado es el vector nulo, es la trivial. Es decir, sean

$(V, +, K, \cdot)\ y\ A \subset V$, A es Linealmente Independiente $\Leftrightarrow \forall i$: $\sum_{i=1}^{n} \alpha_i\ \overrightarrow{v_i} = \overrightarrow{0} \Longrightarrow \alpha_i = 0$

Se dice que el conjunto $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\}$ es linealmente dependiente si existen los escalares $\alpha_1, \alpha_2, \ \dots, \alpha_n$ no todos nulos tales que $\sum_{i=1}^{n} \alpha_i\ \overrightarrow{v_i} = \overrightarrow{0}$ (1).

La ecuación (1) es la relación de dependencia lineal entre $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}$ cuando no todos los coeficientes son ceros. Un conjunto A de vectores es linealmente dependiente si y solo si no es linealmente independiente.

La independencia lineal de un conjunto finito y no vacío de vectores significa que no puede darse una combinación lineal de dicho conjunto que dé el vector nulo, con algún escalar distinto de cero.

Convenimos en que el conjunto vacío es linealmente independiente.

Ejemplo 1: determinar si el conjunto $A = \{(1,2,3);(4,5,6);(2,1,0)\} \subset \mathbb{R}^3$ es linealmente independiente. De no ser así, encuentre la relación de dependencia entre los vectores de A.

La ecuación vectorial es $\alpha_1(1,2,3) + \alpha_2(4,5,6) + \alpha_3(2,1,0) = (0,0,0)$, de donde podemos escribir directamente la matriz ampliada del sistema de ecuaciones lineales.

$$\left(\begin{array}{ccc|c} \textcircled{1} & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{array}\right) \sim \begin{matrix} \\ -2F_1 + F_2 \\ -3F_1 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -6 & -6 & 0 \end{array}\right) \sim \begin{matrix} \\ -\frac{1}{3}F_2 \\ -\frac{1}{6}F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 4 & 2 & 0 \\ 0 & \textcircled{1} & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right)$$

$$\sim \begin{matrix} -4F_2 + F_1 \\ \\ -F_2 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Se trata de un sistema compatible indeterminado, por lo tanto A es un conjunto linealmente dependiente (consecuentemente, no es linealmente independiente).

Para evidenciar la relación de dependencia, reconstruimos el sistema de ecuaciones.

$\begin{cases} \alpha_1 - 2\alpha_3 = 0 \\ \alpha_2 + \alpha_3 = 0 \end{cases}$ entonces si tomamos α_3 como variable libre, tenemos

$\begin{cases} \alpha_1 = 2\alpha_3 \\ \alpha_2 = -\alpha_3 \end{cases}$

Seleccionamos valores no nulos para la variable libre, por ejemplo $\alpha_3 = 1$ por lo tanto $\alpha_1 = 2 \ y \ \alpha_2 = -1$. Sustituimos en (1) y obtenemos $2\overrightarrow{v_1} - \overrightarrow{v_2} + \overrightarrow{v_3} = \vec{0}$ (i). Podríamos elegir $\alpha_3 = 5$ entonces $\alpha_1 = 10 \ y \ \alpha_2 = -5$ y obtener $10\overrightarrow{v_1} - 5\overrightarrow{v_2} + 5\overrightarrow{v_3} = \vec{0}$ (ii) y así sucesivamente. Las ecuaciones (i) y (ii) son dos –entre una infinidad– de las posibles relaciones de dependencia lineal entre los vectores de A. Se sugiere al estudiante comprobar (i) y (ii) y proponer otros valores para α_3 y verificar nuevamente.

Ejemplo 2: sea el espacio vectorial $(\mathbb{R}^{2\times 2}, +, \mathbb{R}, \cdot)$ determinar si los siguientes vectores $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ son linealmente independiente.

$$\alpha_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} \alpha_1 + \alpha_2 & 0 \\ \alpha_2 + \alpha_3 & \alpha_1 + \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{cases} \alpha_1 + \alpha_2 = 0 \implies \alpha_1 = -\alpha_2 \\ \alpha_2 + \alpha_3 = 0 \implies \alpha_3 = -\alpha_2 \\ \alpha_1 + \alpha_3 = 0 \ \therefore -2\alpha_2 = 0 \implies \alpha_2 = 0 \end{cases}$$

Como $\alpha_2 = 0 \implies \alpha_1 = 0 \ y \ \alpha_3 = 0$

Esto significa que la única combinación lineal que satisface la relación propuesta es la trivial, por lo tanto, el conjunto de vectores es linealmente independiente (y, en consecuencia, no es linealmente dependiente).

Ejemplo 3: sea $(\mathbb{P}(x), +, \mathbb{R}, \cdot)$ con $\mathbb{P}(x)$, polinomio de variable x, coeficientes reales y grado menor o igual a 2. Determinar si el conjunto $A = \{(1-x)^2, (1-x), 1\}$ es linealmente independiente.

$$\alpha(1-x)^2 + \beta(1-x) + \gamma = 0x^2 + 0x + 0$$

$\alpha - 2\alpha x + \alpha x^2 + \beta - \beta x + \gamma = 0x^2 + 0x + 0$

$\alpha x^2 + (-2\alpha - \beta)x + (\alpha + \beta + \gamma) = 0x^2 + 0x + 0$

$$\begin{cases} \alpha = 0 \\ -2\alpha - \beta = 0 \implies \beta = 0 \\ \alpha + \beta + \gamma = 0 \implies \gamma = 0 \end{cases}$$

Por lo tanto, el conjunto A es linealmente independiente.

Obsérvense dos cuestiones.

i) En muchos casos es más fácil resolver el sistema de ecuaciones por manipulación algebraica, mientras que en otros conviene resolverlo por algún método estudiado en el capítulo 6. Sistema de Ecuaciones Lineales.

ii) El uso de los nombres para el cálculo de los coeficientes es indistinto.

Ejemplo 4: sea $(\mathbb{R}^{2\times 2}, +, \mathbb{R}, \cdot)$ determinar si los vectores $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix}$ son linealmente independiente.

$$\alpha \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \beta \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} \alpha - \beta & 0 \\ \alpha - \beta & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$\alpha - \beta = 0 \implies \alpha = \beta = k \quad \forall\, k \in \mathbb{R}$

Esto significa que existen escalares no nulos que satisfacen la relación propuesta, por lo tanto, el conjunto de vectores no es linealmente independiente. Podríamos elegir un valor no nulo de k, por ejemplo $k = 3$ y la relación de dependencia será $3A + 3B = N$ (recordar que en este caso N es la matriz nula de orden 2) o bien $k = -5$ y la relación de dependencia será $-5A - 5B = N$.

Verificación:

$3A + 3B = N$

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

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

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

Similar situación ocurre para $-5A - 5B = N$.

Se sugiere proponer otros valores para k y verificar la relación de dependencia.

El ejemplo 1 y el ejemplo 4 nos ayudan para poder dar otra definición de vectores linealmente dependientes.

Definición: dado un conjunto de n vectores, $A = \{ \overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n} \}$ del espacio vectorial $(V, +, K, \cdot)$ se dice que A es linealmente dependiente si y solo si existe una combinación lineal no trivial de A cuyo resultado es el vector nulo. Simbólicamente, sean

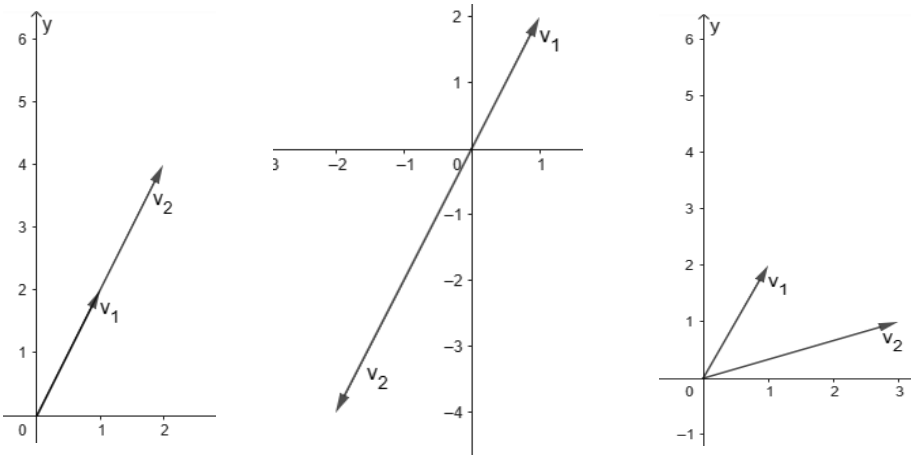
$(V, +, K, \cdot)$ y $A \subset V$, A es Linealmente Dependiente $\Leftrightarrow \exists i / \sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} = \overrightarrow{0} \wedge \alpha_i \neq 0$

La dependencia lineal de un conjunto finito y no vacío de vectores significa que tiene que existir, al menos, una combinación lineal de dicho conjunto que dé el vector nulo y que no sea la trivial.

Interpretación geométrica de la dependencia lineal

La dependencia lineal posee algunas interpretaciones geométricas útiles en $\mathbb{R}^2$ y $\mathbb{R}^3$.

En $\mathbb{R}^2$ un conjunto de dos vectores es linealmente independiente si y solo si los vectores no están en la misma recta cuando se colocan con sus orígenes en el origen del sistema.



a) b) c)

Figura 7.3. Interpretación geométrica, $\mathbb{R}^2$, de la dependencia lineal: a) y b) Linealmente dependientes; c) Linealmente independientes.

Las dos primeras figuras, a) y b), son una consecuencia del hecho de que dos vectores son linealmente independientes si y solo si ninguno de ellos es múltiplo escalar del otro; geométricamente, esto equivale a afirmar que los vectores no están en la misma recta cuando se hace coincidir sus orígenes con el origen de coordenadas.

En $\mathbb{R}^3$ un conjunto de tres vectores es linealmente independiente si y solo si los vectores no están en el mismo plano cuando se colocan con sus

orígenes en el origen del sistema. Una consecuencia de esto es el hecho de que tres vectores son linealmente independientes si ninguno de ellos es combinación lineal de los otros dos, de otra forma, que los tres vectores no sean coplanares –que no están en un mismo plano– cuando se hace coincidir sus orígenes con el origen de coordenadas.

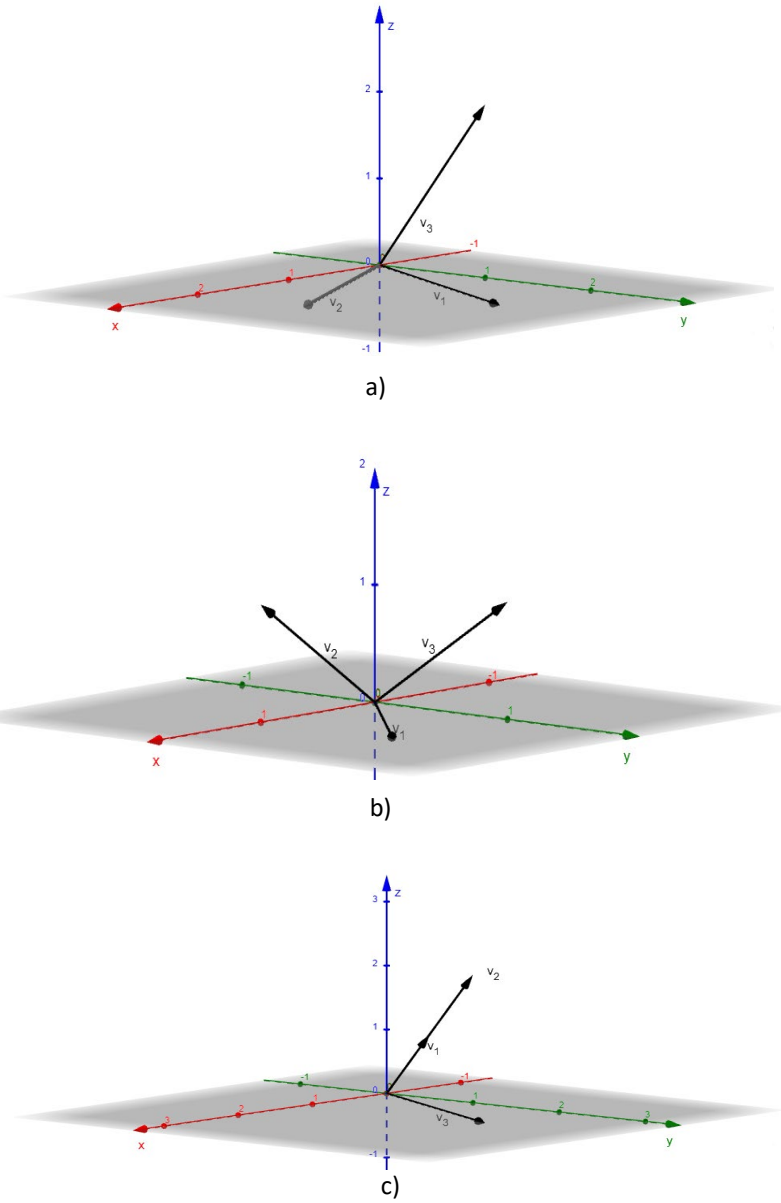


Figura 7.4. Interpretación geométrica, $\mathbb{R}^3$, de la dependencia lineal: a) y b) Linealmente independientes; c) Linealmente dependientes.

5.1. Propiedades de la dependencia lineal

Para finalizar esta sección, se demostrarán algunas propiedades respecto a la dependencia lineal que en muchos casos ayudan a determinar rápidamente si un conjunto de vectores es linealmente independiente o dependiente.

P_1) Si un vector es combinación lineal de un conjunto de vectores linealmente independiente, entonces dicha combinación lineal es única.

Sean el espacio vectorial $(V, +, K, \cdot)$, $\overrightarrow{x} \in V$ combinación lineal de $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}\}$ y éste linealmente independiente.

D) Como $\overrightarrow{x}$ es combinación lineal de los vectores de A, entonces existen los escalares α_i tales que $\overrightarrow{x} = \sum_{i=1}^{n} \alpha_i \overrightarrow{v_i}$.

Suponemos que existen otros escalares β_i tales que $\overrightarrow{x} = \sum_{i=1}^{n} \beta_i \overrightarrow{v_i}$. Por lo tanto

$\sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} = \sum_{i=1}^{n} \beta_i \overrightarrow{v_i}$

$\sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} - \sum_{i=1}^{n} \beta_i \overrightarrow{v_i} = \overrightarrow{0}$

$\sum_{i=1}^{n} (\alpha_i \overrightarrow{v_i} - \beta_i \overrightarrow{v_i}) = \overrightarrow{0}$

$\sum_{i=1}^{n} (\alpha_i - \beta_i) \overrightarrow{v_i} = \overrightarrow{0}$

Como los vectores de A son linealmente independiente, se deduce que

$\alpha_i - \beta_i = 0 \Longrightarrow \alpha_i = \beta_i \; \forall i = 1, 2 \ldots, n.$

Por lo tanto, la combinación lineal es única.

Ejemplo: en el espacio vectorial $(\mathbb{R}^{2x2}, +, \mathbb{R}, \cdot)$ se considera el conjunto M linealmente independiente (se sugiere al estudiante comprobar esto).

$M = \left\{A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}\right\}$, vamos a escribir $D = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}$ como combinación lineal de los vectores de M.

La ecuación vectorial es: $\alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}$, mientras que el sistema de ecuaciones es

$$\begin{cases} \alpha + \beta = -1 \\ \beta + \delta = 0 \\ \alpha + \delta = 5 \end{cases}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 5 \end{array}\right) \sim \overset{-F_1 + F_3}{} \left(\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 6 \end{array}\right) \sim \overset{\begin{matrix} -F_2 + F_1 \\ F_2 + F_3 \end{matrix}}{} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 6 \end{array}\right) \sim$$

$$\sim \overset{\frac{1}{2}F_3}{} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right) \sim \overset{\begin{matrix} F_3 + F_1 \\ -F_3 + F_2 \end{matrix}}{} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Es un sistema compatible determinado, única solución $\alpha = 2, \beta = -3$ y $\delta = 3$. De allí que los únicos coeficientes que permiten escribir D como combinación lineal de los vectores de M, sean $\alpha = 2, \beta = -3$ y $\delta = 3$, Se sugiere al estudiante hacer la verificación de la ecuación vectorial propuesta al iniciar el ejercicio.

P2) Todo vector no nulo de un espacio vectorial constituye un conjunto linealmente independiente.

Sea $\vec{x} \neq \vec{0}$ en $(V, +, K, \cdot)$. Por propiedades de los espacios vectoriales se sabe que

$\alpha\vec{x} = \vec{0} \Longrightarrow \alpha = 0 \vee \vec{x} = \vec{0}$. Como $\vec{x} \neq \vec{0}$ se deduce que $\alpha = 0$ y en consecuencia $\{\vec{x}\}$ es linealmente independiente.

Ejemplo: sea el espacio vectorial $(\mathbb{P}[x], +, \mathbb{R}\cdot)$ en donde $\mathbb{P}[x]$ son los polinomios de una indeterminada con coeficientes reales y de grado menor a n.

Tomemos de este espacio vectorial $P(x) = 5x^2 - 3x + 4$ y propongamos $\alpha\,(5x^2 - 3x + 4) = 0x^2 + 0x + 0$ la única forma de que esta igualdad se cumpla, es que $\alpha = 0$ lo que significa que $\{P(x)\}$ es linealmente independiente. Esto es válido para cualquier $P(x) = ax^2 + bx + c$ del espacio vectorial.

P_3) El vector nulo de cualquier espacio vectorial constituye un conjunto linealmente dependiente.

Por propiedades de los espacios vectoriales se sabe que $\alpha.\overrightarrow{0} = \overrightarrow{0}\ \forall\alpha \in K$, entonces cualquier escalar nulo, o no, satisface esta relación y, por lo tanto, $\{\overrightarrow{0}\}$ es linealmente dependiente.

Ejemplo: sea el espacio vectorial $(\mathbb{R}^4, +, \mathbb{R}\cdot)$, consideremos de él, el vector nulo y propongamos:

$\alpha(0,0,0,0) = (0,0,0,0)$ para que se cumpla esta igualdad, el escalar α no necesariamente debe ser nulo, puede tomar cualquier valor y la igualdad se verificará. Lo que indica que el vector nulo es un conjunto linealmente dependiente.

P_4) Todo conjunto al que pertenezca el vector nulo es linealmente dependiente.

Sea $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\} \subset (V, +, K, \cdot)$ y $con\ \overrightarrow{v_j} = \overrightarrow{0}$ se verifica que

$0\,\overrightarrow{v_1} + 0\,\overrightarrow{v_2} + \ \ldots + \alpha_j \overrightarrow{v_j} + \ \ldots + 0\,\overrightarrow{v_n} = \overrightarrow{0}$ o sea que:

$\alpha_j \overrightarrow{v_j} = \alpha_j \overrightarrow{0} = \overrightarrow{0}$ con α_j no necesariamente nulo. Por lo tanto, A es linealmente dependiente.

Ejemplo: en el espacio vectorial $(\mathbb{R}^{2x2}, +, \mathbb{R}, \cdot)$ se considera el conjunto

$$M = \left\{A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right\}$$

La ecuación vectorial es: $\alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, mientras que el sistema de ecuaciones es

$$\begin{cases} \alpha + \beta = 0 \\ \beta + \delta = 0 \\ \alpha + \delta = 0 \end{cases}$$

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{array}\right) \sim \begin{matrix} \\ \\ -F_1 + F_3 \end{matrix} \left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \end{array}\right) \sim$$

$$\sim \begin{matrix} -F_2 + F_1 \\ \\ F_2 + F_3 \end{matrix} \left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{array}\right) \sim \begin{matrix} \\ \\ \frac{1}{2}F_3 \end{matrix} \left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right) \sim$$

$$\sim \begin{matrix} F_3 + F_1 \\ -F_3 + F_2 \\ \\ \end{matrix} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right)$$

Es un sistema compatible indeterminado porque el rango de la matriz del sistema es 3 al igual que la ampliada, pero el número de incógnitas es cuatro.

$\alpha = 0,\ \beta = 0,\ \delta = 0\ y\ \gamma \in \mathbb{R}$. De esta forma M es linealmente dependiente.

P5) Un conjunto finito y no vacío de vectores es linealmente dependiente si y solo si algún vector es combinación lineal de los demás.

Sea $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\} \subset (V, +, K, \cdot)$

D)

$\Longrightarrow$) A es linealmente dependiente, entonces $\exists j / v_j = \sum_{i\neq j}^{r} \beta_i \overrightarrow{v_i}$

Como A es linealmente dependiente, por definición de dependencia lineal

$\sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} = \overrightarrow{0} \ \wedge \alpha_j \neq 0$

$\alpha_j \overrightarrow{v_j} + \sum_{i\neq j}^{r} \alpha_i \overrightarrow{v_i} = \overrightarrow{0} \ \wedge \alpha_j \neq 0$

$\alpha_j \overrightarrow{v_j} = -\sum_{i\neq j}^{r} \alpha_i \overrightarrow{v_i} \wedge \alpha_j \neq 0$

Premultiplicando por α_j^{-1} que existe pues $\alpha_j \neq 0$. Se tiene

$(\alpha_j^{-1})\alpha_j \overrightarrow{v_j} = (-\alpha_j^{-1}) \sum_{i\neq j}^{r} \alpha_i \overrightarrow{v_i}$

$\left(\alpha_j^{-1} \alpha_j\right)\overrightarrow{v_j} = \sum_{i\neq j}^{r}\left(-\alpha_j^{-1}\right) (\alpha_i \overrightarrow{v_i})$

$\left(\alpha_j^{-1} \alpha_j\right)\overrightarrow{v_j} = \sum_{i\neq j}^{r}\left(-\alpha_j^{-1}\right) (\alpha_i \overrightarrow{v_i})$

$\overrightarrow{v_j} = \sum_{i\neq j}^{r}\left(-\alpha_j^{-1} \alpha_i\right) \overrightarrow{v_i}$

$\overrightarrow{v_j} = \sum_{i\neq j}^{r} \beta_i \overrightarrow{v_i}$ con $\beta_i = -\alpha_j^{-1} \alpha_i$

Por lo tanto, v_j existe y es combinación lineal de los demás vectores de A.

$\Longleftarrow$) $v_j = \sum_{i\neq j}^{r} \alpha_i \overrightarrow{v_i} \Longrightarrow$ A es linealmente dependiente

$v_j = \sum_{i\neq j}^{r} \alpha_i \overrightarrow{v_i}$

$\sum_{i\neq j}^{r} \alpha_i \overrightarrow{v_i} - v_j = \overrightarrow{0}$ con $\alpha_j = -1$.

Como $\alpha_j \neq 0$, el conjunto A es linealmente dependiente.

Ejemplo: sea el conjunto $A = \{(-1, 1), (1, 2), (-3, 0)\} \subset \mathbb{R}^2$. Para ver la dependencia lineal, hacemos

$$\left(\begin{array}{ccc|c} -1 & 1 & -3 & 0 \\ 1 & 2 & 0 & 0 \end{array}\right) \sim^{-F_1} \left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 1 & 2 & 0 & 0 \end{array}\right) \sim_{-F_1+F_2} \left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 3 & -3 & 0 \end{array}\right) \sim$$

$$\sim_{\frac{1}{3}F_2} \left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right) \sim^{F_2+F_1} \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right)$$

Es un sistema compatible indeterminado y, por lo tanto, A es linealmente dependiente.

Reconstruimos el sistema y evidenciamos la relación de dependencia.

$\begin{cases} \alpha + 2\delta = 0 \\ \beta - \delta = 0 \end{cases}$ de donde $\begin{cases} \alpha = -2\delta \\ \beta = \delta \end{cases}$

Damos un valor no nulo a δ, por ejemplo $\delta = -1$con lo que $\alpha = 2\ y\ \beta = -1$ que son justamente los coeficientes de $(-3, 0)$ como combinación lineal de $(-1, 1), (1, 2)$.

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

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

$(-3, 0) = (-3, 0).$

Muchas veces la relación entre $(-3, 0)$ como combinación lineal de $(-1, 1), (1, 2)$ no es tan evidente.

P6) Un conjunto finito y no vacío de vectores es linealmente independiente si y solo si ningún vector es combinación lineal de los demás.

Sea $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\} \subset (V, +, K, \cdot)$

A es linealmente independiente si y solo si $\forall j$: $v_j \neq \sum_{i \neq j}^{r} \beta_i \overrightarrow{v_i}$

D) Lo mismo que en la propiedad P5) hay que demostrar dos implicaciones

i) A es linealmente independiente, entonces $\forall j$: $v_j \neq \sum_{i \neq j}^{r} \beta_i \overrightarrow{v_i}$

ii) $\forall j$: $v_j \neq \sum_{i \neq j}^{r} \beta_i \overrightarrow{v_i}$ entonces A es linealmente independiente.

En la P_5) se demostró que,

A es linealmente dependiente, entonces $\exists j / v_j = \sum_{i \neq j}^{r} \beta_i \, \overrightarrow{v_i}$ la contrarrecíproca[43] de esta implicación tiene el mismo valor de verdad, por lo tanto, también es verdadera, o sea

$\forall j: v_j \neq \sum_{i \neq j}^{r} \beta_i \, \overrightarrow{v_i}$ entonces A es linealmente independiente. Quedando demostrada ii).

De manera similar en la P_5) se demostró que,

$v_j = \sum_{i \neq j}^{r} \beta_i \, \overrightarrow{v_i} \implies A$ es linealmente dependiente, cuya contrarrecíproca se enuncia de la siguiente manera:

A es linealmente independiente, entonces $\forall j: v_j \neq \sum_{i \neq j}^{r} \beta_i \, \overrightarrow{v_i}$. Quedando demostrado i) y por consiguiente también la propiedad.

Ejemplo: el conjunto de vectores $\{ u_1 = (-1, 1) \; y \; u_2 = (1, 0) \}$ es linealmente independiente ya que ninguno de los dos puede expresarse como combinación lineal del otro. Si esto no fuese así, existiría $\alpha \in \mathbb{R}$ tal que

$u_1 = \alpha \, u_2$

$(-1, 1) = \alpha(1, 0)$ lo que implica

$\alpha = -1 \; y \; 1 = 0$ que es imposible. Luego, el conjunto es linealmente independiente.

[43] Sea la implicación $p \implies q$ su contrarrecíproca (negar e intercambiar antecedente por consecuente y viceversa) $\sim q \implies \sim p$ tiene el mismo valor de verdad.

Las dos siguientes propiedades también están relacionadas, por lo tanto en primer lugar establecemos: sea $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\} \subset (V, +, K, \cdot)$ y $A' \subset A$.

P7) Un conjunto de vectores que contenga un subconjunto linealmente dependiente es linealmente dependiente. Es decir, si $A' \subset A$ es linealmente dependiente, $\Longrightarrow A$ es linealmente dependiente.

D1) Como la propiedad de ser linealmente dependiente no depende del orden de los elementos del conjunto, suponemos que los p primeros vectores son linealmente dependientes. Sean $\alpha_1, \dots, \alpha_p$ los coeficientes no todos nulos tales que $\sum_{i=1}^{p} \alpha_i \overrightarrow{v_i} = \overrightarrow{0}$.

Pongamos que $\alpha_{p+1} = \alpha_{p+2} = \dots = \alpha_n = 0$. Por lo tanto

$$\sum_{i=1}^{n} \alpha_i \overrightarrow{v_i} = \underbrace{\sum_{i=1}^{p} \alpha_i \overrightarrow{v_i}}_{\overrightarrow{0}} + \underbrace{\sum_{p+1}^{n} \underbrace{\alpha_i}_{0} \overrightarrow{v_i}}_{\overrightarrow{0}} = \overrightarrow{0}$$

Esta combinación lineal no es la trivial, pues hay coeficientes no nulos entre los primeros coeficientes $\alpha_1, \dots, \alpha_p$. Por lo tanto, A es linealmente dependiente.

D2) Como A' es linealmente dependiente, existe una relación de dependencia no trivial en A', esa misma relación de dependencia nos sirve para probar que A es linealmente dependiente ya que todo elemento de A' pertenece a A, pues $A' \subset A$.

Ejemplo: sea el conjunto $M = \left\{\mathrm{A} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix}, C = \begin{pmatrix} 3 & 0 \\ 1 & 0 \end{pmatrix}\right\}$ del espacio vectorial $(\mathbb{R}^{2x2}, +, \mathbb{R}, \cdot)$. Por el ejemplo 4 del apartado 5,

Dependencia lineal, sabemos que el subconjunto de M, $N = \left\{A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix}\right\}$ es linealmente dependiente; comprobemos que M es también linealmente dependiente.

$$\alpha \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \beta \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 3 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} \alpha - \beta + 3\delta & 0 \\ \alpha - \beta + \delta & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ de donde}$$

$$\begin{cases} \alpha - \beta + 3\delta = 0 \\ \alpha - \beta + \delta = 0 \end{cases} \text{ de donde } \delta = 0.$$

Si $\delta = 0$ entonces $\alpha - \beta = 0 \implies \alpha = \beta = k \quad \forall\, k \in \mathbb{R}$

Lo que indica que el conjunto M es linealmente dependiente, teniendo la misma relación de dependencia del subconjunto N.

P_8) Si un conjunto de vectores es linealmente independiente, todo subconjunto de éste es también linealmente independiente. Si A es linealmente independiente $\implies A' \subset A$ es linealmente independiente.

D_1) Si con n elementos (los vectores de A) no se puede formar una combinación lineal nula con coeficientes no todos nulos, con un elemento menos (subconjunto) tampoco se podrá obtenerla. Por lo tanto, todo subconjunto de A es linealmente independiente.

D_2) La verdad de esta proposición se puede asegurar por ser la contrarrecíproca de la P_7) que ya se demostró que es verdadera.

La P_7) indica, si $A' \subset A$ es linealmente dependiente, $\implies A$ es linealmente dependiente. Negando e intercambiando antecedente y consecuente, tenemos:

A es linealmente independiente $\Longrightarrow$ todo $A' \subset A$ es linealmente independiente.

Ejemplo: sea el conjunto $A = \{(1-x)^2, (1-x), 1\}$ del $(\mathbb{P}(x), +, \mathbb{R}, \cdot)$ con $\mathbb{P}(x)$, polinomio de variable x, coeficientes reales y grado menor o igual a 2. En el ejemplo 3 del apartado 5, Dependencia lineal, se demostró que A es linealmente independiente.

Tomemos el subconjunto $A' = \{(1-x)^2, (1-x)\}$ y verifiquemos que es también linealmente independiente.

$\alpha(1-x)^2 + \beta(1-x) = 0x^2 + 0x + 0$

$\alpha - 2\alpha x + \alpha x^2 + \beta - \beta x = 0x^2 + 0x + 0$

$\alpha x^2 + (-2\alpha - \beta)x + (\alpha + \beta) = 0x^2 + 0x + 0$

$$\begin{cases} \alpha = 0 \\ -2\alpha - \beta = 0 \Longrightarrow \beta = 0 \\ \alpha + \beta = 0 \end{cases}$$

Por lo tanto, el conjunto A' es linealmente independiente.

Podríamos considerar otro subconjunto $A'' = \{(1-x), 1\}$ y también es linealmente independiente.

$\alpha(1-x) + \beta = 0x^2 + 0x + 0$

$\alpha - \alpha x + \beta = 0x^2 + 0x + 0$

$-\alpha x + (\alpha + \beta) = 0x^2 + 0x + 0$

$$\begin{cases} -\alpha = 0 \Longrightarrow \alpha = 0 \\ \alpha + \beta = 0 \Longrightarrow \beta = 0 \end{cases}$$

Por lo tanto, el conjunto A'' es linealmente independiente.

P9) Si un conjunto tiene más vectores que entradas en cada vector, entonces el conjunto es linealmente dependiente. Es decir, el conjunto $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_p}\} \subset \mathbb{R}^n$ es linealmente dependiente si $p > n$.

Para comprender mejor esta propiedad, haremos en primer lugar una aclaración para los vectores en $\mathbb{R}^n$.

Sean los vectores $\overrightarrow{u} = (u_1, u_2) \in \mathbb{R}^2$, $\overrightarrow{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$ y de esta manera en forma general $\overrightarrow{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$, se suele decir que $\overrightarrow{u}$ tiene "dos entradas" $u_1, y\, u_2$ mientras que $\overrightarrow{v}$ tiene "tres entradas", $v_1, v_2\ y\ v_3$, y así sucesivamente $\overrightarrow{z}$ tiene "n entradas".

D) Sea la matriz $M = (\overrightarrow{v_1} \quad \overrightarrow{v_2} \quad \ldots \quad \overrightarrow{v_p})$ cuyas columnas son los vectores de A, entonces la dimensión de M es $n\ x\ p$ y la ecuación $Mx = N$ (N matriz nula de orden n) corresponde a un sistema de n ecuaciones con p incógnitas. Si $p > n$ significa que hay más variables que ecuaciones por lo que debe existir al menos una variable libre. De esta manera, $Mx = N$ tiene una solución no trivial y las columnas de M son linealmente dependientes, es decir que A es linealmente dependiente.

Ejemplo: sean los vectores $(1, 0), (2, 1), (-1, 0)$ de $\mathbb{R}^2$. Aquí $n = 2$ (todos los vectores tienen 2 entradas) y $p = 3$ (número de vectores), es decir $p > n$.

Comprobemos que los vectores dados son linealmente dependientes.

Directamente podemos escribir la matriz ampliada del sistema.

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right) \sim^{-2F_2 + F_1} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right)$$

Es un sistema compatible indeterminado. Por lo tanto, el conjunto de vectores es linealmente dependiente. Reconstruimos el sistema.

$\begin{cases} \alpha - \delta = 0 \\ \beta = 0 \end{cases}$ entonces $\alpha = \delta$.

Para finalizar esta sección enunciaremos, de acuerdo a los conceptos anteriores, las siguientes proposiciones equivalentes.

Sea $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\} \subset (V, +, K, \cdot)$

i) A es linealmente independiente.

ii) Toda combinación lineal de A, cuyo resultado sea el vector nulo, es la trivial.

iii) Ningún vector de A es combinación lineal de los demás.

Análogamente:

i) A es linealmente dependiente.

ii) Existe una combinación lineal de A con escalares no todos nulos, cuyo resultado es el vector nulo.

iii) Algún vector de A es combinación lineal de los demás.

6. Sistema de generadores

Una de las ideas más importantes que tenemos que asociar con la de espacio vectorial es la de conjunto generador. Si un conjunto no vacío de vectores de un espacio vectorial es tal que todo vector de ese espacio vectorial puede expresarse como combinación lineal de dicho conjunto, entonces se dice que éste es un sistema de generadores del espacio vectorial. Esto equivale a decir que el subespacio generado por tal conjunto es el mismo espacio vectorial.

Esta idea nos produce dos definiciones. La primera está relacionada con la idea de que todo vector del espacio vectorial es expresado como combinación lineal del conjunto dado, esto se expresa en la siguiente:

Definición: sea el espacio vectorial $(V,+,K,\cdot)$ y $A=\{\overrightarrow{v_1},\overrightarrow{v_2},\dots,\overrightarrow{v_n}\}$ un subconjunto finito de V, se dice que A es sistema de generadores de V si y solo si todo elemento de V se puede expresar como combinación lineal de los vectores de A. Es decir:

Sean $(V,+,K,\cdot)$ y $A=\{\overrightarrow{v_1},\overrightarrow{v_2},\dots,\overrightarrow{v_n}\}\subset V$, A es S. G. de $V \Leftrightarrow \forall\,\overrightarrow{v}\in V: \overrightarrow{v}=\sum_{i=1}^{n}\alpha_i\,\overrightarrow{v_i}$

La idea –en la práctica– con esta definición es plantear una combinación lineal de los vectores de A que tenga por resultado un vector genérico del espacio vectorial y averiguar si existen los escalares propuestos que verifiquen la combinación lineal.

La segunda definición se relaciona con el conjunto de todas las combinaciones lineales de los vectores de A o sea el subespacio generado, $\overline{A}$, el cual debe coincidir con el mismo espacio vectorial V. Esta idea se expresa en la siguiente:

Definición: sea el espacio vectorial $(V,+,K,\cdot)$ y $A=\{\overrightarrow{v_1},\overrightarrow{v_2},\dots,\overrightarrow{v_n}\}$ un subconjunto finito de V, se dice que A es sistema de generadores de V si y solo si el subespacio generado por A es igual V. Es decir:

Sean $(V,+,K,\cdot)$ y $A=\{\overrightarrow{v_1},\overrightarrow{v_2},\dots,\overrightarrow{v_n}\}\subset V$, A es S. G. de $V \Leftrightarrow \overline{A}=V$.

En la práctica, aplicar esta definición significa verificar que el espacio generado por A es igual al espacio vectorial con el que se está trabajando.

El concepto de sistema de generadores es independiente del concepto de dependencia lineal, es decir, un sistema de generadores puede ser linealmente dependiente o linealmente independiente.

Ejemplo 1: determinar si el conjunto $A = \{(1,0),(0,1),(1,1)\}$ es un sistema de generadores de $\mathbb{R}^2$.

La ecuación vectorial es $\alpha(1,0) + \beta(0,1) + \delta(1,1) = (x,y)$. Podemos escribir directamente la matriz ampliada del sistema de ecuaciones lineales.

$\begin{pmatrix} 1 & 0 & 1 & x \\ 0 & 1 & 1 & y \end{pmatrix}$ la que directamente nos indica que se trata de un sistema compatible indeterminado.

$\begin{cases} \alpha + \delta = x \\ \beta + \delta = y \end{cases}$ Si δ es variable libre, la solución sería $\alpha = x - \delta, \beta = y - \delta \; y \; \delta \in \mathbb{R}$.

Se puede calcular también $\overline{A}$.

$\overline{A} = \{\alpha(1,0) + \beta(0,1) + \delta(1,1)\}$

$\overline{A} = \{\ (\alpha + \delta,\ \beta + \delta)\}$

$\overline{A} = \{\ (x,y)/x = \alpha + \delta,\ y = \beta + \delta, \delta \in \mathbb{R}\}$

Estos resultados nos indican que efectivamente A es un sistema de generadores de $\mathbb{R}^2$ y además que cualquier vector de $\mathbb{R}^2$ puede expresarse de infinitas maneras como combinación lineal de los vectores de A. Por otra parte, es muy fácil verificar que A es un conjunto linealmente dependiente, pues $(1,1) = 1\ (1,0) + 1(0,1)$.

Ejemplo 2: sea $(\mathbb{P}(x), +, \mathbb{R}, \cdot)$ con $\mathbb{P}(x)$, polinomio de variable x, coeficientes reales y grado menor o igual a 2. Determinar si el conjunto $A = \{(1-x)^2, (1-x), 1\}$ es sistema de generadores.

$\alpha(1-x)^2 + \beta(1-x) + \gamma = ax^2 + bx + c$

$\alpha - 2\alpha x + \alpha x^2 + \beta - \beta x + \gamma = ax^2 + bx + c$

$\alpha x^2 + (-2\alpha - \beta)x + (\alpha + \beta + \gamma) = ax^2 + bx + c$

$$\begin{cases} \alpha = a \\ -2\alpha - \beta = b \implies \beta = -b - 2a \\ \alpha + \beta + \gamma = c \implies \gamma = a + b + c \end{cases}$$

Por lo tanto, el conjunto A es sistema de generadores. Por otra parte, se demostró en el apartado 5. Dependencia lineal, mediante el ejemplo 3 que A es un conjunto linealmente independiente.

Ejemplo 3: determinar si $A = \{(1, -3, 2), (3, 0, 0), (-1, -6, 4)\}$ es sistema de generadores de $\mathbb{R}^3$.

$\alpha(1, -3, 2) + \beta(3, 0, 0) + \gamma(-1, -6, 4) = (a, b, c)$

$$\begin{cases} \alpha + 3\beta - \gamma = a \\ -3\alpha - 6\gamma = b \\ 2\alpha + 4\gamma = c \end{cases}$$

Tomando la 2° y 3° ecuación:

$$\begin{cases} -3\alpha - 6\gamma = b \\ 2\alpha + 4\gamma = c \end{cases} \implies \begin{cases} -6\alpha - 12\gamma = 2b \\ 6\alpha + 12\gamma = 3c \end{cases}$$

$2b + 3c = 0 \implies c = -\frac{2b}{3}$

Esto significa que los vectores de A no generan cualquier vector de $\mathbb{R}^3$, es decir que, según lo definido, no es un sistema de generadores de $\mathbb{R}^3$, pero generan las ternas de números reales tales que la tercera componente es igual a $-\frac{2}{3}$ de la segunda.

En estos casos se dice que los vectores dados generan un subespacio, S, de $\mathbb{R}^3$ tal que:

$S = \left\{(a,b,c) \in \mathbb{R}^3 / \; c = -\frac{2b}{3}\right\}$

En la práctica, también se pueden combinar las dos definiciones de sistema de generadores como se muestra en el siguiente ejemplo.

Ejemplo 4: sean $(\mathbb{R}^{2x2}, +\mathbb{R}, \cdot)$ el espacio vectorial de las matrices reales de orden dos con traza nula y sea el conjunto $M = \left\{V_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, V_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, V_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right\}$. Determinar si M es sistema de generadores o determinar $\overline{M}$.

$\overline{M} = \left\{\sum_{i=1}^{3} \alpha_i \, \overrightarrow{v_i} / \alpha_i \in \mathbb{R} \wedge \overrightarrow{v_i} \in M\right\}$

Todo vector de $\overline{M}$ es una matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ tal que

$$\alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$\begin{cases} \alpha = a \\ \beta = b \\ \delta = c \\ -\alpha = d \end{cases}$ de la primera y última ecuación se deduce que $a = -d$

Luego $a = -d, b \; y \; c$ son números reales.

Entonces M genera el conjunto de matrices $\left\{\begin{pmatrix} a & b \\ c & -a \end{pmatrix} / a,b,c \in \mathbb{R}\right\} = \overline{M}$, o sea, genera $(\mathbb{R}^{2x2}, +\mathbb{R}, \cdot)$ el espacio vectorial de las matrices reales de orden dos con traza nula, lo que es lo mismo decir que M es sistema de generadores.

Es muy directo determinar que M es linealmente independiente.

$$\alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$\begin{cases} \alpha = 0 \\ \beta = 0 \\ \delta = 0 \\ -\alpha = 0 \end{cases}$ esto nos indica que $\alpha = \beta = \delta = 0$ con lo que M es linealmente independiente.

Luego de estos ejemplos podríamos preguntarnos si para cada espacio vectorial hay un conjunto finito de vectores que lo generen. La respuesta es NO y se puede justificar con un contraejemplo. Específicamente, el espacio de las funciones reales de variable real no tiene un conjunto finito de funciones que lo generen, lo mismo ocurre con el espacio vectorial de polinomios de coeficientes reales y una variable.

En el ejemplo 2 de este apartado se demostró que el conjunto $A = \{(1-x)^2, (1-x), 1\}$ del espacio vectorial $(\mathbb{P}(x), +, \mathbb{R}, \cdot)$ con $\mathbb{P}(x)$, polinomio de variable x, coeficientes reales y grado menor o igual a 2, es sistema de generadores mientras que, en el apartado 5. Dependencia lineal y mediante el ejemplo 3, se demostró que el mismo conjunto es linealmente independiente.

Se demostraron las mismas características –sistema de generadores y linealmente independiente– para el conjunto $M = \left\{V_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, V_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, V_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right\}$ del espacio vectorial de las matrices reales de orden dos con traza nula $(\mathbb{R}^{2x2}, + \mathbb{R}, \cdot)$.

Los sistemas de generadores que tienen la propiedad de ser linealmente independiente son los conjuntos que se estudiaran en el próximo apartado.

7. Base de un espacio vectorial

En geometría plana aprendimos a asociar un par de números reales (x, y) –llamados coordenadas– con un punto P en el plano al proyectar P sobre un par de ejes de coordenadas perpendiculares, tal cual se muestra en la Figura 7.5.

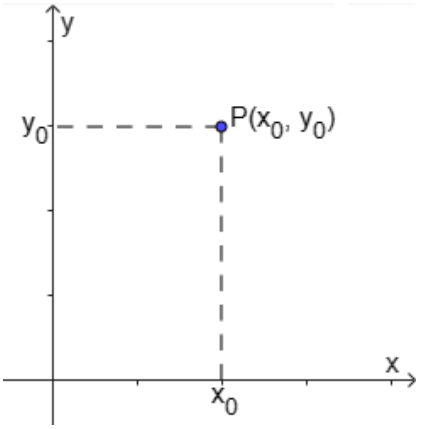


Figura 7.5. Coordenadas de un punto.

Por medio de este proceso, a cada punto en el plano se le asigna un conjunto de coordenadas (par de números reales) único y recíprocamente, a cada par de coordenadas se asocia un único punto en el plano. Esto suele explicarse, afirmando que el sistema de coordenadas establece una correspondencia biunívoca o uno a uno entre los puntos en el plano y los pares ordenados de números reales.

A pesar de que los ejes de coordenadas perpendiculares son los más comunes para definir un sistema de coordenadas en el plano, se puede utilizar cualquier par de rectas no paralelas. Por ejemplo, en la Figura 7.6.

al punto P se han asociado las coordenadas (x_0, y_0) al proyectar P en forma paralela a los ejes de coordenadas no perpendiculares.

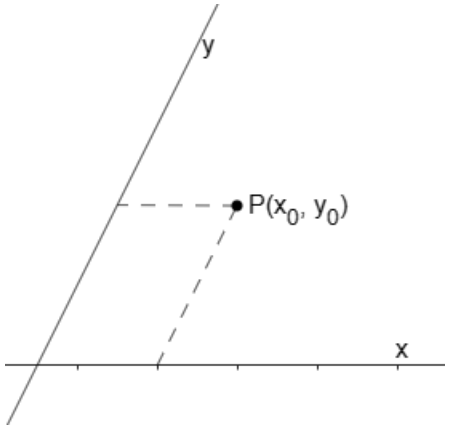


Figura 7.6. Coordenadas con ejes no perpendiculares.

De manera similar, para definir un sistema de coordenadas en el espacio tridimensional es posible usar cualquier terna de ejes de coordenadas no coplanares, como se muestra en la Figura 7.7.

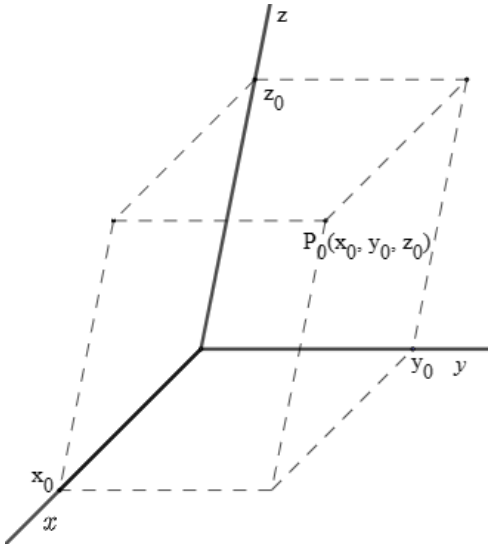


Figura 7.7. Coordenadas con ejes no perpendiculares.

Antes de formalizar cualquier definición en este apartado, ampliaremos el concepto de sistema de coordenadas a espacios vectoriales generales. Para comenzar, será de mucha utilidad volver a plantear el concepto de sistema de coordenadas en el espacio bidimensional o en el tridimensional usando vectores en vez de ejes de coordenadas para especificar el sistema de coordenadas.

Esto se puede lograr si sustituimos cada eje de coordenadas por un vector de longitud 1 que apunte en la dirección positiva del eje. En la Figura 7.8, por ejemplo, $\overrightarrow{v_1}\ y\ \overrightarrow{v_2}$ son esos vectores, tal como se ilustra allí, si P es cualquier punto en el plano, el vector $\overrightarrow{OP}$ se puede escribir como una combinación lineal de $\overrightarrow{v_1}\ y\ \overrightarrow{v_2}$ proyectando P en forma paralela a $\overrightarrow{v_1}\ y\ \overrightarrow{v_2}$ a fin de que $\overrightarrow{OP}$ sea la diagonal del paralelogramo determinado por los vectores $\alpha\ \overrightarrow{v_1}\ y\ \beta\ \overrightarrow{v_2}$ es decir $\overrightarrow{OP} = \alpha\ \overrightarrow{v_1} + \beta\ \overrightarrow{v_2}$.

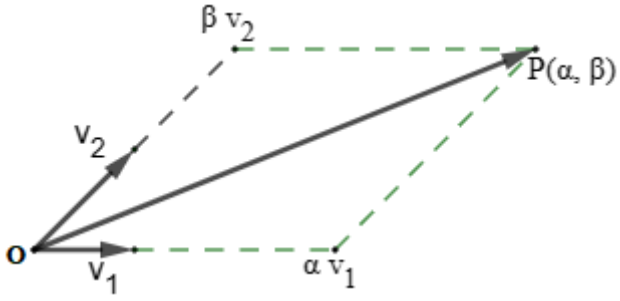


Figura 7.8. Coordenadas con vectores en el plano.

En la ecuación vectorial $\overrightarrow{OP} = \alpha\ \overrightarrow{v_1} + \beta\ \overrightarrow{v_2}$ resulta evidente que los números $\alpha\ y\ \beta$ son precisamente las coordenadas de P en el sistema de coordenadas de la figura. De manera similar, las coordenadas (α, β, δ) del

punto P en la Figura 7.9. se pueden obtener al expresar $\overrightarrow{OP}$ como una combinación lineal de los vectores $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ y $\overrightarrow{v_3}$.

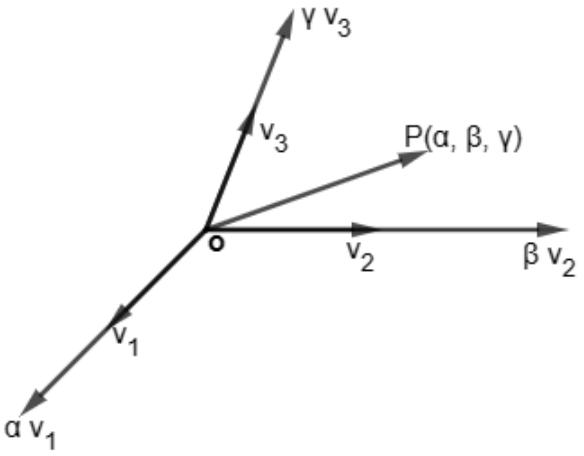


Figura 7.9. Coordenadas con vectores en el espacio.

De manera informal podríamos llamar a los vectores que especifican un sistema de coordenadas "vectores básicos" de ese sistema.

Otra cuestión importante en este análisis es que, si bien se utilizaron vectores básicos de longitud 1, en realidad esto no es esencial puesto que se pueden utilizar vectores no nulos de cualquier longitud. Esto nos lleva a la siguiente idea: las escalas.

En general, se intentará usar la misma escala en cada eje y situar los puntos enteros sobre los ejes a una distancia de una unidad entre sí. Cuando esto no sea posible, por cuestiones de practicidad, se puede usar escalas en las cuales las distancias entre los puntos enteros sean mayor o menor que 1 unidad. Esto se suma a lo anterior respecto de los ejes,

perpendiculares o no. Para el caso de los ejes perpendiculares, se lo muestra en la Figura 7.10.

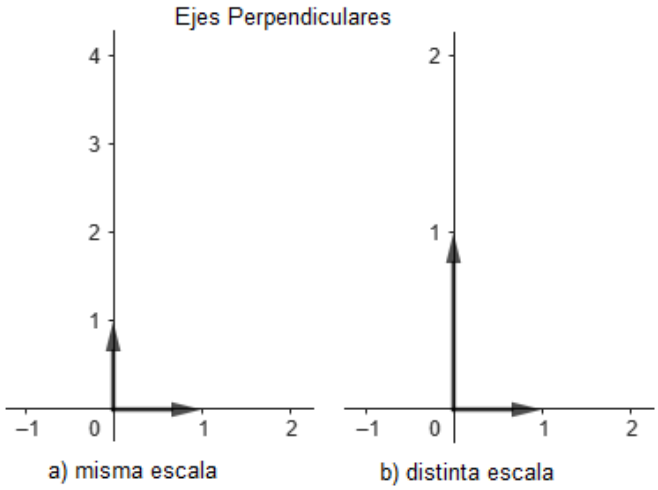


Figura 7.10. Coordenadas ejes perpendiculares y escala en $\mathbb{R}^2$.

La misma idea se puede aplicar en el caso de los ejes no perpendiculares u oblicuos, esto se muestra en la Figura 7.11.

Cuando un sistema de coordenadas se específica mediante un conjunto de vectores básicos, entonces las longitudes en esos vectores corresponden a las distancias entre puntos enteros consecutivos sobre los ejes de coordenadas.

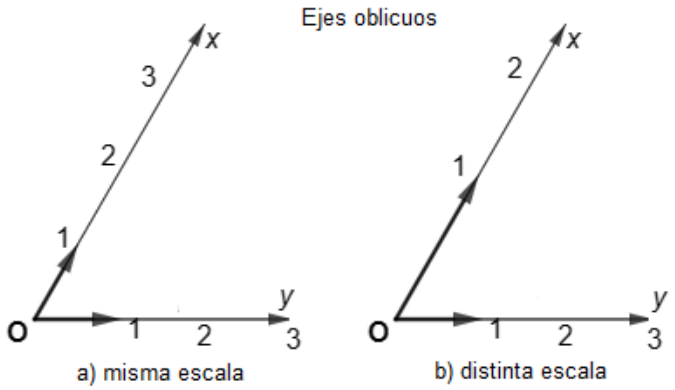


Figura 7.11. Coordenadas ejes oblicuos y escala en $\mathbb{R}^2$.

De esta manera lo que define las direcciones positivas de los ejes de coordenadas son las direcciones de los vectores básicos y lo que establece las escalas de medición son las longitudes de dichos vectores.

Para hacer más precisos los conceptos anteriores y ampliar el concepto de sistema de coordenadas a espacios vectoriales, establecemos la siguiente:

Definición: dado un conjunto de n vectores $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\}$ del espacio vectorial $(V, +, K, \cdot)$ se dice que A es una base si y solo si es linealmente independiente y sistema de generadores de V. Es decir:

Dados $(V, +, K, \cdot)$ y $A = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\}$, $A \subset V$ es base de $V \Leftrightarrow A$ es L I y S de G de V.

Ejemplo 1: una base de $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$ es el conjunto $\{(1,0), (0,1)\}$ y una base de $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$ es $\{(1,0,0), (0,1,0), (0,0,1)\}$. Se puede comprobar fácilmente que los conjuntos dados son linealmente independientes y sistema de generadores de sus respectivos espacios vectoriales. Se sugiere al estudiante comprobarlo.

Estos conjuntos reciben el nombre de bases canónicas. En general una base se denotará como [B].

Ejemplo 2: se demostró que el conjunto $A = \{(1-x)^2, (1-x), 1\}$ de $(\mathbb{P}(x), +, \mathbb{R}, \cdot)$ con $\mathbb{P}(x)$, polinomio de variable x, coeficientes reales y grado menor o igual a 2, es linealmente independiente y sistema de generadores, por lo tanto también constituye una base del respectivo espacio vectorial.

Una base es la generalización de un sistema de coordenadas de un espacio vectorial. Esto se puede poner en evidencia al demostrar el siguiente:

Teorema. Si $[B] = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}\}$ es una base del espacio vectorial $(V, +, K, \cdot)$ entonces todo vector $\overrightarrow{v}$ de V puede expresarse de forma única como combinación lineal de los vectores de [B].

D) Como [B] es sistema de generadores de V significa que todo vector $\overrightarrow{v}$ de V se puede escribir como combinación lineal de los vectores de [B]. Como [B] también es linealmente independiente, significa que esta combinación lineal es única.

Ejemplo 3: construir una base para el subespacio vectorial

$T = \{(x, y, z) \in \mathbb{R}^3 / z = x + y\} \subset (\mathbb{R}^3, +, \mathbb{R}, \cdot)$

Sea (x, y, z) un vector genérico de S, entonces su tercera componente debe ser igual a la primera más la segunda componente, en consecuencia

$(x, y, x + y) = (x, 0, x) + (0, y, y) = x(1, 0, 1) + y(0, 1, 1)$

Este resultado nos indica que el conjunto $\{(1, 0, 1), (0, 1, 1)\}$ es un sistema de generadores de T, además es linealmente independiente ya que:

$\alpha(1,0,1) + \beta\ (0,1,1) = (0,0,0)$

$$\begin{cases} \alpha = 0 \\ \beta = 0 \\ \alpha + \beta = 0 \end{cases}$$

Por lo tanto, el conjunto $\{(1,0,1),(0,1,1)\}$ constituye una base de T.

7.1. Coordenadas respecto a una base

Si $[\ B\ ] = \{\ \overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\ \}$ es una base de $(V,+,K,\cdot)$ entonces, por propiedad, cada vector de V puede expresarse de modo único como combinación lineal de los vectores de la base ya que los vectores de ésta son linealmente independientes y sistema de generadores.

Es decir que si $\overrightarrow{x} \in V$ existen y son únicos los escalares $\alpha_1, \alpha_2, \dots, \alpha_n$ tales que:

$\overrightarrow{x} = \sum_{i=1}^{n} \alpha_i\ \overrightarrow{v_i} = \alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ \dots + \alpha_n\overrightarrow{v_n}$

El vector $\overrightarrow{x} \in V$ queda caracterizado por los coeficientes de la combinación linel, o sea, por la n–tupla de elementos de K: $(\alpha_1, \alpha_2, \dots, \alpha_n)$.

Los escalares α_i se llaman **coordenadas** del vector $\overrightarrow{x} \in V$ **respecto de la base** $[\ B\ ]$.

Si se elige otra base en el espacio V, entonces el mismo vector $\overrightarrow{x}$ admite otras coordenadas $(\alpha'_1, \alpha'_2, \dots, \alpha'_n)$.

Dada una base $[\ B\ ] = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}\}$ del espacio $(V,+,K,\cdot)$ podemos expresar cada vector de V como una matriz fila, cuyos elementos sean las coordenadas del vector respecto de la base $[\ B\ ]$. En tal caso escribiremos $X_{[\ B\ ]} = (\alpha_1, \alpha_2, \dots, \alpha_n\ )$ y este vector recibe el nombre de **vector de coordenadas de $\overrightarrow{x}$ con respecto a** $[\ B\ ]$.

Aquí debemos hacer una observación, nótese que los vectores de coordenadas no solo dependen de la base [B], sino también del orden en que se escriben los vectores básicos, un cambio en el orden de éstos da por resultado un cambio correspondiente en el orden de los elementos en los vectores de coordenadas.

Ejemplo 1: determinar las coordenadas del vector $\vec{x} = (-2, 3) \in (\mathbb{R}^2, +, \mathbb{R}, \cdot)$ respecto de las bases:

a) $[B_1] = \{(1, 1), (1, 0)\}$

$\alpha(1, 1) + \beta(1, 0) = (-2, 3)$

$\begin{cases} \alpha + \beta = -2 \\ \alpha = 3 \end{cases} \Longrightarrow \beta = -5$

Por lo tanto $X_{[B_1]} = (3, -5)$

b) $[B_2] = \{(1, 0), (-1, 3)\}$

$\alpha(1, 0) + \beta(-1, 3) = (-2, 3)$

$\begin{cases} \alpha - \beta = -2 \\ 3\beta = 3 \end{cases} \Longrightarrow \beta = 1 \wedge \alpha = -1$

Por lo tanto $X_{[B_2]} = (-1, 1)$

c) Canónica $[B_3] = \{(1, 0), (0, 1)\}$

$\alpha(1, 0) + \beta(0, 1) = (-2, 3)$

$\begin{cases} \alpha = -2 \\ \beta = 3 \end{cases}$

Por lo tanto $X_{[B_3]} = (-2, 3)$

d) $[B_4] = \{(1, 0), (1, 1)\}$

$\alpha(1, 0) + \beta(1, 1) = (-2, 3)$

$\begin{cases} \alpha + \beta = -2 \\ \beta = 3 \end{cases} \Longrightarrow \alpha = -5$

Por lo tanto $X_{[B_4]} = (-5, 3)$

Compárese los resultados obtenidos con las bases $X_{[B_1]}$ y $X_{[B_4]}$.

La representación de las coordenadas del vector $(-2, 3)$ en las bases $X_{[B_1]}, X_{[B_2]}$ y $X_{[B_3]}$ se representan en la Figura 7.12.

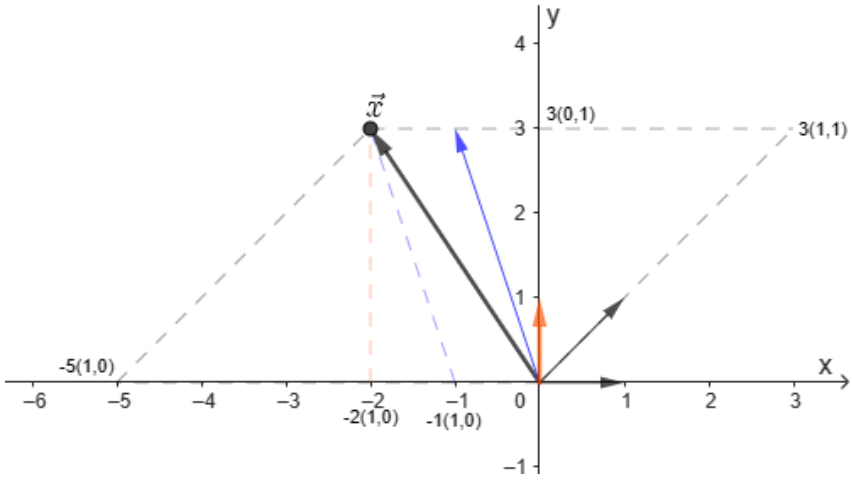


Figura 7.12. Coordenadas en distintas bases.

Ejemplo 2: relación entre sistema de coordenadas cartesianas ortogonales y la base canónica en $\mathbb{R}^3$.

Sabemos –por lo estudiado en el capítulo Vectores en $\mathbb{R}^n$– que los vectores canónicos de $\mathbb{R}^3$ son $\breve{\imath} = (1, 0, 0)$, $\breve{\jmath} = (0, 1, 0)$, $\breve{k} = (0, 0, 1)$.

En el apartado 7. Base de un espacio vectorial, en el ejemplo 1 se sugirió comprobar que el conjunto formado por estos tres vectores forman una base de $\mathbb{R}^3$. Esto equivale a decir que $[\,B\,] = \{(1, 0, 0),\ (0, 1, 0),\ (0, 0, 1).\}$, lo que significa que cualquier vector de $\mathbb{R}^3$ se puede expresar de manera única como:

$\overrightarrow{v} = (v_1, v_2, v_3) = v_1(1,0,0) + v_2(0,1,0) + v_3(0,0,1) = v_1\check{\imath} + v_2\check{\jmath} + v_3\check{k}$ (1)

De esta forma $[B] = \{(1,0,0),\ (0,1,0),\ (0,0,1)\}$ Se denomina base canónica de $\mathbb{R}^3$. Al observar los coeficientes de $\check{\imath}$, $\check{\jmath}$, $\check{k}$ de (1) se concluye que las coordenadas de $\overrightarrow{v}$ respecto de la base canónica son $v_1, v_2\ y\ v_3$, de tal modo que $\overrightarrow{v_{[B]}} = (v_1, v_2, v_3)$. Tal como se muestra en la Figura 7.13.

Comparando este resultado con (1) se observa que $\overrightarrow{v} = \overrightarrow{v_{[B]}}$.

Esta ecuación significa que las componentes de un vector $\overrightarrow{v} \in \mathbb{R}^3$ con respecto a un sistema de coordenadas cartesianas ortogonales, $x\ y\ z$, y las coordenadas del mismo vector con respecto a la base canónica, son las mismas. De esta forma, el sistema de coordenadas y la base canónica producen precisamente la misma correspondencia uno a uno entre puntos en el espacio tridimensional y ternas de números reales.

Sin embargo, hay que destacar que un vector y su vector de coordenadas no son los mismos, la igualdad $\overrightarrow{v} = \overrightarrow{v_{[B]}}$ es una situación especial que ocurre solamente con la base canónica –en este caso de $\mathbb{R}^3$– de $\mathbb{R}^n$.

El resultado de este ejemplo se puede generalizar y establecer que la base canónica de $\mathbb{R}^n$ es $\{e_1 = (1,0,0,\ldots,0), e_2 = (0,1,0,\ldots,0),\ \ldots, e_n = (0,0,0,\ldots,1)\}$.

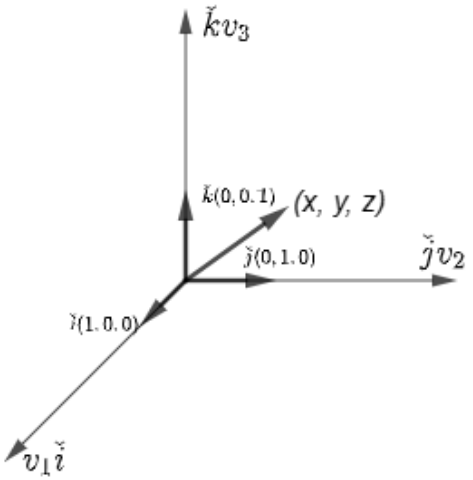


Figura 7.13. Coordenadas en la base canónica.

Ejemplo 3: sea $[B] = \{\overrightarrow{v_1} = (1,2,1), \overrightarrow{v_2} = (2,9,0), \overrightarrow{v_3} = (3,3,4)\}$ una base de $\mathbb{R}^3$.

a) Encontrar las coordenadas (vector) de $\overrightarrow{v} = (5,-1,9)$ con respecto a $[B]$.

Es necesario encontrar los coeficientes $\alpha, \beta\ y\ \delta$ tales que

$\overrightarrow{v} = \alpha\overrightarrow{v_1} + \beta\overrightarrow{v_2} + \delta\overrightarrow{v_2}$

$(5,-1,9) = \alpha(1,2,1) + \beta(2,9,0) + \delta(3,3,4)$

De esta ecuación vectorial, escribimos directamente la matriz ampliada del sistema de ecuaciones lineales:

$\begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 9 & 3 & -1 \\ 1 & 0 & 4 & 9 \end{pmatrix}$ que al aplicar algún método de resolución nos da como única solución $\alpha = 1, \beta = -1, \delta = 2$ (se sugiere al estudiante, comprobar).

Por lo tanto, podemos afirmar $\overrightarrow{v_{[B]}} = (1,-1,2)$.

b) Encontrar el vector $\overrightarrow{v} \in \mathbb{R}^3$ cuyo vector de coordenadas respecto de $[B]$ es:

$\overrightarrow{v_{[B]}} = (-1, 3, 2)$.

En esta situación debemos encontrar el vector que satisface:

$\overrightarrow{v} = \alpha\overrightarrow{v_1} + \beta\overrightarrow{v_2} + \delta\overrightarrow{v_2}$

$(x, y, z) = -1(1, 2, 1) + 3(2, 9, 0) + 2(3, 3, 4)$ de donde $x = 11, y = 31, z = 7$.

Es decir que $\overrightarrow{v} = (11, 31, 7)$.

Para verificar este resultado se sugiere calcular las coordenadas de $\overrightarrow{v}$ en la base $[B]$, debiéndose obtener $\overrightarrow{v_{[B]}} = (-1, 3, 2)$.

Ejemplo 4: sea $M = \left\{M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ comprobar que M es una base de $(\mathbb{R}^{2x2}, +\mathbb{R}, \cdot)$.

Para demostrar que M es sistema de generadores, basta con establecer

$\alpha\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \delta\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \gamma\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de donde

$\alpha = a,\ \beta = b,\ \delta = c\ y\ \gamma = d$

Para demostrar que M es linealmente independiente, proponemos

$\alpha\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \delta\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \gamma\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ de donde

$\alpha = \beta = \delta = \gamma = 0$

Por lo tanto, M es base de $(\mathbb{R}^{2x2}, +\mathbb{R}, \cdot)$. Se la denomina base canónica de $\mathbb{R}^{2x2}$. De manera más general la base canónica de $\mathbb{K}^{m \times n}$ consta de $m \,.\, n$ matrices que tienen un solo 1 y cuyos restantes elementos son ceros.

Las bases de un espacio vectorial tienen una propiedad intrínseca (llamada dimensión) del espacio V que no tiene que ver con la elección particular de la base. Esto se estudiará en el próximo apartado.

7.2. Dimensión de un espacio vectorial

La dimensión de un espacio vectorial es muy importante porque indica cuántos vectores se necesitan para generar dicho espacio. Antes de estudiar la definición de dimensión de un espacio vectorial, necesitamos otros resultados, de los cuales, iniciamos con el siguiente:

Teorema. Sea el espacio vectorial $(V,+,K,\cdot)$ y $[B]=\{\overrightarrow{v_1},\overrightarrow{v_2},\ldots,\overrightarrow{v_n}\}$ una base de V. Si $\overrightarrow{v_1},\overrightarrow{v_2},\ldots,\overrightarrow{v_k}$ son k vectores de V $(k=n+1)$, entonces los k vectores son linealmente dependientes.

D) Resolvemos el sistema:

$\beta_1\overrightarrow{v_1},+\beta_2\overrightarrow{v_2}+\ldots+\beta_k\overrightarrow{v_k}=\overrightarrow{0}$ para $\beta_i\in K, i=1,2,\ldots,k$ (1)

Como $[B]$ es base, existen los escalares $\alpha_{ij}, i=1,2,\ldots,k, j=1,2,\ldots,n$ tales que

$\overrightarrow{v_1}=\alpha_{11}\overrightarrow{v_1}+\alpha_{12}\overrightarrow{v_2}+\ldots+\alpha_{1n}\overrightarrow{v_n}$

$\overrightarrow{v_2}=\alpha_{21}\overrightarrow{v_1}+\alpha_{22}\overrightarrow{v_2}+\ldots+\alpha_{2n}\overrightarrow{v_n}$ (2)

$\vdots = \qquad\qquad \vdots$

$\overrightarrow{v_k}=\alpha_{k1}\overrightarrow{v_1}+\alpha_{k2}\overrightarrow{v_2}+\ldots+\alpha_{kn}\overrightarrow{v_n}$

Sustituimos cada $\overrightarrow{v_i}$ de (2) en (1)

$\beta_1(\alpha_{11}\overrightarrow{v_1}+\alpha_{12}\overrightarrow{v_2}+\ldots+\alpha_{1n}\overrightarrow{v_n})+\beta_2(\alpha_{21}\overrightarrow{v_1}+$

$\alpha_{22}\overrightarrow{v_2}+\ldots+\alpha_{2n}\overrightarrow{v_n})+\ldots+\beta_k(\alpha_{k1}\overrightarrow{v_1}+\alpha_{k2}\overrightarrow{v_2}+\ldots+\alpha_{kn}\overrightarrow{v_n})=\overrightarrow{0}$

Haciendo el producto por un escalar y reagrupando con factor común $\overrightarrow{v_i}$ tenemos

$(\beta_1\alpha_{11}+\beta_2\alpha_{21}+\ldots+\beta_k\alpha_{k1})\overrightarrow{v_1}+(\beta_1\alpha_{12}+$

$\beta_2\alpha_{22}+\ldots+\beta_k\alpha_{k2})\overrightarrow{v_2}+\ldots+(\beta_1\alpha_{1n}+\beta_2\alpha_{2n}+\ldots+\beta_k\alpha_{kn})\overrightarrow{v_n}=\overrightarrow{0}$

Como los $\overrightarrow{v_i}$ son linealmente independiente, se debe cumplir

$$\beta_1\alpha_{11} + \beta_2\alpha_{21} + \ldots + \beta_k\alpha_{k1} = 0$$
$$\beta_1\alpha_{12} + \beta_2\alpha_{22} + \ldots + \beta_k\alpha_{k2} = 0$$
$$\vdots \quad + \quad \vdots \quad + \ldots + \quad \vdots \quad = 0$$
$$\beta_1\alpha_{1n} + \beta_2\alpha_{2n} + \ldots + \beta_k\alpha_{kn} = 0$$

Que es un sistema lineal homogéneo con más incógnitas (k) que ecuaciones (n) puesto que por hipótesis $k > n$, por lo tanto, tiene solución no trivial, lo que significa que existen $\beta_i, i = 1, 2, \ldots, k$ no todos nulos tales que (1) se cumple.

Luego los k vectores $\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_k}$ son linealmente dependientes.

De este teorema se puede deducir un resultado un poco más general "*en un espacio vectorial V que posee una base con n elementos, cualesquiera p vectores de V, con $p > n$ son linealmente dependientes*". Basta con observar que $(n + 1)$ de los p vectores dados son linealmente dependientes por el teorema anterior, por lo tanto los p vectores son linealmente dependientes pues contienen un subconjunto linealmente dependiente.

Este último resultado se aplica al siguiente:

Teorema. Todas las bases de un mismo espacio vectorial poseen el mismo número de elementos.

D) Sea el espacio vectorial $(V, +, K, \cdot)$, $[B] = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}\}$ y $[B'] = \left\{\overrightarrow{v'_1}, \overrightarrow{v'_2}, \ldots, \overrightarrow{v'_m}\right\}$ dos bases de V. Por el razonamiento anterior $m \leq n$, puesto que, en caso contrario, los vectores de la segunda base serían linealmente dependientes. De manera similar $n \leq m$ pues, en caso contrario, los vectores de la primera base serían linealmente dependientes. Por lo tanto $m = n$.

Estos dos resultados nos permiten dar la siguiente:

Definición: el número de elementos que posee una base cualquiera de un espacio vectorial V recibe el nombre de dimensión; este número será designado mediante $dim(V)$.

Si el espacio vectorial contiene un solo elemento, que necesariamente ha de ser su elemento neutro, o sea $V = \{ \vec{0} \}$, decimos que V tiene dimensión cero.

Ejemplo 1: se determinó que una base de $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$ es el conjunto $\{(1,0),(0,1)\}$ por lo tanto $dim(\mathbb{R}^2) = 2$

Ejemplo 2: se determinó que una base de $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$ es el conjunto $V\{(1,0,0),(0,1,0),(0,0,1)\}$, por lo tanto se puede afirmar que $dim(\mathbb{R}^3) = 3$

Se puede demostrar que, en general, $dim(\mathbb{K}^n) = n$.

Ejemplo 3: se demostró que el conjunto $A = \{(1-x)^2, (1-x), 1\}$ de $(\mathbb{P}(x), +, \mathbb{R}, \cdot)$ con $\mathbb{P}(x)$, polinomio de variable x, coeficientes reales y grado menor o igual a 2, es una base del respectivo espacio vectorial; por lo tanto, se puede afirmar que $\dim\big(\mathbb{P}(\mathrm{x})\big) = 3$

Ejemplo 4: se puede demostrar que, en general, $dim(\mathbb{K}^{n \times m}) = n \times m$.

Ejemplo 5: se puede demostrar que, en general, $dim(\mathbb{P}_{\mathbb{K}}^{n}[x]) = n + 1$. Siendo $\mathbb{P}_{\mathbb{K}}^{n}[x]$ los polinomios de grado menor o igual que n y de una indeterminada.

En muchas ocasiones conocemos la dimensión de un espacio vectorial, digamos n, en ese caso el siguiente teorema nos puede ayudar a concluir que un sistema de generadores de n elementos son linealmente independiente y, por consiguiente, base del espacio vectorial, o que n vectores linealmente independientes generan al espacio vectorial y, también, son una base del mismo. De esta manera, nos evitamos de probar la independencia lineal de los vectores o de verificar si generan al espacio. Estos resultados se expresan en el siguiente:

Teorema. Sea el espacio vectorial $(V, +, K, \cdot)$ y $dim(V) = n$, entonces las siguientes condiciones son equivalentes:

i) $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}$ son linealmente independientes.

ii) $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n}$ generan V.

iii) $\{ \overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n} \}$ es una base de V.

Mostraremos que $i) \Longrightarrow ii) \Longrightarrow iii) \Longrightarrow i)$ y tendremos la equivalencia entre las tres condiciones.

Sea $[B]$ una base de V, entonces $[B] = \{ \overrightarrow{u_1}, \overrightarrow{u_2}, \dots, \overrightarrow{u_n} \}$ pues $dim(V) = n$.

$i) \Longrightarrow ii)$. Suponemos que $\overrightarrow{w} \in V$, entonces hemos demostrado que $\{ \overrightarrow{u_1}, \overrightarrow{u_2}, \dots, \overrightarrow{u_n}, \overrightarrow{w} \}$ son linealmente dependientes. Entonces, existen los escalares $\alpha_i, i = 1, 2, \dots, n \; y \; \beta \in \mathrm{K}$ no todos nulos tales que:

$\alpha_1 \overrightarrow{u_1} + \alpha_2 \overrightarrow{u_2} + \dots + \alpha_n \overrightarrow{u_n} + \beta \overrightarrow{w} = \vec{0}$

Si $\beta = 0$, entonces algún $\alpha_i \neq 0$, pero esto es imposible pues los $\overrightarrow{u_i}$ son linealmente independientes, por lo tanto $\beta \neq 0$. Luego $\overrightarrow{w}$ es combinación

lineal de los vectores $\overrightarrow{u_i}$. Dado que $\overrightarrow{w}$ es arbitrario, hemos demostrado que los vectores $\overrightarrow{u_i}$ generan V.

$ii) \Longrightarrow iii)$ Si los vectores $\overrightarrow{u_i}$ son linealmente dependientes, entonces podemos reducir este conjunto a uno linealmente independiente que genere a V con m elementos y $m < n$. Pero esto implicaría que hay una base de V con un número de elementos distinto de n, lo cual es imposible pues se ha demostrado que todas las bases de un mismo espacio vectorial tienen la misma cantidad de elementos. Por lo tanto, los vectores $\overrightarrow{u_i}$ deben ser linealmente independiente, luego forman una base de V.

$iii) \Longrightarrow i)$ es evidente por la definición de base de un espacio vectorial.

Las siguientes son interpretaciones de los resultados de este último teorema:

a) n vectores linealmente independientes de un espacio vectorial n–dimensional constituyen una base del mismo;

b) todo sistema de generadores de n vectores de un espacio vectorial n–dimensional constituye una base del mismo;

c) y, obviamente, toda base de n vectores de un espacio n–dimensional constituye un conjunto linealmente independiente y sistema de generadores del espacio vectorial.

CAPÍTULO 8

TRANSFORMACIONES LINEALES

1. Introducción

La idea de espacio vectorial nos indica que este es básicamente un conjunto de vectores (vectores en $\mathbb{R}^n$, números complejos, polinomios, matrices, entre otros) en los que se definieron dos operaciones que cumplen diez propiedades. La siguiente etapa del estudio consiste en analizar las relaciones que se pueden establecer entre estos conjuntos. El área del Álgebra Lineal que se desarrolló para explicar estas relaciones es lo que hoy se conoce como transformaciones lineales.

Los primeros trabajos sobre transformaciones lineales se remontan hacia 1830. Carl Gustav Jacob Jacobi (1804-1851) y Kronecker fueron pioneros en trabajar con esta noción que empezaba a surgir por esa época; años más tarde, entre 1850 y 1860, se sumó Karl Theodor W. Weierstrass. Sin embargo, el matemático alemán que mayor influencia ejerció sobre el desarrollo del concepto de transformaciones lineales fue Ferdinand Georg Frobenius (1849-1917), que había empezado a estudiarlo gracias a los trabajos de Cauchy, Weierstrass y Kronecker, entre otros, a pesar de esto, no fue sino hasta 1918 que las transformaciones lineales adoptaron su

forma actual de la mano del matemático alemán Hermann Weyl (1885-1955)[44].

Una vez establecida la definición de transformaciones lineales, sin lugar a dudas, una de las mayores contribuciones es aquella en la que se trabaja con un mismo espacio vectorial. En estas transformaciones son de mucho interés los valores y vectores propios. Estos conceptos, según Morris Kline (1908-1992) surgieron en el estudio de formas cuadráticas y el movimiento de los planetas. En primera instancia, los valores propios se conocieron como raíces características de la llamada ecuación escalar. Sin embargo, se sabe que, desde 1740, Euler ya usaba –implícitamente– los valores propios para describir de manera geométrica las formas cuadráticas en tres variables; mientras que, en 1760, Lagrange estudió un sistema de seis ecuaciones diferenciales del movimiento de los planetas (solo se conocían seis) y, a partir de esto, dedujo una ecuación polinomial de sexto grado, cuyas raíces eran los valores propios de una matriz de orden seis. Pero no fue sino hasta el año 1840 que Cauchy fue uno de los primeros en utilizar los términos valores característicos y ecuación característica para indicar los valores propios y la ecuación polinomial básica a la que satisfacen.

En la actualidad, los valores y vectores propios también son conocidos bajo el nombre de eigen valores y eigen vectores (eigen: palabra alemana que se puede traducir como "propio"), valores y vectores característicos

[44] Ver Gutiérrez González, E. y Ochoa García, S. (2014). *Álgebra lineal y sus aplicaciones*. Grupo Editorial Patria S.A. D. F. México.

o como autovalores y autovectores; sin duda, se ha convertido en un tema de gran importancia y utilidad del álgebra lineal. Los valores y vectores propios se usan en diversas áreas de las matemáticas, física, mecánica, ingeniería eléctrica y nuclear, hidrodinámica y aerodinámica, entre otras. Solo, a modo de ejemplo, en la solución de sistemas de ecuaciones diferenciales lineales se puede establecer la ecuación característica que tiene como raíces a los valores propios; o en robótica se requiere el uso constante de rotaciones de cuerpos rígidos. Por medio de transformaciones lineales de los sistemas de coordenadas siempre hay al menos un conjunto de tres direcciones ortogonales alrededor de las cuales el cuerpo puede rotar; para su cálculo se usan vectores propios.

De esta forma iniciamos el capítulo mediante un escueto repaso de funciones y su interpretación; para continuar, en el apartado dos, y luego del análisis de tres ejemplos, con la definición de transformaciones lineales y unas reglas que nos permitirán establecer cuándo una transformación no es lineal. A partir de esto se enuncian y demuestran unas propiedades de las transformaciones lineales y un teorema cuya importancia es fundamental debido a que permite determinar completamente una transformación a partir de conocer las imágenes de los vectores de una base del espacio vectorial dominio.

El apartado tres está dedicado al estudio –definición y propiedades– de dos subespacios vectoriales muy destacados en las transformaciones:

núcleo e imagen; mientras que, en el apartado cuatro detallaremos la representación matricial de las transformaciones lineales.

En los apartados siguientes estudiaremos un tema muy vinculado a las transformaciones, la diagonalización de matrices. En el apartado cinco introducimos el trabajo con endomorfismos para luego definir autovalores, autovectores, polinomio y ecuación característica y estudiar algunas propiedades de estos. Continuamos con la definición de matrices semejantes y su relación con la diagonalización, por medio de distintos teoremas que se demuestran. Para tener una visión general del capítulo y de los temas que se presentan en el mismo, se finaliza con una serie de distintos ejemplos que integran los conceptos vertidos.

1.1. Conceptos básicos

Como el tema central de este capítulo es el estudio de una clase especial de funciones llamadas transformaciones lineales, recordaremos muy brevemente y en primer lugar lo que es una función mediante la siguiente:

Definición: Sean $A\, y\, B$ dos conjuntos no vacíos, una función f con dominio en A y codominio en B es una regla de asignación tal que a cada elemento x de A le asigna un único elemento $y = f(x)$ de B.

Para denotar que f es una función con dominio en A y codominio en B, por lo general se suele escribir como $f: A \longrightarrow B$.

La función (regla de asignación) es f, no debemos confundirla con el valor de ésta en $x: y = f(x)$. En esta última notación, a y se la suele conocer

como variable dependiente o imagen de x bajo f, x es la variable independiente o argumento de la función f.

Ejemplo 1: sea la función $f: \mathbb{Z} \longrightarrow \mathbb{R} / y = 2x - 1$.

a) Calcular $f(-2), f(0), f(3)$. ¿Se puede calcular $f(0{,}5)$ o $f(\pi)$?

En primer lugar, destaquemos que el dominio de f es el conjunto de números enteros, es decir que a x se le puede asignar valores de este campo, mientras que el codominio de f son los reales, vale decir que y puede asumir valores en este campo numérico.

$f(-2) = -5, f(0) = -1$ y $f(3) = 5$. No se puede calcular $f(0{,}5)$ ni $f(\pi)$ pues $0{,}5$ y $\pi \notin \mathbb{Z}$

b) Representar gráficamente f.

Recordemos que, en este caso, por ser el dominio $\mathbb{Z}$ en la representación gráfica se obtienen puntos alineados. Esto se muestra en la Figura 8.1.

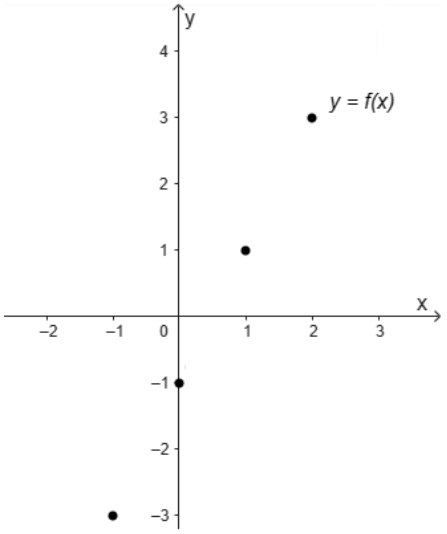


Figura 8.1. Representación gráfica de la función f.

Ejemplo 2: representar gráficamente la función $f: \mathbb{R} \longrightarrow \mathbb{R} \,/ y = x^2 + 2x - 1$

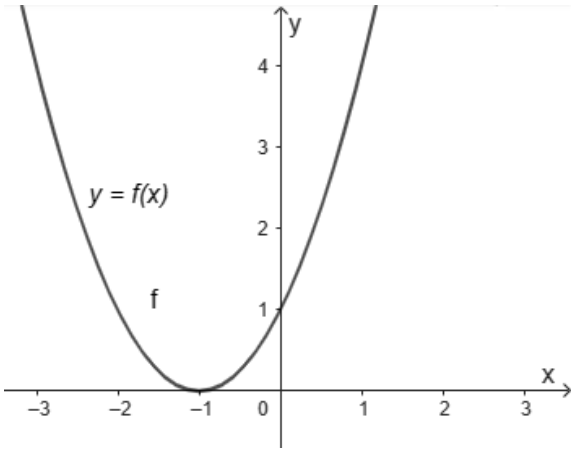


Figura 8.2. Representación gráfica de la función f.

Una manera muy conveniente de interpretar una función f es como un conjunto de procedimientos; una máquina que transforma cada elemento x (materia prima) de A en un producto $y = f(x)$ elemento de B. El resultado de esta transformación –el elemento $y = f(x)$– depende de la materia prima x que se introduce cada vez a la máquina; además, esta transformación está perfectamente determinada por esta máquina o conjunto de procedimientos f (la función o transformación). Es en este sentido que interpretaremos, en este capítulo, el significado de una

función entre espacios vectoriales: como una transformación de un espacio en otro[45].

2. Transformaciones Lineales

Como dijimos anteriormente, los espacios vectoriales son conjuntos, en consecuencia, es posible establecer relaciones entre sus elementos; precisamente, la relación que deseamos establecer entre espacios vectoriales la llamaremos transformación y la denotaremos por f, de modo que la forma de representar que dos espacios vectoriales, $V\ y\ W$, se relacionan es mediante: $f: V \longrightarrow W$ o bien como $T: V \longrightarrow W$. Usaremos indistintamente cualquiera de las notaciones.

Ejemplo 1: la reflexión de un vector en $\mathbb{R}^2$ con respecto al eje x se representa en la Figura 8.3. En términos de transformaciones, la forma de obtener la reflexión de cualquier vector de $\mathbb{R}^2$ con respecto al eje x es mediante la función $T(x, y) = (x, -y)$, esta es una transformación que va de $\mathbb{R}^2$a $\mathbb{R}^2$ y se representa por $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$.

[45] Del Valle Sotelo, J. C. (2011). *Álgebra lineal para estudiantes de ingeniería y ciencias*. D. F. México: Mac Graw Hill Interamericana Editores S. A.

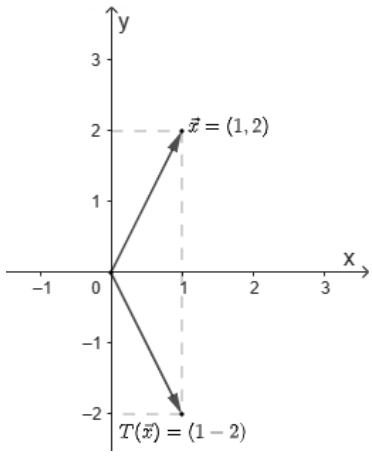


Figura 8.3. Transformación $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$.

Ejemplo 2: en la Figura 8.4. se puede observar la proyección del vector $\overrightarrow{x} = (1,2,3)$ sobre el plano coordenado XY.

La forma de proyectar cualquier vector de $\mathbb{R}^3$ sobre el plano XY es mediante la función $T(x,y,z) = (x,y)$ en donde a $T(x,y,z) = (x,y)$ se la suele llamar expresión de la transformación T. En general esta transformación va de $\mathbb{R}^3$ a $\mathbb{R}^2$ y se la representa por medio de $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$.

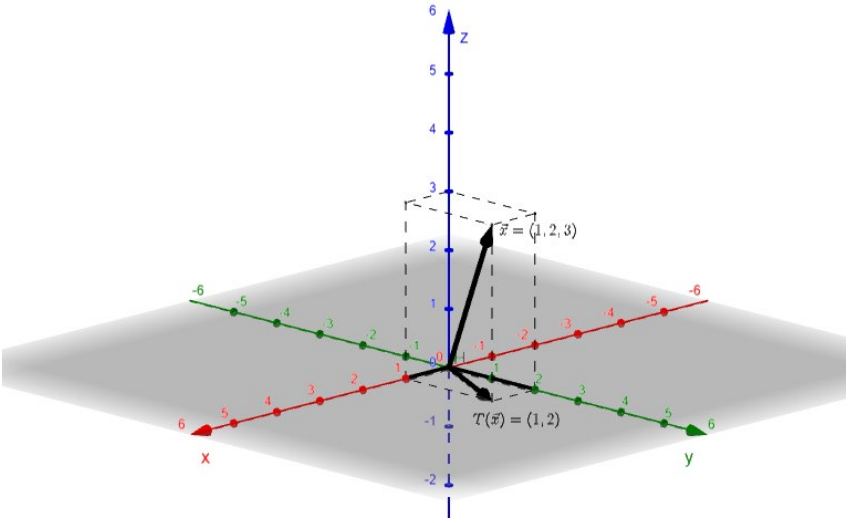


Figura 8.4. Transformación $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$.

Ejemplo 3: la diferencia entre una ecuación matricial $A\,\vec{x} = B$ y la ecuación vectorial asociada $\vec{x_1}a_1 + \ldots + \vec{x_n}a_n = B$ es tan solo cuestión de notación. Sin embargo, es posible encontrar una ecuación matricial $A\,\vec{x} = B$ en álgebra lineal que no esté directamente relacionada con combinaciones lineales de vectores; esto se puede lograr si se piensa en la matriz A como un objeto que "actúa" sobre el vector $\vec{x}$ multiplicándolo para producir un nuevo vector $\overrightarrow{Ax}$. En las ecuaciones

$$\underbrace{\begin{pmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} 5 \\ 8 \end{pmatrix}}_{B} \quad \text{y} \quad \underbrace{\begin{pmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 \\ 4 \\ -1 \\ 3 \end{pmatrix}}_{\vec{v}} = \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{C}$$

Se observa que la multiplicación por la matriz A transforma al vector $\vec{x}$ en B y a $\vec{v}$ en el vector nulo. Desde este nuevo punto de vista, resolver la

ecuación matricial $A\,\vec{x} = B$ equivale a encontrar todos los vectores $\vec{x}$ en $\mathbb{R}^4$ que se transforman en el vector B de $\mathbb{R}^2$ como resultado de la "acción" de la multiplicación por A.

Como podemos observar, no es difícil establecer relaciones entre dos espacios vectoriales. Ahora formalicemos el objeto de estudio –para poder analizar y desarrollar sus propiedades– mediante la siguiente:

Definición: sean $(V, +, K, \cdot)$ y $(W, +, K, \cdot)$ dos espacios vectoriales sobre un mismo cuerpo K, la función $f: V \to W$ es una transformación lineal si y solo si:

i) la imagen de la suma de dos vectores cualesquiera de V es igual a la suma de sus imágenes en W.

ii) la imagen del producto de cualquier escalar por todo vector de V es igual al producto del escalar por la imagen de dicho vector. Es decir:

Sean $(V, +, K, \cdot)$ y $(W, +, K, \cdot)$ dos espacios vectoriales, $f: V \to W$ es una transformación lineal $\Leftrightarrow$

$i)\ f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})\ \ \forall\, \vec{x}, \vec{y} \in V.$

$ii)\ f(\alpha\, \vec{x}) = \alpha f(\vec{x})\ \ \forall\, \vec{x} \in V, \forall \alpha \in K.$

También suele simbolizarse la definición de transformación lineal de la siguiente manera:

Sean $(V, +, K, \cdot)$ y $(W, +, K, \cdot)$ dos espacios vectoriales, $f: V \to W$ es una transformación lineal $\Leftrightarrow$ $f(\alpha\vec{x} + \beta\vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y})\ \ \forall\, \vec{x}, \vec{y} \in V, \forall \alpha, \beta \in K.$

En la literatura matemática, a las transformaciones lineales también se las conoce como aplicaciones lineales u operadores lineales.

Debemos destacar que en la condición ii) de la definición, el primer producto por un escalar ocurre en V, mientras que el segundo producto por un escalar se realiza en W.

La interpretación gráfica de la definición de transformación lineal se presenta en la Figura 8.5.

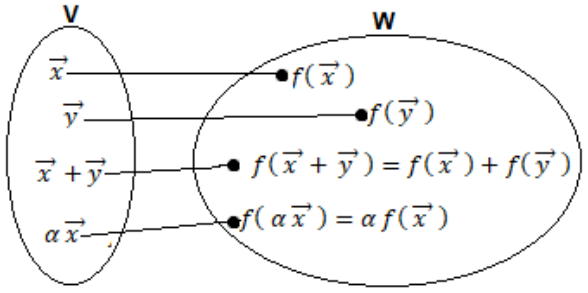


Figura 8.5. Interpretación gráfica de una transformación.

Para comprobar si una transformación $f: V \to W$ entre dos espacios vectoriales V y W es lineal debemos verificar que se cumplan las dos propiedades de la linealidad enunciadas en la definición. Se puede lograr esto de distintas maneras, una de ellas es iniciar la comprobación con la elección de dos vectores arbitrarios, $\vec{x}$ e $\vec{y}$, de V, luego obtenemos el vector suma, $\vec{x}+\vec{y}$, y aplicamos la transformación a la suma, $f(\vec{x}+\vec{y})$. Si se llevan a cabo operaciones algebraicas y se puede concluir $f(\vec{x})+f(\vec{y})$, entonces f cumple con la condición de linealidad de la suma. Se procede de manera similar para la segunda condición de linealidad.

Ejemplo 1: sean los espacios vectoriales $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$ y $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$.

Determinar si $f(x_1, x_2, x_3) = (x_1 - x_3, x_2 - x_3)$, es o no transformación lineal.

$i)\ \overrightarrow{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \Longrightarrow f(\overrightarrow{x}) = (x_1 - x_3, x_2 - x_3)$

$\overrightarrow{y} = (y_1, y_2, y_3) \in \mathbb{R}^3 \Longrightarrow f(\overrightarrow{y}) = (y_1 - y_3, y_2 - y_3)$

$\overrightarrow{x} + \overrightarrow{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3) \Longrightarrow$

$f(\overrightarrow{x} + \overrightarrow{y}) = (x_1 + y_1 - x_3 - y_3, x_2 + y_2 - x_3 - y_3)$

$f(\overrightarrow{x} + \overrightarrow{y}) = (x_1 - x_3 + y_1 - y_3, x_2 - x_3 + y_2 - y_3)$

$f(\overrightarrow{x} + \overrightarrow{y}) = (x_1 - x_3, x_2 - x_3) + (y_1 - y_3 + y_2 - y_3)$

$f(\overrightarrow{x} + \overrightarrow{y}) = f(\overrightarrow{x}) + f(\overrightarrow{y})$

$ii)\ \vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \Longrightarrow f(\overrightarrow{x}) = (x_1 - x_3, x_2 - x_3)$

$\alpha\ \overrightarrow{x}(\alpha x_1, \alpha x_2, \alpha x_3) \in \mathbb{R}^3 \Longrightarrow f(\alpha\ \overrightarrow{x}) = (\alpha x_1 - \alpha x_3, \alpha x_2 - \alpha x_3)$

$f(\alpha\ \overrightarrow{x}) = (\alpha x_1 - \alpha x_3, \alpha x_2 - \alpha x_3)$

$f(\alpha\ \overrightarrow{x}) = \alpha(x_1 - x_3,\ x_2 - x_3)$

$f(\alpha\ \overrightarrow{x}) = \alpha\ f(\overrightarrow{x})$

Por $i)$ y $ii)$ la función dada es una transformación lineal.

Otra forma de comprobar que una función es transformación lineal consiste en iniciar la comprobación con la elección de dos vectores arbitrarios, $\overrightarrow{x}$ e $\overrightarrow{y}$, de V; luego, a cada uno le aplicamos la transformación y después las sumamos, $f(\overrightarrow{x}) + f(\overrightarrow{y})$. Por separado, calculamos la suma $\overrightarrow{x} + \overrightarrow{y}$, y después aplicamos la transformación propuesta $f(\overrightarrow{x} + \overrightarrow{y})$, si ambos resultados son iguales, f cumple con la linealidad de la suma. Para la segunda condición se procede de manera similar.

Ejemplo 2: sean los espacios vectoriales $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$ y $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$.

Determinar si $f(x_1, x_2, x_3) = (x_1 - x_3, x_2 + 1)$, es o no transformación lineal.

$i)\ \overrightarrow{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \Longrightarrow f(\overrightarrow{x}) = (x_1 - x_3, x_2 + 1)$

$\overrightarrow{y} = (y_1, y_2, y_3) \in \mathbb{R}^3 \Longrightarrow f(\overrightarrow{y}) = (y_1 - y_3, y_2 + 1)$

$f(\overrightarrow{x}) + f(\overrightarrow{y}) = (x_1 - x_3, x_2 + 1) + (y_1 - y_3, y_2 + 1)$

$f(\overrightarrow{x}) + f(\overrightarrow{y}) = ((x_1 - x_3) + (y_1 - y_3); (x_2 + 1) + (y_2 + 1))$

$f(\overrightarrow{x}) + f(\overrightarrow{y}) = ((x_1 + y_1) - (x_3 + y_3); (x_2 + y_2) + 2)$ (1)

$\overrightarrow{x} + \overrightarrow{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$

$f(\overrightarrow{x} + \overrightarrow{y}) = [(x_1 + y_1) - (x_3 + y_3), (x_2 + y_2) + 1]$ (2)

Como (1) y (2) son distintos, concluimos que en este caso f no cumple con la linealidad de la suma y, por lo tanto, no es una transformación lineal. No hace falta comprobar la condición ii) pues la definición establece que para ser transformación lineal se deben cumplir ambas condiciones.

Este ejemplo nos permite destacar que en ciertas transformaciones podemos decidir de manera intuitiva cuando no es lineal si consideramos las siguientes reglas.

Sean $V\ y\ W$ dos espacios vectoriales relacionados por la transformación $f: V \to W$, para probar que f no es lineal:

i) Se propone un ejemplo de $\overrightarrow{x}\ e\ \overrightarrow{y}$, de V o algún $\alpha \in K$ que no cumplan con alguna de las dos condiciones de linealidad.

Ejemplo: sea $f: \mathbb{R}^3 \longrightarrow \mathbb{P}_2(x) / f(a, b, c) = (a^2 + bx + c^2x^2)$, verificaremos que no es una transformación lineal por medio de un contraejemplo.

Sean $\overrightarrow{u} = (-1,\ 2,\ 2) \in \mathbb{R}^3 \Longrightarrow f(-1,\ 2,\ 2) = (-1)^2 + 2x + 2^2x^2 = 1 + 2x + 4x^2$

$\overrightarrow{v} = (3, -1, 4) \in \mathbb{R}^3 \Longrightarrow f(3, -1, 4) = 3^2 + (-1)x + 4^2x^2 = 9 - x + 16x^2$

Por lo tanto $f(\overrightarrow{u}) + f(\overrightarrow{v}) = 10 + x + 20x^2$

Pero $f(\overrightarrow{u} + \overrightarrow{v}) = f(2, 1, 6) = 2^2 + 1\,x + 6^2x^2 = 4 + x + 36x^2$

Por lo tanto $f(\overrightarrow{u}) + f(\overrightarrow{v}) \neq f(\overrightarrow{u}) + f(\overrightarrow{v})$ y f no es una transformación lineal.

ii) Si en la expresión que define f aparecen componentes de los elementos de V multiplicándose, por ejemplo $f(x, y) = xy$.

Ejemplo: sea $f: \mathbb{P}_3(x) \longrightarrow \mathbb{P}_1(x) / f(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0a_3 + (a_1 + a_2)x$, una transformación. Decidir si es lineal o no.

Sea $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \in \mathbb{P}_3(x)$ el polinomio que se va a transformar, cuyas componentes son $a_0, a_1, a_2\, y\, a_3$. Como en la transformación aparece el producto a_0a_3 entre sus componentes, entonces por ii) se puede afirmar que f no es lineal.

iii) Si en la expresión que define f aparecen componentes de los elementos de V elevados a una potencia diferente de 1, por ejemplo $f(x, y) = y^3 - 3x$.

iv) Si en la expresión que define f aparecen componentes de los elementos de V que se suman con constantes diferentes de cero, por ejemplo $f(x, y) = 2y - x + 2$.

Ejemplo: sea la transformación $f: \mathbb{P}_2(x) \longrightarrow \mathbb{R}^4 / f(a_0 + a_1x + a_2x^2) = (a_0, a_1, a_2, -2)$, decir si es o no lineal justificando la respuesta.

Sea $p(x) = a_0 + a_1x + a_2x^2 \in \mathbb{P}_2(x)$ el polinomio que se va a transformar, cuyas componentes son $a_0, a_1, y\ a_2$. En la transformación aparece la constante no nula, el –2, que se suma en la cuarta componente de la transformación, por lo tanto por iv) se puede afirmar que f no es lineal.

v) Si en la expresión que define f aparecen componentes de los elementos de V con operaciones diferentes a la suma y producto por un escalar, por ejemplo $f(x, y) = \frac{x+y}{y-3x}$

Ejemplo: sea $f: \mathbb{M}_{2x2} \longrightarrow \mathbb{R}$ una transformación definida por $f: \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{a+b}{c+d}$. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ la matriz que se va a transformar, cuyas componentes son $a, b, c, y\ d$. Como en transformación aparece una división entre componentes, entonces por v) podemos afirmar que f no es una transformación lineal.

vi) Si el vector nulo de V no se transforma en el vector nulo de W. Por ejemplo, se define $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2 / f(x, y) = (2x, y + 1)$, se tiene $f(0, 0) = (0, 1) \neq (0, 0)$.

Una vez analizada una transformación para saber si esta es lineal o no, corresponde demostrar algunas propiedades de las transformaciones lineales que son de utilidad para ampliar las aplicaciones de estos objetos matemáticos, además de que en ciertos casos se pueden usar para resolver problemas de una forma más sencilla.

2.1. Propiedades de las transformaciones lineales

Sean V y W dos espacios vectoriales, $f: V \to W$ una transformación lineal, $\alpha_1, \ldots, \alpha_n \in K$ escalares y $\overrightarrow{v_1}, \ldots, \overrightarrow{v_n} \in V$, entonces se cumple:

P1) La imagen del vector nulo del primer espacio por toda transformación lineal es el vector nulo del segundo espacio. Es decir $f(\overrightarrow{0_V}) = \overrightarrow{0_W}$, siendo $\overrightarrow{0_V}$ y $\overrightarrow{0_W}$ los vectores nulos de V y W respectivamente.

D) Sea $\overrightarrow{v} \in V$ entonces

$\overrightarrow{v} + \overrightarrow{0_V} = \overrightarrow{v}$

$f(\overrightarrow{v} + \overrightarrow{0_V}) = f(\overrightarrow{v})$

$f(\overrightarrow{v}) + f(\overrightarrow{0_V}) = f(\overrightarrow{v})$

$f(\overrightarrow{v}) + f(\overrightarrow{0_V}) = f(\overrightarrow{v}) + \overrightarrow{0_W}$

$f(\overrightarrow{0_V}) = \overrightarrow{0_W}$

P2) La imagen del opuesto de todo vector del primer espacio es igual al opuesto de su imagen.

D) $f(-\overrightarrow{v}) = -f(\overrightarrow{v}) \; \forall \overrightarrow{v} \in V$

$f(-\overrightarrow{v}) = f[(-1)\overrightarrow{v}]$

$f(-\overrightarrow{v}) = (-1)f(\overrightarrow{v})$

$f(-\overrightarrow{v}) = -f(\overrightarrow{v})$

P3) Las transformaciones lineales preservan las combinaciones lineales, es decir, $f(\sum_{i=1}^{n} \alpha_i \overrightarrow{v_i}) = \sum_{i=1}^{n} \alpha_i f(\overrightarrow{v_i}) \;\; \forall i \in \mathbb{N}$.

D) $f(\sum_{i=1}^{n} \alpha_i \overrightarrow{v_i}) = \sum_{i=1}^{n} \alpha_i f(\overrightarrow{v_i}) \;\; \forall i \in \mathbb{N}$.

$f(\alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\overrightarrow{v_n}) = \alpha_1 f(\overrightarrow{v_1}) + \alpha_2 f(\overrightarrow{v_2}) + \ldots + \alpha_n f(\overrightarrow{v_n})$

$f(\alpha_1\overrightarrow{v_1}) + f(\alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\overrightarrow{v_n}) = \alpha_1 f(\overrightarrow{v_1}) + \alpha_2 f(\overrightarrow{v_2}) + \ldots + \alpha_n f(\overrightarrow{v_n})$

$\alpha_1 f(\overrightarrow{v_1}) + f(\alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\overrightarrow{v_n}) = \alpha_1 f(\overrightarrow{v_1}) + \alpha_2 f(\overrightarrow{v_2}) + \ldots + \alpha_n f(\overrightarrow{v_n})$

$\alpha_1 f(\overrightarrow{v_1}) + f(\alpha_2\overrightarrow{v_2}) + f(\alpha_3\overrightarrow{v_3} \ldots + \alpha_n\overrightarrow{v_n}) = \alpha_1 f(\overrightarrow{v_1}) +$

$\alpha_2 f(\overrightarrow{v_2}) + \ldots + \alpha_n f(\overrightarrow{v_n})$

$\alpha_1 f(\overrightarrow{v_1}) + \alpha_2 f(\overrightarrow{v_2}) + f(\alpha_3\overrightarrow{v_3} \ldots + \alpha_n\overrightarrow{v_n}) = \alpha_1 f(\overrightarrow{v_1}) +$

$\alpha_2 f(\overrightarrow{v_2}) + \ldots + \alpha_n f(\overrightarrow{v_n})$

Continuamos con este proceso de utilizar la linealidad de la transformación por pares de vectores hasta terminar con $\alpha_{n-1}\overrightarrow{v_{n-1}}$ y $\alpha_n\overrightarrow{v_n}$. Quedando de este modo demostrada la propiedad.

Ejemplo: supongamos que $f: \mathbb{R}^2 \to \mathbb{R}^3$ es una transformación lineal de la cual se sabe que $f(1,0) = (1,2,3)$ y $f(0,1) = (0,-1,2)$. Nos piden calcular $f(2,-3)$

$f(2,-3) = f[(2,0) + (0,-3)]$

$f(2,-3) = f[2(1,0) + (-3)(0,1)]$

$f(2,-3) = f[2(1,0)] + f[(-3)(0,1)]$

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

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

$f(2,-3) = (2,4,6) + (0,3,-6)$

$f(2,-3) = (2,7,0)$

Destaquemos que la P1) justifica la regla vi) del apartado anterior y muestra un método muy simple para detectar cuándo una transformación no es lineal, es decir que podemos hacer la siguiente:

Observación: sea $f: V \to W$. Si al aplicar la transformación f al neutro de V no resulta el neutro de W, entonces f no es lineal.

Ejemplo: sea $f:\mathbb{R}^3 \to \mathbb{R}^2 \,/\, f(x,y,z) = (x+2y, z+4)$, no es una transformación lineal puesto que el neutro de V es $(0,0,0)$ pero $f(0,0,0) = (0,4) \neq (0,0)$ que es el neutro de W.

Nótese que en el ejemplo de la P3) los vectores de $\mathbb{R}^2$ son los vectores canónicos y forman justamente la base canónica de ese espacio; esta observación se puede, en primer lugar, generalizar y luego demostrar. La generalización nos indicaría que si conocemos las imágenes –por una transformación lineal $f:V \to W$– de los vectores de una base de V, entonces se puede conocer la imagen de cualquier vector de V.

Veamos otro ejemplo para aclarar aún más la idea.

Sea $f:\mathbb{R}^3 \to \mathbb{R}^2$ tal que $f(1,0,0) = (2,3), f(0,1,0) = (-1,4)\ y\ f(0,0,1) = (5,-3)$ calcular $f(3,-4,5)$.

Observemos que conocemos las imágenes de los vectores de una base de $\mathbb{R}^3$, que por simplicidad recurrimos a la base canónica, pero puede ser cualquier base del espacio dominio.

Lo que hacemos es colocar el vector cuya imagen nos piden que calculemos, como combinación lineal de los vectores de la base, lo que se puede hacer por P3).

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

Aplicamos la transformación lineal en ambos miembros de la igualdad y aplicamos las condiciones de linealidad:

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

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

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

$f(3,-4,5) = (6,9) + (4,-16) + (25,-15)$

$f(3,-4,5) = (35,-22)$

Aquí podría surgir otra pregunta, si $\overrightarrow{w_1}, \overrightarrow{w_2}, \ldots, \overrightarrow{w_n}$ son vectores de W ¿existirá una transformación lineal tal que $f(\overrightarrow{v_i}) = \overrightarrow{w_i}$ para $i = 1, 2, \ldots, n$? La respuesta a este interrogante es afirmativa y se concreta mediante la demostración del siguiente:

Teorema. Sean los espacios vectoriales V de dimensión finita con base $[B] = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}\}$ y W que contiene los vectores $\{\overrightarrow{w_1}, \overrightarrow{w_2}, \ldots, \overrightarrow{w_n}\}$; entonces existe una transformación lineal única $f: V \to W$ tal que $f(\overrightarrow{v_i}) = \overrightarrow{w_i}$ para $i = 1, 2, \ldots, n$.

D) Se define la función f como sigue:

i) $f(\overrightarrow{v_i}) = \overrightarrow{w_i}$

ii) Si $\overrightarrow{v} = \alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\overrightarrow{v_n}$ entonces,

$f(\overrightarrow{v}) = \alpha_1\overrightarrow{w_1} + \alpha_2\overrightarrow{w_2} + \ldots + \alpha_n\overrightarrow{w_n}$ (1)

Como B es una base de V, f está definida para todos $\overrightarrow{v} \in V$ y como W es un espacio vectorial, $f(\overrightarrow{v}) \in W$. Por lo tanto, f existe.

Ahora hay que demostrar que f es lineal. Lo que se puede deducir de (1).

Sean $\overrightarrow{u} = \alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\overrightarrow{v_n}$ y $\overrightarrow{v} = \beta_1\overrightarrow{v_1} + \beta_2\overrightarrow{v_2} + \ldots + \beta_n\overrightarrow{v_n}$ entonces

i) $f(\overrightarrow{u} + \overrightarrow{v}) = f((\alpha_1 + \beta_1)\overrightarrow{v_1} + (\alpha_2 + \beta_2)\overrightarrow{v_2} + \ldots + (\alpha_n + \beta_n)\overrightarrow{v_n})$

$f(\overrightarrow{u} + \overrightarrow{v}) = (\alpha_1 + \beta_1)\overrightarrow{w_1} + (\alpha_2 + \beta_2)\overrightarrow{w_2} + \ldots + (\alpha_n + \beta_n)\overrightarrow{w_n}$

$f(\overrightarrow{u} + \overrightarrow{v}) = (\alpha_1\overrightarrow{w_1} + \alpha_2\overrightarrow{w_2} + \ldots + \alpha_n\overrightarrow{w_n}) + (\beta_1\overrightarrow{w_1} + \beta_2\overrightarrow{w_2} + \ldots + \beta_n\overrightarrow{w_n})$

$f(\overrightarrow{u} + \overrightarrow{v}) = f(\overrightarrow{u}) + f(\overrightarrow{v})$

ii) Sea $\lambda \in K$ al aplicar la transformación al producto $\lambda\overrightarrow{u}$ tenemos

$f(\lambda\overrightarrow{u}) = f(\lambda(\alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\overrightarrow{v_n}))$

$f(\lambda\,\overrightarrow{u}) = f((\lambda\,\alpha_1)\overrightarrow{v_1} + (\lambda\,\alpha_2)\overrightarrow{v_2} + \ldots + (\lambda\,\alpha_n)\,\overrightarrow{v_n})$

$f(\lambda\,\overrightarrow{u}) = (\lambda\,\alpha_1)\overrightarrow{w_1} + (\lambda\,\alpha_2)\overrightarrow{w_2} + \ldots + (\lambda\,\alpha_n)\,\overrightarrow{w_n}$

$f(\lambda\,\overrightarrow{u}) = \lambda(\alpha_1\overrightarrow{w_1} + \alpha_2\overrightarrow{w_2} + \ldots + \alpha_n\,\overrightarrow{w_n})$

$f(\lambda\,\overrightarrow{u}) = \lambda\,f(\overrightarrow{v})$

Como se cumplen i) y ii), las dos condiciones de linealidad, se concluye que la transformación construida es lineal.

Resta solamente demostrar la unicidad de la transformación construida.

Suponemos que existen dos transformaciones lineales, $f\ y\ g$. Como B es una base para cualquier $\overrightarrow{v} \in V$ existen únicos escalares $\alpha_1, \alpha_2, \ldots + \alpha_n$ tal que $\overrightarrow{v} = \alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\,\overrightarrow{v_n}$.

Aplicamos las dos transformaciones a $\overrightarrow{v}$:

$f(\overrightarrow{v}) = f(\alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\,\overrightarrow{v_n}) = \alpha_1 f(\overrightarrow{v_1}) + \alpha_2 f(\overrightarrow{v_2}) +$

$\ldots + \alpha_n\,f(\overrightarrow{v_n}) = \alpha_1\overrightarrow{w_1} + \alpha_2\overrightarrow{w_2} + \ldots + \alpha_n\,\overrightarrow{w_n}$

$g(\overrightarrow{v}) = g(\alpha_1\overrightarrow{v_1} + \alpha_2\overrightarrow{v_2} + \ldots + \alpha_n\,\overrightarrow{v_n}) = \alpha_1 g(\overrightarrow{v_1}) + \alpha_2 g(\overrightarrow{v_2}) +$

$\ldots + \alpha_n\,g(\overrightarrow{v_n}) = \alpha_1\overrightarrow{w_1} + \alpha_2\overrightarrow{w_2} + \ldots + \alpha_n\,\overrightarrow{w_n}$

Restando miembro a miembro, resulta:

$f(\overrightarrow{v}) - g(\overrightarrow{v}) = \overrightarrow{0}$

De donde se concluye que $f(\overrightarrow{v}) = g(\overrightarrow{v})$ para cualquier vector de V. Quedando de esta manera demostrado el teorema.

Este teorema nos permite afirmar que para determinar completamente una transformación lineal basta con conocer las imágenes de los elementos de una base del espacio dominio. O sea, el teorema demostrado asegura que toda transformación lineal entre dos espacios vectoriales

queda unívocamente determinada por los valores que toma sobre cualquier base del primero.

Con esta información podemos obtener la imagen de cualquier otro vector o bien la expresión genérica de la transformación lineal.

Observemos que, en este teorema, los vectores $\{\overrightarrow{w_1}, \overrightarrow{w_2}, \ldots, \overrightarrow{w_n}\}$ no tienen que ser linealmente independientes y ni siquiera ser distintos. Es más, se hace hincapié en que el teorema se cumple si V es cualquier espacio vectorial de dimensión finita y nada se dice de la dimensión de W.

Ejemplo: se sabe que una transformación lineal es tal que:

$f(1,0) = (2,1,3)$ y $f(0,1) = (4,-1,0)$. Determinar la transformación lineal.

En primer lugar, se trata de una transformación lineal $f: \mathbb{R}^2 \to \mathbb{R}^3$ y puesto que conocemos las imágenes de la transformación sobre una base de $\mathbb{R}^2$, es decir, conocemos las imágenes de $\{(1,0),(0,1)\}$, podemos calcular la imagen de cualquier vector (x_1, x_2).

En efecto:

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

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

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

$f(x_1, x_2) = x_1(2,1,3) + x_2(4,-1,0)$

$f(x_1, x_2) = (2x_1 + 4x_2, x_1 - x_2, 3x_1)$

Para simplificar el ejemplo y tener una clara idea de lo que se quiere explicar, tomamos la base canónica del espacio dominio, pero de no ser ésta, simplemente se debe comprobar que el conjunto de vectores, de los

cuales conocemos sus imágenes, es linealmente independiente y sistema de generadores.

Veamos cómo trabajar cuando tenemos la información de una base que no es la canónica.

Ejemplo: sea la transformación lineal $f: \mathbb{R}^3 \to \mathbb{R}^2$ tal que asigna a los vectores de la base $\{(1,1,1,),(1,1,0),(1,0,0)\}\, en\ \mathbb{R}^3$, los vectores $(1,2),(1,2)\ y\ (-1,1)\ \ en\ \mathbb{R}^2$ respectivamente. Determinar la expresión genérica de f.

Necesitamos determinar la imagen de un vector genérico (x,y,z) de $\mathbb{R}^3$. Expresamos a este como combinación lineal de los vectores de la base dada.

$(x,y,z) = \alpha(1,1,1) + \beta(1,1,0) + \gamma(1,0,0)$

$\begin{cases} \alpha+\beta+\gamma = x \\ \alpha+\beta = y \\ \alpha = z \end{cases}$ de donde $\begin{cases} \gamma = x-y \\ \beta = y-z \\ \alpha = z \end{cases}$

Por lo tanto:

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

$f(x,y,z) = f[z(1,1,1) + (y-z)(1,1,0) + (x-y)(1,0,0)]$

$f(x,y,z) = zf(1,1,1) + (y-z)f(1,1,0) + (x-y)f(1,0,0)$

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

$f(x,y,z) = (z,2z) + (y-z,2y-2z) + (y-x,x-y)$

$f(x,y,z) = (z+y-z+y-x,2z+2y-2z+x-y)$

$f(x,y,z) = (2y-x,y+x)$

$f: \mathbb{R}^3 \to \mathbb{R}^2\ / f(x,y,z) = (2y-x,y+x)$ que es lo pedido.

3. Núcleo e imagen de una transformación lineal

Toda transformación lineal tiene asociados dos importantes subespacios: el núcleo y la imagen, a partir de estos conjuntos se pueden obtener de manera inmediata algunas propiedades de las transformaciones. En particular interesa destacar la forma de calcular estos dos subespacios y destacar una relación entre las dimensiones de los mismos y la dimensión del espacio dominio. Iniciamos este apartado con la siguiente:

Definición: el núcleo de una transformación lineal f entre dos espacios vectoriales, V y W, sobre un mismo cuerpo, K, es el conjunto de vectores del primer espacio cuyas imágenes por f son el vector nulo del segundo espacio. Es decir:

Sea $f: V \to W$ una transformación lineal, $\Longrightarrow N_f = \{ \vec{x} \in V / f(\vec{x}) = \overrightarrow{0_W} \}$

Gráficamente:

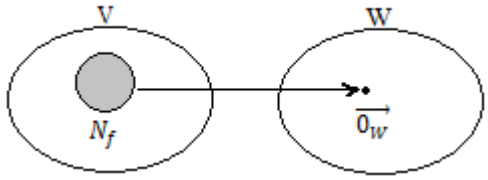


Figura 8.6. Interpretación gráfica del núcleo.

Esta definición nos indica que un vector de V es un elemento del núcleo si y solo si su imagen es el vector nulo de W. Es decir que: $\vec{x} \in N_f \Leftrightarrow f(\vec{x}) = \overrightarrow{0_W}$, esto significa que básicamente el núcleo está constituido por

aquellos vectores del espacio dominio a los cuales la transformación le hace corresponder el vector nulo del espacio codominio.

Observamos, en primer lugar, que el núcleo es subconjunto de V, $N_f \subset V$ y, por otra parte, las propiedades de las transformaciones lineales establecen que $f(\overrightarrow{0_V}) = \overrightarrow{0_W}$, quiere decir que siempre $\overrightarrow{0_V} \in N_f$. Por lo tanto, el núcleo de toda transformación lineal f, es no vacío.

Propiedad del núcleo

El núcleo de toda transformación lineal, f, entre dos espacios vectoriales, $V\ y\ W$ es un subespacio vectorial del primero.

Sea $f: V \to W$ una transformación lineal, $\Longrightarrow (N_f, +, K, \cdot)$ es subespacio vectorial de $(V, +, K, \cdot)$

Para la demostración de esta propiedad recurrimos al criterio de subespacio.

D) $i)\ \overrightarrow{v_1} \in N_f \implies f(\overrightarrow{v_1}) = \overrightarrow{0_W}$

$\overrightarrow{v_2} \in N_f \implies f(\overrightarrow{v_2}) = \overrightarrow{0_W}$

¿ $(\overrightarrow{v_1} + \overrightarrow{v_2}) \in N_f$?

$f(\overrightarrow{v_1} + \overrightarrow{v_2}) = f(\overrightarrow{v_1}) + f(\overrightarrow{v_2})$

$f(\overrightarrow{v_1} + \overrightarrow{v_1}) = \overrightarrow{0_W} + \overrightarrow{0_W}$

$f(\overrightarrow{v_1} + \overrightarrow{v_2}) = \overrightarrow{0_W}$

Por lo tanto $(\overrightarrow{v_1} + \overrightarrow{v_2}) \in N_f$

$ii)\ \overrightarrow{v} \in N_f \implies f(\overrightarrow{v}) = \overrightarrow{0_W}$

$\alpha \in K \implies \alpha = \alpha$

¿ $\alpha\, \overrightarrow{v} \in N_f$?

$f(\alpha\,\vec{v}\,) = \alpha f(\,\vec{v}\,)$

$f(\alpha\,\vec{v}\,) = \alpha\,\overrightarrow{0_W}$

$f(\alpha\,\vec{v}\,) = \overrightarrow{0_W}$

Por lo tanto $\alpha\,\vec{v} \in N_f$

Por $i)\ y\ ii)$ $(N_f, +, K, \cdot)$ es subespacio vectorial de $(V, +, K, \cdot)$

Si deseamos determinar el subespacio núcleo de $f: V \to W$, debemos considerar un vector genérico $\vec{x} \in V$ y evaluar $f(\,\vec{x}\,) = \overrightarrow{0_W}$, de donde resulta un sistema de ecuaciones lineales homogéneo.

Ejemplo: determinar el núcleo de la transformación lineal, una base y dimensión del mismo.

$f: \mathbb{R}^3 \to \mathbb{R}^2 / f(x_1, x_2, x_3) = (x_1 - x_3, x_2 - x_3)$

$N_f = \{(x_1, x_2, x_3)\ \in \mathbb{R}^3 / f(x_1, x_2, x_3) = (0, 0)\ \}$

$(x_1, x_2, x_3)\ \in N_f \Leftrightarrow f(x_1, x_2, x_3) = (0, 0)$

$(x_1 - x_3, x_2 - x_3) = (0, 0)$

$\begin{cases} x_1 - x_3 = 0 \\ x_2 - x_3 = 0 \end{cases}$ Entonces

$\left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right)$

El sistema es compatible indeterminado, consideramos x_3 como variable libre. Entonces,

$\begin{cases} x_1 = x_3 \\ x_2 = x_3 \end{cases}$ o sea que $x_1 = x_2 = x_3$ con $x_3 \in \mathbb{R}$.

Por lo tanto,

$N_f = \{(x_1, x_2, x_3)\ \in \mathbb{R}^3 / x_1 = x_2 = x_3 = k\ \}$

$N_f = \{(k, k, k)\ /\ k \in \mathbb{R}\ \}$

Para determinar la base de N_f hacemos:

$(k, k, k) = k(1, 1, 1)$

Esto quiere decir que cualquier vector del núcleo es generado por el vector $(1, 1, 1)$ es decir que el conjunto $\{(1, 1, 1)\}$ es sistema de generadores; además, por ser un conjunto formado por un vector no nulo es linealmente independiente. Por lo tanto, una base del núcleo es $\left[B_{N_f}\right] = \{(1, 1, 1)\}$ cuya dimensión expresamos como: $dim(N_f) = 1$

La interpretación gráfica del núcleo de esta transformación lineal sería la siguiente

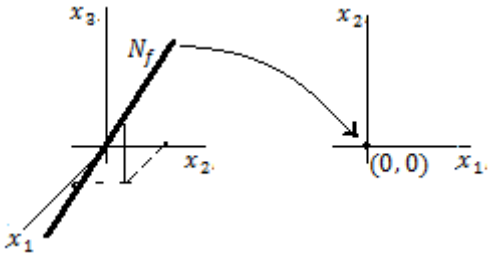


Figura 8.7. Interpretación gráfica del núcleo de f.

Imagen de una transformación lineal $f: V \to W$ es el conjunto imagen del dominio, o sea, es la totalidad de las imagenes de los vectores del espacio dominio. Esto significa que la imagen está constituida por los vectores del espacio codominio que son imagen de algún vector del espacio dominio. Esto se expresa en la siguiente,

Definición: la imagen de una transformación lineal f entre dos espacios vectoriales, V y W, sobre un mismo cuerpo, K, es el conjunto imagen del dominio, es decir, es la totalidad de las imágenes de los vectores del primer espacio. Es decir,

Sea $f: V \to W$ una transformación lineal, $\Longrightarrow I_f = \{ \overrightarrow{y} \in W / \exists \overrightarrow{x} \in \mathrm{V} \wedge f(\overrightarrow{x}) = \overrightarrow{y} \}$

Gráficamente:

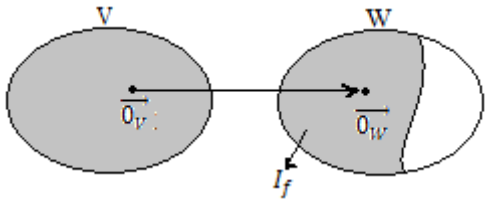


Figura 8.8. Interpretación gráfica de la imagen.

Observamos, en primer lugar, que la imagen es subconjunto de W, $I_f \subset W$ y, por otra parte, las propiedades de las transformaciones lineales establecen que $f(\overrightarrow{0_V}) = \overrightarrow{0_W}$, quiere decir que siempre $\overrightarrow{0_W} \in I_f$. Por lo tanto, la imagen de toda transformación lineal f, es un conjunto no vacío.

Propiedad de la imagen

La imagen de toda transformación lineal, f, entre dos espacios vectoriales, V y W, es un subespacio vectorial del segundo espacio.

Sea $f: V \to W$ una transformación lineal, $\Longrightarrow (I_f, +, K, \cdot)$ es subespacio vectorial de $(W, +, K, \cdot)$

D) $i)$ $\overrightarrow{w_1} \in I_f \Longrightarrow \exists \overrightarrow{v_1} \in V / f(\overrightarrow{v_1}) = \overrightarrow{w_1}$

$\overrightarrow{w_2} \in I_f \Longrightarrow \exists \overrightarrow{v_2} \in V / f(\overrightarrow{v_2}) = \overrightarrow{w_2}$

¿ $(\overrightarrow{w_1} + \overrightarrow{w_2}) \in I_f$?

$\overrightarrow{w_1} + \overrightarrow{w_2} = f(\overrightarrow{v_1}) + f(\overrightarrow{v_2})$

$\overrightarrow{w_1} + \overrightarrow{w_2} = f(\overrightarrow{v_1} + \overrightarrow{v_2})$

Como $\overrightarrow{v_1} \in V \wedge \overrightarrow{v_2} \in V \Longrightarrow (\overrightarrow{v_1} + \overrightarrow{v_2}) \in V$

$\therefore \exists (\overrightarrow{v_1} + \overrightarrow{v_2}) \in V / f(\overrightarrow{v_1} + \overrightarrow{v_2}) = \overrightarrow{w_1} + \overrightarrow{w_2}$

Por lo tanto $(\overrightarrow{w_1} + \overrightarrow{w_2}) \in I_f$

$ii)\ \overrightarrow{w} \in I_f \implies \exists \overrightarrow{v} \in V / f(\overrightarrow{v}) = \overrightarrow{w}$

$\alpha \in K \implies \alpha = \alpha$

$¿\alpha\ \overrightarrow{w} \in I_f\ ? \implies \exists \alpha\overrightarrow{v} \in V / f(\alpha\ \overrightarrow{v}) = \alpha\overrightarrow{w}$

$\alpha\overrightarrow{w} = \alpha f(\overrightarrow{v})$

$\alpha\overrightarrow{w} = f(\alpha\ \overrightarrow{v})$

$\overrightarrow{v} \in V \implies \alpha\overrightarrow{v} \in V$

$\therefore \exists \alpha\overrightarrow{v} \in V / f(\alpha\ \overrightarrow{v}) = \alpha\overrightarrow{w}$

Por lo tanto $\alpha\ \overrightarrow{w} \in I_f$

Por $i)\ y\ ii)$ $(I_f, +, K, \cdot)$ es subespacio vectorial de $(W, +, K, \cdot)$

Cuando un vector $\overrightarrow{y} \in W$ pertenece a la imagen de f debemos buscar un vector $\overrightarrow{x} \in V$ tal que $f(\overrightarrow{x}) = \overrightarrow{y}$. Si el vector $\overrightarrow{x}$ se encuentra entonces $\overrightarrow{y} \in I_f$. Al igual que en el núcleo el vector $\overrightarrow{x}$ se busca en su forma general, de donde resulta un sistema de ecuaciones lineales, si este tiene solución, entonces $\overrightarrow{y} \in I_f$.

Ejemplo: determinar la imagen de la transformación lineal, una base y dimensión del mismo. $f: \mathbb{R}^3 \to \mathbb{R}^2 / f(x_1, x_2, x_3) = (x_1 - x_3, x_2 - x_3)$

Es la misma transformación utilizada en el núcleo.

$I_f = \{(y_1, y_2)\ \in \mathbb{R}^2 / \exists (x_1, x_2, x_3) \in \mathbb{R}^3 \wedge f(x_1, x_2, x_3) = (y_1, y_2)\ \}$

$f(x_1, x_2, x_3) = (y_1, y_2)$

$(x_1 - x_3, x_2 - x_3) = (y_1, y_2)$

$\begin{cases} x_1 - x_3 = y_1 \\ x_2 - x_3 = y_2 \end{cases}$ Entonces,

$\left(\begin{array}{ccc|c} 1 & 0 & -1 & y_1 \\ 0 & 1 & -1 & y_2 \end{array}\right)$ Es un sistema compatible indeterminado; consideramos x_3 como variable libre.

$\begin{cases} x_1 = y_1 + x_3 \\ x_2 = y_2 + x_3 \end{cases}$ Restando miembro a miembro,

$x_1 - x_2 = y_1 - y_2$. Por lo tanto,

$I_f = \{(y_1, y_2) \in \mathbb{R}^2 / y_1 \in \mathbb{R} \ \wedge y_2 \in \mathbb{R}\}$

Gráficamente, la imagen de la transformación lineal es todo el plano, o sea $\mathbb{R}^2$.

Para calcular una base de I_f hacemos,

$(y_1, y_2) = (y_1, 0) + (0, y_2)$

$(y_1, y_2) = y_1(1, 0) + y_2(0, 1)$

Lo que significa que cualquier vector de la imagen es generado por el conjunto $\{(1, 0), (0, 1)\}$, es decir que este conjunto es sistema de generadores, además sabemos que la base canónica de $\mathbb{R}^2$ es linealmente independiente. Por lo tanto, el conjunto $\{(1, 0), (0, 1)\}$ es una base de la imagen y su dimensión es 2, es decir $dim(I_f) = 2$

Hemos calculado el núcleo, la imagen y una base para cada subespacio, de la misma transformación lineal, $f: \mathbb{R}^3 \to \mathbb{R}^2 / f(x_1, x_2, x_3) = (x_1 - x_3, x_2 - x_3)$. Consideremos ahora el siguiente detalle:

$dim(N_f) = 1,\ dim(I_f) = 2,\ dim(\mathbb{R}^3) = 3\ y\ dim(\mathbb{R}^2) = 2$

Observamos que $dim(N_f) + dim(I_f) = dim\mathbb{R}^3$

Esto –la suma de las dimensiones del núcleo y de la imagen es igual a la dimensión del espacio dominio– se verifica para cualquier transformación lineal definida en un espacio vectorial de dimensión finita. Esto se demuestra en la siguiente:

Propiedad: si $(V,+,K,\cdot)$ es un espacio vectorial de dimensión finita y $f: V \to W$ es una transformación lineal, entonces la suma de las dimensiones del núcleo y de la imagen de f es igual a la dimensión del primer espacio. Es decir:

Sea $f: V \to W$ una transformación lineal y $dim(V) = n \implies dim(N_f) + dim(I_f) = dim(V)$

D) Partimos de una base de V, $[B] = \{\overrightarrow{v_1}, \ldots, \overrightarrow{v_k}, \overrightarrow{v_{k+1}} \ldots, \overrightarrow{v_n}\}$ tal que $\{\overrightarrow{v_{k+1}} \ldots, \overrightarrow{v_n}\}$ es una base del N_f, es decir que, $f(\overrightarrow{v_{k+1}}) = \ldots = f(\overrightarrow{v_n}) = \overrightarrow{0_W}$.

Todo vector $\overrightarrow{v} \in V$ puede escribirse como combinación lineal de los vectores de la base

$\overrightarrow{v} = \alpha_1\overrightarrow{v_1} + \ldots + \alpha_k\overrightarrow{v_k} + \alpha_{k+1}\overrightarrow{v_{k+1}} + \ldots + \alpha_n\overrightarrow{v_n}$, aplicamos f a ambos miembros de la igualdad

$f(\overrightarrow{v}) = f(\alpha_1\overrightarrow{v_1} + \ldots + \alpha_k\overrightarrow{v_k} + \alpha_{k+1}\overrightarrow{v_{k+1}} + \ldots + \alpha_n\overrightarrow{v_n})$

$f(\overrightarrow{v}) = \alpha_1 f(\overrightarrow{v_1}) + \ldots + \alpha_k f(\overrightarrow{v_k}) + \alpha_{k+1} f(\overrightarrow{v_{k+1}}) + \ldots + \alpha_n f(\overrightarrow{v_n})$

$f(\overrightarrow{v}) = \alpha_1 f(\overrightarrow{v_1}) + \ldots + \alpha_k f(\overrightarrow{v_k})$ por lo tanto,

$\{f(\overrightarrow{v_1}), \ldots, f(\overrightarrow{v_k})\}$ genera I_f.

Además $\{f(\overrightarrow{v_1}), \ldots, f(\overrightarrow{v_k})\}$ es linealmente independiente pues,

$\overrightarrow{0_W} = \alpha_1 f(\overrightarrow{v_1}) + \ldots + \alpha_k f(\overrightarrow{v_k})$

$\overrightarrow{0_W} = f(\alpha_1\overrightarrow{v_1} + \ldots + \alpha_k\overrightarrow{v_k})$ entonces,

$(\alpha_1\overrightarrow{v_1} + \ldots + \alpha_k\overrightarrow{v_k}) \in N_f$ y entonces,

$\alpha_1\overrightarrow{v_1} + \ldots + \alpha_k\overrightarrow{v_k} = \beta_{k+1}\overrightarrow{v_{k+1}} + \ldots + \beta_n\overrightarrow{v_n}$ Pero por ser los $\overrightarrow{v_i}$ linealmente independientes,

$\alpha_1 = \ldots = \alpha_k = \beta_{k+1} = \ldots = \beta_n = 0$ Por lo tanto,

$dim\, I_f + dim N_f = k + (n - k)$

$dim\, I_f + dim N_f = n = dim\, V$ quedando de esta forma demostrada la propiedad.

4. Representación matricial de una transformación lineal

Hemos vistos que para determinar completamente una transformación lineal es suficiente con conocer las imágenes de una base del espacio dominio. A partir de esta idea existe otra eficaz forma de describir una aplicación lineal, mediante una matriz. Más precisamente, cada matriz de $\mathbb{K}^{m \times n}$ determina una única transformación lineal $f: V \to W$ y, recíprocamente, cada transformación lineal, $f: V \to W$, se puede asociar a una matriz de $\mathbb{K}^{m \times n}$ única si fijamos una base en cada espacio.

En primer lugar, veamos que a toda transformación lineal se le puede asociar una matriz.

Consideraremos una transformación lineal f entre dos espacios vectoriales $V\, y\, W$ de dimensiones finitas $n\, y\, m$, respectivamente. Probaremos que, si se fija una base en cada espacio, entonces f queda caracterizada por una matriz de dimensión $m \times n$, que llamaremos matriz de la transformación lineal respecto del par de bases dado.

Sean $(V, +, K, \cdot)\, y\, (W, +, K, \cdot)$ dos espacios vectoriales de dimensión $n\, y\, m$ respectivamente, $f: V \to W$ una transformación lineal y, además, $[B_1] = \{\overrightarrow{v_1}, \dots\ \overrightarrow{v_n}\}$ una base de V y una base de W, $[B_2] = \{\overrightarrow{w_1}, \dots\ \overrightarrow{w_m}\}$.

Sea un vector cualquiera $\overrightarrow{x} \in V$, por ser $[B_1]$ una base, existen los escalares $\alpha_1, \dots, \alpha_n$ únicos tales que: $\overrightarrow{x} = \sum_{j=1}^{n} \alpha_j \overrightarrow{v_j}$ (1).

Estos escalares caracterizan al vector $\overrightarrow{x}$ respecto de la base $[B_1]$ y, como vimos, se llaman las coordenadas de $\overrightarrow{x}$ respecto de dicha base. Podemos escribir:

$$X_{[B_1]} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Si la imagen de $\overrightarrow{x} \in V$ es $\overrightarrow{y} \in W$ se tiene: $\overrightarrow{y} = f(\overrightarrow{x})$

Si $\overrightarrow{y} \in W$, por ser $[B_2]$ una base de W, existen los escalares $\alpha'_1, \dots, \alpha'_m$ únicos tales que podemos escribir $\overrightarrow{y} = \sum_{i=1}^{m} \alpha'_i \overrightarrow{w_i}$ (2) donde los escalares $\alpha'_1, \dots, \alpha'_m$ son las coordenadas de la imagen de $\overrightarrow{x}$ respecto de la base $[B_2]$.

En consecuencia,

$$Y_{[B_2]} = \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \vdots \\ \alpha'_m \end{pmatrix}$$

En el apartado 2.1. se demostró que f queda caracterizada unívocamente por los valores que toma sobre cualquier base de V. O sea, es suficiente obtener las imágenes, por f, de cada vector de la base $[B_1]$. Ahora bien, cada $j = 1, \dots, n$ de $f(\overrightarrow{v_j}) \in W$, puede expresarse como combinación lineal de la base $[B_2]$, es decir,

$f(\overrightarrow{v_j}) = \sum_{i=1}^{m} a_{ij} \overrightarrow{w_i}$ (3) $\forall\, j = 1, \dots, n$

Cada escalar lleva dos subíndices, el primero asociado a cada vector de la base $[B_2]$ y el segundo en correspondencia con el vector de la base $[B_1]$.

Así,

$f(\overrightarrow{v_1}) = a_{11}\overrightarrow{w_1} + a_{21}\overrightarrow{w_2} + \cdots + a_{m1}\overrightarrow{w_m}$

$f(\overrightarrow{v_2}) = a_{12}\overrightarrow{w_1} + a_{22}\overrightarrow{w_2} + \cdots + a_{m2}\overrightarrow{w_m}$

$\vdots$

$f(\overrightarrow{v_n}) = a_{1n}\overrightarrow{w_1} + a_{2n}\overrightarrow{w_2} + \cdots + a_{mn}\overrightarrow{w_m}$

Los $(n\,.\,m)$ escalares a_{ij} que figuran en las combinaciones de los vectores que son imágenes de los elementos de la base $[B_1]$, constituyen una matriz cuya traspuesta, que denotamos con A, recibe el nombre de matriz de la transformación lineal f respecto de la base $[B_1]$ y $[B_2]$. Es decir,

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

La matriz de la transformación lineal es de orden $(m \times n)$ donde m es la dimensión del espacio codominio y n la del espacio dominio.

Por lo tanto, para determinar la matriz asociada a una transformación lineal $f: V \to W$ respecto de una base en cada espacio, se deben seguir los siguientes pasos:

P_1) Determinar las imágenes dadas por f, de los vectores de la base del espacio dominio.

P_2) Expresar estas imágenes como combinación lineal de los vectores de la base del espacio codominio.

P_3) La traspuesta de la matriz de coeficientes es la matriz de la transformación lineal respecto de las bases dadas en ambos espacios.

Ejemplo:

a) Obtener la matriz asociada a la transformación lineal,

f: $\mathbb{R}^{2\times2} \to \mathbb{R}^2$ / $f\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = (x_1 + x_4, x_2 - x_3)$ respecto de las bases canónicas en ambos espacios.

P1) $f\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = (1,0)$

$f\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = (0,1)$

$f\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = (0,-1)$

$f\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = (1,0)$

P2) $(1,0) = \alpha(1,0) + \beta(0,1) \implies \alpha = 1 \ y \ \beta = 0$

$(0,1) = \gamma(1,0) + \delta(0,1) \implies \gamma = 0 \ y \ \delta = 1$

$(0,-1) = \mu(1,0) + \pi(0,1) \implies \mu = 0 \ y \ \pi - 1$

$(1,0) = \sigma(1,0) + \tau(0,1) \implies \sigma = 1 \ y \ \tau = 0$

P3) $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix}$ que es la matriz asociada a f y cuya dimensión es 2×4.

Como segunda instancia, probaremos que si A es la matriz asociada a la transformación lineal $f: V \to W$, respecto de las bases $[B_1] \subset V \ y \ [B_2] \subset W$ y si $X_{[B_1]}$ denota la matriz columna correspondiente al vector $\overrightarrow{x} \in V$, cuyos elementos son las coordenadas de este respecto de la base $[B_1]$, entonces la imagen de $\overrightarrow{x}$, expresada en términos de la base $[B_2]$, se obtiene premultiplicando por A al vector columna $X_{[B_1]}$, es decir:

$$f(\overrightarrow{x}) = A\, X_{[B_1]} = Y_{[B_2]}$$

En efecto, dado el vector $\overrightarrow{x} \in V$, es:

$\overrightarrow{x} = \sum_{j=1}^{n} \alpha_j \overrightarrow{v_j}$ (1) por lo tanto $f(\overrightarrow{x}) = \overrightarrow{y} = \sum_{i=1}^{m} \alpha'_i \overrightarrow{w_i}$ (2). En consecuencia, teniendo presente la igualdad (1) y la definición de transformación lineal, es

$f(\overrightarrow{x}) = f\left(\sum_{j=1}^{n} \alpha_j \overrightarrow{v_j}\right) = \sum_{j=1}^{n} \alpha_j f(\overrightarrow{v_j})$, introduciendo la expresión (3), se obtiene:

$f(\overrightarrow{x}) = \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{m} a_{ij} \overrightarrow{w_i}$. Podemos colocar α_j en la segunda sumatoria pues es constante respecto del índice i y también permutar las sumatorias, obteniendo

$f(\overrightarrow{x}) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}\ \alpha_j\right) \overrightarrow{w_i}$ (4)

De las igualdades (2) y (4), por la unicidad de la combinación lineal que expresa a $f(\overrightarrow{x})$ en términos de la base $[B_2]$, es $\alpha'_i = \sum_{j=1}^{n} a_{ij}\ \alpha_j$.

Esto significa que la i–esima fila o componente de $Y_{[B_2]}$, o sea α'_i es igual a la suma de los productos de los elementos de la fila i de A por los correspondientes elementos de la única columna de $X_{[B_1]}$ que, por la definición de producto de matrices, resulta

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_i \\ \vdots \\ \alpha'_m \end{pmatrix}$$ o lo que es lo mismo

$$A\ X_{[B_1]} = Y_{[B_2]}$$

Con lo que podemos reiterar lo ya enunciado, la imagen de cualquier vector del espacio dominio es igual al producto de la matriz de la transformación lineal por la matriz de coordenadas de dicho vector,

respecto de la base del espacio dominio. Esta imagen resulta expresada en términos de la base del espacio codominio.

Continuando con el último ejemplo:

b) Obtener la imagen del vector $\overrightarrow{M} = \begin{pmatrix} 2 & -3 \\ 1 & -1 \end{pmatrix}$ utilizando la matriz asociada a f.

Como se utilizaron las bases canónicas en ambos espacios, las coordenadas del vector $\overrightarrow{M}$ se pueden expresar de la siguiente manera:

$M_{[B_1]} = \begin{pmatrix} 2 \\ -3 \\ 1 \\ -1 \end{pmatrix}$ y su imagen por f, utilizando A, es:

$f(\overrightarrow{M}) = A\, M_{[B_1]}$

$f(\overrightarrow{M}) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 1 \\ -1 \end{pmatrix}$

$f(\overrightarrow{M}) = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$

Esta imagen está expresada en la base canónica de $\mathbb{R}^2$ que coincide con la que se obtiene al aplicar directamente la transformación lineal f.

$f \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = (x_1 + x_4, x_2 - x_3)$

$f \begin{pmatrix} 2 & -3 \\ 1 & -1 \end{pmatrix} = (2 + (-1),\ (-3) - 1) = (1, -4)$

5. Diagonalización

En el apartado anterior hemos establecido la relación biunívoca entre transformaciones lineales y matrices, es decir que a toda transformación lineal se le asocia una matriz y viceversa. Estas matrices cambian según

las bases que se elijan para representarla. Ahora bien, hay transformaciones lineales que trabajan sobre un mismo espacio vectorial, es decir son de la forma $f: V \to V$ y son conocidos como endomorfismos. El objetivo de este apartado es tratar de obtener una matriz que represente a un endomorfismo dado, respecto de una cierta base, que sea lo más sencilla posible. Esta representación nos proporcionará información interesante sobre el funcionamiento de una aplicación lineal.

Para entender mejor el objetivo propuesto, veamos el siguiente ejemplo.

Sea la transformación lineal $f: \mathbb{R}^2 \to \mathbb{R}^2 \,/\, f(x_1, x_2) = (6x_1 - 2x_2; 6x_1 - x_2)$, calcular la matriz asociada a f respecto de la base canónica.

Notemos que f es un endomorfismo y, al ser el mismo espacio vectorial de dominio y codominio, nos referimos a una sola base.

$f(1,0) = (6,6) = 6\,(1,0) + 6\,(0,1)$

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

Por lo tanto $A = \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}$

Ahora calculemos la matriz asociada a f pero respecto de la base $[B] = \{\,(1,2), (2,3)\,\}$

$f(1,2) = (2,4) = 2\,(1,2) + 0\,(2,3)$

$f(2,3) = (6,9) = 0\,(1,2) + 3\,(2,3)$

Por lo tanto $A' = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

Podemos observar que en esta transformación tenemos una base respecto de la cual la transformación lineal dada tiene asociada una matriz diagonal, estas matrices cuentan, entre sus ventajas, las de ser fáciles de

operar (multiplicar, invertir, etc.), pero también proporcionan una mejor comprensión de lo que hace la transformación lineal.

Específicamente, si atendemos a la transformación lineal respecto de la base $[B]$ y recordando matriz de una transformación respecto a una base dada, observamos que el vector $(2,4)$ se puede escribir como $(2,4) = 2\,(1,2)$ y $(6,9) = 3\,(2,3)$, lo cual significa que la imagen de cada uno de los vectores de la base elegida corresponde a un múltiplo de ellos mismos. Esto se puede apreciar mejor desde el punto de vista geométrico. En la Figura 8.9. vemos una muestra de esta idea. Supongamos que los vectores de la base, llamémoslos genéricamente $\overrightarrow{v_1}$ y $\overrightarrow{v_2}$, son los vectores de los ejes, entonces la transformación "estira" los vectores en la dirección $\overrightarrow{v_1}$ hasta el doble de su tamaño y los de la dirección $\overrightarrow{v_2}$, al triple de su longitud. De esta manera un círculo con centro en el origen de coordenadas se transformará en una elipse de semiejes 2 y 3.

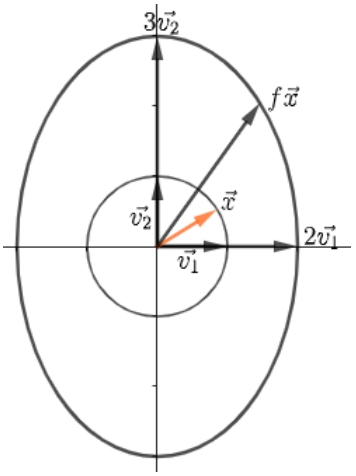


Figura 8.9. Interpretación geométrica de f.

En la Figura 8.10. podemos observar cómo se transforma un vector cualquiera: descomponemos el vector en sus componentes horizontal y vertical y la imagen tendrá su componente horizontal multiplicada por 2 y la vertical por 3.

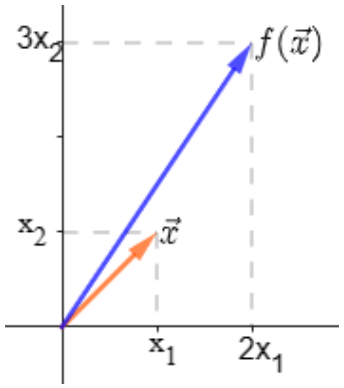


Figura 8.10. Descomposición de un vector genérico.

Obsérvese que el hecho de cada uno de los vectores de la base se transforme en un múltiplo suyo (esencialmente el mismo vector) es lo que nos ha permitido describir geométricamente la transformación. Estos son los vectores "especiales" que trataremos ahora.

Autovalores y autovectores

Los autovalores y autovectores juegan un papel muy importante a la hora de entender el comportamiento de los endomorfismos, lo que se pondrá de manifiesto al estudiar el tema: diagonalización. Por ahora sigamos prestando atención a aquellas transformaciones que precisamente convierten un vector $\overrightarrow{x}$ en otro paralelo a él, es decir en un vector $\lambda\,\overrightarrow{x}$, con

λ escalar. Es evidente que el vector nulo es solución a este problema puesto que es paralelo a cualquier otro vector, pero no es este el caso de interés. Por este motivo trabajaremos con aquellas transformaciones que respondan a la forma $f: V \to V$ con $dim(V) = n$ (finito) de tal modo que $f(\overrightarrow{x}) = \lambda \overrightarrow{x}\ con\ \overrightarrow{x} \neq \overrightarrow{0}$ (1). De esta forma, el problema que vamos a resolver consiste en determinar el conjunto de vectores diferentes del vector nulo que cumplan con la condición (1) para algún escalar λ y está acotado para los espacios vectoriales de dimensión finita. De esta manera, partimos del teorema que nos afirma que existe una única matriz A, en este caso de orden $n \times n$, que se asocia a la transformación lineal f. Haciendo intervenir a la matriz A, el problema se facilita y lo podemos expresar como un sistema de ecuaciones, donde se tiene que encontrar el escalar λ, con su correspondiente vector $\overrightarrow{x}$que cumplan,

$f(\overrightarrow{x}) = \lambda \overrightarrow{x}$ como $A\ X_{[B]} = Y_{[\mathrm{B}]}$ pues es la misa base, entonces $A\overrightarrow{x} = f(\overrightarrow{x})$. Es decir

$A\overrightarrow{x} = \lambda \overrightarrow{x}$ (2)

$A\overrightarrow{x} - \lambda \overrightarrow{x} = \overrightarrow{0}$

$(A - \lambda I)\overrightarrow{x} = \overrightarrow{0}$ (3), sistema de ecuaciones lineales homogéneo que tiene solución múltiple cuando

$| A - \lambda I | = 0$ (4)

Este planteamiento nos permite dar las siguientes definiciones:

Definición: sea $f: V \to V$ una transformación lineal. Un número real λ es un autovalor de la transformación lineal si para algún vector $\overrightarrow{x} \in V$ no

nulo, se verifica que $f(\overrightarrow{x}) = \lambda \overrightarrow{x}$. El vector $\overrightarrow{x}$ se denomina autovector de la transformación lineal relativo al autovalor λ.

Se debe descartar al vector nulo como posible autovector porque, por propiedades de las transformaciones lineales, $f(\overrightarrow{0_V}) = \overrightarrow{0_V}$, como V es espacio vectorial, tiene la propiedad de que $\overrightarrow{0_V} = \lambda \overrightarrow{0_V}$ $\forall \lambda \in \mathbb{R}$, es decir $f(\overrightarrow{0_V}) = \lambda \overrightarrow{0_V}$ con lo que cualquier número real sería un autovalor de la transformación lineal.

Al escalar λ, que es la constante entre vectores paralelos, aparte de autovalor, se le conoce también como valor propio, eigen valor y valor característico de la ecuación (4). No es necesario, pero nosotros trabajaremos solo con $\lambda \in \mathbb{R}$.

De manera similar, el vector que corresponde a un autovalor se llama autovector o bien vector propio o eigen vector y vector característico de la ecuación (4).

Definición: el polinomio $P(\lambda) = |A - \lambda I|$ de grado n se denomina polinomio característico de A. Equivalentemente: sea A una matriz cuadrada de orden n. El polinomio característico de A es el polinomio de grado n que resulta al desarrollar el determinante de la matriz $(A - \lambda I)$.

Definición: la ecuación polinomial $P(\lambda) = |A - \lambda I| = 0$ con respecto a λ de grado n, se llama ecuación característica de A.

A partir de estas definiciones veamos un resultado teórico que las relaciona mediante el siguiente:

Teorema. Sea $f: V \to V$ con $dim(V) = n$ (finito) y A la matriz asociada a f de tal manera que $f(\overrightarrow{x}) = \lambda \overrightarrow{x}$, entonces,

i) $P(\lambda)$ es un polinomio de grado n.

ii) Existe al menos un valor propio de A.

iii) Existe al menos un vector propio de A diferente del vector nulo.

D)

i) Se definió $P(\lambda) = |A - \lambda I|$. Por otro lado, por la definición de determinante establecemos que $P(\lambda)$ contiene a todos los productos de fila y columna sin repetir en un mismo término. Entonces, el término que tiene más factores con λ es el de la diagonal principal:

$(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda) = (-1)^n \lambda^n + \dots$

Después de desarrollar todo el determinante, los otros términos tendrán: $n-1, n-2, \dots 0$ factores con λ. Por lo tanto, no hay forma de que se elimine $(-1)^n \lambda^n$, de donde se concluye que $P(\lambda)$ es un polinomio de grado n.

ii) En este punto recordaremos una consecuencia del teorema fundamental del álgebra estudiado en el capítulo de polinomios "*Un polinomio $P \in \mathbb{C}[x]$ de grado n tiene exactamente n raíces complejas si las contamos de acuerdo con su multiplicidad*", entonces para $P(\lambda) = 0$ existe al menos un valor de λ raíz de la ecuación.

iii) Esta afirmación resulta por definición de vector propio ya que al sustituir el valor propio λ, tenemos un sistema de ecuaciones lineales homogéneo $(A - \lambda I)\overrightarrow{x} = \overrightarrow{0}$ con determinante de la matriz del sistema

$|A - \lambda I| = 0$, es decir $\rho(A - \lambda I) = k < n$ y sabemos que, si en un sistema de ecuaciones lineales el rango de la matriz del sistema es igual al rango de la matriz ampliada, menor que el número de incógnitas, el sistema es compatible indeterminado, es decir, tiene múltiples soluciones.

Luego de estos resultados, estamos en condiciones de calcular los autovalores y autovectores de la matriz asociada a una transformación lineal. Para esto recordemos que si A es la matriz asociada a una transformación, $A\vec{x} = \lambda\vec{x}$, entonces λ y $\vec{x}$ se pueden calcular, resumidamente, de la siguiente manera:

P_1) Calcular $|A - \lambda I|$ para encontrar el polinomio característico $P(\lambda) = |A - \lambda I|$.

P_2) Resolver la ecuación característica $P(\lambda) = 0$, encontrando los n autovalores.

P_3) Sustituir cada autovalor λ en el sistema homogéneo $(A - \lambda I)\vec{x} = \vec{0}$ y resolverlo llevando la matriz ampliada a su forma escalonada o escalonada reducida $(A - \lambda I \mid \vec{0})$. Las soluciones son los autovectores asociados al λ asignado.

Ejemplo: sea $A = \begin{pmatrix} 5 & -3 \\ -1 & 3 \end{pmatrix}$ la matriz asociada a la transformación f. Calcular los autovalores y autovectores de A.

Para calcular $|A - \lambda I|$ necesitamos $A - \lambda I$, entonces,

$$A - \lambda I = \begin{pmatrix} 5 & -3 \\ -1 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5-\lambda & -3 \\ -1 & 3-\lambda \end{pmatrix}$$

$$P(\lambda) = |A - \lambda I| = \begin{vmatrix} 5-\lambda & -3 \\ -1 & 3-\lambda \end{vmatrix} = (5-\lambda)(3-\lambda) - 3 = \lambda^2 - 8\lambda + 12$$

Resolvemos la ecuación $P(\lambda) = 0$

$\lambda^2 - 8\lambda + 12 = (\lambda - 6)(\lambda - 2) = 0$ por lo tanto $\lambda_1 = 6$ y $\lambda_2 = 2$

Encontramos los autovectores para cada valor de λ.

Para $\lambda_1 = 6$

$$(A - 6\,I \mid \vec{0}) = \begin{pmatrix} -1 & -3 & \vdots & 0 \\ -1 & -3 & \vdots & 0 \end{pmatrix} \sim^{-F_1} \begin{pmatrix} 1 & 3 & \vdots & 0 \\ -1 & -3 & \vdots & 0 \end{pmatrix} \sim_{F_1 + F_2} \begin{pmatrix} 1 & 3 & \vdots & 0 \\ 0 & 0 & \vdots & 0 \end{pmatrix}$$

Sistema compatible indeterminado, reconstruimos el sistema,

$x + 3y = 0$, elegimos a y como variable libre, entonces $x = -3y$ con $y \in \mathbb{R}$.

Un autovector para $\lambda_1 = 6$ puede ser

$(x, y) = (-3y, y) = y(-3, 1)$ es decir el autovector $(-3, 1)$ está asociado al autovalor $\lambda_1 = 6$.

Antes de continuar, observemos que cuando se busca un autovector hay que tomar en cuenta $|\,A - \lambda\,I\,| = 0$ lo que significa que al menos un renglón (ecuación) se puede representar como una combinación lineal de las otras, en otras palabras, se puede eliminar. Cuando son dos ecuaciones se elimina cualquiera, cuando son tres o más, se debe visualizar una dependencia en las relaciones y después de identificarla se puede eliminar. O sea, siempre se puede eliminar un renglón para simplificar los cálculos. Para $\lambda_2 = 2$, se tiene

$(A - 2I \mid \vec{0}) = \begin{pmatrix} 3 & -3 & 0 \\ -1 & 1 & 0 \end{pmatrix}$ considerando la observación anterior podemos ver la dependencia entre las dos filas de la matriz ampliada: $F_1 = -3F_2$, decidimos eliminar la primera fila (ecuación) y trabajar con la restante, es decir,

$-x + y = 0$ de donde $y = x$ con $x \in \mathbb{R}$.

Un autovector para $\lambda_2 = 2$ puede ser

$(x, y) = (x, x) = x(1, 1)$ es decir, el autovector $(1, 1)$ está asociado al autovalor $\lambda_2 = 2$.

Finalmente, los autovalores de A son $\lambda_1 = 6$ y $\lambda_2 = 2$ y un par de vectores asociado a cada autovalor, respectivamente, son $\{(-3, 1), (1, 1)\}$.

A partir de las definiciones y el ejemplo brindado, mencionamos tres observaciones. La primera es que de la definición de autovalor se deduce que dos autovalores distintos no pueden tener asociado el mismo autovector o equivalentemente: "*un mismo autovector no puede estar asociado a autovalores diferentes*".

D) Proponemos que el autovector $\vec{x}$ esté asociado a dos autovalores distintos, λ_1 y λ_2, por lo tanto,

$f(\vec{x}) = \lambda_1 \vec{x}$ y $f(\vec{x}) = \lambda_2 \vec{x}$ de donde

$\lambda_1 \vec{x} = \lambda_2 \vec{x}$

$(\lambda_1 - \lambda_2) \vec{x} = \vec{0}$

Por definición de autovalor $\vec{x} \neq \vec{0}$ entonces,

$\lambda_1 - \lambda_2 = 0 \implies \lambda_1 = \lambda_2$, quedando demostrada la afirmación.

La segunda observación tiene que ver con el ejemplo realizado, la transformación lineal tiene dos autovalores y para cada uno de estos, dimos un ejemplo de autovector para cada uno, en realidad a un autovalor están asociados un conjunto de autovectores. Esta afirmación también es una consecuencia de lo demostrado anteriormente. Esto se presenta en el siguiente:

Teorema. Si λ es un autovalor de la transformación lineal $f: V \to V$, entonces el conjunto $L(\lambda) = \{\overrightarrow{x} \in V \,/ f(\overrightarrow{x}) = \lambda \overrightarrow{x}\}$ es un subespacio vectorial de V denominado subespacio asociado al autovalor λ.

D)

i) Sea $\overrightarrow{x} \in L(\lambda) \implies f(\overrightarrow{x}) = \lambda \overrightarrow{x}$

$\overrightarrow{y} \in L(\lambda) \implies f(\overrightarrow{y}) = \lambda \overrightarrow{y}$

$\overrightarrow{x} + \overrightarrow{y} \implies f(\overrightarrow{x} + \overrightarrow{y}) = f(\overrightarrow{x}) + f(\overrightarrow{y})$

$f(\overrightarrow{x} + \overrightarrow{y}) = \lambda \overrightarrow{x} + \lambda \overrightarrow{y}$

$f(\overrightarrow{x} + \overrightarrow{y}) = \lambda(\overrightarrow{x} + \overrightarrow{y})$

$\overrightarrow{x} + \overrightarrow{y} \in L(\lambda)$

ii) Sea $\overrightarrow{x} \in L(\lambda) \implies f(\overrightarrow{x}) = \lambda \overrightarrow{x}$ y $\alpha \in \mathbb{R}$

$f(\alpha \overrightarrow{x}) = \alpha f(\overrightarrow{x})$

$f(\alpha \overrightarrow{x}) = \alpha \lambda \overrightarrow{x}$

$f(\alpha \overrightarrow{x}) = \lambda (\alpha \overrightarrow{x})$

$\alpha \overrightarrow{x} \in L(\lambda)$

Por lo tanto, $L(\lambda)$ es subespacio vectorial de V.

Ejemplo (continuación). Sabemos que la transformación lineal f, cuya matriz asociada de orden 2, es $A = \begin{pmatrix} 5 & -3 \\ -1 & 3 \end{pmatrix}$, tiene dos autovalores: $\lambda_1 = 6$ y $\lambda_2 = 2$. Calcular los subespacios asociados a cada autovalor.

Para $\lambda_1 = 6$

$$(A - 6\,I \mid \overrightarrow{0}) = \left(\begin{array}{cc|c} -1 & -3 & 0 \\ -1 & -3 & 0 \end{array}\right) \sim^{-F_1} \left(\begin{array}{cc|c} 1 & 3 & 0 \\ -1 & -3 & 0 \end{array}\right) \sim_{F_1 + F_2} \left(\begin{array}{cc|c} 1 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Sistema compatible indeterminado, reconstruimos el sistema $x + 3y = 0$ elegimos a y como variable libre, entonces $x = -3y$ con $y \in \mathbb{R}$. Por lo tanto,

$L(6) = \{(x, y) \in \mathbb{R}^2 \,/\, x = -3y\} = \{(-3y, y)\}$, esto significa que al subespacio vectorial asociado al autovalor $\lambda_1 = 6$, pertenecen todos aquellos vectores en los que la primera componente es el triplo de la segunda, cambiada de signo.

Para $\lambda_2 = 2$, se tiene

$(A - 2I \mid \vec{0}) = \left(\begin{array}{cc|c} 3 & -3 & 0 \\ -1 & 1 & 0 \end{array}\right)$, considerando la proporcionalidad de sus filas, decidimos eliminar la primera fila (ecuación) y trabajar con la restante, es decir,

$-x + y = 0$ de donde $y = x$ con $x \in \mathbb{R}$.

$L(2) = \{(x, y) \in \mathbb{R}^2 \,/\, y = x\} = \{(x, x)\}$, esto significa que al subespacio vectorial asociado al autovalor $\lambda_2 = 2$, pertenecen todos aquellos vectores de componentes iguales.

La tercera observación tiene que ver con la definición de autovector, en la cual se excluye el vector nulo como posible autovector, pero no se hace lo mismo con el autovalor cero, es decir que puede ser, $\lambda = 0$. Cuando esto ocurre, el subespacio vectorial asociado al autovalor cero es igual al núcleo de la transformación lineal, cuya demostración se da mediante el siguiente:

Teorema. Si $\lambda = 0$ es un autovalor de la transformación lineal $f: V \to V$, entonces el subespacio asociado a $\lambda = 0$ es igual al núcleo de la transformación lineal.

D) $L(\lambda) = \{\vec{x} \in V \,/ f(\vec{x}) = \lambda \vec{x}\}$

$L(0) = \{\vec{x} \in V \,/ f(\vec{x}) = 0 \vec{x}\}$

$L(0) = \{\overrightarrow{x} \in V \,/ f(\overrightarrow{x}) = \overrightarrow{0}\}$

$L(0) = N_f$

Antes de estudiar el tema diagonalización de matrices, veamos un par de importantes resultados que nos serán de gran utilidad.

Teniendo en cuenta que dos autovalores diferentes λ_1 y λ_2 no tienen autovectores comunes, la intersección de los respectivos subespacios asociados se reduce al vector nulo. Asimismo se deduce que dos autovectores cualesquiera correspondientes a λ_1 y λ_2, respectivamente, son linealmente independiente.

Se puede demostrar que lo enunciado anteriormente también se cumple si se considera cualquier conjunto finito de autovectores asociados a autovalores diferentes. Esto se realiza mediante el siguiente:

Teorema. Sea $f: V \to V$ una transformación lineal y $\{\overrightarrow{x_1}, \overrightarrow{x_2}, \dots, \overrightarrow{x_n}\}$ n autovectores correspondientes a los autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$. Si los autovalores son diferentes entre sí, es decir $i \neq j \Longrightarrow \lambda_{\mathrm{i}} \neq \lambda_{\mathrm{j}}$ entonces los n autovectores son linealmente independientes.

La demostración se hará por inducción sobre el número n de autovectores.

PB) $n = 1$, entonces $\overrightarrow{x_1} \neq \overrightarrow{0}$ por lo tanto $\overrightarrow{x_1}$ es linealmente independiente.

PI) $\{\overrightarrow{x_1}, \overrightarrow{x_2}, \dots, \overrightarrow{x_h}\}$ es linealmente independiente, entonces $\{\overrightarrow{x_1}, \overrightarrow{x_2}, \dots, \overrightarrow{x_h}, \overrightarrow{x_{h+1}}\}$ es linealmente independiente.

D) Consideremos una combinación lineal cualquiera de los autovectores igual al vector nulo.

$\sum_{i=1}^{h+1} \alpha_i \overrightarrow{x_i} = \overrightarrow{0}$ (1)

Multiplicamos a (1) por λ_{h+1}

$\alpha_1 \lambda_{h+1}\overrightarrow{x_1} + \alpha_2 \lambda_{h+1}\overrightarrow{x_2} + \ldots + \alpha_h \lambda_{h+1}\overrightarrow{x_h} + \alpha_{h+1} \lambda_{h+1}\overrightarrow{x_{h+1}} = \overrightarrow{0}$

Aplicando f a (1)

$\alpha_1 \lambda_1\overrightarrow{x_1} + \alpha_2 \lambda_2\overrightarrow{x_2} + \ldots + \alpha_h \lambda_h\overrightarrow{x_h} + \alpha_{h+1} \lambda_{h+1}\overrightarrow{x_{h+1}} = \overrightarrow{0}$

Restando miembro a miembro las dos últimas igualdades

$\alpha_1(\lambda_{h+1} - \lambda_1)\overrightarrow{x_1} + \alpha_2(\lambda_{h+1} - \lambda_2)\overrightarrow{x_2} + \cdots + \alpha_h(\lambda_{h+1} - \lambda_h)\overrightarrow{x_h} = \overrightarrow{0}$

Por PB) resulta

$\alpha_i(\lambda_{h+1} - \lambda_i) = 0 \quad \forall i = 1, 2, \ldots, h$

Y como $\lambda_{h+1} - \lambda_i \neq 0$ resulta

$\alpha_i = 0 \quad \forall i = 1, 2, \ldots, h$

Sustituyendo en (1), se tiene

$\alpha_{h+1}\overrightarrow{x_{h+1}} = \overrightarrow{0}$ o sea $\alpha_{h+1} = 0$ ya que $\overrightarrow{x_{h+1}} \neq \overrightarrow{0}$. En consecuencia

$\alpha_1 = \alpha_2 = \ldots = \alpha_h = \alpha_{h+1} = 0$ y, por lo tanto,

$\{\overrightarrow{x_1}, \overrightarrow{x_2}, \ldots, \overrightarrow{x_h}, \overrightarrow{x_{h+1}}\}$ es linealmente independiente.

El segundo resultado está referido a los autovalores de una matriz triangular y está dado por el siguiente:

Teorema. Los autovalores de una matriz triangular son las componentes diagonales de dicha matriz.

D) Si $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}$ entonces

$$A - \lambda I = \begin{pmatrix} a_{11-\lambda} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22-\lambda} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn-\lambda} \end{pmatrix}$$

Y como el determinante de una matriz triangular es igual al producto de los elementos de la diagonal principal, se ve que $Det(A-\lambda I) = (a_{11}-\lambda)(a_{22}-\lambda)\ldots\ (a_{nn}-\lambda)$ de donde se evidencian los autovalores de A. La demostración para una matriz triangular inferior es muy similar.

Ejemplo: calcular los autovalores de $A = \begin{pmatrix} 2 & 5 & 6 \\ 0 & -3 & 2 \\ 0 & 0 & 5 \end{pmatrix}$.

$Det(A-\lambda I) = \begin{vmatrix} 2-\lambda & 5 & 6 \\ 0 & -3-\lambda & 2 \\ 0 & 0 & 5-\lambda \end{vmatrix} = (2-\lambda)(-3-\lambda)(5-\lambda)$ por lo tanto

$(2-\lambda)(-3-\lambda)(5-\lambda) = 0$, consecuentemente los autovalores son $\lambda_1 = 2, \lambda_2 = -3\ y\ \lambda_3 = 5$.

5.1. Matrices semejantes y diagonalización

En este apartado definiremos una importante relación que se puede cumplir entre dos matrices, la semejanza. La importancia de esta relación radica en el hecho de que las matrices semejantes tienen varias propiedades interesantes comunes y además la mayoría de las matrices son semejantes a las matrices diagonales. Entonces, iniciemos el apartado con la siguiente:

Definición: se dice que dos matrices $A\ y\ B$ de orden n son semejantes si existe una matriz C de orden n tal que $B = C^{-1}AC$ (1).

Nótese que si a (1) la premutiplicamos por C, tenemos

$CB = C\ C^{-1}AC$ es decir

$CB = AC$ (2), esta expresión suele, en muchos casos, tomarse como una definición alternativa de semejanza de matrices pues tiene la ventaja de

que no es necesario calcular C^{-1} sino simplemente saber que es invertible, es decir que su determinante sea no nulo. Esto significa que "*dos matrices A y B son semejantes si y solo si existe una matriz invertible C tal que $CB = AC$*".

Ejemplo 1: sean las matrices $A = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix}$ y $C = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ podemos comprobar que

$Det(C) = 1 \neq 0$ por lo tanto, C es invertible.

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

$CB = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} = B = \begin{pmatrix} 3 & -1 \\ 1 & -1 \end{pmatrix}$, esto muestra por (2) que A y B son semejantes.

Ejemplo 2: sean las matrices $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, A = \begin{pmatrix} -6 & -3 & -25 \\ 2 & 1 & 8 \\ 2 & 2 & 7 \end{pmatrix}$ y $C = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix}$ podemos comprobar que

$Det(C) = 3 \neq 0$ por lo tanto, C es invertible.

$DC = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & -1 & 1 \\ 6 & 10 & 14 \end{pmatrix}$

$CA = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix}\begin{pmatrix} -6 & -3 & -25 \\ 2 & 1 & 8 \\ 2 & 2 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & -1 & 1 \\ 6 & 10 & 14 \end{pmatrix}$, esto muestra por (2) que D y A son semejantes.

Como se dijo anteriormente, las matrices semejantes tienen propiedades comunes, una de estas propiedades se enuncia en el siguiente:

Teorema. Si las matrices $A\ y\ B$ de orden n son semejantes, entonces $A\ y\ B$ tienen el mismo polinomio característico y por lo tanto tienen los mismos autovalores.

D) Como $A\ y\ B$ son semejantes, entonces $B = C^{-1}AC$ y

$Det(B - \lambda\, I) = Det(C^{-1}AC - \lambda\, I)$

$Det(B - \lambda\, I) = Det(C^{-1}AC - C^{-1}(\lambda\, I)C)$

$Det(B - \lambda\, I) = Det(C^{-1}(A - \lambda\, I)\, C)$

$Det(B - \lambda\, I) = Det(C^{-1}\,)\ Det\ (A - \lambda\, I)\ Det(C)$

$Det(B - \lambda\, I) = Det(C^{-1}\,)Det(C)\ Det\ (A - \lambda\, I)$

$Det(B - \lambda\, I) = Det(C^{-1}C\,)\ Det\ (A - \lambda\, I)$

$Det(B - \lambda\, I) = Det(I)\ Det\ (A - \lambda\, I)$

$Det(B - \lambda\, I) = Det\ (A - \lambda\, I)$

Esto significa que $A\ y\ B$ tienen la misma ecuación característica y como los autovalores son raíces de la ecuación característica, tiene los mismos autovalores.

En el ejemplo 2, anterior, vimos que $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, A = \begin{pmatrix} -6 & -3 & -25 \\ 2 & 1 & 8 \\ 2 & 2 & 7 \end{pmatrix}$ son semejantes.

Por el último teorema del apartado anterior, sabemos que los tres autovalores de la matriz D son $\lambda_1 = 1, \lambda_2 = -1\ y\ \lambda_3 = 2$, por lo tanto, los autovalores de $A = \begin{pmatrix} -6 & -3 & -25 \\ 2 & 1 & 8 \\ 2 & 2 & 7 \end{pmatrix}$ son también $\lambda_1 = 1, \lambda_2 = -1\ y\ \lambda_3 =$ 2. Esta afirmación se puede verificar, calculando los autovalaores de A – lo que se sugiere al estudiante– o bien viendo que se cumple

$$Det(A-I) = Det(A+I) = Det(A-2I) = 0$$

$$\begin{vmatrix} -7 & -3 & -25 \\ 2 & 0 & 8 \\ 2 & 2 & 6 \end{vmatrix} = \begin{vmatrix} -5 & -3 & -25 \\ 2 & 2 & 8 \\ 2 & 2 & 8 \end{vmatrix} = \begin{vmatrix} -8 & -3 & -25 \\ 2 & -1 & 8 \\ 2 & 2 & 5 \end{vmatrix} = 0$$

En muchas aplicaciones es de gran utilidad diagonalizar una matriz A, esto significa encontrar una matriz diagonal semejante a A. Para poder hacer esto, veamos en primer lugar la siguiente:

Definición: la mariz $A \in \mathbb{K}^{n \times n}$ es diagonalizable si existe una matriz diagonal D tal que A es semejante a D.

Observemos que si D es una matriz diagonal, entonces sus autovalores son los elementos de la diagonal principal. Si A es semejante a D, entonces $A\ y\ D$ tienen los mismos autovalores. Si unimos estos dos hechos se observa que si A es diagonalizable, entonces A es semejante a una matriz diagonal cuyos elementos en la diagonal principal son los autovalaores de A.

Entonces, vamos a establecer cuando una matriz es diagonalizable, mediante el siguiente:

Teorema. La mariz $A \in \mathbb{K}^{n \times n}$ es diagonalizable si y solo si tiene n autovectores linealmente independientes. En tal caso, la matriz diagonal D semejante a A está dada de la siguiente forma $D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$

donde $\lambda_1, \lambda_2, \dots, \lambda_n$ son los autovalores de A. Si C es una matriz cuyas columnas son los autovectores linealmente independientes de A, entonces $D = C^{-1}AC$.

D) $\Longrightarrow$) Si A tiene n autovectores linealmente independientes, entoces es diagonalizable.

Primero se supone que A tiene n autovectores linealmente independientes $\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}$ que corresponden a los autovalores (no necesariamente diferentes) $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Sean $\overrightarrow{v_1} = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix}$, $\overrightarrow{v_2} = \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{n2} \end{pmatrix}, \ldots, \overrightarrow{v_n} = \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{nn} \end{pmatrix}$ y

$$C = \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

Construida C de esta manera, es invertible pues sus columnas son linealmente independiente. Ahora bien

$$AC = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$ de donde se sabe que la columna

i de AC es $A\begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{pmatrix} = A\overrightarrow{v_i} = \lambda_i \overrightarrow{v_i}$[46]. Así AC es la matriz cuya columna i es $\lambda_i \overrightarrow{v_i}$ y

$$AC = \begin{pmatrix} \lambda_1 c_{11} & \lambda_2 c_{21} & \cdots & \lambda_n c_{1n} \\ \lambda_1 c_{21} & \lambda_2 c_{22} & \cdots & \lambda_n c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \lambda_1 c_{n1} & \lambda_2 c_{n2} & \cdots & \lambda_n c_{nn} \end{pmatrix}$$

[46] Recordar la igualdad (2) al iniciar el estudio de autovalores y autovectores.

Pero

$$CD = \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} =$$

$$\begin{pmatrix} \lambda_1 c_{11} & \lambda_2 c_{21} & \ldots & \lambda_n c_{1n} \\ \lambda_1 c_{21} & \lambda_2 c_{22} & \ldots & \lambda_n c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \lambda_1 c_{n1} & \lambda_2 c_{n2} & \cdots & \lambda_n c_{nn} \end{pmatrix}$$

Luego $AC = CD$ y como C es invertible se puede premultiplicar ambos miembros por C^{-1} para obtener

$C^{-1}AC = C^{-1}CD$

$D = C^{-1}AC$ (1)

⟸) Si A es diagonlizable, entonces A tiene n autovectores linealmente independientes.

Como A es diagonalizable, se cumple $D = C^{-1}AC$ para alguna matriz C. Sean $\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}$ las columnas de C. Entonces $AC = CD$, e invirtiendo los argumentos anteriores, se ve de inmediato que $A\overrightarrow{v_i} = \lambda_i \overrightarrow{v_i}$ para $i = 1, 2, \ldots, n$. Entonces $\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}$ son los autovectores de A y son linealmente independientes porque C es invertible.

Este teorema tiene un resultado muy útil que se deduce directamente del teorema, que se demostró en el apartado anterior, referido a la independencia lineal de autovectores relativos a autovalores diferentes. Esto se enuncia en el siguiente:

Corolario. Si la matriz $A \in \mathbb{K}^{n\,x\,n}$ tiene n autovalores diferentes, entonces A es diagonalizable.

Finalmente al tratar de determinar si una matriz es diagonalizable o no, es necesario tener en cuenta la siguiente:

Observación: si λ es una raíz del polinomio característico de una matriz A y su multiplicidad es m, la dimensión del subespacio vectorial asociado a ese autovalor puede ser igual o menor que m. Es decir $dim(L(\lambda)) \leq m$.

Por lo tanto resulta útil hallar en primer lugar los subespacios de los autovectores asociados a los autovalores que son raíces múltiples del polinomio característico, para determinar su dimensión. Si para alguno de ellos se verifica que $di\,m(L(\lambda)) < m$, entonces la matriz no es diagonalizable.

Ejemplo 1: sea $A = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$.

a) Calcular los autovalores y los autovectores (subespacios asociados)

$Det(A - \lambda I) = \begin{vmatrix} 4-\lambda & 2 \\ 3 & 3-\lambda \end{vmatrix} = (4-\lambda)(3-\lambda) - 6 = \lambda^2 - 7\lambda + 6$

$\lambda^2 - 7\lambda + 6 = 0 \Leftrightarrow (\lambda - 1)(\lambda - 6) = 0$. Por lo tanto, los autovalores son $\lambda_1 = 1 \;\; y \;\; \lambda_2 = 6$

Observemos que $m(\lambda_1) = 1$ y $m(\lambda_2) = 1$

Para $\lambda_1 = 1$

$(A - \; I \mid \vec{0}) = \left(\begin{array}{cc|c} 3 & 2 & 0 \\ 3 & 2 & 0 \end{array}\right)$ eliminamos una fila, reconstruimos el sistema $3x + 2y = 0$, elegimos a y como variable libre, entonces $x = -\frac{2}{3}y$ con $y \in \mathbb{R}$. Por lo tanto

$L(1) = \left\{(x, y) \in \mathbb{R}^2 \,/\, x = -\frac{2}{3}y \;\; con\; y \in \mathbb{R}\right\} = \left\{\left(-\frac{2}{3}y, y\right)\right\}$

Construimos una base para $L(1)$

$\left(-\frac{2}{3}y, y\right) = y(-\frac{2}{3}, 1)$ entonces $\left\{(-\frac{2}{3}, 1)\right\}$ es sistema de generadores y linealmente independiente. $Dim\big(L(1)\big) = 1$.

Para $\lambda_2 = 6$

$(A - 6I \mid \overrightarrow{0}) = \left(\begin{matrix} -2 & 2 \\ 3 & -3 \end{matrix} \vdots \begin{matrix} 0 \\ 0 \end{matrix}\right)$ eliminamos la segunda fila, reconstruimos el sistema y obtenemos $-2x + 2y = 0$, elegimos a y como variable libre, entonces $x = y$ con $y \in \mathbb{R}$. Por lo tanto,

$L(6) = \{(x, y) \in \mathbb{R}^2 \,/\, x = y \; con \; y \in \mathbb{R}\} = \{(x, x)\}$

Construimos una base para $L(6)$

$(x, x) = x(1, 1)$ entonces $\{(1, 1)\}$ es sistema de generadores y linealmente independiente. $Dim\big(L(6)\big) = 1$.

b) Decir si la transformación lineal es diagonalizable o no. Justificar.

Sí es diagonalizable porque $m(\lambda_1) = dim\big(L(1)\big) = 1 \; y \; m(\lambda_2) = dim(L(6)) = 1$ y además porque $A \in \mathbb{R}^{2\times 2}$ tiene dos autovalores distintos.

c) Hallar una base respecto de la cual A sea diagonalizable.

Tomamos un vector de $L(1)$, por ejemplo $(2, -3)$ que tiene un manejo algebraico más fácil. Hacemos lo mismo para $L(6)$, por ejemplo el ya propuesto u obtenido $(1, 1)$. Es decir $\{(2, -3), (1, 1)\}$.

d) Encontrar la matriz de diagonalización C.

Hacemos $C = \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}$ encontramos que

$C^{-1}AC = \frac{1}{5}\begin{pmatrix}1 & -1\\ 3 & 2\end{pmatrix}\begin{pmatrix}4 & 2\\ 3 & 3\end{pmatrix}\begin{pmatrix}2 & 1\\ -3 & 1\end{pmatrix} = \frac{1}{5}\begin{pmatrix}1 & -1\\ 3 & 2\end{pmatrix}\begin{pmatrix}2 & 6\\ -3 & 6\end{pmatrix} = \frac{1}{5}\begin{pmatrix}5 & 0\\ 0 & 30\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 6\end{pmatrix}$ que es la matriz cuyas componentes en la diagonal principal son los autovalores de A.

Obsérvese que como existe un número infinito de maneras en las cuales se puede elegir un autovector, existe un número infinito de formas para seleccionar una matriz de diagonalización C. Se aconseja elegir los autovectores y la matriz C que sean más sencillo en el manejo aritmético, es decir, en términos generales debe insertarse el mayor número de ceros y unos posibles.

Ejemplo 2: sea $A = \begin{pmatrix}1 & -1 & 4\\ 3 & 2 & -1\\ 2 & 1 & -1\end{pmatrix}$

a) Calcular los autovalores y los autovectores (subespacios asociados)

$Det(A - \lambda I) = \begin{vmatrix}1-\lambda & -1 & 4\\ 3 & 2-\lambda & -1\\ 2 & 1 & -1-\lambda\end{vmatrix} = -(\lambda^3 - 2\lambda^2 - 5\lambda + 6)$

$-(\lambda^3 - 2\lambda^2 - 5\lambda + 6) = 0 \Leftrightarrow -(\lambda - 1)(\lambda + 2)(\lambda - 3) = 0$ Por lo tanto los autovalores y sus respectivos orden de multiplicidad son $\lambda_1 = 1$ con $m(\lambda_1) = 1$, $\lambda_2 = -2$ con $m(\lambda_2) = 1$ y por último $\lambda_3 = 3$ con $m(\lambda_3) = 1$,

Para $\lambda_1 = 1$

$(A - I \mid \vec{0}) =$

$$\left(\begin{array}{ccc|c}0 & -1 & 4 & 0\\ 3 & 1 & -1 & 0\\ 2 & 1 & -2 & 0\end{array}\right) \sim \left(\begin{array}{ccc|c}2 & 1 & -2 & 0\\ 3 & 1 & -1 & 0\\ 0 & -1 & 4 & 0\end{array}\right) \sim^{\frac{1}{2}F_1} \left(\begin{array}{ccc|c}1 & \frac{1}{2} & -1 & 0\\ 3 & 1 & -1 & 0\\ 0 & -1 & 4 & 0\end{array}\right)$$

$$\sim -3F_1 + F_2 \begin{pmatrix} 1 & \frac{1}{2} & -1 & 0 \\ 0 & -\frac{1}{2} & 2 & 0 \\ 0 & -1 & 4 & 0 \end{pmatrix}$$

Eliminamos la segunda fila, pues es proporcional con la tercera. Reconstruimos el sistema

$$\begin{cases} x + \frac{1}{2}y - z = 0 \\ -y + 4z = 0 \end{cases} \Longrightarrow \begin{cases} x + \frac{1}{2}y - z = 0 \\ y = 4z \end{cases} \Longrightarrow \begin{cases} x = -z \\ y = 4z \end{cases}$$

$L(1) = \{(x, y, z) \in \mathbb{R}^3 \,/\, x = -z, y = 4z \;\; con\; z \in \mathbb{R}\} = \{(-z, 4z, z)\}$

Construios una base para $L(1)$

$(-z, 4z, z) = z(-1, 4, 1)$ entonces $\{(-1, 4, 1)\}$ es sistema de generadores y linealmente independiente. $Dim\big(L(1)\big) = 1$.

Para $\lambda_2 = -2$

$(A + 2I \mid \overrightarrow{0}) =$

$$\begin{pmatrix} 3 & -1 & 4 & 0 \\ 3 & 4 & -1 & 0 \\ 2 & 1 & 1 & 0 \end{pmatrix} \sim \begin{matrix} \\ -F_1 + F_2 \\ -\frac{2}{3}F_1 + F_3 \end{matrix} \begin{pmatrix} 3 & -1 & 4 & 0 \\ 0 & 5 & -5 & 0 \\ 0 & \frac{5}{3} & -\frac{5}{3} & 0 \end{pmatrix}$$ Eliminamos la tercera fila pues es proporcional con la segunda. Reconstruimos el sistema

$$\begin{cases} 3x - y + 4z = 0 \\ 5y - 5z = 0 \end{cases} \Longrightarrow \begin{cases} 3x - y + 4z = 0 \\ y = z \end{cases} \Longrightarrow \begin{cases} x = -z \\ y = z \end{cases}$$

$L(-2) = \{(x, y, z) \in \mathbb{R}^3 \,/\, x = -z, y = z \;\; con\; z \in \mathbb{R}\} = \{(-z, z, z)\}$

Construimos una base para $L(-2)$

$(-z, z, z) = z(-1, 1, 1)$ entonces $\{(1, -1, -1)\}$ es sistema de generadores y linealmente independiente. $Dim\big(L(-2)\big) = 1$.

Para $\lambda_3 = 3$

$(A - 3I \mid \overrightarrow{0}) =$

$\begin{pmatrix} -2 & -1 & 4 & 0 \\ 3 & -1 & -1 & 0 \\ 2 & 1 & -4 & 0 \end{pmatrix} \sim_{\substack{\frac{3}{2}F_1 + F_2 \\ F_1 + F_3}} \begin{pmatrix} -2 & -1 & 4 & 0 \\ 0 & -\frac{5}{2} & 5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Reconstruimos el sistema

$$\begin{cases} -2x - y + 4z = 0 \\ -\frac{5}{2}y + 5z = 0 \end{cases} \Rightarrow \begin{cases} -2x - y + 4z = 0 \\ y = 2z \end{cases} \Rightarrow \begin{cases} x = z \\ y = 2z \end{cases}$$

$L(3) = \{(x, y, z) \in \mathbb{R}^3 \,/\, x = z, y = z \;\; con\, z \in \mathbb{R}\} = \{(z, 2z, z)\}$

Construimos una base para $L(3)$

$(z, 2z, z) = z(1, 2, 1)$ entonces $\{(1, 2, 1)\}$ es sistema de generadores y linealmente independiente. $Dim\big(L(3)\big) = 1$.

b) Decir si la transformación lineal es diagonalizable o no. Justificar.

Es diagonalizable porque $m(\lambda_1) = di\, m\big(L(1)\big) = 1$, $m(\lambda_2) = di\, m\big(L(-2)\big) = 1$ y también y $m(\lambda_3) = di\, m\big(L(3)\big) = 1$ además porque $A \in \mathbb{R}^{3\,x\,3}$ tiene tres autovalores distintos.

c) Hallar una base respecto de la cual A sea diagonalizable.

$\{(-1, 4, 1), (1, -1, -1), (1, 2, 1)\}$

d) Encontrar la matriz de diagonalización C.

Hacemos $C = \begin{pmatrix} -1 & 1 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{pmatrix}$ encontramos que

$$C^{-1}AC = -\frac{1}{6}\begin{pmatrix} 1 & -2 & 3 \\ -2 & -2 & 6 \\ -3 & 0 & -3 \end{pmatrix}\begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{pmatrix}$$

$$C^{-1}AC = -\frac{1}{6}\begin{pmatrix} 1 & -2 & 3 \\ -2 & -2 & 6 \\ -3 & 0 & -3 \end{pmatrix}\begin{pmatrix} -1 & -2 & 3 \\ 4 & 2 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

$$C^{-1}AC = -\frac{1}{6}\begin{pmatrix} -6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & -18 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Ejemplo 3: sea $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$

a) Calcular los autovalores y los autovectores (subespacios asociados)

$$Det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 + 15\lambda + 8$$

$-\lambda^3 + 6\lambda^2 + 15\lambda + 8 = 0 \Leftrightarrow -(\lambda + 1)^2(\lambda - 8) = 0$. Por lo tanto, los autovalores y sus respectivos órdenes de multiplicidad son $\lambda_1 = -1$ con $m(\lambda_1) = 2$, $\lambda_2 = 8$ con $m(\lambda_2) = 1$.

Para $\lambda_1 = -1$

$(A + I \mid \overrightarrow{0}) =$

$\left(\begin{array}{ccc|c} 4 & 2 & 4 & 0 \\ 2 & 1 & 2 & 0 \\ 4 & 2 & 4 & 0 \end{array}\right) \sim \begin{matrix} \\ -\frac{1}{2}F_1 + F_2 \\ -F_1 + F_3 \end{matrix} \left(\begin{array}{ccc|c} 4 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$. Reconstruimos el sistema y dividimos en dos la única ecuación, obtenemos:

$2x + y + 2z = 0$ entonces $y = -2x - 2z$

$L(-1) = \{(x, y, z) \in \mathbb{R}^3 \,/\, y = -2x - 2z \;\; con \; x, z \in \mathbb{R}\} = \{(x, -2x - 2z, z)\}$

Construimos una base para $L(-1)$

$(x, -2x - 2z, z) = (x, -2x, 0) + (0, -2z, z)$ entonces

$\{(1, -2, 0), (0, -2, 1)\}$ es sistema de generadores y se sugiere al estudiante verificar que el conjunto es linealmente independiente. $Dim\big(L(-1)\big) = 2$.

Para $\lambda_1 = 8$

$(A - 8\, I \mid \overrightarrow{0}) =$

$\left(\begin{array}{ccc|c} -5 & 2 & 4 & 0 \\ 2 & -8 & 2 & 0 \\ 4 & 2 & -5 & 0 \end{array}\right) \sim \begin{matrix} \\ 4F_1 + F_2 \\ -F_1 + F_3 \end{matrix} \left(\begin{array}{ccc|c} -5 & 2 & 4 & 0 \\ -18 & 0 & 18 & 0 \\ 9 & 0 & -9 & 0 \end{array}\right)$. Eliminamos la segunda fila por ser proporcional con la tercera. Reconstruimos el sistema:

$$\begin{cases} -5x + 2y + 4z = 0 \\ 9x - 9z = 0 \end{cases} \Rightarrow \begin{cases} -5x + 2y + 4z = 0 \\ x = z \end{cases} \Rightarrow \begin{cases} y = \frac{1}{2}z \\ x = z \end{cases}$$

$$L(8) = \left\{(x, y, z) \in \mathbb{R}^3 \,/\, y = \tfrac{1}{2}z, x = z \;\; con\, z \in \mathbb{R}\right\} = \left\{\left(z, \tfrac{1}{2}z, z\right)\right\}$$

Construimos una base para $L(8)$

$\left(z, \frac{1}{2}z, z\right) = z(1, \frac{1}{2}, 1)$ entonces $\{(2, 1, 2)\}$ es sistema de generadores y linealmente independiente. $Dim\big(L(8)\big) = 1$.

b) Decir si la transformación lineal es diagonalizable o no. Justificar.

Es diagonalizable porque $m(\lambda_1) = di\, m\big(L(-1)\big) = 2,\; m(\lambda_2) = di\, m\big(L(8)\big) = 1$

c) Hallar una base respecto de la cual A sea diagonalizable

Tomamos dos vectores de $L(-1)$ y un vector de $L(8)$.

$\{(1, -2, 0), (0, -2, 1), (2, 1, 2)\}$

d) Encontrar la matriz de diagonalización C.

Hacemos $C = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -2 & -2 \\ 2 & 0 & 1 \end{pmatrix}$ encontramos que

$$C^{-1}AC = -\frac{1}{9}\begin{pmatrix} -2 & -1 & -2 \\ -5 & 2 & 4 \\ 4 & 2 & -5 \end{pmatrix}\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 & 0 \\ 1 & -2 & -2 \\ 2 & 0 & 1 \end{pmatrix}$$

$$C^{-1}AC = -\frac{1}{9}\begin{pmatrix} -2 & -1 & -2 \\ -5 & 2 & 4 \\ 4 & 2 & -5 \end{pmatrix}\begin{pmatrix} 16 & -1 & 0 \\ 8 & 2 & 2 \\ 16 & 0 & -1 \end{pmatrix}$$

$$C^{-1}AC = -\frac{1}{9}\begin{pmatrix} -72 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Obsévese que este ejemplo evidencia que la matriz A es diagonalizable aún cuando no tenga los autovalores diferentes.

Ejemplo 4: sea la matriz $A = \begin{pmatrix} 4 & -1 \\ 0 & 4 \end{pmatrix}$.

a) Calcular los autovalores y los autovectores (subespacios asociados).

$Det(A - \lambda I) = \begin{vmatrix} 4-\lambda & -1 \\ 0 & 4-\lambda \end{vmatrix} = (4-\lambda)^2$

$(4-\lambda)^2 = 0$. Por lo tanto, el único autovalor es $\lambda_1 = 4$

Observemos que $m(\lambda_1) = 2$

Para $\lambda_1 = 4$

$(A - 4\,I \mid \overrightarrow{0}) = \left(\begin{array}{cc|c} 0 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right)$. Reconstruimos el sistema $y = 0$. Por lo tanto,

$L(4) = \{(x, y) \in \mathbb{R}^2 \,/\, y = 0 \;\; con\; x \in \mathbb{R}\} = \{(x, 0)\}$

Construimos una base para $L(4)$

$(x, 0) = x(1, 0)$ entonces $\{(1, 0)\}$ es sistema de generadores y linealmente independiente. $Dim\big(L(4)\big) = 1$.

b) Decir si la transformación lineal es diagonalizable o no. Justificar.

No es diagonalizable porque $di\, m(L(\lambda)) < m$.

Otro razonamiento sería:

Suponer que A fuera diagonalizable, entonces $D = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$ y existiría una matriz invertible C tal que $C^{-1}AC = D$, ahora bien

$CC^{-1}AC = CD$

$AC = CD$

$ACC^{-1} = CDC^{-1}$

$A = CDC^{-1}$ como $D = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$

$A = C \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} C^{-1}$

$A = C(4\, I)C^{-1}$

$A = 4C\ IC^{-1}$

$A = 4C\ C^{-1}$

$A = 4I$

$A = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = D$. Pero $A \neq D$

Por lo tanto, no existe tal C.

BIBLIOGRAFÍA

Acevedo Cardona, M. A. (2013). *Álgebra para la Física*. Reporte de Actividad Docente. Universidad Nacional Autónoma de México. Facultad de Ciencias.

Anton, H. (2003). *Introducción al álgebra lineal*. (Tercera Edición). México D. F.: Editorial Limusa.

Aranda Ortega, E. (2013). *Álgebra lineal con aplicaciones y Python*. (Primera Edición). Impreso por Lulu.com. Disponible en descarga gratuita. http://matematicas.uclm.es/earanda/?page_id=152

Arce, C.; Castillo, W. y González, J. (2003). *Álgebra Lineal*. (Tercera Edición). Escuela de Matemática. Universidad de Costa Rica.

Argomedo Cornejo, S.; Herrera Tobar, J.; Molina Alfaro, K. y Relos Paco, S. (2014). *Álgebra Lineal para Ingeniería.* (Primera Edición). Iniciativa Latinoamericana de Libros de Texto Abiertos.

Barrera Mora, F. (2014). *Álgebra Lineal.* Primera Edición Ebook. México: Grupo Editorial Patria.

Buitrago Alemán, R. (2009). *Álgebra Lineal*. Facultad de Ciencias Básicas. Departamento de Matemáticas. Bogotá, D.C. Colombia: Editor: Universidad Militar Nueva Granada.

Cobos Gavala, F. J.; Osuna Lucena, A.; Robles Arias, R. y Silva Gallardo, B. (s/f). *Apuntes de Álgebra Lineal*. Departamento de Matemática Aplicada I. Universidad de Sevilla. España.

Costa, V.; Rossignoli, R.; Sorichetti, C. y Vampa, V. (2018). *Álgebra Lineal con Aplicaciones*. Parte I. Universidad Nacional de la Plata. Facultad de Ingeniería. Buenos Aires. Argentina: Libros de Cátedra. Edulp.

Departamento de Álgebra. Universidad de Sevilla (2009). *Álgebra Lineal y Geometría*. Tema 2. Disponible en http//www.departamento.us.es/da.

Eguía Ribero, M. I.; Alberdi Celaya, E. y Unzueta Inchaurbe, A. (2015). *Fundamentos del Álgebra lineal. Ejercicios y cuestiones. Soluciones con Mathematica.*

González Meneses López, J. (2009). *Álgebra Lineal*. Departamento de Álgebra. Universidad de Sevilla.

Grossman, S. I. (2008). *Álgebra lineal*. (Sexta Edición). Revisión y adaptación: José Job Flores Godoy. México D. F.: Mc. Graw Hill Educación.

Gutiérrez González, E. y Ochoa García, S. I. (2014). *Álgebra lineal y sus aplicaciones*. México. D. F.: Grupo Editorial Patria, S.A.

Howard, A. (2008). *Introducción al Álgebra Lineal*. (4.ta Edición). Buenos Aires Argentina: Editorial Limusa.

Isaacs, R. y Sabogal, S. (2009). *Aproximación al Álgebra Lineal. Un Enfoque Geométrico*. Bucaramanga. Colombia.

Jerónimo, G.; Sabia, J. y Tesauri, S. (2008). *Álgebra Lineal*. (Fascículo 2 Cursos de grado). Departamento de Matemática. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. Argentina.

Larson, R. y otros. (2008). *Cálculo y Geometría Analítica.* (6.ta Edición). Buenos Aires. Argentina: Mc Graw Hill Editorial.

Lay, D. C. (2012). *Álgebra lineal y sus aplicaciones*. (Cuarta edición). (Ana Elizabeth García Hernández [Trad.], Javier Alfaro Pastor [Rev. Técnica]). México: Pearson Educación.

López García, J. D. y Rodríguez Santamaría, J. C. (2010). *Álgebra Lineal*. Universidad de los Andes.

Martínez, H. J. y Sanabria, A. M. (2014). *Álgebra Lineal.* Primera Edición. Colección Ciencias Naturales y Exactas. Programa Editorial Universidad del Valle. Santiago de Cali. Colombia.

Poole, D. (2011). *Álgebra lineal. Una introducción moderna*. (Tercera Edición). México, D.F.: Cengage Learning Editores.

Saldarriaga, O. y Giraldo, H. (2010). *Álgebra Lineal con el uso de Matlab.* Instituto de Matemáticas. Universidad de Antioquia. Colombia.

Sotelo, J. C. (2011). *Álgebra Lineal para estudiantes de ingeniería y ciencias*. Mc. Graw Hill Educación. México D. F.

Torres Bugeau, C.; Lasserre, A. y García, A. (2009). *Elementos de Álgebra y Geometría Analítica*. (Vol. I). Jujuy: EDIUNJu.

Torres Bugeau, C.; Lasserre, A. y García, A. (2009). *Elementos de Álgebra y Geometría Analítica*. (Vol. II). Jujuy: EDIUNJu.

Wiley, J. (2001). *Introducción al Álgebra Lineal*. (Segunda Edición). México D. F.: Editorial Limusa.

SOBRE LOS AUTORES

Héctor Ramón Tarifa. Doctor en Ciencias Sociales por la Universidad Nacional de Jujuy (UNJu). Especialista en Investigación Educativa por la Universidad Nacional de Tucumán, Licenciado en Enseñanza de la Matemática, egresado de la Facultad de Ingeniería de la UNJu y profesor de Matemática egresado del Instituto Nacional de Enseñanza Superior "José E. Tello" de la provincia de Jujuy. Se desempeña como profesor adjunto ordinario de la Cátedra Álgebra y Geometría Analítica de carreras de ingenierías y profesor Adjunto Ordinario de las Cátedras Álgebra I y Álgebra II de la carrera Analista Programador Universitario de la Facultad de Ingeniería. Es investigador categoría III de la Secretaría de Políticas Universitarias, en el área Análisis de Redes Sociales orientado al estudio de sistemas complejos.

Patricio Omar Condori. Especialista en Docencia en Entornos Virtuales por la Universidad Nacional de Quilmes. Licenciado en Enseñanza de la Matemática e ingeniero en Informática, egresado de la Universidad Nacional de Jujuy. Profesor de Matemática, egresado del Instituto de Educación Superior "José Eugenio Tello" de San Salvador de Jujuy. Se desempeña como profesor adjunto en la cátedra de Álgebra y Geometría Analítica de la Facultad de Ingeniería de la UNJu y es investigador de la Secretaría de Políticas Universitarias.

José Luis Medina. Licenciado en Física, egresado de la Facultad de Ciencias Exactas y Tecnología de la Universidad Nacional de Tucumán. Se desempeña como profesor adjunto en la Cátedra de Física I y en la Cátedra de Álgebra y Geometría Analítica de la Facultad de Ingeniería de la UNJu. Es investigador categorizado en la Secretaría de Ciencia y Técnica y Estudios Regionales de la UNJu. Actualmente dirige un proyecto de investigación relacionado con el desarrollo y análisis de tecnologías sustentables basadas en Energías Renovables y, también, participa como investigador en un proyecto cuyo objetivo es analizar, mediante ciencias de datos, la retención y progresión académica de

estudiantes que cursan las carreras dictadas en la Facultad de Ingeniería de la UNJu.

Víctor Eduardo Mérida. Licenciado en Enseñanza de la Matemática, egresado de la Universidad Nacional de Jujuy. Profesor en Matemática y Cosmografía, egresado del Instituto de Educación Superior N° 5 "José Eugenio Tello" de la provincia de Jujuy. Es experto en Educación Universitaria por la Universidad Católica de Santiago del Estero. Se desempeña como profesor adjunto en las Cátedras de Introducción a la Matemática y Matemática I en la Facultad de Ciencias Económicas de la UNJu, desarrollando actividades docentes; también es ayudante de primera en la asignatura de Álgebra y Geometría Analítica en la Facultad de Ingeniería de la UNJu. Es investigador novel en proyectos de iniciación en investigaciones científicas relacionadas con la enseñanza y el aprendizaje de la matemática en la Facultad de Ingeniería y la Facultad de Ciencias Económicas de la UNJu.



Álgebra Lineal para Ingenierías. Álgebra y Geometría Analítica, de Héctor Ramón Tarifa, Patricio Omar Condorí, José Luis Medina, Víctor Eduardo Mérida se terminó de imprimir en la primera quincena del mes de noviembre de 2025, en los Talleres Gráficos de la Imprenta de la UNJu.

Jujuy - Argentina.
Tirada: 100 ejemplares.

Editorial de la Universidad Nacional de Jujuy
ediunju

El libro ***Álgebra Lineal para Ingenierías*** es una herramienta fundamental para estudiantes y profesionales, diseñado para fortalecer las bases teóricas y aplicaciones prácticas en diversas ramas de la ingeniería. Consta de ocho capítulos que profundizan conceptos del nivel medio hasta temas avanzados, todos ellos estructurados cuidadosamente para facilitar el aprendizaje progresivo.

El texto inicia con los conceptos básicos de vectores en Rn para luego profundizar en números complejos y polinomios. Seguidamente se lleva a cabo un estudio detallado de tres temas íntimamente vinculados: matrices, determinantes y sistemas de ecuaciones lineales; para finalizar con espacios vectoriales y transformaciones lineales.

El enfoque del escrito combina demostraciones teóricas con ejemplos prácticos y ejercicios desafiantes, promoviendo no solo la comprensión conceptual, sino también el pensamiento crítico y la resolución de problemas. El objetivo del libro no es solo consolidar conceptos matemáticos, sino también destacar la relevancia del álgebra lineal como herramienta esencial en la innovación y desarrollo tecnológico. Es una guía indispensable para quienes buscan aplicar las matemáticas en soluciones reales de ingeniería.