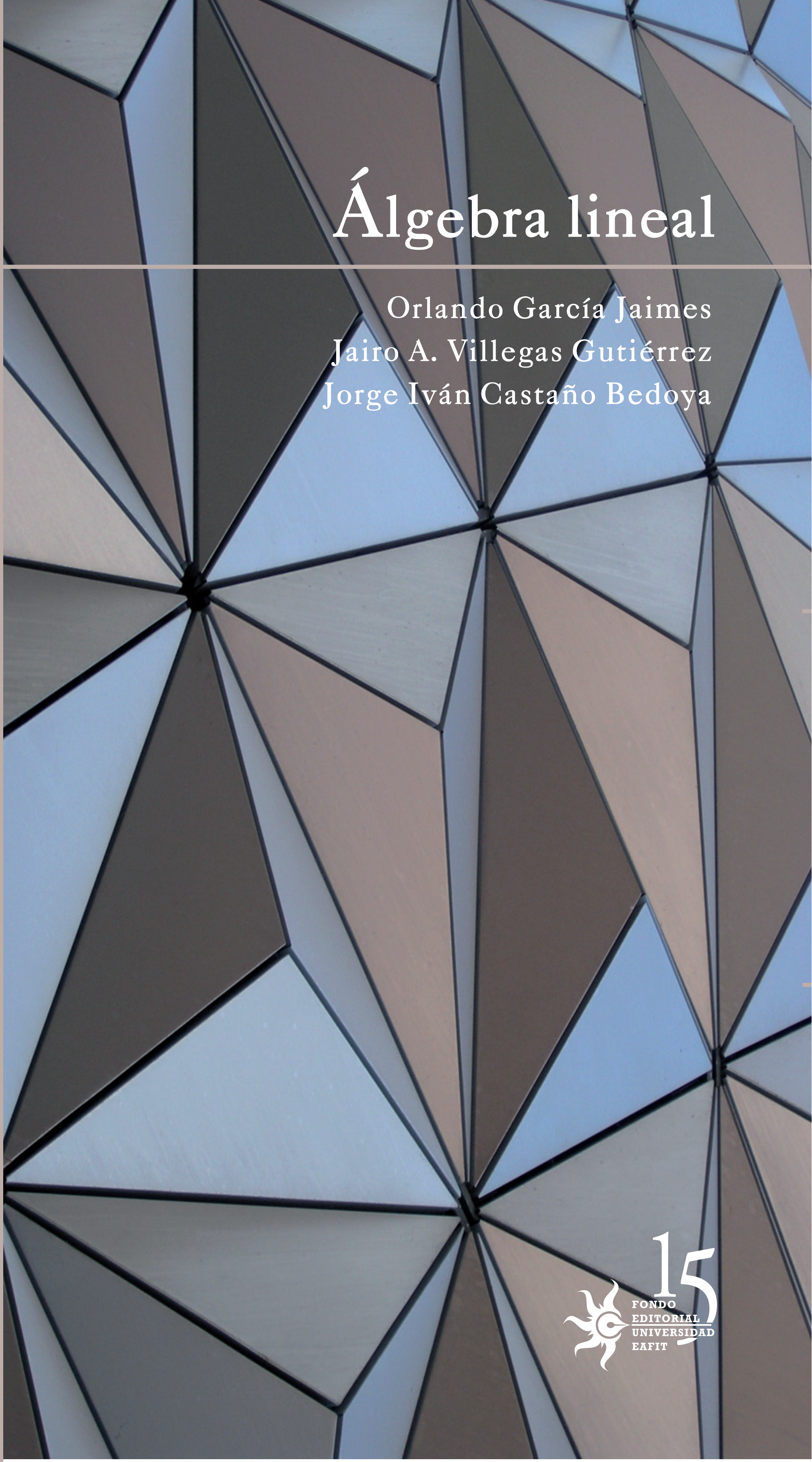
C
A
COLECCIÓN ACADÉMICA
EAFIT



Álgebra lineal

Orlando García Jaimes
Jairo A. Villegas Gutiérrez
Jorge Iván Castaño Bedoya

ESCUELA DE CIENCIAS Y HUMANIDADES

15
FONDO EDITORIAL UNIVERSIDAD EAFIT

Álgebra lineal

Orlando García Jaimes
Jairo A. Villegas Gutiérrez
Jorge Iván Castaño Bedoya

15
FONDO
EDITORIAL
UNIVERSIDAD
EAFIT

García Jaimes, Orlando
Álgebra lineal / Orlando García Jaimes, Jairo Alberto Villegas Gutiérrez, Jorge Iván Castaño Bedoya. -- Medellín : Fondo Editorial Universidad EAFIT, 2012.
520 p. ; 24 cm. -- (Colección académica)

1.Algebras lineales. 2. Matrices (Matemáticas). 3. Espacios vectoriales.
I. Villegas Gutiérrez, Jairo Alberto. II. Castaño Bedoya, Jorge Iván. III. Tít. IV. Serie
512.5 cd 21 ed.
G216
Universidad EAFIT-Centro Cultural Biblioteca Luis Echavarría Villegas

Álgebra lineal

Primera edición: noviembre de 2012
© Orlando García Jaimes
© Jairo A. Villegas Gutiérrez
© Jorge Iván Castaño Bedoya
© Fondo Editorial Universidad EAFIT
Cra. 48A No. 10 sur - 107. Tel. 261 95 23
www.eafit.edu.co/fondoeditorial
Correo electrónico: fonedit@eafit.edu.co

Imagen de carátula: *Epcot Polygons,* subido por: turbidity,
disponible en: www.sxc.hu, consulta: 7 de noviembre del 2012

Editado en Medellín, Colombia

Contenido

Índice de figuras

Índice de tablas

Capítulo 1

Matrices y sistemas de ecuaciones lineales

1.1 Introducción

Desde el punto de vista lógico, un curso de Álgebra Lineal debería iniciarse con el concepto de *espacio vectorial*, presentar luego las transformaciones lineales entre espacios vectoriales y la cuestión de las matrices, y a partir de esto atacar el problema de los sistemas de ecuaciones. Sin embargo, desde el punto de vista pedagógico, resulta más productivo introducir al estudiante en los conceptos del Álgebra Lineal comenzando con los temas de matrices y sistemas de ecuaciones lineales, con los cuales él ya está un poco más familiarizado.

Para resolver sistemas de ecuaciones lineales, las matrices constituyen una poderosa herramienta para la formulación de algoritmos de solución. Además, ellas nos permiten expresar de forma simple gran variedad de problemas relacionados con procesos de información, introducción de sistemas lineales en un computador, etc. En el campo del álgebra lineal y sus temas afines, las matrices constituyen una herramienta fundamental.

Históricamente, la referencia a problemas algebraicos aparece en épocas muy remotas. Ya en las civilizaciones de la antigua

Mesopotamia se muestran algunos problemas matemáticos usando escritura cuneiforme.

Mejor suerte se ha tenido con los antiguos documentos egipcios, puesto que el descubrimiento de la piedra trilingüe de Roseta (griego, demótico y jeroglífico) en 1779, en el transcurso de una de las expediciones napoleónicas, ha permitido descifrar gran parte de documentos e inscripciones egipcias. Uno de los más importantes, con carácter matemático, que se ha encontrado es el denominado papiro Rhind o papiro de Ahmes. Este es un rollo de papiro de 30 cm de ancho por 6 m de largo, que fue comprado en 1858 por el coleccionista de antigüedades escocés Henry Rhind. Ahmes, el escriba que transcribió originalmente el papiro alrededor del 1650 a. C., explica que el material es una copia de otro manuscrito que data de entre los años 2000 y 1800 a. C. y que es muy probable que sean conocimientos que provengan de Imhotep, arquitecto y médico del faraón Zoser.

En el papiro Rhind hallamos una gran variedad de problemas matemáticos, que van desde representaciones para los números hasta problemas geométricos y de razones trigonométricas. En el caso del álgebra, aparecen problemas en los cuales se plantea la solución de una ecuación lineal de la forma $x + ax = b$ o $x + ax + bx = c$. Es necesario aclarar que ellos no emplearon este tipo de escritura; para representar las incógnitas usaban la expresión "aha" o montón, y con esta terminología nos encontramos con problemas como este: "Calcular el montón si el montón y un séptimo del montón son 19".

Por otro lado, los chinos también se plantearon el problema de resolver sistemas lineales. En el texto *Nueve capítulos de arte matemático*, escrito durante la dinastía Han alrededor de los años 200 a. C., nos encontramos con el siguiente problema: "Hay tres tipos de cereal, de los cuales tres fardos del primero, dos del segundo y uno del tercero hacen 39 medidas. Dos del primero, tres del segundo y uno del tercero hacen 34 medidas. Y uno del primero, dos del segundo y tres del tercero hacen 26 medidas. ¿Cuántas medidas de cereal son contenidas en un fardo de cada tipo?".

Usando la notación actual para las ecuaciones, tenemos un sistema de tres ecuaciones con tres incógnitas: x, y y z, que puede escribirse en la forma

$$\begin{aligned} 3x + 2y + z &= 39 \\ 2x + 3y + z &= 34 \\ x + 2y + 3z &= 26. \end{aligned}$$

En cuanto a la solución, lo que se propone es escribir las ecuaciones en un tablero de la forma

1	2	3
2	3	2
3	1	1
26	34	39

Para resolver el sistema de ecuaciones, el autor de los *Nueve capítulos de arte matemático* explica el procedimiento en los siguientes términos:

1. Multiplique la columna dos por tres y la columna tres por dos; reste la nueva columna tres a la columna dos, generando una nueva columna dos. Luego, multiplique la columna uno por tres y quítele la columna tres, generando una nueva columna uno. De esta forma se obtiene el siguiente tablero:

0	2	3
4	5	2
8	1	1
39	24	39

2. Multiplique la columna uno por cinco y la columna dos por cuatro; reste la nueva columna dos a la columna uno, obteniendo una nueva columna uno. El tablero quedará finalmente de la siguiente forma:

0	0	3
0	5	2
36	1	1
99	24	39

En este tablero, la primera columna permite calcular el valor de z. Usando este valor y la segunda columna se puede calcular y. Por último, con estos valores y la ecuación correspondiente a la tercera columna, se puede determinar el valor de x. Este procedimiento constituye un antecesor chino del método de eliminación de Gauss.

Entre los métodos que se conocen para resolver sistemas de ecuaciones lineales, es curioso que lo que hoy conocemos como método de Cramer para un sistema de la forma

$$\begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned}$$

fuese presentado por primera vez en la forma

$$x = \frac{ce - bf}{ae - bd} \qquad y = \frac{af - dc}{ae - bd}$$

por Colin Maclaurin en 1729, en el texto *Treatise of Algebra,* que se pretendía fuese una introducción al libro de Isaac Newton *Arithmetica universalis.* El libro de Colin Maclaurin se publico en el año 1748, dos años antes de que viera la luz el texto de Gabriel Cramer sobre líneas algebraicas, en el que aparecen las mencionadas fórmulas (Boyer, 2001: 540-542).

1.2 Matrices

Definición 1.2.1. *Una matriz A es un arreglo rectangular de $m \times n$ elementos de $\mathbb{R}$ o $\mathbb{C}$, distribuidos en m filas y n columnas, y que denotamos con $A = (a_{ij})_{m\times n}$.*

$$A = (a_{ij})_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}.$$

En este caso decimos que A es una matriz de orden $m \times n$.

En el elemento a_{ij} de dicho arreglo, el subíndice i señala la fila a la que aquel elemento pertenece, mientras que el subíndice j señala la columna a la que pertenece el elemento. Por otro lado, al conjunto de todas las matrices de m filas y n columnas, con elementos en el conjunto de los números reales, lo denotamos con $M_{m\times n}(\mathbb{R})$. En el caso en el cual $m = n$, decimos que la matriz es *cuadrada*.

Ejemplo 1.2.1. Sean las matrices

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & -4 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & 5 & 2 \\ 2 & 2 & 1 \\ 7 & -4 & 0 \end{bmatrix}.$$

A es una matriz de orden 2×3, mientras que B es una matriz de orden 3×3. Además, $a_{23} = -4$ y $b_{13} = 2$.

1.2.1 Algunos tipos de matrices

En esta sección establecemos algunos tipos especiales de matrices que aparecen en un primer curso de Álgebra Lineal y que por su estructura y propiedades cumplen un papel importante en la teoría.

Matriz diagonal. Sea $A \in M_{n\times n}$; se dice que A es diagonal si y solo si $a_{ij} = 0$ cuando $i \neq j$.

Ejemplo 1.2.2. La matriz A, dada por

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

es una matriz diagonal.

Matriz identidad. Sea $A \in M_{n\times n}$; se dice que A es la matriz identidad si y solo si $a_{ij} = 0$ cuando $i \neq j$ y $a_{ij} = 1$ cuando $i = j$. La matriz identidad se suele representar mediante la expresión I_n.

Ejemplo 1.2.3. La matriz

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

es la matriz identidad de orden 3×3.

Matriz escalar. Sea $A \in M_{n\times n}$; se dice que A es una matriz escalar si para cada elemento a_{ij} se cumple que:

$$a_{ij} = \begin{cases} 0, & i \neq j; \\ k, & i = j. \end{cases}$$

Ejemplo 1.2.4. La matriz

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

es una matriz escalar. Toda matriz escalar es diagonal, pero no toda matriz diagonal es escalar.

Matriz nula. Sea $A \in M_{m\times n}$; se dice que A es la matriz nula si $a_{ij} = 0$ para todo $i = 1, \ldots m$ y para todo $j = 1, \ldots n$. Esta matriz se denota con 0.

Ejemplo 1.2.5. La matriz

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

es la matriz nula de orden 3×3.

Matriz triangular superior. Sea $A \in M_{n\times n}$; se dice que A es triangular superior si $a_{ij} = 0$, para todo $i > j$. Si A es un matriz de este tipo, los elementos bajo la diagonal son nulos.

Ejemplo 1.2.6. La matriz

$$A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 4 & 4 & 5 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 7 \end{bmatrix}$$

es una matriz triangular superior.

Matriz triangular inferior. Sea $A \in M_{n\times n}$; se dice que A es triangular inferior si $a_{ij} = 0$, para todo $i < j$. Si A es un matriz de este tipo, los elementos sobre la diagonal son nulos.

Ejemplo 1.2.7. La matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 4 & 0 & 5 & 0 \\ 7 & 6 & 2 & 7 \end{bmatrix}$$

es una matriz triangular inferior.

1.3 Operaciones con matrices

Esta sección la dedicamos a definir y estudiar distintos tipos de operaciones con matrices. Realizamos la suma de matrices del mismo orden, sumando los elementos que ocupan la misma posición en ambas matrices; además, el producto de una matriz por un escalar corresponde a la multiplicación de cada uno de los elementos de la matriz por dicho escalar.

Definición 1.3.1. *Dadas las matrices $A \in M_{m\times n}$ y $B \in M_{m\times n}$, su suma, que denotamos con $A+B$, es una matriz $M_{m\times n}$, cuyos elementos se obtienen de la siguiente forma:*

$$\begin{aligned}
A+B &= (a_{ij}) + (b_{ij}) \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix};
\end{aligned}$$

es decir, la suma de matrices se puede ver como una función

$$\begin{aligned}
+ \ : \ M_{m\times n} \times M_{m\times n} &\longrightarrow M_{m\times n} \\
(A,B) &\longrightarrow A+B.
\end{aligned}$$

Definición 1.3.2. *Llamamos producto de un escalar λ por una matriz $A = (a_{ij})_{m\times n}$ a la nueva matriz dada por*

$$\lambda A = \lambda \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{bmatrix};$$

es decir, el producto de una matriz por un escalar se puede ver como una función

$$\begin{aligned}
. \ : \ \mathbb{R} \times M_{m\times n} &\longrightarrow M_{m\times n} \\
(\lambda, A) &\longrightarrow \lambda A.
\end{aligned}$$

Ejemplo 1.3.1. Dadas las matrices

$$A = \begin{bmatrix} -1 & 4 & 1 & 0 \\ 2 & 0 & 1 & 2 \\ 0 & 0 & -1 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & -1 & 0 & 3 \\ 3 & -2 & 1 & 0 \\ -1 & -1 & -2 & 4 \end{bmatrix},$$

entonces:

$$A+B = \begin{bmatrix} 1 & 3 & 1 & 3 \\ 5 & -2 & 2 & 2 \\ -1 & -1 & -3 & 6 \end{bmatrix} \quad \text{y} \quad -2A = \begin{bmatrix} 2 & -8 & -2 & 0 \\ -4 & 0 & -2 & -4 \\ 0 & 0 & 2 & -4 \end{bmatrix}.$$

Sean las matrices A, B y C en $M_{m\times n.}$ La suma de matrices cumple las siguientes propiedades:

1. $A + B$ es una matriz en $M_{m\times n}$ (Cerradura).

2. $A + B = B + A$ (Conmutativa).

3. $A + (B + C) = (A + B) + C$ (Asociativa).

4. $A + 0 = A$ para toda matriz $A \in M_{m\times n}$ (Elemento neutro).

5. En $M_{m\times n}$ se define $-A = (-a_{ij})$. A esta matriz se le llama *matriz opuesta* de A y cumple que para cada $A \in M_{m\times n}$, $A + (-A) = 0$.

Sean las matrices A y B en $M_{m\times n}$ y α y β dos escalares; el producto de una matriz por un escalar cumple las siguientes propiedades:

1. αA es una matriz en $M_{m\times n}$ (Cerradura).

2. $\alpha(\beta A) = (\alpha\beta)A$ (Asociativa).

3. $\alpha(A + B) = \alpha A + \alpha B$ (Distributiva).

4. $(\alpha + \beta)A = \alpha A + \beta A$ (Distributiva).

5. $1A = A$ (Elemento neutro del producto por escalar).

Demostración. Vamos a probar el tercer numeral del primer grupo de propiedades. Sean las matrices $A = (a_{ij})$, $B = (b_{ij})$ y $C = (c_{ij})$ en $M_{m\times n}$; entonces:

$$\begin{aligned}
A + (B + C) &= (a_{ij}) + [(b_{ij}) + (c_{ij})] \\
&= (a_{ij}) + [b_{ij} + c_{ij}] \\
&= (a_{ij} + [b_{ij} + c_{ij}]) \\
&= ([a_{ij} + b_{ij}] + c_{ij}) \\
&= (a_{ij} + b_{ij}) + (c_{ij}) \\
&= [(a_{ij}) + (b_{ij})] + (c_{ij}) \\
&= (A + B) + C.
\end{aligned}$$

Para el tercer numeral del segundo grupo de propiedades tenemos que para las matrices $A = (a_{ij})$ y $B = (b_{ij})$ en $M_{m\times n}$ y el escalar $\alpha \in \mathbb{R}$:

$$\begin{aligned}
\alpha(A + B) &= \alpha[(a_{ij}) + (b_{ij})] \\
&= \alpha(a_{ij} + b_{ij}) \\
&= [\alpha(a_{ij} + b_{ij})] \\
&= [(\alpha a_{ij} + \alpha b_{ij})] \\
&= [(\alpha a_{ij}) + (\alpha b_{ij})] \\
&= (\alpha a_{ij}) + (\alpha b_{ij}) \\
&= \alpha(a_{ij}) + \alpha(b_{ij}) \\
&= \alpha A + \alpha B.
\end{aligned}$$

Las demás propiedades se demuestran de modo similar.

En general, un conjunto de objetos dotado de una operación binaria interna (suma) y de una operación binaria externa (producto por escalar), que cumplen las diez propiedades que hemos mencionado, constituye una estructura matemática conocida como *espacio vectorial.* Así, el conjunto de $M_{m\times n}$, dotado de las operaciones de suma y multiplicación de una matriz por un escalar, es un *espacio vectorial.* En el capítulo 3 presentamos otros ejemplos de espacios vectoriales y generalizamos algunas propiedades comunes a ellos.

1.4 Multiplicación de matrices

En la sección anterior se expusieron las operaciones de suma y producto por escalar para las matrices. Como se mencionó al final de la sección, uno de los aspectos importantes de estas operaciones consiste en la posibilidad de pensar en estructuras algebraicas generales; ahora presentamos la noción de *producto de matrices* y mostramos que aun cuando no tiene el poder de generalización que proporcionan los espacios vectoriales, permite agrupar, dentro de un mismo concepto, gran número de problemas y aplicaciones.

Como una manera de justificar la operación de multiplicación de matrices, consideremos el siguiente ejemplo.

Ejemplo 1.4.1. Una empresa requiere los servicios de tres empleados para ocupar las jefaturas de tres departamentos. A la convocatoria para dichos cargos se presentan cinco aspirantes. Para el proceso de selección se tiene en cuenta las calificaciones de los aspirantes en tres asignaturas: Cálculo, Álgebra y Administración. Además, cada departamento asigna a cada materia un determinado valor porcentual o peso.

Los pesos de las materias en cada departamento están dados en la Tabla 1.1.

	Depto. 1	**Depto. 2**	**Depto. 3**
Cálculo	0,6	0,4	0,3
Álgebra	0,3	0,3	0,4
Admón.	0,1	0,3	0,3

Tabla 1.1. Pesos de las materias por departamento

En la Tabla 1.2 aparecen las calificaciones de los aspirantes en cada materia.

	Cálculo	**Álgebra**	**Admón.**
Pedro	4,0	3,5	4,0
Juan	4,0	4,0	4,5
Luis	3,5	4,0	3,5
Roberto	4,0	3,5	3,5
María	3,5	4,0	4,0

Tabla 1.2. Calificaciones por aspirante

Determinar la calificación de cada aspirante por departamento. ¿Cuál es el aspirante mejor evaluado en cada departamento?

Solución

De acuerdo con la información suministrada en las tablas, podemos construir dos matrices A y B, cuyas filas y columnas representan los datos de cada tabla. Para establecer la calificación de un aspirante en un departamento se multiplican los datos de la fila de notas del aspirante por los datos de la columna de pesos del departamento y luego se suman estos productos. Así, por ejemplo, si queremos saber cuál es la calificación de Luis en el departamento 2, tomamos la tercera fila de la matriz A y la segunda columna de la matriz B y las combinamos en la forma:

$$\begin{bmatrix} 3,5 & 4,0 & 3,5 \end{bmatrix} \begin{bmatrix} 0,4 \\ 0,3 \\ 0,3 \end{bmatrix} = 3,5 \times 0,4 + 4,0 \times 0,3 + 3,5 \times 0,3 = 3,65.$$

De modo general, si queremos determinar la calificación del aspirante que ocupa la fila i de la matriz A en el departamento j, lo que debemos hacer es multiplicar la fila i de A por la columna j de B en la forma

$$\begin{bmatrix} a_{i1} & a_{i2} & a_{i3} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ b_{3j} \end{bmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j},$$

donde los $i = 1, 2, 3, 4, 5$ y $j = 1, 2, 3$, lo que significa que debemos tener como resultado una matriz de cinco filas y tres columnas. Todos los resultados los podemos escribir así:

$$\begin{bmatrix} 4,0 & 3,5 & 4,0 \\ 4,0 & 4,0 & 4,5 \\ 3,5 & 4,0 & 3,5 \\ 4,0 & 3,5 & 3,5 \\ 3,5 & 4,0 & 4,0 \end{bmatrix} \begin{bmatrix} 0,6 & 0,4 & 0,3 \\ 0,3 & 0,3 & 0,4 \\ 0,1 & 0,3 & 0,3 \end{bmatrix} = \begin{bmatrix} 3,85 & 3,85 & 3,80 \\ 4{,}05 & 4{,}15 & 4{,}15 \\ 3,65 & 3,65 & 3,70 \\ 3,80 & 3,70 & 3,65 \\ 3,70 & 3,80 & 3,85 \end{bmatrix}.$$

En la matriz del producto, el elemento c_{ij} representa la calificación del aspirante i en el departamento j. Según esto, Juan es el que obtiene la nota más alta en el primer departamento, mientras que la mejor calificación en los otros dos departamentos es también para Juan, como se indica en la matriz del producto.

Definición 1.4.1. *Dadas las matrices* $A \in M_{m\times n}$ *y* $B \in M_{n\times p}$, *entonces el producto de* A *con* B, *que denotamos con* AB, *es una matriz* $C \in M_{m\times p}$, *dada por*

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj},$$

con $i = 1, \ldots m$ *y* $j = 1 \ldots p$.

La definición anterior establece que el elemento que ocupa la posición i, j del producto AB es igual a la suma de los productos de las componentes que ocupan la misma posición en la fila i de A y la columna j de B.

Ejemplo 1.4.2. Sean las matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix};$$

entonces:

$$C = AB = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix},$$

donde los elementos c_{ij} están dados por las relaciones:

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ c_{12} &= a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \\ c_{22} &= a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32}. \end{aligned}$$

En el caso del producto BA la matriz resultante es de tamaño 3×3, lo que permite ver que no se cumple la ley conmutativa.

Ejemplo 1.4.3. Sean las matrices

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & 1 & 1 & 4 \\ -1 & 0 & 1 & 2 \end{bmatrix};$$

entonces:

$$\begin{aligned} AB &= \begin{bmatrix} (-1)2 + 0(-1) & (-1)1 + 0(0) & (-1)1 + 0(2) & (-1)4 + 2(0) \\ (0)2 + 2(-1) & (0)1 + 2(0) & (0)1 + 2(1) & (0)4 + 2(2) \end{bmatrix} \\ &= \begin{bmatrix} -2 & -1 & -1 & -4 \\ -2 & 0 & 2 & 4 \end{bmatrix}. \end{aligned}$$

Es importante recalcar que el producto de matrices no es conmutativo, como puede notarse en el ejemplo 1.4.3:

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

Sin embargo, el producto BA no está definido. Lo anterior nos conduce a afirmar que para que el producto BA esté definido se requiere que el número de columnas de la primera matriz coincida con el número de filas de la segunda matriz.

Como hemos dicho previamente, la suma de matrices posee las mismas propiedades que la suma de escalares. Para el producto de matrices las cosas son diferentes, es decir, algunas de las propiedades usuales del producto entre escalares no son válidas para la multiplicación de matrices. Veamos una lista de las propiedades del producto de matrices:

1. *Ley distributiva.* Si B y C son matrices de tamaño $m \times n$ y A es una matriz de tamaño $r \times m$, entonces $A(B+C) = AB+AC$.

2. *Ley distributiva.* Si B y C son matrices de tamaño $m \times n$ y A es una matriz de tamaño $n \times r$, entonces $(B+C)A = BA+CA$.

3. Si α es un escalar y A es una matriz de tamaño $m \times n$ y B es una matriz de tamaño $n \times r$, entonces $A(\alpha B) = (\alpha A)B = \alpha(AB)$.

4. Si A, B y C son matrices, entonces $A(BC) = (AB)C$, siempre y cuando el producto de estas matrices se pueda realizar.

Demostración. Vamos a efectuar la prueba de la primera propiedad, las demás se proponen como ejercicio para el lector. Notemos que $A(B+C)$ y $AB+AC$ son matrices de tamaño $r \times n$. Ahora, si llamamos $D = B + C$, $E = AD$, $F = AB$, $G = AC$ y $H = F + G$, entonces tenemos que demostrar que $E = H$.

Si $E = [e_{ij}]$ y $H = [h_{ij}]$, para probar que estas matrices son iguales, demostremos que $e_{ij} = h_{ij}$ para $i = 1, 2, \ldots r$ y $j = 1, 2, \ldots n$. En efecto,

$$e_{ij} = \sum_{k=1}^{m} a_{ik} d_{kj};$$

pero al ser $d_{kj} = b_{kj} + c_{kj}$, al reemplazar d_{kj} en la expresión e_{ij}, tenemos que

$$\begin{aligned} e_{ij} &= \sum_{k=1}^{m} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{m}(a_{ik}b_{kj} + a_{ik}c_{kj}) \\ &= \sum_{k=1}^{m} a_{ik}b_{kj} + \sum_{k=1}^{m} a_{ik}c_{kj}. \end{aligned}$$

Por otro lado, a partir de las definiciones de H, F y G obtenemos que $h_{ij} = f_{ij} + g_{ij}$, con lo cual

$$h_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj} + \sum_{k=1}^{m} a_{ik}b_{kj},$$

que es la misma expresión que hemos determinado para e_{ij}, mostrando entonces que $E = H$ y, por lo tanto, $A(B + C) = AB + AC$.

1.5 Transpuesta de una matriz

Definición 1.5.1. *Dada una matriz de tamaño $m \times n$, la transpuesta de A, que denotamos con A^t, es la matriz de tamaño $n \times m$, cuyas columnas se forman a partir de las filas de A. Esto es, si $A = [a_{ij}]$, entonces $A^t = [a_{ji}]$.*

Ejemplo 1.5.1. Para el caso de una matriz A de tamaño 3×3, dada por

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

tenemos que

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

Obsérvese que los elementos de la diagonal son invariantes.

Ejemplo 1.5.2. Dadas las matrices

$$A = \begin{bmatrix} -2 & 1 & 4 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 \\ -1 & 4 \\ 1 & 3 \end{bmatrix},$$

entonces:

$$A^t = \begin{bmatrix} -2 & 0 \\ 1 & 1 \\ 4 & 2 \end{bmatrix} \quad \text{y} \quad B^t = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 4 & 3 \end{bmatrix}.$$

Al combinar la transpuesta de una matriz con las operaciones que hemos definido para las matrices, obtenemos las siguientes propiedades:

1. $(A^t)^t = A$.
2. $(A + B)^t = A^t + B^t$.
3. Para cualquier escalar α, $(\alpha A)^t = \alpha A^t$.
4. $(AB)^t = B^t A^t$.

Demostración. Las demostraciones de las tres primeras propiedades es inmediata; solo consideraremos el cuarto numeral. Sea $A = [a_{ij}]$ una matriz de tamaño $m \times p$ y sea $B = [b_{ij}]$ una matriz de tamaño $p \times n$; si llamamos $C = AB$, entonces el elemento c_{ij} de la matriz C^t está dado por c^t_{ij}. Además,

$$\begin{aligned} c^t_{ij} = c_{ji} &= \text{(fila j de A).(columna i de B)} \\ &= (a_{j1}, a_{j2}, \ldots, a_{jp}).(b_{1i}, a_{2i}, \ldots, a_{pi}) \\ &= a_{j1}b_{1i} + a_{j2}b_{2i} + \ldots + a_{jp}b_{pi} \\ &= a^t_{1j}b^t_{i1} + a^t_{2j}b^t_{i2} + \ldots + a^t_{pj}b^t_{ip} \\ &= b^t_{i1}a^t_{1j} + b^t_{i2}a^t_{2j} + \ldots + b^t_{ip}a^t_{pj} \\ &= (\text{fila } i \text{ de } B^t\ ).(\text{columna } j \text{ de } A^t) \\ &= \text{elemento } (i, j) \text{ de } B^t A^t, \end{aligned}$$

con lo cual se concluye la prueba.

Ejemplo 1.5.3. Dadas las matrices

$$A = \begin{bmatrix} -1 & 4 & 0 \\ 2 & 1 & 0 \\ 1 & 3 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 1 & 4 \end{bmatrix},$$

entonces:

$$AB = \begin{bmatrix} -1 & 4 & 0 \\ 2 & 1 & 0 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} -7 & 6 \\ 5 & 6 \\ 2 & 16 \end{bmatrix},$$

a partir de lo cual concluimos que

$$(AB)^t = \begin{bmatrix} -7 & 5 & 2 \\ 6 & 6 & 16 \end{bmatrix}.$$

Si ahora calculamos por separado las transpuestas, obtenemos:

$$A^t = \begin{bmatrix} -1 & 2 & 1 \\ 4 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{y} \quad B^t = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 2 & 4 \end{bmatrix},$$

y de esta forma,

$$B^t A^t = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} -1 & 2 & 1 \\ 4 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -7 & 5 & 2 \\ 6 & 6 & 16 \end{bmatrix},$$

como se afirma en las propiedades de la transpuesta.

Ejemplo 1.5.4. Dada la matriz

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

y el escalar $\alpha = -2$, entonces:

$$\alpha A = -2A = \begin{bmatrix} 4 & -2 & -8 & -2 \\ -4 & 0 & 0 & -2 \\ -2 & -2 & 0 & -4 \end{bmatrix},$$

con lo cual

$$(\alpha A)^t = (-2A)^t = \begin{bmatrix} 4 & -4 & -2 \\ -2 & 0 & -2 \\ -8 & 0 & 0 \\ -2 & -2 & -4 \end{bmatrix}.$$

Además, tenemos que

$$\alpha A^t = -2A^t = -2 \begin{bmatrix} -2 & 2 & 1 \\ 1 & 0 & 1 \\ 4 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -4 & -2 \\ -2 & 0 & -2 \\ -8 & 0 & 0 \\ -2 & -2 & -4 \end{bmatrix},$$

es decir, $(-2A)^t = -2A^t$.

Definición 1.5.2 (Matriz simétrica). *Una matriz* $A = [a_{ij}]$ *cuadrada, de tamaño* $n \times n$*, se denomina simétrica si* $A = A^t$.

Definición 1.5.3 (Matriz antisimétrica). *Una matriz* $A = [a_{ij}]$ *cuadrada, de tamaño* $n \times n$*, se denomina antisimétrica si* $A^t = -A$.

Ejemplo 1.5.5. Las matrices

$$A = \begin{bmatrix} -2 & 2 & 1 \\ 2 & 0 & -3 \\ 1 & -3 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 2 & 1 \\ -2 & 0 & 2 \\ -1 & -2 & 0 \end{bmatrix}$$

son simétricas y antisimétricas respectivamente, puesto que

$$A^t = \begin{bmatrix} -2 & 2 & 1 \\ 2 & 0 & -3 \\ 1 & -3 & 1 \end{bmatrix} = A \quad \text{y} \quad B^t = \begin{bmatrix} 0 & -2 & -1 \\ 2 & 0 & -2 \\ 1 & 2 & 0 \end{bmatrix} = -B.$$

En el capítulo 7 ("Valores y vectores propios"), al estudiar la diagonalización veremos algunos aspectos que muestran la importancia de las matrices simétricas.

Ejemplo 1.5.6 (Uso de computador). En el programa DERIVE se tienen dos formas para definir matrices:

1. Usando la secuencia de comandos de AUTOR MATRIZ.

2. Empleando el comando AUTOR para editar expresiones.

Veamos cómo definir la matriz $\begin{bmatrix} 1 & 2 & x+1 \\ 2 & 4 & x-1 \end{bmatrix}$ utilizando los dos métodos anteriores.

Si aplicamos la secuencia AUTOR MATRIZ, el programa muestra una pantalla preguntando por el número de filas y el número de columnas de la matriz. En este caso indicamos que nuestra matriz tiene dos filas y tres columnas, como se ve en la Figura 1.1.[1]

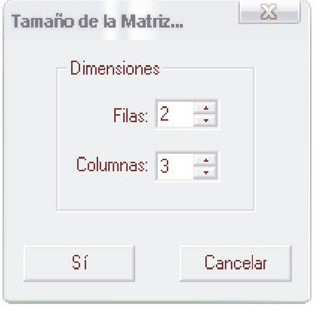


Figura 1.1. Introducción de matrices con DERIVE

A continuación introducimos uno por uno los elementos de la matriz, pulsando la tecla del tabulador después de escribir cada componente. Al finalizar, en la ventana aparece la expresión que se ve en la Figura 1.2.

[1] Ésta pantalla y las restantes que aparecen en el texto, se generaron con el programa DERIVE 6 de Texas Instruments Incorporated.

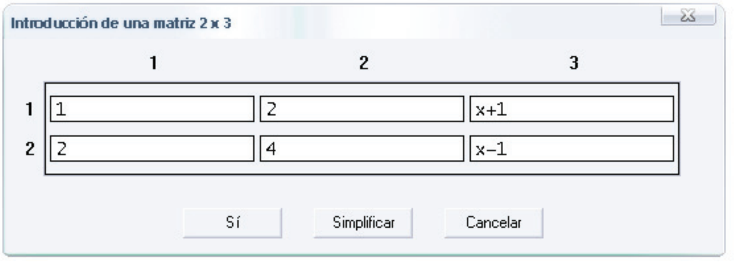


Figura 1.2. Pantalla para matrices 2×3 con DERIVE

El segundo método consiste en crear la matriz como un vector de vectores fila, utilizando el comando AUTOR. En este caso escribimos la expresión $[[1, 2, x + 1], [2, 4, x - 1]]$ al entrar los datos. Por cualquiera de los dos métodos obtenemos la expresión

$$\begin{bmatrix} 1 & 2 & x+1 \\ 2 & 4 & x-1 \end{bmatrix},$$

que representa a la matriz dada.

Ejemplo 1.5.7 (Uso de computador). Si vamos a utilizar una matriz en una serie de operaciones, en DERIVE podemos asignar un nombre a dicha matriz; para ello basta utilizar la sintaxis:

Nombre de la matriz := Matriz.

Así, si queremos definir las matrices

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -1 & 0 & 2 \\ 5 & 6 & 7 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & 1 & 7 \\ 5 & 0 & -3 \\ 3 & 1 & 0 \end{bmatrix},$$

podemos hacerlo utilizando el comando AUTOR y luego introducimos la expresión $A := [[1, 2, 4], [-1, 0, 2], [5, 6, 7]]$; después, para la matriz B, la expresión $B := [[2, 1, 7], [5, 0, -3], [3, 1, 0]]$, lo cual se muestra en DERIVE como

$$\#1 \qquad A := \begin{bmatrix} 1 & 2 & 4 \\ -1 & 0 & 2 \\ 5 & 6 & 7 \end{bmatrix}$$

$$\#2 \qquad B := \begin{bmatrix} 2 & 1 & 7 \\ 5 & 0 & -3 \\ 3 & 1 & 0 \end{bmatrix}.$$

Como se afirmó previamente, podemos realizar operaciones con estas matrices usando la variable correspondiente a cada una de ellas.

En el siguiente ejemplo mostramos cómo utilizar las variables almacenadas en el programa para realizar diversas operaciones con ellas.

Ejemplo 1.5.8 (Uso de computador). Dadas las matrices

$$A = \begin{bmatrix} 1 & 6 & 0 & 1 \\ 3 & 4 & 1 & 5 \\ 1 & -2 & 4 & 5 \\ 2 & 6 & 1 & 2 \\ 2 & 3 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & -1 & -1 & 3 & 5 \\ 2 & -4 & 1 & 0 & 13 \\ -1 & 12 & 1 & 4 & 5 \\ 3 & 3 & 2 & 1 & 0 \end{bmatrix} \quad \text{y}$$

$$C = \begin{bmatrix} 0 & 6 & -2 & 4 & 3 \\ 3 & 1 & 1 & 5 & 4 \\ -2 & 1 & -1 & 2 & 3 \\ 7 & 4 & 1 & 2 & 23 \end{bmatrix},$$

calcular las siguientes expresiones usando DERIVE:

1. B^tC.

2. $A(2B - 3C)$.

3. $A + 5C^t$.

Solución

Para calcular cada una de las expresiones, introducimos las matrices A, B y C; luego, solo basta con editar la operación correspondiente y simplificar, con lo cual obtenemos:

1. Desde AUTOR escribimos la expresión $B'C$, la entramos y después de simplificar se tiene:

$$\begin{bmatrix} 29 & 49 & -6 & 38 & 92 \\ -15 & 14 & -11 & 6 & 86 \\ 15 & 4 & 4 & 7 & 50 \\ -1 & 26 & -9 & 22 & 44 \\ 29 & 48 & -2 & 95 & 82 \end{bmatrix}.$$

2. Desde AUTOR escribimos la expresión $A(2B-3C)$, la entramos y después de simplificar se obtiene:

$$\begin{bmatrix} -33 & -92 & -1 & -100 & 16 \\ -55 & -113 & 18 & -96 & -285 \\ -37 & 56 & 31 & 12 & -368 \\ -32 & -97 & 9 & -108 & -51 \\ -9 & -22 & 22 & -59 & -91 \end{bmatrix}.$$

3. Desde AUTOR escribimos la expresión $A + 5C'$, la entramos y después de simplificar se obtiene:

$$\begin{bmatrix} 1 & 21 & -10 & 36 \\ 33 & 9 & 6 & 25 \\ -9 & 3 & -1 & 10 \\ 22 & 31 & 11 & 12 \\ 17 & 23 & 18 & 117 \end{bmatrix}.$$

Como lo ilustra el ejercicio, la realización de las operaciones básicas es supremamente sencilla.

1.6 Sistemas de ecuaciones lineales

En la forma usual de expresar una ecuación se suelen presentar las variables teniendo en cuenta la relación causa-efecto; así, al escribir la ecuación $\alpha = 2x$, los valores de la variable α se determinan inmediatamente si conocemos el valor de la x. En las llamadas *ecuaciones lineales*, al variar la causa, el efecto se altera en la misma proporción; por ejemplo, en la ecuación $\alpha = 2x$, si triplicamos el valor de la x, el valor de la variable α queda triplicado.

En los problemas directos, calculamos el valor de la variable dependiente a partir de los valores de las variables independientes. En el caso de los problemas inversos, nos interesa conocer los valores de las variables independientes para un valor fijo de la variable dependiente. De este modo, si consideramos un sistema de la forma

$$\begin{aligned} 3x + y &= \alpha \\ 2x - 2y &= \beta, \end{aligned}$$

nos interesa calcular los valores de las variables x y y que conducen a los valores dados de α y β.

Definición 1.6.1. *Un sistema de m ecuaciones lineales con n incógnitas $x_1, \ldots, x_n$ es un conjunto de relaciones de la forma*

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &=& b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &=& b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad &=& \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &=& b_m \end{array}$$

donde los coeficientes a_{ij} y b_i son escalares (en este texto, los escalares serán números reales).

El conjunto de valores $\alpha_1, \alpha_2, \ldots \alpha_n$ es una solución del sistema si al sustituir $x_1, x_2, \ldots x_n$ por $\alpha_1, \alpha_2, \ldots \alpha_n$ en las m ecuaciones del sistema, se obtienen m igualdades.

Si el sistema tiene al menos una solución, se dice que es *consistente*; en caso contrario se denomina *inconsistente*. Cuando la solución del sistema es única, el sistema se conoce como *sistema determinado*; si existe más de una solución, el sistema se denomina *sistema indeterminado.*

Ejemplo 1.6.1. El sistema de ecuaciones

$$\begin{aligned} 3x + 2y - z &= 6 \\ x - y + z &= 8 \\ 2x + y - z &= 6 \end{aligned}$$

es un sistema determinado, cuya solución está dada por $x = 14/3$, $y = -14/3$ y $z = -4/3$.

El sistema de ecuaciones

$$\begin{aligned} 3x + 2y - z &= 6 \\ x - y + z &= 8 \\ 4x + y &= 7 \end{aligned}$$

es un sistema inconsistente, mientras que el sistema

$$\begin{aligned} x - y + z &= 8 \\ -4x - 5y + 5z &= -2 \\ 2x + y - z &= 6 \end{aligned}$$

es un sistema indeterminado, con soluciones que pueden escribirse en la forma $x = 14/3$, $y = t - 10/3$ y $z = t$. Posteriormente, en el teorema 1.8.2, se enuncian criterios que nos permitan decidir cuándo un sistema dado es consistente, inconsistente o indeterminado.

1.7 Forma matricial de un sistema

Dado el sistema lineal de m ecuaciones y n incógnitas

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad & = & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m, \end{array}$$

este queda completamente representado si conocemos matrices con los $m \times n$ coeficientes de las incógnitas y los m términos independientes b_i, situados en las mismas posiciones relativas que estos ocupan en el sistema. A partir de este razonamiento podemos establecer la siguiente definición.

Definición 1.7.1. *Dado el sistema lineal de m ecuaciones y n incógnitas*

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad & = & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m, \end{array}$$

se llaman matriz de coeficientes del sistema, matriz de términos independientes y matriz aumentada a las siguientes matrices:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad y$$

$$A|B = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Así mismo, tal sistema se escribirá de modo abreviado $AX = B$, donde X es el vector de incógnitas dado por

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Si todos los términos independientes son nulos, es decir, $b_i = 0$, para $i = 1, \ldots, n$, el sistema se denomina *sistema homogéneo.* Es inmediato que si el sistema es homogéneo, siempre tiene al menos una solución que es $x_1 = \cdots = x_n = 0$ y que se suele llamar la *solución trivial.* Además, si se conocen dos soluciones diferentes α_i y β_i, entonces la expresión $\lambda\alpha_i + \gamma\beta_i$, con λ y γ en $\mathbb{R}$, es también solución del sistema homogéneo.

Ejemplo 1.7.1. Dado el sistema

$$\begin{aligned} x - y + z &= 8 \\ -4x - 4y + 5z &= -2 \\ 2x + y - z &= 6, \end{aligned}$$

la matriz de coeficientes del sistema es:

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

mientras que la matriz aumentada del sistema es:

$$A|B = \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ -4 & -4 & 5 & -2 \\ 2 & 1 & -1 & 6 \end{array}\right];$$

y la forma matricial del sistema puede escribirse como

$$\begin{bmatrix} 1 & -1 & 1 \\ -4 & -4 & 5 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ -2 \\ 6 \end{bmatrix}.$$

1.8 Sistemas equivalentes

En definitiva, para resolver sistemas de ecuaciones lineales lo que se hace es reducirlos a otros sistemas más simples, es decir, se buscan sistemas que resulten más fáciles de resolver que el sistema inicial, pero que tengan las mismas soluciones que el original. Como los sistemas están conformados por igualdades, es claro que estas igualdades no se alteran cuando realizamos sobre ellas una o varias de las siguientes transformaciones:

- Cambiar el orden de las ecuaciones.
- Multiplicar una de las ecuaciones por un número distinto de cero.
- Sumar a una de las ecuaciones otra ecuación multiplicada por un número.

Con estas aclaraciones en mente, podemos establecer la siguiente definición:

Definición 1.8.1. *Dos sistemas de ecuaciones, con las mismas incógnitas, son equivalentes si tienen las mismas soluciones.*

Con esta definición, podemos establecer el resultado principal de la equivalencia de sistemas: un sistema de ecuaciones lineales es equivalente a cualquier otro que resulte después de aplicarle al sistema original una secuencia finita de operaciones elementales, entendiendo por *operación elemental* una de las siguientes operaciones:

1. Cambiar el orden de las ecuaciones.
2. Multiplicar una de las ecuaciones por un número distinto de cero.
3. Sumar a una de las ecuaciones otra ecuación multiplicada por un número.

Ejemplo 1.8.1. Consideremos el siguiente sistema lineal de ecuaciones:

$$\left.\begin{array}{lr} 1 \rightarrow & x+y-z=4 \\ 2 \rightarrow & 2x-y+z=5 \\ 3 \rightarrow & 3x+2y-3z=10 \end{array}\right\} S;$$

mediante las transformaciones indicadas, se obtiene:

$$\left.\begin{array}{lcr} 1'=1 & \rightarrow & x+y-z=4 \\ 2'=2-2(1) & \rightarrow & -3y+3z=-3 \\ 3'=3-3(1) & \rightarrow & -y=-2 \end{array}\right\} S',$$

el cual puede finalmente reducirse a

$$\left.\begin{array}{lcr} 1''=1' & \rightarrow & x+y-z=4 \\ 2''=-(3') & \rightarrow & y=2 \\ 3''=\frac{1}{3}(2'-3(3')) & \rightarrow & z=1 \end{array}\right\} S''.$$

Los sistemas S, S' y S'' son equivalentes, pero el sistema S'' tiene como solución $y=2$, $z=1$ y $x=3$. Obviamente, estos valores son soluciones para los sistemas S y S'.

El procedimiento que se aplicó en el ejemplo anterior puede establecerse trabajando con la matriz aumentada del sistema, lo cual permite un manejo más fácil para las operaciones.

Desde el punto de vista de las matrices, la noción de *equivalencia* corresponde a la siguiente definición.

Definición 1.8.2. *Dos matrices son equivalentes por filas, si es posible obtener una de ellas mediante la aplicación de una secuencia finita de operaciones elementales sobre las filas de la otra matriz. Se consideran como operaciones elementales las siguientes transformaciones:*

- *Cambiar el orden de las filas. Si la fila i se intercambia con la fila j, lo indicamos mediante la expresión $F_i \leftrightarrow F_j$.*

- *Multiplicar una de las filas por un número distinto de cero. Si la fila i se multiplica por el escalar $c \neq 0$, lo indicamos mediante la expresión cF_i.*

- *Sumar, a una de las filas, otra fila multiplicada por un número. Si a la fila F_i se le suma cF_j, lo indicamos con $F_i + cF_j$ y es claro que los nuevos valores quedan en la fila i.*

Al considerar la matriz aumentada de un sistema de ecuaciones, estas pueden hacerse equivalentes a matrices que representan sistemas de muy fácil solución. Para aclarar este aspecto, vamos a introducir la siguiente definición.

Definición 1.8.3 (Matriz escalonada reducida). *Se dice que una matriz A de tamaño $m \times n$ está en la forma escalonada reducida si satisface las siguientes condiciones:*

1. *Las filas nulas de A, si las hay, están debajo de las filas no nulas.*

2. *El primer elemento no nulo de izquierda a derecha en cada fila no nula es un uno. A este uno se le llama pivote.*

3. *El número de ceros a la izquierda de los pivotes aumenta a medida que se desciende.*

4. *El único elemento distinto de cero en la columna donde figura el pivote es precisamente este uno.*

Si la matriz cumple solo las tres primeras condiciones se dice que la matriz está en la forma escalonada.

Ejemplo 1.8.2. Las siguientes matrices son escalonadas:

$$A = \begin{bmatrix} \mathbf{1} & 2 & -3 & 4 & 4 & 3 \\ 0 & 0 & \mathbf{1} & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \mathbf{1} & -1 & 3 \\ 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ejemplo 1.8.3. Las siguientes matrices son escalonadas reducidas

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

El siguiente teorema nos proporciona un algoritmo bastante eficaz para determinar las soluciones de un sistema de ecuaciones lineales.

Teorema 1.8.1. *Toda matriz A de tamaño $m \times n$ es equivalente por filas a una única matriz en la forma escalonada reducida.*

Ejemplo 1.8.4. Dada la matriz

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

al aplicar las operaciones elementales por fila, se obtiene:

$$\begin{bmatrix} 4 & 2 & 2 & 2 \\ 1 & 0 & 0 & 1 \\ 0 & 2 & 2 & -2 \\ 5 & 0 & 0 & -1 \end{bmatrix} \xrightarrow[F_1 \leftrightarrow F_2]{} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 4 & 2 & 2 & 2 \\ 0 & 2 & 2 & -2 \\ 5 & 0 & 0 & -1 \end{bmatrix}$$

$$\xrightarrow[F_2 - 4F_1]{F_4 - 5F_1} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 2 & -2 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 0 & -6 \end{bmatrix} \xrightarrow[\frac{1}{2} F_2]{} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 0 & -6 \end{bmatrix}$$

$$\xrightarrow[F_3 - 2F_2]{} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -6 \end{bmatrix} \xrightarrow[F_3 \leftrightarrow F_4]{} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow[-\frac{1}{6}F_3]{} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[F_1 - F_3]{F_2 + F_3} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Esta última matriz es equivalente por filas a la matriz A y se encuentra en la forma escalonada reducida.

Veamos cómo usar este tipo de matrices para la solución de sistemas de ecuaciones lineales. Dado el sistema lineal de ecuaciones $AX = B$, al considerar la matriz aumentada del sistema $A|B$ se dice que un elemento $a_{ij} \neq 0$ se utiliza de pivote hacia abajo, si en las filas de la matriz aumentada se realizan operaciones que transforman a a_{ij} en 1 y a los elementos de la columna j, por debajo de la fila i, en ceros.

Si se toman como pivotes a ciertos elementos de A, es posible transformar la matriz aumentada del sistema en una matriz escalonada reducida. En general, se puede establecer el siguiente algoritmo para transformar la matriz aumentada del sistema en una matriz escalonada reducida.

1. En la primera columna de A debe existir un elemento no nulo (de lo contrario, se eliminaría una incognita). Mediante alguna operación elemental se transforma la posición $(1, 1)$ en 1 y se pivotean con este elemento los demás elementos de la columna.

2. Si todos los elementos de la segunda columna que quedan por debajo de la posición $(1, 2)$ son todos nulos, se deja la

segunda columna tal cual. Si hay algún elemento no nulo en esta columna por debajo de la posición $(1, 2)$, mediante alguna operación elemental se transforma el elemento de la posición $(2, 2)$ en 1 y se pivotea con él los demás elementos de la columna.

3. Con los pasos anteriores, las columnas primera y segunda ya tienen la forma deseada. Se pasa ahora a la tercera columna, se procede de manera similar y se reitera el procedimiento con las columnas siguientes hasta la última columna de la matriz.

Este algoritmo se conoce como *proceso de escalonamiento* y como veremos al finalizar esta sección, permite transformar un sistema lineal de ecuaciones en un sistema equivalente que tiene una solución sencilla.

Ejemplo 1.8.5. Transformar la matriz aumentada del sistema dado en una matriz escalonada reducida.

$$\begin{aligned} 3x_1 + x_2 + x_3 &= 7 \\ x_1 - x_2 + x_3 &= 8 \\ 2x_1 + x_2 - x_3 &= 4. \end{aligned}$$

Escribir el sistema equivalente y determinar su solución.

Solución

La matriz aumentada del sistema es

$$\left[\begin{array}{rrr|r} 3 & 1 & 1 & 7 \\ 1 & -1 & 1 & 8 \\ 2 & 1 & -1 & 4 \end{array}\right].$$

Como el primer paso del proceso de escalonamiento es obtener un 1 en la posición a_{11}, una alternativa simple para realizar esto es el intercambio de las filas primera y segunda. Es costumbre indicar la operación u operaciones elementales en la forma que mostramos a continuación:

$$\xrightarrow[F_1 \leftrightarrow F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 3 & 1 & 1 & 7 \\ 2 & 1 & -1 & 4 \end{array}\right].$$

En el ejemplo se ha encerrado en un círculo el elemento que se va a utilizar para transformar en ceros los demás elementos de la primera columna; las operaciones que se requieren para esto se indican antes de la matriz.

$$\xrightarrow[F_3 - 2F_1]{F_2 - 3F_1} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 4 & -2 & -17 \\ 0 & 3 & -3 & -12 \end{array}\right]$$

Como los elementos de la segunda columna por debajo da la posición a_{12} no son nulos, es necesario transformar el elemento de la posición a_{22} en 1. Esto puede realizarse mediante la aplicación de diversas operaciones elementales, pero es recomendable utilizar aquella que sea más fácil de aplicar y dé resultados más cómodos para cálculos posteriores. Para el ejemplo, escogemos restar a la segunda fila la tercera, con lo cual

$$\xrightarrow[F_2 - F_3]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 1 & 1 & -5 \\ 0 & 3 & -3 & -12 \end{array}\right].$$

Con esto conseguimos el pivote para los elementos de la segunda columna por debajo de la posición a_{22}, obteniendo:

$$\xrightarrow[F_3 - 3F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 1 & 1 & -5 \\ 0 & 0 & -6 & 3 \end{array}\right].$$

La inspección de la matriz resultante nos muestra ahora que hay términos no nulos por debajo de la posición a_{23}. Para este caso, la opción que tenemos para conseguir el pivote de la posición a_{33}, es dividir la tercera fila por -6, con lo cual

$$\xrightarrow[-\frac{1}{6}F_3]{} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 8 \\ 0 & 1 & 1 & -5 \\ 0 & 0 & 1 & -\frac{1}{2} \end{array}\right]. \tag{1.8.1}$$

El sistema correspondiente a esta matriz es:

$$\begin{aligned} x_1 - x_2 + x_3 &= 8 \\ x_2 + x_3 &= -5 \\ x_3 &= -\frac{1}{2}. \end{aligned}$$

Este sistema puede resolverse sustituyendo, inicialmente, los valores de x_3 en la segunda ecuación y, luego, los valores de x_2 y x_3 en la primera ecuación. De esta forma, la solución del sistema es:

$$\begin{aligned} x_3 &= -\frac{1}{2} \\ x_2 &= -5 - x_3 = -\frac{9}{2} \\ x_1 &= 8 + x_2 - x_3 = 4. \end{aligned}$$

El procedimiento que se acabó de aplicar se conoce con el nombre de *sustitución en reversa* o también como el *método de eliminación de Gauss.*

Si llevamos la matriz 1.8.1 a la forma escalonada reducida mediante las operaciones elementales indicadas, obtenemos

$$\xrightarrow[F_1+F_2]{} \left[\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 0 & 1 & 1 & -5 \\ 0 & 0 & 1 & -\frac{1}{2} \end{array}\right]$$

$$\xrightarrow[F_2-F_3]{F_1-2F_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -\frac{9}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \end{array}\right].$$

En este punto es importante aclarar que si se quiere llevar la matriz a la forma escalonada reducida, lo más conveniente sería eliminar los elementos superiores de cada pivote, comenzando el proceso de escalonamiento en sentido contrario al que se aplicó para obtener la forma escalonada; es decir, en nuestro ejemplo, usar el pivote de la tercera fila para eliminar los demás elementos de dicha columna y luego eliminar usando el pivote de la segunda columna. El sistema equivalente a esta matriz es

$$\begin{aligned} x_3 &= -\frac{1}{2} \\ x_2 &= -\frac{9}{2} \\ x_1 &= 4. \end{aligned}$$

Cuando la matriz aumentada del sistema se lleva hasta la forma escalonada reducida estamos hablando del *método de eliminación de Gauss-Jordan*.

Ejemplo 1.8.6 (Uso de computador). En DERIVE existe un comando que permite resolver sistemas de ecuaciones (no solamente lineales). A este comando se puede acceder desde el menú principal (con SOLVE o RESOLVER) o utilizando la combinación de teclas *Ctrl+May+y*. En tal situación se despliega una pantalla preguntando por el número de ecuaciones del sistema; si en este caso marcamos 3, obtenemos la Figura 1.3.

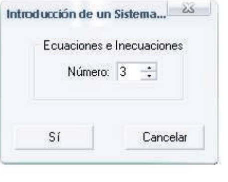


Figura 1.3. Pantalla para insertar el número de ecuaciones en DERIVE

Una vez aceptamos, aparece una tabla para escribir las ecuaciones que conforman el sistema (véase Figura 1.4). Para editar cada fila utilizamos el tabulador; además, es necesario resaltar, en el cuadro de “Variables”, cuáles son las del sistema.

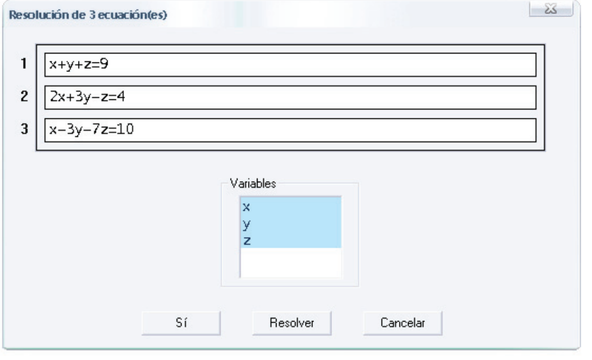


Figura 1.4. Pantalla para introducir ecuaciones en DERIVE

Luego de entrar el sistema y simplificar, obtenemos la solución del mismo en la forma

$$\text{SOLVE}\,([x+y+z=9,\ 2x+3y-z=4,\ x-3y-7z=10],\ [x,y,z])$$
$$\left[x=12 \wedge y=-\tfrac{23}{4} \wedge z=\tfrac{11}{4}\right].$$

Si el sistema no posee solución, no aparecerá nada en el corchete de soluciones.

Para determinar si un sistema de m ecuaciones lineales con n incógnitas tiene solución única, no tiene solución o tiene infinitas soluciones, definimos lo que es el rango de una matriz.

Definición 1.8.4 (Rango de una matriz). *Sea A una matriz de tamaño $m \times n$; entonces, el rango de A, que denotamos con $\rho(A)$, se define como el número de unos pivote que aparecen en la forma escalonada reducida de A.*

De la definición se desprende que $\rho(A) \leq m$ y $\rho(A) \leq n$.

Ejemplo 1.8.7. Hallar el rango de la matriz

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

Solución

En este caso, lo que se requiere es determinar la forma escalonada reducida de la matriz dada. Para ello procedemos a escalonar la matriz original, mediante la aplicación de operaciones elementales por fila.

$$\begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & 2 & 2 & 2 \\ -1 & -2 & 0 & 1 \end{bmatrix} \xrightarrow{F_3+F_1} \begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & 2 & 2 & 2 \\ 0 & 2 & 2 & 2 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}F_2} \begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \end{bmatrix} \xrightarrow{F_3-2F_2} \begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

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

Esta última matriz está en la forma escalonada reducida y el número de unos pivote es 2, con lo cual $\rho(A) = 2$.

Ejemplo 1.8.8. Hallar el rango de la matriz

$$B = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \\ 0 & 4 & 1 \end{bmatrix}.$$

Solución

Como en el caso anterior, lo que se requiere es determinar la forma escalonada reducida de la matriz dada, para lo cual escalonamos la matriz original mediante la aplicación de operaciones elementales por fila.

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \\ 0 & 4 & 1 \end{bmatrix} \xrightarrow{\frac{1}{2}F_2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 4 & 1 \end{bmatrix}$$

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

Esta última matriz está en la forma escalonada reducida y el número de unos pivote es 3; en consecuencia, $\rho(B) = 3$.

El siguiente teorema nos permite afirmar si un sistema de m ecuaciones lineales con n incógnitas tiene solución única, infinitas soluciones o ninguna.

Teorema 1.8.2. *Sea $AX = B$ un sistema de m ecuaciones lineales con n incógnitas; entonces:*

1. *Si $\rho(A|B) = \rho(A)$, el sistema $AX = B$ es consistente. En este caso:*

   - *Si $\rho(A|B) = \rho(A) < n$, el sistema tiene infinitas soluciones, es decir, es indeterminado.*
   - *Si $\rho(A|B) = \rho(A) = n$, el sistema tiene solución única.*

2. *Si $\rho(A|B) > \rho(A)$, el sistema no tiene solución, es decir, es inconsistente.*

Corolario 1.8.1. *Sea $AX = 0$ un sistema homogéneo de m ecuaciones lineales con n incógnitas; entonces:*

1. *El sistema es consistente, esto es, tiene al menos una solución, conocida como solución trivial, la cual es $x_i = 0$ para todo valor de i.*

    - *Si $\rho(A) < n$, el sistema tiene infinitas soluciones, es decir, es indeterminado.*
    - *Si $\rho(A) = n$, el sistema tiene solución única.*

2. *Como $\rho(A|0) \leq \rho(A)$, el sistema siempre es consistente.*

En los siguientes ejemplos vamos a tomar los sistemas de ecuaciones presentados en la página 35 y veremos cómo se aplica el teorema precedente para determinar si el sistema dado es consistente o inconsistente.

Ejemplo 1.8.9. Consideremos el sistema de ecuaciones lineales

$$\begin{aligned} 3x + 2y - z &= 6 \\ x - y + z &= 8 \\ 2x + y - z &= 6. \end{aligned}$$

Determinar si el sistema es consistente o inconsistente.

Solución

Para dicho sistema, la matriz aumentada es

$$\left[\begin{array}{rrr|r} 3 & 2 & -1 & 6 \\ 1 & -1 & 1 & 8 \\ 2 & 1 & -1 & 6 \end{array}\right].$$

Vamos a calcular la forma escalonada reducida de dicha matriz.

$$\xrightarrow[F_1 \leftrightarrow F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 3 & 2 & -1 & 6 \\ 2 & 1 & -1 & 6 \end{array}\right] \xrightarrow[F_3 - 2F_1]{F_2 - 3F_1} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 5 & -4 & -18 \\ 0 & 3 & -3 & -10 \end{array}\right]$$

$$\xrightarrow[\frac{1}{5}F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 1 & -\frac{4}{5} & -\frac{18}{5} \\ 0 & 3 & -3 & -10 \end{array}\right] \xrightarrow[F_1+F_2]{F_3-3F_2} \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{5} & \frac{22}{5} \\ 0 & 1 & -\frac{4}{5} & -\frac{18}{5} \\ 0 & 0 & -\frac{3}{5} & \frac{4}{5} \end{array}\right]$$

$$\xrightarrow[-\frac{5}{3}F_3]{} \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{5} & \frac{22}{5} \\ 0 & 1 & -\frac{4}{5} & -\frac{18}{5} \\ 0 & 0 & 1 & -\frac{4}{3} \end{array}\right] \xrightarrow[F_2+\frac{4}{5}F_3]{F_1-\frac{1}{5}F_3} \left[\begin{array}{rrr|r} 1 & 0 & 0 & \frac{14}{3} \\ 0 & 1 & 0 & -\frac{14}{3} \\ 0 & 0 & 1 & -\frac{4}{3} \end{array}\right].$$

En este caso tenemos que $\rho(A|B) = \rho(A) = n$, donde n es el número de incógnitas. Por lo tanto, el sistema tiene solución única y sus valores están dados por

$$x = \frac{14}{3}, \quad y = -\frac{14}{3}, \quad z = -\frac{4}{3}.$$

Ejemplo 1.8.10. Determinar si el siguiente sistema de ecuaciones lineales es consistente o inconsistente.

$$\begin{aligned} 3x + 2y - z &= 6 \\ x - y + z &= 8 \\ 4x + y &= 7. \end{aligned}$$

Solución

Su matriz aumentada es

$$\left[\begin{array}{rrr|r} 3 & 2 & -1 & 6 \\ 1 & -1 & 1 & 8 \\ 4 & 1 & 0 & 7 \end{array}\right].$$

Vamos a calcular la forma escalonada reducida para esta matriz:

$$\left[\begin{array}{rrr|r} 3 & 2 & -1 & 6 \\ 1 & -1 & 1 & 8 \\ 4 & 1 & 0 & 7 \end{array}\right] \xrightarrow[F_1 \leftrightarrow F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 3 & 2 & -1 & 6 \\ 4 & 1 & 0 & 7 \end{array}\right]$$

$$\xrightarrow[F_3-4F_1]{F_2-3F_1} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 5 & -4 & -18 \\ 0 & 5 & -4 & -25 \end{array}\right] \xrightarrow[F_3-F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 5 & -4 & -18 \\ 0 & 0 & 0 & -7 \end{array}\right].$$

En este punto se puede ver que el sistema es inconsistente, puesto que el sistema equivalente es

$$\begin{aligned} x - y + z &= 8 \\ 5y - 4z &= -18 \\ 0 &= -7. \end{aligned}$$

Sin embargo, como se trata de ilustrar el teorema, continuamos con el proceso de llevar la matriz aumentada del sistema a su forma escalonada reducida, así:

$$\xrightarrow[\frac{1}{5}F_2]{-\frac{1}{7}F_3} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 1 & -\frac{4}{5} & -\frac{18}{5} \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[F_1+F_2]{} \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{5} & \frac{22}{5} \\ 0 & 1 & -\frac{4}{5} & -\frac{18}{5} \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[F_2+\frac{18}{5}F_3]{F_1-\frac{22}{5}F_3} \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{5} & 0 \\ 0 & 1 & -\frac{4}{5} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

En este caso tenemos que $\rho(A|B) = 3$, mientras que $\rho(A) = 2$. Por lo tanto, $\rho(A|B) > \rho(A)$ y, en consecuencia, el sistema no tiene solución.

Ejemplo 1.8.11. Dado el siguiente sistema de ecuaciones:

$$\begin{aligned} x - y + z &= 8 \\ -4x - 5y + 5z &= -2 \\ 2x + y - z &= 6, \end{aligned}$$

determine si el sistema tiene solución única, infinitas soluciones o si es inconsistente.

Solución

De nuevo, debemos considerar la forma escalonada reducida de la matriz aumentada del sistema:

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ -4 & -5 & 5 & -2 \\ 2 & 1 & -1 & 6 \end{array}\right].$$

Para calcular dicha forma, realizamos la siguiente secuencia de operaciones elementales por filas:

$$\xrightarrow[F_2+4F_1]{F_3-2F_1} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & -9 & 9 & 30 \\ 0 & 3 & -3 & -10 \end{array}\right] \xrightarrow[F_2\leftrightarrow F_3]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 3 & -3 & -10 \\ 0 & -9 & 9 & 30 \end{array}\right]$$

$$\xrightarrow[F_3+3F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 3 & -3 & -10 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow[\frac{1}{3}F_2]{} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 8 \\ 0 & 1 & -1 & -\frac{10}{3} \\ 0 & 0 & 0 & 0 \end{array}\right];$$

finalmente,

$$\xrightarrow[F_1+F_2]{} \left[\begin{array}{rrr|r} 1 & 0 & 0 & \frac{14}{3} \\ 0 & 1 & -1 & -\frac{10}{3} \\ 0 & 0 & 0 & 0 \end{array}\right],$$

lo cual es equivalente al sistema

$$\begin{aligned} x &= 14/3 \\ y - z &= -10/3. \end{aligned}$$

En este caso, es posible despejar las variables correspondientes a los pivotes para obtener como solución

$$x = 14/3, \;\; y = z - 10/3.$$

Aquí, las variables que corresponden a los pivotes pueden resolverse en términos de las demás variables. A las asociadas a los pivotes se les denomina *variables básicas,* mientras que a las otras se les suele llamar *variables libres.*

Puesto que los valores de las variables básicas dependen de los valores que les asignemos a ellas, la solución puede escribirse en la forma

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 14 \\ -10 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

donde t es un escalar cualquiera. En este caso tenemos que $\rho(A|B) = \rho(A) = 2 < 3$ (número de incógnitas) y, en consecuencia, el sistema tiene infinitas soluciones.

Ejemplo 1.8.12. Consideremos el sistema

$$\begin{aligned} x - y + z &= 0 \\ y + z &= 0. \end{aligned}$$

La matriz aumentada del sistema es

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

la cual es equivalente por filas a la matriz

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

que está en la forma escalonada reducida.

Como $\rho(A|B) = \rho(A) = 2 < 3$ (número de incógnitas), entonces el sistema tiene infinitas soluciones. La solución del sistema puede escribirse como: $z = t$, $y = -t$ y $x = -2t$, o en forma vectorial:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}.$$

Nótese que, en este caso, z es un parámetro del cual dependen los valores de las otras dos variables.

En el ejemplo anterior se presentó un sistema homogéneo con una solución distinta de la trivial. El siguiente teorema nos muestra el caso cuando este tipo de solución se presenta.

Teorema 1.8.3. *Sea $AX = 0$ un sistema homogéneo de m ecuaciones lineales con n incógnitas. Si $m < n$ (el número de ecuaciones es menor que el número de incógnitas), entonces el sistema tiene una solución no trivial; en consecuencia, tiene infinitas soluciones.*

Demostración. Sea C la forma escalonada reducida de A; entonces, los sistemas $AX = 0$ y $CX = 0$ son equivalentes.

Sea $\rho(C) = r$ el rango de C; en consecuencia, $r \leq m$. Como $m < n$, entonces $r < n$.

Si ahora despejamos las variables correspondientes a los pivotes, tendremos r incógnitas en términos de las $n - r$ restantes, donde estas últimas pueden tomar cualquier valor. Si una de estas $n - r$ variables es distinta de cero, tendremos una solución no trivial del sistema $CX = 0$ y, por lo tanto, una solución no trivial del sistema $AX = 0$.

En particular, si X_0 es una solución del sistema $AX = 0$, es decir, si $AX_0 = 0$, entonces, para cualquier escalar λ se cumple que $A(\lambda X_0) = \lambda AX_0 = 0$, lo cual nos permite concluir que cualquier múltiplo escalar de X_0 es también solución del sistema $AX = 0$, esto es, dicho sistema tiene infinitas soluciones.

Ejemplo 1.8.13. Consideremos el sistema de ecuaciones

$$\begin{aligned} x + y - z - w &= 0 \\ -y + 2z &= 0 \\ x + z + w &= 0. \end{aligned}$$

Como es un sistema homogéneo con más incógnitas que ecuaciones, tiene infinitas soluciones; para verificarlo, llevemos la matriz del sistema a la forma escalonada reducida.

$$\left[\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 0 \\ 0 & -1 & 2 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{array}\right] \xrightarrow[F_3-F_1]{} \left[\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 0 \\ 0 & -1 & 2 & 0 & 0 \\ 0 & -1 & 2 & 2 & 0 \end{array}\right]$$

$$\xrightarrow[-F_2]{} \left[\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & -1 & 2 & 2 & 0 \end{array}\right] \xrightarrow[F_1-F_2]{F_3+F_2} \left[\begin{array}{rrrr|r} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{array}\right]$$

$$\xrightarrow[\frac{1}{2}F_3]{} \left[\begin{array}{rrrr|r} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \xrightarrow[F_1+F_3]{} \left[\begin{array}{rrrr|r} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

El sistema equivalente, en este caso, es:

$$\begin{aligned} x + z &= 0 \\ y - 2z &= 0 \\ w &= 0. \end{aligned}$$

Al analizar este sistema, notamos que hay tres variables básicas (los tres pivotes) y, por lo tanto, todas las variables básicas se pueden expresar en términos de las variables no básicas o parámetros. De esta forma obtenemos:

$$\begin{aligned} x &= -z \\ y &= 2z \\ w &= 0, \end{aligned}$$

con lo cual concluimos que el sistema posee infinitas soluciones, que dependen del valor que le asignemos a la variable z.

El teorema anterior solo es válido para sistemas homogéneos. Para ver esto, a continuación presentamos un ejemplo de un sistema no homogéneo con un número menor de ecuaciones que incógnitas, pero que no posee solución.

Ejemplo 1.8.14. Sea el sistema

$$\begin{aligned} -2x + y + z &= 10 \\ 4x - 2y - 2z &= 40. \end{aligned}$$

La matriz aumentada de este sistema es

$$\left[\begin{array}{rrr|r} -2 & 1 & 1 & 10 \\ 4 & -2 & -2 & 40 \end{array}\right],$$

cuya forma escalonada reducida se obtiene así:

$$\xrightarrow[-\frac{1}{2}F_1]{} \left[\begin{array}{rrr|r} 1 & -\frac{1}{2} & -\frac{1}{2} & -5 \\ 4 & -2 & -2 & 40 \end{array}\right] \xrightarrow[F_2-4F_1]{} \left[\begin{array}{rrr|r} 1 & -\frac{1}{2} & -\frac{1}{2} & -5 \\ 0 & 0 & 0 & 60 \end{array}\right]$$

$$\xrightarrow[\frac{1}{60}F_2]{} \left[\begin{array}{rrr|r} 1 & -\frac{1}{2} & -\frac{1}{2} & -5 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[F_1+5F_2]{} \left[\begin{array}{rrr|r} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Como en este caso $\rho(A|B) = 2$ y $\rho(A) = 1$, entonces $\rho(A|B) > \rho(A)$ y, en consecuencia, el sistema no tiene solución.

El siguiente teorema sobre sistemas de ecuaciones lineales tiene una importante aplicación en la teoría de las ecuaciones diferenciales cuando se estudian las ecuaciones diferenciales no homogéneas.

Teorema 1.8.4. *Sea $AX = B$ un sistema de m ecuaciones lineales con n incógnitas; se tiene que:*

1. *Si X_1 y X_2 son soluciones de $AX = B$, entonces $X_1 - X_2$ es solución del sistema $AX = 0$.*

2. *Si X_0 es una solución particular de $AX = B$, toda solución X^* de $AX = B$ es de la forma $X^* = X_0 + H$, donde H es solución del sistema $AX = 0$.*

Demostración. Para la primera parte, sean X_1 y X_2 soluciones de $AX = B$; entonces, $AX_1 = B$ y $AX_2 = B$. Por lo tanto, $AX_1 - AX_2 = A(X_1 - X_2) = B - B = 0$ y, en consecuencia, $(X_1 - X_2)$ es solución de $AX = 0$.

Para la segunda parte, supongamos que X_0 es un solución del sistema $AX = B$, es decir, $AX_0 = B$, y sea X^* cualquier otra solución de $AX = B$; entonces, $AX^* = B$ y, por lo tanto, $AX^* - AX_0 = A(X^* - X_0) = 0$, con lo cual probamos que $X^* - X_0$ es una solución de $AX = 0$. Así, existe un cierto H, con $H = X^* - X_0$, tal que $AH = 0$. Para este H se tiene que $X^* = X_0 + H$.

Ejemplo 1.8.15. Se puede ver que

$$X_0 = \begin{bmatrix} 1 \\ 13 \\ 4 \end{bmatrix}$$

es una solución particular del siguiente sistema de tres ecuaciones lineales con tres incógnitas:

$$\begin{aligned} x + y - z &= 10 \\ 2x + z &= 6 \\ 4x + 2y - z &= 26. \end{aligned}$$

Para ver que toda solución X^* del sistema no homogéneo $AX = B$ es de la forma $X^* = X_0 + H$, con H como solución del sistema homogéneo $AX = 0$, vamos a determinar las soluciones de este último, es decir, vamos a resolver el sistema:

$$\begin{aligned} x + y - z &= 0 \\ 2x + z &= 0 \\ 4x + 2y - z &= 0. \end{aligned}$$

Aplicando el método de eliminación de Gauss-Jordan.

$$\left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 2 & 0 & 1 & 0 \\ 4 & 2 & -1 & 0 \end{array}\right] \xrightarrow[F_3-4F_1]{F_2-2F_1} \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & -2 & 3 & 0 \\ 0 & -2 & 3 & 0 \end{array}\right]$$

$$\xrightarrow[F_3-F_2]{} \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & -2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow[-\frac{1}{2}F_2]{} \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & 1 & -\frac{3}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

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

De esta forma, la solución del sistema lineal homogéneo corresponde a las soluciones del sistema equivalente

$$\begin{aligned} x + \frac{1}{2}z &= 0 \\ y - \frac{3}{2}z &= 0, \end{aligned}$$

la cual puede escribirse en la forma

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}z \\ \frac{3}{2}z \\ z \end{bmatrix}$$

y, en consecuencia, según el teorema anterior, toda solución del sistema

$$\begin{aligned} x + y - z &= 10 \\ 2x + z &= 6 \\ 4x + 2y - z &= 26 \end{aligned}$$

es de la forma

$$X^* = \begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}z \\ \frac{3}{2}z \\ z \end{bmatrix} + \begin{bmatrix} 1 \\ 13 \\ 4 \end{bmatrix} = z \begin{bmatrix} -\frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 13 \\ 4 \end{bmatrix},$$

siendo z un escalar cualquiera. En este ejemplo puede verse claramente que la solución del sistema no homogéneo se compone de la solución general del sistema homogéneo asociado y una solución particular del sistema no homogéneo dado.

1.9 Matrices invertibles

Cuando definimos el producto de matrices y estudiamos sus propiedades hicimos notar que la operación no es conmutativa. Hay también otra propiedad muy útil en el caso del producto de dos números reales que nos permite resolver el tipo más sencillo de ecuación lineal, esto es, una ecuación de la forma $ax = b$, con $a \neq 0$, cuya solución es $x = a^{-1}b$ y depende de la existencia del número a^{-1}, es decir, depende de la existencia del inverso multiplicativo del número a. De esta manera, en esta sección nos ocupamos de matrices cuadradas que poseen inversa; para ello establecemos las condiciones para calcular la inversa, sus propiedades y un algoritmo para determinarla.

Definición 1.9.1. *Una matriz A de tamaño $n \times n$ es invertible o no singular si existe otra matriz B, de tamaño $n \times n$ tal que $AB = BA = I_n$. Si tal matriz existe se dice que es la inversa de A y la denotamos con A^{-1}. Si A no posee inversa se dice que es singular.*

Ejemplo 1.9.1. Sean las matrices

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Como se puede verificar, $AB = BA = I_n$; entonces, A tiene inversa y $B = A^{-1}$.

El siguiente teorema resume las principales propiedades de la inversa de una matriz.

Teorema 1.9.1. *Si A y B son matrices de tamaño $n \times n$, entonces:*

1. *Si A es una matriz no singular, la inversa de A es única.*

2. *Si A es una matriz no singular, A^{-1} es no singular y $(A^{-1})^{-1} = A$.*

3. *Si A y B son matrices de tamaño $n \times n$ no singulares, AB es no singular y $(AB)^{-1} = B^{-1}A^{-1}$.*

4. *Si A es una matriz no singular, A^t es no singular y, además, $(A^t)^{-1} = (A^{-1})^t$.*

Demostración. La prueba de la primera propiedad se deja como ejercicio.

Para la segunda propiedad se tiene que A es no singular; entonces,

$$AA^{-1} = A^{-1}A = I_n,$$

lo cual significa también que A^{-1} es no singular y, además,

$$(A^{-1})^{-1} = A.$$

En el caso de la propiedad tres:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n,$$

con lo cual se prueba que AB es no singular y que

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Para la cuarta propiedad, tenemos que al ser A no singular, entonces

$$AA^{-1} = A^{-1}A = I_n.$$

Calculando la transpuesta en ambos lados se obtiene

$$(AA^{-1})^t = (A^{-1}A)^t = I_n;$$

luego, aplicando las propiedades de la transpuesta,

$$(A^{-1})^t A^t = A^t(A^{-1})^t = I_n,$$

con lo cual

$$(A^t)^{-1} = (A^{-1})^t.$$

La cuarta propiedad puede escribirse de forma general como

$$(A_1A_2 \ldots A_n)^{-1} = A_n^{-1} \ldots A_2^{-1}A_1^{-1}.$$

Ejemplo 1.9.2. Dada la matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

¿qué condición deben cumplir los elementos de A para que A sea invertible? ¿Cuál es la expresión para A^{-1}?

Solución

Supongamos que

$$A^{-1} = \begin{bmatrix} x & z \\ y & w \end{bmatrix};$$

entonces, debe cumplirse que:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x & z \\ y & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

lo cual es equivalente a

$$\begin{bmatrix} ax+by & az+bw \\ cx+dy & cz+dw \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Esta igualdad de matrices representa realmente un sistema de cuatro ecuaciones con cuatro incógnitas, que puede escribirse así:

$$\begin{aligned} ax+by &= 1 \\ cx+dy &= 0 \\ az+bw &= 0 \\ cz+dw &= 1. \end{aligned}$$

Por la forma que presenta el sistema, para su solución, este puede descomponerse en dos sistemas de dos ecuaciones con dos incógnitas cada uno. Estos sistemas son:

$$\begin{aligned} ax+by=1 \qquad & az+bw=0 \\ cx+dy=0 \qquad & cz+dw=1. \end{aligned}$$

Para resolver estos sistemas podemos utilizar el método de eliminación de Gauss-Jordan. Lo que se debe hacer es llevar las matrices aumentadas de cada sistema a la forma escalonada reducida, esto es, debemos escalonar las matrices

$$\left[\begin{array}{cc|c} a & b & 1 \\ c & d & 0 \end{array}\right] \qquad \left[\begin{array}{cc|c} a & b & 0 \\ c & d & 1 \end{array}\right].$$

Como la matriz de coeficientes de cada sistema es la misma, las operaciones elementales para escalonar dichas matrices también son las mismas. De esta manera, podemos realizar simultáneamente estas operaciones sobre la última columna de cada matriz aumentada por separado. Esto puede lograrse llevando la matriz

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right]$$

a la forma escalonada reducida.

En primer lugar, notemos que a y c no pueden ser nulos a la vez. Por comodidad, supongamos que $a \neq 0$:

$$\xrightarrow[\frac{1}{a}F_1]{} \left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ c & d & 0 & 1 \end{array}\right] \xrightarrow[F_2 - cF_1]{} \left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & \frac{da-bc}{a} & -\frac{c}{a} & 1 \end{array}\right].$$

Aquí se nota que para poder resolver los sistemas se requiere que la expresión

$$\Delta = da - bc \neq 0,$$

con lo cual

$$\xrightarrow[\frac{a}{\Delta}F_2]{} \left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & -\frac{c}{\Delta} & \frac{a}{\Delta} \end{array}\right] \xrightarrow[F_1 - \frac{b}{a}F_2]{} \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{a} + \frac{bc}{a\Delta} & -\frac{b}{\Delta} \\ 0 & 1 & -\frac{c}{\Delta} & \frac{a}{\Delta} \end{array}\right].$$

Pero como

$$\frac{1}{a} + \frac{bc}{a\Delta} = \frac{\Delta + bc}{a\Delta} = \frac{da - bc + bc}{a\Delta} = \frac{d}{\Delta},$$

entonces, la forma escalonada reducida de la matriz original es

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{\Delta} & -\frac{b}{\Delta} \\ 0 & 1 & -\frac{c}{\Delta} & \frac{a}{\Delta} \end{array}\right];$$

esto nos indica que

$$A^{-1} = \begin{bmatrix} \frac{d}{\Delta} & -\frac{b}{\Delta} \\ -\frac{c}{\Delta} & \frac{a}{\Delta} \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

siempre y cuando

$$\Delta = da - bc \neq 0.$$

Este ejemplo plantea que la inversa de una matriz de tamaño 2×2 se obtiene intercambiando los elementos de la diagonal, multiplicando los elementos que están fuera de ella por -1 y multiplicando esta nueva matriz por el escalar $1/\Delta$.

El ejemplo anterior nos muestra que el cálculo de la inversa de una matriz está estrechamente relacionado con el problema de la solución de sistemas de ecuaciones. De hecho, si cierto sistema tiene infinitas soluciones o ninguna, ello determina que la matriz no sea invertible. En el siguiente teorema generalizamos el proceso realizado en el ejemplo anterior y damos un criterio para decidir cuándo una matriz dada es invertible.

Teorema 1.9.2. *A es una matriz de tamaño $n \times n$ equivalente a I_n si y solo si A es no singular.*

Demostración. Solo vamos a probar una de las implicaciones del teorema. Consideremos la matriz $M = [A|I_n]$ de tamaño $n \times 2n$. Como A es equivalente por filas a la matriz identidad, entonces, después de un número finito de operaciones elementales por filas, la matriz M es equivalente por filas a la matriz $[I_n|B]$. Al eliminar todas las columnas de la parte derecha de la matriz $M = [A|I_n]$, exceptuando la columna de la posición j, obtenemos la matriz $M = [A|e_j]$, donde

$$e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}_{n \times 1}, \quad \text{con 1 en la posición } j.$$

De la misma secuencia de operaciones elementales por fila obtenemos que la matriz $M = [A|e_j]$ es equivalente a la matriz $[I_n|B_j]$, donde B_j es la columna de la posición j de la matriz B.

A partir de los resultados anteriores concluimos que B_j es la solución del sistema de ecuaciones $AX = [e_j]$ y que, además,

$$AB = [AB_1|AB_2| \dots AB_n] = [e_1|e_2| \dots e_n] = I_n.$$

De la misma manera se prueba que $BA = I_n$ y concluimos que A es no singular y que su inversa es B.

La prueba del teorema anterior nos proporciona un método constructivo para determinar la inversa de una matriz no singular. Este algoritmo puede sintetizarse así:

Paso 1. Se construye la matriz $[A|I_n]$ de tamaño $n \times 2n$.

Paso 2. Por medio de operaciones elementales por filas, se lleva la matriz $[A|I_n]$ a la matriz $[C|D]$ en forma escalonada reducida.

1. Si $C = I_n$, entonces A es no singular y $A^{-1} = D$.
2. Si $C \neq I_n$, entonces A es singular.

Ejemplo 1.9.3. Dada la matriz

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

determinar si es o no invertible.

Solución

Iniciamos el procedimiento construyendo la matriz $[A|I_3]$ de tamaño 3×6 :

$$[A|I_3] = \left[\begin{array}{rrr|rrr} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 1 \end{array}\right].$$

Luego determinamos la forma escalonada reducida, mediante la aplicación de operaciones elementales por filas.

$$\left[\begin{array}{rrr|rrr} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow[F_1 \leftrightarrow F_3]{} \left[\begin{array}{rrr|rrr} -1 & 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[-F_1]{} \left[\begin{array}{rrr|rrr} 1 & 0 & -2 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \end{array}\right] \xrightarrow[F_3 - 2F_2]{} \left[\begin{array}{rrr|rrr} 1 & 0 & -2 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 4 & 1 & 0 & 2 \end{array}\right]$$

$$\xrightarrow[F_3-F_2]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & 2 \end{array}\right] \xrightarrow[\frac{1}{3}F_3]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{array}\right]$$

$$\xrightarrow[F_1+2F_3]{F_2-F_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 1 & 0 & -\frac{1}{3} & \frac{4}{3} & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{array}\right].$$

Como en este caso $C = I_3$, entonces concluimos que A es no singular y, además, $D = A^{-1}$; es decir:

$$A^{-1} = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{4}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Ejemplo 1.9.4. Dada la matriz

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

determinar si es o no singular.

Solución

Iniciamos el algoritmo construyendo la matriz $[A|I_3]$ de tamaño 3×6 :

$$[A|I_3] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ -1 & 0 & 2 & 0 & 1 & 0 \\ -1 & 2 & 7 & 0 & 0 & 1 \end{array}\right].$$

Luego pasamos a determinar la forma escalonada reducida de la matriz, mediante la aplicación de operaciones elementales por filas.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ -1 & 0 & 2 & 0 & 1 & 0 \\ -1 & 2 & 7 & 0 & 0 & 1 \end{array}\right] \xrightarrow[F_2+F_1]{F_3+F_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 5 & 1 & 1 & 0 \\ 0 & 4 & 10 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[F_3-2F_2]{} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 & -2 & 1 \end{array}\right] \xrightarrow[-F_3]{} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & -1 \end{array}\right]$$

$$\xrightarrow[F_1-F_2]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 2 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & -1 \end{array}\right] \xrightarrow[\frac{1}{2}F_2]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 1 & \frac{5}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & 2 & -1 \end{array}\right]$$

$$\xrightarrow[F_2-\frac{1}{2}F_3]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 1 & \frac{5}{2} & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 2 & -1 \end{array}\right].$$

Como en este caso

$$C = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & \frac{5}{2} \\ 0 & 0 & 0 \end{bmatrix} \neq I_3,$$

entonces concluimos que A es singular. Es importante aclarar que aun antes de llegar a la forma escalonada reducida, ya es posible saber que la matriz es singular. Esto se ve desde que en el proceso de escalonamiento arribamos a la matriz

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 & -2 & 1 \end{array}\right],$$

en la cual se ve que es imposible llegar a I_n en la parte izquierda, puesto que ya la fila última es nula.

Ejemplo 1.9.5 (Uso de computador). Determine si las siguientes matrices son o no invertibles:

$$A=\begin{bmatrix} 1 & 3 & 4 \\ 2 & 3 & -1 \\ 0 & 1 & 3 \end{bmatrix},\ B=\begin{bmatrix} 2 & 1 & 3 & 2 \\ 3 & 0 & 1 & -2 \\ 1 & -1 & 4 & 3 \\ 2 & 2 & -1 & 1 \end{bmatrix} \text{ y } C = \begin{bmatrix} -1 & 4 & 3 & 0 \\ 0 & 0 & 1 & 2 \\ 2 & 4 & 5 & 2 \\ 1 & 8 & 8 & 2 \end{bmatrix}.$$

En caso de que alguna de ellas posea inversa, calcularla.

Solución

Como se explicó en el ejemplo 1.5.6 sobre edición de matrices con DERIVE, usando los comandos AUTOR MATRIZ O AUTOR definimos las matrices.

Un primer cálculo que puede realizarse es el de verificar si cada matriz es invertible, lo cual puede llevarse a cabo determinando la forma escalonada reducida de la matriz dada.

En DERIVE, el cálculo de la forma escalonada reducida puede realizarse de manera automática usando el comando ROW_REDUCE(M), el cual genera la forma escalonada reducida de la matriz M. En nuestro caso tenemos:

$$\begin{array}{ll} \#4 & \textit{ROW\_REDUCE(A)} \\ \#5 & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{array}$$

De lo anterior concluimos que el rango de la matriz A es 3 y, en consecuencia, se puede calcular su inversa.

Para calcular la inversa de la matriz en AUTOR, editamos la expresión $A^{\wedge(-1)}$, con lo cual conseguimos, luego de simplificar:

$$\begin{array}{ll} \#6 & A^{-1} \\ \#7 & \begin{bmatrix} -\frac{7}{3} & \frac{5}{3} & \frac{11}{3} \\ 2 & -1 & -3 \\ -\frac{2}{3} & \frac{1}{3} & \frac{4}{3} \end{bmatrix}. \end{array}$$

Es posible también determinar la inversa de una matriz A calculando la forma escalonada de la matriz aumentada $A|I_n$, lo cual se puede realizar en DERIVE con la instrucción *Row_Reduce*(A, *Identity_Matrix(3))*. Esto permite obtener la matriz $A|I_n$ y luego su forma escalonada reducida. El resultado que se obtiene es:

#8 *Row_Reduce*$(A,$ *Identity_Matrix(3))*

#9 $$\begin{bmatrix} 1 & 0 & 0 & -\frac{7}{3} & \frac{5}{3} & \frac{11}{3} \\ 0 & 1 & 0 & 2 & -1 & -3 \\ 0 & 0 & 1 & -\frac{2}{3} & \frac{1}{3} & \frac{4}{3} \end{bmatrix}$$

Realizando estos cálculos para las matrices B y C, obtenemos:

#10 *Row_Reduce*$(B,$ *Identity_Matrix(4))*

#11 $$\begin{bmatrix} 1 & 0 & 0 & 0 & -\frac{21}{55} & \frac{14}{55} & \frac{17}{55} & \frac{19}{55} \\ 0 & 1 & 0 & 0 & \frac{4}{5} & -\frac{1}{5} & -\frac{3}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & 0 & \frac{29}{55} & -\frac{1}{55} & -\frac{13}{55} & -\frac{21}{55} \\ 0 & 0 & 0 & 1 & -\frac{17}{55} & -\frac{7}{55} & \frac{19}{55} & \frac{18}{55} \end{bmatrix}$$

#12 *Row_Reduce*$(C,$ *Identity_Matrix(4))*

#13 $$\begin{bmatrix} 1 & 0 & 0 & -\frac{2}{3} & 0 & -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 & -\frac{5}{3} & 0 & -\frac{11}{12} & -\frac{1}{12} & \frac{1}{6} \\ 0 & 0 & 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -1 \end{bmatrix}.$$

De los resultados obtenidos podemos concluir que el rango de la matriz B es 4 y, por lo tanto, posee inversa, mientras que el rango de la matriz C es 3 y, en consecuencia, no tiene inversa. La matriz inversa de B es:

$$\begin{bmatrix} -\frac{21}{55} & \frac{14}{55} & \frac{17}{55} & \frac{19}{55} \\ \frac{4}{5} & -\frac{1}{5} & -\frac{3}{5} & -\frac{1}{5} \\ \frac{29}{55} & -\frac{1}{55} & -\frac{13}{55} & -\frac{21}{55} \\ -\frac{17}{55} & -\frac{7}{55} & \frac{19}{55} & \frac{18}{55} \end{bmatrix}.$$

Si queremos calcular directamente la inversa de C, que no existe, DERIVE no ejecuta ninguna operación y devuelve en la pantalla la operación indicada sin realizar cálculos.

Cuando queremos resolver un sistema lineal de n ecuaciones con n incógnitas, recordemos que su forma matricial es $AX = B$. Si la matriz A de coeficientes del sistema es no singular, entonces al multiplicar ambos lados de la ecuación $AX = B$ por A^{-1}, preservando el orden de multiplicación, obtenemos:

$$\begin{aligned} A^{-1}(AX) &= A^{-1}B \\ (A^{-1}A)X &= A^{-1}B \\ I_nX &= A^{-1}B \\ X &= A^{-1}B, \end{aligned}$$

lo cual significa que el sistema tiene solución única.

Ejemplo 1.9.6. Usar la inversa para resolver el sistema

$$\begin{aligned} 2x + y &= 6 \\ y + z &= 9 \\ -x + 2z &= 10. \end{aligned}$$

Solución

La forma matricial de este sistema es $AX = B$, donde:

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 6 \\ 9 \\ 10 \end{bmatrix} \quad \text{y} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Usando el resultado del ejemplo 1.9.3, la solución del sistema es:

$$X = A^{-1}B = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ -1 & 4 & -2 \\ 1 & -1 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ 9 \\ 10 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 4 \\ 10 \\ 17 \end{bmatrix}$$

y, por lo tanto, el valor de cada incógnita es

$$x = \frac{4}{3},\ y = \frac{10}{3} \text{ y } z = \frac{17}{3}.$$

Ejemplo 1.9.7. Resolver el sistema de ecuaciones

$$\begin{aligned} x - y &= 4 \\ 2x + y &= 26. \end{aligned}$$

Solución

En este caso, la forma matricial del problema es

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 26 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Para calcular la inversa, utilizamos el método explicado en los ejemplos anteriores.

$$\left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right] \xrightarrow[F_2-2F_1]{} \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -2 & 1 \end{array}\right]$$

$$\xrightarrow[\frac{1}{3}F_2]{} \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{2}{3} & \frac{1}{3} \end{array}\right] \xrightarrow[F_1+F_2]{} \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & -\frac{2}{3} & \frac{1}{3} \end{array}\right].$$

Esto nos indica que

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

y, en consecuencia, la solución del sistema es:

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}B = \frac{1}{3} \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 26 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \end{bmatrix}.$$

Teorema 1.9.3. *Si A es una matriz de tamaño $n \times n$ no singular, entonces el sistema lineal homogéneo $AX = 0$, de n ecuaciones con n incógnitas, tiene una solución no trivial si y solo si A es singular.*

Demostración. Sea A una matriz de tamaño $n \times n$.

Supongamos que el sistema lineal homogéneo $AX = 0$, de n ecuaciones con n incógnitas, tiene una solución no trivial. Si A es no singular, existe A^{-1} y, en consecuencia,

$$\begin{aligned} A^{-1}(AX) &= A^{-1}0 \\ (A^{-1}A)X &= 0 \\ I_nX &= 0 \\ X &= 0. \end{aligned}$$

O sea que la única solución del sistema $AX = 0$ es $X = 0$, lo cual es absurdo.

Supongamos que A es singular. Uno de los teoremas anteriores nos muestra que A no es equivalente por filas a la matriz identidad y, por lo tanto, $\rho(A) < n$. Entonces, el sistema tiene infinitas soluciones y, en consecuencia, posee soluciones no triviales.

Corolario 1.9.1. *Sea A una matriz de tamaño $n \times n$. El sistema homogéneo $AX = 0$ tiene solución única si y solo si A es no singular.*

Ejemplo 1.9.8. Dado el sistema de ecuaciones

$$\begin{aligned} x + 2y &= 0 \\ 2x + 4y &= 0, \end{aligned}$$

analice si tiene solución única utilizando el criterio de la inversa de la matriz de coeficientes. En caso de que tenga infinitas soluciones, determine una pareja de valores de x y y que sean una solución particular.

Solución

En este caso, la forma matricial del problema es

$$A=\begin{bmatrix} 1 & 2\\ 2 & 4\end{bmatrix},\quad B=\begin{bmatrix} 0\\ 0\end{bmatrix},\quad X=\begin{bmatrix} x\\ y\end{bmatrix}.$$

Probemos que A es singular. Aplicando operaciones elementales por filas, se tiene:

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0\\ 2 & 4 & 0 & 1\end{array}\right] \xrightarrow[F_2-2F_1]{} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0\\ 0 & 0 & -2 & 1\end{array}\right]$$

$$\xrightarrow[-\frac{1}{2}F_2]{} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0\\ 0 & 0 & 1 & -\frac{1}{2}\end{array}\right] \xrightarrow[F_1-F_2]{} \left[\begin{array}{cc|cc} 1 & 2 & 0 & \frac{1}{2}\\ 0 & 0 & 1 & -\frac{1}{2}\end{array}\right].$$

Esta última matriz está en la forma escalonada reducida y como

$$\begin{bmatrix} 1 & 2\\ 0 & 0\end{bmatrix} \neq I_n,$$

se concluye que A es singular y, por lo tanto, el sistema tiene una solución no trivial.

La solución general de este sistema es de la forma $x=-2t$, $y=t$. Una posible pareja para la solución es $x=4$ y $y=-2$.

Para sistemas no homogéneos de la forma $AX=B$, donde A es una matriz de tamaño $n\times n$, es posible establecer teoremas semejantes al anterior.

Teorema 1.9.4. *Sea A una matriz de tamaño $n\times n$. Entonces, A es no singular si y solo si el sistema lineal $AX=B$ tiene una solución única para cada matriz B de tamaño $n\times 1$.*

Ejemplo 1.9.9. Analizar el sistema de ecuaciones

$$\begin{aligned} x-y &= 4\\ 2x+y &= 26.\end{aligned}$$

Solución

En este caso, la forma matricial del problema es

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 26 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Como el sistema tiene solución única, la matriz A es no singular.

Los resultados establecidos hasta el momento, para matrices y sistemas $n \times n$, pueden resumirse en el siguiente teorema.

Teorema 1.9.5. *Las siguientes afirmaciones son equivalentes:*

1. *A es no singular.*
2. *El sistema lineal $AX = 0$ tiene solución única $X = 0$.*
3. *El sistema lineal $AX = B$ tiene solución única para cada matriz B de tamaño $n \times 1$.*
4. *A es equivalente por filas a I_n.*
5. *$\rho(A) = n$.*

1.10 Ejercicios resueltos

Ejercicio 1.10.1. Mostrar que toda matriz cuadrada de orden $n \times n$ se puede escribir de manera única como la suma de una matriz simétrica y otra antisimétrica.

Solución

Sea A una matriz cuadrada. Es claro que

$$A = \frac{A + A^t}{2} + \frac{A - A^t}{2}.$$

Si definimos las matrices

$$B = \frac{A + A^t}{2} \quad \text{y} \quad C = \frac{A - A^t}{2},$$

se tiene que B es simétrica y C es antisimétrica, puesto que

$$\begin{aligned} B^t &= \left(\frac{A+A^t}{2}\right)^t = \frac{(A^t+(A^t)^t)}{2} = \frac{A^t+A}{2} = \frac{A+A^t}{2} = B \\ C^t &= \left(\frac{A-A^t}{2}\right)^t = \frac{A^t-(A^t)^t}{2} = \frac{A^t-A}{2} = -\frac{A-A^t}{2} = -C. \end{aligned}$$

Supongamos ahora que A se puede escribir de manera diferente como la suma de una matriz simétrica D y otra antisimétrica E; entonces:

$$A = B + C = D + E,$$

o lo que es lo mismo:

$$B - D = E - C.$$

Pero la matriz $B - D$ es simétrica, mientras que la matriz $E - C$ es antisimétrica; en consecuencia, $B - D$ y $E - C$ son simultáneamente simétricas y antisimétricas.

Por otro lado, se sabe que la única matriz que es simétrica y antisimétrica a la vez es la matriz nula; por lo tanto, $B - D = 0$ y $E - C = 0$, o de forma equivalente, $B = D$ y $E = C$.

Ejercicio 1.10.2. Una tienda de helados vende solo helados con soda y malteadas. Un helado con soda se prepara con 1 onza de jarabe y 4 onzas de helado, mientras que una malteada se prepara con 1 onza de jarabe y 3 onzas de helado. Si la tienda usa 4 galones de helado y 5/4 de jarabe en un día, ¿cuántos helados con soda y cuántas malteadas vende? (Un galón equivale a 128 onzas).

Solución

Sea x el número de helados con soda y y el número de malteadas. Las restricciones del problema se pueden escribir como:

$$\begin{aligned} 1x + 1y &= 160 && \text{Restricción para las onzas de jarabe} \\ 4x + 3y &= 512 && \text{Restricción para las onzas de helado.} \end{aligned}$$

La solución de este sistema se puede realizar fácilmente utilizando el proceso de eliminación de Gauss-Jordan, es decir, llevando la matriz aumentada del sistema a la forma escalonada reducida.

$$\left[\begin{array}{cc|c} 1 & 1 & 160 \\ 4 & 3 & 512 \end{array}\right] \xrightarrow[F_2-4F_1]{} \left[\begin{array}{cc|c} 1 & 1 & 160 \\ 0 & -1 & -128 \end{array}\right] \xrightarrow[-F_2]{} \left[\begin{array}{cc|c} 1 & 1 & 160 \\ 0 & 1 & 128 \end{array}\right]$$

$$\xrightarrow[F_1-F_2]{} \left[\begin{array}{cc|c} 1 & 0 & 32 \\ 0 & 1 & 128 \end{array}\right].$$

Por lo tanto, la solución del sistema es $x = 32$ helados de soda y $y = 128$ malteadas.

Ejercicio 1.10.3. Sean A, B y C matrices. Determinar si los siguientes enunciados son verdaderos o falsos, justificando la respuesta:

1. Si $AB = 0$, entonces $A = 0$ o $B = 0$.
2. Si $AB = AC$, entonces $B = C$, siempre y cuando el producto tenga sentido.

Solución

1. Falso, ya que basta escoger dos matrices no nulas A y B, tales que $AB = 0$. En este caso:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix},$$

mientras que su producto es la matriz nula, esto es:

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

2. Falso, puesto que si tomamos $A = 0$, y B y C como dos matrices cualquiera distintas, obtenemos que $AB = AC = 0$ y, sin embargo, $B \neq C$.

Una consecuencia de estos enunciados es que, dado un sistema de m ecuaciones lineales con n incógnitas, si $AX = 0$, no podemos concluir que $A = 0$ o $X = 0$.

Ejercicio 1.10.4 (Interés compuesto). Calcular la cantidad de dinero que se tiene al cabo de n años si invertimos \$100 a un interés de 4, 6 y 7 %.

Solución

Si colocamos P pesos durante un año a un interés r, entonces el valor que se tiene al final del año es

$$P_f = P + rP = (1 + r)P.$$

El producto

$$AB = \begin{bmatrix} 1,04 & 0 & 0 \\ 0 & 1,06 & 0 \\ 0 & 0 & 1,07 \end{bmatrix} \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 104 \\ 106 \\ 107 \end{bmatrix}$$

da la cantidad de dinero que se tiene al cabo de un año si invertimos \$100 a un interés de 4, 6 y 7 % respectivamente. Al final del segundo período, el monto estará dado por

$$\begin{aligned} C_2 &= \begin{bmatrix} 1,04 & 0 & 0 \\ 0 & 1,06 & 0 \\ 0 & 0 & 1,07 \end{bmatrix} \begin{bmatrix} 104 \\ 106 \\ 107 \end{bmatrix} \\ &= \begin{bmatrix} 1,04 & 0 & 0 \\ 0 & 1,06 & 0 \\ 0 & 0 & 1,07 \end{bmatrix} \begin{bmatrix} 1,04 & 0 & 0 \\ 0 & 1,06 & 0 \\ 0 & 0 & 1,07 \end{bmatrix} \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix} \\ &= \begin{bmatrix} 1,04 & 0 & 0 \\ 0 & 1,06 & 0 \\ 0 & 0 & 1,07 \end{bmatrix}^2 \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix}. \end{aligned}$$

En general, el monto al final de n años estará dado por:

$$C_f = \begin{bmatrix} 1,04 & 0 & 0 \\ 0 & 1,06 & 0 \\ 0 & 0 & 1,07 \end{bmatrix}^n \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix}.$$

Ejercicio 1.10.5. Calcule la inversa de la matriz diagonal

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

Solución

Si aplicamos el método de escalonamiento de Gauss-Jordan, obtenemos:

$$\left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 4 & 0 & 0 & 1 \end{array}\right] \xrightarrow[1/2F_1;1/3F_2]{1/4F_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{4} \end{array}\right]$$

y, por lo tanto,

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

que corresponde a un matriz diagonal cuyos elementos son los inversos de los números de la diagonal de la matriz inicial.

Ejercicio 1.10.6. Determine si la siguiente matriz es o no singular:

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

Solución
Para responder esto, basta con analizar la forma escalonada reducida de la matriz $[A|I_3]$:

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 0 & 1 & 0 \\ 4 & 6 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[F_2+2F_1]{F_3-4F_1} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 6 & 1 & -4 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[F_2\leftrightarrow F_3]{} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 6 & 1 & -4 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 & 0 \end{array}\right] \xrightarrow[\frac{1}{6}F_2]{} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{6} & -\frac{4}{6} & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 2 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[\frac{1}{2}F_3]{} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{6} & -\frac{4}{6} & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 1 & \frac{1}{2} & 0 \end{array}\right] \xrightarrow[F_2+\frac{4}{6}F_3]{F_1-F_3} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 1 & \frac{1}{2} & 0 \end{array}\right].$$

Esta última matriz está en la forma escalonada reducida; pero en la parte de la izquierda no se obtuvo la matriz identidad; luego, A^{-1} no existe, es decir, A es singular.

Ejercicio 1.10.7. Sea A una matriz $m \times n$. Probar que:

1. AA^t es una matriz simétrica de tamaño $m \times m$.

2. A^tA es una matriz simétrica de tamaño $n \times n$.

Solución
Para el primer caso:

$$(AA^t)^t = (A^t)^t A^t = AA^t;$$

luego AA^t es simétrica. El segundo caso se demuestra en forma similar.

Ejercicio 1.10.8. Encuentre los números a y b tales que la matriz

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

sea simétrica.

Solución

Para que la matriz A sea simétrica se debe cumplir que $A = A^t$; esto es:

$$\begin{bmatrix} 1 & 2 & a \\ b & 3 & -1 \\ 4a & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & b & 4a \\ 2 & 3 & -1 \\ a & -1 & 0 \end{bmatrix},$$

a partir de lo cual se concluye que $b = 2$ y $4a = a$, con lo cual $b = 2$ y $3a = 0$. En consecuencia, $b = 2$ y $a = 0$.

Ejercicio 1.10.9. Determine la forma escalonada reducida de la matriz

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

y establezca su inversa en caso de que esta exista.

Solución

Como en definitiva se trata de determinar la inversa de A, aplicamos el procedimiento usual para calcular la inversa, es decir, analizamos la forma escalonada reducida de la matriz $[A|I_3]$:

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

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

$$\xrightarrow[F_2\leftrightarrow F_3]{} \left[\begin{array}{rrr|rrr} 1 & -3 & -1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 6 & 6 & 1 & 2 & 0 \end{array}\right] \xrightarrow[F_3-6F_2]{F_1+3F_2} \left[\begin{array}{rrr|rrr} 1 & 0 & 5 & 0 & -1 & 3 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & -6 & 1 & 2 & -6 \end{array}\right]$$

$$\xrightarrow[-1/6F_3]{} \left[\begin{array}{rrr|rrr} 1 & 0 & 5 & 0 & -1 & 3 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\frac{1}{6} & -\frac{1}{3} & 1 \end{array}\right] \xrightarrow[F_1-5F_3]{F_2-2F_3} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{5}{6} & \frac{2}{3} & -2 \\ 0 & 1 & 0 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 1 & -\frac{1}{6} & -\frac{1}{3} & 1 \end{array}\right].$$

Por lo tanto, la forma escalonada reducida de la matriz A es

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

por lo que la matriz es invertible y, además,

$$A^{-1} = \begin{bmatrix} \frac{5}{6} & \frac{2}{3} & -2 \\ \frac{1}{3} & \frac{2}{3} & -1 \\ -\frac{1}{6} & -\frac{1}{3} & 1 \end{bmatrix}.$$

Ejercicio 1.10.10. Responder: ¿para qué valores de k el rango de la matriz

$$A = \begin{bmatrix} 1 & 1 & -1 & 3 \\ 1 & -1 & 3 & 4 \\ 1 & 1 & k^2-10 & k \end{bmatrix}$$

es 1, 2 o 3?

Solución

Llevando la matriz A a la forma escalonada reducida, obtenemos:

$$\begin{bmatrix} 1 & 1 & -1 & 3 \\ 1 & -1 & 3 & 4 \\ 1 & 1 & k^2-10 & k \end{bmatrix} \xrightarrow[F_3-F_1]{F_2-F_1} \begin{bmatrix} 1 & 1 & -1 & 3 \\ 0 & -2 & 4 & 1 \\ 0 & 0 & k^2-9 & k-3 \end{bmatrix}$$

$$\xrightarrow[-\frac{1}{2}F_1]{} \begin{bmatrix} 1 & 1 & -1 & 3 \\ 0 & 1 & -2 & -\frac{1}{2} \\ 0 & 0 & k^2-9 & k-3 \end{bmatrix} \xrightarrow[F_1-F_2]{} \begin{bmatrix} 1 & 0 & 1 & \frac{7}{2} \\ 0 & 1 & -2 & -\frac{1}{2} \\ 0 & 0 & k^2-9 & k-3 \end{bmatrix}.$$

Aun cuando la matriz no está en la forma escalonada reducida, es posible hacer las siguientes afirmaciones:

1. En la última matriz hay por lo menos dos pivotes, con lo cual el rango es, como mínimo, 2. Por lo tanto, es imposible que el rango sea 1.
2. Para que el rango sea 2 se debe tener que la última fila sea nula, lo cual sucede cuando $k = 3$.
3. Si $k = -3$, $k^2 - 9 = 0$; pero $k - 3 = -6$ y, por lo tanto, el rango de A es 3. En este caso el elemento de la posición $(3, 4)$ se puede convertir en un 1 pivote y con él eliminar los demás elementos de la cuarta columna, para obtener la forma escalonada reducida.

En el siguiente ejercicio se presenta un problema similar, pero esta vez desde el punto de vista de los sistemas de ecuaciones no homogéneos.

Ejercicio 1.10.11. Determine todos los valores de k para los cuales el sistema de ecuaciones lineales

$$\begin{aligned} x+y-z &= 3 \\ x-y+3z &= 4 \\ x+y+(k^2-10)z &= k: \end{aligned}$$

1. No tenga solución.
2. Tenga solución única.
3. Tenga infinitas soluciones.

Solución

La matriz aumentada del sistema es:

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 3 \\ 1 & -1 & 3 & 4 \\ 1 & 1 & k^2-10 & k \end{array}\right].$$

Usando los resultados del ejercicio anterior, esta puede llevarse hasta la matriz

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & \frac{7}{2} \\ 0 & 1 & -2 & -\frac{1}{2} \\ 0 & 0 & k^2-9 & k-3 \end{array}\right].$$

De esta matriz podemos hacer las siguientes afirmaciones:

1. Si $k=-3$, continuando con el proceso de escalonamiento, se tiene:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & \frac{7}{2} \\ 0 & 1 & -2 & -\frac{1}{2} \\ 0 & 0 & 0 & -6 \end{array}\right] \xrightarrow[-\frac{1}{6}F_3]{} \left[\begin{array}{ccc|c} 1 & 0 & 1 & \frac{7}{2} \\ 0 & 1 & -2 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[F_1-\frac{7}{2}F_3]{F_2+\frac{1}{2}F_3} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right],$$

y en este caso $\rho(A|B)=3$, mientras que $\rho(A)=2$, o sea que $\rho(A|B)>\rho(A)$ y, por lo tanto, el sistema no tiene solución cuando $k=-3$.

2. Si $k=3$, entonces tenemos la matriz

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & \frac{7}{2} \\ 0 & 1 & -2 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right],$$

la cual ya está en la forma escalonada reducida y en la que se puede observar que $\rho(A|B) = \rho(A) = 2 < n$, donde n es el número de incógnitas, en cuyo caso el sistema tiene infinitas soluciones.

3. Si $k \neq 3$ y $k \neq -3$, el elemento que ocupa la posición $(3,3)$, es decir, $k^2 - 9$, es no nulo y, por lo tanto, $\rho(A|B) = \rho(A) = 3$; en consecuencia, el sistema tiene solución única.

Es importante recalcar que solo podemos pensar en transformar el término $k^2 - 9$ en un pivote cuando se garantice que no hay divisiones por cero.

Ejercicio 1.10.12. Dos compañías rivales M y N fabrican cierto producto. Mientras cada año la compañía M conserva $1/3$ de sus clientes, y $2/3$ de sus consumidores cambian a N, en ese mismo período N conserva $1/2$ de sus clientes y el mismo número de sus consumidores cambia a M. Esta información puede escribirse en forma matricial como

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix}.$$

Si al comenzar por primera vez la fabricación del producto, M tiene $1/3$ del mercado, mientras que N posee los otros $2/3$,

1. Determine la distribución del mercado un año después.

2. Determine la distribución estable del mercado.

Solución

Si denotamos con X_0 la distribución inicial del mercado, podemos escribir esto como

$$X_0 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix},$$

1. Un año después, la distribución del mercado es

$$X_1 = AX_0 = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{9} \\ \frac{5}{9} \end{bmatrix}.$$

   Esto puede justificarse como sigue: supongamos que el mercado inicial es de h personas y que este número no se modifica con el paso del tiempo. Entonces, inicialmente M tiene $1/3h$ de los clientes, mientras que N tiene $2/3h$. Al final del primer año, M conserva $1/3$ de sus clientes y gana $1/2$ de los consumidores de N. En consecuencia, M tiene

$$\frac{1}{3}\left(\frac{1}{3}\,h\right) + \frac{1}{2}\left(\frac{2}{3}\,h\right) = \frac{1}{9}\,h + \frac{1}{3}\,h = \frac{4}{9}\,h.$$

   De modo similar, al final del primer año, N conserva $1/2$ de sus clientes y gana $2/3$ de los consumidores de M. En consecuencia, N tiene

$$\frac{1}{2}\left(\frac{2}{3}\,h\right) + \frac{2}{3}\left(\frac{1}{3}\,h\right) = \frac{1}{3}\,h + \frac{2}{9}\,h = \frac{5}{9}\,h.$$

2. Sea

$$X_0 = \begin{bmatrix} a \\ b \end{bmatrix},$$

   con $a + b = 1$. Decimos que X_0 es la distribución estable del mercado si $X_0 = AX_0$, es decir, es aquella distribución inicial del mercado que no se modifica de un año al otro.

   Desde el punto de vista matricial, esta condición equivale a la ecuación

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix},$$

lo cual genera el sistema de ecuaciones

$$\begin{aligned}\frac{1}{3}\,a+\frac{1}{2}\,b &= a\\ \frac{2}{3}\,a+\frac{1}{2}\,b &= b.\end{aligned}$$

Este es equivalente a

$$\begin{aligned}-\frac{2}{3}\,a+\frac{1}{2}\,b &= 0\\ \frac{2}{3}\,a-\frac{1}{2}\,b &= 0,\end{aligned}$$

que en realidad son solo una ecuación.

Como se debe cumplir la condición de que $a+b=1$, entonces $a = 1-b$, con lo cual, al sustituir en cualquiera de las ecuaciones del sistema inicial, obtenemos:

$$-\frac{2}{3}\,(1-b)+\frac{1}{2}\,b=0;$$

de donde concluimos que $b = 4/7$ y por consiguiente $a = 3/7$.

En el caso de la distribución estable, realmente interesa es resolver el sistema de ecuaciones

$$\begin{aligned}-\frac{2}{3}\,a+\frac{1}{2}\,b &= 0\\ a+b &= 1.\end{aligned}$$

Ejercicio 1.10.13. Determine una matriz X de tamaño 3×1 cuyas entradas no sean todas nulas, tal que $AX = 3X$, donde

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

Solución

La ecuación $AX = 3X$ puede escribirse en la forma $AX - 3X = 0$ o también $(A - 3I_3)X = 0$, lo cual quiere decir que nos interesa resolver el sistema homogéneo

$$\left(\begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{bmatrix} - \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}\right)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

que se simplifica como

$$\begin{bmatrix} -2 & 2 & -1 \\ 1 & -3 & 1 \\ 4 & -4 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Al aplicar el método de eliminación de Gauss-Jordan a la matriz aumentada de este sistema se tiene:

$$\left[\begin{array}{rrr|r} -2 & 2 & -1 & 0 \\ 1 & -3 & 1 & 0 \\ 4 & -4 & 2 & 0 \end{array}\right] \xrightarrow[F_1 \leftrightarrow F_2]{} \left[\begin{array}{rrr|r} 1 & -3 & 1 & 0 \\ -2 & 2 & -1 & 0 \\ 4 & -4 & 2 & 0 \end{array}\right]$$

$$\xrightarrow[F_2+2F_1]{F_3-4F_1} \left[\begin{array}{rrr|r} 1 & -3 & 1 & 0 \\ 0 & -4 & 1 & 0 \\ 0 & 8 & -2 & 0 \end{array}\right]$$

$$\xrightarrow[-\frac{1}{4}F_2]{} \left[\begin{array}{rrr|r} 1 & -3 & 1 & 0 \\ 0 & 1 & -\frac{1}{4} & 0 \\ 0 & 8 & -2 & 0 \end{array}\right] \xrightarrow[F_3-8F_2]{F_1+3F_2} \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{4} & 0 \\ 0 & 1 & -\frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

La última matriz está en la forma escalonada reducida. Como el sistema es homogéneo y el rango de la matriz

$$\left[\begin{array}{rrr|r} -2 & 2 & -1 & 0 \\ 1 & -3 & 1 & 0 \\ 4 & -4 & 2 & 0 \end{array}\right]$$

es 2, entonces el sistema tiene infinitas soluciones. El sistema equivalente puede escribirse en la forma

$$\begin{aligned} x + \frac{1}{4}\, z &= 0 \\ y - \frac{1}{4}\, z &= 0. \end{aligned}$$

Como se puede concluir de la forma escalonada reducida de la matriz original, las dos primeras variables pueden escribirse en términos de la tercera variable (variable libre), obteniendo la relación

$$\begin{aligned} x &= -\frac{1}{4}\, z \\ y &= \frac{1}{4}\, z. \end{aligned}$$

Si tomamos $z = t$, la solución finalmente puede escribirse en la forma

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -\frac{1}{4} \\ \frac{1}{4} \\ 1 \end{bmatrix}.$$

Si le damos a t el valor de $t = 4$, obtenemos como solución particular

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}.$$

En este caso, el conjunto solución del sistema homogéneo dado es el conjunto de los múltiplos del vector $[-1, 1, 4]$.

Ejercicio 1.10.14. Determine si los siguientes enunciados son verdaderos o falsos; justifique la respuesta.

1. Si A y B son matrices no singulares, entonces $A + B$ es no singular.

2. Si A y B son matrices no singulares, entonces $A - B$ es no singular.

3. Si A es no singular, entonces $-A$ es no singular.

Solución

1. Falso. Consideremos, por ejemplo, las matrices
$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{y} \qquad B = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}.$$
Ambas son no singulares, puesto que
$$A^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{y} \qquad B^{-1} = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & -1 \end{bmatrix};$$
y, sin embargo, $A + B = 0_2$, que es singular.

2. Falso. Aquí es suficiente con mostrar que si
$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$
y $B = A$, entonces, ambas son no singulares, mientras que $A - B = 0_2$ es singular.

El último numeral se deja como ejercicio.

Ejercicio 1.10.15. Determinar el polinomio cuadrático que pase por los puntos $(2, 9)$, $(1, 2)$ y $(-1, 0)$. Este polinomio es conocido como el *polinomio interpolador*.
Solución
Si denotamos con $P(x) = ax^2 + bx + c$ al polinomio requerido, se cumple que
$$\begin{aligned} P(2) = 4a + 2b + c &= 9 \\ P(1) = a + b + c &= 2 \\ P(-1) = a - b + c &= 0. \end{aligned}$$

Al resolver, usando Gauss-Jordan, se tiene:

$$\left[\begin{array}{rrr|r} 4 & 2 & 1 & 9 \\ 1 & 1 & 1 & 2 \\ 1 & -1 & 1 & 0 \end{array}\right] \xrightarrow[F_1 \leftrightarrow F_2]{} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 4 & 2 & 1 & 9 \\ 1 & -1 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[F_3 - F_1]{F_2 - 4F_1} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & -2 & -3 & 1 \\ 0 & -2 & 0 & -2 \end{array}\right] \xrightarrow[-\frac{1}{2}F_2]{F_3 \leftrightarrow F_2} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & -2 & -3 & 1 \end{array}\right]$$

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

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

En consecuencia, la solución del sistema es:

$$\begin{aligned} a &= 2 \\ b &= 1 \\ c &= -1, \end{aligned}$$

con lo cual concluimos que el polinomio pedido es $P(x) = 2x^2 + x - 1$.

1.11 Ejercicios propuestos capítulo 1

1. Dadas las matrices

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

$$C = \begin{bmatrix} -2 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 2 & 1 & 1 & -1 \\ 3 & 0 & 1 & -2 \end{bmatrix},$$

realizar, cuando sea posible, las siguientes operaciones:

a) CAB
b) $A^{-1}B$
c) $(B+D)A$
d) CBA
e) $A+B$
f) $A^{-1}C^t$
g) B^tC^t
h) $(A^{-1}C^t)^t$
i) $(B-D)A$
j) $(B-A)D$.

2. Dado el sistema lineal de ecuaciones

$$\begin{aligned} -2x+3y-z &= 1 \\ x+2y-z &= 4 \\ -2x-y+z &= -3, \end{aligned}$$

a) Hallar la matriz de coeficientes.

b) Hallar la matriz aumentada.

c) Hallar la inversa de la matriz de coeficientes.

d) Hallar la solución del sistema usando la inversa.

3. Dado el sistema lineal de ecuaciones

$$\begin{aligned} x_1+3x_2+x_3 &= -2 \\ x_1-2x_2-2x_3 &= 4 \\ x_1-x_2+x_3 &= 0, \end{aligned}$$

a) Hallar la matriz de coeficientes.

b) Hallar la matriz aumentada.

c) Hallar la solución del sistema usando eliminación de Gauss-Jordan.

d) Hallar, si existe, la inversa de la matriz de coeficientes.

e) Hallar, si es posible, la solución del sistema, usando A^{-1}.

4. Resolver los siguientes sistemas de ecuaciones utilizando el método de eliminación de Gauss.

   a)
$$\begin{aligned} x + 5y + 11z &= -5 \\ 2x + 3y + 8z &= 4 \\ -x + 2y + 3z &= -9. \end{aligned}$$

   b)
$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 3 \\ x_1 + 2x_2 + x_3 &= 1 \\ -x_1 + 4x_2 &= -2. \end{aligned}$$

   c)
$$\begin{aligned} 2x + 4y + 4z &= 6 \\ 3x + 6y - z &= 5 \\ x + y + z &= 6. \end{aligned}$$

5. Resolver los siguientes sistemas de ecuaciones utilizando el método de eliminación de Gauss-Jordan.

   a)
$$\begin{aligned} 2x - y - z &= 4 \\ 3x - 2y + 4z &= 11 \\ 6x + 8y - 4z &= 22. \end{aligned}$$

   b)
$$\begin{aligned} x + y - z &= 0 \\ 2x - 4y + 3z &= 0 \\ 5x + 13y - 10z &= 0. \end{aligned}$$

c)

$$\begin{aligned}
-x_1 - 2x_2 - 4x_4 &= 4 \\
3x_2 + 5x_3 + 10x_4 &= 2 \\
4x_2 + 8x_3 + 16x_4 &= 0 \\
x_2 + x_3 + 2x_4 &= 2.
\end{aligned}$$

d)

$$\begin{aligned}
-x_1 - 2x_2 - 4x_4 + x_5 &= 2 \\
x_1 + 3x_2 + 5x_3 + 10x_4 - 3x_5 &= 1 \\
2x_1 + 4x_2 + 8x_3 + 16x_4 + 5x_5 &= 3 \\
-5x_1 - 2x_2 - 4x_4 - x_5 &= 2 \\
4x_1 + x_2 + x_3 + 2x_4 - 3x_5 &= -1.
\end{aligned}$$

6. En una distribuidora de café se venden tres tipos de mezclas. La mezcla especial tiene 300 g de café colombiano y 200 g de café francés. La mezcla refinada se compone de 200 g de café colombiano, 200 g de keniano y 100 g de francés. La mezcla sencilla comprende 100 g de café colombiano, 200 g de keniano y 200 g de francés. Si en la tienda se dispone de 30 kg de café colombiano, 15 kg de keniano y 25 kg de francés, ¿cuántas mezclas de cada tipo se pueden realizar para consumir la totalidad del café?

7. Para el control de cierta enfermedad de una planta se usan tres productos químicos en las siguientes proporciones: 10 unidades del químico A, 12 del químico B y 8 del químico C. Las marcas X, Y y Z son atomizadores comerciales que se venden en el mercado. Un galón de la marca X contiene los químicos A, B y C en las cantidades 1, 2 y 1 unidades, respectivamente. Un galón de la marca Y los contiene en las cantidades 2, 1 y 2 unidades, mientras un galón de la marca Z en las cantidades 3, 2 y 1 unidades, respectivamente. ¿Que cantidad de cada marca debe emplearse para fumigar la planta con las medidas exactas de los químicos requeridos para el control de la enfermedad?

8. En un estanque para la cría de peces se colocan diariamente 400 unidades del alimento A, 600 del B y 600 del C. Los datos del consumo diario de alimento para cada tipo de pez se muestran en la siguiente tabla:

Alimento	**Trucha**	**Carpa**	**Mojarra**
A	1	2	0
B	2	1	1
C	1	1	2

¿Cuantos ejemplares de cada tipo pueden tenerse en el estanque de modo que consuman todo el alimento?

9. Ana y Sebastián desean ir a comprar frutas. Los tipos y las cantidades de frutas que van a comprar se muestran en la siguiente tabla:

	Manzanas	**Peras**	**Bananos**
Ana	6	4	5
Sebastián	4	3	6

En su vecindario existen dos mercados (el Baratón y el Superior) y los precios por unidad son los dados en la siguiente tabla:

	El Baratón	**El Superior**
Manzanas ($)	1.000	1.100
Peras ($)	1.200	1.250
Bananos ($)	150	200

¿Cuánto le costará a Ana y Sebastián realizar sus compras en cada uno de los mercados?

10. Hallar la inversa, si existe, para cada una de las siguientes matrices:

$$A = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad C = \begin{bmatrix} 3 & 3 & 4 \\ 1 & 2 & 1 \\ 1 & 6 & 6 \end{bmatrix}$$

$$D=\begin{bmatrix}1&0&0&0\\0&2&2&0\\0&0&3&0\\0&0&0&4\end{bmatrix},\qquad E=\begin{bmatrix}4&0&1\\2&3&6\\6&-3&-4\end{bmatrix}.$$

11. Determinar todos los valores de λ para los cuales A es singular, donde
$$A=\begin{bmatrix}\lambda+2&-1&3\\2&\lambda-1&2\\0&0&\lambda+4\end{bmatrix}.$$

12. Una empresa fabrica, en su planta de Medellín, tres tipos de ventanas: de 1,40, 2,10 y 2,50 m. Los almacenes principales se encuentran en Bogotá, Cali, Barranquilla y Cartagena. Las ventas durante el año 2006 del almacén de Bogotá se cifraron en 400, 100 y 500 ventanas de 1,40, 2,10 y 2,50 m respectivamente; las del almacén de Cali, en 300, 150 y 400; las del almacén de Barranquilla, en 100, 100 y 200, y las del almacén de Cartagena, en 200, 150 y 300. Los precios de venta de las ventanas en el año 2006 fueron de \$25.000 para las de 1,40 m, de \$50.000 para las de 2,10 m y \$80.000 para las de 2,50 m.

    a) Expresar las ventas de la empresa mediante una matriz A de orden 4×3.

    b) Expresar, mediante un vector columna, el precio de cada tipo de ventana.

13. Si A y B son matrices $n\times n$ y AB es no singular, probar que A y B son matrices no singulares.

14. Hallar los espacios en blanco para que las siguientes matrices sean simétricas.
$$D=\begin{bmatrix}4&2&-3&(\ )\\2&0&1&3\\-3&(\ )&2&1\\4&(\ )&(\ )&6\end{bmatrix},\qquad C=\begin{bmatrix}2&(\ )&4\\6&-6&2\\(\ )&2&4\end{bmatrix}.$$

15. Si A y B son matrices antisimétricas de tamaño $n \times n$, demuestre que $(AB)^t = BA$, de manera que AB es simétrica si y solo si A y B conmutan.

16. Probar que toda matriz A antisimétrica de tamaño $n \times n$ tiene nulos los elementos de la diagonal.

17. Demuestre que si A y B son matrices simétricas de tamaño $n \times n$, entonces $A + B$ también lo es.

18. Demuestre que si A y B son matrices antisimétricas de tamaño $n \times n$, entonces $A + B$ también lo es.

19. Hallar los valores de a, de manera que el rango de la matriz

$$A = \begin{bmatrix} 1 & 1 & -1 & 2 \\ 1 & 2 & 1 & 3 \\ 1 & 1 & a^2 - 5 & a \end{bmatrix},$$

sea 1, 2 o 3.

20. Demuestre que si A es una matriz cuadrada que satisface la ecuación $A^2 - 3A + I = 0$, entonces $A^{-1} = 3I - A$.

21. Determine todos los valores de a para los cuales el sistema

$$\begin{aligned} x + y + z &= 2 \\ 2x + 3y + 2z &= 5 \\ 2x + 3y + (a^2 + 1)z &= a + 4 \end{aligned}$$

   a) No tenga solución.
   b) Tenga solución única.
   c) Tenga infinitas soluciones.

22. Determinar si las siguientes afirmaciones son falsas o verdaderas. Justifique sus respuestas:

   a) ( ) Si A es una matriz de orden $n \times n$, entonces $A - A^T$ es antisimétrica.

b) ( ) Si A es una matriz de orden 3×3 y

$$X = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$$

es solución para $AX = 0$, entonces A es singular.

c) ( ) Si A y B son matrices cuadradas, entonces

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

d) ( ) Si A y B son matrices cuadradas, entonces se cumple que $(2A)(2B) = 4(BA)$.

23. Determine el polinomio cuadrático que pase por los puntos $(2, 2)$, $(4, 3)$ y $(5, 6)$.

24. Hallar la inversa de la matriz

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

25. Dada la matriz

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

a) Calcule $P(A)$, sabiendo que $P(x) = x^2 - 5x + 2I_2$.

b) Calcule $Q(A)$, si $Q(x) = x^2 - x + 2I_2$.

26. Determinar si las siguientes afirmaciones son falsas o verdaderas. Justifique sus respuestas:

a) ( ) $(A^{-1}BC^t)^t = CB^t(A^t)^{-1}$.

b) ( ) Si AB es singular, una de ellas es singular.

c) ( ) La matriz A está en la forma escalonada reducida.

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

d) ( ) Toda matriz de tamaño $n \times n$ tiene inversa.

e) ( ) $((A^{-1})^t)^{-1} = A$.

27. Determine $P(A)$ sabiendo que $P(x) = x^2 - 3x - 14I_2$ y que

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

28. Determine todos los valores de a para los cuales la matriz A tiene inversa, siendo

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

29. Determine una constante k tal que $(kA)^T(kA) = [2]$, donde $A = [1, 2, 0]^t$. ¿Cuántas soluciones existen?

30. Determine una matriz X de tamaño 2×1 cuyas entradas no sean todas nulas, tal que $AX = 4X$, donde A está dada por:

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

31. Sea la matriz

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

a) Calcule $A^2, A^3, \ldots, A^7$.

b) ¿Qué será A^{2001}?

32. Sea la matriz

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

Determine una fórmula para A^n.

33. Dadas las siguientes matrices, utilice DERIVE para determinar las soluciones de la ecuación $AX + B = C$:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & -1 & 2 \\ 0 & 3 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 1 & 2 & 1 \\ -6 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & -1 & 2 \\ 6 & 1 & 0 & 0 \end{bmatrix} \text{ y } X = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \\ x_9 & x_{10} & x_{11} & x_{12} \end{bmatrix}.$$

34. Utilice DERIVE para resolver los siguientes sistemas:

a)
$$\begin{aligned} 2x_1 - x_2 + 3x_3 - x_4 + x_5 + 2x_6 &= 0 \\ 2x_2 - x_3 + 7x_5 &= 0 \\ x_1 + 3x_4 + 7x_6 &= 0 \\ x_3 - 5x_5 + 7x_6 &= 0. \end{aligned}$$

b)
$$\begin{aligned} 3x_1 - 2x_2 + x_3 - x_4 &= 7 \\ 2x_2 + x_2 - 2x_3 + x_4 &= -5 \\ -2x_2 + x_3 &= 7 \\ 5x_1 - x_2 - x_3 &= 9. \end{aligned}$$

c)
$$\begin{aligned} x_1 + 2x_2 + x_3 - 4x_4 &= 8 \\ 2x_2 - x_2 + 2x_3 + x_4 + x_5 &= 6 \\ -2x_2 + x_3 + 4x_5 &= 2 \\ 5x_1 - x_2 - x_3 &= 9 \\ 5x_1 - x_2 - x_3 + x_4 - 2x_5 &= 7. \end{aligned}$$

c)
$$\begin{aligned} 2x_1 - 2x_2 + x_3 + 4x_4 &= 3 \\ 2x_1 - x_2 + 2x_3 + x_4 + x_5 &= 4 \\ -2x_2 + x_4 + 4x_5 &= 1 \\ 5x_1 - x_3 - x_4 - x_5 &= 3 \\ 5x_1 - 2x_2 - 3x_3 + x_4 - 2x_5 &= 4. \end{aligned}$$

Capítulo 2

Determinantes

2.1 Introducción

En este capítulo definimos el *determinante* de una matriz de tamaño $n \times n$. La definición de este concepto se puede hacer de diversas maneras, pero aquí presentamos una que permite desarrollar un procedimiento relativamente fácil para el cálculo de determinantes. Con el propósito de permitir una mejor comprensión de dicho concepto, precisamos una serie de propiedades y aplicaciones de este.

El estudio del determinante de una matriz se justifica en la medida en que simplifica el problema de la resolución de sistemas lineales, proporciona otro método para el cálculo de la inversa de una matriz y cuenta con una gran variedad de aplicaciones geométricas.

En Occidente, la primera referencia a problemas relacionados con los determinantes aparece en unas cartas enviadas por Gottfried Leibniz en 1693 al marqués de L'Hôpital. En estas cartas Leibniz se refiere al problema de resolver un sistema de ecuaciones lineales. El japonés Seki Kowa introdujo los determinantes de orden 3 y 4 en la misma época que Leibniz. Los trabajos de este último sobre determinantes sólo aparecieron publicados en el año 1850.

La teoría de los determinantes surgió en el ámbito matemático alrededor de un siglo antes que el concepto de *matriz*, y fue precisamente James Joseph Sylvester quien acuñó el nombre de "matriz", al afirmar que ésta era "la madre de los determinantes".

El matemático que más resultados aportó al cálculo de los determinantes fue Luis Auguste Cauchy, quien publicó por primera vez el teorema $\det(AB) = (\det A)(\det B)$. El desarrollo de un determinante por cofactores fue empleado por vez primera por el matemático francés Pierre S. Laplace (1749-1827). La regla de Cramer aparece en el año 1750, aunque, curiosamente, ya Colin Maclaurin la había publicado en 1748.

El matemático alemán Carl Gustav Jacobi (1804-1851) fue quien en definitiva formalizó el concepto de *determinante* y, además, lo aplicó a otras ramas de la matemática. De hecho, a partir de sus aportes a las funciones de varias variables, encontramos el concepto de *jacobiano*. Arthur Cayley fue el inventor de la notación de los determinantes mediante barras verticales en 1841 y en 1858 estableció la fórmula para el cálculo de la inversa de una matriz mediante determinantes.

2.2 Determinante de una matriz

Para introducir la definición del determinante de una matriz, comencemos presentando el concepto de *permutación.*

Definición 2.2.1 (Permutación). *Dado un conjunto finito con todos sus elementos diferentes, llamamos permutación a cada una de las posibles ordenaciones de los elementos de dicho conjunto.*

Según la definición, una permutación difiere de otra en el orden de colocación de sus elementos. Para el caso que nos interesa, vamos a considerar el conjunto finito $I_n = \{1, 2, \ldots n\}$; en este caso, la definición anterior se transforma en la siguiente.

Definición 2.2.2. *Una permutación $p : I_n \to I_n$, es una función tal que si $i, j \in I_n$, con $i \neq j$, entonces $p(i) \neq p(j)$.*

Según la definición anterior, si p es una permutación, entonces el conjunto de enteros $p(1), p(2), \ldots, p(n)$ consta de los mismos n números, pero en un orden diferente.

Ejemplo 2.2.1. En el caso de I_6, una permutación p puede representarse de la manera

$$p = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{bmatrix}$$

o también en la forma

$$p = (2, 1, 3, 5, 6, 4).$$

En ambas representaciones es claro que $p(2) = 1$ o que $p(5) = 6$.

Si p y q son permutaciones en I_n, su composición $p \circ q$ es una nueva permutación en I_n. En el siguiente ejemplo se muestra cómo calcular la composición de estas permutaciones.

Ejemplo 2.2.2. Sean las permutaciones p y q en I_6 definidas por

$$p = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{bmatrix}, \quad q = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 5 & 4 & 6 \end{bmatrix}.$$

Entonces, $p \circ q(1) = p(q(1)) = p(3) = 3$. Los demás elementos se muestran en la forma usual:

$$p \circ q = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 6 & 5 & 4 \end{bmatrix}.$$

Como $p \circ q(5) = p(q(5)) = p(4) = 5$ y, de un modo similar, $q \circ p(5) = q(p(5)) = q(6) = 6$, se ve que la composición de permutaciones no es conmutativa.

Es claro que una permutación en I_n, es biyectiva y, en consecuencia, es posible definir la *inversa de una permutación.* Así, si denotamos con p^{-1} a la permutación inversa de p, se tiene que $p \circ p^{-1} = I = p^{-1} \circ p$, donde I es el orden natural $I = (1, 2, \ldots, n)$ o la *permutación identidad.*

Ejemplo 2.2.3. En el caso de I_6, dada la permutación

$$p = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{bmatrix},$$

entonces su inversa es

$$p^{-1} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 6 & 4 & 5 \end{bmatrix}.$$

Aquí puede verse que $p \circ p^{-1}(5) = p(p^{-1}(5)) = p(4) = 5$ y también que $p^{-1} \circ p(5) = p^{-1}(p(5)) = p^{-1}(6) = 5$.

Definición 2.2.3 (Inversión). *Dada la permutación $p : I_n \to I_n$, una inversión en p es una pareja de índices (i, j) tales que si se tiene que $1 \leq i < j \leq n$, entonces $p(i) > p(j)$. Al número de inversiones de una permutación lo denotamos con $inv(p)$, es decir,*

$$inv(p) = \sum_{i=1}^{n-1} |i < j : p(i) > p(j)|.$$

A cada permutación p se le asigna un número $\sigma(p)$, llamado signo de la permutación, tal que

1. $\sigma(p) = (-1)^{inv(p)}$.
2. *Dadas dos permutaciones p y q, entonces $\sigma(p \circ q) = \sigma(p)\sigma(q)$.*

Si el número de inversiones de p es par, entonces $\sigma(p) = 1$ y en tal caso decimos que la *permutación es par*; pero si $\sigma(p)$ es impar, concluimos que $\sigma(p) = -1$, y la *permutación será impar*.

Si consideramos la última propiedad propuesta en la definición anterior, tenemos:

$$\sigma(p \circ p^{-1}) = \sigma(p)\sigma(p^{-1}) = \sigma(I) = 1.$$

De allí se concluye que una permutación y su inversa tienen la misma paridad.

Ejemplo 2.2.4. En el caso de I_6, dada la permutación

$$p = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 6 & 5 & 3 & 4 \end{bmatrix},$$

las inversiones corresponden a los índices $(1, 2)$, $(3, 4)$, $(3, 5)$, $(3, 6)$, $(4, 5)$, $(4, 6)$. Por lo tanto, se trata de una permutación par.

Ejemplo 2.2.5. Las permutaciones $p = (2, 1, 3)$ y $q = (2, 3, 1)$ son dos permutaciones diferentes de $\{1, 2, 3\}$. En p, el 2 precede a 1 y, en consecuencia, la permutación es impar, mientras que en q el 2 y el 3 preceden a 1, con lo cual concluimos que hay dos inversiones y, por lo tanto, la permutación es par.

Para el caso en el cual n toma valores pequeños, una forma sencilla de calcular el número de inversiones de una permutación se ilustra en la Figura 2.1.

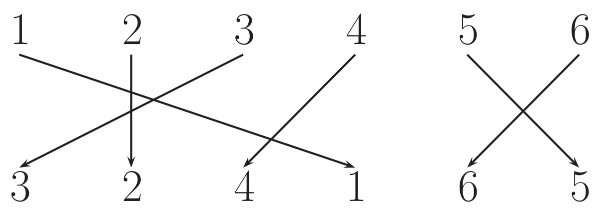


Figura 2.1. Inversiones de una permutación ($n = 6$)

El método consiste en escribir la lista $\{1, 2, \ldots, n\}$ y, en la parte inferior, la permutación $p = (p_1, p_2, \ldots, p_n)$; luego se traza una flecha uniendo el número j de la primera lista con el j de la segunda lista. El número de inversiones de la permutación se obtiene contando el número total de intersecciones de las n flechas. En la Figura 2.1 las flechas tienen 5 intersecciones y, en consecuencia, la permutación dada es impar.

Con base en las anteriores definiciones tenemos la siguiente definición.

Definición 2.2.4. *Sea A una matriz de tamaño $n \times n$. El determinante de A, que denotamos con $\det A$ o $|A|$, se define mediante la relación*

$$\det A = \sum_p \sigma(p) a_{1p_1} a_{2p_2} a_{3p_3} \cdots a_{np_n},$$

donde la suma se toma sobre todas las $n!$ permutaciones $p = (p_1 p_2 \cdots p_n)$ de $\{1, 2, \ldots, n\}$ y $\sigma(p)$ es el signo de la permutación, esto es, $\sigma(p) = 1$ si la permutación es par y $\sigma(p) = -1$ si la permutación es impar.

De la definición precedente se puede concluir que el determinante es una función $|A| : M_{n\times n} \to \mathbb{R}$. Obsérvese que cada término $a_{1p_1} a_{2p_2} a_{3p_3} \cdots a_{np_n}$ contiene un elemento de cada fila y de cada columna de la matriz A y que, en definitiva, se están sumando (con la paridad de la respectiva permutación) todos los factores que se forman usando un elemento de cada fila y de cada columna.

Ejemplo 2.2.6. Para el caso de una matriz

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

de tamaño 2×2, hay $2! = 2$ permutaciones de $\{1, 2\}$, las cuales son $(1, 2)$ y $(2, 1)$. De esta forma, el determinante de A se calcula como

$$\det A = \sigma(1, 2) a_{11} a_{22} + \sigma(2, 1) a_{12} a_{21}.$$

Pero como $\sigma(1, 2) = 1$ y $\sigma(2, 1) = -1$, obtenemos que

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

que es una expresión ya conocida para la solución de sistemas de dos ecuaciones con dos incógnitas.

Ejemplo 2.2.7. Consideremos una matriz A de tamaño 3×3. El número de sumandos que conforman el determinante es $3! = 6$, y si

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

los datos necesarios para poder calcular el determinante están resumidos en la Tabla 2.1.

Permutación	$\sigma(\mathbf{p})$	**Factores**
$(1,2,3)$	1	$a_{11}a_{22}a_{33}$
$(1,3,2)$	-1	$a_{11}a_{23}a_{32}$
$(2,1,3)$	-1	$a_{12}a_{21}a_{33}$
$(2,3,1)$	-1	$a_{12}a_{23}a_{31}$
$(3,1,2)$	1	$a_{13}a_{21}a_{32}$
$(3,2,1)$	-1	$a_{13}a_{22}a_{31}$

Tabla 2.1. Determinante de una matriz 3×3

Con esta información obtenemos la relación:

$$\begin{aligned} \det A &= (a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31}) \\ &\quad -(a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31}). \qquad (2.2.1) \end{aligned}$$

Esta relación puede recordarse mediante la Figura 2.2.

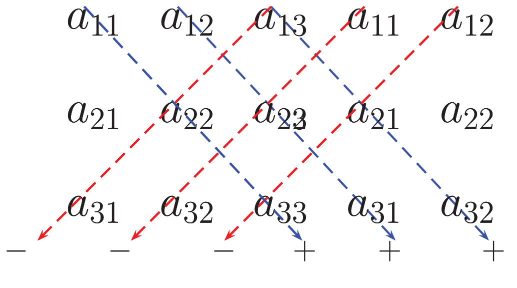


Figura 2.2. Cálculo de $\det A$, si A es de tamaño 3×3

En esta figura se han repetido, al final, las dos primeras columnas de la matriz y las flechas indican los productos que se forman cuando se calcula el determinante, lo mismo que el signo de los respectivos factores.

Para el caso de la matriz

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},$$

los datos para calcular el determinante se encuentran resumidos en la Tabla 2.2.

Permutación	$\sigma(\mathbf{p})$	**Factores**
$(1,2,3)$	1	$a_{11}a_{22}a_{33} = (1)(5)(9) = 45$
$(1,3,2)$	-1	$a_{11}a_{23}a_{32} = (1)(6)(8) = 48$
$(2,1,3)$	-1	$a_{12}a_{21}a_{33} = (2)(4)(9) = 72$
$(2,3,1)$	-1	$a_{12}a_{23}a_{31} = (2)(6)(7) = 84$
$(3,1,2)$	1	$a_{13}a_{21}a_{32} = (2)(4)(8) = 64$
$(3,2,1)$	-1	$a_{13}a_{22}a_{31} = (2)(5)(7) = 70$
$\det A$		$45 - 48 - 72 + 84 + 64 - 70 = 3$

Tabla 2.2. Determinante de la matriz A dada

Esta tabla es solo una manera de presentar el cálculo del determinante. En la práctica es más frecuente aplicar la fórmula 2.2.1 o el método gráfico expuesto en este ejercicio.

2.3 Propiedades de los determinantes

En adelante usamos la notación $A^{(j)}$ para expresar la columna j de la matriz A y con este criterio la matriz puede escribirse en la forma $A = [A^{(1)}, A^{(2)}, \ldots, A^{(n)}]$. Además, nos referimos a ella como la matriz A escrita en términos de sus columnas. Por su parte, la notación A_j se usará para identificar a la fila j de la matriz A, y con este criterio la matriz puede escribirse en la forma $A = [A_1, A_2, \ldots, A_n]$; aludimos a ella como la matriz A escrita en términos de sus filas.

Ahora bien: la definición 2.2.4 permite probar las siguientes propiedades.

Teorema 2.3.1. $\det A = \det A^t$.

Este teorema afirma que el determinante de una matriz y el de su transpuesta son iguales. Con este resultado, todas las propiedades de los determinantes que se establezcan para las filas de una matriz son válidas también con respecto a las columnas.

Demostración. Sea $A = [a_{ij}]$ y $A^t = [b_{ji}]$, con $b_{ji} = a_{ij}$. Entonces,

$$\det A^t = \sum_p \sigma(p) b_{1p_1} b_{2p_2} b_{3p_3} \cdots b_{np_n} = \sum_p \sigma(p) a_{p_1 1} a_{p_2 2} a_{p_3 3} \cdots a_{p_n n}.$$

Si ahora reordenamos los factores de $a_{p_1 1} a_{p_2 2} a_{p_3 3} \cdots a_{p_n n}$ en la forma natural $a_{1q_1} a_{2q_2} a_{3q_3} \cdots a_{nq_n}$, obtenemos que

$$b_{1p_1} b_{2p_2} b_{3p_3} \cdots b_{np_n} = a_{p_1 1} a_{p_2 2} a_{p_3 3} \cdots a_{p_n n} = a_{1q_1} a_{2q_2} a_{3q_3} \cdots a_{nq_n}.$$

Consideremos la permutación q tal que $q(p_i) = i$. Esta permutación es la inversa de la permutación p y como se estableció en la definición 2.2.4 de paridad de una permutación, las permutaciones p y $q = p^{-1}$ tienen la misma paridad, es decir, $\sigma(p) = \sigma(q)$. Esto lo podemos ver particularmente en el siguiente caso:

$$b_{12} b_{24} b_{33} b_{41} = a_{21} a_{42} a_{33} a_{14} = a_{14} a_{21} a_{33} a_{42}.$$

La primera permutación es $(2, 4, 3, 1)$, la cual tiene cuatro inversiones, mientras que la segunda permutación, $(4, 1, 3, 2)$, también tiene cuatro inversiones. De este modo $\sigma(p) = \sigma(q)$. Con este resultado obtenemos que

$$\sum_p \sigma(p) b_{1p_1} b_{2p_2} b_{3p_3} \cdots b_{np_n} = \sum_q \sigma(q) a_{1q_1} a_{2q_2} a_{3q_3} \cdots a_{nq_n},$$

con lo cual $\det A = \det A^t$.

Ejemplo 2.3.1. Consideremos la matriz

$$A = \begin{bmatrix} 4 & 2 \\ 1 & 6 \end{bmatrix}.$$

Entonces, $\det A = 4(6) - 2(1) = 22$; además,

$$A^t = \begin{bmatrix} 4 & 1 \\ 2 & 6 \end{bmatrix},$$

con lo cual $\det A^t = 4(6) - 1(2) = 22 = \det A$.

Teorema 2.3.2. *Si A es una matriz $n \times n$ y k es un número real, entonces* $\det[A_1, \ldots, kA_j, \ldots, A_n] = k \det[A_1, \ldots, A_j, \ldots, A_n]$.

Esta propiedad afirma que si multiplicamos una fila de la matriz A por un escalar k, entonces el determinante de la nueva matriz es igual a k veces el determinante de la matriz original.

Demostración. Sea la matriz

$$B = [A_1, \ldots, kA_j, \ldots, A_n],$$

que se obtiene de la matriz A multiplicando la fila A_j por el escalar k. Usando la definición 2.2.4, del determinante de una matriz, se tiene que:

$$\begin{aligned} \det B &= \sum_p \sigma(p) a_{1p_1} \cdots k a_{jp_j} \cdots a_{np_n} \\ &= k \sum_p \sigma(p) a_{1p_1} \cdots a_{jp_j} \cdots a_{np_n} \\ &= k \det A, \end{aligned}$$

como se había propuesto.

Ejemplo 2.3.2. Sea la matriz

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 3 & 5 & 1 \end{bmatrix},$$

para la cual $\det A = 6$. Calcule el determinante de la matriz

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

Solución

En este caso, usando la notación del teorema 2.3.2, se tiene que

$$\det[A_1, 2A_2, A_3] = 2\det[A_1, A_2, A_3] = 2(6) = 12.$$

Como consecuencia directa del teorema anterior, podemos calcular el determinante de la matriz kA, donde k es un escalar y A es una matriz $n \times n$. En este caso, kA puede escribirse en la forma

$$kA = [kA_1, \ldots, kA_j, \ldots, kA_n]$$

y, por lo tanto,

$$\begin{aligned} \det(kA) &= \det[kA_1, \ldots, kA_j, \ldots, kA_n] \\ &= k^n \det[A_1, \ldots, A_j, \ldots, A_n] \\ &= k^n \det A, \end{aligned}$$

es decir, el determinante de kA es k^n veces el determinante de A.

Ejemplo 2.3.3. Sea la matriz

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 3 & 5 & 1 \end{bmatrix},$$

para la cual $\det A = 6$. Calcule el determinante de la matriz

$$B = \begin{bmatrix} 2 & 8 & 4 \\ 4 & 8 & 4 \\ 6 & 10 & 2 \end{bmatrix}.$$

Solución

En este caso, usando la notación del teorema 2.3.2, se tiene que

$$\det[2A_1, 2A_2, 2A_3] = 2^3 \det[A_1, A_2, A_3] = 8(6) = 48.$$

Teorema 2.3.3. *Sea* $D_n = \{d_{11}, d_{22}, \cdots, d_{nn}\}$ *una matriz diagonal; su determinante está dado por* $\det D_n = d_{11}d_{22}d_{33}\cdots d_{nn}$*, es decir, el determinante de una matriz diagonal es el producto de los elementos de la diagonal.*

Demostración. Para calcular el determinante de una matriz, cada sumando de la forma $a_{1p_1}a_{2p_2}a_{3p_3}\cdots a_{np_n}$ que contenga un término por fuera de la diagonal es nulo, de modo que el único sumando no nulo es $a_{11}a_{22}a_{33}\cdots a_{nn}$, y su signo es positivo, ya que en esta permutación no existen inversiones y, en consecuencia,

$$\det D = \sum_p \sigma(p) d_{1p_1}d_{2p_2}d_{3p_3}\cdots d_{np_n} = d_{11}d_{22}d_{33}\cdots d_{nn}.$$

Corolario 2.3.1. *En el caso de la matriz identidad,* $d_{ii} = 1$*, para cada* $i = 1, 2, \ldots, n$*, con lo cual se tiene que* $\det I_n = 1$.

Ejemplo 2.3.4. Dada la matriz

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

entonces $\det A = (2)(3)(1)(2)(3) = 36$.

Teorema 2.3.4. *Si* $A_j = 0$ *para algún valor de* j*, con* $j = 1, \ldots n$*, entonces* $\det[A_1 \ldots, A_j, \ldots, A_n] = 0$.

El teorema anterior afirma que si hay una fila de ceros, el determinante es cero.

Demostración. Para calcular el determinante de una matriz, cada sumando de la forma $a_{1p_1}a_{2p_2}a_{3p_3}\cdots a_{np_n}$ tiene un factor de la fila j; por lo tanto, todos los factores son nulos y $\det A = 0$.

Ejemplo 2.3.5. Dada la matriz

$$A = \begin{bmatrix} 2 & 3 & 8 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 4 & 1 & 5 & 0 \\ 3 & 5 & 6 & 2 & 0 \\ 0 & 8 & 8 & 9 & 3 \end{bmatrix},$$

entonces $\det A = 0$, ya que la segunda fila es nula.

Teorema 2.3.5. *Dada la matriz* $A = [A_1, \ldots, A_j, \ldots, A_n]$*, se cumple que*

$$\begin{aligned} \det[A_1, \ldots, A_j + B_j, \ldots, A_n] &= \det[A_1, \ldots, A_j, \ldots, A_n] \\ &\quad + \det[A_1, \ldots, B_j, \ldots, A_n]. \end{aligned}$$

Este teorema afirma que si a una fila de una matriz le sumamos una fila arbitraria, el determinante de la matriz resultante es la suma del determinante de la matriz original más el determinante de la matriz que se obtiene de sustituir la fila original por la fila que se le sumó.

Demostración. Aplicando la definición del determinante de una matriz, tenemos que:

$$\det[A_1, \ldots, A_j + B_j, \ldots, A_n] = \sum_p \sigma(p) a_{1p_1} \cdots (a_{jp_j} + b_{jp_j}) \cdots a_{np_n};$$

luego de aplicar la ley distributiva, obtenemos:

$$\begin{aligned}\sum_p \sigma(p) a_{1p_1} \cdots (a_{jp_j} + b_{jp_j}) \cdots a_{np_n} &= \sum_p \sigma(p) a_{1p_1} \cdots a_{jp_j} \cdots a_{np_n} \\ &\quad + \sum_p \sigma(p) a_{1p_1} \cdots b_{jp_j} \cdots a_{np_n},\end{aligned}$$

con lo cual

$$\begin{aligned}\det[A_1, \ldots, A_j + B_j, \ldots, A_n] &= \det[A_1, \ldots, A_j, \ldots, A_n] \\ &\quad + \det[A_1, \ldots, B_j, \ldots, A_n] \\ &= \det A + \det A^*,\end{aligned}$$

donde la matriz A^* se obtiene al cambiar la fila A_j de A por B_j.

Ejemplo 2.3.6. Sea la matriz

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 3 & 5 & 1 \end{bmatrix},$$

para la cual $\det A = 6$. Calcular el determinante de la matriz

$$C = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 6 & 8 & 6 \end{bmatrix}.$$

Solución

En este caso se tiene que $\det C = \det[A_1, A_2, A_3 + B_3]$, donde $B_3 = [3, 3, 5]$ y, por lo tanto,

$$\det C = \begin{vmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 6 & 8 & 6 \end{vmatrix} = \det A + \det A^* = \begin{vmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 3 & 5 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 3 & 3 & 5 \end{vmatrix}.$$

En este caso, solo hay que realizar el cálculo del determinante de la matriz A^*. Y ya que el valor de $\det A$ está dado en el problema, es fácil verificar que $\det A^* = -14$ y, en consecuencia,

$$\det C = \det A + \det A^* = 6 - 14 = -8.$$

Teorema 2.3.6. *Dada la matriz* $B = [A_1, \ldots, A_l, \ldots, A_i, \ldots A_n]$, *que se obtiene al intercambiar las filas* A_i *y* A_l *de* A, *entonces se cumple que* $\det B = -\det A$.

Según este teorema, si se intercambian dos filas de una matriz, el determinante de la matriz resultante es el negativo del determinante de la matriz original.

Demostración. Al aplicar la definición de determinante a la matriz B, se tiene que:

$$\begin{aligned}
\det B &= \sum_q \sigma(q) b_{1q_1} \ldots b_{lq_l} \ldots b_{iq_i} \cdots b_{nq_n} \\
&= \sum_p \sigma(p) a_{p_1 1} \ldots a_{p_i i} \ldots a_{p_l l} \ldots a_{p_n n},
\end{aligned}$$

donde la permutación q se consigue al intercambiar las posiciones i y l de la permutación p. Si son dos filas consecutivas, q se obtiene realizando una inversión, por lo que $\sigma(q) = -\sigma(p)$. Si no son filas consecutivas, se puede llevar el elemento i a la posición l, realizando $l - i$ intercambios entre elementos consecutivos, lo cual genera $l - i$ inversiones en p y, luego, para llevar el elemento l a la posición i, se requieren $l - i - 1$ intercambios consecutivos. En total, el número de inversiones termina siendo $2(l - i) - 1$, lo que indica que q tiene signo contrario a p y, por lo tanto, $\sigma(q) = -\sigma(p)$, con lo cual,

$$\det B = -\sum_p \sigma(p) a_{p_1 1} \ldots a_{p_i i} \ldots a_{p_l l} \ldots a_{p_n n} = -\det A,$$

como queríamos probar.

Ejemplo 2.3.7. Sea la matriz

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 4 & 2 \\ 3 & 5 & 1 \end{bmatrix},$$

para la cual $\det A = 6$. Calcule el determinante de la matriz

$$B = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 5 & 1 \\ 2 & 4 & 2 \end{bmatrix}.$$

Solución

En este caso, la matriz B se obtiene de la inicial intercambiando las filas segunda y tercera; por lo tanto, $\det B = -\det A = -6$.

Teorema 2.3.7. *Si la matriz A tiene dos filas iguales, entonces se cumple que* $\det A = 0$.

Demostración. Si la fila i y la fila j son iguales, entonces, por el teorema 2.3.6, al intercambiar la fila i con la j, se tiene que las matrices resultantes no cambian, pero su determinante cambia de signo; es decir:

$$\det B = \det A = -\det A,$$

con lo cual $\det A = 0$.

Ejemplo 2.3.8. Dada la matriz

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

como la primera y la cuarta fila son iguales, entonces $\det A = 0$.

Teorema 2.3.8. *Si A es una matriz triangular superior o inferior de tamaño $n \times n$, entonces su determinante es el producto de los elementos de la diagonal.*

Demostración. Sea $A = [a_{ij}]$ una matriz triangular superior, es decir, $a_{ij} = 0$, si $i > j$. Entonces, cada uno de los términos $a_{1p_1} \dots a_{ip_i} \dots a_{np_n}$ que aparece en la suma que define el determinante de A solo puede ser distinto de cero cuando $1 \leq p_1, 2 \leq p_2, \dots, n \leq p_n$, en donde $\{p_1, p_2, \dots, p_n\}$ es una permutación de $\{1, 2, \dots, n\}$. Por lo tanto, debemos tener que $1 = p_1, 2 = p_2, \dots, n = p_n$, lo cual nos indica que el único término que puede ser no nulo en la definición del determinante de A es el factor $a_{11} \dots a_{ii} \dots a_{nn}$, que corresponde al producto de los elementos de la diagonal. Además, el signo de la permutación es positivo, ya que no hay ninguna transposición; en consecuencia, obtenemos que $\det A = a_{11} \dots a_{ii} \dots a_{nn}$.

Ejemplo 2.3.9. Dada la matriz

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

como la matriz es triangular superior, entonces el valor de su determinante está dado por $\det A = (1)(4)(2)(4) = 32$.

Teorema 2.3.9. *Si la matriz B se obtiene de la matriz A sumándole a una fila de A un múltiplo de otra fila de A, entonces* $\det B = \det A$.

Demostración. Sea $B = [A_1, \dots, A_i + cA_l, \dots, A_n]$ la matriz que se obtiene al sumarle a la fila i de A, c veces la fila l de A. Entonces, por los teoremas 2.3.5 y 2.3.2 obtenemos que

$$\begin{aligned}
\det B &= \det[A_1, \ldots, A_i + cA_l, \ldots, A_l, \ldots, A_n] \\
&= \det[A_1, \ldots, A_i, \ldots, A_n] + c\det[A_1, \ldots, A_l, \ldots, A_l, \ldots, A_n] \\
&= \det[A_1, \ldots, A_i, \ldots, A_n] \\
&= \det A,
\end{aligned}$$

puesto que $\det[A_1, \ldots, A_l, \ldots, A_l \ldots, A_n] = 0$, debido a que tiene dos filas repetidas.

Ejemplo 2.3.10. Dada la matriz

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 4 & 5 \\ 0 & 0 & 2 & 0 \\ 2 & 4 & 6 & 12 \end{bmatrix},$$

calcular su determinante.

Solución

Como

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 4 & 5 \\ 0 & 0 & 2 & 0 \\ 2 & 4 & 6 & 12 \end{bmatrix}}_{B} \xrightarrow{F_4 - 2F_1} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 4 & 5 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}}_{A}.$$

el teorema anterior afirma que $\det A = \det B$; pero como A es una matriz triangular superior, $\det B = 32$.

Teorema 2.3.10. *Si A y B son matrices $n \times n$, entonces:*

$$\det(AB) = \det A \det B.$$

Esto significa que el determinante de un producto de matrices es el producto de los determinantes de cada matriz.

Ejemplo 2.3.11. Dadas las matrices

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 4 & 5 \\ 0 & 0 & 2 & 0 \\ 2 & 4 & 6 & 12 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 4 & 5 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix},$$

entonces, usando los resultados del ejemplo anterior y el teorema 2.3.10, obtenemos que $\det(AB) = \det A \det B = (32)(32) = 1.024$.

Nota 2.3.1. Como se muestra en el teorema 2.3.1, el determinante de una matriz y su transpuesta son iguales; según esto, todos los teoremas relativos a filas de una matriz son también aplicables a columnas de una matriz.

Teorema 2.3.11. *Si A es una matriz no singular, entonces se cumple que:* $\det(A^{-1}) = (\det A)^{-1}$.

Demostración. Sea A una matriz no singular de tamaño $n \times n$; entonces, se cumple que $AA^{-1} = I_n$. Si aplicamos determinantes en ambos lados y tenemos en cuenta el teorema 2.3.10, obtenemos:

$$\det(AA^{-1}) = (\det A^{-1})(\det A) = \det I_n = 1,$$

con lo cual

$$\det A^{-1} = \frac{1}{\det A},$$

como se quería demostrar.

Los teoremas 2.3.2, 2.3.6, 2.3.8 y 2.3.9, que son aquellos que involucran el cálculo del determinante y las operaciones elementales en las filas de una matriz, proporcionan un método útil para calcular el determinante de una matriz: tome la matriz original y mediante operaciones elementales por filas, transfórmela en una matriz triangular superior o inferior, cuyo determinante es el producto de los elementos de la diagonal. El valor del determinante de la matriz inicial se obtiene teniendo en cuenta las distintas transformaciones que sufre la función determinante. El siguiente ejemplo ilustra este método.

Ejemplo 2.3.12. Sea la matriz

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

Calcule el determinante reduciendo la matriz a otra matriz que sea triangular superior o inferior.

Solución

Como hemos propuesto antes del ejemplo, mediante operaciones elementales por filas transformamos la matriz A en otra matriz triangular superior.

$$\underbrace{\begin{bmatrix} 2 & 4 & 6 & 8 \\ -1 & 2 & 0 & 1 \\ 4 & 3 & 8 & 3 \\ -2 & 0 & 4 & 1 \end{bmatrix}}_{A} \xrightarrow{\frac{1}{2}F_1} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & 2 & 0 & 1 \\ 4 & 3 & 8 & 3 \\ -2 & 0 & 4 & 1 \end{bmatrix}}_{B} \xrightarrow{f_2+F_1} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 4 & 3 & 8 & 3 \\ -2 & 0 & 4 & 1 \end{bmatrix}}_{C}$$

$$\xrightarrow{f_3-4F_1} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & -5 & -4 & -13 \\ -2 & 0 & 4 & 1 \end{bmatrix}}_{D} \xrightarrow{f_4+2F_1} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & -5 & -4 & -13 \\ 0 & 4 & 10 & 9 \end{bmatrix}}_{E}$$

$$\xrightarrow{f_4-F_2} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & -5 & -4 & -13 \\ 0 & 0 & 7 & 4 \end{bmatrix}}_{F} \xrightarrow{f_3\leftrightarrow F_4} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & 0 & 7 & 4 \\ 0 & -5 & -4 & -13 \end{bmatrix}}_{G}$$

$$\xrightarrow{f_4+\frac{5}{4}F_2} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & 0 & 7 & 4 \\ 0 & 0 & -\frac{1}{4} & -\frac{27}{4} \end{bmatrix}}_{H} \xrightarrow{-4f_4} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & 0 & 7 & 4 \\ 0 & 0 & 1 & 27 \end{bmatrix}}_{J}$$

$$\xrightarrow{f_3\leftrightarrow f_4} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & 0 & 1 & 27 \\ 0 & 0 & 7 & 4 \end{bmatrix}}_{R} \xrightarrow{f_4-7f_3} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 3 & 5 \\ 0 & 0 & 1 & 27 \\ 0 & 0 & 0 & -185 \end{bmatrix}}_{T}.$$

Como T es triangular superior, tenemos que $\det T = -740$. Además:

$$\begin{aligned}
\det T &= \det R && \text{(Por el teorema 2.3.9)}\\
\det R &= -\det J && \text{(Por el teorema 2.3.6)}\\
\det J &= -4\det H && \text{(Por el teorema 2.3.2)}\\
\det H &= \det G && \text{(Por el teorema 2.3.9)}\\
\det G &= -\det F && \text{(Por el teorema 2.3.6)}\\
\det F &= \det E && \text{(Por el teorema 2.3.9)}\\
\det E &= \det D && \text{(Por el teorema 2.3.9)}\\
\det D &= \det C && \text{(Por el teorema 2.3.9)}\\
\det C &= \det B && \text{(Por el teorema 2.3.9)}\\
\det B &= \tfrac{1}{2}\det A && \text{(Por el teorema 2.3.2).}
\end{aligned}$$

De las relaciones anteriores podemos concluir que

$$\det A = -2\left(\frac{1}{4}\right)\det T.$$

Por lo tanto, $\det A = 370$.

2.4 Cálculo de determinantes por cofactores

A continuación exponemos un método diferente para calcular el valor de un determinante de una matriz de tamaño $n \times n$. El método consiste en reducir el problema de calcular el determinante de una matriz de tamaño $n \times n$ al cálculo de determinantes de matrices de tamaño $(n-1) \times (n-1)$ y repetir dicho procedimiento hasta reducir el cálculo a matrices 2×2.

Definición 2.4.1. *Sea $A = [a_{ij}]$ una matriz de tamaño $n \times n$. El cofactor C_{ij} de a_{ij} se define como $C_{ij} = (-1)^{i+j} \det M_{ij}$, donde M_{ij} es la matriz de tamaño $(n-1) \times (n-1)$, que se obtiene al eliminar la fila i y la columna j de la matriz A.*

Ejemplo 2.4.1. Sea la matriz

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

entonces, los cofactores de A están dados por

$$C_{11} = (-1)^2 \det \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix} = 1 \qquad C_{12} = (-1)^3 \det \begin{bmatrix} -2 & 1 \\ 0 & 4 \end{bmatrix} = 8$$

$$C_{13} = (-1)^4 \det \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} = -6 \qquad C_{21} = (-1)^3 \det \begin{bmatrix} 4 & 0 \\ 3 & 4 \end{bmatrix} = -16$$

$$C_{22} = (-1)^4 \det \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} = 4 \qquad C_{23} = (-1)^5 \det \begin{bmatrix} 1 & 4 \\ 0 & 3 \end{bmatrix} = -3$$

$$C_{31} = (-1)^4 \det \begin{bmatrix} 4 & 0 \\ 1 & 1 \end{bmatrix} = 4 \qquad C_{32} = (-1)^5 \det \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = -1$$

$$C_{33} = (-1)^6 \det \begin{bmatrix} 1 & 4 \\ -2 & 1 \end{bmatrix} = 9.$$

Las siguientes definiciones son de gran utilidad y nos permiten construir métodos alternativos para calcular el determinante de una matriz.

Definición 2.4.2. *Sea $A = [a_{ij}]$ una matriz de tamaño $n \times n$.*

1. *La matriz de cofactores de A, denotada con $Cof(A)$, es la matriz $Cof(A) = [C_{ij}]$, donde C_{ij} es el cofactor de a_{ij}.*
2. *La matriz adjunta de A, denotada con $adj(A)$, es la transpuesta de la matriz de cofactores, es decir, $adj(A) = (Cof(A))^t$.*

Ejemplo 2.4.2. Sea la matriz

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

entonces, la matriz de cofactores y la adjunta están dadas por:

$$Cof(A) = \begin{bmatrix} 1 & 8 & -6 \\ -16 & 4 & -3 \\ 4 & -1 & 9 \end{bmatrix}, \quad adj(A) = \begin{bmatrix} 1 & -16 & 4 \\ 8 & 4 & -1 \\ -6 & -3 & 9 \end{bmatrix}.$$

El siguiente teorema nos proporciona un método alternativo para calcular determinantes.

Teorema 2.4.1. *Sea $A = [a_{ij}]$ una matriz de tamaño $n \times n$.*

1. *Para cada $1 \leq i \leq n$, se cumple que:*

   $$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$

   Este es el desarrollo por filas para el determinante.

2. *Para cada $1 \leq j \leq n$, se cumple que:*

   $$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

   Este es el desarrollo por columnas para el determinante.

Demostración. Aquí solo presentamos la prueba para el caso de $n = 3$. Sea la matriz

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

El determinante de la matriz A está dado por la ecuación 2.2.1, es decir:

$$\begin{aligned} \det A &= (a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31}) \\ &\quad -(a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31}), \end{aligned}$$

lo cual puede escribirse en la forma

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

Por otro lado, los cofactores de la primera fila son:

$$C_{11} = (-1)^2 \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} = a_{22}a_{33} - a_{23}a_{32}$$

$$C_{12} = (-1)^3 \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} = a_{23}a_{31} - a_{21}a_{33}$$

$$C_{13} = (-1)^4 \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = a_{21}a_{32} - a_{22}a_{31}$$

y, en consecuencia,

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13},$$

que es el desarrollo por los cofactores de la primera fila.

De manera similar, tomando el desarrollo por los cofactores de la tercera columna, obtenemos:

$$C_{13} = (-1)^4 \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = a_{21}a_{32} - a_{22}a_{31}$$

$$C_{23} = (-1)^5 \det \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} = a_{12}a_{31} - a_{11}a_{32}$$

$$C_{33} = (-1)^4 \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

mientras que la ecuación 2.2.1 puede escribirse como

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

En consecuencia,

$$\det A = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33},$$

que es el desarrollo por los cofactores de la tercera columna.

Ejemplo 2.4.3. Sea la matriz

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

Calcular el determinante de A mediante los cofactores de:

1. La primera fila.

2. La segunda columna.

3. La tercera columna.

Solución
A lo largo de los cofactores de la primera fila, tenemos la relación

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13},$$

con lo cual

$$\begin{aligned} \det A &= 2(-1)^2 \begin{vmatrix} 1 & 3 \\ 0 & 2 \end{vmatrix} + 4(-1)^3 \begin{vmatrix} 3 & 3 \\ 1 & 2 \end{vmatrix} + (-1)(-1)^4 \begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} \\ &= 2(2) - 4(6-3) - (-1) \\ &= 4 - 12 + 1 = -7. \end{aligned}$$

A lo largo de los cofactores de la segunda columna, tenemos la relación

$$\det A = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32},$$

con lo cual

$$\begin{aligned}
\det A &= 4(-1)^3 \begin{vmatrix} 3 & 3 \\ 1 & 2 \end{vmatrix} + 1(-1)^4 \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} + 0(-1)^5 \begin{vmatrix} 2 & -1 \\ 3 & 3 \end{vmatrix} \\
&= -4(6-3) + (4+1) + 0 \\
&= -12 + 5 = -7.
\end{aligned}$$

A lo largo de los cofactores de la tercera columna, tenemos la relación

$$\det A = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33},$$

con lo cual

$$\begin{aligned}
\det A &= (-1)(-1)^4 \begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} + 3(-1)^5 \begin{vmatrix} 2 & 4 \\ 1 & 0 \end{vmatrix} + 2(-1)^6 \begin{vmatrix} 2 & 4 \\ 3 & 1 \end{vmatrix} \\
&= -1(-1) - 3(-4) + 2(2-12) \\
&= 1 + 12 - 20 = -7.
\end{aligned}$$

Este ejemplo nos muestra que el determinante de una matriz se puede evaluar mediante los cofactores de cualquier fila o cualquier columna, pero preferiblemente lo hacemos mediante los cofactores de la fila o la columna con mayor número de ceros.

2.5 Adjuntas e inversas

Sea $A = [a_{ij}]$ una matriz de tamaño $n \times n$ y $C = [C_{ij}]$ la matriz de cofactores de A. Nos podemos preguntar, si $i \neq j$, cuál, es el valor de la expresión

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}.$$

Para responder esta pregunta, consideremos la matriz B, que resulta de sustituir en A la fila j por la fila i, con $i \neq j$.

$$B = \begin{array}{c} \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{1n} \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \to \text{fila}\, i \\ \\ \to \text{fila}\, j. \\ \\ \\ \end{array}$$

Como B tiene dos filas iguales, $\det B = 0$; por lo tanto, al hacer el desarrollo por los cofactores de la fila j, obtenemos:

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn} = 0.$$

Esto último responde a la pregunta formulada y teniendo en cuenta la definición del determinante de una matriz, podemos concluir que

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn} = \begin{cases} \det A & \text{si}\, i = j \\ 0 & \text{si}\, i \neq j, \end{cases}$$

lo cual equivale a la igualdad

$$\underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \vdots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix}}_{C^t} = \underbrace{\begin{bmatrix} \det A & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \det A \end{bmatrix}}_{(\det A)I_n};$$

es decir, $AC^t = (\det A)I_n$. Y tomando en cuenta la definición de la adjunta, obtenemos: $A adj(A) = (\det A)I_n$, con lo cual llegamos al siguiente teorema.

Teorema 2.5.1. *Sea A una matriz $n \times n$; entonces,*

$$A\,adj(A) = adj(A)A = (\det A)I_n.$$

Corolario 2.5.1. *Sea A una matriz $n \times n$; si* $\det A \neq 0$, *entonces A es no singular y*

$$A^{-1} = \frac{1}{\det A} adj(A),$$

Como consecuencia del corolario anterior y del hecho que

$$\det A^{-1} = \frac{1}{\det A},$$

obtenemos el siguiente teorema.

Teorema 2.5.2. *Sea A una matriz $n \times n$; A es no singular si y solo si* $\det A \neq 0$.

Corolario 2.5.2. *Sea A una matriz $n \times n$; el sistema homogéneo $AX = 0$ tiene solución no trivial si y solo si* $\det A = 0$.

Demostración. Sabemos que el sistema homogéneo $AX = 0$ tiene solución no trivial si y solo si A es singular; y, por el teorema anterior, podemos concluir que A es singular si y solo si $\det A = 0$, lo cual completa la prueba.

Ejemplo 2.5.1. Sea la matriz

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

Determinar si es invertible. En caso de serlo, calcular su inversa.

Solución

Sabemos que $\det A = 4 - 6 = -2 \neq 0$, con lo cual concluimos que tiene inversa. Para calcular esta, determinamos su adjunta mediante el cálculo de la matriz de cofactores:

$$\begin{aligned} C_{11} &= (-1)^2 \det[4] = 4 \\ C_{12} &= (-1)^3 \det[3] = -3 \\ C_{21} &= (-1)^3 \det[2] = -2 \\ C_{22} &= (-1)^4 \det[1] = 1, \end{aligned}$$

y de esta forma

$$Cof(A) = \begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix} \quad \text{y} \quad adj(A) = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix},$$

lo que finalmente nos conduce a que

$$A^{-1} = \frac{1}{\det A} = \frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}.$$

Ejemplo 2.5.2. Sea la matriz

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

Determinar si A es no singular; en caso de serlo, calcular A^{-1}.
Solución
Procedemos inicialmente a calcular el determinante de la matriz; para ello podemos utilizar los cofactores de la primera fila:

$$\begin{aligned} \det A &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= 1(-1)^2 \det \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix} + 4(-1)^3 \det \begin{bmatrix} -2 & 1 \\ 0 & 4 \end{bmatrix} \\ &= 1 - 4(-8) \\ &= 1 + 32 \\ &= 33. \end{aligned}$$

Como $\det A \neq 0$, entonces la matriz es invertible. Del ejemplo 2.4.2 de la sección anterior sabemos que

$$adj(A) = \begin{bmatrix} 1 & -16 & 4 \\ 8 & 4 & -1 \\ -6 & -3 & 9 \end{bmatrix};$$

por lo tanto, tenemos que

$$A^{-1} = \frac{1}{\det A}\, adj(A) = \frac{1}{33}\begin{bmatrix} 1 & -16 & 4 \\ 8 & 4 & -1 \\ -6 & -3 & 9 \end{bmatrix},$$

como puede verificarse por el método de escalonamiento.

2.6 Regla de Cramer

El método de Cramer nos permite resolver un sistema de n ecuaciones con n incógnitas, cuando la matriz de coeficientes del sistema sea una matriz no singular. Este resultado se resume en el siguiente teorema.

Teorema 2.6.1. *Sea el sistema de n ecuaciones lineales*

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots \qquad\quad \vdots & = & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & b_n \end{array}$$

con n incógnitas y sea $A = [a_{ij}]$ la matriz de coeficientes del sistema; esto es, $AX = B$, con

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad y \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Entonces, si $\det A \neq 0$, el sistema tiene solución única y, además,

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \quad \ldots, \quad x_n = \frac{\det A_n}{\det A,}$$

donde A_j es la matriz que se obtiene reemplazando la columna j de A por los elementos del vector B.

Demostración. Sabemos que si $\det A \neq 0$, entonces A es no singular y, por lo tanto,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A^{-1}B = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \dots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

donde C_{ij} es el cofactor de a_{ij}. De esta última expresión concluimos que

$$x_i = \frac{1}{\det A}(C_{1i}b_1 + C_{2i}b_2 + \ldots + C_{ni}b_n), \quad 1 \leq i \leq n.$$

Consideremos ahora la matriz

$$A_i = \begin{bmatrix} a_{11} & \dots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \dots & a_{1n} \\ a_{21} & \dots & a_{2(i-1)} & b_2 & a_{2(i+1)} & \dots & a_{2n} \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & b_n & a_{n(i+1)} & \dots & a_{nn} \end{bmatrix}.$$

Si calculamos el determinante de A_i mediante los cofactores de la columna i, se obtiene que

$$\det A_i = C_{1i}b_1 + C_{2i}b_2 + \ldots + C_{ni}b_n,$$

con lo cual se establece la relación

$$x_i = \frac{\det A_i}{\det A}, \quad 1 \leq i \leq n.$$

Ejemplo 2.6.1. Resolver el sistema lineal de ecuaciones

$$\begin{aligned} 2x + 4y + 6z &= 2 \\ x + 2z &= 0 \\ 2x + 3y - z &= -5. \end{aligned}$$

Solución

Para este sistema, la matriz de coeficientes viene dada por

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

y, en consecuencia, su determinante, calculado usando los elementos de la segunda fila, corresponde a:

$$\det A = 1C_{21} + 0C_{22} + 2C_{23} = -1 \begin{vmatrix} 4 & 6 \\ 3 & -1 \end{vmatrix} - 2 \begin{vmatrix} 2 & 4 \\ 2 & 3 \end{vmatrix},$$

con lo cual

$$\det A = (-1)(-4-18) - 2(6-8) = 22 + 4 = 26.$$

Entonces, las incógnitas del sistema están dadas por:

$$x_1 = x = \frac{\det \begin{bmatrix} 2 & 4 & 6 \\ 0 & 0 & 2 \\ -5 & 3 & -1 \end{bmatrix}}{26} = -\frac{52}{26} = -2$$

$$x_2 = y = \frac{\det \begin{bmatrix} 2 & 2 & 6 \\ 1 & 0 & 2 \\ 2 & -5 & -1 \end{bmatrix}}{26} = \frac{0}{26} = 0$$

$$x_3 = z = \frac{\det \begin{bmatrix} 2 & 4 & 2 \\ 1 & 0 & 0 \\ 2 & 3 & -5 \end{bmatrix}}{26} = \frac{26}{26} = 1.$$

2.7 Ejercicios resueltos

Ejercicio 2.7.1. Sean A y B matrices de tamaño 3×3, tales que $\det A = -3$ y $\det B = 2$. Calcular el valor de $\det(3A^{-1}B^t(4A)^{-1})$.
Solución

Aplicando las propiedades de los determinantes, obtenemos:

$$\begin{aligned}
\det(3A^{-1}B^t(4A)^{-1}) &= \det(3A^{-1})(\det B^t)\det(4A)^{-1} \\
&= 3^3\det(A^{-1})(\det B)\left(\frac{1}{\det(4A)}\right) \\
&= 3^3\left(\frac{1}{\det A}\right)(\det B)\left(\frac{1}{4^3\det A}\right) \\
&= 27\left(\frac{1}{-3}\right)(2)\left(\frac{1}{64(-3)}\right) = \frac{3}{32}.
\end{aligned}$$

Ejercicio 2.7.2. Hallar los valores de λ para los cuales la matriz

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

sea no singular.
Solución

Sabemos que A es no singular si y solo si $\det A \neq 0$. Para este caso, $\det A = \lambda(\lambda+2)(3\lambda+1) = 0$, con lo cual obtenemos que $\lambda = 0$, $\lambda = -2$ y $\lambda = -1/3$, si y solo si A es singular. Luego A es no singular si y solo si $\lambda \in \mathbb{R} - \{0, -2, -1/3\}$.

Ejercicio 2.7.3. Sea A una matriz antisimétrica de tamaño $n \times n$. Probar que si A es no singular, entonces n es par.
Solución

Como A es antisimétrica, entonces $A^t = -A$ y, en consecuencia, $\det A^t = \det(-A)$. Pero se sabe que $\det A^t = \det A$; por lo tanto, $\det A = \det(-A)$, con lo cual $\det A = (-1)^n \det A$.

A partir de esta última relación concluimos que

$$(1-(-1)^n)\det A = 0.$$

Pero como A es no singular, $\det A \neq 0$; por lo tanto, $(-1)^n = 1$, de donde se concluye que n es par.

Ejercicio 2.7.4. Hallar los valores de k para los cuales la matriz

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

no posee inversa. Encontrar A^{-1} si $k = 5$.

Solución

Uno de los criterios para determinar si una matriz posee o no inversa es el determinante, esto es, la matriz A no posee inversa si $\det A = 0$. Para la matriz A podemos calcular el determinante a lo largo de los elementos de la primera fila, así:

$$\begin{aligned} \det A &= 1\begin{vmatrix} 1 & -1 \\ -1 & k \end{vmatrix} - (-1)\begin{vmatrix} 0 & -1 \\ 2 & k \end{vmatrix} \\ &= (k-1)+2 \\ &= k+1. \end{aligned}$$

Se sabe que A no tiene inversa cuando $\det A = 0$, es decir, cuando $k+1=0$, con lo cual $k=-1$.

Ahora, cuando $k=5$, la matriz A se transforma en

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

Calculemos A^{-1} usando la adjunta de la matriz; para ello es necesario hallar los cofactores de A :

$$C_{11} = (-1)^2 \det \begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix} = 4 \qquad C_{12} = (-1)^3 \det \begin{bmatrix} 0 & -1 \\ 2 & 5 \end{bmatrix} = -2$$

$$C_{13} = (-1)^4 \det \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix} = -2 \qquad C_{21} = (-1)^3 \det \begin{bmatrix} -1 & 0 \\ -1 & 5 \end{bmatrix} = 5$$

$$C_{22} = (-1)^4 \det \begin{bmatrix} 1 & 0 \\ 2 & 5 \end{bmatrix} = 5 \qquad C_{23} = (-1)^5 \det \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} = -1$$

$$C_{31} = (-1)^4 \det \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix} = 1 \qquad C_{32} = (-1)^5 \det \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 1$$

$$C_{33} = (-1)^6 \det \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = 1.$$

Por lo tanto, la matriz de cofactores de A está dada por

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

y la adjunta de A es

$$adj(A) = (Cof(A))^t = \begin{bmatrix} 4 & 5 & 1 \\ -2 & 5 & 1 \\ -2 & -1 & 1 \end{bmatrix}.$$

Finalmente, al ser $\det A = 6$, obtenemos que

$$A^{-1} = \frac{1}{6}\, adj(A) = \frac{1}{6} \begin{bmatrix} 4 & 5 & 1 \\ -2 & 5 & 1 \\ -2 & -1 & 1 \end{bmatrix}.$$

Ejercicio 2.7.5. Sea A una matriz simétrica y no singular; probar que $adj(A)$ es una matriz simétrica.

Solución

Como $adj(A)A = (\det A)I_n$, obtenemos que

$$(adj(A)A)^t = ((\det A)I_n)^t,$$

lo cual, al aplicar propiedades de la transpuesta, se transforma en

$$A^t(adj(A))^t = (\det A)I_n.$$

Si multiplicamos en ambos lados de la igualdad anterior por $adj(A)$, se tiene

$$(adj(A))A^t(adj(A))^t = (adj(A))(\det A)I_n,$$

que es equivalente a:

$$(adj(A))A(adj(A))^t = (\det A)(adj(A)),$$

puesto que la matriz A es simétrica.

Si de nuevo hacemos uso del hecho que $adj(A)A = (\det A)I_n$, la relación anterior se transforma en

$$(\det A)I_n(adj(A))^t = (\det A)(adj(A)),$$

de lo cual concluimos que

$$(\det A)(adj(A))^t = (\det A)(adj(A)).$$

Y al ser $\det A \neq 0$, se tiene:

$$(adj(A))^t = adj(A),$$

lo cual prueba que la matriz adjunta de A es una matriz simétrica.

Ejercicio 2.7.6. Sea la matriz A de tamaño $n \times n$ dada por:

$$A = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 & 1 \\ -1 & 2 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 2 & \ldots & 0 & 0 \\ 0 & 0 & -1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 & 2 \end{bmatrix}.$$

Calcular el valor de su determinante.

Solución

Sea $\Delta_n = \det A$. Desarrollando el determinante de la matriz A, mediante los cofactores de la primera columna se tiene:

$$\Delta_n = \begin{vmatrix} 2 & 0 & 0 & \ldots & 0 \\ -1 & 2 & 0 & \ldots & 0 \\ 0 & -1 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & -1 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ -1 & 2 & 0 & \ldots & 0 \\ 0 & -1 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & -1 & 2 \end{vmatrix}.$$

Y como la primera matriz es triangular inferior, su determinante es el producto de los elementos de la diagonal; por lo tanto,

$$\Delta_n = 2^{n-1} + \Delta_{n-1},$$

donde Δ_{n-1} es un matriz de la misma forma que la original, pero de tamaño $(n-1) \times (n-1)$.

Al calcular de nuevo el determinante Δ_{n-1} a lo largo de los elementos de la primera columna y repitiendo este proceso sucesivamente a matrices de tamaño cada vez menor, se tiene que:

$$\Delta_n = 2^{n-1} + 2^{n-2} + 2^{n-3} + \ldots + 2^2 + 2^1 + 2^0,$$

mientras que

$$2\Delta_n = 2^n + 2^{n-1} + 2^{n-2} + \ldots + 2^3 + 2^2 + 2.$$

De esta forma, si restamos las dos últimas expresiones, conseguimos la relación

$$\Delta_n = 2^n - 2^0 = 2^n - 1.$$

Ejercicio 2.7.7. Resolver el sistema lineal de ecuaciones

$$\begin{aligned} 2x + 3y + z &= 9 \\ -2x + +y - 2z &= 6 \\ x + 3y + 2z &= 12 \end{aligned}$$

utilizando la regla de Cramer.

Solución

Para este sistema, la matriz de coeficientes viene dada por

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

y, en consecuencia, su determinante, calculado con los elementos de la primera fila, está dado por

$$\det A = 2C_{11} - 3C_{12} + 1C_{13}$$

$$\det A = 2\begin{vmatrix} 1 & -2 \\ 3 & 2 \end{vmatrix} - 3\begin{vmatrix} -2 & -2 \\ 1 & 2 \end{vmatrix} + 1\begin{vmatrix} -2 & 1 \\ 1 & 3 \end{vmatrix},$$

con lo cual

$$\det A = 2(2+6) - 3(-4+2) + 1(-6-1) = 16 + 6 - 7 = 15.$$

Entonces, las incógnitas del sistema están dadas por

$$x_1 = x = \frac{\det \begin{bmatrix} 9 & 3 & 1 \\ 6 & 1 & -2 \\ 12 & 3 & 2 \end{bmatrix}}{15}$$

$$x_2 = y = \frac{\det\begin{bmatrix} 2 & 9 & 1 \\ -2 & 6 & -2 \\ 1 & 12 & 2 \end{bmatrix}}{15}$$

$$x_3 = z = \frac{\det\begin{bmatrix} 2 & 3 & 9 \\ -2 & 1 & 6 \\ 1 & 3 & 12 \end{bmatrix}}{15}.$$

Calculando cada uno de estos determinantes mediante los elementos de la primera fila, se llega finalmente al resultado:

$$x = \frac{9\ (2+6) - 3\ (12+24) + 1\,(18-12)}{15} = \frac{72-108+6}{15} = -2$$

$$y = \frac{2\,(12+24) - 9\,(-4+2) + 1\,(-24-6)}{15} = \frac{72+18-30}{15} = 4$$

$$z = \frac{2(12-18) - 3(-24-6) + 9(-6-1)}{15} = \frac{-12+90-63}{15} = 1.$$

Ejercicio 2.7.8. Sea la matriz

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

1. Calcule $\det A$ mediante los cofactores de la segunda fila.
2. Calcule $\det A$ mediante los cofactores de la tercera columna.
3. Calcule el determinante de la matriz utilizando propiedades, es decir, llevando la matriz inicial a una matriz triangular superior o inferior.

Solución

1. A lo largo de los cofactores de la segunda fila, tenemos la relación

$$\det A = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23},$$

con lo cual

$$\begin{aligned} \det A &= -1\begin{vmatrix} -1 & 1 \\ 4 & 3 \end{vmatrix} - 2\begin{vmatrix} 2 & -1 \\ -1 & 4 \end{vmatrix} \\ &= -1(-3-4) - 2(8-1) \\ &= 7 - 14 \\ &= -7. \end{aligned}$$

2. A lo largo de los cofactores de la tercera columna, tenemos la relación

$$\det A = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33},$$

con lo cual

$$\begin{aligned} \det A &= 1(-1)^4\begin{vmatrix} 1 & 0 \\ -1 & 4 \end{vmatrix} + 2(-1)^5\begin{vmatrix} 2 & -1 \\ -1 & 4 \end{vmatrix} + 3(-1)^6\begin{vmatrix} 2 & -1 \\ 1 & 0 \end{vmatrix} \\ &= 1(4-0) - 2(8-1) + 3(0+1) \\ &= 4 - 14 + 3 \\ &= -7. \end{aligned}$$

3. Como hemos propuesto al inicio del ejemplo, mediante operaciones elementales por filas transformamos la matriz A en otra matriz triangular superior.

$$\underbrace{\begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \\ -1 & 4 & 3 \end{bmatrix}}_{A} \xrightarrow{F_1 \leftrightarrow F_2} \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 1 \\ -1 & 4 & 3 \end{bmatrix}}_{B} \xrightarrow[F_2 - 2F_1]{F_3 + F_1} \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -3 \\ 0 & 4 & 5 \end{bmatrix}}_{C}$$

$$\xrightarrow{-F_2} \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 4 & 5 \end{bmatrix}}_{D} \xrightarrow{F_3-4F_2} \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & -7 \end{bmatrix}}_{T}.$$

Como la matriz T es triangular superior, su determinante es el producto de los elementos de la diagonal; por lo tanto, $\det T = -7$.

Si ahora tenemos en cuenta las propiedades del determinante, se tiene:

$$\begin{aligned} \det T &= \det D \\ \det D &= -\det C \\ \det C &= \det B \\ \det B &= -\det A, \end{aligned}$$

a partir de lo cual podemos concluir que $\det A = \det T = -7$.

Ejercicio 2.7.9. (**Uso del computador**) Dado el sistema de ecuaciones lineales

$$\begin{aligned} x + 2y + z + at &= 1 \\ ax + 4y + 2z + 4t &= 2 \\ (a-1)x + ay + z + 2t &= 1 \\ x - 2at &= 1, \end{aligned}$$

determinar para qué valores de a el sistema posee:

1. Infinitas soluciones.
2. Solución única.
3. Ninguna solución.

Solución

Para hallar el número de soluciones que posee un sistema de ecuaciones lineales, podemos calcular el rango de la matriz del sistema y el rango de la matriz aumentada.

Como se mencionó previamente, DERIVE cuenta con la función ROW_REDUCE, que efectúa operaciones elementales por filas y determina la forma escalonada reducida de la matriz dada. Editando la matriz aumentada del sistema, usamos la expresión $M(a)$, la cual nos indica que la matriz M depende del parámetro a. Como puede observarse, al simplificar, en la forma escalonada reducida de la matriz aumentada del sistema, aparece una división por el factor $a - 1$ y, en consecuencia, la matriz solo podrá escalonarse si $a \neq 1$.

$$\#1. \quad M(a) := \begin{bmatrix} 1 & 2 & 1 & a & 1 \\ a & 4 & 2 & 4 & 2 \\ a-1 & a & 1 & 2 & 1 \\ 1 & 0 & 0 & -2a & 1 \end{bmatrix}$$

#2. ROW_REDUCE(M(a))

$$\#3. \quad \begin{bmatrix} 1 & 0 & 0 & 0 & \dfrac{1}{1-a} \\ 0 & 1 & 0 & 0 & \dfrac{1}{2(a-1)} \\ 0 & 0 & 1 & 0 & \dfrac{3a-2}{2(a-1)} \\ 0 & 0 & 0 & 1 & \dfrac{1}{2(1-a)} \end{bmatrix}.$$

Lo anterior nos indica que el procedimiento utilizado para calcular el rango de la matriz no es correcto cuando la matriz dada depende de un parámetro.

Otra alternativa para resolver el problema consiste en calcular el determinante de la matriz de coeficientes del sistema, lo cual puede realizarse en DERIVE con la instrucción $det(M)$.

Nótese que, en este caso, la matriz dada es la matriz aumentada del sistema. Para no editar una nueva matriz, calculamos la expresión $det(M(a)col[1,2,3,4])$, la cual establece el determinante de la matriz que se obtiene al eliminar la quinta columna de $M(a)$.

#4. $\det(M(a)col[1,2,3,4])$

#5. $-2(a^3 - 5a^2 + 8a - 4)$

#6. SOLVE $(-2(a^3 - 5a^2 + 8a - 4))$

#7. $a = 2 \ \vee \ a = 1$

El determinante que se obtiene es un polinomio de grado tres en la variable a, al cual llamamos $P(a)$.

Como es sabido, un sistema tiene solución única cuando el determinante de la matriz de coeficientes del sistema es no nulo. De esta forma, cuando resolvemos la ecuación $P(a) = 0$, estamos calculando los valores de a para los cuales el sistema tiene infinitas soluciones o ninguna. Usando el comando SOLVE, calculamos las raíces del polinomio $P(a)$.

Al sustituir los valores de las raíces calculadas en la matriz $M(a)$ y establecer la forma escalonada reducida para cada una de las matrices, obtenemos los siguientes resultados:

#8. ROW_REDUCE(M(1))

#9. $$\begin{bmatrix} 1 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

#10. ROW_REDUCE(M(2))

$$\#11. \quad \begin{bmatrix} 1 & 0 & 0 & -4 & 1 \\ 0 & 1 & \frac{1}{2} & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

De la expresión #9 se concluye que si $a = 1$, el sistema no tiene solución, mientras que a partir de la expresión 11 se concluye que si $a = 2$, el sistema tiene infinitas soluciones.

Ejercicio 2.7.10. Una matriz es ortogonal si $A^t = A^{-1}$. Demuestre que si A es ortogonal, entonces se cumplen las siguientes propiedades:

1. $\det A^2 = 1$.
2. A^{-1} es ortogonal.
3. A^t es ortogonal.

Solución

1. Como A es ortogonal, $\det A^t = \det A^{-1}$; pero también $\det A^t = \det A$, a partir de lo cual concluimos que $\det A = \det A^{-1}$. En consecuencia,

$$\det A^2 = (\det A)(\det A) = (\det A)(\det A^{-1}) = \det(AA^{-1}) = 1.$$

2. Al ser $A^t = A^{-1}$, entonces

$$\begin{aligned} (A^t)^{-1} &= A \\ (A^{-1})^t &= A \\ (A^{-1})^t &= (A^{-1})^{-1} \end{aligned}$$

y, en consecuencia, A^{-1} es ortogonal.

3. De nuevo, como $A^t = A^{-1}$, se tiene que $(A^t)^t = (A^{-1})^t = (A^t)^{-1}$, lo cual prueba que A^t es ortogonal.

Ejercicio 2.7.11. Determinar los valores de α y β para que las siguientes matrices sean invertibles; en tal caso, hallar dichas inversas.

$$A = \begin{bmatrix} 0 & \alpha \\ \beta & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & \alpha \\ 0 & -1 & 0 \\ \alpha & \beta & 0 \end{bmatrix}.$$

Solución

Para la matriz A, como $\det A = -\alpha\beta$, concluimos que A es no singular si y solo si $\det A = -\alpha\beta \neq 0$, es decir, si $\alpha \neq 0$ y $\beta \neq 0$. Para calcular la inversa establecemos la matriz adjunta de A :

$$\begin{aligned} C_{11} &= (-1)^2(2) = 2 \\ C_{12} &= (-1)^3(\beta) = -\beta \\ C_{21} &= (-1)^3(\alpha) = -\alpha \\ C_{22} &= (-1)^4(0) = 0, \end{aligned}$$

con lo cual concluimos que

$$Cof(A) = \begin{bmatrix} 2 & -\beta \\ -\alpha & 0 \end{bmatrix}$$

y, en consecuencia,

$$adj(A) = (Cof(A))^t = \begin{bmatrix} 2 & -\alpha \\ -\beta & 0 \end{bmatrix}.$$

Finalmente, recordando la fórmula para la inversa obtenemos que

$$A^{-1} = \frac{1}{\det A} adj(A) = \frac{-1}{\alpha\beta} \begin{bmatrix} 2 & -\alpha \\ -\beta & 0 \end{bmatrix},$$

es decir,

$$A^{-1} = \begin{bmatrix} \dfrac{-2}{\alpha\beta} & \dfrac{1}{\beta} \\ \dfrac{1}{\alpha} & 0 \end{bmatrix}.$$

Para la matriz B, usando los cofactores de la segunda fila, se tiene que

$$\det B = (-1)(-1)^4 \det \begin{bmatrix} 1 & \alpha \\ \alpha & 0 \end{bmatrix} = \alpha^2$$

y, en consecuencia, B es no singular si y solo si $\alpha^2 \neq 0$, o lo que es equivalente, $\alpha \neq 0$. Para calcular B^{-1} usamos de nuevo la matriz adjunta de B. Los cofactores de B están dados por:

$$C_{11} = (-1)^2 \det \begin{bmatrix} -1 & 0 \\ \beta & 0 \end{bmatrix} = 0 \qquad C_{12} = (-1)^3 \det \begin{bmatrix} 0 & 0 \\ \alpha & 0 \end{bmatrix} = 0$$

$$C_{13} = (-1)^4 \det \begin{bmatrix} 0 & -1 \\ \alpha & \beta \end{bmatrix} = \alpha \qquad C_{21} = (-1)^3 \det \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix} = \alpha\beta$$

$$C_{22} = (-1)^4 \det \begin{bmatrix} 1 & \alpha \\ \alpha & 0 \end{bmatrix} = -\alpha^2 \quad C_{23} = (-1)^5 \det \begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix} = -\beta$$

$$C_{31} = (-1)^4 \det \begin{bmatrix} 0 & \alpha \\ -1 & 0 \end{bmatrix} = \alpha \quad C_{32} = (-1)^5 \det \begin{bmatrix} 1 & \alpha \\ 0 & 0 \end{bmatrix} = 0$$

$$C_{33} = (-1)^6 \det \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = -1$$

y, por lo tanto, la matriz adjunta es

$$adj(B) = \begin{bmatrix} 0 & \alpha\beta & \alpha \\ 0 & -\alpha^2 & 0 \\ \alpha & -\beta & -1 \end{bmatrix},$$

de tal forma que

$$B^{-1} = \frac{1}{\det B} adj(B) = \frac{1}{\alpha^2} \begin{bmatrix} 0 & \alpha\beta & \alpha \\ 0 & -\alpha^2 & 0 \\ \alpha & -\beta & -1 \end{bmatrix},$$

es decir,

$$B^{-1} = \begin{bmatrix} 0 & \dfrac{\beta}{\alpha} & \dfrac{1}{\alpha} \\ 0 & -1 & 0 \\ \dfrac{1}{\alpha} & -\dfrac{\beta}{\alpha^2} & -\dfrac{1}{\alpha^2} \end{bmatrix}.$$

Ejercicio 2.7.12. Dada la matriz

$$A = \begin{bmatrix} \alpha & 2 & 2 \\ \alpha & \alpha & 2 \\ 1 & \alpha & 1 \end{bmatrix},$$

determine los valores de α para los cuales dicha matriz sea invertible; en tal caso, halle su inversa.

Solución

Sabemos que una matriz es invertible si su determinante es diferente de cero. Si definimos la matriz en DERIVE, podemos calcular su determinante mediante las instrucciones:

#1. $A := \begin{bmatrix} \alpha & 2 & 2 \\ \alpha & \alpha & 2 \\ 1 & \alpha & 1 \end{bmatrix}$

#2. $\det(A)$

#3. $\alpha^2 - 4\alpha + 4$

#4. $\text{SOLVE}(\alpha^2 - 4\alpha + 4, \alpha)$

#5. $\alpha^2 = 2.$

De esta forma obtenemos que $\det A = \alpha^2 - 4\alpha + 4$ y, en consecuencia, la matriz es singular cuando $\det A = \alpha^2 - 4\alpha + 4 = 0$. Usando la instrucción para resolver ecuaciones en DERIVE encontramos, como puede verse en la instrucción #5, que si $\alpha = 2$ no existe la inversa de A.

Para valores diferentes, podemos calcular la inversa mediante la matriz de cofactores y su traspuesta, es decir, la matriz adjunta. Para generar la matriz de cofactores en DERIVE usamos la instrucción COFACTOR(A,i,j), la cual establece el cofactor correspondiente al elemento que ocupa la fila i y la columna j de la matriz A. Por su parte, la instrucción VECTOR(f(i),i,1,n) crea un vector de la forma $[f(1), f(2), \ldots, f(n)]$.

Los cofactores de A se muestran en la Figura 2.3 en la cual se calculó también la matriz adjunta de A.

Cabe señalar que en DERIVE existe la instrucción ADJOINT(A) que permite calcular directamente $adj(A)$, de tal forma que la inversa de A está dada por la expresión

$$A^{-1} = \frac{1}{\det A} adj(A),$$

la cual pudo haberse calculado directamente por la instrucción A^{-1} en DERIVE .

```
#4:    VECTOR(VECTOR(COFACTOR(A, i, j), j, 3), i, 1, 3)
```

#5:
$$\begin{bmatrix} -\alpha & 2 - \alpha & \alpha^2 - \alpha \\ 2\cdot(\alpha - 1) & \alpha - 2 & 2 - \alpha^2 \\ 4 - 2\cdot\alpha & 0 & \alpha^2 - 2\cdot\alpha \end{bmatrix}$$

Figura 2.3. Cálculo de los cofactores de una matriz

En la Figura 2.4 incluimos los cálculos realizados para obtener la adjunta de la matriz A, la cual corresponde a la traspuesta de la matriz de cofactores, es decir: $adj(A) = (Cof(A))^t$.

$$\#6: \quad \begin{bmatrix} -\alpha & 2\cdot(\alpha - 1) & 4 - 2\cdot\alpha \\ 2 - \alpha & \alpha - 2 & 0 \\ \alpha^2 - \alpha & 2 - \alpha^2 & \alpha^2 - 2\cdot\alpha \end{bmatrix}$$

$$\#7: \quad \mathrm{ADJOINT(A)}$$

$$\#8: \quad \begin{bmatrix} -\alpha & 2\cdot(\alpha - 1) & 4 - 2\cdot\alpha \\ 2 - \alpha & \alpha - 2 & 0 \\ \alpha^2 - \alpha & 2 - \alpha^2 & \alpha^2 - 2\cdot\alpha \end{bmatrix}$$

Figura 2.4. Cálculo de la adjunta de una matriz

Con ello obtenemos que

$$adj(A) = \begin{bmatrix} -\alpha & 2(\alpha - 1) & 4 - 2\alpha \\ 2 - \alpha & \alpha - 2 & 0 \\ \alpha^2 - \alpha & 2 - \alpha^2 & \alpha^2 - 2\alpha \end{bmatrix}.$$

Por último, se muestra el valor de la inversa de la matriz A, para todo $\alpha \neq 2$, que se obtiene usando DERIVE .

$$\#9. \quad \frac{1}{\det(A)}\mathrm{ADJOINT(A)}$$

$$\#10. \quad \begin{bmatrix} -\dfrac{\alpha}{\alpha^2 - 4\alpha + 4} & \dfrac{2(\alpha - 1)}{\alpha^2 - 4\alpha + 4} & \dfrac{2}{2 - \alpha} \\ \dfrac{1}{2 - \alpha} & \dfrac{1}{\alpha - 2} & 0 \\ \dfrac{\alpha(\alpha - 1)}{\alpha^2 - 4\alpha + 4} & \dfrac{2 - \alpha^2}{\alpha^2 - 4\alpha + 4} & \dfrac{\alpha}{\alpha - 2} \end{bmatrix}$$

Ejercicio 2.7.13. Sean $A(x_1, y_1)$, $B(x_2, y_2)$ y $C(x_3, y_3)$ tres puntos del plano xy. Demuestre que el área del triángulo con vértices en A, B y C está dada por

$$A = \left| \frac{1}{2} \det \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} \right|.$$

Solución

Consideremos el triángulo con vértices en A, B y C como se muestra en la Figura 2.5.

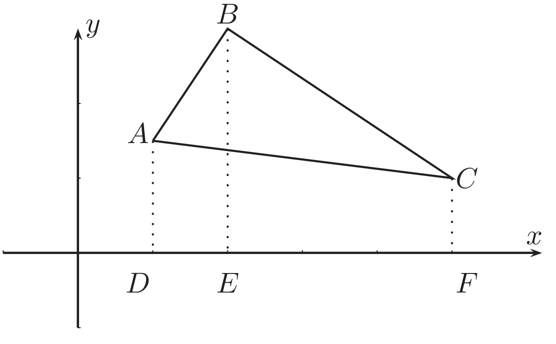


Figura 2.5. Área del triángulo con vértices en A, B y C

El área S, del triángulo ABC, se puede calcular sumando el área del trapecio $DABE$ con el área del trapecio $EBCF$ y restándole el área del trapecio $DACF$. Estas áreas pueden establecerse de la siguiente forma:

$$S = \frac{(x_2 - x_1)(y_2 + y_1)}{2} + \frac{(x_3 - x_2)(y_2 + y_3)}{2} - \frac{(x_3 - x_1)(y_3 + y_1)}{2},$$

la cual puede escribirse en la forma

$$S = \frac{x_2y_1}{2} - \frac{x_1y_2}{2} + \frac{x_3y_2}{2} - \frac{x_2y_3}{2} - \frac{x_3y_1}{2} + \frac{x_1y_3}{2}.$$

Esta última expresión es igual a

$$\left|\frac{1}{2}\det\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}\right|.$$

Ejercicio 2.7.14. Sea la matriz

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix},$$

para la cual $\det A = 3$. Calcular el determinante de las matrices B, C y D definidas de la siguiente manera:

$$B = \begin{bmatrix} a_1 + 2b_1 - 3c_1 & a_2 + 2b_2 - 3c_2 & a_3 + 2b_3 - 3c_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix},$$

$$C = \begin{bmatrix} a_1 & 3a_2 & a_3 \\ b_1 & 3b_2 & b_3 \\ c_1 & 3c_2 & c_3 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$

Solución

La matriz B se obtiene de la matriz A mediante operaciones elementales:

$$\underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}}_{A} \xrightarrow[F_1-3F_3]{F_1+2F_2} \underbrace{\begin{bmatrix} a_1 + 2b_1 - 3c_1 & a_2 + 2b_2 - 3c_2 & a_3 + 2b_3 - 3c_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}}_{B};$$

por tanto, $\det B = \det A = 3$.

Para la matriz C se tiene que

$$\underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}}_{A} \xrightarrow{3C_2} \underbrace{\begin{bmatrix} a_1 & 3a_2 & a_3 \\ b_1 & 3b_2 & b_3 \\ c_1 & 3c_2 & c_3 \end{bmatrix}}_{C}$$

y, en consecuencia, $\det C = 3 \det A = 9$.

Para la matriz D se tiene que

$$\underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}}_{A} \xrightarrow{F_3 \leftrightarrow F_2} \underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \end{bmatrix}}_{D}$$

y, por lo tanto, $\det D = -\det A = -3$.

2.8 Ejercicios propuestos capítulo 2

1. Sea A una matriz de tamaño $n \times n$ antisimétrica. Probar que si n es impar, entonces $\det A = 0$.

2. El determinante de Vandermonde está dado por

$$D_3 = \begin{bmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{bmatrix}.$$

   Demostrar que $D_3 = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2)$.

3. Sean A, B y C matrices de tamaño 3×3 tales que $\det A = -2$, $\det B = 3$ y $\det C = -1$. Calcular $\det(A^t(-3B)^{-1}(4C^{-1}))$.

4. Use la regla de Cramer para resolver los siguientes sistemas de ecuaciones lineales.

a)

$$\begin{aligned} x_1 + x_2 + x_3 &= 8 \\ 4x_2 - x_3 &= 2 \\ 3x_1 - x_2 + 2x_3 &= 0 \end{aligned}$$

b)

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 5 \\ 2x_1 - x_2 - 2x_3 &= -1 \\ x_1 + 3x_2 + x_3 &= 0 \end{aligned}$$

c)

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= -2 \\ 3x_1 + x_2 + 3x_3 &= 6 \\ 2x_1 + x_2 - 2x_3 &= -1. \end{aligned}$$

5. Hallar la inversa de la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a & 1 & 0 & \dots & 0 \\ a^2 & a & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a^n & a^{n-1} & a^{n-2} & \dots & 1 \end{bmatrix}.$$

6. Calcular el determinante de la matriz

$$A = \begin{bmatrix} n & 1 & 1 & 1 & 1 & \dots & 1 \\ n & 2 & 1 & 1 & 1 & \dots & 1 \\ n & 1 & 3 & 1 & 1 & \dots & 1 \\ n & 1 & 1 & 4 & 1 & \dots & 1 \\ \vdots & & & & & & \\ n & 1 & 1 & 1 & 1 & \dots & n \end{bmatrix}.$$

7. Dada la matriz

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

hallar el determinante de dicha matriz a lo largo de los cofactores de:

a) La primera fila.

b) La tercera fila.

c) La segunda columna.

8. Sean $A(-4,2)$, $B(3,5)$ y $C(0,8)$ tres puntos del plano xy. Hallar el área del triángulo con vértices en A, B y C.

9. Dada la matriz

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

calcular su inversa usando la matriz adjunta de A.

10. Determine todos los valores de α para los cuales el sistema dado tiene solución única.

a)

$$\begin{aligned} \alpha x + y + z &= 1 \\ x + \alpha y + z &= \alpha \\ x + y + \alpha z &= \alpha^2 \end{aligned}$$

b)

$$\begin{aligned} \alpha x + y - z &= 10 \\ (\alpha - 2)y + 3z &= 1 \\ 4z &= 10. \end{aligned}$$

11. ¿Para qué valores de α la matriz A no tiene inversa?

$$A = \begin{bmatrix} -\alpha & \alpha - 1 & \alpha + 1 \\ 1 & 2 & 3 \\ 2 - \alpha & \alpha + 3 & \alpha + 7 \end{bmatrix}.$$

12. Determine si los siguientes enunciados son falsos o verdaderos. Justifique la respuesta.

    a) Sean A y B dos matrices de tamaño $n \times n$. Si $\det A = 0$, entonces $\det(AB) = 0$.

    b) Si A es una matriz cuadrada y $A^k = 0$, con $k \in \mathbb{Z}^+$, entonces $\det A = 0$.

    c) Sean A y B dos matrices de tamaño $n \times n$. Entonces se cumple que $\det(A + B) = \det A + \det B$.

    d) Si $A^t = A^{-1}$, entonces $\det A = \pm 1$.

    e) Si A es una matriz cuadrada de tamaño $n \times n$ y c es un número real, entonces $\det(cA) = c \det A$.

    f) Sean A y B dos matrices de tamaño 3×3 tales que $\det A = -2$ y $\det B = 3$; entonces, se cumple que: $\det(A^{-1}(3B^t)(4A)^{-1}) = 107$.

    g) Si A es una matriz no singular, entonces $adj(A)$ es no singular y, además,

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

    h) Si $B = P^{-1}AP$, con P invertible, entonces $\det A = \det B$.

    i) Si $\det A = 5$, entonces el sistema $AX = 0$ tiene como única solución la trivial $X = 0$.

    j) Si $\det(AB) = 0$, entonces A es singular o B es singular.

    k) $\det(BB^t) = \det(B^2)$, donde B es una matriz de tamaño $n \times n$.

    l) Los valores de α para los cuales la matriz

    $$A = \begin{bmatrix} \alpha - 1 & \alpha \\ 1 & \alpha \end{bmatrix}$$

    es no singular son $\alpha = 0$ y $\alpha = 2$.

13. Si A es una matriz de tamaño $n \times n$, demostrar que $\det(adj(A)) = (\det A)^{n-1}$.

14. ¿Para que valores de α se cumple que

$$\det \begin{bmatrix} 2 & 1 & 0 \\ 0 & -1 & 3 \\ 0 & 1 & \alpha \end{bmatrix} + \det \begin{bmatrix} 0 & \alpha & 1 \\ 1 & 3\alpha & 0 \\ -2 & \alpha & 2 \end{bmatrix} = 14?$$

15. Use la definición de determinante para calcular el valor del determinante de las siguientes matrices:

$$A = \begin{bmatrix} 3 & -2 & 1 \\ -5 & 4 & 0 \\ 2 & 1 & 6 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 & 1 \\ 6 & 2 & 1 \\ -2 & 2 & 1 \end{bmatrix}.$$

16. Sea la matriz

$$A = \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{bmatrix}.$$

Calcular su determinante. (Sugerencia: Sea $D_n = \det A$; derive la fórmula $D_n = 2D_{n-1} - D_{n-2}$).

17. Sea la matriz

$$A = \begin{bmatrix} 1 & 2 & -3 & 4 \\ 4 & 8 & 12 & -8 \\ 2 & 3 & 2 & 1 \\ -3 & -1 & 1 & -4 \end{bmatrix}.$$

Calcular su determinante usando las propiedades del determinante, es decir, transformando la matriz A en una matriz triangular superior o inferior.

18. Demuestre que la matriz

$$A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta & 0 \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

es invertible para cualquier valor de θ. Encuentre A^{-1}.

19. Demuestre que $(adj(A))^t = adj(A^t)$.

20. Demuestre que si A es una matriz simétrica, entonces $adj(A)$ es simétrica.

21. Demuestre que si $AA^t = I_n$, entonces $\det A = \pm 1$.

22. Si

$$\det \begin{bmatrix} a & a_1 & a_2 \\ b & b_1 & b_2 \\ c & c_1 & c_2 \end{bmatrix} = k,$$

halle el valor de

$$\det \begin{bmatrix} b+c & b_1+c_1 & b_2+c_2 \\ c+a & c_1+a_1 & c_2+a_2 \\ a+b & a_1+b_1 & a_2+b_2 \end{bmatrix}.$$

23. Obtenga el valor de

$$\det \begin{bmatrix} 3a & 3b & 3c \\ 4a & 4b & 4c \\ c & 2c & 3c \end{bmatrix}.$$

24. Establezca el valor de

$$\det \begin{bmatrix} a & a+2 & a+4 \\ a+1 & a+3 & a+5 \\ a+6 & a+8 & a+10 \end{bmatrix}.$$

25. Si
$$\det\begin{bmatrix} a & a_1 & a_2 \\ b & b_1 & b_2 \\ c & c_1 & c_2 \end{bmatrix} = 2,$$
halle el valor de
$$\det\begin{bmatrix} 2a & 2a_1 & 2a_2 \\ 3b & 3b_1 & 3b_2 \\ -5c & -5c_1 & -5c_2 \end{bmatrix}$$
y de
$$\det\begin{bmatrix} 2a & 2a_1 & 2a_2 \\ a+c & a_1-c_1 & a_2-c_2 \\ -5c & -5c_1 & -5c_2 \end{bmatrix}.$$

26. Sea la matriz
$$A = \begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 2 & 1 & -8 \\ 5 & 2 & 2 & 0 \\ 1 & -1 & 1 & 2 \end{bmatrix}.$$
Calcular su determinante usando las propiedades del determinante, es decir, transformando la matriz A en una matriz triangular superior o inferior.

Capítulo 3

Espacios vectoriales

3.1 Introducción

En este capítulo se pretende formular, de manera abstracta, las diferentes propiedades de las matrices de tamaño $m \times n$, conjuntamente con las operaciones de suma y producto por escalar. Estas dos operaciones se ciñen a un conjunto de axiomas que generalizan las propiedades comunes de las n-tuplas de números reales.

La principal estructura de la que nos ocupamos acá es la de *espacio vectorial*. Este objeto matemático se remonta en la historia hasta el siglo XVII. En ese entonces, Leibniz lo introduce tratando de encontrar objetos algebraicos que permitieran facilitar los cálculos realizados sobre objetos geométricos (origen de la geometría analítica). Sin embargo, hubo que esperar cerca de 200 años hasta que William Rowan Hamilton y Hermann Grassmann presentaran la definición de vector. Un espacio vectorial (o espacio lineal) es el objeto básico de estudio del álgebra lineal. A los elementos de los espacios vectoriales se les llama *vectores*, independientemente de su naturaleza. La primera formulación moderna y axiomática de la teoría de los espacios vectoriales se debe a Giuseppe Peano, a finales del siglo XIX (Boyer, 2001: 709-740).

3.2 Espacios vectoriales

Antes de dar la definición formal de lo que es un espacio vectorial es necesario recordar las nociones de *campo*, *operación binaria interna* y de *operación binaria externa*. Estos conceptos hacen parte de la axiomatización correspondiente a los espacios vectoriales.

Definición 3.2.1. *Un* campo *es un conjunto K en el que se han definido dos operaciones, $+$ y $*$, llamadas suma y multiplicación respectivamente, que cumplen las siguientes propiedades:*

1. *K es cerrado para la suma y la multiplicación. Para todo a, b en K, $a+b$ y $a*b$ pertenecen a K.*

2. *Asociatividad para la suma y la multiplicación. Para todo a, b en K, $a+(b+c)=(a+b)+c$ y $a*(b*c)=(a*b)*c$.*

3. *Conmutatividad de la suma y la multiplicación. Para a, b en K, $a+b=b+a$ y $a*b=b*a$.*

4. *Existencia de un elemento neutro para la suma y la multiplicación. Existe un elemento 0 en K, tal que para todo a en K, $a+0=a$. Existe un elemento 1 en K diferente a 0, tal que para todo a en K, $a*1=a$.*

5. *Existencia de elemento opuesto y de inversos. Para cada a en K, existe un elemento $-a$ en K, tal que $a+(-a)=0$. Para cada $a\neq 0$ en K, existe un elemento a^{-1} en K, tal que $a*a^{-1}=1$.*

6. *Distributividad de la multiplicación respecto de la suma. Para todo a, b, c en K, $a*(b+c)=(a*b)+(a*c)$.*

El requisito $a\neq 0$ en la quinta propiedad asegura que el conjunto que contiene solamente un cero no sea un *cuerpo* o *campo*, y de paso elimina la posibilidad de que en el cuerpo existan divisores de cero distintos de 0. Entre los ejemplos de *campos* o *cuerpos* más comunes encontramos el conjunto de los números reales $\mathbb{R}$, el de los números complejos $\mathbb{C}$ y el de los números racionales $\mathbb{Q}$. Estos campos tienen en común, aparte de su estructura algebraica, que poseen un número infinito de términos.

El estudio de los campos con un número finito de términos se inició con los trabajos de Évariste Galois (1811-1832.) En el año de 1893, Eliakim Hastings Moore (1862-1932) demostró que los campos finitos son todos del tipo de los estudiados por Galois.

Una operación binaria interna en un conjunto no vacío V es una relación de la forma

$$\oplus : V \times V \to V;$$

esto significa que la combinación de dos objetos de V genera un objeto del conjunto V. De forma semejante, una operación binaria externa es una relación de la forma

$$\odot : \mathbb{K} \times V \to V$$

y su significado es que la combinación de un elemento del campo $\mathbb{K}$ con un elemento de V es nuevamente un elemento de V.

Definición 3.2.2. *Un espacio vectorial sobre un campo $\mathbb{K}$ es un conjunto de objetos V, no vacío, con una operación binaria interna $\oplus$, llamada suma, y una operación binaria externa $\odot$, llamada producto por escalar, que satisfacen las propiedades que se enuncian a continuación:*

Para el caso de la suma:

1. *Es asociativa, es decir, para cualquier u, v y w,*

$$(u \oplus v) \oplus w = u \oplus (v \oplus w).$$

2. *Tiene elemento neutro, esto es, existe un elemento $\boldsymbol{0}$ tal que para todo u en V,*

$$u \oplus \boldsymbol{0} = u.$$

3. *Tiene elemento inverso, esto es, para cada elemento u existe un elemento −u tal que*

$$u \oplus -u = \boldsymbol{0}.$$

4. *Es conmutativa, es decir, para cualquier u y v en V,*

$$u \oplus v = v \oplus u.$$

En el caso del producto por escalar, se tiene:

1. *Es distributivo, es decir, para todo escalar α y u, y v en V, se cumple que*
$$\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot .v$$

2. *Es distributivo bajo la suma por escalar, es decir, para escalares α y β, y u en V, se cumple que*

$$(\alpha + \beta) \odot u = \alpha \odot u \oplus \beta \odot u.$$

3. *Es asociativo: para escalares α y β, y u en V, se cumple que*

$$\alpha \odot (\beta \odot u) = (\alpha\beta) \odot u.$$

4. *Tiene neutro para la multiplicación por escalar, es decir, para cualquier u en V,*
$$1 \odot u = u.$$

Si el conjunto de los escalares es $\mathbb{R}$, se dice de V que se trata de un *espacio vectorial real*; si es $\mathbb{C}$, el espacio V es un *espacio vectorial complejo.*

En adelante, para representar a los escalares se utilizan letras griegas minúsculas, y para los vectores, letras latinas minúsculas; además, para representar la suma de vectores $u \oplus v$, utilizamos $u+v$ y λa en vez de $\lambda \odot a$.

Las propiedades algebraicas de los elementos de un espacio vectorial arbitrario son muy semejantes a las de los vectores libres en $\mathbb{R}^3$. Como consecuencia, también se acostumbra a llamar *vectores* a los elementos de un espacio vectorial arbitrario. En los axiomas que hemos presentado debe quedar claro que las operaciones definidas en V deben definir nuevos elementos en V; estas dos condiciones se presentan con el nombre de *leyes de cerradura o de clausura.*

Ejemplo 3.2.1 (El conjunto de las n-tuplas). Sea $\mathbb{K}$ un campo y sea V el conjunto formado por todas las n-tuplas

$$a = (\alpha_1, \ldots, \alpha_n),$$

donde los α_i, con $i = 1, \ldots n$ son números en el campo $\mathbb{K}$.

Dos elementos $a = (\alpha_1, \ldots, \alpha_n)$ y $b = (\beta_1, \ldots, \beta_n)$ se consideran iguales si $\alpha_i = \beta_i$, con $i = 1 \ldots n$. Además, su suma y producto por escalar quedan definidos por

$$\begin{aligned} a + b &= (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \\ \lambda a &= (\lambda\alpha_1, \ldots, \lambda\alpha_n), \qquad \lambda \in \mathbb{K}. \end{aligned}$$

Veamos que las operaciones dadas satisfacen los axiomas de espacio vectorial.

Para el caso de la suma se tiene:

1. Es asociativa, es decir, para cualquier a, b y c,

$$\begin{aligned} (a + b) + c &= (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) + (\psi_1, \ldots, \psi_n) \\ &= (\alpha_1 + \beta_1 + \psi_1, \ldots, \alpha_n + \beta_n + \psi_n) \\ &= (\alpha_1, \ldots, \alpha_n) + (\beta_1 + \psi_1, \ldots, \beta_n + \psi_n) \\ &= a + (b + c). \end{aligned}$$

2. Tiene elemento neutro, esto es, existe un elemento $\mathbf{0}$ tal que, para todo a en V,

$$a + \mathbf{0} = (\alpha_1 + 0, \ldots, \alpha_n + 0) = a.$$

3. Tiene elemento inverso, esto es, para cada elemento a existe un elemento $-a$ tal que

$$a + (-a) = (\alpha_1 - \alpha_1, \ldots, \alpha_n - \alpha_n) = (0, \ldots, 0) = \mathbf{0}.$$

4. Es conmutativa, es decir, para cualquier a y b en V,

$$\begin{aligned} a + b &= (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \\ &= (\beta_1 + \alpha_1, \ldots, \beta_n + \alpha_n) \\ &= b + a. \end{aligned}$$

En el caso del producto por escalar, se tiene:

1. Es distributiva, es decir, para todo escalar α y los vectores a y b en V, se cumple que

$$\begin{aligned} \alpha(a + b) &= \alpha(\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \\ &= (\alpha\alpha_1 + \alpha\beta_1, \ldots, \alpha\alpha_n + \alpha\beta_n) \\ &= (\alpha\alpha_1, \ldots, \alpha\alpha_n) + (\alpha\beta_1, \ldots, \alpha\beta_n) \\ &= \alpha a + \alpha b. \end{aligned}$$

2. Es distributiva bajo la suma por escalar, es decir, para escalares α y β, y a en V, se cumple que

$$\begin{aligned} (\alpha + \beta)a &= ((\alpha + \beta)\alpha_1, \ldots, (\alpha + \beta)\alpha_n) \\ &= (\alpha\alpha_1, \ldots, \alpha\alpha_n) + (\alpha\beta_1, \ldots, \alpha\beta_n) \\ &= \alpha a + \beta a. \end{aligned}$$

3. Es asociativa: para escalares α y β, y a en V, se cumple que

$$\begin{aligned} \alpha(\beta a) &= \alpha\left(\beta\alpha_1, \ldots, \beta\alpha_n\right) \\ &= (\alpha\beta\alpha_1, \ldots, \alpha\beta\alpha_n) \\ &= (\alpha\beta)a. \end{aligned}$$

4. Tiene neutro para la multiplicación por escalar, es decir, para cualquier a en V,

$$\begin{aligned} 1a &= (1\alpha_1, \ldots, 1\alpha_n) \\ &= (\alpha_1, \ldots, \alpha_n) \\ &= a. \end{aligned}$$

Con estos resultados es posible concluir que el conjunto de n-tuplas de números en $\mathbb{K}$, con las operaciones definidas previamente, es un espacio vectorial sobre el campo $\mathbb{K}$. Este espacio vectorial se representa por $\langle \mathbb{K}^n, +, . \rangle$ o simplemente $\mathbb{K}^n$. En el caso en el cual el campo es el de los números reales nos referimos a este espacio vectorial como $\mathbb{R}^n$.

Ejemplo 3.2.2 (El espacio vectorial de los números complejos). Sea el conjunto de los números de la forma $z = a + bi$, donde a y b son números reales e $i = \sqrt{-1}$. Estos objetos constituyen el conjunto de los *números complejos* y se denota con $\mathbb{C}$.

En el número complejo $z = a + bi$, el número a se denomina *parte real* de z mientras que b es la *parte imaginaria* de z. Si la parte imaginaria de z es cero, el número es un número real, mientras que si la parte real es nula, el número se conoce como un *imaginario puro*.

Dos números complejos son iguales si sus partes reales e imaginarias lo son respectivamente; es decir, si $z = a + bi$ y $w = c + di$, entonces $z = w$ si y solo si $a = c$ y $b = d$.

La suma y el producto por escalar quedan definidos así:

$$\begin{aligned} z + w &= (a + c) + (b + d)i \\ \alpha z &= \alpha a + \alpha bi \quad \alpha \in \mathbb{R}. \end{aligned}$$

Con estas definiciones, el conjunto de los números complejos constituye un espacio vectorial.

La primera interpretación útil de los números complejos se debe al topógrafo noruego Caspar Wessel (1745-1818) en el año 1797. Wessel presenta de manera formal lo que ahora conocemos, erróneamente, como el diagrama de Argand, que no es más que una representación cartesiana de los números complejos. Infortunadamente, el trabajo de Wessel fue publicado en una revista muy poco consultada por los matemáticos de la época y pasó desapercibido hasta 1897 (Bell, 2002: 186-188). Este hecho permitió que fuese Carl F. Gauss quien hiciera la formalización de los números complejos.

Ejemplo 3.2.3 (El espacio vectorial de las matrices). Dadas las matrices $A \in M_{m\times n}$ y $B \in M_{m\times n}$, su suma, que denotamos con $A + B$, es una matriz $M_{m\times n}$, cuyos elementos son:

$$\begin{aligned}
A + B &= (a_{ij}) + (b_{ij}) \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}.
\end{aligned}$$

Es decir, la suma de matrices se puede ver como una función

$$\begin{aligned}
+ \ : \ & M_{m\times n} \times M_{m\times n} \longrightarrow M_{m\times n} \\
& (A, B) \longrightarrow A + B
\end{aligned}$$

A su vez, llamamos producto de un escalar λ por una matriz

$A = (a_{ij})_{m\times n}$ a la nueva matriz dada por

$$\lambda A = \lambda \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{bmatrix};$$

es decir, el producto de una matriz por un escalar se puede ver como una función

$$\begin{aligned} . \ : \ \mathbb{R} \times M_{m\times n} &\longrightarrow M_{m\times n} \\ (\lambda, A) &\longrightarrow \lambda A. \end{aligned}$$

Se ha probado en la sección 1.3, dedicada a las matrices, que estas operaciones cumplen los axiomas de un espacio vectorial.

Ejemplo 3.2.4 (**El conjunto de las funciones polinómicas**). Una función polinómica, en la variable x y con coeficientes en el campo $\mathbb{K}$, es una aplicación $p : \mathbb{K} \to \mathbb{K}$ definida para todo x mediante

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0,$$

donde los a_i, con $i = 0, \ldots n$, son números en el campo $\mathbb{K}$.

Si $a_n \neq 0$, se dice que el polinomio tiene grado n.

Si

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

y

$$q(x) = b_n x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0$$

son dos polinomios, $p(x)$ y $q(x)$ se consideran iguales si tienen el mismo grado y $a_i = b_i$, con $i = 0 \ldots n$.

Además, su suma y producto por escalar quedan definidos por

$$\begin{aligned} p(x) + q(x) &= (a_n + b_n)x^n + \ldots + (a_1 + b_1)x + (a_0 + b_0) \\ \alpha q(x) &= (\alpha a_n)x^n + \ldots + (\alpha a_1)x + (\alpha a_0). \end{aligned}$$

En la definición de suma de polinomios no es necesario que estos tengan el mismo grado; en general, si $p(x)$ tiene grado n y $q(x)$ tiene grado m, con $n > m$, entonces la suma tendrá la forma:

$$p(x)+q(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + (a_m + b_m)x^m + \ldots + (a_0 + b_0).$$

Veamos que, con estas definiciones, el conjunto de los polinomios de grado menor o igual a n, con las operaciones dadas, satisfacen los axiomas de espacio vectorial.

Con respecto a la suma:

1. Es asociativa, es decir, para cualesquier $p(x)$, $q(x)$ y $r(x)$,

$$\begin{aligned}(p(x)+q(x))+r(x) &= ((a_n+b_n)+c_n)x^n + \ldots + ((a_0+b_0)+c_0) \\ &= (a_n+(b_n+c_n))x^n + \ldots + (a_0+(b_0+c_0)) \\ &= p(x)+(q(x)+r(x)).\end{aligned}$$

2. Tiene elemento neutro, esto es, existe un elemento $\mathbf{0}$ tal que para todo $p(x)$ en V,

$$p(x)+\mathbf{0} = (a_n+0)x^n + \ldots + (a_1+0)x + (a_0+0) = p(x).$$

3. Tiene elemento inverso, esto es, para cada elemento $p(x)$ existe un elemento $-p(x)$, tal que

$$p(x)+(-p(x)) = (a_n-a_n)x^n + \ldots + (a_1-a_1)x + (a_0-a_0) = \mathbf{0}.$$

4. Es conmutativa, es decir, para cualquier $p(x)$ y $q(x)$ en V,

$$\begin{aligned}p(x)+q(x) &= (a_n+b_n)x^n + \ldots + (a_1+b_1)x + (a_0+b_0) \\ &= (b_n+a_n)x^n + \ldots + (b_1+a_1)x + (b_0+a_0) \\ &= q(x)+p(x).\end{aligned}$$

Para el producto por escalar se tiene:

1. Es distributivo, es decir, para todo escalar α, y $p(x)$ y $q(x)$ en V, se cumple que

$$\begin{aligned}
\alpha(p(x)+q(x)) &= (\alpha(a_n+b_n))x^n+\ldots+(\alpha(a_0+b_0)) \\
&= (\alpha a_n)x^n+\ldots+(\alpha a_0)+(\alpha b_n)x^n+\ldots+(\alpha b_0) \\
&= \alpha p(x)+\alpha q(x).
\end{aligned}$$

2. Es distributivo bajo la suma de escalares, es decir, para escalares α y β, y $p(x)$ en V, se cumple que

$$\begin{aligned}
(\alpha+\beta)p(x) &= ((\alpha+\beta)a_n)x^n+\ldots+((\alpha+\beta)a_1)x+((\alpha+\beta)a_0) \\
&= (\alpha a_n)x^n+\ldots+(\alpha a_1)x+(\alpha a_0) \\
&\quad +(\beta a_n)x^n+\ldots+(\beta a_1)x+(\beta a_0) \\
&= \alpha p(x)+\beta p(x).
\end{aligned}$$

3. Es asociativo: para escalares α y β, y $p(x)$ en V, se cumple que

$$\begin{aligned}
\alpha(\beta p(x)) &= \alpha\left((\beta a_n)x^n+\ldots+(\beta a_1)x+(\beta a_0)\right) \\
&= (\alpha(\beta a_n))x^n+\ldots+(\alpha(\beta a_1))x+(\alpha(\beta a_0)) \\
&= (\alpha\beta a_n)x^n+\ldots+(\alpha\beta a_1)x+(\alpha\beta a_0) \\
&= (\alpha\beta)p(x).
\end{aligned}$$

4. Tiene neutro para la multiplicación por escalar, es decir, para cualquier $p(x)$ en V,

$$\begin{aligned}
1p(x) &= (1a_n)x^n+\ldots+(1a_1)x+(1a_0) \\
&= a_nx^n+\ldots+a_1x+a_0 \\
&= p(x).
\end{aligned}$$

Los resultados anteriores permiten concluir que el conjunto de las funciones polinómicas de grado menor o igual a n, con las operaciones definidas, es un espacio vectorial sobre el campo $\mathbb{K}$.

Es importante recalcar que si consideramos los polinomios de grado n exclusivamente, el conjunto no constituye un espacio vectorial, ya que no cumple la cerradura para la suma, es decir, no siempre la suma de dos polinomios de grado n es nuevamente un polinomio de grado n. Por ejemplo, $ax^n + (-ax^n) = 0$, el cual no es un polinomio de grado n.

Ejemplo 3.2.5. Considere el conjunto de elementos de $\mathbb{R}^3$ dado por $V = \{(x, y, z) : 2x + 3y - z = 0\}$, con las operaciones usuales de $\mathbb{R}^3$. Determine si V es un espacio vectorial.

Solución

Sean $u = (x_1, y_1, z_1)$ y $v = (x_2, y_2, z_2)$ elementos de V; entonces, se cumple que

$$2x_1 + 3y_1 - z_1 = 0 \quad \text{y} \quad 2x_2 + 3y_2 - z_2 = 0.$$

Ahora, vamos a verificar que $u + v \in V$. Como

$$u + v = (x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2).$$

Pero

$$\begin{aligned} 2(x_1 + x_2) + 3(y_1 + y_2) - (z_1 + z_2) &= \underbrace{(2x_1 + 3y_1 - z_1)}_{0} \\ &\quad + \underbrace{(2x_2 + 3y_2 - z_2)}_{0} = 0 \end{aligned}$$

y, por lo tanto,

$$2(x_1 + x_2) + 3(y_1 + y_2) - (z_1 + z_2) = 0,$$

es decir, $u + v \in V$.

Sea ahora $w = (x_3, y_3, z_3)$. Como

$$\begin{aligned} (u + v) + w &= (x_1 + x_2, y_1 + y_2, z_1 + z_2) + (x_3, y_3, z_3) \\ &= ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3, (z_1 + z_2) + z_3) \in V \\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3), z_1 + (z_2 + z_3)) \in V. \end{aligned}$$

Por otro lado,

$$\begin{aligned} u+(v+w) &= (x_1,y_1,z_1)+(x_2+x_3,y_2+y_3,z_2+z_3) \\ &= (x_1+(x_2+x_3), y_1+(y_2+y_3), z_1+(z_2+z_3)) \in V. \end{aligned}$$

De este modo tenemos

$$(u+v)+w = u+(v+w),$$

es decir, la suma es asociativa.

Al ser

$$\begin{aligned} u+v &= (x_1,y_1,z_1)+(x_2,y_2,z_2) \\ &= (x_1+x_2, y_1+y_2, z_1+z_2) \\ &= (x_2+x_1, y_2+y_1, z_2+z_1) \\ &= v+u, \end{aligned}$$

con ello se prueba la conmutatividad.

Veamos ahora que si $u \in V$, entonces $u+0 \in V$. Como

$$u+0 = (x_1,y_1,z_1)+(0,0,0) = (x_1+0, y_1+0, z_1+0),$$

entonces

$$2(x_1+0)+3(y_1+0)-(z_1+0) = \underbrace{2x_1+3y_1-z_1}_{0} = 0.$$

Luego $u+0 \in V$.

Probemos ahora que cada elemento de V posee un inverso aditivo:

$$u+(-u) = (x_1,y_1,z_1)+(-x_1,-y_1,-z_1) = (0,0,0) = 0.$$

Claramente cumple la ecuación

$$2(0)+3(0)-(0) = 0.$$

De manera similar podemos ver que $cu \in V$, donde $cu = (cx_1, cy_1, cz_1)$, puesto que

$$2(cx_1) + 3(cy_1) - (cz_1) = c\underbrace{(2x_1 + 3y_1 - z_1)}_{0} = c0 = 0.$$

Para las leyes distributivas:

$$\begin{aligned}
c(u+v) &= c(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\
&= (cx_1 + cx_2, cy_1 + cy_2, cz_1 + cz_2) \\
&= (c(x_1, y_1, z_1) + c(x_2, y_2, z_2)) \\
&= cu + cv.
\end{aligned}$$

Para el caso de la otra ley distributiva, tenemos:

$$\begin{aligned}
(c+d)u &= (c+d)(x_1, y_1, z_1) \\
&= ((c+d)x_1, (c+d)y_1, (c+d)z_1) \\
&= (cx_1 + dx_1, cy_1 + dy_1, cz_1 + dz_1) \\
&= (cx_1, cy_1, cz_1) + (dx_1, dy_1, dz_1) \\
&= c(x_1, y_1, z_1) + d(x_1, y_1, z_1) \\
&= cu + du.
\end{aligned}$$

En el caso de la asociatividad para la multiplicación por escalares tenemos:

$$c(du) = c(dx_1, dy_1, dz_1) = (cdx_1, cdy_1, cdz_1) = (cd)u.$$

Nuevamente, esto es verdad si

$$2(cdx_1) + 3(cdy_1) - (cdz_1) = 0;$$

pero, $2(cdx_1) + 3(cdy_1) - (cdz_1) = cd(2x_1 + 3y_1 - z_1) = 0.$

Por último,

$$1u = 1(x_1, y_1, z_1) = (1x_1, 1y_1, 1z_1) = u,$$

es decir, existe un neutro para la multiplicación por escalar.

Con esto queda probado que $V = \{(x, y, z) : \ 2x + 3y - z = 0\}$ es un espacio vectorial real.

El siguiente teorema resume algunas de las principales propiedades de los espacios vectoriales.

Teorema 3.2.1. *En un espacio vectorial sobre un campo* $\mathbb{K}$, *para cualesquiera que sean los vectores* u *y* v, *y los escalares* λ *y* ψ, *se cumple que:*

1. $0u = \mathbf{0}$.
2. $\lambda\mathbf{0} = \mathbf{0}$.
3. *Si* $\lambda u = \mathbf{0}$, *entonces* $\lambda = 0$ *o* $u = \mathbf{0}$.
4. $(-\lambda)u = -\lambda u$.
5. *Si* $\lambda u = \lambda v$ *y* $\lambda \neq 0$, *entonces* $u = v$.
6. *Si* $\lambda u = \psi u$ *y* $u \neq \mathbf{0}$, *entonces* $\lambda = \psi$.
7. $\mathbf{0}$ *y* $-a$ *son únicos.*

Demostración. Sean los vectores u y v y los escalares λ y ψ.

1. $0u = (0 + 0)u = 0u + 0u$, luego $0u = 0$.
2. $\lambda\mathbf{0} = \lambda(\mathbf{0} + \mathbf{0}) = \lambda\mathbf{0} + \lambda\mathbf{0}$, luego $\lambda\mathbf{0} = \mathbf{0}$.
3. Si fuese $\lambda \neq 0$, entonces existiría λ^{-1} y se tendría que $\lambda^{-1}(\lambda u) = \lambda^{-1}\mathbf{0} = \mathbf{0}$. Pero $\lambda^{-1}(\lambda u) = (\lambda^{-1}\lambda)u = u$; en consecuencia, $u = \mathbf{0}$.
4. $(-\lambda)u + \lambda u = (-\lambda + \lambda)u = 0u = \mathbf{0}$; luego $-\lambda u = (-\lambda)u$.
5. Si $\lambda u = \lambda v$, entonces $\lambda u - \lambda v = \lambda(u - v) = \mathbf{0}$. Como $\lambda \neq 0$, por el numeral 3, $u - v = 0$, es decir, $u = v$.
6. Si $\lambda u = \psi u$ y $u \neq \mathbf{0}$, entonces $\lambda u - \psi u = \mathbf{0} = (\lambda - \psi)u$; en consecuencia, $\lambda - \psi = 0$, es decir, $\lambda = \psi$.

La última propiedad se deja como ejercicio.

3.3 Subespacios vectoriales

En esta sección se establecen algunas propiedades de los espacios vectoriales. Para iniciar, se toman subconjuntos de un espacio vectorial, con las operaciones del espacio, que a su vez poseen las propiedades que caracterizan a los espacios vectoriales. De esta forma tenemos la siguiente definición.

Definición 3.3.1. *Sea V un espacio vectorial y $W \neq \phi$ un subconjunto de V. Si W es un espacio vectorial con respecto a las operaciones definidas en V, entonces W es un subespacio de V.*

Ejemplo 3.3.1. Todo espacio vectorial V tiene al menos dos subespacios, él mismo y el *subespacio cero* formado por $\{\mathbf{0}\}$ y se denominan *subespacios triviales* de V.

Ejemplo 3.3.2. Sea $W = \{(a, b, 0) : a, b \in \mathbb{R}\}$. Para determinar si W es un subespacio de $\mathbb{R}^3$, se debe verificar que W cumple con todas las propiedades que definen un espacio vectorial. En particular, es necesario ver que es cerrado para las dos operaciones definidas en $\mathbb{R}^3$, esto es, si $u = (a_1, b_1, 0)$ y $v = (a_2, b_2, 0)$, entonces $u + v = (a_1 + a_2, b_1 + b_2, 0)$ es nuevamente un elemento de W. De forma similar, si λ es un escalar, $\lambda u = \lambda(a_1, b_1, 0) = (\lambda a_1, \lambda b_1, 0)$, que es un elemento de W.

Las demás propiedades que caracterizan un espacio vectorial no es necesario verificarlas, puesto que los elementos de W son, a su vez, elementos de $\mathbb{R}^3$ y, en consecuencia, heredan todas estas propiedades.

El ejemplo anterior permite ver que para determinar si un subconjunto de un espacio vectorial es un subespacio, no es necesario verificar todas las propiedades, puesto que hay propiedades que las operaciones heredan del espacio original. Al fin y al cabo, las operaciones con elementos del subconjunto son las mismas del espacio vectorial original. En definitiva, W es subespacio si para todos los elementos u y v en W, el vector $u + v$ está en W; el vector $\mathbf{0}$ está en W, al igual que los vectores $-u$ y λu. Este comentario se puede resumir en el siguiente teorema.

Teorema 3.3.1. *Sea V un espacio vectorial y W un subconjunto no vacío de V. W es un subespacio de V si y solo si para cualquier par de vectores u y v en W, y el escalar λ, los vectores $u + v$ y λu están en W.*

Demostración. Supóngase que W es un subconjunto no vacío de V tal que, para cualquier par de vectores u y v en W, y el escalar λ, los vectores $u+v$ y λu están en W. Entonces, como para todo vector v en W, λV está en W, con $\lambda = -1$, tenemos que al ser $-v = (-1)v$, luego $-v$ está en W. Además, para todo v en W existe $-v$, tal que $v + (-v) = 0$ está en W; por lo tanto, W es un subespacio.

El recíproco es inmediato, ya que si W es subespacio para cualquier par de vectores u y v en W, y el escalar λ, los vectores $u + v$ y λu están en W.

A continuación se presentan dos formas diferentes de obtener subespacios: la primera permite obtener un subespacio a partir de otros conocidos; la segunda muestra la construcción de un subespacio teniendo como punto de partida un conjunto arbitrario de vectores de un espacio vectorial dado.

Teorema 3.3.2. *Si S y T son dos subespacios de un espacio vectorial V, entonces su intersección, $S \cap T$, es subespacio de V.*

Demostración. Sean u y v elementos de $S \cap T$ y λ un escalar. Entonces, u y v están en S; y como S es un subespacio, los vectores $u + v$ y λu están en S. De la misma forma, los vectores $u + v$ y λu están en T. Con ello se concluye que los vectores $u + v$ y λu están en $S \cap T$, y por el teorema 3.3.2 se concluye que $S \cap T$ es un subespacio.

Ejemplo 3.3.3. Sea el espacio vectorial $\mathbb{R}^3$. Si S es el conjunto de vectores de la forma $(\alpha, \beta, 0)$ y los vectores de T son de la forma $(\psi, 0, \lambda)$, es fácil verificar que S y T son subespacios vectoriales. El subespacio $S \cap T$ es el conjunto de vectores de la forma $(\phi, 0, 0)$.

Definición 3.3.2. *Sean v_1, $v_2, \ldots, v_k$ vectores en un espacio vectorial. Un vector v en V es una combinación lineal de v_1, $v_2, \ldots, v_k$ si*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

para ciertos escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$.

Ejemplo 3.3.4. Probar que el vector $(5, 2)$ es una combinación lineal de los vectores $(-1, 1)$ y $(1, 1)$.

Solución

En este caso, se trata de resolver la ecuación vectorial

$$(5, 2) = \alpha_1(-1, 1) + \alpha_2(1, 1),$$

la cual es equivalente al sistema de ecuaciones

$$\begin{aligned} -\alpha_1 + \alpha_2 &= 5 \\ \alpha_1 + \alpha_2 &= 2. \end{aligned}$$

Al sumar estas dos ecuaciones se obtiene que $2\alpha_2 = 7$, con lo cual

$$\alpha_2 = \frac{7}{2}, \quad \alpha_1 = 2 - \frac{7}{2} = -\frac{3}{2},$$

esto es:

$$(5, 2) = -\frac{3}{2}(-1, 1) + \frac{7}{2}(1, 1),$$

que es la combinación lineal pedida.

Ejemplo 3.3.5. El vector $(5,1)$ de $\mathbb{R}^2$ es una combinación lineal de los vectores $(1,-1)$ y $(1,2)$, puesto que $(5,1) = 3(1,-1) + 2(1,2)$. Los escalares de la combinación lineal son los números 3 y 2 respectivamente, y suelen llamarse las *coordenadas* del vector $(5,1)$ respecto a los vectores $(1,-1)$ y $(1,2)$. Esta situación queda ilustrada en la Figura 3.1.

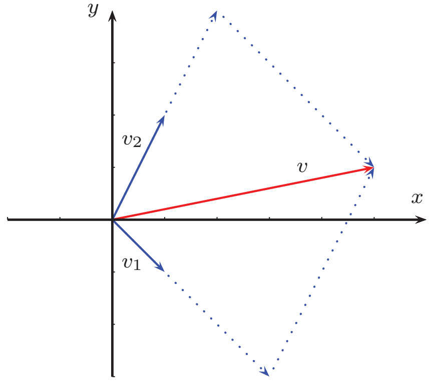


Figura 3.1. Combinación lineal de los vectores $(1,-1)$ y $(1,2)$

Definición 3.3.3. *Si $S = \{v_1, v_2, \ldots, v_k\}$ es un conjunto de vectores de un espacio vectorial V, entonces el conjunto generado por S, es el conjunto de todos los vectores que son combinaciones lineales de los vectores de S y se denota con $gen\{S\}$.*

De la definición se concluye que si $S \neq \phi$ es un subconjunto de un espacio vectorial V y $S \subseteq T$, entonces $gen\{S\} \subseteq gen\{T\}$.

El siguiente ejemplo permite ver la interpretación geométrica del conjunto generado por un par de vectores del espacio tridimensional.

Ejemplo 3.3.6. Sea el espacio de los vectores libres de $\mathbb{E}^3$, es decir, todo vector del espacio que tienen las mismas características: igual módulo, dirección y sentido. Sean los vectores libres $\overrightarrow{v}_1$ y $\overrightarrow{v}_2$;

entonces, la combinación lineal $\overrightarrow{v} = \alpha_1\overrightarrow{v}_1 + \alpha_2\overrightarrow{v}_2$ es el vector que se observa en la Figura 3.2.

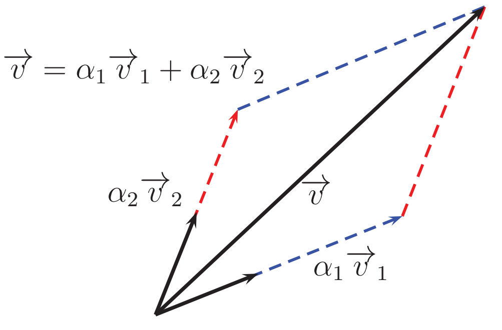


Figura 3.2. Combinación lineal de dos vectores

En el siguiente teorema se muestra la principal propiedad del conjunto de todas las combinaciones lineales de un subconjunto de vectores de un espacio vectorial.

Teorema 3.3.3. *Sea $S = \{v_1, v_2, \ldots, v_k\}$ un conjunto de vectores en un espacio vectorial V. Entonces, $gen\{S\}$ es un subespacio de V.*

Demostración. Sea $S = \{v_1,\ v_2, \ldots, v_k\}$ y sean los vectores v y u en $gen\{S\}$; entonces, si

$$u = \alpha_1 v_1 + \ldots + \alpha_k v_k \quad \text{y} \quad v = \beta_1 v_1 + \ldots + \beta_k v_k,$$

se tiene que

$$\begin{aligned} u + v &= (\alpha_1 + \beta_1)v_1 + \ldots + (\alpha_k + \beta_k)v_k \quad \text{y} \\ \lambda u &= (\lambda\alpha_1)v_1 + \ldots + (\lambda\alpha_k)v_k. \end{aligned}$$

De esta manera, $u + v$ y λu están en $gen\{S\}$, ya que ambos son también combinaciones lineales de los vectores de S. Luego $gen\{S\}$ es un subespacio vectorial.

En el siguiente ejemplo se muestra una interpretación geométrica del espacio generado por un vector en el espacio $\mathbb{R}^2$.

Ejemplo 3.3.7. Sea el vector $v = (3, 4)$ en el espacio $\mathbb{R}^2$; entonces, el subespacio generado por v, $gen\{v\}$, es el conjunto de vectores de la forma $\lambda(3, 4)$ (véase la Figura 3.3).

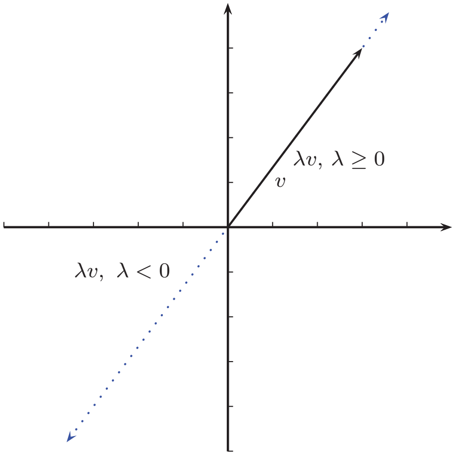


Figura 3.3. Subespacio generado por el vector $(3, 4)$

De esta forma, si $a = (\alpha, \beta)$ es un vector de $gen\{v\}$, entonces

$$(\alpha, \beta) = (3\lambda, 4\lambda),$$

lo cual implica que

$$gen\{v\} = \left(\alpha, \frac{4\alpha}{3}\right).$$

Geométricamente, esto corresponde a una recta que pasa por el origen y tiene pendiente $4/3$.

Ejemplo 3.3.8. Sean los vectores $a = (1, 1)$ y $b = (2, 3)$ de $\mathbb{R}^2$. Probar que el subespacio generado por estos dos vectores es el mismo $\mathbb{R}^2$, es decir, $gen\{a, b\} = \mathbb{R}^2$.

Solución
Por la definición de subespacio generado, $gen\{a, b\} \subseteq \mathbb{R}^2$.

Ahora se probará la inclusión contraria. Sea el vector $c = (\alpha, \beta)$ de $\mathbb{R}^2$; lo que se debe mostrar es que este vector puede escribirse como una combinación lineal de los vectores a y b, esto es, que existen escalares x y y tales que

$$(\alpha, \beta) = x(1, 1) + y(2, 3).$$

A partir de esta relación se obtiene el sistema de ecuaciones

$$\begin{aligned} x + 2y &= \alpha \\ x + 3y &= \beta, \end{aligned}$$

del cual se concluye que

$$\begin{aligned} x &= 3\alpha - 2\beta \\ y &= \beta - \alpha. \end{aligned}$$

Así, si (α, β) es un vector de $\mathbb{R}^2$, existen escalares $x = 3\alpha - 2\beta$ y $y = \beta - \alpha$ tales que

$$(\alpha, \beta) = (3\alpha - 2\beta)(1, 1) + (\beta - \alpha)(2, 3),$$

con lo cual $\mathbb{R}^2 \subseteq gen\{a, b\}$ y, en consecuencia, $gen\{a, b\} = \mathbb{R}^2$.

Teorema 3.3.4. *Sea V un espacio vectorial y sea $\{v_1, \ldots, v_n\} \neq \phi$. Un vector $u \in gen\{v_1, \ldots, v_n\}$ si y solo si $gen\{v_1, \ldots, v_n\} = gen\{v_1, \ldots, v_n, u\}$.*

Demostración. Sea u tal que $u \in gen\{v_1, \ldots, v_n\}$. Como

$$\{v_1, \ldots, v_n\} \subset \{v_1, \ldots, v_n, u\},$$

se concluye que

$$gen\{v_1, \ldots, v_n\} \subseteq gen\{v_1, \ldots, v_n, u\}.$$

Sea ahora $y \in gen\{v_1, \ldots, v_n, u\}$,

$$y = \alpha_1 v_1 + \ldots + \alpha_n v_n + \gamma u;$$

como $u \in gen\{v_1, \ldots, v_n\}$, entonces

$$u = \beta_1 v_1 + \ldots + \beta_n v_n;$$

en consecuencia,

$$y = \alpha_1 v_1 + \ldots + \alpha_n v_n + \gamma(\beta_1 v_1 + \ldots + \beta_n v_n),$$

a partir de lo cual se concluye que

$$y = (\alpha_1 + \gamma\beta_1)v_1 + \ldots + (\alpha_n + \gamma\beta_n)v_n,$$

con lo que se muestra que y es una combinación lineal de los vectores $\{v_1, \ldots, v_n\}$, es decir,

$$gen\{v_1, \ldots, v_n, u\} \subseteq gen\{v_1, \ldots, v_n\}$$

y, por lo tanto,

$$gen\{v_1, \ldots, v_n, u\} = gen\{v_1, \ldots, v_n\}.$$

Si se supone ahora que $gen\{v_1, \ldots, v_n, u\} = gen\{v_1, \ldots, v_n\}$, entonces como $u \in gen\{v_1, \ldots, v_n, u\}$, se concluye que $u \in gen\{v_1, \ldots, v_n\}$.

El teorema anterior implica que si se añade o se quita un vector que es combinación lineal de los otros, el subespacio generado por estos no cambia.

Ejemplo 3.3.9. Sea el espacio vectorial $\mathbb{R}^3$ y sea el conjunto S dado por $S = \{(1,2,3), (0,-2,4), (1,0,7)\}$. Como tenemos que $(1,0,7) = (1,2,3) + (0,-2,4)$, entonces

$$gen\{(1,2,3), (0,-2,4), (1,0,7)\} = gen\{(1,2,3), (0,-2,4)\},$$

puesto que el vector $(1,0,7)$ es combinación lineal de los otros dos.

3.4 Bases de un espacio vectorial

Cuando se tiene un espacio vectorial, algunos de los elementos del espacio permiten reconstruir a cualquier otro elemento del espacio. De esta forma, es importante destacar aquellos elementos que pueden generar a los demás vectores del espacio. En esta sección se dan las definiciones tendientes a la caracterización de tales elementos.

Definición 3.4.1 (Independencia lineal). *Los vectores v_1, $v_2, \ldots, v_k$ en un espacio vectorial V son linealmente dependientes si existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$, no todos nulos, tales que*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0.$$

En el caso contrario se dice que los vectores $v_1, v_2, \ldots, v_k$ son linealmente independientes. Esto es, la solución de la ecuación

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$$

es la solución trivial, es decir: $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.

Ejemplo 3.4.1. Sea el conjunto $S = \{(1,2),(3,4)\}$ de vectores de $\mathbb{R}^2$. Probar que es un conjunto linealmente independiente.

Solución

Sea la ecuación vectorial

$$\alpha(1,2) + \beta(3,4) = (0,0).$$

Esta ecuación es equivalente al sistema lineal de ecuaciones

$$\begin{aligned} \alpha + 3\beta &= 0 \\ 2\alpha + 4\beta &= 0. \end{aligned}$$

Si a la segunda ecuación se le resta dos veces la primera, se consigue que $-2\beta = 0$, esto es, $\beta = 0$. Al sustituir este valor en la primera ecuación se obtiene que $\alpha = 0$. De esta forma, como ambos escalares son nulos, el conjunto S es linealmente independiente.

De la definición de independencia lineal se desprenden los siguientes hechos:

- Un conjunto formado por un solo vector no nulo es linealmente independiente.
- Si un conjunto contiene al vector cero, este conjunto es linealmente dependiente.
- Si un conjunto tiene dos vectores y es linealmente dependiente, entonces uno de ellos es un múltiplo escalar del otro.
- Si un conjunto tiene dos vectores y estos no son paralelos, entonces el conjunto es linealmente independiente.

Ejemplo 3.4.2. Sea V el espacio vectorial formado por las funciones de una variable x. Probar que las funciones $f(x) = e^x$ y $g(x) = e^{-x}$ son linealmente independientes.
Solución
Sean α y β escalares tales que

$$\alpha e^x + \beta e^{-x} = 0.$$

Si se derivan ambos lados de la ecuación anterior se consigue

$$\alpha e^x - \beta e^{-x} = 0.$$

Al sumar las dos últimas relaciones se concluye que

$$2\alpha e^x = 0.$$

Como la función exponencial es siempre positiva, entonces $\alpha = 0$.

Al sustituir este valor en la primera ecuación se concluye también que $\beta = 0$ y, en consecuencia, las dos funciones son linealmente independientes.

Los conceptos de *dependencia lineal* y *combinación lineal* se relacionan mediante el siguiente teorema.

Teorema 3.4.1. *Sea V un espacio vectorial sobre un campo $\mathbb{K}$. Sea $S = \{v_1, v_2, \ldots, v_n\}$ un conjunto de vectores de V, con $S \neq \phi$. El conjunto S es linealmente dependiente si y solo si existe al menos un vector en S que puede expresarse como una combinación lineal de los vectores restantes.*

Demostración. Sea el conjunto S un conjunto linealmente dependiente. Entonces, existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, no todos nulos, tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0.$$

Si α_k es el primer escalar distinto de cero que satisface la ecuación anterior, entonces:

$$\alpha_k v_k = -\alpha_1 v_1 - \cdots - \alpha_{k-1} v_{k-1} - \alpha_{k+1} v_{k+1} - \ldots - \alpha_n v_n,$$

con lo cual

$$v_k = -\frac{\alpha_1}{\alpha_k} v_1 - \cdots - \frac{\alpha_{k-1}}{\alpha_k} v_{k-1} - \frac{\alpha_{k+1}}{\alpha_k} v_{k+1} - \ldots - \frac{\alpha_n}{\alpha_k} v_n,$$

lo que permite concluir que uno de los vectores es una combinación lineal de los demás.

Sea ahora un vector v_k que es combinación lineal de los restantes; entonces,

$$v_k = -\alpha_1 v_1 - \cdots - \alpha_{k-1} v_{k-1} - \alpha_{k+1} v_{k+1} - \ldots - \alpha_n v_n;$$

luego se tiene que

$$\alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} + v_k + \alpha_{k+1} v_{k+1} + \ldots + \alpha_n v_n = 0;$$

pero en esta ecuación hay por lo menos un escalar $\alpha_k = 1 \neq 0$, a partir de lo cual se obtiene que el conjunto dado es linealmente dependiente.

Teorema 3.4.2. *Sea V un espacio vectorial sobre un campo $\mathbb{K}$ y sean A y B subconjuntos de V tales que $A \subset B$, $A \neq \phi$ y B finito; entonces, se cumple que*

1. *Si A es un subconjunto linealmente dependiente, B es linealmente dependiente.*

2. *Si B es un conjunto linealmente independiente, A es linealmente independiente.*

Demostración. Sean los subconjuntos $A = \{v_1,\ v_2, \ldots, v_n\}$ y $B = \{v_1,\ v_2, \ldots, v_n, \ldots, v_{n+k}\}$; como A es linealmente dependiente, existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, no todos nulos, tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0.$$

Luego existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, no todos nulos, tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n + 0 v_{n+1} + \ldots + 0 v_{n+k} = 0.$$

En consecuencia, B es linealmente dependiente.

La segunda parte es válida, ya que la proposición es el contrarrecíproco del primer numeral.

Definición 3.4.2. *Los vectores $v_1,\ v_2, \ldots, v_n$ en un espacio vectorial V forman una base para V si:*

- *$v_1,\ v_2, \ldots, v_n$ generan a V.*

- *$v_1,\ v_2, \ldots, v_n$ son linealmente independientes.*

Ejemplo 3.4.3. Sea el conjunto $S = \{(1,3), (3,5)\}$ de vectores de $\mathbb{R}^2$. Probar que estos vectores determinan una base para $\mathbb{R}^2$.
Solución
Primero se demuestra que el conjunto es linealmente independiente.

Sea la ecuación vectorial

$$\alpha(1,3) + \beta(3,5) = (0,0).$$

Esta ecuación es equivalente al sistema lineal de ecuaciones

$$\begin{aligned} \alpha + 3\beta &= 0 \\ 3\alpha + 5\beta &= 0, \end{aligned}$$

cuya matriz de coeficientes es

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

La forma escalonada reducida de esta matriz es

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

y, por lo tanto, $\beta = 0$ y $\alpha = 0$. De esta forma, como ambos escalares son nulos, los dos vectores son linealmente independientes.

Ahora debe probarse que estos vectores generan a $\mathbb{R}^2$, esto es, dado el vector (a, b), este puede escribirse como una combinación lineal de los dos vectores dados, es decir, se debe demostrar que la ecuación

$$\alpha(1, 3) + \beta(3, 5) = (a, b)$$

tiene solución. En este caso se presenta el siguiente sistema:

$$\begin{aligned} \alpha + 3\beta &= a \\ 3\alpha + 5\beta &= b; \end{aligned}$$

para este sistema no homogéneo, la matriz aumentada es

$$\begin{bmatrix} 1 & 3 & a \\ 3 & 5 & b \end{bmatrix}$$

y su forma escalonada reducida es

$$\begin{bmatrix} 1 & 0 & \dfrac{3b - 5a}{4} \\ 0 & 1 & \dfrac{3a - b}{4} \end{bmatrix},$$

con lo cual

$$\alpha = \frac{3b-5a}{4}$$

y

$$\beta = \frac{3a-b}{4}.$$

De esta forma, el vector (a, b) puede escribirse como

$$(a, b) = \frac{3b-5a}{4}\Big(1, 3\Big) + \frac{3a-b}{4}\Big(3, 5\Big).$$

Esta última relación muestra que cualquier vector de $\mathbb{R}^2$ puede escribirse como una combinación lineal de los vectores $(1, 3)$ y $(3, 5)$. Por lo tanto, ellos forman un base del espacio $\mathbb{R}^2$.

Dado el conjunto de vectores $S = \{v_1,\, v_2, \ldots, v_n\}$ en un espacio vectorial; si S es una base para V, entonces todo vector u se puede expresar como una combinación lineal de los vectores de S en la forma

$$u = \alpha_1 v_1 + \ldots + \alpha_n v_n.$$

El siguiente teorema indica que cuando se tiene una base para un espacio vectorial, cada vector del espacio tiene una representación única en términos de los vectores de la base.

Teorema 3.4.3. *Si $S = \{v_1, v_2, \ldots, v_n\}$ es una base para un espacio vectorial V, entonces todo vector en V se puede escribir en forma única como una combinación lineal de los vectores de S.*

Demostración. Si el vector u de V tuviese dos representaciones diferentes $u = \alpha_1 v_1 + \ldots + \alpha_n v_n$ y $u = \beta_1 v_1 + \ldots + \beta_n v_n$, al restar estas dos representaciones se obtiene

$$(\alpha_1 - \beta_1)v_1 + \ldots + (\alpha_n - \beta_n)v_n = 0;$$

pero como los vectores v_i son linealmente independientes, se tiene que $\alpha_i - \beta_i = 0$, es decir, $\alpha_i = \beta_i$ para $i = 1, \ldots, n$.

Si S es una base para V, entonces todo vector u se puede expresar como una combinación lineal de los vectores de S en la forma

$$u = \alpha_1 v_1 + \ldots + \alpha_n v_n.$$

En tal caso se dice que la n-tupla $(\alpha_1, \ldots, \alpha_n)$ son las *coordenadas* del vector u respecto a la base S.

Definición 3.4.3. *Sea $\{v_1, v_2, \ldots, v_n\}$ un conjunto de vectores de un espacio vectorial V. Sea r un entero positivo tal que $r \leq n$. Se dice que $S = \{v_1, v_2, \ldots, v_r\}$ es un subconjunto máximo de vectores linealmente independientes si los elementos $v_1, v_2, \ldots, v_r$ son linealmente independientes y si, además, dado cualquier v_i, con $i > r$, los elementos $v_1, v_2, \ldots, v_r, v_i$ son linealmente dependientes.*

Con esta definición es posible utilizar un criterio nuevo para determinar si un conjunto dado de vectores es una base para un espacio vectorial.

Teorema 3.4.4. *Sea $\{v_1, v_2, \ldots, v_n\}$ un conjunto de generadores de un espacio vectorial V. Sea $\{v_1, v_2, \ldots, v_r\}$ un subconjunto máximo de vectores linealmente independientes. Entonces, $\{v_1, v_2, \ldots, v_r\}$ es una base para el espacio V.*

Demostración. Como el conjunto $\{v_1, v_2, \ldots, v_r\}$ es linealmente independiente, solo resta probar que estos vectores generan a V. Para demostrar esto último, se probará que todo vector v_i, con $r < i \leq n$, es una combinación lineal de los vectores $v_1, v_2, \ldots, v_r$.

Como el conjunto $\{v_1, v_2, \ldots, v_r\}$ es un subconjunto máximo de vectores linealmente independientes, se tiene que si $i > r$, existen escalares $\alpha_{1,i}, \ldots, \alpha_{r,i}, \alpha_i$, no todos nulos, tales que

$$\alpha_{1,i} v_1 + \ldots + \alpha_{r,i} v_r + \alpha_i v_i = 0,$$

donde $\alpha_i \neq 0$, pues, de lo contrario, los $v_1, v_2, \ldots, v_r$ serían linealmente dependientes.

La relación anterior hace posible despejar cada uno de los vectores v_i como una combinación lineal de los vectores $v_1, v_2, \ldots, v_r$; de esta forma, si u es un elemento de V,

$$u = \alpha_1 v_1 + \ldots + \alpha_r v_r + \alpha_i v_i + \ldots + \alpha_n v_n;$$

pero como cada v_i, con $i > r$, es una combinación lineal de los vectores $v_1, v_2, \ldots, v_r$, al sustituirlos en la ecuación anterior se concluye que u es una combinación lineal de los vectores $v_1, v_2, \ldots, v_r$, lo que muestra que estos vectores también generan a V y, por lo tanto, forman una base.

Ejemplo 3.4.4. A partir de los vectores de la Figura 3.4, escribir los vectores $\vec{b}$ y $\vec{c}$ en términos de los vectores de la base $\{\vec{a}_1, \vec{a}_2\}$.

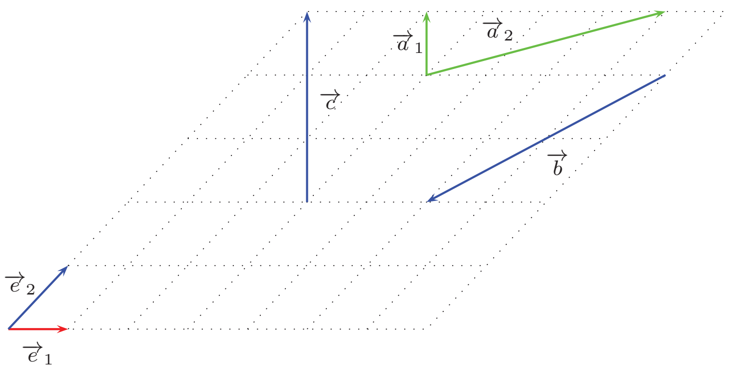


Figura 3.4. Bases para $\mathbb{R}^2$

Solución

Como

$$\vec{b} = -2(\vec{e}_1 + \vec{e}_2)$$

y, además,

$$\begin{aligned} \vec{a}_1 &= \vec{e}_2 - \vec{e}_1 \\ \vec{a}_2 &= 3\vec{e}_1 + \vec{e}_2, \end{aligned}$$

con ello

$$\begin{aligned}\vec{e}_1 &= \frac{1}{4}(\vec{a}_2 - \vec{a}_1)\\ \vec{e}_2 &= \frac{1}{4}(\vec{a}_2 + 3\vec{a}_1).\end{aligned}$$

Sustituyendo las relaciones anteriores en la expresión, para $\vec{b}$ se concluye que

$$\begin{aligned}\vec{b} &= -2(\vec{e}_1 + \vec{e}_2)\\ &= -2\left[\frac{1}{4}(\vec{a}_2 - \vec{a}_1) + \frac{1}{4}(\vec{a}_2 + 3\vec{a}_1)\right]\\ &= -\vec{a}_2 - \vec{a}_1.\end{aligned}$$

De modo semejante,

$$\vec{c} = -3\vec{e}_1 + 3\vec{e}_2;$$

y usando nuevamente las relaciones entre los vectores de las dos bases, se consigue que

$$\begin{aligned}\vec{c} &= -3(\vec{e}_1 - \vec{e}_2)\\ &= -3\left[\frac{1}{4}(\vec{a}_2 - \vec{a}_1) - \frac{1}{4}(\vec{a}_2 + 3\vec{a}_1)\right]\\ &= 3\vec{a}_1.\end{aligned}$$

Este último resultado muestra que los vectores $\vec{a}_1$ y $\vec{c}$ son paralelos.

Ejemplo 3.4.5. Hallar una base para el subespacio generado por el conjunto de vectores $S = \{(-1, 3, 1), (3, -1, 1), (4, 0, 2)\}$.

Solución

La base para $gen\{S\}$ corresponde al máximo número de vectores linealmente independientes que posea el conjunto S. Para determinar este número, se inicia analizando la independencia o dependencia lineal de todos los vectores de S, es decir, se analiza la solución de la ecuación

$$\alpha(-1,3,1)+\beta(3,-1,1)+\gamma(4,0,2)=(0,0,0),$$

la cual da origen al sistema de ecuaciones

$$\begin{aligned} -\alpha+3\beta+4\gamma &= 0\\ 3\alpha-\beta &= 0\\ \alpha+\beta+2\gamma &= 0. \end{aligned}$$

La matriz de coeficientes de este sistema es

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

y su forma escalonada reducida es

$$\begin{bmatrix} 1 & 0 & \frac{1}{2}\\ 0 & 1 & \frac{3}{2}\\ 0 & 0 & 0 \end{bmatrix},$$

lo cual significa que el sistema inicial tiene infinitas soluciones y que, por lo tanto, los tres vectores son linealmente dependientes.

Según lo expuesto anteriormente, el último vector puede escribirse como una combinación lineal de los dos primeros. Esta combinación se consigue haciendo $\gamma=1$ y resolviendo el sistema inicial, con lo cual se obtiene

$$\begin{aligned} 3\alpha-\beta &= 0\\ \alpha+\beta &= -2. \end{aligned}$$

Al sumar estas dos ecuaciones se tiene que $4\alpha=-2$; por lo tanto, $\alpha=-1/2$ y $\beta=-3/2$. De esta forma,

$$(4,0,2) = -\alpha(-1,3,1) - \beta(3,-1,1) = \frac{1}{2}(-1,3,1) + \frac{3}{2}(3,-1,1).$$

Ahora, deberá estudiarse la independencia lineal de los vectores $(-1,3,1)$ y $(3,-1,1)$, es decir, analizar la solución de la ecuación

$$\alpha(-1,3,1) + \beta(3,-1,1) = (0,0,0),$$

que da origen al sistema de ecuaciones

$$\begin{aligned} -\alpha + 3\beta &= 0 \\ 3\alpha - \beta &= 0 \\ \alpha + \beta &= 0. \end{aligned}$$

Al sumar la tercera ecuación y la primera se llega a que $4\beta = 0$, lo cual obliga a que $\alpha = \beta = 0$, probando así que los vectores $(-1,3,1)$ y $(3,-1,1)$ son linealmente independientes y que, por lo tanto, la base para $gen\{S\}$ es el conjunto de vectores formado por los vectores {(-1,3,1),(3,-1,1) }.

3.5 Dimensión de un espacio vectorial

El resultado principal de esta sección es la afirmación que dadas dos bases cualesquiera de un mismo espacio vectorial, estas tienen el mismo número de vectores. Por el momento, se establece el siguiente teorema.

Teorema 3.5.1. *Sean V un espacio vectorial sobre un campo $\mathbb{K}$; $S = \{v_1, v_2, \ldots, v_n\}$ una base de V y $T = \{w_1, w_2, \ldots, w_m\}$ un conjunto de vectores de V. Si tenemos que $m > n$, entonces T es linealmente dependiente.*

Demostración. Vamos a demostrar que si T es linealmente independiente, entonces $m \leq n$. Supongamos que T es un conjunto linealmente independiente; entonces, para todo $i = 1, \ldots, m$, $w_i \neq 0$ y, en consecuencia, al ser S una base de V, existen escalares, no todos nulos, $\alpha_1, \ldots, \alpha_n$, tales que

$$w_1 = \alpha_1 v_1 + \ldots + \alpha_n v_n.$$

Después de reenumerar, si es necesario, se puede suponer sin pérdida de generalidad que $\alpha_1 \neq 0$ y, en consecuencia, se puede despejar el vector v_1, obteniendo la expresión

$$v_1 = \frac{1}{\alpha_1} w_1 - \frac{\alpha_2}{\alpha_1} v_2 - \ldots - \frac{\alpha_n}{\alpha_1} v_n;$$

luego $v_1 \in gen\{w_1, v_2, \ldots, v_n\} = V$.

De este modo, se pueden ir reemplazando los $v_2, v_3 \ldots, v_n$, por $w_2, w_3 \ldots, w_n$, hasta sustituir los elementos $v_1, v_2, \ldots, v_n$, si es el caso.

Sea ahora un entero r tal que $1 \leq r \leq n$; después de una adecuada reorganización de los $v_1, v_2, \ldots, v_n$, los elementos $w_1, w_2, \ldots, w_r, v_{r+1}, \ldots, v_n$ generan a V (esto significa que no es necesario sustituir todos los $v_1, v_2, \ldots, v_n$ para generar a V). De esta forma, existen escalares $\beta_1, \beta_2, \ldots, \beta_r, \gamma_{r+1}, \ldots, \gamma_n$ tales que

$$w_{r+1} = \beta_1 w_1 + \beta_2 w_2 + \ldots + \beta_r w_r + \gamma_{r+1} v_{r+1} + \ldots + \gamma_n v_n.$$

No puede ocurrir que $\gamma_j = 0$, para $j = r+1, \ldots, n$, ya que entonces desde w_{r+1} sería combinación lineal de $w_1, w_2, \ldots, w_r$, lo que es contrario al supuesto de que los elementos de T son linealmente independientes.

Luego de reenumerar si es necesario, se puede suponer que $\gamma_{r+1} \neq 0$ y es posible despejar v_{r+1} en la forma

$$v_{r+1} = \gamma_{r+1}^{-1} w_{r+1} - \gamma_{r+1}^{-1} \beta_1 w_1 - \ldots - \gamma_{r+1}^{-1} \beta_r w_r - \ldots - \gamma_{r+1}^{-1} \gamma_n v_n,$$

lo que muestra que v_{r+1} está en el subespacio generado por el conjunto de vectores $w_1, w_2, \ldots, w_r, w_{r+1}, v_{r+2}, \ldots, v_n$.

Así, si fuera $m > n$, existen escalares $\psi_1, \psi_2, \ldots, \psi_r, \ldots, \psi_n$, no todos nulos, tales que

$$w_m = \psi_1 w_1 + \psi_2 w_2 + \ldots + \psi_r w_r + \ldots + \psi_n w_n$$

y, en consecuencia, los vectores $w_1, w_2, \ldots, w_r, \ldots, w_n, w_m$ serían linealmente dependientes, contradiciendo la hipótesis de que el conjunto T es linealmente independiente. Luego debe ser $m \leq n$.

El teorema anterior indica que si una base posee n elementos y un conjunto de vectores linealmente independientes posee m elementos, entonces, $m \leq n$.

Teorema 3.5.2. *Sea V un espacio vectorial. Supongamos que una base de V tiene n elementos y que otra base del mismo espacio tiene m elementos; entonces, $n = m$.*

Demostración. Por el teorema anterior, como los elementos de cada base son linealmente independientes, entonces $m \leq n$ y $n \leq m$, con lo cual $n = m$.

Definición 3.5.1 (Dimensión). *Sea V un espacio vectorial. Si V tiene una base con n elementos, se dice que el espacio tiene dimensión n. Si V consta solo del vector nulo, entonces V no tiene base y su dimensión es cero.*

Ejemplo 3.5.1. Como los vectores $\{e_1, \ldots, e_n\}$ –siendo el vector $e_i = (0, \ldots, 1, \ldots, 0)$, donde el 1 ocupa la posición i– son linealmente independientes y generan a $\mathbb{R}^n$, ellos forman una base de $\mathbb{R}^n$; entonces, la dimensión de este espacio es n. Esta base es conocida como la *base canónica* de $\mathbb{R}^n$.

Ejemplo 3.5.2. Sea el espacio vectorial de las funciones polinómicas. Los elementos 1, x, $\ldots, x^n$ forman una base, puesto que son linealmente independientes y cualquier polinomio de grado menor o igual que n puede generarse con ellos. Por tanto la dimensión de este espacio es $n+1$.

Teorema 3.5.3. *Sea $\{v_1, \ldots, v_n\}$ un conjunto máximo de vectores de V linealmente independientes; entonces, el conjunto formado por los vectores $\{v_1, \ldots, v_n\}$, es una base de V.*

Demostración. Para mostrar que dicho conjunto es una base se debe verificar que es linealmente independiente, y luego, que cualquier vector del espacio se puede escribir como una combinación lineal de los vectores del conjunto.

Como aquí se trata de un conjunto linealmente independiente, solo resta probar que generan a V.

Para esto, sea un vector u cualquiera en V. Como el conjunto dado es un conjunto máximo de vectores linealmente independientes, entonces los vectores $v_1, \ldots, v_n, u$ son linealmente dependientes, lo que significa que existen escalares $\alpha_1, \ldots, \alpha_n, \alpha_{n+1}$, no todos nulos, tales que

$$\alpha_1 v_1 + \ldots + \alpha_n v_n + \alpha_{n+1} u = 0.$$

Como $\alpha_{n+1} \neq 0$, entonces se puede despejar u para obtener que

$$u = -\alpha_{n+1}^{-1} \alpha_1 v_1 - \ldots - \alpha_{n+1}^{-1} \alpha_n v_n,$$

lo cual demuestra que el conjunto $\{v_1, \ldots, v_n\}$ genera a V y, en consecuencia, es una base del espacio V.

Teorema 3.5.4. *Sea V un espacio vectorial de dimensión n y sea $\{v_1, \ldots, v_n\}$ un conjunto de vectores de V linealmente independientes; entonces, $\{v_1, \ldots, v_n\}$ es una base de V.*

Demostración. Como la dimensión del espacio es n, al añadir un vector más a $\{v_1, \ldots, v_n\}$ se obtiene un conjunto de vectores linealmente dependientes, lo cual prueba que el conjunto $\{v_1, \ldots, v_n\}$ es un conjunto máximo de vectores de V linealmente independientes y, por el teorema anterior, es una base del espacio vectorial V.

Corolario 3.5.1. *Sea V un espacio vectorial de dimensión n y W un subespacio de V también de dimensión n; entonces, $W = V$.*

Corolario 3.5.2. *Sean V un espacio vectorial de dimensión n, $r < n$ y $\{v_1, \ldots, v_r\}$ un conjunto de r vectores linealmente independientes; entonces, es posible hallar vectores $v_{r+1}, \ldots, v_n$ tales que $\{v_1, \ldots, v_n\}$ formen una base de V.*

Demostración. Al ser $r < n$, $\{v_1, \ldots, v_r\}$ no es una base de V y no pueden formar un conjunto máximo de vectores de V linealmente independientes; luego existe v_{r+1} tal que $\{v_1, \ldots, v_{r+1}\}$ es linealmente independiente. Si $r + 1 < n$, se repite el proceso anterior, hasta obtener el conjunto de vectores $\{v_1, \ldots, v_n\}$ que sea un conjunto máximo de vectores de V linealmente independientes y, en consecuencia, una base para V.

Ejemplo 3.5.3. Sea el conjunto $S = \{(1, 0, 1), (0, 2, 1)\}$ de vectores de $\mathbb{R}^3$. Completar una base para dicho espacio.

Solución

Estos dos vectores son linealmente independientes, ya que la ecuación

$$\alpha(1, 0, 1) + \beta(0, 2, 1) = (0, 0, 0),$$

que es equivalente al sistema

$$\begin{aligned} \alpha &= 0 \\ 2\beta &= 0 \\ \alpha + \beta &= 0, \end{aligned}$$

tiene por solución $\alpha = 0$ y $\beta = 0$.

Para completar la base, basta con agregar un vector que no esté en el espacio generado por estos dos vectores. Como tenemos que $gen\{S\} = \{v/v = (\alpha, 2\beta, \alpha + \beta)\}$, para completar la base debemos tomar un vector cualquiera que no tenga esta forma; en este caso, podemos tomar el vector $v_3 = (1, 2, 0)$.

Si queremos verificar que este sistema es linealmente independiente, consideremos la solución de la ecuación vectorial

$$\alpha(1, 0, 1) + \beta(0, 2, 1) + \gamma(1, 2, 0) = (0, 0, 0),$$

la cual da origen al sistema de ecuaciones lineales

$$\begin{aligned} \alpha + \gamma &= 0 \\ 2\beta + 2\gamma &= 0 \\ \alpha + \beta &= 0. \end{aligned}$$

De la primera ecuación y de la última se tiene que $\alpha = -\beta$ y $\alpha = -\gamma$, con lo cual $-4\alpha = 0$, y así concluimos que $\alpha = \beta = \gamma = 0$, lo que prueba que los tres vectores son linealmente independientes y, por tanto, determinan una base para $\mathbb{R}^3$.

Teorema 3.5.5. *Sean V un espacio vectorial de dimensión n, $r < n$ y $\{v_1, \ldots, v_r\}$ un conjunto de r vectores; es posible hallar vectores $v_s, \ldots, v_n$ tales que $\{v_1, \ldots, v_n\}$ formen una base de V.*

Demostración. Si los vectores $\{v_1, \ldots, v_r\}$ son linealmente independientes, el resultado anterior indica que es posible hallar vectores $v_{r+1}, \ldots, v_n$ tales que $\{v_1, \ldots, v_n\}$ formen una base de V.

Si los vectores $\{v_1, \ldots, v_r\}$ son linealmente dependientes, existe un $s < r$ tal que el conjunto de vectores $\{v_1, \ldots, v_s\}$ es linealmente independiente y, en consecuencia, es posible hallar vectores $v_{s+1}, \ldots, v_n$ tales que $\{v_1, \ldots, v_n\}$ formen una base de V.

Ejemplo 3.5.4. Hallar la dimensión del subespacio generado por los vectores $v_1 = (0, 1, 2)$, $v_2 = (2, -2, 6)$ y $v_3 = (-1, 2, -1)$.

Solución

Para hallar la dimensión de este subespacio basta con determinar una base para él. Esto se hace identificando el número máximo de vectores linealmente independientes en el conjunto $\{v_1, v_2, v_3\}$.

Se comienza verificando si los tres vectores dados son linealmente independientes, es decir, se buscan las soluciones de la ecuación vectorial

$$\alpha(0,1,2) + \beta(2,-2,6) + \gamma(-1,2,-1) = (0,0,0),$$

lo cual da origen al sistema de ecuaciones

$$\begin{aligned} 2\beta - \gamma &= 0 \\ \alpha - 2\beta + 2\gamma &= 0 \\ 2\alpha + 6\beta - \gamma &= 0. \end{aligned}$$

Para hallar la solución de este sistema se puede usar la eliminación de Gauss-Jordan. En este caso, la matriz de coeficientes del sistema es

$$\begin{bmatrix} 0 & 2 & -1 \\ 1 & -2 & 2 \\ 2 & 6 & -1 \end{bmatrix},$$

cuya forma escalonada reducida es

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

Por lo tanto, el sistema posee infinitas soluciones.

Para expresar las infinitas soluciones se puede escoger una de las incógnitas como parámetro. En este ejemplo, si $\beta = t$, entonces $\gamma = 2t$ y $\alpha = 2\beta - 2\gamma = -2t$. Lo anterior muestra que los tres vectores dados son linealmente dependientes.

Puede observarse que si $t = 1$, entonces $\beta = 1$, $\gamma = 2$ y $\alpha = 2\beta - 2\gamma = -2$. Estos valores permiten despejar el segundo vector como combinación lineal de los otros dos, esto es:

$$(2, -2, 6) = 2(0, 1, 2) - 2(-1, 2, -1).$$

En consecuencia, el subespacio $gen\{(0, 1, 2), (2, -2, 6), (-1, 2, -1)\}$ es el mismo que $gen\{(0, 1, 2), (-1, 2, -1)\}$.

Si en el sistema inicial se hace $\beta = 0$, se obtiene el sistema

$$\begin{aligned} -\gamma &= 0 \\ \alpha + 2\gamma &= 0 \\ 2\alpha - \gamma &= 0, \end{aligned}$$

el cual tiene por solución única $\alpha = 0$, $\gamma = 0$, lo cual significa que los vectores $\{(0, 1, 2), (-1, 2, -1)\}$ son linealmente independientes y, en consecuencia, ellos forman una base para el subespacio $gen\{(0, 1, 2), (2, -2, 6), (-1, 2, -1)\}$. Como esta base posee dos vectores, la dimensión del subespacio es 2.

Ejemplo 3.5.5. Probar que los vectores $v_1 = (1, 2, 3)$, $v_2 = (0, 1, 1)$ y $v_3 = (2, 0, 1)$ determinan una base para $\mathbb{R}^3$. Escribir el vector $v = (a, b, c)$ como una combinación lineal de los vectores v_1, v_2 y v_3.

Solución

Como la dimensión del espacio $\mathbb{R}^3$ es 3 y se tienen tres vectores, para probar que ellos determinan una base solo es necesario probar que los vectores son linealmente independientes. Esto se hace hallando la solución de la ecuación vectorial

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

la cual da origen al sistema homogéneo

$$\begin{aligned} x + 2z &= 0 \\ 2x + y &= 0 \\ 3x + y + z &= 0. \end{aligned}$$

La matriz de coeficientes de este sistema es

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

para la cual se tiene que $\det A = -1 \neq 0$, que nos indica que el sistema homogéneo tiene solución única $x = y = z = 0$. De esta manera se prueba que los tres vectores dados son linealmente independientes y, por lo tanto, constituyen una base para $\mathbb{R}^3$.

Si se quiere escribir un vector genérico $v = (a, b, c)$ como una combinación lineal de los tres vectores, se requiere resolver el sistema

$$\begin{aligned} x + 2z &= a \\ 2x + y &= b \\ 3x + y + z &= c. \end{aligned}$$

En este caso, vamos a utilizar el método de eliminación de Gauss-Jordan para determinar los coeficientes x, y y z de la combinación lineal. La matriz aumentada del sistema es

$$A = \left[\begin{array}{ccc|c} 1 & 0 & 2 & a \\ 2 & 1 & 0 & b \\ 3 & 1 & 1 & c \end{array}\right],$$

que al llevarla a la forma escalonada reducida se expresa como

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -a - 2b + 2c \\ 0 & 1 & 0 & 2a + 5b - 4c \\ 0 & 0 & 1 & a + b - c \end{array}\right],$$

es decir, los coeficientes de la combinación lineal son:

$$\begin{aligned} x &= -a - 2b + 2c \\ y &= 2a + 5b - 4c \\ z &= a + b - c \end{aligned}$$

y, por lo tanto, obtenemos la combinación lineal

$$(-a-2b+2c)(1,2,3)+(2a+5b-4c)(0,1,1)+(a+b-c)(2,0,1)=(a,b,c)$$

para cualquier vector (a,b,c) de $\mathbb{R}^3$.

Ejemplo 3.5.6. Probar que los vectores $v_1=(1,0,0)$, $v_2=(0,1,0)$ y $v_3=(0,0,1)$ determinan una base para $\mathbb{R}^3$. Escribir el vector $v=(a,b,c)$ como una combinación lineal de los vectores v_1, v_2 y v_3.
Solución

Como en el ejercicio anterior, para probar que ellos determinan una base solo es necesario probar que los vectores son linealmente independientes. Esto se hace hallando la solución de la ecuación vectorial

$$x(1,0,0)+y(0,1,0)+z(0,0,1)=(0,0,0),$$

la cual da origen al sistema homogéneo

$$\begin{aligned} x &= 0 \\ y &= 0 \\ z &= 0, \end{aligned}$$

que nos indica que tiene solución única $x=y=z=0$, probando así que los tres vectores dados son linealmente independientes y, por lo tanto, constituyen una base para $\mathbb{R}^3$.

Si se quiere escribir un vector genérico $v=(a,b,c)$ como una combinación lineal de los tres vectores, se requiere resolver el sistema

$$\begin{aligned} x &= a \\ y &= b \\ z &= c, \end{aligned}$$

que obviamente tiene solución única.

La forma de elección de los vectores para la construcción de una base de $\mathbb{R}^3$ usada en los dos ejercicios anteriores, sugiere que si se toman tres vectores libres que determinen un paralelepípedo, estos forman una base para $\mathbb{R}^3$, tal como puede verse en la Figura 3.5.

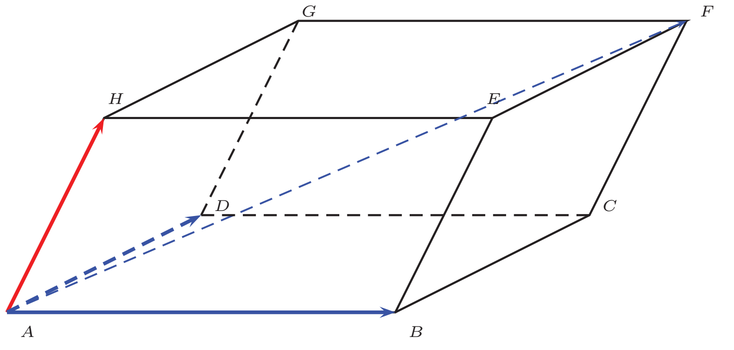


Figura 3.5. Tres vectores libres determinan una base para $\mathbb{R}^3$

Como ya se ha discutido, dado un espacio vectorial, este puede tener diferentes bases; sin embargo, en el capítulo 4 se verá que existen bases con características especiales que hacen más cómodo el problema de la búsqueda de las coordenadas de un vector respecto a una de estas bases.

3.6 Ejercicios resueltos

Ejercicio 3.6.1. Hallar la ecuación implícita del subespacio de $\mathbb{R}^3$ generado por los vectores $v_1 = (1, 1, 1)$ y $v_2 = (2, -1, 2)$. Determinar si $v = (3, 2, 1)$ está en dicho subespacio.

Solución

Sea $v = (x, y, z)$ un vector de $\mathbb{R}^3$. Sabemos que v está en el subespacio generado por v_1 y v_2 si y solo si existen escalares α y β tales que

$$(x, y, z) = \alpha(1, 1, 1) + \beta(2, -1, 2).$$

Esta ecuación vectorial da origen al sistema de ecuaciones

$$\begin{aligned} \alpha + 2\beta &= x \\ \alpha - \beta &= y \\ \alpha + 2\beta &= z. \end{aligned}$$

La matriz aumentada de este sistema es

$$\left[\begin{array}{rr|r} 1 & 2 & x \\ 1 & -1 & y \\ 1 & 2 & z \end{array}\right]$$

y su forma escalonada reducida es

$$\left[\begin{array}{cc|c} 1 & 0 & \dfrac{x+2y}{3} \\ 0 & 1 & \dfrac{x-y}{3} \\ 0 & 0 & z-x \end{array}\right],$$

lo cual indica que el sistema tiene solución si se cumple que $z - x = 0$, es decir, (x, y, z) está en el subespacio generado por los vectores v_1 y v_2 si y solo si $z - x = 0$, o en forma equivalente, $z = x$. Con esta condición se nota inmediatamente que el vector $v = (3, 2, 1)$ no está en el subespacio generado por v_1 y v_2.

Ejercicio 3.6.2. Hallar el subespacio $A \cap B$ donde A y B son los subespacios generados por los vectores $v_1 = (1, 1, 2)$ y $v_2 = (2, 3, 4)$, y $w_1 = (1, 0, 1)$ y $w_2 = (3, -1, 4)$ respectivamente.

Solución

Sea $v = (x, y, z)$ un vector de $\mathbb{R}^3$. Sabemos que v está en el subespacio generado por v_1 y v_2 si y solo si existen escalares α y β tales que

$$(x, y, z) = \alpha(1, 1, 2) + \beta(2, 3, 4).$$

Esta ecuación vectorial da origen al sistema de ecuaciones

$$\begin{aligned}\alpha + 2\beta &= x\\ \alpha + 3\beta &= y\\ 2\alpha + 4\beta &= z.\end{aligned}$$

La matriz aumentada de este sistema es

$$\left[\begin{array}{cc|c} 1 & 2 & x\\ 1 & 3 & y\\ 2 & 4 & z \end{array}\right]$$

y su forma escalonada reducida es

$$\left[\begin{array}{cc|c} 1 & 0 & 3x - 2y\\ 0 & 1 & y - x\\ 0 & 0 & z - 2x \end{array}\right],$$

lo cual indica que el sistema tiene solución si se cumple que $z - 2x = 0$, es decir, (x, y, z) está en el subespacio generado por los vectores v_1 y v_2 si y solo si $z - 2x = 0$.

De forma similar, el vector $v = (x, y, z)$ de $\mathbb{R}^3$ está en el subespacio generado por w_1 y w_2 si y solo si existen escalares α y β tales que

$$(x, y, z) = \alpha(1, 0, 1) + \beta(3, -1, 4).$$

Esta ecuación vectorial da origen al sistema de ecuaciones

$$\begin{aligned}\alpha + 3\beta &= x\\ -\beta &= y\\ \alpha + 4\beta &= z.\end{aligned}$$

La matriz aumentada de este sistema es

$$\left[\begin{array}{cc|c} 1 & 3 & x\\ 0 & -1 & y\\ 1 & 4 & z \end{array}\right]$$

y su forma escalonada reducida es

$$\left[\begin{array}{cc|c} 1 & 0 & x+3y \\ 0 & 1 & -y \\ 0 & 0 & z-x+y \end{array}\right],$$

lo cual indica que el sistema tiene solución si se cumple que $z-x+y=0$, es decir, (x,y,z) está en el subespacio generado por los vectores w_1 y w_2 si y solo si $z-x+y=0$. En consecuencia, un vector $v=(x,y,z)$ está en el subespacio $A\cap B$ si cumple simultáneamente el sistema de ecuaciones

$$\begin{aligned} -2x+z &= 0 \\ -x+y+z &= 0, \end{aligned}$$

que son las que caracterizan a cada uno de los subespacios.

Para resolver este sistema homogéneo basta escalonar la matriz de coeficientes del sistema, la cual es:

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

y su forma escalonada reducida es

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

La solución de este sistema, en forma paramétrica, está dada por las ecuaciones

$$z=t, \quad y=-\frac{1}{2}t, \quad x=\frac{1}{2}t,$$

las cuales pueden escribirse así:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{t}{2}\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}.$$

El resultado anterior es equivalente a afirmar que $A \cap B$ es el subespacio generado por el vector $(1, -1, 2)$, que es la recta de intersección de los dos planos A y B.

Ejercicio 3.6.3. Determine una base para el subespacio generado por los vectores $v_1 = (1, -1, 1)$, $v_2 = (3, 2, 1)$ y $v_3 = (2, 3, 0)$.
Solución
Si los tres vectores son linealmente independientes, ellos constituyen una base para dicho subespacio, que en este caso es el mismo $\mathbb{R}^3$. Recordemos que los vectores $v_1 = (1, -1, 1)$, $v_2 = (3, 2, 1)$ y $v_3 = (2, 3, 0)$ son linealmente independientes si la ecuación

$$\alpha(1, -1, 1) + \beta(3, 2, 1) + \gamma(2, 3, 0) = (0, 0, 0)$$

tiene como única solución $\alpha = \beta = \gamma = 0$.

Esta ecuación da origen al sistema de ecuaciones homogéneo

$$\begin{aligned} \alpha + 3\beta + 2\gamma &= 0 \\ -\alpha + 2\beta + 3\gamma &= 0 \\ \alpha + \beta &= 0, \end{aligned}$$

cuya matriz de coeficientes es

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

El determinante de esta matriz es

$$\det A = 1(-3) - 3(-3) + 2(-3) = 0$$

y, por lo tanto, los tres vectores son linealmente dependientes.

Para establecer los vectores que determinan una base para dicho subespacio podemos aplicar dos procedimientos. En primer lugar, eliminamos un vector que sea una combinación de los demás y, luego, verificamos que los restantes sean linealmente independientes. El otro camino consiste en llevar la matriz

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

a la forma escalonada reducida. En este caso el número de pivotes o unos capitales, nos da el número de vectores linealmente independientes que generan el subespacio y, además, los vectores que conforman la base son los que corresponden a las columnas donde se encuentran los pivotes. En el ejemplo se tiene que la forma escalonada reducida de la matriz A es

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

con lo cual se concluye que la base para este subespacio está conformada por los vectores $v_1 = (1, -1, 1)$ y $v_2 = (3, 2, 1)$.

Ejercicio 3.6.4. Sea V el espacio vectorial de las funciones polinómicas de $\mathbb{R}$ en $\mathbb{R}$ y grado menor o igual que dos. Probar que los polinomios $g_1(t) = 1$, $g_2(t) = t + a$, y $g_3(t) = (t + a)^2$, con a fijo, determinan una base para V. Establecer las coordenadas del polinomio $p(t) = bt^2 + ct + d$ respecto a dicha base.

Solución

Como la dimensión del espacio es 3, para probar que los tres vectores conforman una base, solo basta demostrar que ellos son linealmente independientes. Consideremos la ecuación

$$\alpha g_1(t) + \beta g_2(t) + \gamma g_3(t) = 0,$$

la cual equivale a

$$\alpha(1) + \beta(t + a) + \gamma(t + a)^2 = \gamma t^2 + (\beta + 2a\gamma)t + (\alpha + a\beta + a^2\gamma) = 0.$$

Como se trata de una ecuación polinómica, donde uno de los polinomios es nulo, cada uno de los coeficientes del polinomio del lado izquierdo debe ser nulo, dando origen al sistema de ecuaciones

$$\begin{aligned}\gamma &= 0\\ \beta + 2a\gamma &= 0\\ \alpha + a\beta + a^2\gamma &= 0.\end{aligned}$$

La matriz de coeficientes de este sistema es

$$A = \begin{bmatrix} 0 & 0 & 1\\ 0 & 1 & 2a\\ 1 & a & a^2\end{bmatrix}.$$

Usando los cofactores de la primera fila se obtiene que $\det A = -1 \neq 0$ y, por lo tanto, el sistema homogéneo tiene como única solución la trivial, esto es, $\alpha = \beta = \gamma = 0$. En consecuencia, los tres polinomios son linealmente independientes, determinando así una base para V.

Para hallar las coordenadas de $p(t) = bt^2 + ct + d$ respecto a la base ordenada $\{g_1(t), g_2(t), g_3(t)\}$ debemos resolver la ecuación

$$\begin{aligned}\alpha + \beta(t+a) + \gamma(t+a)^2 &= bt^2 + ct + d\\ \gamma t^2 + (\beta + 2a\gamma)t + (\alpha + a\beta + a^{2\gamma}) &= bt^2 + ct + d,\end{aligned}$$

al igualar los coeficientes de los términos semejantes, obtenemos el sistema de ecuaciones:

$$\begin{aligned}\gamma &= b\\ \beta + 2a\gamma &= c\\ \alpha + a\beta + a^2\gamma &= d.\end{aligned}$$

La matriz aumentada de este sistema es

$$A = \left[\begin{array}{ccc|c} 0 & 0 & 1 & b\\ 0 & 1 & 2a & c\\ 1 & a & a^2 & d\end{array}\right].$$

La forma escalonada reducida de esta matriz es

$$\left[\begin{array}{ccc|l} 1 & 0 & 0 & a^2b - ac + d \\ 0 & 1 & 0 & c - 2ab \\ 0 & 0 & 1 & b \end{array}\right]$$

y, en consecuencia, las coordenadas del vector $p(t)$ respecto a la base ordenada $\{g_1(t), g_2(t), g_3(t)\}$ son

$$\begin{aligned} \alpha &= a^2b - ac + d \\ \beta &= c - 2ab \\ \gamma &= b, \end{aligned}$$

con lo cual

$$(a^2b - ac + d) + (c - 2ab)(t + a) + b(t + a)^2 = bt^2 + ct + d.$$

Ejercicio 3.6.5. Sea

$$W = gen\{(1, 1, 1, 2), (2, -1, 3, 4), (0, 1, 1, -1), (3, -1, 3, 7)\}$$

y sea

$$U = gen\{(-1, 2, 3, 2)\}.$$

Definimos $W + U$ por

$$W + U = \{x = w + u/w \in W, u \in U\}.$$

Verificar que $W + U$ es un subespacio vectorial. Además, hallar una base para $W + U$ y una para $W \cap U$.

Solución

Una base para el subespacio $W + U$ está formada por $B_1 \cup B_2$, donde B_1 y B_2 son bases para W y U respectivamente. Es obvio que la base para U es el vector $(-1, 2, 3, 2)$. Para determinar una base para el otro subespacio debemos seleccionar cuáles de los vectores dados son linealmente independientes. Para hacer dicha selección, como ya se ha visto, debemos buscar la forma escalonada reducida (o escalonada) de la matriz, cuyas columnas son los vectores que generan el subespacio, es decir, se debe escalonar la matriz:

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

Después de realizar las operaciones elementales correspondientes, se tiene que la forma escalonada reducida de dicha matriz es

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

lo que nos indica que los cuatro vectores originales son linealmente dependientes y que los tres primeros son linealmente independientes; por lo tanto, la base para el primer subespacio está conformada por los vectores $\{(1,1,1,2),(2,-1,3,4),(0,1,1,-1)\}$ y, en consecuencia, una base para el subespacio $W+U$ estaría constituida por $\{(1,1,1,2),(2,-1,3,4),(0,1,1,-1),(-1,2,3,2)\}$.

Veamos que estos vectores son linealmente independientes. Para ello consideremos la matriz

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

cuya forma escalonada reducida es

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

lo cual muestra que los cuatro vectores son linealmente independientes y, por lo tanto, ellos conforman una base para el subespacio $W + U$.

Para describir el subespacio $W \cap U$ establecemos primero las condiciones que debe cumplir un vector para estar en cada uno de los subespacios dados.

Un vector (x, y, z, w) está en W si existen escalares α_1, α_2 y α_3 tales que

$$\alpha_1(1, 1, 1, 2) + \alpha_2(2, -1, 3, 4) + \alpha_3(0, 1, 1, -1) = (x, y, z, w).$$

A partir de esta relación obtenemos un sistema lineal de ecuaciones cuya matriz aumentada es

$$\left[\begin{array}{rrr|c} 1 & 2 & 0 & x \\ 1 & -1 & 1 & y \\ 1 & 3 & 1 & z \\ 2 & 4 & -1 & w \end{array}\right].$$

Esta matriz, al llevarla a la forma escalonada, se transforma en

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 7x - 2z - 2w \\ 0 & 1 & 0 & -3x + z + w \\ 0 & 0 & 1 & 2x - w \\ 0 & 0 & 0 & -12x + y + 3z + 4w \end{array}\right].$$

De este modo, para que este sistema tenga solución se debe cumplir que $-12x + y + 3z + 4w = 0$.

De manera similar, un vector (x, y, z, w) está en U si existe un escalar α tal que $(x, y, z, w) = \alpha(-1, 2, 3, 2)$. Dicha condición genera un sistema de ecuaciones cuya matriz aumentada es

$$\left[\begin{array}{c|c} -1 & x \\ 2 & y \\ 3 & z \\ 2 & w \end{array}\right]$$

y cuya forma escalonada es

$$\left[\begin{array}{c|c} 1 & -x \\ 0 & y+2x \\ 0 & z+3x \\ 0 & w+2x \end{array}\right].$$

Para que este sistema tenga solución se requiere que

$$\begin{aligned} y+2x &= 0 \\ z+3x &= 0 \\ w+2x &= 0. \end{aligned}$$

Así, un vector de U debe tener la forma $(x, -2x, -3x, -2x)$.

En conclusión, para que un vector esté en $W \cap U$, simultáneamente debe cumplir las condiciones $-12x + y + 3z + 4w = 0$ y

$$\begin{aligned} x &= x \\ y &= -2x \\ z &= -3x \\ w &= -2x, \end{aligned}$$

con lo cual se obtiene que

$$-12x + (-2x) + 3(-3x)) + 4(-2x) = 0$$

y, en consecuencia, $x = y = z = w = 0$, es decir, $W \cap U$ es el subespacio nulo.

Ejercicio 3.6.6 (**Uso del computador**). Dado el subespacio vectorial

$$W = \left\{(x_1, x_2, x_3, x_4, x_5, x_6) \in \mathbb{R}^6\right\},$$

donde las variables x_i, $i = 1, \ldots 6$ se toman de tal forma que cumplen las siguientes condiciones:

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 0 \\ 4x_2 - x_3 + x_4 &= 0 \end{aligned}$$

$$x_1 + x_2 + x_3 - x_5 + x_6 = 0,$$

hallar un conjunto de generadores para este subespacio.

Solución

Para que DERIVE identifique variables de la forma x_i, es decir, variables con más de un carácter, es necesario cambiar la configuración inicial. Esto se logra editando con AUTOR la expresión INPUT MODE:=WORD, aplicamos «Enter» y a continuación ya podemos escribir las ecuaciones que definen el subespacio vectorial, introduciendo con AUTOR la expresión

$$[x_1 + 3x_2 - 5x_3 = 0, 4x_2 - x_3 + x_4 = 0,$$
$$x_1 + x_2 + x_3 - x_5 + x_6 = 0, 0, 0, 0]$$

y luego «Enter». Por cada cero que hemos agregado a las ecuaciones anteriores, DERIVE interpreta que hemos añadido una nueva ecuación trivial $0 = 0$. Así, hemos construido un sistema que consta de seis ecuaciones con seis incógnitas.

Si ahora resolvemos con la instrucción SOLVE, se obtiene la expresión

$$\left[x_1 = @_1, x_2 = @_2, x_3 = \frac{@_1 + 3@_2}{5}, x_4 = \frac{@_1 - 17@_2}{5},\right.$$
$$\left. x_5 = @_3, x_6 = \frac{5@_3 - 8@_2 - 6@_1}{5}\right].$$

Estas constituyen las ecuaciones paramétricas a partir de las ecuaciones cartesianas del subespacio. Los indicadores $@_1$, $@_2$ y $@_3$ son los parámetros del sistema. De este modo, las ecuaciones paramétricas escritas en la forma usual con $@_1 = \lambda$, $@_2 = \beta$ y

$@_3 = \gamma$, quedan determinadas por las relaciones:

$$\begin{aligned}
x_1 &= \lambda \\
x_2 &= \beta \\
x_3 &= \frac{1}{5}\lambda + \frac{3}{5}\beta \\
x_4 &= \frac{1}{5}\lambda - \frac{17}{5}\beta \\
x_5 &= \gamma \\
x_6 &= -\frac{6}{5}\lambda - \frac{8}{5}\beta + \gamma.
\end{aligned}$$

A partir de las ecuaciones paramétricas de W escritas en su forma vectorial $\overline{x} = \lambda\overline{u_1} + \beta\overline{u_2} + \gamma\overline{u_3}$ se tiene directamente un conjunto de generadores $\{u_1, u_2, u_3\}$. Para obtener estos vectores, basta con dar valores a los parámetros; de esta manera, si hacemos $\lambda = 1$, $\beta = 0$ y $\gamma = 0$ ($\overline{u_1}$); $\lambda = 0$, $\beta = 1$ y $\gamma = 0$ ($\overline{u_2}$) y $\lambda = 0$, $\beta = 0$ y $\gamma = 1$ ($\overline{u_3}$). Así, en la última expresión de DERIVE, realizamos la sustitución $@_1 = 1$, $@_2 = 0$ y $@_3 = 0$ mediante la secuencia de comandos «Manage-sustitute»; luego de simplificar resulta:

$$\left[x_1 = 1, x_2 = 0, x_3 = \frac{1}{5}, x_4 = \frac{1}{5}, x_5 = 0, x_6 = -\frac{6}{5}\right].$$

De forma similar, para $@_1 = 0$, $@_2 = 1$ y $@_3 = 0$ y $@_1 = 0$, $@_2 = 0$ y $@_3 = 1$, se obtienen respectivamente

$$\left[x_1 = 0, x_2 = 1, x_3 = \frac{3}{5}, x_4 = -\frac{17}{5}, x_5 = 0, x_6 = -\frac{8}{5}\right],$$

$$[x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0, x_5 = 1, x_6 = 1].$$

Por lo tanto, un conjunto de generadores de W es:

$$S = \left\{\left(1, 0, \frac{1}{5}, \frac{1}{5}, 0, -\frac{6}{5}\right), \left(0, 1, \frac{3}{5}, -\frac{17}{5}, 0, -\frac{8}{5}\right), (0, 0, 0, 0, 1, 1)\right\}.$$

Ejercicio 3.6.7 (Uso del computador). Calcular los valores de a y b para los cuales el vector $\overline{v} = (2, a, -3, 2, 18, b)$ pertenezca al subespacio generado por el conjunto de vectores $\overline{v}_1 = (3, 5, -2, 1, 7, -4)$, $\overline{v}_2 = (1, 9, 0, 0, -4, 3)$ y $\overline{v}_3 = (3, 6, -1, 0, 0, 0)$.

Solución

Definamos en primer lugar en DERIVE, los vectores v_1, v_2 y v_3. El vector $\overline{v} \in gen\{v_1, v_2, v_3\}$, si se verifica que $\overline{v}$ es una combinación lineal de los vectores v_1, v_2 y v_3. Bastará entonces con verificar que existe solución para la ecuación vectorial $\overline{v} = p * v_1 + q * v_2 + r * v_3$ editada con AUTOR.

Al simplificar esta expresión obtenemos el sistema

$$[2 = 3p+q+3r, a = 5p+9q+6r, -3 = -2p-r, 2 = p, 18 = 7p-4q,$$
$$b = -4p + 3q].$$

Al aplicar el comando SOLVE se halla la solución

$$[a = -5, b = -11, p = 2, q = -1, r = -1].$$

Luego, para $a = -5$ y $b = -11$, el vector dado pertenece al subespacio generado por v_1, v_2, y v_3 y, además, $\overline{v} = 2v_1 - v_2 - v_3$.

Ejercicio 3.6.8 (Uso del computador). Determine si los siguientes vectores son linealmente independientes o linealmente dependientes. Además, el *rango* de un conjunto de vectores es el máximo número de vectores linealmente independientes que hay en dicho conjunto, es decir, la dimensión del espacio vectorial que generan. ¿Cuál es el rango del conjunto dado?

1. $u_1 = (1, -1, 0, -2, 5)$, $u_2 = (1, 3, -5, 1, 3)$,
   $u_3 = (1, 3, 1, -2, 5)$, $u_4 = (1, 1, 1, 0, -2)$, $u_5 = (9, 0, 1, 2, -1)$,

2. $v_1 = (2, 3, 4, -3, 1)$, $v_2 = (1, 0, -3, 2, 5)$, $v_3 = (7, 9, 9, -7, 8)$,
   $v_4 = (-2, 3, 16, -11, -19)$.

Solución
Como en los ejercicios anteriores, comenzamos definiendo todos los vectores con ayuda de AUTOR.

1. Editamos la expresión

$$a_1 * u_1 + a_2 * u_2 + a_3 * u_3 + a_4 * u_4 + a_5 * u_5 = [0, 0, 0, 0, 0],$$

utilizando el comando AUTOR. Al simplificar resulta el sistema

$$\begin{aligned}[][a_1 + a_2 + a_3 + a_4 + 9a_5 = 0, \quad -a_1 + 3a_2 + 3a_3 + a_4 = 0,\\ -5a_2 + a_3 + a_4 + a_5 = 0, \quad -2a_1 + a_2 - 2a_3 + 2a_5 = 0,\\ 5a_1 + 3a_2 + 5a_3 - 2a_4 - a_5 = 0].\end{aligned}$$

Al aplicar SOLVE se obtiene

$$[a_1 = 0, a_2 = 0, a_3 = 0, a_4 = 0, a_5 = 0],$$

a partir de lo cual concluimos que el conjunto de vectores $\{u_1, u_2, u_3, u_4, u_5\}$ es linealmente independiente y de rango 5.

2. Editamos la expresión

$$b_1 * v_1 + b_2 * v_2 + b_3 * v_3 + b_4 * v_4 = [0, 0, 0, 0, 0]$$

mediante el comando AUTOR. Al simplificar resulta el sistema

$$\begin{array}{rcl} [2b_1 + b_2 + 7b_3 - 2b_4 = 0 & , & 3b_1 + 9b_3 + 3b_4 = 0 \\ 4b_1 - 3b_2 + 9b_3 + 16b_4 = 0 & , & -3b_1 + 2b_2 - 7b_3 - 112b_4 = 0 \\ & & b_1 + 5b_2 + 8b_3 - 19b_4 = 0]. \end{array}$$

Mediante SOLVE obtenemos que la solución del sistema es

$$\left[b_1 = @_1, b_2 = @_2, b_3 = -\frac{4@_1 + @_2}{13}, b_4 = \frac{3@_2 - @_1}{13}\right].$$

En consecuencia, existen infinitas soluciones para el sistema y, por lo tanto, el conjunto $\{v_1, v_2, v_3, v_4\}$ es linealmente dependiente.

Como el conjunto es linealmente dependiente, existe un vector que es combinación lineal de los restantes. Se puede ver, por ejemplo, que

$$v_4 = \frac{13}{4}v_1 - \frac{13}{4}v_2 - \frac{3}{4}v_3$$

Con el fin de establecer el rango del conjunto dado, eliminamos el vector v_4 y determinamos si los vectores restantes son linealmente independientes. Procediendo como antes, editamos la expresión

$$c_1 * v_1 + c_2 * v_2 + c_3 * v_3 = [0, 0, 0, 0, 0]$$

y, simplificando, tenemos:

$$\begin{aligned}
&[2c_1 + c_2 + 7c_3 = 0 \quad , \quad 3c_1 + 9c_3 = 0 \\
&4c_1 - 3c_2 + 9c_3 = 0 \quad , \quad -3c_1 + 2c_2 - 7c_3 = 0 \\
&\qquad\qquad\qquad\qquad\qquad c_1 + 5c_2 + 8c_3 = 0]\,.
\end{aligned}$$

Al aplicar SOLVE se obtiene de nuevo que el sistema tiene infinitas soluciones y que estas están dadas por:

$$\left[c_1 = @_3, c_2 = \frac{@_3}{3}, c_3 = \frac{@_3}{3}\right].$$

Así, el conjunto $\{v_1, v_2, v_3\}$ es de nuevo linealmente dependiente.

Se puede ver que el vector v_3 es una combinación lineal de los vectores v_1 y v_2 (compruébelo); entonces, repetimos el proceso eliminando el vector v_3.

Desde AUTOR editamos la expresión

$$d_1 * v_1 + d_2 * v_2 = [0, 0, 0, 0, 0]$$

que al simplificar nos da

$$[2d_1 + d_2 = 0, 3d_1 = 0, 4d_1 - 3d_2 = 0, 2d_2 - 3d_1 = 0, d_1 + 5d_2 = 0].$$

Mediante SOLVE encontramos que la solución del sistema está dada por $[d_1 = 0, d_2 = 0]$ y que, por lo tanto, el conjunto $\{v_1, v_2\}$ es linealmente independiente y, en consecuencia, el rango de $\{v_1, v_2, v_3, v_4\}$ es 2.

Nota: el cálculo del rango se puede simplificar utilizando directamente la función *rank* definida por DERIVE en el archivo de utilidades *vector.mth.* Para emplear esta función debemos cargar previamente el archivo en mención, utilizando la secuencia *transfer-load-utility* sobre el archivo *vector.mth.*

Una vez hecho esto, será suficiente construir un vector con todos los vectores que estamos considerando, e introducirlos como los argumentos de la función *rank,* es decir, hay que editar con AUTOR la expresión $rank([v_1, v_2, v_3, v_4])$. Al simplificar obtenemos 2, que es el rango del conjunto $\{v_1, v_2, v_3, v_4\}$.

Ejercicio 3.6.9 (Uso del computador). Calcular $A \cap B$ para los siguientes subespacios vectoriales de $\mathbb{R}^3$:

$$A = \{(x, y, z) \diagup x + 2y + z = 0\},$$
$$B = \{(x, y, z) \diagup 2x + y - z = 0, 2x - 4y + 8z = 0\}.$$

Solución

Para hallar el subespacio intersección será suficiente determinar los vectores que satisfacen simultáneamente las ecuaciones que definen los subespacios A y B. Por lo tanto, editamos, mediante AUTOR, el sistema de ecuaciones

$$[x + 2y + z = 0, 2x + y - z = 0, 2x - 4y + 8z = 0];$$

con SOLVE obtenemos la solución $[x = 0, y = 0, z = 0]$, por lo que $A \cap B = \{(0, 0, 0)\}$.

Ejercicio 3.6.10 (Uso del computador). Calcular $A + B$ para los siguientes subespacios vectoriales de $\mathbb{R}^3$:

$$A = \{(x, y, z) \diagup x + 2y + z = 0\},$$

$$B = \{(x, y, z) \diagup 2x + y - z = 0, 2x - 4y + 8z = 0\}.$$

Solución

Para obtener las ecuaciones del subespacio suma, extraemos un conjunto de generadores de ambos espacios a partir de sus ecuaciones paramétricas.

En el caso del subespacio A, basta con editar el sistema $[x + 2y + z = 0, 0, 0]$, cuyas soluciones son

$$[x = @_1, y = @_2, z = -@_1 - 2@_2].$$

Así pues, la ecuación vectorial del subespacio A es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ -\alpha - 2\beta \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix};$$

por lo tanto, un conjunto de generadores de A es

$$S_A = \{(1, 0, -1), (0, 1, -2)\}.$$

De igual forma, las ecuaciones paramétricas del subespacio B se obtienen editando con AUTOR el sistema

$$[2x + y - z = 0, 2x - 4y + 8z = 0, 0],$$

cuyas soluciones son

$$\left[x = @_3, y = -\frac{9@_3}{2}, z = -\frac{5@_3}{2}\right].$$

Luego

$$S_B = \left\{\left(1, -\frac{9}{2}, -\frac{5}{2}\right)\right\}$$

es un conjunto de generadores del subespacio B. El subespacio vectorial suma se determina considerando el subespacio vectorial generado por la unión de ambos conjuntos de generadores, es decir,

$$A + B = gen\left\{(1, 0, -1), (0, 1, -2), \left(1, -\frac{9}{2}, -\frac{5}{2}\right)\right\}.$$

Ejercicio 3.6.11. Hallar una base para el subespacio de las matrices antisimétricas de tamaño 3×3.

Solución

Una matriz antisimétrica de tamaño 3×3 tiene la forma

$$\begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$$

y puede escribirse como la siguiente combinación lineal:

$$\begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix} = a \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{M_1} + b \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}}_{M_2} + c \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}}_{M_3}.$$

De esta última relación se concluye que el conjunto de matrices $\{M_1, M_2, M_3\}$ genera al conjunto de las matrices antisimétricas.

Veamos ahora que $\{M_1, M_2, M_3\}$ es un conjunto linealmente independiente. Para ello necesitamos determinar las soluciones de la ecuación

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = x \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{M_1} + y \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}}_{M_2} + z \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}}_{M_3}.$$

Esta relación da origen a las ecuaciones $x = -x = 0$, $y = -y = 0$ y $z = -z = 0$, lo cual significa que la única solución de dicho sistema es la trivial y, por lo tanto, las matrices $\{M_1, M_2, M_3\}$ son linealmente independientes, constituyendo de esta forma una base para el espacio dado.

3.7 Ejercicios propuestos capítulo 3

1. En los siguientes ejercicios, determine si el conjunto dado, con las operaciones de adición y multiplicación por escalar, es un espacio vectorial. En caso de no serlo, especifique los axiomas que no se cumplen.

   a) El conjunto de todos los vectores de $\mathbb{R}^2$ de la forma (x, x), con las operaciones vectoriales habituales de suma y multiplicación por un escalar.

   b) El conjunto de todos los vectores de $\mathbb{R}^2$ de la forma (x, y), con $x \geq 0$ y $y \geq 0$, con las operaciones vectoriales habituales de suma y multiplicación por un escalar.

   c) El conjunto de todos los vectores de $\mathbb{R}^2$ de la forma (x, y), donde $xy \geq 0$, con las operaciones vectoriales habituales de suma y multiplicación por un escalar.

   d) El conjunto de todos los vectores de $\mathbb{R}^2$, de la forma (x, y), donde $x \geq y$, con las operaciones vectoriales habituales de suma y multiplicación por un escalar.

   e) El conjunto de todos los vectores de $\mathbb{R}^2$, con la suma habitual y la multiplicación por un escalar definida por $\lambda(x, y) = (\lambda x, y)$.

   f) El conjunto de todos los vectores de $\mathbb{R}^2$, con la multiplicación habitual y la suma definida por $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2 + 1, y_1 + y_2 + 1)$.

   g) El conjunto de todos los números reales positivos, con la suma definida por $\alpha \oplus \beta = \alpha\beta$ y la multiplicación por escalar definida como $\alpha \odot x = x^\alpha$.

   h) El conjunto de todos los vectores de $\mathbb{R}^2$, con la multiplicación y la suma definidas por

$$\begin{aligned} (\alpha_1, \beta_1) \oplus (\alpha_2, \beta_2) &= \left(\frac{\alpha_1 + \alpha_2}{2}, \frac{\beta_1 + \beta_2}{2} \right) \\ \lambda \odot (\alpha_1, \beta_1) &= (\lambda\alpha_1, \lambda\beta_1). \end{aligned}$$

i) El conjunto de todos los vectores de $\mathbb{R}^2$, con la multiplicación y la suma definidas por $\lambda \odot (\alpha_1, \beta_1) = (\alpha_1, 0)$ y $(\alpha_1, \beta_1) \oplus (\alpha_2, \beta_2) = (\alpha_1 + \alpha_2, 0)$.

2. Determine cuáles de los siguientes subconjuntos de $\mathbb{R}^2$ son subespacios.

   a) $W = \{(\alpha, \beta) \in \mathbb{R}^2 / \alpha = 1\}$
   b) $W = \{(\alpha, \beta) \in \mathbb{R}^2 / \alpha + \beta = 0\}$
   c) $W = \{(\alpha, \beta) \in \mathbb{R}^2 / \alpha = \lambda\beta; \lambda \neq 0\}$
   d) $W = \{(\alpha, \beta) \in \mathbb{R}^2 / \alpha = \beta + 1\}$.

3. Determine cuáles de los siguientes subconjuntos de $\mathbb{R}^3$ son subespacios.

   a) $W = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 / |\alpha| = |\gamma|\}$
   b) $W = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 / \alpha \geq \gamma\}$
   c) $W = \{(\alpha + 2\beta - 3, 2\alpha - \beta, \alpha + \beta) \in \mathbb{R}^3 / \alpha, \beta \in \mathbb{R}\}$
   d) $W = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 / \beta = |\gamma|\}$
   e) $W = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 / \alpha + \beta + \gamma = 0\}$
   f) $W = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 / \alpha + \beta = 3\gamma\}$.

4. Determine el valor del escalar α de tal forma que el vector $(\alpha, 1, 2)$ sea una combinación lineal de $(1, 2, 3)$ y $(2, -3, 4)$.

5. Si $S = \{(1, \sqrt{2}), (-1, \sqrt{3})\}$, demuestre que $gen\{S\} = \mathbb{R}^2$. Escribir el vector $(1, 2)$ como combinación lineal de los elementos de S.

6. Sean los espacios dados por $S = gen\{(1, 0, 2), (2, 3, -5)\}$ y $W = gen\{(-1, -1, 2), (2, 1, 2)\}$. Hallar $S \cap W$.

7. Sean W y U dos subespacios de V; demostrar que $W + U = \{x = w + u / w \in W, u \in U\}$ es un subespacio vectorial de V.

8. Demostrar que los siguientes vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$ son linealmente independientes.

   a) $(1,1)$ y $(0,1)$

   b) $(-1,1,0)$ y $(0,1,2)$

   c) $(1,1,0)$, $(1,1,1)$ y $(0,1,-1)$

   d) $(1,1)$ y $(0,-2)$

   e) $(0,1,1)$, $(0,2,1)$ y $(1,5,3)$.

9. Dado el vector x, expresarlo como una combinación lineal de los vectores a y b. Hallar sus coordenadas respecto a a y b.

   a) $x=(1,0)$, $a=(1,1)$ y $b=(0,1)$

   b) $x=(2,1)$, $a=(1,1)$ y $b=(-1,1)$

   c) $x=(4,3)$, $a=(2,1)$ y $b=(0,-1)$

   d) $x=(1,1)$, $a=(2,1)$ y $b=(-1,1)$.

10. Dado el vector x, expresarlo como combinación lineal de los vectores a, b y c. Hallar sus coordenadas respecto a a, b y c.

    a) $x=(1,0,0)$, $a=(1,1,1)$, $b=(0,1,1)$ y $c=(1,0,1)$

    b) $x=(1,1,1)$, $a=(0,1,-1)$, $b=(1,1,0)$ y $c=(1,0,2)$

    c) $x=(0,0,1)$, $a=(1,1,1)$, $b=(1,-1,0)$ y $c=(1,0,-1)$.

11. Determine si el conjunto $\{-1+2x+x^2, 1+x^2, x+x^2\}$ es linealmente dependiente o independiente.

12. Demuestre que si α, β y γ son números reales distintos, $\{(1,\alpha,\alpha^2),(1,\beta,\beta^2),(1,\gamma,\gamma^2)\}$ es linealmente independiente.

13. Sea V un espacio vectorial en $\mathbb{R}$. Sean los vectores $\{a,b,c\}$ en V un conjunto de vectores linealmente independientes. Si α, β y γ están en $\mathbb{R}$, determinar la dependencia o independencia lineal de los conjuntos:

a) $\{a+\alpha b+\beta c, b+\lambda c, c\}$

b) $\{a, b+\alpha c, c+\beta b\}$.

14. Sea V un espacio vectorial en $\mathbb{R}$. Si los vectores $\{a, b, c\} \in V$ son un conjunto linealmente independiente, pruebe que:

    a) $\{a+b, a+c, b+c\}$ es un conjunto independiente.

    b) $\{a-b, a+c, b+c\}$ es un conjunto linealmente dependiente.

15. Sea V un espacio vectorial. Sean los vectores $\{a_1, \ldots, a_n\}$ en V un conjunto de vectores linealmente independientes. Sea z un vector en V tal que $z \notin gen\{a_1, \ldots, a_n\}$. Demostrar que $\{a_1+z, \ldots, a_n+z\}$ es un conjunto linealmente independiente.

16. Sea $S=\{v_1, v_2, v_3, v_4\}$, donde $v_1=(1,2,2)$, $v_2=(11,10,7)$, $v_3=(3,2,1)$ y $v_4=(7,6,6)$. Determine una base para el espacio generado por S. ¿Cuál es su dimensión?

17. Sea $S=\{v_1, v_2, v_3\}$, donde $v_1=(1,1,2)$, $v_2=(2,-1,3)$ y $v_3=(3,0,5)$. Determine una base para el espacio generado por S. ¿Cuál es la dimensión de tal subespacio?

18. Determine una base para cada uno de los siguientes subespacios de $\mathbb{R}^3$ e indique la dimensión del subespacio generado por todos los vectores de la forma (a, b, c),

    a) Con $a=b$.

    b) Donde $a+b+c=0$.

    c) Con $a=2b+c$.

    d) Con $b=0$.

19. Determine una base para el siguiente subconjunto de los polinomios de grado menor o igual que tres (P_3). Indique la dimensión del subespacio dado.

$$S=\{t^3+t^2-2t+1, t^2+1, t^3-2t, 2t^3+3t^2-4t+3\}.$$

Capítulo 4

Producto escalar

4.1 Introducción

En las secciones previas se ha definido el concepto de *espacio vectorial*. Este concepto resulta ser una generalización de las propiedades que tienen las operaciones de suma y producto por escalar para los vectores libres de $\mathbb{R}^3$. Al hacer tal generalización se ha pasado por alto algunos otros conceptos asociados a los vectores geométricos, como son los conceptos de *ángulo entre vectores* y la *longitud de un vector*. Lo anterior significa que hasta el momento, para los espacios vectoriales en general, carecen de significado las nociones geométricas de *ángulo* y *distancia*.

En adelante, el campo sobre el que se define el espacio vectorial es el campo de los números reales. En este capítulo se establece la axiomática general que define el concepto de *producto interno* o *escalar*. Es importante recalcar que a pesar de presentarse la teoría para el caso de espacios vectoriales en general, el origen de esta noción inicialmente se presentó para el caso de espacios euclídeos de dimensión inferior a tres. En la literatura de la matemática, los espacios dotados de un producto interno se conocen con el nombre de *espacios prehilbertianos* y se verá que los resultados obtenidos a partir de está generalización son aplicables al caso de los vectores libres de $\mathbb{R}^3$.

4.2 Producto escalar

Sea V un espacio vectorial sobre $\mathbb{R}$. Se llama *producto escalar* sobre él a una operación de $V \times V \longrightarrow \mathbb{R}$, tal que a cada par de vectores x y y le asocia un número real que se denota con $\langle x, y\rangle$ y cumple las siguientes propiedades:

1. *Conmutatividad*: $\langle x, y\rangle = \langle y, x\rangle, \forall x, y \in V$.

2. *Distributividad*: $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle, \quad \forall x, y, z \in V$.

3. $\langle x, \lambda y\rangle = \lambda\langle y, x\rangle,\ \forall x, y \in V$ y $\lambda \in \mathbb{R}$.

4. $\langle x, x\rangle \geq 0,\ \forall x \in V$.

5. $\langle x, x\rangle = 0$ si y solo si $x = \mathbf{0}$.

Las propiedades anteriores caracterizan de manera única el concepto de *producto interior*, permitiendo generalizar de modo algebraico algunas propiedades de carácter geométrico del plano y el espacio. De la cuarta propiedad, que nos garantiza la positividad del producto interior de un vector por sí mismo, es posible definir la norma o magnitud de un vector, $\|\| : V \to \mathbb{R}^+ \cup \{0\}$, mediante la relación $\|x\| = \sqrt{\langle x, x\rangle}$.

Los siguientes ejemplos nos muestran algunos productos internos de uso frecuente.

Ejemplo 4.2.1. Consideremos el espacio vectorial conformado por las funciones reales continuas sobre el intervalo $[a, b]$. Dadas dos funciones $f(x)$ y $g(x)$, se define una operación de $V \times V \longrightarrow \mathbb{R}$ por

$$\langle f(x), g(x)\rangle = \langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Pruebe que, en este espacio vectorial, la operación definida es un producto interno.

Solución

Es necesario probar que la operación establecida cumple con los axiomas que definen el producto interno. Esto se hace aplicando las propiedades de la integral, con lo cual obtenemos:

1. $\langle f, g\rangle = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \langle g, f\rangle, \quad \forall f, g \in V.$

2. Para la ley distributiva tenemos que

$$\begin{aligned}
\langle f, g+h\rangle &= \int_a^b f(x)(g(x)+h(x))dx \\
&= \int_a^b (f(x)g(x)+f(x)h(x))dx \\
&= \int_a^b f(x)g(x)dx + \int_a^b f(x)h(x)dx \\
&= \langle f, g\rangle + \langle f, h\rangle, \quad \forall f, g, h \in V.
\end{aligned}$$

3. $\forall f, g \in V$ y $\lambda \in \mathbb{R}$, se cumple que

$$\langle f, \lambda g\rangle = \int_a^b f(x)\lambda g(x)dx = \lambda \int_a^b f(x)g(x)dx = \lambda\langle f, g\rangle.$$

4. $\langle f, f\rangle = \int_a^b f(x)f(x)dx \geq 0$, $\forall f \in V$.

5. $\langle f, f\rangle = \int_a^b f(x)f(x)dx = 0$, $\forall f \in V$ si y solo si $f(x) = \mathbf{0}$.

En la última propiedad, se deja propuesto como ejercicio verificar que si $\langle f, f\rangle = 0$, entonces $f = \mathbf{0}$.

Ejemplo 4.2.2. Consideremos el espacio vectorial $\mathbb{R}^n$, conformado por las n-tuplas ordenadas de números reales. Dados dos vectores $x = (\alpha_1, \ldots, \alpha_n)$ y $y = (\beta_1, \ldots, \beta_n)$, se define una operación de $V \times V \longrightarrow \mathbb{R}$ por

$$\langle x, y\rangle = \sum_{i=1}^{n} \alpha_i\beta_i.$$

Pruebe que, en este espacio vectorial, la operación definida es un producto interno.

Solución

Es necesario demostrar que la operación definida cumple con los axiomas que definen el producto interno. Esto se hace aplicando las propiedades de la suma, con lo cual obtenemos:

1. $\langle x, y\rangle = \sum_{i=1}^{n} \alpha_i\beta_i = \sum_{i=1}^{n} \beta_i\alpha_i = \langle y, x\rangle,\ \forall x, y \in \mathbb{R}^n.$

2. Para la ley distributiva tenemos:

$$\begin{aligned}\langle x, y+z\rangle &= \sum_{i=1}^{n} \alpha_i(\beta_i + \lambda_i) \\ &= \sum_{i=1}^{n} \alpha_i\beta_i + \sum_{i=1}^{n} \alpha_i\lambda_i \\ &= \langle x, y\rangle + \langle x, z\rangle,\ \forall x, y, z \in \mathbb{R}^n.\end{aligned}$$

3. $\forall x, y \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$ se cumple que

$$\langle x, \lambda y\rangle = \sum_{i=1}^{n} \alpha_i\lambda\beta_i = \lambda\sum_{i=1}^{n} \alpha_i\beta_i = \lambda\langle x, y\rangle.$$

4. $$\langle x, x\rangle = \sum_{i=1}^{n} \alpha_i\alpha_i \geq 0,\ \forall x \in \mathbb{R}^n.$$

5. $$\langle x, x\rangle = \sum_{i=1}^{n} \alpha_i^2 = 0,$$

   si y solo si $\alpha_i = 0$ para todo i; luego $x = (0_1, \ldots, 0_n) = \mathbf{0}$,

con lo cual se prueba que la operación efectivamente es un producto escalar.

El primer grupo de propiedades del producto escalar puede resumirse en el siguiente teorema.

Teorema 4.2.1. *Sea V un espacio vectorial sobre $\mathbb{R}$; si $\langle x, y\rangle$ es un producto escalar, entonces se cumplen las siguientes propiedades:*

1. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle,\ \forall x, y, z \in V.$
2. $\langle \lambda x, y\rangle = \lambda\langle y, x\rangle,\ \forall x, y \in V$ *y* $\lambda \in \mathbb{R}.$
3. $\langle x, \mathbf{0}\rangle = 0,\ \forall x \in V.$
4. *Si* $\langle x, y\rangle = 0\ \forall y \in V,$ *entonces* $x = \mathbf{0}.$

Demostración. La primera propiedad se sigue inmediatamente de la conmutatividad y de la distributividad del producto escalar, mientras que la segunda hace uso de la conmutatividad y del hecho de que $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$. Pasemos a probar ahora 3 y 4.

3. Como $\langle x, \mathbf{0}\rangle = \langle x, \mathbf{0}+\mathbf{0}\rangle = \langle x, \mathbf{0}\rangle + \langle x, \mathbf{0}\rangle$, entonces $\langle x, \mathbf{0}\rangle = 0$.
4. Como $\langle x, y\rangle = 0\ \forall y \in V$. En particular, si $y = x$, entonces $\langle x, x\rangle = 0$ y, en consecuencia, $x = \mathbf{0}$.

La siguiente definición nos permite introducir el concepto de *norma* de un vector, a partir de un producto interno definido en un espacio vectorial.

Definición 4.2.1. *Un espacio vectorial euclídeo real es un espacio vectorial real en el cual se ha definido un producto escalar, es decir, un par $(V, \langle,\rangle)$, donde V es un espacio vectorial real y $\langle,\rangle : V \times V \to \mathbb{R}$ es un producto escalar.*

Definición 4.2.2. *En un espacio vectorial euclídeo $(V, \langle,\rangle)$, la norma o longitud de un vector x, que se denota con $\|x\|$, es el número real no negativo*

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

De la definición anterior, se infiere que:

- Como $\langle x, x\rangle \geq 0$, todo vector en V tiene una norma.
- Como $\langle x, x\rangle = 0$ si y solo si $x = \mathbf{0}$, entonces el único vector con $\|x\| = 0$ es $x = \mathbf{0}$.
- Si $\|x\| = 1$, se dice que el vector es unitario o que está normalizado.
- Si $x \neq \mathbf{0}$, $1/\|x\|x$ es un vector unitario en la dirección de x.

Teorema 4.2.2 (Desigualdad de Schwarz). *Para cualquier par de vectores de un espacio euclídeo se cumple que*

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

Demostración. Para cualquier $\lambda \in \mathbb{R}$ se cumple que

$$\langle \lambda x + y, \lambda x + y\rangle \geq 0.$$

Al aplicar las propiedades del producto escalar se consigue que

$$\begin{aligned}\langle \lambda x + y, \lambda x + y\rangle &= \lambda^2\langle x, x\rangle + 2\lambda\langle x, y\rangle + \langle y, y\rangle \\ &= \lambda^2\|x\|^2 + 2\lambda\langle x, y\rangle + \|y\|^2 \geq 0.\end{aligned}$$

Como esta última relación es una ecuación cuadrática en la variable λ y no es negativa, tiene a lo sumo al 0 como una raíz real; en consecuencia, el discriminante de esta ecuación no debe ser positivo, es decir:

$$4\langle x, y\rangle^2 - 4\|x\|^2\|y\|^2 \leq 0,$$

con lo cual se consigue que

$$|\langle x, y\rangle| \leq \|x\|\|y\|,$$

como se pretendía demostrar.

A partir de este teorema es posible precisar un nuevo criterio para establecer la independencia lineal; esto se resume en el siguiente teorema.

Teorema 4.2.3. *En un espacio euclídeo, dos vectores x y y son linealmente dependientes si y solo si* $|\langle x, y\rangle| = \|x\|\|y\|$.

Demostración. Supongamos que los vectores x y y son linealmente dependientes; entonces, existen escalares α y β no nulos tales que $\alpha x + \beta y = 0$, con lo cual concluimos que

$$x = \frac{\beta}{\alpha} y = \lambda y$$

y, por lo tanto,

$$\begin{aligned} \langle x, y\rangle &= \langle \lambda y, y\rangle = \lambda\langle y, y\rangle = \lambda\|y\|^2 \\ \langle x, y\rangle &= \langle x, \frac{1}{\lambda}x\rangle = \frac{1}{\lambda}\langle x, x\rangle = \frac{1}{\lambda}\|x\|^2. \end{aligned}$$

En consecuencia,

$$\langle x, y\rangle^2 = \|y\|^2\|x\|^2,$$

con lo cual

$$|\langle x, y\rangle| = \|y\|\|x\|.$$

Supongamos ahora que $|\langle x, y\rangle| = \|y\|\|x\|$, donde $\|y\| \neq 0$ y $\|x\| \neq 0$, y que $x + \lambda y = 0$. Entonces,

$$\begin{aligned} \langle x + \lambda y, x + \lambda y\rangle &= \langle x, x\rangle + 2\lambda\langle x, y\rangle + \lambda^2\langle y, y\rangle \\ &= \|x\|^2 \pm 2\lambda\|y\|\|x\| + \lambda^2\|y\|^2 = 0 \\ &= (\|x\| \pm \lambda\|y\|)^2 = 0 \\ &= (\|x\| \pm \lambda\|y\|) = 0, \end{aligned}$$

con lo cual concluimos que

$$\|x\| = \mp\lambda\|y\|$$

y, en consecuencia,

$$\lambda = \mp \frac{\|y\|}{\|x\|}.$$

Esto obliga a que $\lambda \neq 0$, con lo cual se prueba que x y y son linealmente dependientes.

Además de las propiedades establecidas hasta el momento para la norma de un vector, en el siguiente teorema se presentan otros resultados que son de importancia en el caso de los espacios euclídeos.

Teorema 4.2.4. *Sean x y y vectores de un espacio vectorial euclídeo V sobre $\mathbb{R}$ y sea λ un escalar; entonces, se cumple que:*

1. $\|x\| \geq 0, \; \forall x$ *en* V.
2. $\|x\| = 0$ *si y solo si* $x = \boldsymbol{0}$.
3. $\|\lambda x\| = |\lambda| \|x\|, \forall x$ *en* V *y* λ *en* $\mathbb{R}$.
4. $\|x + y\| \leq \|x\| + \|y\|, \forall x, \; y$ *en* V.

El último resultado es conocido como la *desigualdad de Minkowski* o *desigualdad triangular*.

Demostración. Los dos primeros enunciados son obvios a partir de la definición de norma de un vector. Para los otros dos se tiene:

3. $\|\lambda x\|^2 = \langle \lambda x, \lambda x \rangle = \lambda^2 \|x\|^2$; en consecuencia,

$$\|\lambda x\| = \sqrt{\lambda^2 \|x\|^2} = |\lambda| \|x\|.$$

4. En el caso de la desigualdad de Minkowski tenemos que, por la desigualdad de Schwarz,

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle \\
&= \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&\leq (\|x\| + \|y\|)^2,
\end{aligned}$$

a partir de lo cual, tomando raíz cuadrada, concluimos la desigualdad pedida.

El siguiente teorema tiene una implicación importante cuando se mira desde el punto de vista geométrico en $\mathbb{R}^3$ y será utilizado para establecer diferencias entre los espacios vectoriales euclídeos y los espacios vectoriales normados.

Teorema 4.2.5. *Sean x y y vectores de un espacio vectorial euclídeo; entonces,*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

El significado geométrico de esta relación corresponde al hecho de que la suma de los cuadrados de las diagonales de un paralelogramo es igual al doble de la suma de los cuadrados de los lados. Esto es conocido como la *igualdad del paralelogramo.*

Demostración. Como

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \\
\|x-y\|^2 &= \langle x-y, x-y\rangle = \|x\|^2 - 2\langle x,y\rangle + \|y\|^2,
\end{aligned}$$

al sumar estas dos relaciones conseguimos

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2,$$

como queríamos demostrar.

El concepto de *norma* que se ha obtenido aparece como una consecuencia del producto escalar; sin embargo, es posible establecer normas en un espacio vectorial sin recurrir al producto escalar. La estrategia que se emplea en estos casos está ligada a una serie de axiomas que son equivalentes a las propiedades que cumplen las normas derivadas de productos escalares. Estos comentarios quedan aclarados con la siguiente definición.

Definición 4.2.3. *Sea V un espacio vectorial real; se llama norma en V a toda aplicación $\| \| : V \to \mathbb{R}$ que satisface los siguientes axiomas:*

1. $\|x\| \geq 0$, $\forall x$ *en* V.
2. $\|x\| = 0$ *si y solo si* $x = \boldsymbol{0}$.
3. $\|\lambda x\| = |\lambda|\|x\|$, $\forall x$ *en* V *y* λ *en* $\mathbb{R}$.
4. $\|x + y\| \leq \|x\| + \|y\|$, $\forall x$ *y* y *en* V.

Un espacio vectorial en el que se ha definido una norma se llama *espacio vectorial normado.*

En este momento es conveniente recalcar dos hechos importantes

- Todo espacio euclídeo es un espacio vectorial normado.
- No toda norma en un espacio vectorial normado proviene de un producto escalar. El siguiente ejemplo ilustra este hecho.

Ejemplo 4.2.3. Sea el espacio vectorial $\mathbb{R}^2$ y sea la aplicación $\| \| : \mathbb{R}^2 \to \mathbb{R}$, definida como $\|(x, y)\| = \text{máx}\{|x|, |y|\}$. Demostrar que es una norma, pero que no proviene de un producto escalar.

Solución

Veamos que cumple con los axiomas que definen una norma.

1. $\|(x, y)\| \geq 0$, $\forall (x, y)$ en $\mathbb{R}^2$.

2. $\|(x,y)\| = 0$ si y solo si $|x| = 0$ y $|y| = 0$; luego $(x,y) = \mathbf{0}$.

3. $\|\lambda(x,y)\| = \text{máx}\{|\lambda x|, |\lambda y|\} = |\lambda| \,\text{máx}\{|x|, |y|\} = |\lambda|\|(x,y)\|$, $\forall (x,y)$ en $\mathbb{R}^2$ y λ en $\mathbb{R}$.

4. Como

$$\|(x,y) + (u,v)\| = \|(x+u, y+v)\| = \text{máx}\{|x+u|, |y+v|\}$$

y si ahora tenemos en cuenta las relaciones

$$\begin{array}{rcl} |x+u| & \leq & |x| + |u| \leq \text{máx}\{|x|,|y|\} + \text{máx}\{|u|,|v|\} \\ |y+v| & \leq & |y| + |v| \leq \text{máx}\{|x|,|y|\} + \text{máx}\{|u|,|v|\}; \end{array}$$

y como

$$\text{máx}\{|x|,|y|\} + \text{máx}\{|u|,|v|\} = \|(x,y)\| + \|(u,v)\|,$$

entonces se cumple que

$$\begin{array}{rcl} |x+u| & \leq & \|(x,y)\| + \|(u,v)\| \\ |y+v| & \leq & \|(x,y)\| + \|(u,v)\|; \end{array}$$

por lo tanto,

$$\|(x,y) + (u,v)\| = \text{máx}\{|x+u|, |y+v|\} \leq \|(x,y)\| + \|(u,v)\|,$$

$\forall (x,y)$ y (u,v) en $\mathbb{R}^2$.

Los resultados anteriores aclaran que la aplicación dada es una norma. Para mostrar que esta no proviene de un producto escalar, es suficiente probar que no cumple la igualdad del paralelogramo para los productos escalares. Así,

$$\|(1,3)\| = \text{máx}\{|1|,|3|\} = 3, \;\; \|(-1,2)\| = \text{máx}\{|-1|,|2|\} = 2,$$

mientras que

$$\|(1,3) + (-1,2)\| = \|(0,5)\| = \text{máx}\{|0|,|5|\} = 5$$

y, además,

$$\|(1,3)-(-1,2)\| = \|(2,1)\| = \text{máx}\{|2|,|1|\} = 2,$$

con lo cual

$$\|(1,3)+(-1,2)\|^2 + \|(1,3)-(-1,2)\|^2 = 29,$$

mientras que

$$2(\|(-1,2)\|^2 + \|(1,3)\|^2) = 26;$$

esto significa que

$$\|x+y\|^2 + \|x-y\|^2 \neq 2\|x\|^2 + 2\|y\|^2$$

y, en consecuencia, esta norma no proviene de un producto escalar.

Definición 4.2.4. *Sea V un espacio vectorial normado, con una norma denotada por $\| \ \|$. Se llama distancia del vector x al vector y, que se representa por $d(x,y)$, a la norma del vector diferencia, esto es,*

$$d(x,y) = \|x-y\|.$$

De la definición de distancia y de las propiedades de la norma se tienen los siguientes resultados:

1. $d(x,y) \geq 0$, $\forall x$ en V.

2. $d(x,y) = 0$ si y solo si $x = y$.

3. $d(x,y) = d(y,x)$, $\forall x$ y y en V.

4. $d(x,y) \leq d(x,z) + d(z,y)$, $\forall x$, y y z en V. Esta última relación se conoce como la *desigualdad triangular.*

5. Como $d(x,z) \leq d(x,y) + d(y,z)$, entonces se tiene que

$$d(x,z) - d(y,z) \leq d(x,y);$$

además,

$$d(y,z) \leq d(y,x) + d(x,z),$$

con lo cual, al ser $d(y,x) = d(x,y)$,

$$d(y,z) - d(x,z) \leq d(x,y).$$

De las relaciones anteriores se concluye que

$$-d(x,y) \leq d(y,z) - d(x,z) \leq d(x,y),$$

lo que es equivalente a la relación

$$|d(y,z) - d(x,z)| \leq d(x,y), \ \forall x,y,z \in V.$$

Teorema 4.2.6. *En un espacio vectorial normado V, la distancia es invariante bajo traslaciones, es decir,*

$$d(y+t, x+t) = d(y,x) \ \forall x,y \in V.$$

Demostración. Para todo $x, y \in V$ tenemos que

$$d(y+t, x+t) = \|(y+t) - (x+t)\| = \|y - x\| = d(y,x),$$

como se quería probar.

Hasta el momento se ha presentado una visión general del concepto de *producto escalar*, ahora se muestra un caso particular, que tiene ver con el espacio de los vectores libres de $\mathbb{R}^3$. Para ello se enuncia la siguiente definición.

Definición 4.2.5. *Sea $\langle \mathbb{E}^3, +, . \rangle$ el espacio vectorial conformado por los vectores libres de $\mathbb{R}^3$. Si α es el ángulo dirigido desde el vector $\vec{a}$ al vector $\vec{b}$, se define el producto escalar entre $\vec{a}$ y $\vec{b}$ por:*

$$\vec{a} \cdot \vec{b} = \|\vec{a}\|\|\vec{b}\| \cos\alpha.$$

Teorema 4.2.7. *En el espacio vectorial $\langle \mathbb{E}^3, +, .\rangle$, conformado por los vectores libres de $\mathbb{R}^3$, el producto escalar $\vec{a} \cdot \vec{b}$ es efectivamente un producto escalar.*

Demostración. Para demostrar este teorema, se debe probar que el producto definido cumple con los axiomas. Sean los vectores $\vec{a}, \vec{b}, \vec{c} \in \langle \mathbb{E}^3, +, .\rangle$ y $\alpha \in \mathbb{R}$; entonces:

1. $\vec{a} \cdot \vec{b} = \|\vec{a}\|\|\vec{b}\| \cos\alpha = \|\vec{b}\|\|\vec{a}\| \cos\alpha = \vec{b} \cdot \vec{a}$.

2. Para el caso en el cual $\lambda \geq 0$,

$$\lambda\vec{a} \cdot \vec{b} = \|\lambda\vec{a}\|\|\vec{b}\| \cos\alpha = \lambda\|\vec{a}\|\|\vec{b}\| \cos\alpha = \lambda(\vec{a} \cdot \vec{b}).$$

Cuando $\lambda < 0$,

$$\lambda\vec{a} \cdot \vec{b} = |\lambda|\|\vec{a}\|\|\vec{b}\| \cos(\alpha + \pi) = \lambda(\vec{a} \cdot \vec{b}),$$

como se ve en la Figura 4.1.

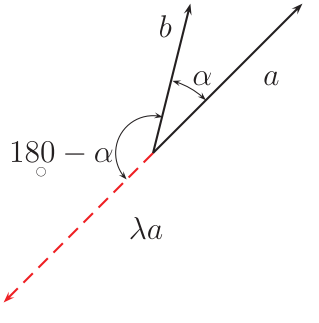


Figura 4.1. Prueba del axioma $\lambda\vec{a} \cdot \vec{b} = \lambda(\vec{a} \cdot \vec{b})$ para el producto escalar

3. $\vec{a} \cdot \vec{a} = \|\vec{a}\|\|\vec{a}\| = \|\vec{a}\|^2 \geq 0$ y, además, $\vec{a} \cdot \vec{a} = \|\vec{a}\|\|\vec{a}\| = \|\vec{a}\|^2 = 0$ si y solo si $\vec{a} = \mathbf{0}$.

4. Para la ley distributiva, consideremos la Figura 4.2.

Sea $\overrightarrow{AB} = \vec{a}$, $\overrightarrow{AH} = \vec{b}$, $\overrightarrow{AD} = \vec{c}$, y sean α y β los ángulos entre los vectores $\vec{a}$ y $\vec{b}$ y $\vec{a}$ y $\vec{c}$ respectivamente.

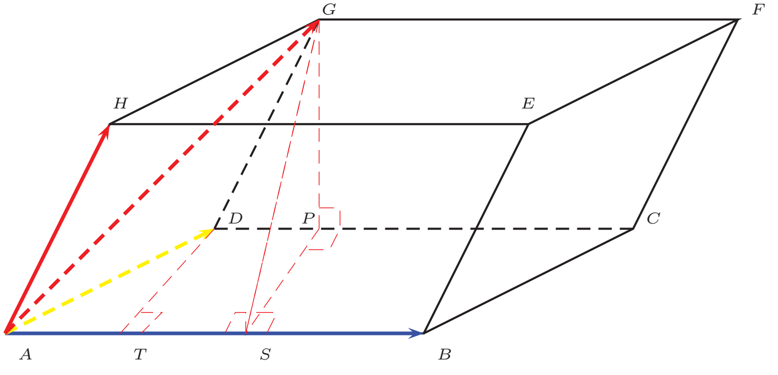


Figura 4.2. Ley distributiva para el producto escalar

El vector $\overrightarrow{AG}$ es la resultante de los vectores $\vec{b}$ y $\vec{c}$, es decir, $\overrightarrow{AG} = \vec{b} + \vec{c}$, y forma un ángulo ϕ con el vector $\vec{a}$.

Por G tracemos las líneas GP perpendicular a DC y GS perpendicular a AB. De esta forma:

$$\|\overrightarrow{DP}\| = \|\overrightarrow{DG}\| \cos \widehat{PDG} = \|\vec{b}\| \cos \alpha$$

$$\|\overrightarrow{AS}\| = \|\overrightarrow{AG}\| \cos \widehat{SAG} = \|\vec{b} + \vec{c}\| \cos \phi.$$

Por la forma como se tomaron los puntos P y S, el plano determinado por G, S y P es perpendicular al plano que contiene al paralelogramo $ABCD$ y, en consecuencia, el segmento PS es perpendicular a DC y a AB. Por el punto D tracemos, sobre el plano ABC, la línea DT paralela a PS. Con esta construcción, $DPST$ es un rectángulo en el cual

$$\|\overrightarrow{DP}\| = \|\overrightarrow{TS}\| \text{ y } \|\overrightarrow{PS}\| = \|\overrightarrow{DT}\|.$$

Además, al considerar el triángulo ADP,

$$\|\overrightarrow{AT}\| = \|\overrightarrow{AD}\| \cos \widehat{TAD} = \|\vec{c}\| \cos \beta.$$

Por último, notemos que

$$\|\overrightarrow{AS}\| = \|\overrightarrow{AT}\| + \|\overrightarrow{TS}\|,$$

a partir de lo cual se concluye que

$$\|\vec{b}+\vec{c}\|\cos\phi = \|\vec{b}\|\cos\alpha + \|\vec{c}\|\cos\beta.$$

Al multiplicar ambos lados de la igualdad anterior por $\|\vec{a}\|$ se obtiene

$$\|\vec{a}\|\|\vec{b}+\vec{c}\|\cos\phi = \|\vec{a}\|\|\vec{b}\|\cos\alpha + \|\vec{a}\|\|\vec{c}\|\cos\beta,$$

lo que es equivalente a afirmar que

$$\vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b} + \vec{a}\cdot\vec{c}.$$

Queda así probado que el producto definido es un producto escalar.

Una primera consecuencia importante, que se deriva de la definición del producto escalar establecido para los vectores libres del espacio tridimensional, es el hecho de que si los vectores son perpendiculares, su producto escalar es nulo; recíprocamente, si el producto escalar es nulo, los vectores son perpendiculares. A pesar de que en el caso de espacios vectoriales euclídeos arbitrarios la noción de *ángulo entre vectores* pierde su intuición geométrica, el concepto de *perpendicularidad* se toma a partir de la observación anterior sobre el producto escalar, de manera que podemos decir que dos vectores arbitrarios son ortogonales si y solo si su producto escalar es cero.

Teorema 4.2.8 (**Teorema del coseno**). *Dado el triángulo ABC de la Figura 4.3, se tiene que*

$$\|\vec{a}\|^2 = \|\vec{b}\|^2 + \|\vec{c}\|^2 - 2\|\vec{c}\|\|\vec{b}\|\cos\alpha.$$

Demostración. $\vec{b}+\vec{a}=\vec{c}$, con lo cual $\vec{a}=\vec{c}-\vec{b}$; de esta forma, $\vec{a}\cdot\vec{a}=(\vec{c}-\vec{b})\cdot(\vec{c}-\vec{b})$, a partir de lo cual concluimos que

$$\vec{a}\cdot\vec{a} = \vec{c}\cdot\vec{c} - 2\vec{c}\cdot\vec{b} + \vec{b}\cdot\vec{b}.$$

Al aplicar la definición de producto escalar y de norma, se consigue finalmente que

$$\|\vec{a}\|^2 = \|\vec{b}\|^2 + \|\vec{c}\|^2 - 2\|\vec{c}\|\|\vec{b}\| \cos\alpha.$$

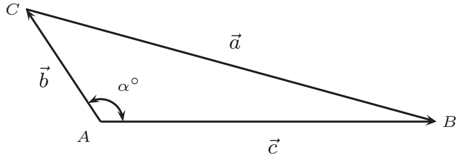


Figura 4.3. Teorema del coseno

Corolario 4.2.1 (Teorema de Pitágoras). *Dado el triángulo ABC de la Figura 4.4, recto en A, entonces* $\|\vec{a}\|^2 = \|\vec{b}\|^2 + \|\vec{c}\|^2$.

Demostración. Si aplicamos el teorema del coseno obtenemos la relación

$$\|\vec{a}\|^2 = \|\vec{b}\|^2 + \|\vec{c}\|^2 - 2\|\vec{c}\|\|\vec{b}\| \cos\alpha;$$

pero como el ángulo α es recto, entonces $\cos\alpha = \cos(\pi/2) = 0$; por lo tanto,

$$\|\vec{a}\|^2 = \|\vec{b}\|^2 + \|\vec{c}\|^2,$$

lo cual es el conocido como el *teorema de Pitágoras.*

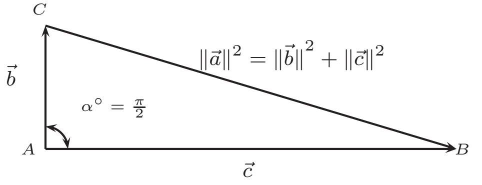


Figura 4.4. Teorema de Pitágoras

Ejemplo 4.2.4. Demostrar que un paralelogramo es un rombo si y solo si sus diagonales son ortogonales.

Solución

Sea el paralelogramo $ABCD$ y sea P el punto de corte de las diagonales (véase Figura 4.5).

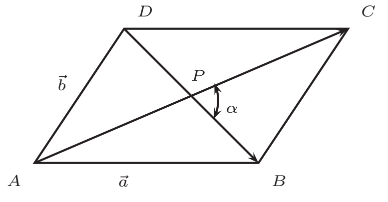


Figura 4.5. Paralelogramo con diagonales ortogonales

Entonces, como $\overrightarrow{AC} = \vec{a} + \vec{b}$ y también $\overrightarrow{DB} = \vec{a} - \vec{b}$, ocurre que si las diagonales son perpendiculares, $\overrightarrow{AC} \cdot \overrightarrow{DB} = 0$, con lo cual

$$(\vec{a} + \vec{b}) \cdot (\vec{a} - \vec{b}) = 0 = \vec{a} \cdot \vec{a} - \vec{b} \cdot \vec{b};$$

en consecuencia, $\|\vec{a}\|^2 = \|\vec{b}\|^2$, lo que implica que $\|\vec{a}\| = \|\vec{b}\|$, comprobando que el paralelogramo tiene los cuatro lados iguales, es decir, es un rombo.

Supongamos ahora que el paralelogramo es un rombo, es decir, que se cumple que $\|\vec{a}\| = \|\vec{b}\|$; entonces, como

$$\overrightarrow{AC} \cdot \overrightarrow{DB} = (\vec{a} + \vec{b}) \cdot (\vec{a} - \vec{b}) = \|\vec{a}\|^2 - \|\vec{b}\|^2 = 0,$$

se deduce así que las diagonales son ortogonales.

4.3 Proyección ortogonal

Sean $\vec{a}$ y $\vec{b}$ dos vectores libres, $\vec{a} \neq \mathbf{0}$. Es posible descomponer el vector $\vec{b}$ como la suma de dos vectores, uno en la dirección del vector $\vec{a}$ y otro en dirección perpendicular a $\vec{a}$, como puede verse en la Figura 4.6.

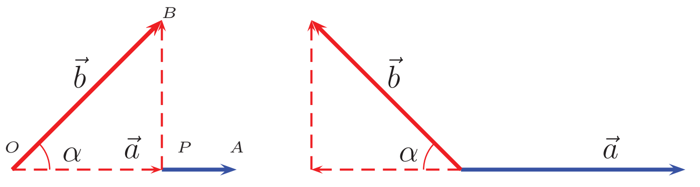


Figura 4.6. Proyección ortogonal

Para hacer esta descomposición, sea O el punto de concurrencia de los vectores $\vec{a}$ y $\vec{b}$, y sean A y B los extremos de dichos vectores. Por el punto B trazamos una perpendicular al vector $\vec{a}$. Si P es el pie de dicha perpendicular, al vector $\overrightarrow{OP}$ se le denomina la *proyección ortogonal* de $\vec{b}$ sobre $\vec{a}$ y se denota con $\overrightarrow{Pr}\left(\vec{b}/\vec{a}\right)$.

Para calcular esta proyección, notemos que ella es paralela al vector $\vec{a}$ y que, por lo tanto, se cumple la relación

$$\overrightarrow{Pr}\left(\vec{b}/\vec{a}\right) = \lambda\vec{a}.$$

La condición adicional nos permite escribir la igualdad

$$\lambda\vec{a} + \overrightarrow{PB} = \vec{b}.$$

Como $\overrightarrow{PB}$ es perpendicular al vector $\vec{a}$, $\overrightarrow{PB} \cdot \vec{a} = 0$, entonces

$$\lambda\vec{a} \cdot \vec{a} + \overrightarrow{PB} \cdot \vec{a} = \vec{b} \cdot \vec{a},$$

con lo cual, como $\vec{a} \neq \mathbf{0}$, se tiene que

$$\lambda = \frac{\vec{b} \cdot \vec{a}}{\vec{a} \cdot \vec{a}}.$$

En consecuencia,

$$\overrightarrow{Pr}\left(\vec{b}/\vec{a}\right) = \left(\frac{\vec{b} \cdot \vec{a}}{\vec{a} \cdot \vec{a}}\right)\vec{a}.$$

De modo similar, si $\vec{b} \neq \mathbf{0}$, entonces

$$\overrightarrow{Pr}\left(\vec{a}/\vec{b}\right) = \left(\frac{\vec{a}\cdot\vec{b}}{\vec{b}\cdot\vec{b}}\right)\vec{b}.$$

La ley distributiva del producto escalar garantiza que

$$\overrightarrow{Pr}\left((\vec{a}+\vec{c})/\vec{b}\right) = \overrightarrow{Pr}\left(\vec{a}/\vec{b}\right) + \overrightarrow{Pr}\left(\vec{c}/\vec{b}\right),$$

puesto que

$$\begin{aligned}\overrightarrow{Pr}\left((\vec{a}+\vec{c})/\vec{b}\right) &= \left(\frac{(\vec{a}+\vec{c})\cdot\vec{b}}{\vec{b}\cdot\vec{b}}\right)\vec{b}\\ &= \left(\frac{\vec{a}\cdot\vec{b}}{\vec{b}\cdot\vec{b}}\right)\vec{b} + \left(\frac{\vec{c}\cdot\vec{b}}{\vec{b}\cdot\vec{b}}\right)\vec{b}\\ &= \overrightarrow{Pr}\left(\vec{a}/\vec{b}\right) + \overrightarrow{Pr}\left(\vec{c}/\vec{b}\right).\end{aligned}$$

Uno de los resultados básicos del producto escalar es la desigualdad de Schwarz, la cual afirma que para cualquier producto escalar

$$\langle a, b\rangle \leq \|a\|\|b\|.$$

Si, en particular, se trata del producto escalar para los vectores libres, entonces

$$|\vec{a}\cdot\vec{b}| \leq \|\vec{a}\|\|\vec{b}\|,$$

con lo cual

$$-1 \leq \frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|\|\vec{b}\|} \leq 1,$$

lo que permite, junto con la definición de $\vec{a}\cdot\vec{b}$, calcular el ángulo entre dos vectores, siempre y cuando se tenga otro medio para computar $\vec{a}\cdot\vec{b}$. De esta forma,

$$\cos\alpha = \frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|\|\vec{b}\|}.$$

Una aplicación geométrica de este resultado nos lo proporciona el siguiente ejemplo.

Ejemplo 4.3.1. Demostrar que las diagonales de un rombo son bisectrices.
Solución
Sea $ABCD$ un rombo, en el cual se cumple que

$$\overrightarrow{AB} = \overrightarrow{DC} = \vec{a} \quad \text{y} \quad \overrightarrow{AD} = \overrightarrow{BC} = \vec{b}.$$

Si $\overrightarrow{AC} = \vec{c}$ es un vector según la diagonal, y α y β son los ángulos que la diagonal forma con los lados del rombo, como se ve en la Figura 4.7, y, como, además, por tratarse de un rombo, $\|\vec{a}\| = \|\vec{b}\|$, entonces, los cosenos para dichos ángulos se pueden calcular mediante las relaciones:

$$\cos\alpha = \frac{\vec{a}\cdot\vec{c}}{\|\vec{a}\|\|\vec{c}\|} = \frac{\vec{a}\cdot(\vec{a}+\vec{b})}{\|\vec{a}\|\|\vec{c}\|}.$$

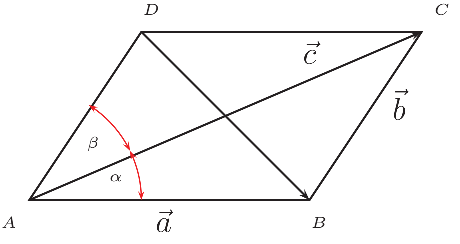


Figura 4.7. Diagonales de un rombo

En consecuencia,

$$\cos\alpha = \frac{\vec{a}\cdot\vec{a}}{\|\vec{a}\|\|\vec{c}\|} + \frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|\|\vec{c}\|} = \frac{\|\vec{a}\|^2}{\|\vec{a}\|\|\vec{c}\|} + \frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|\|\vec{c}\|}.$$

De manera similar, para el ángulo β se tiene

$$\cos\beta = \frac{\vec{b}\cdot\vec{c}}{\|\vec{b}\|\|\vec{c}\|} = \frac{\vec{b}\cdot(\vec{a}+\vec{b})}{\|\vec{b}\|\|\vec{c}\|},$$

con lo cual

$$\cos\beta = \frac{\vec{b}\cdot\vec{b}}{\|\vec{b}\|\|\vec{c}\|} + \frac{\vec{a}\cdot\vec{b}}{\|\vec{b}\|\|\vec{c}\|} = \frac{\|\vec{b}\|^2}{\|\vec{b}\|\|\vec{c}\|} + \frac{\vec{a}\cdot\vec{b}}{\|\vec{b}\|\|\vec{c}\|};$$

pero al ser $\|\vec{a}\| = \|\vec{b}\|$, entonces

$$\cos\beta = \frac{\|\vec{a}\|^2}{\|\vec{a}\|\|\vec{c}\|} + \frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|\|\vec{c}\|} = \cos\alpha$$

y, por lo tanto, los ángulos α y β son iguales, es decir, la diagonal es bisectriz.

Ejemplo 4.3.2. Demostrar que un ángulo inscrito en media circunferencia es recto.

Solución

Se trata de demostrar que el ángulo $\widehat{ABC}$, que se muestra en la Figura 4.8, es un ángulo recto.

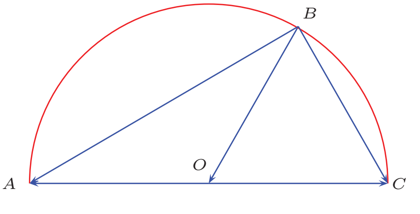


Figura 4.8. Ángulo inscrito en media circunferencia

Una de las maneras de probar esto es verificando que

$$\overrightarrow{BA}\cdot\overrightarrow{BC} = 0.$$

De la Figura 4.8 se tiene que

$$\overrightarrow{BA} \cdot \overrightarrow{BC} = \left(\overrightarrow{BO} + \overrightarrow{OA}\right) \cdot \left(\overrightarrow{BO} + \overrightarrow{OC}\right).$$

Si tenemos en cuenta que $\overrightarrow{OC} = -\overrightarrow{OA}$, entonces

$$\overrightarrow{BA} \cdot \overrightarrow{BC} = \left(\overrightarrow{BO} + \overrightarrow{OA}\right) \cdot \left(\overrightarrow{BO} - \overrightarrow{OA}\right).$$

Al aplicar la distributividad del producto escalar se obtiene:

$$\overrightarrow{BA} \cdot \overrightarrow{BC} = \overrightarrow{BO} \cdot \overrightarrow{BO} + \overrightarrow{OA} \cdot \overrightarrow{BO} - \overrightarrow{OA} \cdot \overrightarrow{BO} - \overrightarrow{OA} \cdot \overrightarrow{OA} = \|\overrightarrow{BO}\|^2 - \|\overrightarrow{OA}\|^2;$$

pero como

$$\|\overrightarrow{BO}\|^2 = \|\overrightarrow{OA}\|^2 = r^2,$$

concluimos que

$$\overrightarrow{BA} \cdot \overrightarrow{BC} = 0,$$

es decir, el ángulo entre ellos es un ángulo recto.

Ejemplo 4.3.3. Demostrar que las alturas de un triángulo se cortan en un punto. Este punto es conocido como el *ortocentro* del triángulo (véase la Figura 4.9).

Solución

Sea el triángulo ABC, con alturas AM y BN. Estas alturas se cortan en el punto O.

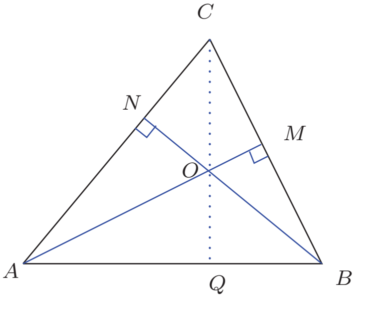


Figura 4.9. Ortocentro de un triángulo

Sea Q el punto de corte del lado AB con la recta determinada por los puntos C y O. Vamos a probar que esta recta es perpendicular al lado AB y que, por lo tanto, CQ es también una altura.

Como MA es la altura trazada desde el vértice A y OA es paralelo a MA, se cumple que

$$\vec{B}C \cdot \vec{O}A = \overrightarrow{OA} \cdot \left(\overrightarrow{BO} + \overrightarrow{OC}\right) = \overrightarrow{OA} \cdot \left(\overrightarrow{OC} - \overrightarrow{OB}\right) = 0.$$

En consecuencia,

$$\overrightarrow{OA} \cdot \overrightarrow{OC} = \overrightarrow{OA} \cdot \overrightarrow{OB}. \tag{4.3.1}$$

De igual manera, como NB es la altura trazada desde el vértice B y OB es paralelo a NB, se cumple que

$$\overrightarrow{AC} \cdot \overrightarrow{OB} = \overrightarrow{OB} \cdot \left(\overrightarrow{AO} + \overrightarrow{OC}\right) = \overrightarrow{OB} \cdot \left(\overrightarrow{OC} - \overrightarrow{OA}\right) = 0.$$

En consecuencia,

$$\overrightarrow{OB} \cdot \overrightarrow{OC} = \overrightarrow{OB} \cdot \overrightarrow{OA}. \tag{4.3.2}$$

De las ecuaciones (4.3.1) y (4.3.2) concluimos que

$$\overrightarrow{OB} \cdot \overrightarrow{OC} = \overrightarrow{OA} \cdot \overrightarrow{OC},$$

con lo cual

$$\overrightarrow{OC} \cdot \left(\overrightarrow{OB} - \overrightarrow{OA}\right) = \overrightarrow{OC} \cdot \left(\overrightarrow{OB} + \overrightarrow{AO}\right) = \overrightarrow{OC} \cdot \overrightarrow{AB} = 0.$$

Por lo tanto, OC es perpendicular a AB y, como resultado, QC es la altura trazada desde el vértice C.

4.4 Producto escalar respecto a una base

Sea $B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ una base ordenada para $\mathbb{E}^3$. Si $\vec{a}$ y $\vec{b}$ son dos vectores cualesquiera, con

$$\vec{a} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \alpha_3\vec{v}_3 \quad \text{y} \quad \vec{b} = \beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \beta_3\vec{v}_3,$$

la ley distributiva permite calcular el producto escalar en términos de los productos escalares de los vectores de la base. Esta expresión es:

$$\begin{aligned}
\vec{a}\cdot\vec{b} &= (\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \alpha_3\vec{v}_3)\cdot(\beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \beta_3\vec{v}_3)\\
&= \alpha_1\beta_1(\vec{v}_1\cdot\vec{v}_1) + \alpha_2\beta_2(\vec{v}_2\cdot\vec{v}_2) + \alpha_3\beta_3(\vec{v}_3\cdot\vec{v}_3)\\
&\quad + (\alpha_1\beta_2 + \alpha_2\beta_1)\,\vec{v}_1\cdot\vec{v}_2 + (\alpha_1\beta_3 + \alpha_3\beta_1)\,\vec{v}_1\cdot\vec{v}_3\\
&\quad + (\alpha_2\beta_3 + \alpha_3\beta_3)\,\vec{v}_2\cdot\vec{v}_3.
\end{aligned}$$

De nuevo, los cálculos se reducen si elegimos bases que tengan la posibilidad de simplificar los productos escalares. La siguiente definición va encaminada en esa dirección.

Definición 4.4.1. *Sea $B = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ una base de un espacio vectorial V; se dice que B es ortonormal si para todo $i, j = 1, \ldots, n$ se cumple que*

$$\vec{v}_i\cdot\vec{v}_j = \delta_{ij} = \begin{cases} 1, & i = j\\ 0, & i \neq j,\end{cases}$$

donde el símbolo δ_{ij} es conocido como el delta de Kronecker.

Si recordamos que

$$\vec{v}_i\cdot\vec{v}_i = \|v_i\|^2,$$

una base ortonormal es aquella en la cual los vectores tienen norma unitaria y son ortogonales dos a dos. Según esta característica, si $B = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es una base ortonormal, la ecuación para el cálculo del producto escalar se transforma en

$$\vec{a}\cdot\vec{b} = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3.$$

Notemos que esta expresión es igual a la que define el producto escalar entre dos vectores de $\mathbb{R}^3$. Esto nos muestra que el isomorfismo que establecimos entre los espacios $\mathbb{E}^3$ y $\mathbb{R}^3$ preserva la noción de *ortogonalidad* en ambos espacios, es decir, si $\vec{a}$ y $\vec{b}$ son dos vectores ortogonales en el espacio $\mathbb{R}^3$, con representaciones $(\alpha_1, \alpha_2, \alpha_3)$ y $(\beta_1, \beta_2, \beta_3)$ respecto a una base ordenada, entonces

$$\langle (\alpha_1, \alpha_2, \alpha_3), (\beta_1, \beta_2, \beta_3) \rangle = 0.$$

Otro aspecto importante que resulta del uso de las bases ortonormales es la posibilidad de calcular, de manera sencilla, la norma de un vector y la distancia entre dos vectores, ya que

$$\|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}} = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2},$$

mientras que

$$d\left(\vec{b}, \vec{a}\right) = \|\vec{b} - \vec{a}\| = \sqrt{(\beta_1 - \alpha_1)^2 + (\beta_2 - \alpha_2)^2 + (\beta_3 - \alpha_3)^2}.$$

De aquí en adelante, a menos que se diga lo contrario, se usará para $\mathbb{E}^3$ la base ordenada ortonormal $B = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, que también suele representarse como $C = \{\vec{i}, \vec{j}, \vec{k}\}$.

Ejemplo 4.4.1. Dados los puntos con coordenadas $A = (2, 1, 0)$, $B = (2, 0, 1)$ y $C = (3, -1, 2)$, hallar el perímetro del triángulo determinado por ellos.

Solución

Lo que se requiere en este problema es el cálculo de las longitudes de cada uno de los lados, así:

$$\begin{aligned}
d\left(\vec{A}, \vec{B}\right) &= \|\vec{A} - \vec{B}\| = \sqrt{(2-2)^2 + (1-0)^2 + (0-1)^2} = \sqrt{2} \\
d\left(\vec{C}, \vec{A}\right) &= \|\vec{C} - \vec{A}\| = \sqrt{(3-2)^2 + (-1-1)^2 + (2-0)^2} = 3 \\
d\left(\vec{C}, \vec{B}\right) &= \|\vec{C} - \vec{B}\| = \sqrt{(3-2)^2 + (-1-0)^2 + (2-1)^2} = \sqrt{3}
\end{aligned}$$

y, en consecuencia. el perímetro es $P = \sqrt{2} + 3 + \sqrt{3}$.

Ejemplo 4.4.2. Dados los puntos con coordenadas $A = (2, 1, 0)$, $B = (2, 0, 1)$ y $C = (3, -1, 2)$, hallar los ángulos del triángulo determinado por ellos.

Solución

Este problema es una de las más frecuentes aplicaciones del producto escalar. Consideremos los vectores

$$\begin{aligned} \overrightarrow{AB} &= \overrightarrow{B} - \overrightarrow{A} = -\vec{e}_2 + \vec{e}_3 \\ \overrightarrow{AC} &= \overrightarrow{C} - \overrightarrow{A} = \vec{e}_1 - 2\vec{e}_2 + 2\vec{e}_3 \\ \overrightarrow{BC} &= \overrightarrow{C} - \overrightarrow{B} = \vec{e}_1 - \vec{e}_2 + \vec{e}_3. \end{aligned}$$

Sea α el ángulo en el vértice A; entonces,

$$\cos\alpha = \frac{\overrightarrow{AB} \cdot \overrightarrow{AC}}{\left\|\overrightarrow{AB}\right\| \left\|\overrightarrow{AC}\right\|} = \frac{4}{3\sqrt{2}} \therefore \alpha = \arccos\left(\frac{4}{3\sqrt{2}}\right) = 19,47^\circ.$$

Sea β el ángulo en el vértice B; entonces,

$$\cos\beta = \frac{\overrightarrow{BA} \cdot \overrightarrow{BC}}{\left\|\overrightarrow{BA}\right\| \left\|\overrightarrow{BC}\right\|} = \frac{-2}{\sqrt{2}\sqrt{3}} \therefore \beta = \arccos\left(\frac{-2}{\sqrt{2}\sqrt{3}}\right) = 144,73^\circ.$$

Por último, sea γ el ángulo en el vértice C, entonces

$$\cos\gamma = \frac{\overrightarrow{CA} \cdot \overrightarrow{CB}}{\left\|\overrightarrow{CA}\right\| \left\|\overrightarrow{CB}\right\|} = \frac{5}{3\sqrt{3}} \therefore \gamma = \arccos\left(\frac{5}{3\sqrt{3}}\right) = 15,79^\circ.$$

Sin embargo, este último cálculo es innecesario ya que, en un triángulo cualesquiera, $\gamma = 180 - (\alpha + \beta)$.

Ejemplo 4.4.3. Hallar el ángulo que la diagonal del cubo forma con una de sus aristas.

Solución

Sea

$$\overrightarrow{OB} = l\vec{e}_1 + l\vec{e}_2 + l\vec{e}_3$$

un vector en la dirección de la diagonal y el vector $\overrightarrow{OA}$ tal que

$$\overrightarrow{OA} = l\vec{e}_2$$

es un vector en la dirección de uno de los lados; entonces, el ángulo entre ellos puede calcularse mediante la relación:

$$\cos\gamma = \frac{\overrightarrow{OB}\cdot\overrightarrow{OA}}{\left\|\overrightarrow{OB}\right\|\left\|\overrightarrow{OA}\right\|} = \frac{l^2}{l^2\sqrt{3}} \therefore \gamma = \arccos\left(\frac{1}{\sqrt{3}}\right) = 54,73^\circ.$$

Este ángulo es el formado por los vectores que van del origen a los puntos A y B de la Figura 4.10.

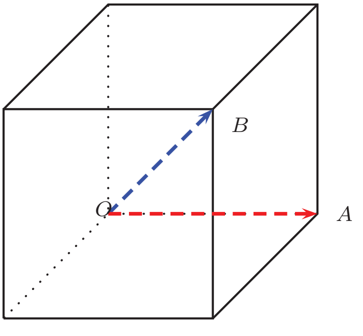


Figura 4.10. Ángulo de la diagonal de un cubo

Ejemplo 4.4.4. Descomponer el vector

$$\vec{b} = 2\vec{e}_1 + 3\vec{e}_2 - \vec{e}_3$$

como la suma de dos vectores, uno paralelo a

$$\vec{a} = \vec{e}_1 - \vec{e}_2 + \vec{e}_3$$

y otro perpendicular a $\vec{a}$.

Solución

Se trata de hallar la proyección ortogonal del vector $\vec{b}$ sobre el vector $\vec{a}$. De esta forma obtenemos

$$\overrightarrow{Pr}\left(\vec{b}/\vec{a}\right) = \left(\frac{\vec{b}\cdot\vec{a}}{\vec{a}\cdot\vec{a}}\right)\vec{a} = -\frac{2}{3}\left(\vec{e}_1 - \vec{e}_2 + \vec{e}_3\right).$$

Esta es la componente de $\vec{b}$ paralela a $\vec{a}$. Si denotamos con $\vec{c}$ a la componente perpendicular, entonces

$$\overrightarrow{Pr}\left(\vec{b}/\vec{a}\right) + \vec{c} = \vec{a},$$

con lo cual

$$\begin{aligned} \vec{c} &= \vec{b} - \overrightarrow{Pr}\left(\vec{b}/\vec{a}\right) = 2\vec{e}_1 + 3\vec{e}_2 - \vec{e}_3 - \left[-\frac{2}{3}\left(\vec{e}_1 - \vec{e}_2 + \vec{e}_3\right)\right] \\ &= \frac{1}{3}\left(8\vec{e}_1 + 7\vec{e}_2 - \vec{e}_3\right). \end{aligned}$$

Es claro que este vector y la proyección son ortogonales.

Ejemplo 4.4.5. Hallar los ángulos que el vector con representación

$$\vec{a} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \alpha_3\vec{e}_3$$

forma con cada uno de los ejes coordenados.

Solución

Si denotamos con α, β y γ los ángulos que el vector $\vec{a}$ forma con los vectores unitarios e_1, e_2 y e_3 respectivamente, entonces, los cosenos de dichos ángulos son:

$$\begin{aligned} \cos\alpha &= \frac{\vec{a}.e_1}{\|\vec{a}\|\|\vec{e}_1\|} = \frac{\alpha_1}{\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}} \\ \cos\beta &= \frac{\vec{a}.e_2}{\|\vec{a}\|\|\vec{e}_2\|} = \frac{\alpha_2}{\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}} \\ \cos\gamma &= \frac{\vec{a}.e_3}{\|\vec{a}\|\|\vec{e}_3\|} = \frac{\alpha_3}{\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}}, \end{aligned}$$

los cuales cumplen la ecuación

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1.$$

Estos números son conocidos como los *cosenos directores* del vector $\vec{a}$ (véase la Figura 4.11).

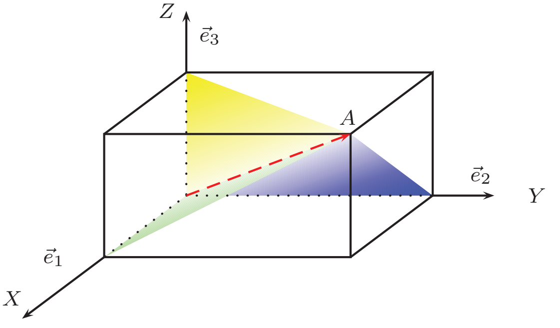


Figura 4.11. Cosenos directores de un vector

4.5 Lugares geométricos

En esta sección mostramos algunos lugares geométricos que pueden expresarse en términos de ecuaciones que involucran las coordenadas de los puntos respecto a una base determinada. Como veremos, el concepto fundamental que está ligado a las aplicaciones del producto escalar es el de *ortogonalidad*.

4.5.1 Rectas en el plano

Sea el punto $A = (\alpha_0, \beta_0)$ un punto sobre un plano Π y sea $\vec{n} = (\alpha, \beta)$ un vector sobre el mismo plano, como en la Figura 4.12.

Denotamos con $L(A, \vec{n})$ a la recta del plano que pasa por el punto A y es normal a $\vec{n}$.

Si X es un punto sobre la recta, tal que $\vec{X} = (x_1, x_2)$, entonces se cumple que

$$\overrightarrow{AX} \cdot \vec{n} = 0.$$

Con lo cual

$$L(A, \vec{n}) = \left\{ X / \overrightarrow{AX} \cdot \vec{n} = 0 \right\}.$$

Si usamos las coordenadas de los puntos respecto a la base dada,

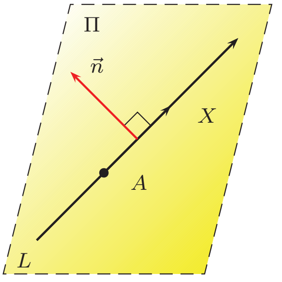


Figura 4.12. Ecuación de una recta en el plano

entonces la ecuación anterior se transforma en:

$$\alpha(x_1 - \alpha_0) + \beta(x_2 - \beta_0) = 0,$$

de donde

$$\alpha x_1 + \beta x_2 = \alpha\alpha_0 + \beta\beta_0.$$

Ejemplo 4.5.1. Hallar la ecuación de la recta normal al vector $\vec{n} = (1, -3)$ y que pasa por el punto $A = (2, 1)$.

Solución

Si $\vec{X} = (x_1, x_2)$, entonces, usando el resultado anterior, encontramos que

$$x_1 - 3x_2 = 2 - 3(1) = -1,$$

que es la ecuación pedida.

4.5.2 Ángulo entre dos rectas

En la sección previa mostramos que si $\vec{n} = (\alpha, \beta)$ es un vector normal a la recta, la ecuación de ésta toma la forma (en el plano)

$$\alpha x_1 + \beta x_2 = \delta,$$

donde $\delta = \alpha\alpha_0 + \beta\beta_0$.

Veamos recíprocamente que si tenemos la ecuación

$$\alpha x_1 + \beta x_2 = \delta,$$

entonces el vector $\vec{n} = (\alpha, \beta)$ es normal a la recta dada. Para esto, tomemos un vector $\overrightarrow{AB}$ determinado por dos puntos sobre la recta y calculemos el producto escalar con el vector $\vec{n} = (\alpha, \beta)$. Para establecer el punto A hacemos $x_1 = 0$, con lo cual

$$x_2 = \delta/\beta,$$

es decir,

$$A = (0, \delta/\beta).$$

Para el punto B hacemos $x_2 = 0$, con lo cual

$$B = (\delta/\alpha, 0);$$

por lo tanto,

$$\overrightarrow{AB} = \left(\frac{\delta}{\alpha}, -\frac{\delta}{\beta}\right)$$

y, en consecuencia,

$$\vec{n} \cdot \overrightarrow{AB} = \frac{\delta}{\alpha}\alpha - \frac{\delta}{\beta}\beta = 0.$$

Este resultado nos permite calcular en forma inmediata el vector normal a una recta dada.

Si ahora consideramos un par de rectas L_1 y L_2, en un mismo plano, el ángulo entre ellas es el mismo que el ángulo formado por los vectores normales a cada recta (véase la Figura 4.13).

Si denotamos con θ al ángulo entre las dos rectas, entonces, usando el producto escalar, tenemos que

$$\cos\theta = \frac{\vec{n}_1 \cdot \vec{n}_2}{\|\vec{n_1}\| \|\vec{n_2}\|}.$$

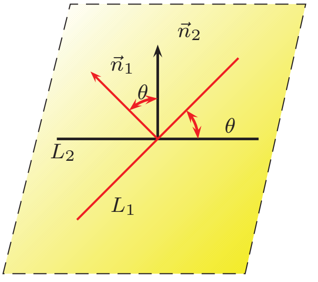


Figura 4.13. Ángulo entre dos rectas de un mismo plano

Finalmente, si las ecuaciones de las rectas son

$$\alpha_1 x_1 + \beta_1 x_2 = \delta_1 \quad \text{y} \quad \alpha_2 x_1 + \beta_2 x_2 = \delta_2,$$

los vectores normales a las rectas son

$$\vec{n}_1 = (\alpha_1, \beta_1) \text{ y } \vec{n}_2 = (\alpha_2, \beta_2),$$

con lo cual

$$\cos\theta = \frac{\alpha_1\alpha_2 + \beta_1\beta_2}{\sqrt{\alpha_1^2 + \beta_1^2}\sqrt{\alpha_2^2 + \beta_2^2}}.$$

Ejemplo 4.5.2. Hallar el ángulo determinado por las rectas con ecuaciones dadas por

$$3x_1 + 2x_2 = 5 \text{ y } 5x_1 - 2x_2 = 6.$$

Solución

En este caso, los vectores normales a cada recta son $\vec{n}_1 = (3, 2)$ y $\vec{n}_2 = (5, -2)$, con lo cual

$$\cos\theta = \frac{\vec{n}_1 \cdot \vec{n}_2}{\|\vec{n}_1\|\|\vec{n}_2\|} = \frac{11}{\sqrt{13}\sqrt{29}};$$

por lo tanto, $\theta = \arccos 0,567 = 55,49°$.

4.5.3 Planos

Sean $A = (\alpha_0, \beta_0, \gamma_0)$ un punto del espacio y un vector libre $\vec{n} = (\alpha, \beta, \gamma)$ no nulo; entonces, existe un único plano Π que pasa por A y es perpendicular al vector $\vec{n}$ (véase la Figura 4.14).

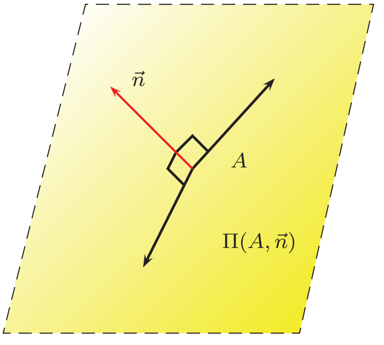


Figura 4.14. Plano normal al vector $\vec{n}$ y que pasa por A

Si denotamos este plano por $\Pi(A, \vec{n})$, es claro que un punto X con coordenadas $X = (x_1, x_2, x_3)$ está sobre el plano si

$$\overrightarrow{AX} \cdot \vec{n} = 0,$$

es decir,

$$\Pi(A, \vec{n}) = \left\{ (x_1, x_2, x_3) \,/\, \overrightarrow{AX} \cdot \vec{n} = 0 \right\}.$$

Si estamos utilizando un sistema de coordenadas ortogonal, entonces

$$\overrightarrow{AX} = (x_1 - \alpha_0, x_2 - \beta_0, x_3 - \gamma_0)$$

y, en consecuencia, la ecuación en coordenadas para el plano toma la forma:

$$\alpha(x_1 - \alpha_0) + \beta(x_2 - \beta_0) + \gamma(x_3 - \gamma_0) = 0,$$

con lo cual

$$\alpha x_1 + \beta x_2 + \gamma x_3 = \alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0.$$

Recíprocamente, si un par de puntos A y B cumplen la ecuación

$$\alpha x_1 + \beta x_2 + \gamma x_3 = \delta,$$

entonces el vector $\vec{n} = (\alpha, \beta, \gamma)$ es perpendicular al vector $\overrightarrow{AB}$. Para verificar esto, sean

$$\begin{aligned} A &= \left(x_1, x_2, \frac{\delta - \alpha x_1 - \beta x_2}{\gamma}\right) \\ B &= \left(x_1', x_2', \frac{\delta - \alpha x_1' - \beta x_2'}{\gamma}\right). \end{aligned}$$

Entonces,

$$\overrightarrow{AB} = \left(x_1' - x_1, x_2' - x_2, \frac{\alpha(x_1 - x_1') + \beta(x_2 - x_2')}{\gamma}\right)$$

y, en consecuencia,

$$\overrightarrow{AB} \cdot \vec{n} = (x_1' - x_1)\alpha + (x_2' - x_2)\beta + \frac{\alpha(x_1 - x_1') + \beta(x_2 - x_2')}{\gamma}\gamma = 0,$$

con lo cual se muestra que los dos vectores son ortogonales.

Ejemplo 4.5.3. Hallar la ecuación del plano normal al vector $\vec{n} = (3, 2, 1)$ y que pasa por el punto $A = (0, 1, 2)$.
Solución
Sea $X = (x_1, x_2, x_3)$; entonces, la ecuación en coordenadas para el plano pedido es:

$$3x_1 + 2x_2 + x_3 = 3(0) + 2(1) + 1(2) = 4.$$

Nótese que los exponentes de todas las variables son uno, lo cual corresponde a una ecuación lineal.

Cuando tenemos dos planos, es posible determinar el ángulo entre ellos. Para esto, hacemos uso de la siguiente definición.

Definición 4.5.1. *Sean* Π_1 *y* Π_2 *dos planos; el ángulo formado por ellos es el ángulo entre sus vectores normales.*

De este modo, si el primer plano tiene por ecuación

$$\alpha_1 x_1 + \beta_1 x_2 + \gamma_1 x_3 = \delta_1,$$

y el segundo plano

$$\alpha_2 x_1 + \beta_2 x_2 + \gamma_2 x_3 = \delta_2,$$

sus vectores normales serán

$$\vec{n}_1 = (\alpha_1, \beta_1, \gamma_1) \text{ y } \vec{n}_2 = (\alpha_2, \beta_2, \gamma_2).$$

Así, si θ es el ángulo formado por estos dos vectores, como puede verse en la Figura 4.15,

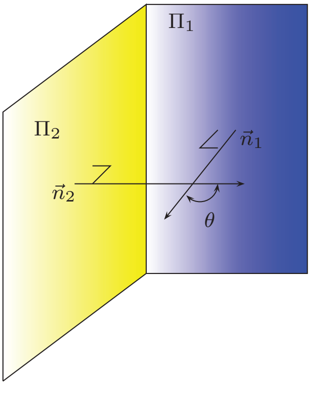


Figura 4.15. Ángulo entre dos planos

entonces, tenemos que:

$$\cos\theta = \frac{\vec{n}_1 \cdot \vec{n}_2}{\|\vec{n}_1\|\|\vec{n}_2\|} = \frac{\alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2}{\sqrt{\alpha_1^2 + \beta_1^2 + \gamma_1^2}\sqrt{\alpha_2^2 + \beta_2^2 + \gamma_2^2}}.$$

Ejemplo 4.5.4. Hallar el ángulo formado por los planos con ecuaciones

$$3x_1 + 2x_2 - x_3 = 4 \ \text{ y } x_1 - x_2 - x_3 = 6.$$

Solución

Los vectores normales a los planos son

$$\vec{n}_1 = (3, 2, -1) \ \text{ y } \vec{n}_2 = (1, -1, -1),$$

con lo cual

$$\|\vec{n}_1\| = \sqrt{3^2 + 2^2 + (-1)^2} = \sqrt{14}; \quad \|\vec{n}_2\| = \sqrt{1^2 + (-1)^2 + (-1)^2} = \sqrt{3}.$$

Así, si θ es el ángulo formado por estos dos vectores, entonces

$$\cos\theta = \frac{\vec{n}_1 \cdot \vec{n}_2}{\|\vec{n}_1\|\|\vec{n}_2\|} = \frac{2}{\sqrt{14}\sqrt{3}},$$

de modo que $\theta = 72,025°$.

Ejemplo 4.5.5. Hallar la intersección de los planos $-2x_1 + x_2 + x_3 = 4$ y $2x_1 + x_2 + x_3 = 6$.

Solución

En este problema, se trata de resolver el sistema de ecuaciones

$$\begin{aligned} -2x_1 + x_2 + x_3 &= 4 \\ 2x_1 + x_2 + x_3 &= 6. \end{aligned}$$

Al sumar estas dos ecuaciones y simplificar, se obtiene

$$x_2 + x_3 = 5,$$

lo cual nos dice que el sistema tiene infinitas soluciones.

Si hacemos $x_3 = \lambda$, entonces $x_2 = 5 - \lambda$ y, en consecuencia, al sustituir estos dos valores en la ecuación del segundo plano, se consigue que $x_1 = 1/2$. De esta manera, la solución puede escribirse en la forma

$$\begin{aligned} x_1 &= 1/2 \\ x_2 &= 5 - \lambda \\ x_3 &= \lambda, \end{aligned}$$

lo cual corresponde a la ecuación paramétrica de una recta.

4.5.4 Circunferencia y superficie esférica

Cuando se considera una ecuación de la forma

$$\|\vec{X} - \vec{C}\| = \rho,$$

el lugar geométrico correspondiente depende de que se estén utilizando vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$. En el primero de los casos, la ecuación representa los puntos de una circunferencia de radio ρ y centro en el punto C; en el segundo caso se trata de una superficie esférica, la cual es el lugar geométrico de los puntos del espacio que equidistan de un punto C llamado *centro*. La distancia del centro a la superficie se denomina *radio* y se representa con ρ. A la circunferencia con centro en C y radio ρ se denota con $C(C, \rho)$. De este modo, un punto X del plano pertenece a la circunferencia si y solo si

$$\|\vec{X} - \vec{C}\| = \rho$$

o, en forma equivalente,

$$C(C, \rho) = \left\{ X \in \mathbb{R}^2 / \|\vec{X} - \vec{C}\| = \rho \right\}.$$

Si se usan las coordenadas respecto a una base ortonormal, en la cual $C = (\alpha_1, \alpha_2)$ y $X = (x_1, x_2)$, la ecuación de la circunferencia es entonces

$$(x_1 - \alpha_1)^2 + (x_2 - \alpha_2)^2 = \rho^2.$$

Para representar el interior de la circunferencia, se tiene la expresión

$$(x_1 - \alpha_1)^2 + (x_2 - \alpha_2)^2 \leq \rho^2.$$

A la superficie esférica con centro en C y radio ρ se la denota con $S(C, \rho)$. Esta puede verse en la Figura 4.16.

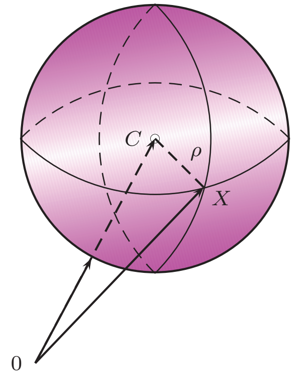


Figura 4.16. Superficie esférica de radio ρ y centro C

De esta forma, un punto X del espacio pertenece a la superficie esférica si y solo si

$$\|\vec{X} - \vec{C}\| = \rho$$

o, en forma equivalente,

$$S(C, \rho) = \left\{ X \in \mathbb{R}^3 / \|\vec{X} - \vec{C}\| = \rho \right\}.$$

Si se usan las coordenadas respecto a una base ortonormal, en la cual el centro tiene coordenadas $C = (\alpha_1, \alpha_2, \alpha_3)$ y $X = (x_1, x_2, x_3)$, y el radio está dado por ρ, la ecuación de la superficie esférica toma la forma

$$(x_1 - \alpha_1)^2 + (x_2 - \alpha_2)^2 + (x_3 - \alpha_3)^2 = \rho^2.$$

Notemos que si $\rho = 0$, el lugar geométrico se reduce al punto C.

Para representar toda la región interior a la superficie esférica, se usa la relación

$$(x_1 - \alpha_1)^2 + (x_2 - \alpha_2)^2 + (x_3 - \alpha_3)^2 \leqq \rho^2.$$

Ejemplo 4.5.6. Hallar la ecuación de las rectas tangentes a la circunferencia con centro en $C(1,1)$ y radio 2, trazadas desde el punto $P(5,4)$, como se muestra en la Figura 4.17.

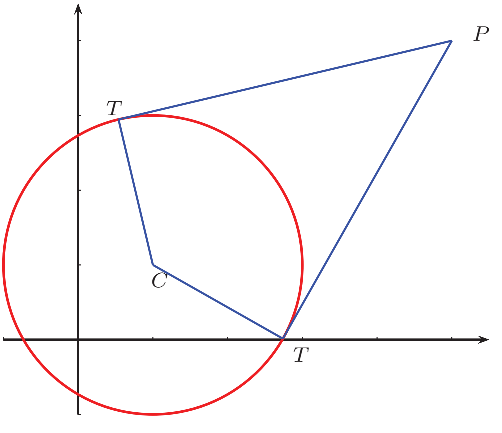


Figura 4.17. Tangentes a una circunferencia

Solución

Una de las formas de resolver este problema es buscando las coordenadas del punto de tangencia y luego usar la ecuación de la recta que pasa por dos puntos dados. La ecuación de la circunferencia del problema es

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

Sea ahora $T(x,y)$ el punto de tangencia; como este punto está sobre la circunferencia, cumple que

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

Consideremos ahora los vectores

$$\overrightarrow{TC} = (1-x, 1-y) \text{ y } \overrightarrow{TP} = (5-x, 4-y).$$

Estos vectores tienen direcciones según el radio y según la recta tangente; en consecuencia, ellos son perpendiculares, lo cual equivale a la ecuación

$$\overrightarrow{TP} \cdot \overrightarrow{TC} = (1-x)(5-x) + (1-y)(4-y) = 0.$$

Esta última ecuación puede escribirse como

$$x^2 - 6x + y^2 - 5y + 9 = 0. \tag{4.5.1}$$

De la ecuación de la circunferencia, se tiene que

$$y = \sqrt{4-(x-1)^2} + 1 = \sqrt{3+2x-x^2} + 1,$$

con lo cual, al elevar al cuadrado,

$$y^2 = 4 + 2x - x^2 + 2\sqrt{3+2x-x^2}.$$

Al sustituir estos valores en 4.5.1, se obtiene

$$x^2 - 6x + (4 + 2x - x^2 + 2\sqrt{3+2x-x^2}) - 5(\sqrt{3+2x-x^2} + 1) + 9 = 0,$$

que es equivalente a

$$8 - 4x = 3\sqrt{3+2x-x^2}.$$

Elevando al cuadrado en ambos lados se consigue

$$64 - 64x + 16x^2 = 9(3 + 2x - x^2),$$

que se transforma en la ecuación cuadrática

$$25x^2 - 82x + 37 = 0,$$

cuyas soluciones son

$$x_1 = \frac{41 + 6\sqrt{21}}{25},$$

$$x_2 = \frac{41 - 6\sqrt{21}}{25}.$$

Al sustituir estos valores en la ecuación para y y simplificar, se obtienen los siguientes valores:

$$y_1 = \frac{37 - 8\sqrt{21}}{25},$$

$$y_2 = \frac{37 + 8\sqrt{21}}{25}.$$

Lo anterior nos muestra que desde el punto exterior P se pueden trazar dos rectas tangentes. La ecuación de una de ellas es:

$$\overrightarrow{PX} = \lambda \overrightarrow{PT} \quad \therefore \quad \vec{X} = \vec{P} + \lambda \overrightarrow{PT}.$$

Usando las coordenadas de los puntos se tiene

$$(x, y) = (5, 4) + \lambda \left(\frac{41 + 6\sqrt{21}}{25} - 5, \frac{37 - 8\sqrt{21}}{25} - 4 \right),$$

con lo cual las coordenadas paramétricas de la recta son

$$\begin{aligned} x &= 5 + \lambda \left(\frac{-84 + 6\sqrt{21}}{25} \right) \\ y &= 4 + \lambda \left(\frac{-63 - 8\sqrt{21}}{25} \right). \end{aligned}$$

En el caso de la otra tangente, la ecuación se obtiene de manera semejante.

4.5.5 Intersección entre circunferencia y recta

Sea P un punto exterior a la circunferencia $\|\vec{X} - \vec{C}\| = \rho$ y sea $\vec{u}$ un vector unitario. Consideremos el problema de determinar el punto de intersección de la circunferencia dada con la recta L que pasa por el punto P y es paralela al vector $\vec{u}$, como se muestra en la Figura 4.18.

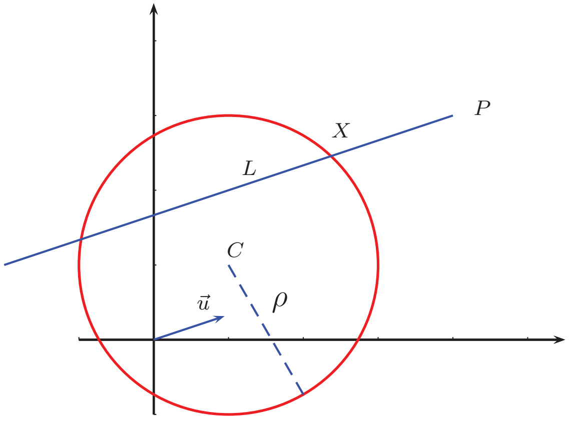


Figura 4.18. Intersección entre una circunferencia y una recta

Un punto X pertenece a la recta y a la circunferencia si satisface simultáneamente las ecuaciones

$$\|\vec{X} - \vec{C}\| = \rho \qquad \text{y} \qquad \vec{X} = \vec{P} + \lambda\vec{u}.$$

A partir de ello conseguimos la relación

$$\|\vec{P} + \lambda\vec{u} - \vec{C}\|^2 = \|\lambda\vec{u} - \overrightarrow{PC}\|^2 = \rho^2,$$

donde $\overrightarrow{PC}$ es el vector que une el centro de la circunferencia con el punto P.

Si utilizamos la definición de norma de un vector (definición 4.2.3), la relación anterior puede escribirse como

$$\|\lambda\vec{u} - \overrightarrow{PC}\|^2 = \left(\lambda\vec{u} - \overrightarrow{PC}\right) \cdot \left(\lambda\vec{u} - \overrightarrow{PC}\right) = \rho^2.$$

Al desarrollar el producto escalar se tiene que

$$\left(\lambda\vec{u} - \overrightarrow{PC}\right) \cdot \left(\lambda\vec{u} - \overrightarrow{PC}\right) = \lambda^2\vec{u} \cdot \vec{u} - 2\lambda\vec{u} \cdot \overrightarrow{PC} + \overrightarrow{PC} \cdot \overrightarrow{PC} = \rho^2.$$

Teniendo en cuenta que $\vec{u}$ es un vector unitario, la ecuación anterior se transforma en

$$\lambda^2 - 2\lambda\vec{u} \cdot \overrightarrow{PC} + \left\|\overrightarrow{PC}\right\|^2 - \rho^2 = 0.$$

Esta es una ecuación cuadrática en λ y su solución es

$$\lambda = \frac{2\vec{u} \cdot \overrightarrow{PC} \pm \sqrt{\left(2\vec{u} \cdot \overrightarrow{PC}\right)^2 - 4\left(\left\|\overrightarrow{PC}\right\|^2 - \rho^2\right)}}{2},$$

que es lo mismo que

$$\lambda = \vec{u} \cdot \overrightarrow{PC} \pm \sqrt{\left(\vec{u} \cdot \overrightarrow{PC}\right)^2 + \rho^2 - \left\|\overrightarrow{PC}\right\|^2}.$$

La solución de esta ecuación depende de los signos del término interior del radical; sea δ definido por

$$\delta = \left(\vec{u} \cdot \overrightarrow{PC}\right)^2 + \rho^2 - \left\|\overrightarrow{PC}\right\|^2;$$

entonces, al estudiar el valor de δ, podemos clasificar el problema de la intersección en tres casos, a saber:

- Si $\delta < 0$, la recta y la circunferencia no se cortan.
- Si $\delta = 0$, la recta y la circunferencia se cortan en un solo punto, el cual corresponde al punto de tangencia. Además, para ese punto sus coordenadas son

$$\vec{X} = \vec{P} + \left(\vec{u} \cdot \overrightarrow{PC}\right)\vec{u}.$$

- Si $\delta > 0$, la recta y la circunferencia se cortan en dos puntos, cuyas coordenadas son

$$\vec{X} = \vec{P} + \left[\vec{u} \cdot \overrightarrow{PC} \pm \sqrt{\left(\vec{u} \cdot \overrightarrow{PC}\right)^2 + \rho^2 - \|\overrightarrow{PC}\|^2}\right] \vec{u}.$$

Ejemplo 4.5.7. Sea una circunferencia con centro en el punto $C(1,1)$ y radio 2. Hallar su intersección con la recta que pasa por el punto $P(4,4)$ y es paralela al vector $\vec{u} = 1/5(3,4)$.

Solución

Para este problema, el vector $\overrightarrow{PC}$, que une el centro de la circunferencia con el punto P, y el vector unitario $\vec{u}$ son:

$$\overrightarrow{PC} = (-3,-3), \qquad \vec{u} = \frac{1}{5}(3,4).$$

Por lo tanto,

$$\overrightarrow{PC} \cdot \vec{u} = -\frac{21}{5}$$

y, en consecuencia,

$$\delta = \left(\vec{u} \cdot \overrightarrow{PC}\right)^2 + \rho^2 - \|\overrightarrow{PC}\|^2 = \left(\frac{-21}{5}\right)^2 + 4 - 18 = \frac{91}{25} > 0,$$

lo cual muestra que la circunferencia y la recta se cortan en dos puntos. Los valores del parámetro λ son

$$\lambda = -\frac{21}{5} \pm \frac{\sqrt{91}}{5},$$

con lo cual, el primer punto de corte es

$$\vec{X}_1 = \vec{P} + \lambda\vec{u} = (4,4) + \left(\frac{\sqrt{91} - 21}{25}\right)(3,5) \approx (2{,}625, 2{,}166),$$

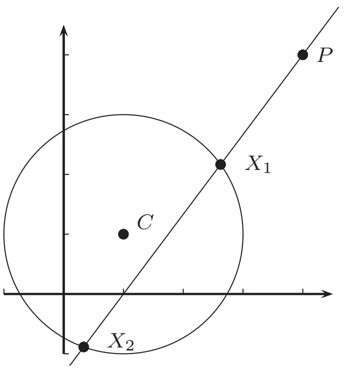


Figura 4.19. Intersección de una recta y una circunferencia

mientras que el segundo punto de corte es

$$\vec{X}_2 = \vec{P} + \lambda\vec{u} = (4,4) + \left(\frac{-\sqrt{91}-21}{25}\right)(3,5) \approx (0,335,-0,886).$$

Estos son los puntos que se muestran en la Figura 4.19.

4.5.6 Superficie cónica

Consideremos dos semirrectas L_1 y l_2 que parten de un punto P y que forman un ángulo α. Si una de ellas rota alrededor de la otra, se describe una superficie llamada *cono circular* o *cono de revolución* (véase la Figura 4.20).

A la recta que permanece fija se le denomina *eje del cono*, mientras que la que lleva a cabo la rotación es conocida como la *generatriz del cono*. Si denotamos con $L(P,\vec{g})$ a la semirrecta con origen en P y que es paralela al vector $\vec{g}$ y con $L(P,\vec{e})$ a la semirrecta que parte del punto P y es paralela al vector $\vec{e}$, entonces un punto X está sobre la superficie cónica si y solo si el ángulo entre los vectores $\overrightarrow{PX}$ y $\vec{e}$ es el mismo que el ángulo entre los vectores $\vec{g}$ y $\vec{e}$ (véase la Figura 4.20).

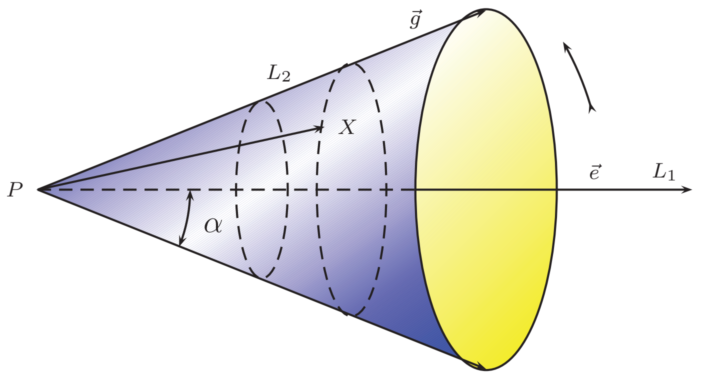


Figura 4.20. Superficie cónica

Esto significa que:

$$\cos\alpha = \frac{\vec{g}\cdot\vec{e}}{\|\vec{g}\|\|\vec{e}\|} = \frac{\overrightarrow{PX}\cdot\vec{e}}{\|\overrightarrow{PX}\|\|\vec{e}\|}.$$

Así, si $C(P,\vec{g},\vec{e})$ representa el cono con vértice en P, generatriz $\vec{g}$ y eje $\vec{e}$, su ecuación vectorial es

$$C(P,\vec{g},\vec{e}) = \left\{ X \Big/ \frac{\overrightarrow{PX}\cdot\vec{e}}{\|\overrightarrow{PX}\|\|\vec{e}\|} = \frac{\vec{g}\cdot\vec{e}}{\|\vec{g}\|\|\vec{e}\|} \right\}.$$

Ejemplo 4.5.8. Hallar la ecuación de la superficie cónica con vértice en el origen, eje $\vec{e} = (0,0,2)$ y generatriz $\vec{g} = (1,2,1)$.

Solución

Sea $X = (x,y,z)$; la ecuación vectorial del cono es

$$\frac{\overrightarrow{PX}\cdot\vec{e}}{\|\overrightarrow{PX}\|\|\vec{e}\|} = \frac{\vec{g}\cdot\vec{e}}{\|\vec{g}\|\|\vec{e}\|},$$

la cual puede simplificarse a la expresión

$$\overrightarrow{PX}\cdot\vec{e} = \frac{\|\overrightarrow{PX}\|}{\|\vec{g}\|}\,\vec{g}\cdot\vec{e}.$$

Como $\overrightarrow{PX} = (x,y,z)$, la ecuación anterior se transforma en

$$2z = \frac{2}{\sqrt{6}}\sqrt{x^2+y^2+z^2};$$

esta, después de simplificarla, se transforma en la expresión

$$z = \frac{1}{\sqrt{5}}\sqrt{x^2 + y^2},$$

que corresponde a un cono con eje sobre el eje z y vértice en el origen.

Ejemplo 4.5.9. Hallar la ecuación de la superficie cónica con vértice en $P = (2, 1, 1)$, eje $\vec{e} = (3, 1, 1)$ y generatriz $\vec{g} = (2, -1, -1)$.
Solución
Sea $X = (x_1, x_2, x_3)$ un punto de la superficie cónica. La ecuación vectorial del cono es

$$\frac{\overrightarrow{PX} \cdot \vec{e}}{\|\overrightarrow{PX}\|\|\vec{e}\|} = \frac{\vec{g} \cdot \vec{e}}{\|\vec{g}\|\|\vec{e}\|},$$

la cual, después de simplificar, se transforma en la expresión

$$\overrightarrow{PX} \cdot \vec{e} = \frac{\|\overrightarrow{PX}\|}{\|\vec{g}\|}\, \vec{g} \cdot \vec{e}.$$

Como el vector $\overrightarrow{PX} = (x_1 - 2, x_2 - 1, x_3 - 1)$, la ecuación anterior se transforma en

$$3(x_1-2)+(x_2-1)+(x_3-1) = \frac{4}{\sqrt{6}}\sqrt{(x_1 - 2)^2 + (x_2 - 1)^2 + (x_3 - 1)^2},$$

que corresponde a una ecuación cuadrática.

4.6 Cálculo de distancias

El concepto de *distancia* está íntimamente ligado a la noción de *ortogonalidad* y, en este caso, el producto escalar proporciona un método práctico para la solución de este tipo de problemas. A continuación presentamos algunas aplicaciones que conducen a fórmulas de uso común en el cálculo y en la geometría analítica.

4.6.1 Distancia de un punto a una recta

La formulación de este problema está restringida inicialmente a la determinación de la distancia de un punto $P = (\alpha_0, \beta_0)$ a la recta que tiene por ecuación $\alpha x_1 + \beta x_2 = \gamma$, teniendo en cuenta que la recta y el punto están contenidos en el mismo plano y que las coordenadas están referidas respecto a una base ortonormal. Estas condiciones se muestran en la Figura 4.21.

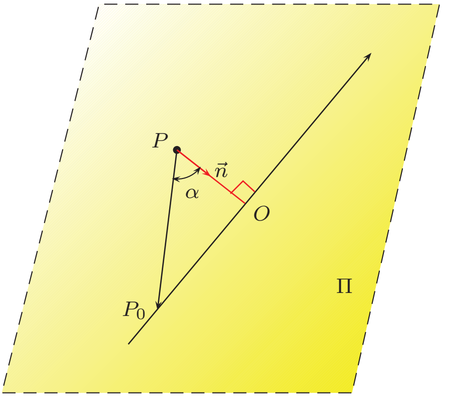


Figura 4.21. Distancia de un punto a una recta

Si $d = \|\overrightarrow{PO}\|$ representa la distancia del punto a la recta, entonces, si P_0 es un punto sobre la recta y α es el ángulo entre el vector $\overrightarrow{PP_0}$ y el vector normal a la recta, obtenemos que

$$d = \|\overrightarrow{PP_0}\| |\cos\alpha| = \|\overrightarrow{PP_0}\| \frac{|\vec{n} \cdot \overrightarrow{PP_0}|}{\|\vec{n}\| \|\overrightarrow{PP_0}\|},$$

con lo cual

$$d = \frac{|\vec{n} \cdot \overrightarrow{PP_0}|}{\|\vec{n}\|}.$$

Además, el vector normal a la recta está dado por $\vec{n} = (\alpha, \beta)$ y, en consecuencia,

$$\vec{n} \cdot \overrightarrow{PP_0} = \vec{n} \cdot \vec{P_0} - \vec{n} \cdot \vec{P} = \alpha\alpha_0 + \beta\beta_0 - (\alpha x_1 + \beta x_2);$$

pero como $(\alpha x_1 + \beta x_2) = \gamma$, obtenemos que $\vec{n} \cdot \overrightarrow{PP_0} = \alpha\alpha_0 + \beta\beta_0 - \gamma$, lo cual nos permite escribir la fórmula de distancia de un punto a una recta como

$$d = \frac{|\alpha\alpha_0 + \beta\beta_0 - \gamma|}{\sqrt{\alpha^2 + \beta^2}}.$$

Ejemplo 4.6.1. Hallar la distancia del punto $P = (1, 2)$ a la recta con ecuación $3x_1 - 2x_2 = 9$.
Solución

El vector normal a la recta es $\vec{n} = (3, -2)$, con lo que

$$d = \frac{|3 \times 1 - 2 \times 2 - 9|}{\sqrt{3^2 + (-2)^2}} = \frac{10}{\sqrt{13}}.$$

Ejemplo 4.6.2. Hallar la distancia del punto $P(3, 1)$ a la recta $2x_1 - x_2 = 7$.
Solución

Si usamos la ecuación para la distancia de un punto a una recta, conseguimos que

$$d = \frac{|(\alpha\alpha_0 + \beta\beta_0) - \gamma|}{\sqrt{\alpha^2 + \beta^2}} = \frac{|2 \times 3 - 1 \times 1 - 7|}{\sqrt{2^2 + (-1)^2}} = \frac{2}{\sqrt{5}}.$$

Esta longitud corresponde a la magnitud del vector perpendicular a la recta trazado desde esta hasta el punto P.

Ejemplo 4.6.3. Hallar el área del triángulo con vértices en $A(1, 1)$, $B(3, -1)$ y $C(2, 4)$.
Solución

Una de las formas de resolver este problema es calcular la longitud de uno de los lados y la de la altura trazada desde el vértice opuesto a este lado. En este caso, si tomamos como base el lado AB, la altura será la distancia del punto C a la recta que pasa por A y B (véase la Figura 4.22).

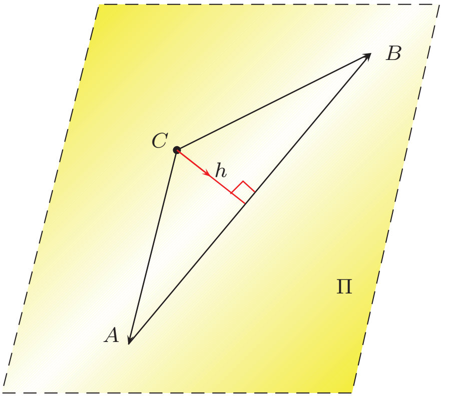


Figura 4.22. Área de un triángulo cualquiera

La ecuación de la recta que pasa por A y B es

$$\vec{X} = \vec{A} + \lambda(\vec{B} - \vec{A}),$$

la cual, expresada en coordenadas, corresponde a

$$(x_1, x_2) = (1, 1) + \lambda(2, -2).$$

Al eliminar el parámetro λ se obtiene la ecuación

$$\frac{x_1 - 1}{2} = \frac{x_2 - 1}{-2}.$$

La ecuación anterior puede simplificarse y escribirse finalmente como

$$x_1 + x_2 = 2,$$

a partir de la cual concluimos que

$$h = \frac{|1 \times 2 + 1 \times 4 - 2|}{\sqrt{1^2 + 1^2}} = \frac{4}{\sqrt{2}},$$

mientras que la base es $\|\overrightarrow{AB}\| = 2\sqrt{2}$. Así, el área A del triángulo es

$$A = \frac{\|\overrightarrow{AB}\| \times h}{2} = \frac{1}{2}(2\sqrt{2})\left(\frac{4}{\sqrt{2}}\right) = 4.$$

Puede verificarse que si se cambia la base y se toma la correspondiente altura, el área es la misma.

4.6.2 Distancia de un punto a un plano

Sean P_0 un punto de coordenadas $(\alpha_0, \beta_0, \gamma_0)$ respecto a una base ortonormal B y un plano Π, con ecuación $\alpha x_1 + \beta x_2 + \gamma x_3 = \delta$ (véase la Figura 4.23).

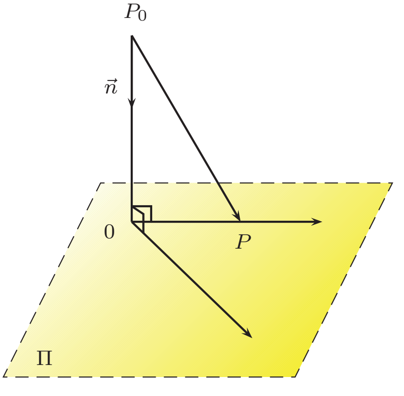


Figura 4.23. Distancia del punto P_0 al plano Π

La distancia del punto P_0 al plano Π es la longitud de la perpendicular al plano trazada desde el punto dado. De esta forma, si $P = (x_1, x_2, x_3)$ es un punto sobre el plano y $\vec{n}$ es el vector normal al plano, entonces

$$d = \|\overrightarrow{P0}\| = \|\overrightarrow{P_0P}\| |\cos \alpha|,$$

donde α es el ángulo entre el vector normal y el vector $\overrightarrow{P_0P}$, con lo cual obtenemos que

$$d = \|\overrightarrow{P0}\| = \|\overrightarrow{P_0P}\| \left| \frac{\overrightarrow{P_0P} \cdot \vec{n}}{\|\overrightarrow{P_0P}\| \, \|\vec{n}\|} \right|,$$

que finalmente se simplifica como

$$d = \left| \frac{\overrightarrow{P_0P} \cdot \vec{n}}{\|\vec{n}\|} \right|.$$

Esta expresión, en términos de las coordenadas de los puntos y del vector normal al plano, puede escribirse en la forma

$$d = \left| \frac{\alpha(x_1 - \alpha_0) + \beta(x_2 - \beta_0) + \gamma(x_3 - \gamma_0)}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}} \right|;$$

pero al ser

$$\alpha x_1 + \beta x_2 + \gamma x_3 = \delta,$$

obtenemos la relación conocida como *fórmula de Lagrange.*

$$d = \left| \frac{\delta - (\alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0)}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}} \right|.$$

Ejemplo 4.6.4. Hallar la distancia del punto $P = (1, 2, 3)$ al plano con ecuación $3x_1 + x_2 - x_3 = 7$.

Solución

Al aplicar la ecuación anterior, obtenemos:

$$d = \left| \frac{\delta - (\alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0)}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}} \right| = \left| \frac{7 - (3 \times 1 + 1 \times 2 - 1 \times 3)}{\sqrt{3^2 + 1^2 + (-1)^2}} \right| = \frac{5}{\sqrt{11}}.$$

Es claro que si el punto está sobre el plano, su distancia es nula.

Ejemplo 4.6.5 (Distancia entre dos planos). En este ejemplo consideramos el problema de determinar la distancia entre dos planos paralelos, en cuyo caso sus ecuaciones tendrán la forma

$$\alpha x_1 + \beta x_2 + \gamma x_3 = \delta_1 \text{ y } \alpha x_1 + \beta x_2 + \gamma x_3 = \delta_2.$$

Para la solución de este problema, basta considerar un punto P_0 de coordenadas $(\alpha_0, \beta_0, \gamma_0)$ sobre el primer plano y reducir el problema al cálculo de la distancia de un punto a un plano, con lo cual obtenemos entonces la relación

$$d = \left| \frac{\delta_2 - (\alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0)}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}} \right|;$$

pero como el punto P_0 está sobre el primer plano,

$$\alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0 = \delta_1,$$

obteniéndose finalmente la ecuación

$$d = \left| \frac{\delta_2 - \delta_1}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}} \right|.$$

Dado el caso de que se trate del mismo plano, $\delta_1 = \delta_2$ y, obviamente, la distancia es cero.

Ejemplo 4.6.6 (Distancia entre rectas). Consideremos las rectas L_1 y L_2 con ecuaciones paramétricas dadas por

$$\begin{aligned} x_1 &= 3t + 1 \\ x_2 &= t - 1 \\ x_3 &= t + 1 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x_1 &= 2s + 3 \\ x_2 &= -s - 3 \\ x_3 &= s + 1. \end{aligned}$$

Estas rectas no se cortan, ya que si fuese de este modo, el sistema de ecuaciones

$$\begin{aligned} 3t + 1 &= 2s + 3 \\ t - 1 &= -s - 3 \\ t + 1 &= s + 1 \end{aligned}$$

tendría al menos una solución; pero de la última de las ecuaciones concluimos que $t = s$, y al sustituir este valor en la primera de la ecuaciones se obtiene que $t = 2$, mientras que si se sustituye en la segunda ecuación, el resultado es $2t = -2$, con lo cual $t = -1$, diferente del valor previamente hallado.

En este caso, la distancia entre las dos rectas corresponde al mínimo de todas las longitudes de los vectores $\overrightarrow{PQ}$, donde el punto P está en la recta L_1 y el punto Q está sobre la recta L_2, es decir:

$$d = \text{mín}\left\{\|\overrightarrow{PQ}\| \,/ P \in L_1; Q \in L_2\right\}.$$

Según lo anterior, lo que se pretende es minimizar la función

$$L(P, Q) = (x_1 - x_1')^2 + (x_2 - x_2')^2 + (x_3 - x_3')^2,$$

donde (x_1, x_2, x_3) son las coordenadas de un punto sobre L_1, mientras que (x_1', x_2', x_3') son las coordenadas de un punto sobre L_2.

Usando las ecuaciones paramétricas de cada una de las rectas, debemos entonces minimizar la función

$$L(t, s) = [(3t+1)-(2s+3)]^2+[(t-1)-(-s-3)]^2+[(t+1)-(s+1)]^2,$$

expresión que al simplificarla corresponde a

$$L(t, s) = (3t - 2s - 2)^2 + (t + s + 2)^2 + (t - s)^2.$$

Aquí es importante resaltar el hecho de que lo que se pretende minimizar es la expresión que corresponde a la distancia entre puntos, es decir,

$$d = \sqrt{L(t, s)};$$

pero como

$$\text{mín}\{\sqrt{L(s, t)}\} = \sqrt{\text{mín}\{L(s, t)\}},$$

es suficiente que calculemos el valor

$$\text{mín}\{L(s, t)\}.$$

Para minimizar esta expresión, se calculan las derivadas parciales respecto a cada variable, se igualan a cero y se resuelve el sistema simultáneo de ecuaciones así obtenido.

Las soluciones de este sistema corresponden a los valores extremos de la función. En el ejemplo, estas ecuaciones son:

$$\begin{aligned}\frac{\partial L}{\partial t} &= 6(3t-2s-2)+2(t+s+2)+2(t-s)\\ \frac{\partial L}{\partial s} &= -4(3t-2s-2)+2(t+s+2)-2(t-s).\end{aligned}$$

Al simplificar estas ecuaciones e igualar cada una a cero, conseguimos el sistema de ecuaciones

$$\begin{aligned}22t-12s-8 &= 0\\ -12t+12s+12 &= 0\end{aligned}$$

o, en forma equivalente,

$$\begin{aligned}11t-6s &= 4\\ t-s &= 1,\end{aligned}$$

cuyas soluciones son

$$t=\frac{\begin{vmatrix}4 & -6\\ 1 & -1\end{vmatrix}}{\begin{vmatrix}11 & -6\\ 1 & -1\end{vmatrix}}=-\frac{2}{5}; \qquad s=\frac{\begin{vmatrix}11 & 4\\ 1 & 1\end{vmatrix}}{\begin{vmatrix}11 & -6\\ 1 & -1\end{vmatrix}}=-\frac{7}{5}.$$

Para determinar si estos valores extremos corresponden a un mínimo, determinamos el signo del hessiano de la función en este punto, es decir, calculamos el signo de

$$H(s,t)=L_{ss}L_{tt}-L_{st}^2.$$

Como en este problema $L_{ss}=12$, $L_{tt}=22$ y $L_{st}=-12$, entonces

$$H(s,t)=12\times 22-(-12)^2=120>0,$$

con lo cual se muestra que para los valores de s y t hallados la función es mínima.

Este valor mínimo es

$$L\left(-\frac{2}{5},-\frac{7}{5}\right)=\left(-\frac{6}{5}+\frac{14}{5}-2\right)^2+\left(-\frac{2}{5}-\frac{7}{5}+2\right)^2+\left(-\frac{2}{5}+\frac{7}{5}\right)^2,$$

con lo que

$$L\left(-\frac{2}{5},-\frac{7}{5}\right)=\frac{30}{25}=\frac{6}{5}.$$

En consecuencia, la distancia mínima corresponde a $\sqrt{\frac{6}{5}}$.

Otra forma de resolver este problema es considerando las propiedades geométricas que determinan la distancia mínima entre las rectas. En este caso, si P está sobre L_1 y Q está sobre L_2 y si $\vec{v}_1$ y $\vec{v}_2$ son vectores paralelos a L_1 y L_2 respectivamente, entonces la distancia mínima entre P y Q se presenta cuando $\overrightarrow{PQ}$ es perpendicular a los vectores $\vec{v}_1$ y $\vec{v}_2$, lo cual es equivalente al par de condiciones:

$$\overrightarrow{PQ}\cdot\vec{v}_1=0 \quad \text{y} \quad \overrightarrow{PQ}\cdot\vec{v}_2=0.$$

Como

$$\overrightarrow{PQ}=(3t-2s-2,t+s+2,t-s) \;\text{ y } \vec{v}_1=(3,1,1),\; \vec{v}_2=(2,-1,1),$$

los dos productos escalares anteriores se transforman en

$$9t-6s-6+t+s+2+t-s=0 \quad \text{y} \quad 6t-4s-4-t-s-2+t-s=0.$$

Estas dos expresiones, luego de simplificarlas, se reducen al sistema

$$\begin{aligned} 11t-6s &= 4 \\ t-s &= 1, \end{aligned}$$

que es exactamente el mismo sistema resuelto anteriormente.

Usando los valores de las soluciones para s y t, concluimos que el vector $\overrightarrow{PQ}$ tiene por coordenadas

$$\overrightarrow{PQ} = \left(-\frac{2}{5}, \frac{1}{5}, 1\right),$$

por lo que su magnitud es $\sqrt{\frac{6}{5}}$, como ya se había establecido.

4.7 Bases ortonormales

Hasta el momento se ha visto que algunas operaciones entre vectores se llevan a cabo de manera más sencilla si las bases son ortonormales; tal es el caso del producto escalar y la longitud de un vector. En esta sección se presenta un método para construir bases ortonormales para un espacio vectorial, a partir de una base conocida para dicho espacio; además, se enuncian algunas otras características notables de las bases ortonormales.

Definición 4.7.1. *Un conjunto de vectores* $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ *es ortonormal si:*

- $\vec{u}_i \cdot \vec{u}_j = 0$ *para* $i \neq j$
- $\|\vec{u}_i\| = 1$ *para* $i = 1, 2 \ldots, k$.

La primera propiedad indica que todos los vectores son ortogonales (perpendiculares) dos a dos, mientras que la segunda establece que todos los vectores son unitarios.

Ejemplo 4.7.1. Sea $S = \{\vec{u}_1, \vec{u}_2\}$, donde

$$\vec{u}_1 = \frac{1}{\sqrt{2}}(1, 0, 1) \quad \text{y} \quad \vec{u}_2 = \frac{1}{\sqrt{2}}(-1, 0, 1).$$

Este es un conjunto ortonormal de $\mathbb{R}^3$.

Ejemplo 4.7.2. Sea $S = \{\vec{u}_1, \vec{u}_2.\vec{u}_3\}$, donde

$$\vec{u}_1 = (1, 0, 0)\,, \;\; \vec{u}_2 = (0, 1, 0) \text{ y } \vec{u}_3 = (0, 0, 1)\,.$$

Este es un conjunto ortonormal de $\mathbb{R}^3$.

El siguiente teorema muestra una de las propiedades importantes de los conjuntos ortonormales de vectores.

Teorema 4.7.1. *Sea $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ un conjunto ortonormal de vectores en $\mathbb{R}^n$; entonces, S es un conjunto linealmente independiente.*

Demostración. Consideremos la ecuación

$$\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \ldots + \alpha_k\vec{u}_k = 0.$$

Al multiplicar escalarmente ambos lados de la ecuación dada por $\vec{u}_i$, con $1 \leq i \leq k$, obtenemos

$$(\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \ldots + \alpha_k\vec{u}_k) \cdot \vec{u}_i = 0.$$

Por las propiedades del producto escalar, la ecuación anterior puede escribirse en la forma

$$\alpha_1\vec{u}_1 \cdot \vec{u}_i + \alpha_2\vec{u}_2 \cdot \vec{u}_i + \ldots + \alpha_k\vec{u}_k \cdot \vec{u}_i = 0.$$

De todos los productos escalares, como el conjunto es ortonormal, solo es diferente de cero el producto $\vec{u}_i \cdot \vec{u}_i = 1$, con lo cual obtenemos la relación

$$\alpha_i\vec{u}_i \cdot \vec{u}_i = \alpha_i = 0,$$

lo que nos muestra que todos los α_i son nulos, es decir, el conjunto S es linealmente independiente.

Corolario 4.7.1. *Un conjunto de n vectores ortonormales forma una base de $\mathbb{R}^n$. En este caso se dice que la base es ortonormal.*

Corolario 4.7.2. *Si* $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ *es un conjunto ortonormal de vectores en* $\mathbb{R}^n$*, entonces para cualquier vector* $\vec{v}$ *en* $\mathbb{R}^n$ *se tiene que*

$$\vec{v} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \ldots + \alpha_n\vec{u}_n,$$

donde $\alpha_i = \vec{v} \cdot \vec{u}_i$*,* $i = 1, \ldots, n$*.*

Demostración. Como S es una base de $\mathbb{R}^n$, cualquier vector de $\mathbb{R}^n$ se puede escribir como una combinación lineal de los vectores de S, esto es,

$$\vec{v} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \ldots + \alpha_n\vec{u}_n.$$

Al multiplicar ambos lados de la relación anterior por $\vec{u}_i$ y utilizando el hecho de que $\vec{u}_i \cdot \vec{u}_j = 0$ si $j \neq i$ y que $\vec{u}_i \cdot \vec{u}_i = 1$, tenemos que

$$\begin{aligned} \vec{v} \cdot \vec{u}_i &= \alpha_1\vec{u}_1 \cdot \vec{u}_i + \alpha_2\vec{u}_2 \cdot \vec{u}_i + \ldots + \alpha_n\vec{u}_n \cdot \vec{u}_i \\ &= \alpha_i\vec{u}_i \cdot \vec{u}_i = \alpha_i \end{aligned}$$

que es la relación pedida.

El corolario anterior nos muestra que cuando estamos utilizando bases ortonormales, el problema de la determinación de las componentes de un vector, respecto a los vectores de la base, es mucho más sencillo.

Ejemplo 4.7.3. Sea $S = \{\vec{u}_1, \vec{u}_2.\vec{u}_3\}$, donde

$$\vec{u}_1 = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right), \ \vec{u}_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right) \text{ y } \vec{u}_3 = \left(\frac{2}{3}, \frac{2}{3}, \frac{1}{3}\right).$$

Escribir el vector $\vec{v} = (1, 5, 7)$ como una combinación lineal de los vectores de S.

Solución

Como el conjunto S es ortonormal, constituye una base para $\mathbb{R}^3$; en consecuencia,

$$\vec{v} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \alpha_3\vec{u}_3,$$

donde $\alpha_i = \vec{v} \cdot \vec{u}_i$. De esta forma,

$$\begin{aligned}
\alpha_1 &= \vec{v} \cdot \vec{u}_1 = \frac{1}{3}(1 - 10 + 14) &= \frac{5}{3} \\
\alpha_2 &= \vec{v} \cdot \vec{u}_2 = \frac{1}{3}(-2 + 5 + 14) &= \frac{17}{3} \\
\alpha_3 &= \vec{v} \cdot \vec{u}_3 = \frac{1}{3}(2 + 10 + 7) &= \frac{19}{3}.
\end{aligned}$$

La otra forma de resolver el problema consiste en determinar las soluciones del sistema de tres ecuaciones con tres incógnitas que se obtiene al plantear la ecuación vectorial

$$(1, 5, 7) = \alpha_1 \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3} \right) + \alpha_2 \left(-\frac{2}{3}, \frac{1}{3}, \frac{2}{3} \right) + \alpha_3 \left(\frac{2}{3}, \frac{2}{3}, \frac{1}{3} \right),$$

que es equivalente al sistema de ecuaciones lineales

$$\begin{aligned}
\frac{1}{3}\alpha_1 - \frac{2}{3}\alpha_2 + \frac{2}{3}\alpha_3 &= 1 \\
-\frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2 + \frac{2}{3}\alpha_3 &= 5 \\
\frac{2}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \frac{1}{3}\alpha_3 &= 7,
\end{aligned}$$

cuya solución corresponde a los valores calculados usando el producto escalar.

Como se ha visto previamente, un espacio vectorial puede tener diferentes bases; a continuación presentamos un teorema que permite construir bases ortonormales a partir de una base cualquiera del espacio. Este procedimiento es conocido como el *proceso de ortonormalización de Gram-Schmidt.*

Teorema 4.7.2 (Proceso de Gram-Schmidt). *Sea W un subespacio no nulo de $\mathbb{R}^n$, con una base $S = \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$; entonces, existe una base ortonormal $T = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ para W.*

Demostración. La prueba muestra la forma como se construye la base ortonormal para W. En el proceso, primero se construye una base $R = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, constituida por vectores ortogonales, y luego se llega a la base $T = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$, normalizando cada uno de los vectores de la base ortogonal.

Sea $\vec{v}_1 = \vec{w}_1$. Ahora buscamos un segundo vector $\vec{v}_2$ en el subespacio generado por los vectores $S_1 = \{\vec{w}_1, \vec{w}_2\}$, que sea ortogonal a $\vec{v}_1$.

Como $\vec{v}_1 = \vec{w}_1$, S_1 es también el generado por $\{\vec{v}_1, \vec{w}_2\}$ y, en consecuencia,

$$\vec{v}_2 = \alpha_1 \vec{v}_1 + \alpha_2 \vec{w}_2.$$

Como la condición que queremos es que $\vec{v}_1$ y $\vec{v}_2$ sean ortogonales, al multiplicar ambos lados de la ecuación anterior por $\vec{v}_1$ conseguimos la relación

$$\vec{v}_2 \cdot \vec{v}_1 = \alpha_1 \vec{v}_1 \cdot \vec{v}_1 + \alpha_2 \vec{w}_2 \cdot \vec{v}_1 = 0;$$

como $\vec{v}_1$ es uno de los vectores de la base, $\vec{v}_1 \neq \vec{0}$ y, en consecuencia,

$$\alpha_1 = -\frac{\alpha_2}{\vec{v}_1 \cdot \vec{v}_1}\, \vec{w}_2 \cdot \vec{v}_1.$$

Este es un sistema de una ecuación con dos incógnitas y, por lo tanto, tiene infinitas soluciones que dependen de un parámetro. Si hacemos $\alpha_2 = 1$, entonces

$$\alpha_1 = -\frac{\vec{w}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\,,$$

con lo cual

$$\vec{v}_2 = \vec{w}_2 - \left(\frac{\vec{w}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \right) \vec{v}_1.$$

Estos dos vectores pueden verse en la Figura 4.24. Como puede observarse, el vector $\vec{v}_2$ corresponde a la componente de $\vec{w}_2$, que es perpendicular a $\vec{v}_1$. Ahora buscamos un tercer vector $\vec{v}_3$ en el

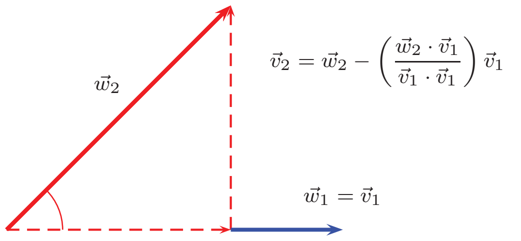


Figura 4.24. Proceso de Gram-Schmidt

subespacio generado por los vectores $S_2 = \{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$, que sea ortogonal a $\vec{v}_1$ y $\vec{v}_2$.

Como S_2 es también el generado por $\{\vec{v}_1, \vec{v}_2, \vec{w}_3\}$, se puede escribir en la forma

$$\vec{v}_3 = \beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \beta_3\vec{w}_3.$$

Como en este caso, se impone a $\vec{v}_3$ es que sea ortogonal tanto a $\vec{v}_1$ como a $\vec{v}_2$, mientras que $\vec{v}_1$ y $\vec{v}_2$ son ortogonales, lo cual equivale a las condiciones:

$$\vec{v}_1 \cdot \vec{v}_2 = 0, \;\; \vec{v}_3 \cdot \vec{v}_1 = 0 \text{ y } \vec{v}_3 \cdot \vec{v}_2 = 0.$$

Las dos últimas condiciones nos proporcionan el siguiente sistema de ecuaciones:

$$\begin{aligned} \vec{v}_3 \cdot \vec{v}_1 &= \beta_1\vec{v}_1 \cdot \vec{v}_1 + \beta_3\vec{w}_3 \cdot \vec{v}_1 = 0 \\ \vec{v}_3 \cdot \vec{v}_2 &= \beta_2\vec{v}_2 \cdot \vec{v}_2 + \beta_3\vec{w}_3 \cdot \vec{v}_2 = 0. \end{aligned}$$

Este es un sistema de dos ecuaciones con tres incógnitas, con lo cual las infinitas soluciones pueden escribirse en términos de un parámetro, en este caso, β_3.

Puesto que $\vec{v}_1 \cdot \vec{v}_1 \neq 0$ y $\vec{v}_2 \cdot \vec{v}_2 \neq 0$, es posible despejar β_1 y β_2

para obtener

$$\begin{aligned}\beta_1 &= -\beta_3 \frac{\vec{w}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\\ \beta_2 &= -\beta_3 \frac{\vec{w}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}.\end{aligned}$$

Haciendo $\beta_3 = 1$, se tiene que

$$\begin{aligned}\beta_1 &= -\frac{\vec{w}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\\ \beta_2 &= -\frac{\vec{w}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2},\end{aligned}$$

de donde

$$\vec{v}_3 = \vec{w}_3 - \left(\frac{\vec{w}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 - \left(\frac{\vec{w}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\right)\vec{v}_2.$$

Gráficamente, el vector $\vec{v}_3$ corresponde a la componente $\vec{w}_3$ en la dirección del vector normal al plano determinado por $\vec{v}_1$ y $\vec{v}_2$, tal como lo muestra la Figura 4.25.

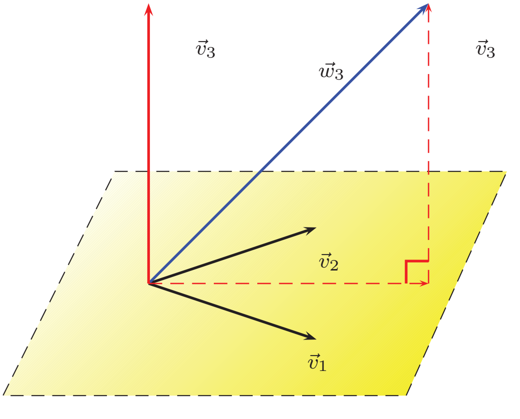


Figura 4.25. Vector ortogonal a otros dos

Continuando con este procedimiento, conseguimos la siguiente expresión para el vector $\vec{v}_j$, con $j = 1, \ldots, k$.

$$\vec{v}_j = \vec{w}_j - \left(\frac{\vec{w}_j \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 - \left(\frac{\vec{w}_j \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\right)\vec{v}_2 - \cdots - \left(\frac{\vec{w}_j \cdot \vec{v}_{j-1}}{\vec{v}_{j-1} \cdot \vec{v}_{j-1}}\right)\vec{v}_{j-1}.$$

El conjunto de vectores $T = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ es una base ortogonal para W. La base ortonormal $T = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ para W se obtiene normalizando los vectores $\vec{u}_j$, es decir, haciendo

$$\vec{u}_j = \frac{1}{\|\vec{v}_j\|}\vec{v}_j,$$

lo cual nos proporciona la base pedida.

El siguiente ejemplo ilustra la forma recursiva como se construye una base ortonormal a partir de una base cualquiera de un espacio vectorial.

Ejemplo 4.7.4. Sea la base dada por los vectores $B = \{w_1, w_2, w_3\}$ de $\mathbb{R}^3$, donde

$$w_1 = (1, 0, 1); \quad w_2 = (1, -2, 1) \text{ y } w_3 = (0, 1, 1).$$

Aplicar el proceso de ortogonalización de Gram-Schmidt para transformar esta base en una base ortonormal.

Solución

En primer lugar, vamos a determinar los vectores de la base ortogonal. En este caso hacemos $v_1 = w_1 = (1, 0, 1)$. Para el siguiente vector tenemos la relación

$$\vec{v}_2 = \vec{w}_2 - \left(\frac{\vec{w}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 = (1, -2, 1) - \frac{2}{2}(1, 0, 1) = (0, -2, 0).$$

La relación para el siguiente vector es

$$\begin{aligned}
\vec{v}_3 &= \vec{w}_3 - \left(\frac{\vec{w}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 - \left(\frac{\vec{w}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\right)\vec{v}_2 \\
&= (0,1,1) - \frac{1}{2}(1,0,1) - \frac{-2}{4}(0,-2,0) \\
&= \frac{1}{2}(-1,0,1).
\end{aligned}$$

El conjunto conformado por los vectores $T = \{v_1, v_2, v_3\}$ constituyen una base ortogonal, mientras que los vectores

$$\begin{aligned}
\vec{u}_1 &= \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{\sqrt{2}}(1,0,1), \\
\vec{u}_2 &= \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \frac{1}{2}(0,-2,0) \text{ y} \\
\vec{u}_3 &= \frac{1}{\|\vec{v}_3\|}\vec{v}_3 = \frac{1}{2\sqrt{2}}(-1,0,1)
\end{aligned}$$

forman la base ortonormal pedida.

4.8 Ejercicios propuestos capítulo 4

1. Demuestre que la suma de los cuadrados de las diagonales de un paralelogramo es igual a la suma de los cuadrados de sus cuatro lados.

2. Demuestre que los segmentos trazados desde los pies de dos de las alturas de un triángulo al punto medio del tercer lado son iguales.

3. Si $\vec{a}$ y $\vec{b}$ son dos vectores linealmente independientes, pruebe que
$$\frac{\|\vec{a}\|\vec{b} + \|\vec{b}\|\vec{a}}{\|\vec{a}\|\|\vec{b}\|}$$

tiene la misma dirección que la bisectriz del ángulo formado por los dos vectores $\vec{a}$ y $\vec{b}$.

4. Demuestre el teorema de Apolonio: si ABC es un triángulo cualquiera, la suma de los cuadrados de dos lados es igual al doble del cuadrado de la mediana relativa al tercer lado, más la mitad del cuadrado del tercer lado.

5. Demuestre que los vectores

$$\vec{c} = \|\vec{a}\|\vec{b} + \|\vec{b}\|\vec{a} \quad \text{y} \quad \vec{d} = \|\vec{a}\|\vec{b} - \|\vec{b}\|\vec{a}$$

son ortogonales.

6. Dado un triángulo isósceles, pruebe que la mediana relativa a la base (lado desigual) es perpendicular a dicho lado.

7. Demuestre que

$$\vec{a} \cdot \vec{b} = \frac{1}{4}\left[\|\vec{a} + \vec{b}\|^2 - \|\vec{a} - \vec{b}\|^2\right].$$

8. Si desde un punto exterior a una circunferencia de radio r se traza una tangente y una secante que pase por el centro, pruebe que la tangente es media proporcional entre la secante y su segmento externo.

9. Sea el triángulo ABC; demostrar que

$$\|\overrightarrow{AB}\| = \|\overrightarrow{AC}\| \cos\alpha + \|\overrightarrow{BC}\| \cos\beta,$$

donde α y β son los ángulos en los vértices A y B respectivamente.

10. Demuestre que el punto medio de la hipotenusa de un triángulo rectángulo equidista de los tres vértices.

11. Dado el producto interno para funciones $C(-1,1)$:

$$\langle f(x), g(x)\rangle = \langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx,$$

considere las funciones $f(x) = 1$, $g(x) = x$ y $h(x) = x + 1$. Pruebe que dos de ellas son ortogonales, dos forman un ángulo de $\pi/3$ y dos forman un ángulo de $\pi/6$.

12. Sean $x = (x_1, \dots, x_n)$ y $y = (y_1, \dots, y_n)$ elementos de $\mathbb{R}^n$. Determine si las siguientes fórmulas corresponden a productos internos:

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i|y_i|;\ \langle x, y\rangle = \left(\sum_{i=1}^{n} x_i^2 y_i^2\right)^{1/2};$$

$$\langle x, y\rangle = \left|\sum_{i=1}^{n} x_i y_i\right|;\ \langle x, y\rangle = \sum_{i=1}^{n} |x_i y_i|.$$

13. En cada uno de los casos, halle una base ortonormal para el subespacio generado por los vectores dados.

    a) $w_1 = (1,1,1)$, $w_2 = (1,1,0)$ y $w_3 = (0,1,-1)$.

    b) $w_1 = (2,1,-1)$, $w_2 = (1,1,0)$ y $w_3 = (2,1,1)$.

14. En cada uno de los casos, halle una base ortonormal para el subespacio generado por los vectores dados.

    a) $w_1 = (1,1,1,0)$, $w_2 = (1,1,0,0)$, $w_3 = (0,1,-1,-1)$ y $w_4 = (1,0,1,0)$

    b) $w_1 = (2,1,-1,0)$, $w_2 = (1,1,0,0)$, $w_3 = (2,1,1,0)$ y $w_4 = (-1,0,-1,0)$.

    c) $w_1 = (2,0,-1,1)$, $w_2 = (-1,1,0,0)$, $w_3 = (-2,1,-1,0)$ y $w_4 = (1,0,-1,1)$.

15. Dados los siguientes vectores, halle el ángulo entre cada pareja de ellos:

    a) $v_1 = (1, 1, 1)$, $v_2 = (1, 1, 0)$ y $v_3 = (0, 2, -1)$.

    b) $u_1 = (-2, 1, -1)$, $u_2 = (-1, 1, 0)$ y $u_3 = (2, -1, 1)$.

16. Dados los vectores

$$\begin{array}{ll} v_1 = (1, 1, 1) & v_4 = (-1, 0, 1) \\ v_2 = (2, 2, 2) & v_5 = (-1, 1, 0) \\ v_3 = (3, 1, 1) & v_6 = (-1, 0, 3), \end{array}$$

    ¿cuáles de ellos son perpendiculares?

17. Calcule las proyecciones escalares y vectoriales del vector $v_1 = (1, -3, 1)$ sobre $v_2 = (-2, -5, 1)$.

18. Descomponer el vector $v_1 = (1, -3, 1)$ como la suma de un vector paralelo a $v_2 = (2, 5, -1)$ y otro perpendicular a v_2.

19. De ser posible, calcule los valores de a, b y c de modo que el vector $v = (a, b, c)$ sea ortogonal a los vectores $v_1 = (1, 3, 1)$ y $v_2 = (-1, 4, 2)$.

20. Verifique que el triángulo con vértices en $A(3, 1, 2)$, $B(-3, 0, 4)$ y $C(2, 3, -4)$ es isósceles.

21. Halle la ecuación del plano normal al vector $n = (-2, 1, -4)$ que pasa por el punto $P(2, 1, 8)$.

22. Halle el ángulo entre los planos $x+y+z = 2$ y $2x-4y+z = 5$.

23. Determine los valores de la constante a de tal modo que $v = (a, -a, 8)$ sea ortogonal a $u = (1, 8, 2)$.

24. Halle la distancia entre las rectas L_1 y L_2 con ecuaciones paramétricas dadas por

$$\begin{aligned} x &= 3t+2 \\ y &= t \\ z &= 2t+1 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x &= s+3 \\ y &= s-3 \\ z &= s+1. \end{aligned}$$

25. Halle la distancia del punto $P = (2, 5)$ a la recta con ecuación $4x - 5y = 9$.

26. Halle la ecuación de la superficie cónica con vértice en el origen, eje $\vec{e} = (1, 0, 0)$ y generatriz $\vec{g} = (1, 2, 3)$.

Capítulo 5

Producto vectorial

5.1 Introducción

Desde el punto de vista de las operaciones entre vectores, aún falta por definir una operación que combine dos vectores y dé por resultado un nuevo vector. Desde la óptica de la física es bien conocido el hecho de que al aplicar una fuerza sobre un sólido rígido, el efecto de la fuerza depende del punto de aplicación de esta. Tal observación da origen al concepto de *momento o torque de una fuerza* con respecto a un punto fijo. Estos dos comentarios, justifican la definición del producto vectorial entre dos vectores. Como se ve más adelante, las aplicaciones de este producto se extienden a la representación vectorial de planos y al cálculo de áreas. El concepto de *producto vectorial* que se maneja en este capítulo está restringido al caso de los vectores libres de $\mathbb{R}^3$; además, se usa como base el conjunto ortonormal determinado por los vectores $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Respecto a la nomenclatura, en adelante los vectores se denotan con letras en negrita.

5.2 Producto vectorial

Definición 5.2.1. *Dados los vectores* $\mathbf{a}$ *y* $\mathbf{b}$ *de* $\mathbb{E}^3$, *el producto vectorial entre* $\mathbf{a}$ *y* $\mathbf{b}$, *que se denota con* $\mathbf{a} \times \mathbf{b}$, *se define como un nuevo vector que posee las siguientes características:*

1. *El vector* $\mathbf{a} \times \mathbf{b}$ *es perpendicular al plano que contiene a los vectores* $\mathbf{a}$ *y* $\mathbf{b}$.
2. *El módulo de* $\mathbf{a} \times \mathbf{b}$ *está dado por*

   $$\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\|\mathbf{b}\| \operatorname{sen} \theta,$$

   siendo θ *el ángulo determinado por los dos vectores.*
3. *El sentido de* $\mathbf{a} \times \mathbf{b}$ *lo determina la regla de la mano derecha (si con el dedo índice se marca el sentido del primer vector y con el dedo del corazón se indica el sentido del segundo vector, entonces* $\mathbf{a} \times \mathbf{b}$ *tiene el sentido del dedo pulgar). Esta última regla es equivalente a afirmar que los vectores* $\mathbf{a}$, $\mathbf{b}$ *y* $\mathbf{a} \times \mathbf{b}$ *determinan un triedro a derechas.*

De la definición anterior se concluye inmediatamente que:

- El producto vectorial es anticonmutativo, es decir $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.
- Si $\mathbf{a}$ y $\mathbf{b}$ son vectores paralelos, $\mathbf{a} \times \mathbf{b} = 0$.
- Para cualquier vector $\mathbf{a}$, $\mathbf{a} \times \mathbf{a} = 0$.

Para los vectores ortonormales $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ se tienen los siguientes resultados:

$$\begin{array}{lll} \mathbf{i} \times \mathbf{j} = \mathbf{k} & \mathbf{j} \times \mathbf{k} = \mathbf{i} & \mathbf{k} \times \mathbf{i} = \mathbf{j} \\ \mathbf{j} \times \mathbf{i} = -\mathbf{k} & \mathbf{k} \times \mathbf{j} = -\mathbf{i} & \mathbf{i} \times \mathbf{k} = -\mathbf{j}. \end{array}$$

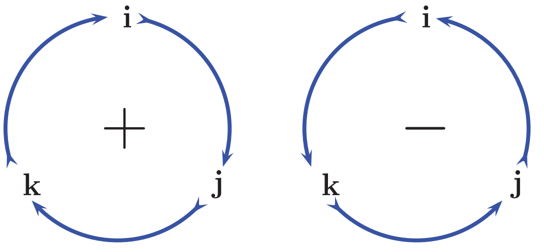


Figura 5.1. Sentidos positivo y negativo del producto vectorial

Estas relaciones se recuerdan fácilmente mediante los diagramas de la Figura 5.1.

Una de las propiedades más significativas del producto vectorial es la ley distributiva, es decir, dados los vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$, entonces se cumple que: $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$. Para demostrar dicha propiedad, veamos primero el siguiente teorema.

Teorema 5.2.1. *Sea el vector* $\mathbf{v_1}$ *y sean los puntos* $A_1, A_2, \ldots, A_k$ *tomados sobre una paralela al soporte de* $\mathbf{v_1}$*; entonces, si* O *es el origen del vector* $\mathbf{v_1}$*, se cumple que:*

$$\mathbf{v_1} \times \mathbf{OA_1} = \mathbf{v_1} \times \mathbf{OA_2} = \ldots = \mathbf{v_1} \times \mathbf{OA_k} = \textit{constante}.$$

El teorema anterior afirma que todos vectores que sean concurrentes con $\mathbf{v_1}$ y tengan su extremo sobre una paralela al soporte de $\mathbf{v_1}$ tienen el mismo producto vectorial con $\mathbf{v_1}$, como se puede observar en el diagrama representado en la Figura 5.2.

En el caso del vector $\mathbf{OA_1}$, que es perpendicular a $\mathbf{v_1}$, el producto vectorial se puede interpretar como un operador que actúa sobre $\mathbf{OA_1}$, transformándolo de la siguiente manera:

1. Le da a $\mathbf{OA_1}$ un giro de 90°.

2. Amplifica a $\mathbf{OA_1}$ en un factor de $\|\mathbf{v_1}\|$, es decir,

$$\|\mathbf{v_1} \times \mathbf{OA_1}\| = \|\mathbf{v_1}\|\|\mathbf{OA_1}\|.$$

Demostración. Según la definición de producto vectorial, tenemos lo siguiente:

- Como los vectores $\mathbf{OA_1}, \mathbf{OA_2}, \ldots, \mathbf{OA_k}$ están en el mismo plano, entonces todos los productos que tienen la forma $\mathbf{v_1} \times \mathbf{OA_1}, \mathbf{v_1} \times \mathbf{OA_2}, \ldots, \mathbf{v_1} \times \mathbf{OA_k}$ son paralelos.

- $\|\mathbf{v_i} \times \mathbf{OA_i}\| = \|\mathbf{v_1}\|\|\mathbf{OA_i}\| \operatorname{sen}\theta_i = \|\mathbf{v_1}\|\|\mathbf{OA_1}\|$, donde θ_i es el ángulo entre los vectores $\mathbf{OA_i}$ y $\mathbf{v_1}$, para cualquier $i = 1, \ldots k$.

- Todos los vectores $\mathbf{v_1} \times \mathbf{OA_1}, \mathbf{v_1} \times \mathbf{OA_2}, \ldots, \mathbf{v_1} \times \mathbf{OA_k}$ tienen el mismo sentido de giro.

Por lo tanto,

$$\mathbf{v_1} \times \mathbf{OA_1} = \mathbf{v_1} \times \mathbf{OA_2} = \ldots = \mathbf{v_1} \times \mathbf{OA_k} = \textit{constante}.$$

Con este teorema a disposición, podemos establecer el siguiente teorema.

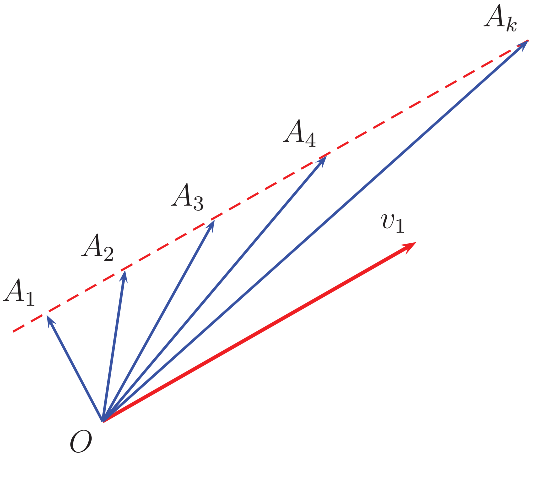


Figura 5.2. Invarianza del producto vectorial

Teorema 5.2.2. *Dados los vectores* $\mathbf{a}$, $\mathbf{b}$ *y* $\mathbf{c}$, *entonces se cumple que:*

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}.$$

Demostración. Para la prueba vamos a considerar tres casos diferentes:

1. Vectores en el mismo plano. En la Figura 5.3 se muestran los vectores $\mathbf{v_2}$, $\mathbf{v_3}$ y $\mathbf{v_2} + \mathbf{v_3}$, así como sus proyecciones sobre la recta L perpendicular al soporte de $\mathbf{v_1}$. Aplicando el teorema anterior se tiene:

$$\begin{aligned} \mathbf{v_1} \times \mathbf{v_2} &= \mathbf{v_1} \times \mathbf{OA} \\ \mathbf{v_1} \times \mathbf{v_3} &= \mathbf{v_1} \times \mathbf{OB} \\ \mathbf{v_1} \times (\mathbf{v_2} + \mathbf{v_3}) &= \mathbf{v_1} \times \mathbf{OC}. \end{aligned}$$

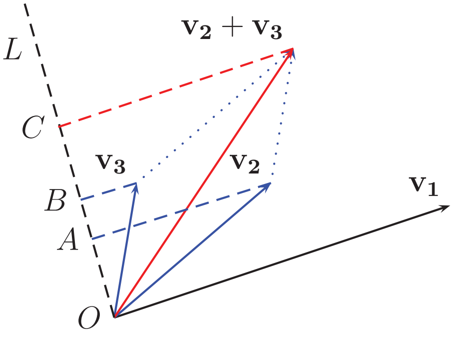


Figura 5.3. Distributividad para vectores en el mismo plano

Como todos estos productos vectoriales están sobre el mismo eje y apuntan en el mismo sentido (saliendo del plano de la hoja) y como, además, $\mathbf{OC} = \mathbf{OA} + \mathbf{OB}$, entonces se cumple que

$$\mathbf{v_1} \times \mathbf{OC} = \mathbf{v_1} \times \mathbf{OA} + \mathbf{v_1} \times \mathbf{OB},$$

con lo cual

$$\mathbf{v_1} \times (\mathbf{v_2} + \mathbf{v_3}) = \mathbf{v_1} \times \mathbf{v_2} + \mathbf{v_1} \times \mathbf{v_3}.$$

2. El vector $\mathbf{v_1}$ *perpendicular a* $\mathbf{v_2}$ *y* $\mathbf{v_3}$. En este caso, los productos $\mathbf{v_1}\times\mathbf{v_2}$ y $\mathbf{v_1}\times\mathbf{v_3}$ están en el plano determinado por los vectores $\mathbf{v_2}$ y $\mathbf{v_3}$. Si recordamos la interpretación que se dio al producto vectorial por $\mathbf{v_1}$, los productos $\mathbf{v_1} \times \mathbf{v_2}$ y $\mathbf{v_1} \times \mathbf{v_3}$ son rotaciones de 90° de los vectores $\mathbf{v_2}$ y $\mathbf{v_3}$, y los módulos de los productos vectoriales son $\|\mathbf{v_1} \times \mathbf{v_2}\| = \|\mathbf{v_1}\|\|\mathbf{v_2}\|$ y $\|\mathbf{v_1} \times \mathbf{v_3}\| = \|\mathbf{v_1}\|\|\mathbf{v_3}\|$.

Como puede apreciarse en la Figura 5.4, tenemos que

$$\begin{array}{rcl} \|\mathbf{OA'}\| & = & \|\mathbf{v_1}\|\|\mathbf{OA}\| \\ \|\mathbf{OC'}\| & = & \|\mathbf{v_1}\|\|\mathbf{OC}\| \end{array}$$

y, además,

$$\begin{array}{rcl} \mathbf{OA'} & \perp & \mathbf{OA} \\ \mathbf{OC'} & \perp & \mathbf{OC}, \end{array}$$

con lo cual podemos concluir que los paralelogramos $OABC$ y $OA'B'C'$ son semejantes y, en consecuencia,

$$\|\mathbf{OB'}\| = \|\mathbf{v_1}\|\|\mathbf{OB}\| \quad \text{y} \quad \mathbf{OB'} \perp \mathbf{OB}.$$

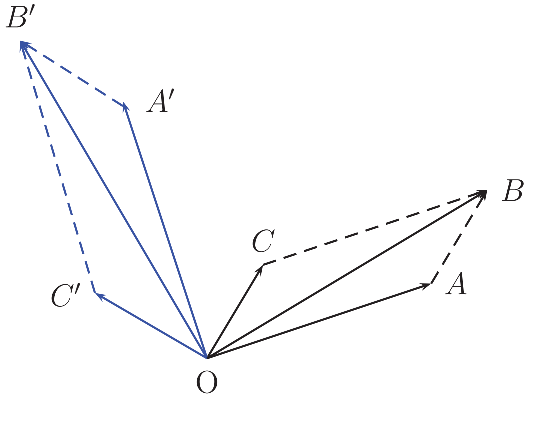


Figura 5.4. Distributividad para $\mathbf{v_1}$ perpendicular a $\mathbf{v_2}$ y $\mathbf{v_3}$

Esta última ecuación significa que

$$\mathbf{OB}' = \mathbf{v_1} \times \mathbf{OB};$$

pero como

$$\mathbf{OB}' = \mathbf{OA}' + \mathbf{OC}' = \mathbf{v_1} \times \mathbf{OA} + \mathbf{v_1} \times \mathbf{OC} = \mathbf{v_1} \times \mathbf{v_2} + \mathbf{v_1} \times \mathbf{v_3},$$

mientras que

$$\mathbf{OB} = \mathbf{OA} + \mathbf{OC} = \mathbf{v_2} + \mathbf{v_3},$$

a partir de ello, concluimos que

$$\mathbf{v_1} \times (\mathbf{v_2} + \mathbf{v_3}) = \mathbf{v_1} \times \mathbf{v_2} + \mathbf{v_1} \times \mathbf{v_3}.$$

3. Vectores arbitrarios. Consideremos los vectores $\mathbf{v_1}$, $\mathbf{v_2}$ y $\mathbf{v_3}$ con origen en el punto O y con extremos en puntos arbitrarios del espacio, como puede verse en la Figura 5.5.

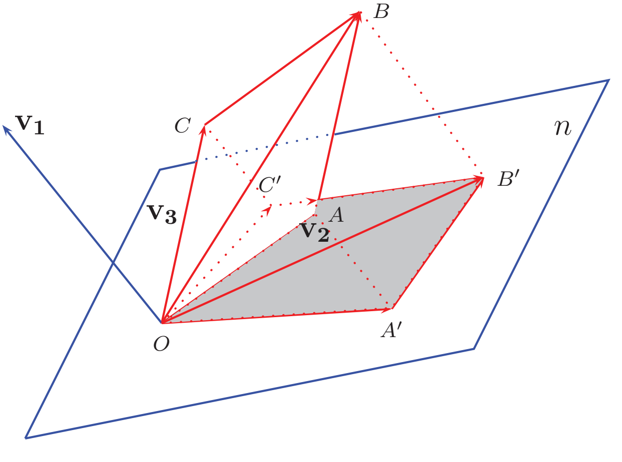


Figura 5.5. Ley distributiva, caso general

Sea n un plano normal al vector $\mathbf{v_1}$, pasando por el punto de concurrencia de los tres vectores; establezcamos, además, las proyecciones ortogonales de los vectores $\mathbf{v_2} = \mathbf{OA}$, $\mathbf{v_3} = \mathbf{OC}$ y $\mathbf{v_2} + \mathbf{v_3} = \mathbf{OB}$ sobre el plano n.

El cuadrilátero $OA'B'C'$, que es la proyección del paralelogramo $OABC$, es también un paralelogramo y, según el teorema 5.2.1, tenemos que:

$$\mathbf{v_1} \times \mathbf{v_2} = \mathbf{v_1} \times \mathbf{OA'}; \quad \mathbf{v_1} \times \mathbf{v_3} = \mathbf{v_1} \times \mathbf{OC'} \text{ y } \mathbf{v_1} \times \mathbf{OB} = \mathbf{v_1} \times \mathbf{OB'}.$$

Como los vectores $\mathbf{OA'}$ y $\mathbf{OC'}$ son ambos perpendiculares a $\mathbf{v_1}$, del caso anterior se concluye que

$$\mathbf{v_1} \times \mathbf{OA'} + \mathbf{v_1} \times \mathbf{OC'} = \mathbf{v_1} \times (\mathbf{OA'} + \mathbf{OC'}) = \mathbf{v_1} \times \mathbf{OB'},$$

con lo cual

$$\mathbf{v_1} \times (\mathbf{v_2} + \mathbf{v_3}) = \mathbf{v_1} \times \mathbf{v_2} + \mathbf{v_1} \times \mathbf{v_3}.$$

La ley distributiva puede ampliarse para el caso de una cantidad arbitraria de vectores en cada uno de los factores. En particular, tenemos que

$$(\mathbf{v_1} + \mathbf{v_2}) \times (\mathbf{v_3} + \mathbf{v_4}) = \mathbf{v_1} \times \mathbf{v_3} + \mathbf{v_1} \times \mathbf{v_4} + \mathbf{v_2} \times \mathbf{v_3} + \mathbf{v_2} \times \mathbf{v_4}.$$

Otra propiedad útil del producto vectorial es aquella que permite asociar los escalares al comienzo del producto, es decir:

$$\alpha\mathbf{v_1} \times \beta\mathbf{v_2} = (\alpha\beta)\mathbf{v_1} \times \mathbf{v_2},$$

la cual se deduce inmediatamente de los siguientes hechos:

- Dados los vectores $\mathbf{v_1}$ y $\mathbf{v_2}$ y los escalares α y β, entonces

  $$\|\alpha\mathbf{v_1} \times \beta\mathbf{v_2}\| = |\alpha|\|\mathbf{v_1}\||\beta|\|\mathbf{v_2}\| \operatorname{sen}\theta',$$

  siendo θ' el ángulo formado por los vectores $\alpha\mathbf{v_1}$ $\beta\mathbf{v_2}$, es decir, θ o $180° - \theta$, según sean los signos de los escalares α y β. En ambos casos se tiene que $\operatorname{sen}\theta' = \operatorname{sen}\theta$, con lo que

  $$\|\alpha\mathbf{v_1} \times \beta\mathbf{v_2}\| = |\alpha\beta|\|\mathbf{v_1}\|\|\mathbf{v_2}\| \operatorname{sen}\theta.$$

- Los productos $\alpha\mathbf{v_1} \times \beta\mathbf{v_2}$ y $(\alpha\beta)\mathbf{v_1} \times \mathbf{v_2}$ están sobre el mismo eje y tienen siempre la misma orientación.

Las dos propiedades anteriores permiten establecer relaciones sencillas para calcular el producto vectorial entre dos vectores cuando ambos están referidos a una base ortonormal. De ahora en adelante, al referirnos a vectores en $\mathbb{E}^3$, las coordenadas de estos vectores estarán dadas con respecto a la base canónica $B = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, la cual conforma un triedro a derechas.

De esta forma, si

$$\mathbf{a} = \alpha_1\mathbf{i} + \alpha_2\mathbf{j} + \alpha_3\mathbf{k} \quad \text{y} \quad \mathbf{b} = \beta_1\mathbf{i} + \beta_2\mathbf{j} + \beta_3\mathbf{k},$$

entonces

$$\begin{aligned}
\mathbf{a} \times \mathbf{b} &= (\alpha_1\mathbf{i} + \alpha_2\mathbf{j} + \alpha_3\mathbf{k}) \times (\beta_1\mathbf{i} + \beta_2\mathbf{j} + \beta_3\mathbf{k}) \\
&= \alpha_1\beta_1\mathbf{i} \times \mathbf{i} + \alpha_1\beta_2\mathbf{i} \times \mathbf{j} + \alpha_1\beta_3\mathbf{i} \times \mathbf{k} + \alpha_2\beta_1\mathbf{j} \times \mathbf{i} + \alpha_2\beta_2\mathbf{j} \times \mathbf{j} \\
&\quad + \alpha_2\beta_3\mathbf{j} \times \mathbf{k} + \alpha_3\beta_1\mathbf{k} \times \mathbf{i} + \alpha_3\beta_2\mathbf{k} \times \mathbf{j} + \alpha_3\beta_3\mathbf{k} \times \mathbf{k}.
\end{aligned}$$

Como hemos probado antes, para los vectores ortonormales $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$, al calcular los productos vectoriales, se tienen los siguientes resultados:

$$\begin{array}{lll}
\mathbf{i} \times \mathbf{j} = \mathbf{k} & \mathbf{j} \times \mathbf{k} = \mathbf{i} & \mathbf{k} \times \mathbf{i} = \mathbf{j} \\
\mathbf{j} \times \mathbf{i} = -\mathbf{k} & \mathbf{k} \times \mathbf{j} = -\mathbf{i} & \mathbf{i} \times \mathbf{k} = -\mathbf{j} \\
\mathbf{i} \times \mathbf{i} = \mathbf{0} & \mathbf{j} \times \mathbf{j} = \mathbf{0} & \mathbf{k} \times \mathbf{k} = \mathbf{0},
\end{array}$$

con lo cual obtenemos la relación

$$\mathbf{a} \times \mathbf{b} = (\alpha_2\beta_3 - \alpha_3\beta_2)\mathbf{i} - (\alpha_1\beta_3 - \alpha_3\beta_1)\mathbf{j} + (\alpha_1\beta_2 - \alpha_2\beta_1)\mathbf{k}.$$

Al comparar el desarrollo de un determinante de una matriz 3×3 podemos ver la analogía que existe entre este y las relaciones establecidas para el producto vectorial. Según esto, es posible escribir el cálculo del producto vectorial en la siguiente forma:

$$\mathbf{a}\times\mathbf{b}=\det\begin{bmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ \alpha_1 & \alpha_2 & \alpha_3\\ \beta_1 & \beta_2 & \beta_3\end{bmatrix}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ \alpha_1 & \alpha_2 & \alpha_3\\ \beta_1 & \beta_2 & \beta_3\end{vmatrix}.$$

Debemos aclarar que esta expresión es solo una manera de memorizar uno de los modos de calcular el producto vectorial, pero realmente este es un vector, mientras que la definición establecida para el determinante es un número real. Además, es importante mencionar que en esta forma de calcular el producto vectorial, deben usarse los cofactores de la primera fila.

Ejemplo 5.2.1. Dados los vectores

$$\mathbf{a}=3\mathbf{i}-2\mathbf{j}+\mathbf{k}\quad \text{y}\quad \mathbf{b}=\mathbf{i}+\mathbf{j}+2\mathbf{k},$$

determine su producto vectorial.

Solución

Aplicando la última de las fórmulas establecidas, obtenemos que

$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 3 & -2 & 1\\ 1 & 1 & 2\end{vmatrix},$$

con lo cual

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= \begin{vmatrix}-2 & 1\\ 1 & 2\end{vmatrix}\mathbf{i}-\begin{vmatrix}3 & 1\\ 1 & 2\end{vmatrix}\mathbf{j}+\begin{vmatrix}3 & -2\\ 1 & 1\end{vmatrix}\mathbf{k}\\ &= -5\mathbf{i}-5\mathbf{j}+5\mathbf{k}.\end{aligned}$$

Puede verificarse de modo sencillo que $\mathbf{a}\times\mathbf{b}$ es perpendicular a cada uno de los vectores $\mathbf{a}$ y $\mathbf{b}$.

Dados los vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ de $\mathbb{E}^3$ referidos a la base canónica, se puede calcular el triple producto vectorial. Una de las formas de obtener este producto la proporciona la expresión $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$. Como

$\mathbf{b} \times \mathbf{c}$ es normal al plano de $\mathbf{b}$ y $\mathbf{c}$, entonces $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ es normal a $(\mathbf{b} \times \mathbf{c})$ y, en consecuencia, yace sobre el plano determinado por $\mathbf{b}$ y $\mathbf{c}$ (si ellos son linealmente independientes).

Como $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ está sobre el plano determinado por $\mathbf{b}$ y $\mathbf{c}$, este producto puede escribirse como una combinación lineal de los vectores $\mathbf{b}$ y $\mathbf{c}$; este resultado es conocido como la *relación de Gibbs*, la cual constituye la propiedad más importante del triple producto vectorial y es el enunciado del siguiente teorema.

Teorema 5.2.3. *Sean los vectores* $\mathbf{a}$, $\mathbf{b}$ *y* $\mathbf{c}$ *de* $\mathbb{E}^3$ *referidos a la base canónica. Entonces,*

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}.$$

Demostración. En caso de que los vectores $\mathbf{b}$ y $\mathbf{c}$ sean linealmente dependientes, se tiene que $\mathbf{b} = \alpha\mathbf{c}$ y, en consecuencia,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \alpha(\mathbf{a} \cdot \mathbf{c})\mathbf{c} - \alpha(\mathbf{a} \cdot \mathbf{c})\mathbf{c} = 0.$$

Veamos ahora el caso cuando los vectores $\mathbf{b}$ y $\mathbf{c}$ son linealmente independientes. Como $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ está sobre el plano determinado por los vectores $\mathbf{b}$ y $\mathbf{c}$, tenemos que

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \alpha\mathbf{b} + \beta\mathbf{c}.$$

Calculemos los coeficientes α y β. Para ello consideremos una base ortonormal $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ a derechas, de tal forma que $\mathbf{e}_2$ sea paralelo a $\mathbf{b}$, mientras que $\{\mathbf{e}_2, \mathbf{e}_3\}$ conforman un plano paralelo al plano determinado por $\mathbf{b}$ y $\mathbf{c}$. Con esta elección de la base, los tres vectores originales pueden escribirse en la forma

$$\begin{aligned}
\mathbf{a} &= \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 \\
\mathbf{c} &= \gamma_2\mathbf{e}_2 + \gamma_3\mathbf{e}_3 \\
\mathbf{b} &= \beta_2\mathbf{e}_2,
\end{aligned}$$

con lo cual, si aprovechamos la distributividad del producto vectorial y el hecho de que $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ es una base ortonormal, concluimos que:

$$\mathbf{b} \times \mathbf{c} = \beta_2\gamma_3\mathbf{e}_2 \times \mathbf{e}_3 = \beta_2\gamma_3\mathbf{e}_1$$

y, por lo tanto,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \alpha_2\beta_2\gamma_3\mathbf{e}_2 \times \mathbf{e}_1 + \alpha_3\beta_2\gamma_3\mathbf{e}_3 \times \mathbf{e}_1.$$

Como $\mathbf{e}_2 \times \mathbf{e}_1 = -\mathbf{e}_3$ y $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$, obtenemos la siguiente expresión para el triple producto vectorial:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = -\alpha_2\beta_2\gamma_3\mathbf{e}_3 + \alpha_3\beta_2\gamma_3\mathbf{e}_2.$$

Igualando las dos expresiones propuestas para $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ se obtiene

$$\alpha\mathbf{b} + \beta\mathbf{c} = -\alpha_2\beta_2\gamma_3\mathbf{e}_3 + \alpha_3\beta_2\gamma_3\mathbf{e}_2,$$

de donde

$$\alpha\beta_2\mathbf{e}_2 + \beta(\gamma_2\mathbf{e}_2 + \gamma_3\mathbf{e}_3) = -\alpha_2\beta_2\gamma_3\mathbf{e}_3 + \alpha_3\beta_2\gamma_3\mathbf{e}_2,$$

que puede escribirse en la forma

$$(\alpha\beta_2 + \beta\gamma_2 - \alpha_3\beta_2\gamma_3)\mathbf{e}_2 + (\beta\gamma_3 + \alpha_2\beta_2\gamma_3)\mathbf{e}_3 = 0.$$

Pero como los vectores $\mathbf{e}_2$ y $\mathbf{e}_3$ son linealmente independientes, sus coeficientes son nulos, es decir,

$$\begin{aligned} \alpha\beta_2 + \beta\gamma_2 - \alpha_3\beta_2\gamma_3 &= 0 \qquad (5.2.1)\\ \beta\gamma_3 + \alpha_2\beta_2\gamma_3 &= 0. \qquad (5.2.2) \end{aligned}$$

De la ecuación 5.2.2, como $\gamma_3 \neq 0$ puesto que $\mathbf{c}$ y $\mathbf{b}$ son linealmente independientes, se obtiene que

$$\beta = -\alpha_2\beta_2 = -\mathbf{a} \cdot \mathbf{b}.$$

Al sustituir este valor en la ecuación 5.2.1, conseguimos la relación

$$\alpha\beta_2 - \alpha_2\beta_2\gamma_2 - \alpha_3\beta_2\gamma_3 = \beta_2(\alpha - \alpha_2\gamma_2 - \alpha_3\gamma_3) = 0.$$

De nuevo, como $\beta_2 \neq 0$, se obtiene que

$$\alpha - \alpha_2\gamma_2 - \alpha_3\gamma_3 = 0,$$

con lo cual

$$\alpha = \alpha_2\gamma_2 + \alpha_3\gamma_3 = \mathbf{a} \cdot \mathbf{c}.$$

Por lo tanto,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \alpha\mathbf{b} + \beta\mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c},$$

que era la relación que queríamos demostrar.

De acuerdo con las propiedades del producto vectorial,

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = -\mathbf{c} \times (\mathbf{a} \times \mathbf{b});$$

en este caso, la relación de Gibbs toma la forma

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = -\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = -(\mathbf{c} \cdot \mathbf{b})\mathbf{a} + (\mathbf{a} \cdot \mathbf{c})\mathbf{b}.$$

Ejemplo 5.2.2. Dados los vectores $\mathbf{a} = 3\mathbf{i} + 2\mathbf{j} - \mathbf{k}$, $\mathbf{b} = \mathbf{i} - \mathbf{j} + \mathbf{k}$ y $\mathbf{c} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$, determinar su triple producto vectorial.
Solución
Aplicando la definición tenemos:

$$\mathbf{b} \times \mathbf{c} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -1 & 1 \\ 2 & -3 & 1 \end{vmatrix} = \mathbf{2i} + \mathbf{j} - \mathbf{k}$$

y, por lo tanto,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{vmatrix} = -\mathbf{i} + \mathbf{j} - \mathbf{k}.$$

El mismo cálculo, pero usando la relación de Gibbs, se reduce a

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} = (-1)(\mathbf{i}-\mathbf{j}+\mathbf{k}) - (0)(3\mathbf{i}+2\mathbf{j}-\mathbf{k}),$$

con lo cual

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (-1)(\mathbf{i}-\mathbf{j}+\mathbf{k}) = -\mathbf{i}+\mathbf{j}-\mathbf{k},$$

como se había calculado previamente.

5.3 Aplicaciones del producto vectorial

En esta sección presentamos algunos problemas que pueden resolverse fácilmente usando el concepto de *producto vectorial*. Algunos de estos problemas ya han sido tratados mediante otros métodos, pero sirven para ilustrar un nuevo enfoque en cuanto a su forma de solución.

5.3.1 Ecuación del plano

En primer lugar consideremos el problema de la determinación de la ecuación del plano que pasa por un punto y es paralelo a dos vectores dados, como se ve en la Figura 5.6.

Sea $\Pi(A,\mathbf{u},\mathbf{v})$ el plano que pasa por el punto A tal que $\mathbf{A} = (\alpha_1,\alpha_2,\alpha_3)$ y es paralelo a los vectores $\mathbf{u}$ y $\mathbf{v}$ (no paralelos entre sí) cuyas representaciones son

$$\mathbf{u} = u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k} \;\text{ y }\; \mathbf{v} = v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k}.$$

Un punto X del espacio tal que $\vec{X} = (x_1,x_2,x_3)$ está sobre el plano $\Pi(A,\mathbf{u},\mathbf{v})$ si y solo si

$$\textbf{\textsc{ax}}.\mathbf{n} = 0,$$

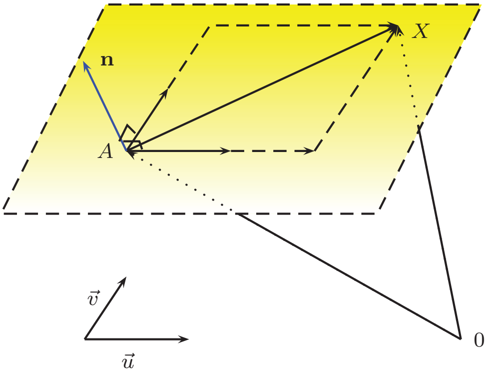


Figura 5.6. Plano que pasa por un punto A y es paralelo a dos vectores $\mathbf{u}$ y $\mathbf{v}$

siendo $\mathbf{n}$ el vector normal al plano y el vector $\textsc{ax}$ dado por la relación

$$\textsc{ax} = (x_1 - \alpha_1)\mathbf{i} + (x_2 - \alpha_2)\mathbf{j} + (x_3 - \alpha_3)\mathbf{k}.$$

Como $\mathbf{n}$ es el vector normal a los vectores $\mathbf{u}$ y $\mathbf{v}$, entonces,

$$\mathbf{n} = \mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

Al desarrollar este último determinante obtenemos la relación

$$\mathbf{n} = \mathbf{u} \times \mathbf{v} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \mathbf{k}$$

y, en consecuencia, la ecuación del plano toma la forma

$$\textsc{ax}\cdot\mathbf{n} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} (x_1 - \alpha_1) - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} (x_2 - \alpha_2) + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} (x_3 - \alpha_3) = 0,$$

lo cual es equivalente a la relación

$$\begin{vmatrix} (x_1 - \alpha_1) & (x_2 - \alpha_2) & (x_3 - \alpha_3) \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0.$$

Ejemplo 5.3.1. Hallar la ecuación del plano que pasa por el punto $A(2,1,0)$ y es paralelo a los vectores $\mathbf{u} = 3\mathbf{i} - \mathbf{j} + \mathbf{k}$ y $\mathbf{v} = \mathbf{i} - \mathbf{j} - 2\mathbf{k}$.

Solución

Sea $P(x_1, x_2, x_3)$ un punto sobre dicho plano; la ecuación pedida corresponde a la simplificación de la expresión

$$\begin{vmatrix} (x_1-2) & (x_2-1) & x_3 \\ 3 & -1 & 1 \\ 1 & -1 & -2 \end{vmatrix} = 0.$$

Resolviendo el determinante, la ecuación se reduce a

$$3(x_1-2) + 7(x_2-1) - 2x_3 = 0$$

o, simplemente, $3x_1 + 7x_2 - 2x_3 = 13$.

A partir del resultado anterior, es posible resolver el problema de determinar la ecuación del plano que pasa por tres puntos dados. En este caso, se toma el primer punto como aquel por el cual pasa el plano y los vectores paralelos al plano, corresponden a los vectores trazados desde el primer punto hasta los otros dos; es decir, si los puntos dados son $A(\alpha_1, \alpha_2, \alpha_3)$, $B(\beta_1, \beta_2, \beta_3)$ y $C(\gamma_1, \gamma_2, \gamma_3)$, entonces los vectores paralelos al plano son:

$$\begin{aligned} \mathbf{u} &= (\beta_1-\alpha_1)\mathbf{i} + (\beta_2-\alpha_2)\mathbf{j} + (\beta_3-\alpha_3)\mathbf{k} \\ \mathbf{v} &= (\gamma_1-\alpha_1)\mathbf{i} + (\gamma_2-\alpha_2)\mathbf{j} + (\gamma_3-\alpha_3)\mathbf{k} \end{aligned}$$

y, en consecuencia, la ecuación del plano que pasa por estos tres puntos toma la forma:

$$\begin{vmatrix} x_1-\alpha_1 & x_2-\alpha_2 & x_3-\alpha_3 \\ \beta_1-\alpha_1 & \beta_2-\alpha_2 & \beta_3-\alpha_3 \\ \gamma_1-\alpha_1 & \gamma_2-\alpha_2 & \gamma_3-\alpha_3 \end{vmatrix} = 0.$$

Ejemplo 5.3.2. Determinar la ecuación del plano que pasa por los puntos $A(1,2,3)$, $B(0,1,4)$ y $C(2,3,5)$.
Solución
En este caso, los vectores paralelos al plano son:

$$\begin{aligned}\mathbf{u} &= -\mathbf{i}-\mathbf{j}+\mathbf{k}\\ \mathbf{v} &= \mathbf{i}+\mathbf{j}+2\mathbf{k},\end{aligned}$$

lo cual nos conduce a la ecuación

$$\begin{vmatrix} x_1-1 & x_2-2 & x_3-3 \\ -1 & -1 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 0.$$

Resolviendo el determinante, la ecuación se reduce a

$$-3(x_1-1)+3(x_2-2)=0$$

que, luego de simplificar, corresponde a la ecuación $x_1-x_2+1=0$.

5.3.2 Distancia entre dos rectas

Consideremos ahora el problema de determinar la distancia entre dos rectas que se cruzan en el espacio. Sean las rectas L_1 y L_2, con ecuaciones paramétricas dadas por

$$\begin{aligned} x_1 &= \alpha_1 t+\beta_1 & x_1' &= \gamma_1 s+\delta_1\\ x_2 &= \alpha_2 t+\beta_2 & x_2' &= \gamma_2 s+\delta_2\\ x_2 &= \alpha_2 t+\beta_3 & x_3' &= \gamma_3 s+\delta_3. \end{aligned}$$

Un vector que une un punto cualquiera de la primera recta, con un punto de la segunda recta, está dado por:

$$\mathbf{u}=(x_1-x_1')\mathbf{i}+(x_2-x_2')\mathbf{j}+(x_3-x_3')\mathbf{k}.$$

Si utilizamos las ecuaciones paramétricas de ambas rectas, el vector $\mathbf{u}$ puede escribirse en la forma

$$\mathbf{u} = t\mathbf{a} - s\mathbf{c} + \mathbf{d},$$

donde

$$\begin{aligned}
\mathbf{a} &= \alpha_1\mathbf{i} + \alpha_2\mathbf{j} + \alpha_3\mathbf{k} \\
\mathbf{c} &= \gamma_1\mathbf{i} + \gamma_2\mathbf{j} + \gamma_3\mathbf{k} \\
\mathbf{d} &= (\beta_1 - \delta_1)\mathbf{i} + (\beta_2 - \delta_2)\mathbf{j} + (\beta_3 - \delta_3)\mathbf{k}.
\end{aligned}$$

El cuadrado de la distancia entre estos dos puntos, que denotamos mediante L, es una función de t y s, y puede calcularse como el producto escalar de $\mathbf{u}$ consigo mismo, es decir,

$$L(t, s) = \mathbf{u} \cdot \mathbf{u}.$$

El problema que nos interesa es el de determinar los valores de t y s que minimizan la función $L(t, s)$. Para establecer estos valores, lo que se debe hacer es calcular las soluciones del sistema de ecuaciones

$$\begin{aligned}
L_t &= 0 \\
L_s &= 0.
\end{aligned}$$

Usando la expresión para el vector $\mathbf{u}$, la distancia entre puntos sobre cada recta toma la forma

$$\begin{aligned}
L(t, s) &= (t\mathbf{a} - s\mathbf{c} + \mathbf{d}) \cdot (t\mathbf{a} - s\mathbf{c} + \mathbf{d}) \\
&= \|\mathbf{a}\|^2 t^2 + \|\mathbf{c}\|^2 s^2 - 2ts\,\mathbf{a} \cdot \mathbf{c} + 2t\,\mathbf{a} \cdot \mathbf{d} - 2s\,\mathbf{c} \cdot \mathbf{d} + \|\mathbf{d}\|^2,
\end{aligned}$$

con lo cual

$$\begin{aligned}
L_t &= 2\|\mathbf{a}\|^2 t - 2s\,\mathbf{a} \cdot \mathbf{c} + 2\,\mathbf{a} \cdot \mathbf{d} \\
L_s &= 2\|\mathbf{c}\|^2 s - 2t\,\mathbf{a} \cdot \mathbf{c} - 2\,\mathbf{c} \cdot \mathbf{d}
\end{aligned}$$

y, en consecuencia, el sistema pedido es:

$$\begin{aligned} \|\mathbf{a}\|^2 t - \mathbf{a}\cdot\mathbf{c}\, s &= -\mathbf{a}\cdot\mathbf{d} \\ \mathbf{a}\cdot\mathbf{c}\, t - \|\mathbf{c}\|^2 s &= -\mathbf{c}\cdot\mathbf{d}, \end{aligned}$$

cuyas soluciones son

$$t = \frac{\begin{vmatrix} -\mathbf{a}\cdot\mathbf{d} & -\mathbf{a}\cdot\mathbf{c} \\ -\mathbf{c}\cdot\mathbf{d} & -\|\mathbf{c}\|^2 \end{vmatrix}}{\begin{vmatrix} \|\mathbf{a}\|^2 & -\mathbf{a}\cdot\mathbf{c} \\ \mathbf{a}\cdot\mathbf{c} & -\|\mathbf{c}\|^2 \end{vmatrix}} = \frac{(\mathbf{a}\cdot\mathbf{c})(\mathbf{c}\cdot\mathbf{d}) - \mathbf{a}\cdot\mathbf{d}\|\mathbf{c}\|^2}{\|\mathbf{a}\|^2\|\mathbf{c}\|^2 - (\mathbf{a}\cdot\mathbf{c})^2}$$

$$s = \frac{\begin{vmatrix} \|\mathbf{a}\|^2 & -\mathbf{a}\cdot\mathbf{d} \\ \mathbf{a}\cdot\mathbf{c} & -\mathbf{c}\cdot\mathbf{d} \end{vmatrix}}{\begin{vmatrix} \|\mathbf{a}\|^2 & -\mathbf{a}\cdot\mathbf{c} \\ \mathbf{a}\cdot\mathbf{c} & -\|\mathbf{c}\|^2 \end{vmatrix}} = \frac{\mathbf{c}\cdot\mathbf{d}\|\mathbf{a}\|^2 - (\mathbf{a}\cdot\mathbf{c})(\mathbf{a}\cdot\mathbf{d})}{\|\mathbf{a}\|^2\|\mathbf{c}\|^2 - (\mathbf{a}\cdot\mathbf{c})^2}.$$

Además, tenemos que:

$$\begin{aligned} L_{tt} &= 2\|\mathbf{a}\|^2 \\ L_{ss} &= 2\|\mathbf{c}\|^2 \\ L_{ts} &= -2\mathbf{a}\cdot\mathbf{c} \end{aligned}$$

y, por lo tanto, el hessiano de la función de distancia es:

$$L_{tt}L_{ss} - (L_{ts})^2 = 4(\|\mathbf{a}\|^2\|\mathbf{c}\|^2 - \mathbf{a}\cdot\mathbf{c});$$

pero por la desigualdad de Schwartz, $\|\mathbf{a}\|\|\mathbf{c}\| \geq \mathbf{a}\cdot\mathbf{c}$, con lo cual el hessiano de $L(t,s)$ es siempre positivo y, en consecuencia, la función toma el valor mínimo en los valores críticos calculados.

De esta forma, para calcular la distancia mínima, reemplazamos los valores críticos hallados en cualquiera de las expresiones para $L(t,s)$. Antes de hacer estos reemplazos, si recordamos la definición del producto vectorial, obtenemos la siguiente relación:

$$\|\mathbf{a}\times\mathbf{c}\|^2 = \|\mathbf{a}\|^2\|\mathbf{c}\|^2 \operatorname{sen}^2\theta = \|\mathbf{a}\|^2\|\mathbf{c}\|^2(1-\cos^2\theta) = \|\mathbf{a}\|^2\|\mathbf{c}\|^2-(\mathbf{a}\cdot\mathbf{c})^2,$$

con lo cual los valores críticos pueden escribirse como

$$t = \frac{(\mathbf{a}\cdot\mathbf{c})(\mathbf{c}\cdot\mathbf{d}) - \mathbf{a}\cdot\mathbf{d}\|\mathbf{c}\|^2}{\|\mathbf{a}\times\mathbf{c}\|^2}$$

$$s = \frac{\mathbf{c}\cdot\mathbf{d}\|\mathbf{a}\|^2 - (\mathbf{a}\cdot\mathbf{c})(\mathbf{a}\cdot\mathbf{d})}{\|\mathbf{a}\times\mathbf{c}\|^2}.$$

Luego del proceso de simplificación, se llega a que la expresión para el valor mínimo de la función $L(t,s)$ es:

$$L(t,s) = \frac{[\mathbf{d}\cdot(\mathbf{a}\times\mathbf{c})]^2}{\|\mathbf{a}\times\mathbf{c}\|^2},$$

con lo cual la distancia entre las dos rectas se reduce a la fórmula

$$d = \frac{|\mathbf{d}\cdot(\mathbf{a}\times\mathbf{c})|}{\|\mathbf{a}\times\mathbf{c}\|}. \tag{5.3.1}$$

Una forma geométrica y fácil de resolver este problema, que se ilustra en la Figura 5.7, consiste en determinar un vector que sea perpendicular a las dos rectas dadas. Esto se logra realizando el producto vectorial entre los vectores paralelos a cada una de las rectas del problema y, una vez establecido este vector, calcular la proyección escalar de un vector arbitrario con extremos en cada una de las rectas.

Si utilizamos las ecuaciones paramétricas de ambas rectas, los vectores $\mathbf{a}$, $\mathbf{c}$, paralelos a las rectas, y el vector $\mathbf{d} = \mathbf{\scriptstyle PQ}$, que une los puntos P y Q sobre cada recta, pueden escribirse en la forma

$$\begin{aligned}
\mathbf{a} &= \alpha_1\mathbf{i} + \alpha_2\mathbf{j} + \alpha_3\mathbf{k}\\
\mathbf{c} &= \gamma_1\mathbf{i} + \gamma_2\mathbf{j} + \gamma_3\mathbf{k}\\
\mathbf{d} &= (\beta_1-\delta_1)\mathbf{i} + (\beta_2-\delta_2)\mathbf{j} + (\beta_3-\delta_3)\mathbf{k}.
\end{aligned}$$

El vector normal a las dos rectas, $\mathbf{n}$, es:

$$\mathbf{n} = \mathbf{a}\times\mathbf{c},$$

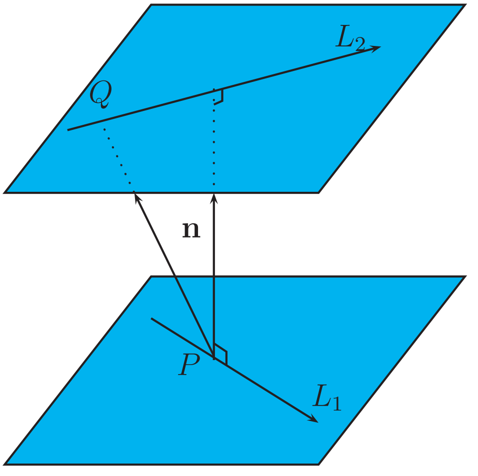


Figura 5.7. Distancia entre rectas cruzadas en el espacio

y la proyección de $\mathbf{d}$ sobre $\mathbf{n}$ es:

$$d = \|\mathbf{d}\| | \cos \phi|,$$

siendo ϕ el ángulo formado por los vectores $\mathbf{n}$ y $\mathbf{d}$. Según la definición del producto escalar,

$$d = \|\mathbf{d}\| |\cos \phi| = \|\mathbf{d}\| \left| \frac{\mathbf{d} \cdot \mathbf{n}}{\|\mathbf{d}\| \|\mathbf{n}\|} \right| = \frac{|\mathbf{d} \cdot \mathbf{n}|}{\|\mathbf{n}\|} = \frac{|\mathbf{d} \cdot (\mathbf{a} \times \mathbf{c})|}{\|\mathbf{a} \times \mathbf{c}\|}$$

como se había indicado en la ecuación 5.3.1.

Ejemplo 5.3.3. Hallar la distancia entre las rectas con ecuaciones:

$$\begin{aligned} x_1 &= 2 + 3t \qquad & x_1' &= 1 + s \\ x_2 &= 1 - t \qquad & x_2' &= 1 - s \\ x_2 &= 2t \qquad & x_3' &= 3 + 4s. \end{aligned}$$

En este caso, los vectores $\mathbf{a}$ y $\mathbf{c}$ paralelos a las rectas y el vector

PQ que une un punto de L_1 con un punto de L_2 están dados por

$$\begin{aligned}\mathbf{a} &= 3\mathbf{i}-\mathbf{j}+2\mathbf{k}\\ \mathbf{c} &= \mathbf{i}-\mathbf{j}+4\mathbf{k}\\ \mathrm{PQ}=\mathbf{d} &= -\mathbf{i}+3\mathbf{k},\end{aligned}$$

con lo que el vector normal a las dos rectas es

$$\mathbf{a}\times\mathbf{c}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 3 & -1 & 2\\ 1 & -1 & 4\end{vmatrix}=-2\mathbf{i}-10\mathbf{j}-2\mathbf{k}$$

y, en consecuencia, la distancia entre las rectas está dada por

$$d=\frac{|\mathbf{d}\cdot(\mathbf{a}\times\mathbf{c})|}{\|\mathbf{a}\times\mathbf{c}\|}=\frac{|(-\mathbf{i}+3\mathbf{k})\cdot(-2\mathbf{i}-10\mathbf{j}-2\mathbf{k})|}{\|-2\mathbf{i}-10\mathbf{j}-2\mathbf{k}\|}=\frac{4}{\sqrt{108}}=\frac{2}{3\sqrt{3}}.$$

Si utilizáramos el otro procedimiento, el problema que se presenta es el optimizar la función

$$L(t,s)=14t^2+18s^2-24ts+6t-22s+10,$$

con lo cual los valores críticos pueden escribirse como

$$t=\frac{(\mathbf{a}\cdot\mathbf{c})(\mathbf{c}\cdot\mathbf{d})-\mathbf{a}\cdot\mathbf{d}\|\mathbf{c}\|^2}{\|\mathbf{a}\times\mathbf{c}\|^2}=\frac{(12)(11)-3(18)}{108}=\frac{13}{18}$$

$$s=\frac{\mathbf{c}\cdot\mathbf{d}\|\mathbf{a}\|^2-(\mathbf{a}\cdot\mathbf{c})(\mathbf{a}\cdot\mathbf{d})}{\|\mathbf{a}\times\mathbf{c}\|^2}=\frac{(11)(14)-12(3)}{108}=\frac{59}{54},$$

que corresponden a la solución del sistema de ecuaciones

$$\begin{aligned}L_t &= 28t-24s+6=0\\ L_s &= 24t-36s+22=0.\end{aligned}$$

Con ello, obtenemos que:

$$L\left(\frac{13}{18},\frac{59}{54}\right)=\frac{4}{27}$$

y, así, la distancia es $2/(3\sqrt{3})$, como se había calculado al inicio de este ejemplo.

5.3.3 Área de un triángulo

Consideremos el triángulo determinado por los puntos del espacio $A(\alpha_1, \alpha_2, \alpha_2)$, $B(\beta_1, \beta_2, \beta_2)$ y $C(\gamma_1, \gamma_2, \gamma_2)$ con coordenadas referidas a una base ortonormal (véase la Figura 5.8).

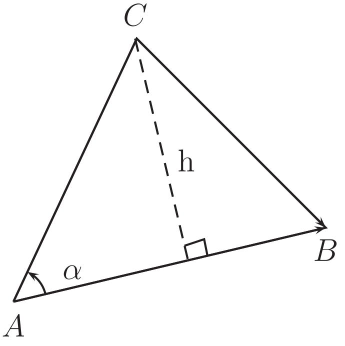


Figura 5.8. Área de un triángulo

De la Figura 5.8 puede concluirse que el área a está dada por

$$a = \frac{1}{2}\|\mathbf{AB}\|h;$$

pero si recordamos la definición del producto vectorial:

$$\|\mathbf{AB} \times \mathbf{AC}\| = \|\mathbf{AB}\|\|\mathbf{AC}\| \operatorname{sen} \alpha \quad \text{y, además,} \quad h = \|\mathbf{AC}\| \operatorname{sen} \alpha,$$

entonces

$$a = \frac{1}{2}\|\mathbf{AB} \times \mathbf{AC}\|.$$

Si contamos con dos vectores concurrentes $\mathbf{AB}$ y $\mathbf{AC}$, el paralelogramo determinado por ellos dos tiene por área el doble de la del triángulo determinado por los extremos de los dos vectores. Por lo tanto, si llamamos a_p al área del paralelogramo, concluimos que:

$$a_p = \|\mathbf{AB} \times \ \mathbf{AC}\|.$$

Ejemplo 5.3.4. Hallar el área del triángulo determinado por los puntos $A(2, 1, 3)$, $B(1, 0, -1)$ y $C(1, 1, 0)$ con coordenadas referidas a una base ortonormal.

Solución
En este caso, los vectores que corresponden a dos de los lados del triángulo son:

$$\begin{aligned} \textsc{ab} &= -\mathbf{i} - \mathbf{j} - 4\mathbf{k} \\ \textsc{ac} &= -\mathbf{i} - 3\mathbf{k} \end{aligned}$$

y, en consecuencia, su producto vectorial está dado por

$$\|\textsc{ab} \times \ \textsc{ac}\| = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & -1 & -4 \\ 1 & 0 & -3 \end{vmatrix} = 3\mathbf{i} + \mathbf{j} - \mathbf{k},$$

con lo cual el área del triángulo es:

$$a = \frac{1}{2}\|3\mathbf{i} + \mathbf{j} - \mathbf{k}\| = \frac{\sqrt{11}}{2}.$$

Este resultado permite calcular, indirectamente, la longitud de la altura relativa al lado AB del triángulo.

Si se tiene un grupo de puntos sobre un plano, se puede aplicar el procedimiento anterior para establecer el área del polígono determinado por estos puntos, siempre y cuando el polígono sea convexo, es decir, si y solamente si cualquier línea que contiene un lado de este deja a todo el polígono completamente en uno de los semiplanos definidos por la recta, como se ve en la Figura 5.9.

De esta manera, si tenemos n puntos que constituyen una poligonal cerrada y convexa, al trazar los vectores $\mathbf{A_1A_j}$, con $j = 2, 3, \ldots, n$, se forman $n - 2$ triángulos con el punto A_1 como vértice común. Sin pérdida de generalidad, podemos suponer que los puntos están sobre el plano xy y que las coordenadas del punto A_j son $(\alpha_j, \beta_j, 0)$, de tal modo que el área del primer triángulo, a_1, esta dada por la expresión:

$$a_1 = \frac{1}{2}\|\mathbf{A_1A_2} \times \mathbf{A_1A_3}\|.$$

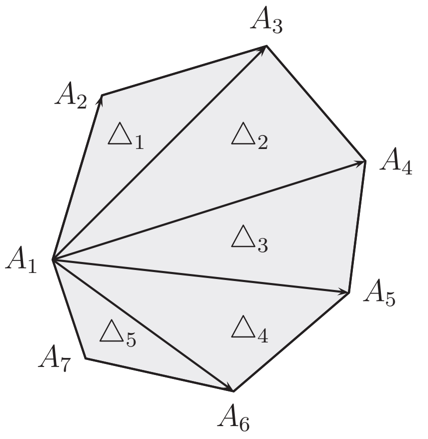


Figura 5.9. Área de una poligonal convexa

Al ser

$$\mathbf{A_1A_2} \times \mathbf{A_1A_3} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ (\alpha_2 - \alpha_1) & (\beta_2 - \beta_1) & 0 \\ (\alpha_3 - \alpha_1) & (\beta_3 - \beta_1) & 0 \end{vmatrix},$$

calculando este determinante, obtenemos la expresión:

$$\mathbf{A_1A_2} \times \mathbf{A_1A_3} = [(\alpha_2 - \alpha_1)(\beta_3 - \beta_1) - (\alpha_3 - \alpha_1)(\beta_2 - \beta_1)]\mathbf{k},$$

de donde se concluye que

$$a_1 = \frac{1}{2}|(\alpha_2 - \alpha_1)(\beta_3 - \beta_1) - (\alpha_3 - \alpha_1)(\beta_2 - \beta_1)|.$$

Una simplificación adicional se obtiene si el punto A_1 se toma como el origen de coordenadas, caso en el cual la expresión para el área del primer triángulo se reduce a

$$a_1 = \frac{1}{2}|\alpha_2\beta_3 - \alpha_3\beta_2|.$$

Para el área del segundo triángulo, a_2, obtenemos

$$a_2 = \frac{1}{2}\|\mathbf{A_1A_3} \times \mathbf{A_1A_4}\|.$$

Al ser

$$\mathbf{A_1A_3} \times \mathbf{A_1A_4} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha_3 & \beta_3 & 0 \\ \alpha_4 & \beta_4 & 0 \end{vmatrix} = (\alpha_3\beta_4 - \alpha_4\beta_3)\mathbf{k},$$

se concluye que

$$a_2 = \frac{1}{2}|\alpha_3\beta_4 - \alpha_4\beta_3|.$$

Continuando de esta forma, para el área del último triángulo, a_{n-2}, obtenemos

$$a_{n-2} = \frac{1}{2}\|\mathbf{A_1A_{n-1}} \times \mathbf{A_1A_n}\|.$$

Al ser

$$\mathbf{A_1A_{n-1}} \times \mathbf{A_1A_n} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha_{n-1} & \beta_{n-1} & 0 \\ \alpha_n & \beta_n & 0 \end{vmatrix},$$

el valor de este determinante es

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha_{n-1} & \beta_{n-1} & 0 \\ \alpha_n & \beta_n & 0 \end{vmatrix} = (\alpha_{n-1}\beta_n - \alpha_n\beta_{n-1})\mathbf{k},$$

de lo cual se concluye que

$$a_{n-2} = \frac{1}{2}|\alpha_{n-1}\beta_n - \alpha_n\beta_{n-1}|.$$

La anterior información puede recogerse fácilmente como se muestra en la Tabla 5.1.

El área total de la poligonal está dada por:

$$A = \frac{1}{2}\sum_{j=2}^{n-1}|\alpha_j\beta_{j+1} - \alpha_{j+1}\beta_j|,$$

que corresponde a la semisuma de los elementos de la columna 3. En el siguiente ejemplo ilustramos cómo aplicar este procedimiento.

Coord. x	**Coord. y**	**Área j -paralelogramo**
0	0	
α_2	β_2	
α_3	β_3	$\lvert\alpha_2\beta_3-\alpha_3\beta_2\rvert$
α_4	β_4	$\lvert\alpha_3\beta_4-\alpha_4\beta_4\rvert$
$\vdots$	$\vdots$	$\vdots$
α_{n-1}	β_{n-1}	$\vdots$
α_n	β_n	$\lvert\alpha_{n-1}\beta_n-\alpha_n\beta_{n-1}\rvert$
		$A=\dfrac{1}{2}\sum_{j=2}^{n-1}\lvert\alpha_j\beta_{j+1}-\alpha_{j+1}\beta_j\rvert$

Tabla 5.1. Área de una poligonal convexa

Ejemplo 5.3.5. Hallar el área de la región determinada por el conjunto de puntos que se muestra en la Figura 5.10.

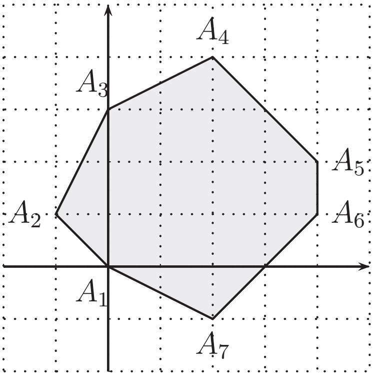


Figura 5.10. Área del polígono convexo del ejercicio 5.3.5

Solución
Las coordenadas de los puntos y las operaciones descritas previamente se presentan en la Tabla 5.2, que no es más que una reproducción numérica de la Tabla 5.1.

Coord. x	**Coord.** y	**Área** j**-paralelogramo**
0	0	
-1	1	
0	3	3
2	4	6
4	2	12
4	1	4
2	-1	6
		31

Tabla 5.2. Área del polígono convexo del ejercicio 5.3.5

El área dentro de la poligonal tiene por valor 31/2 unidades.

5.4 Triple producto escalar

En esta sección presentamos algunas aplicaciones que se obtienen al combinar los diferentes productos entre vectores. Dado que el producto vectorial entre dos vectores es nuevamente un vector, es posible realizar con él alguno de los productos previamente definidos. En realidad, esto ya lo hemos realizado cuando consideramos el problema de la distancia entre dos rectas que se cruzan en el espacio. En esta sección vamos a considerar este producto, conocido como el *triple producto escalar*, al igual que algunas de sus propiedades y aplicaciones.

Definición 5.4.1. *Sean los vectores* $\mathbf{a}$, $\mathbf{b}$ *y* $\mathbf{c}$ *de* $\mathbb{E}^3$ *referidos a la base canónica. El triple producto escalar, también conocido como producto mixto, denotado por* $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$, *se define por*

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}).$$

En la práctica, no es necesario el uso del paréntesis, ya que es obvio que la agrupación $(\mathbf{a} \cdot \mathbf{b}) \times \mathbf{c}$ carece de sentido, puesto que $(\mathbf{a} \cdot \mathbf{b})$ es un escalar. Según la definición, dados los vectores

$$\mathbf{a} = \alpha_1\mathbf{i} + \alpha_2\mathbf{j} + \alpha_3\mathbf{k},\ \ \mathbf{b} = \beta_1\mathbf{i} + \beta_2\mathbf{j} + \beta_3\mathbf{k}\ \text{ y }\ \mathbf{c} = \gamma_1\mathbf{i} + \gamma_2\mathbf{j} + \gamma_3\mathbf{k},$$

entonces, al ser

$$\mathbf{b} \times \mathbf{c} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} = \mathbf{i}\begin{vmatrix} \beta_2 & \beta_3 \\ \gamma_2 & \gamma_3 \end{vmatrix} - \mathbf{j}\begin{vmatrix} \beta_1 & \beta_3 \\ \gamma_1 & \gamma_3 \end{vmatrix} + \mathbf{k}\begin{vmatrix} \beta_1 & \beta_2 \\ \gamma_1 & \gamma_2 \end{vmatrix},$$

se concluye que

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \alpha_1\begin{vmatrix} \beta_2 & \beta_3 \\ \gamma_2 & \gamma_3 \end{vmatrix} - \alpha_2\begin{vmatrix} \beta_1 & \beta_3 \\ \gamma_1 & \gamma_3 \end{vmatrix} + \alpha_3\begin{vmatrix} \beta_1 & \beta_2 \\ \gamma_1 & \gamma_2 \end{vmatrix},$$

con lo cual, teniendo en cuenta la definición de determinante, resulta

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix}.$$

Ejemplo 5.4.1. Hallar el producto triple escalar de los siguientes vectores: $\mathbf{a} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$, $\mathbf{b} = \mathbf{i} - 5\mathbf{j} + 2\mathbf{k}$ y $\mathbf{c} = \mathbf{i} + \mathbf{j} + \mathbf{k}$.

Solución

Teniendo en cuenta la fórmula expuesta, el triple producto escalar está dado por:

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \begin{vmatrix} 2 & -3 & 1 \\ 1 & -5 & 2 \\ 1 & 1 & 1 \end{vmatrix}.$$

El cálculo de este último determinante lo podemos realizar con la ayuda la Figura 5.11.

Con ello, obtenemos que

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = (-10 - 6 + 1) - (-5 + 4 - 3) = (-15) - (-4) = -11.$$

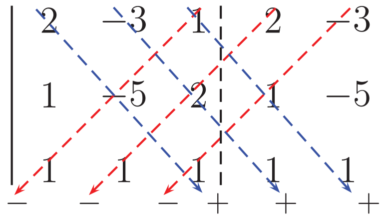


Figura 5.11. Determinante de un matriz 3×3

Esto puede verificarse al calcular el determinante por los cofactores de la primera fila.

Una interpretación geométrica del triple producto escalar es el cálculo del volumen de un paralelepípedo, como puede verse en la Figura 5.12. Para probar esto, veamos las siguientes propiedades:

- Al calcular el producto $\mathbf{b} \times \mathbf{c}$, el módulo de este producto representa el área del paralelogramo determinado por los vectores $\mathbf{b}$ y $\mathbf{c}$, mientras que $\mathbf{b} \times \mathbf{c}$ es un vector normal al plano determinado por los vectores $\mathbf{b}$ y $\mathbf{c}$.

- Al considerar el paralelepípedo determinado por los tres vectores, la altura respecto al plano determinado por los vectores $\mathbf{b}$ y $\mathbf{c}$ puede calcularse mediante la relación

$$h = \|\mathbf{a}\| \cos \alpha, \quad 0 \leq \alpha \leq \frac{\pi}{2},$$

siendo α el ángulo entre los vectores $\mathbf{a}$ y $\mathbf{b} \times \mathbf{c}$. En consecuencia,

$$h = \|\mathbf{a}\| \frac{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}{\|\mathbf{a}\| \|\mathbf{b} \times \mathbf{c}\|} = \frac{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}{\|\mathbf{b} \times \mathbf{c}\|}.$$

De está forma, el volumen del paralelepípedo es

$$Vol = \|\mathbf{b} \times \mathbf{c}\| \frac{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}{\|\mathbf{b} \times \mathbf{c}\|} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}).$$

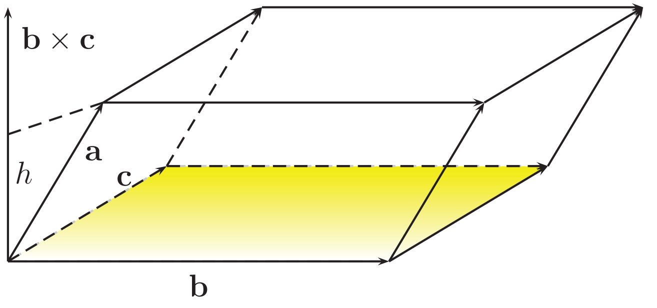


Figura 5.12. Volumen del paralelepípedo

En el caso propuesto, el ángulo α es agudo; de no serlo, el volumen del paralelepípedo será el valor absoluto del triple producto escalar.

Ejemplo 5.4.2. Hallar el volumen del paralelepípedo determinado por los vectores $\mathbf{a} = 3\mathbf{i} - 2\mathbf{j} + 2\mathbf{k}$, $\mathbf{b} = 4\mathbf{i} - \mathbf{j} - \mathbf{k}$ y $\mathbf{c} = \mathbf{i} - \mathbf{j} + \mathbf{k}$.
Solución
En este caso, el calculo del determinante es:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} 3 & -2 & 2 \\ 4 & -1 & -1 \\ 1 & -1 & 1 \end{vmatrix} = 3(-2) + 2(5) + 2(-3) = -2,$$

con lo cual $Vol = |(\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})| = 2$.

En el caso de vectores concurrentes, puede usarse el triple producto escalar como un criterio de coplanariedad, ya que si los tres vectores están en el mismo plano, el producto vectorial de dos de ellos será perpendicular al tercero y, en consecuencia, su producto escalar será nulo (obviamente el volumen es nulo), con lo cual puede afirmarse que tres vectores concurrentes son coplanarios si y solo si

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = 0.$$

Con este criterio es fácil reescribir el problema de la ecuación del plano que pasa por tres puntos dados. Si A, B y C son los puntos dados y X es un punto arbitrario, entonces X está sobre el plano determinado por estos puntos si y solo si

$$\mathbf{AX} \cdot (\mathbf{AB} \times \mathbf{AC}) = 0.$$

En términos de las coordenadas de los puntos, esta ecuación puede escribirse en la forma

$$\begin{vmatrix} x_1 - \alpha_1 & x_2 - \alpha_2 & x_3 - \alpha_3 \\ \beta_1 - \alpha_1 & \beta_2 - \alpha_2 & \beta_3 - \alpha_3 \\ \gamma_1 - \alpha_1 & \gamma_2 - \alpha_2 & \gamma_3 - \alpha_3 \end{vmatrix} = 0.$$

Ejemplo 5.4.3. Hallar la ecuación del plano determinado por los puntos $A(3, 0, 0)$, $B(0, 4, 5)$ y $C(0, 1, 1)$.
Solución
Al emplear la ecuación anterior, la ecuación del plano es

$$\begin{vmatrix} x_1 - 3 & x_2 & x_3 \\ -3 & 4 & 5 \\ -3 & 1 & 1 \end{vmatrix} = 0.$$

Con el desarrollo del determinante obtenemos

$$-(x_1 - 3) - 12x_2 + 9x_3 = 0 \quad \text{o} \quad x_1 + 12x_2 - 9x_3 = 3.$$

Según la definición del triple producto escalar, hay diversas maneras de ordenar los vectores. La siguiente propiedad ilustra cómo cambia el valor de dicho producto para las diferentes ordenaciones de los vectores.

Sean los vectores **a**, **b** y $\mathbf{c} \in \mathbb{E}^3$; entonces, se cumple que

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{c}, \mathbf{a}, \mathbf{b}] = [\mathbf{b}, \mathbf{c}, \mathbf{a}] = -[\mathbf{a}, \mathbf{c}, \mathbf{b}] = -[\mathbf{c}, \mathbf{b}, \mathbf{a}] = -[\mathbf{b}, \mathbf{a}, \mathbf{c}],$$

Las pruebas de estas igualdades se pueden realizar a partir de las definiciones de cada uno de los productos y comparando los términos que aparecen en cada una de las expansiones. Así, para las expresiones que aparecen en los extremos de las igualdades positivas, tenemos que

$$\begin{aligned}
[\mathbf{a}, \mathbf{b}, \mathbf{c}] &= \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} \\
&= \alpha_1\beta_2\gamma_3 - \alpha_1\beta_3\gamma_2 - \alpha_2\beta_1\gamma_3 + \alpha_2\gamma_1\beta_3 + \beta_1\alpha_3\gamma_2 - \alpha_3\beta_2\gamma_1,
\end{aligned}$$

mientras que

$$\begin{aligned}
[\mathbf{b}, \mathbf{c}, \mathbf{a}] &= \begin{vmatrix} \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{vmatrix} \\
&= \alpha_1\beta_2\gamma_3 - \alpha_1\beta_3\gamma_2 - \alpha_2\beta_1\gamma_3 + \alpha_2\gamma_1\beta_3 + \beta_1\alpha_3\gamma_2 - \alpha_3\beta_2\gamma_1,
\end{aligned}$$

lo que nos permite concluir que

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{b}, \mathbf{c}, \mathbf{a}].$$

Las demás igualdades se dejan como ejercicio.

5.5 Momento producido por una fuerza

En la mecánica del sólido rígido, es fundamental el *principio de transmisibilidad,* el cual establece que las condiciones de equilibrio o movimiento de un sólido rígido permanecerán invariantes si una fuerza $\mathbf{F}$ que actúa en un punto del sólido rígido se sustituye por una fuerza $\mathbf{F}'$ del mismo módulo y la misma dirección y sentido, pero que actúa en un punto diferente, siempre que las dos fuerzas tengan la misma línea de acción. Este tipo de vectores se conocen con el nombre de *vectores deslizantes*. Cuando las fuerzas tienen el mismo efecto sobre el sólido se dice que son *equivalentes*, como se observa en la Figura 5.13.

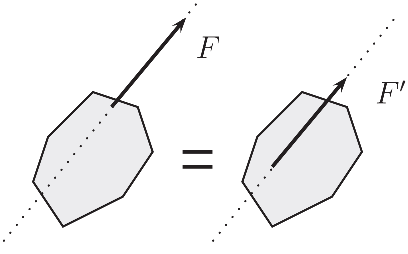


Figura 5.13. Fuerzas equivalentes

Definición 5.5.1. *El momento o torque* $\mathbf{M}$ *de un vector deslizante* $\mathbf{F}$*, respecto al origen de coordenadas, es* $\mathbf{M} = \mathbf{r} \times \mathbf{F}$, *donde* $\mathbf{r}$ *es el vector de posición del origen del vector* $\mathbf{F}$ *(véase la Figura 5.14).*

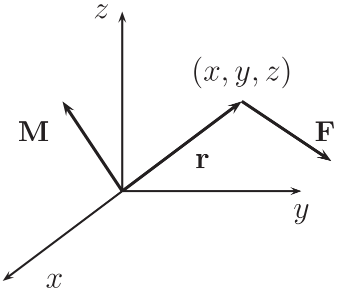


Figura 5.14. Momento de una fuerza

Ejemplo 5.5.1. Hallar el momento respecto al origen producido por una fuerza de 100 N aplicada en el punto A, como se muestra en la Figura 5.15.

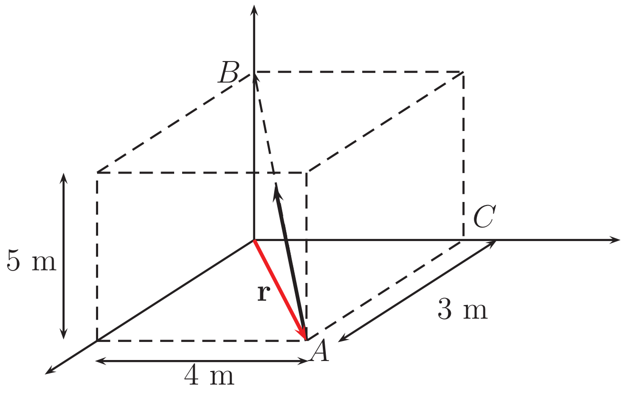


Figura 5.15. Momento de una fuerza respecto al origen

Solución

El vector del origen al punto de aplicación de la fuerza es $\mathbf{r} = 3\mathbf{i}+4\mathbf{j}$, mientras que el vector que representa a la fuerza es

$$\mathbf{F} = \frac{100}{\|\mathbf{r_B} - \mathbf{r_A}\|}(\mathbf{r_B} - \mathbf{r_A}) = \frac{100}{5\sqrt{2}}(-3\mathbf{i} - 4\mathbf{j} + 5\mathbf{k}),$$

con lo cual el momento queda determinado por

$$\mathbf{M} = \mathbf{r} \times \mathbf{F} = \frac{20}{\sqrt{2}} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 4 & 0 \\ -3 & -4 & 5 \end{vmatrix} = 50\sqrt{2}(4\mathbf{i} - 3\mathbf{j}).$$

El momento de una fuerza mide la tendencia de esta para hacer girar el cuerpo y, como lo indica la definición del producto vectorial, es perpendicular al plano determinado por la fuerza y el vector de posición. Si $\mathbf{U_f}$ es un vector unitario en la dirección de la fuerza, entonces

$$\mathbf{M} = \|\mathbf{F}\|\mathbf{r} \times \mathbf{U_f}.$$

Al desplazar la fuerza sobre la recta de soporte, el vector de posición del nuevo punto de aplicación $\mathbf{r_n}$ será $\mathbf{r_n} = \mathbf{r} + t\mathbf{U_f}$, donde

t es un escalar cualquiera. De este modo, el momento estará dado por

$$\mathbf{M} = \|\mathbf{F}\|\mathbf{r_n} \times \mathbf{U_f} = \|\mathbf{F}\|(\mathbf{r} + t\mathbf{U_f}) \times \mathbf{U_f}.$$

Al aplicar la distributividad del producto vectorial, la expresión anterior se transforma en

$$\mathbf{M} = \|\mathbf{F}\|\mathbf{r} \times \mathbf{U_f} + \|\mathbf{F}\|t\mathbf{U_f} \times \mathbf{U_f} = \|\mathbf{F}\|\mathbf{r} \times \mathbf{U_f},$$

puesto que $\mathbf{U_f} \times \mathbf{U_f} = 0$.

Este resultado indica que el momento producido por una fuerza no cambia cuando la fuerza se aplica en puntos distintos sobre su línea de acción.

Si el momento se quiere calcular con respecto a otro punto 0_1 diferente del origen, de acuerdo con la Figura 5.16, se tiene que

$$\mathbf{M_{0_1}} = \overrightarrow{0_1H} \times \mathbf{F} = (\mathbf{r} - \mathbf{r_1}) \times \mathbf{F} = \mathbf{r} \times \mathbf{F} - \mathbf{r_1} \times \mathbf{F}.$$

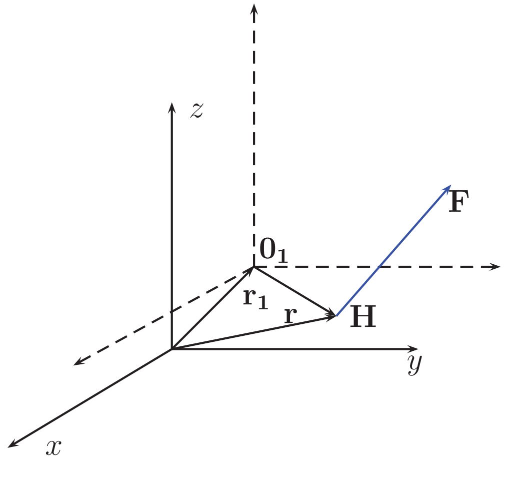


Figura 5.16. Momento respecto a un punto arbitrario

Esta última expresión significa que el momento $\mathbf{M_{0_1}}$ respecto a otro punto de referencia es el momento respecto al origen menos el momento de la fuerza $\mathbf{F}$ aplicada en el punto 0_1.

Ejemplo 5.5.2. Hallar el momento respecto al punto C producido por una fuerza de 100 N aplicada en el punto A, como se muestra en la figura 5.15.

Solución

Tal como hemos probado, para el momento con respecto al punto C tenemos

$$\mathbf{M_C} = \mathbf{r} \times \mathbf{F} - \mathbf{r_C} \times \mathbf{F} = \mathbf{M} - \mathbf{r_C} \times \mathbf{F}.$$

Usando los resultados del ejercicio precedente y teniendo en cuenta que $\mathbf{r_C} = 4\mathbf{j}$, obtenemos que

$$\mathbf{M_C} = 50\sqrt{2}(4\mathbf{i} - 3\mathbf{j}) - \frac{100}{\sqrt{50}} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 4 & 0 \\ -3 & -4 & 5 \end{vmatrix},$$

con lo cual

$$\mathbf{M_C} = 50\sqrt{2}(4\mathbf{i} - 3\mathbf{j}) - \frac{20}{\sqrt{2}}(20\mathbf{i} + 12\mathbf{k}) = -150\sqrt{2}\mathbf{j} - 120\sqrt{2}\mathbf{k};$$

es decir, el momento respecto al punto C está dado por la expresión

$$\mathbf{M_C} = -150\sqrt{2}\mathbf{j} - 120\sqrt{2}\mathbf{k}.$$

Una consecuencia inmediata de la distributividad del producto vectorial es el teorema conocido como *teorema de Varignon*, el cual proporciona un método práctico para el cálculo del momento generado por un conjunto de fuerzas cuyas líneas de acción concurren a un mismo punto del espacio. Este resultado corresponde al siguiente teorema.

Teorema 5.5.1. *Sean $\mathbf{F_1}$ y $\mathbf{F_2}$ dos vectores deslizantes cuyas líneas de acción concurren a un mismo punto A del espacio. El momento resultante de estos dos vectores, con respecto al origen, es la suma vectorial de los momentos correspondientes a cada uno de los vectores, además, es igual al momento producido por la resultante de las fuerzas.*

La prueba de este teorema se sigue de la relación

$$\mathbf{M_r} = \mathbf{M_1} + \mathbf{M_2} = \mathbf{r} \times \mathbf{F_1} + \mathbf{r} \times \mathbf{F_2} = \mathbf{r} \times (\mathbf{F_1} + \mathbf{F_2}) = \mathbf{r} \times \mathbf{F_r}$$

y se ilustra en la Figura 5.17.

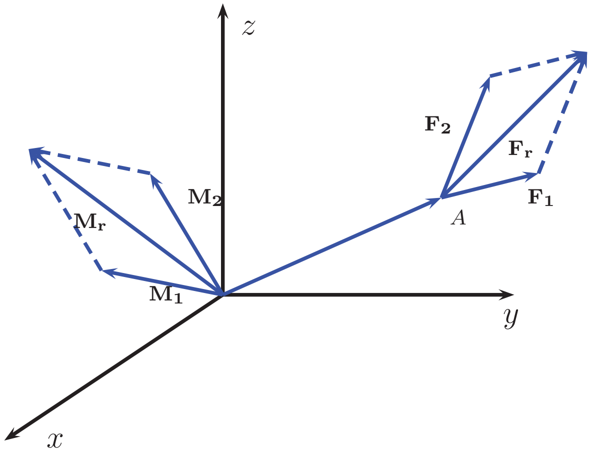


Figura 5.17. Teorema de Varignon

En el caso de varias fuerzas no concurrentes, su estudio requiere de la consideración de dos elementos:

- La resultante del sistema, que corresponde a la suma vectorial de todas las fuerzas que actúan en el sistema.

- El momento resultante, el cual es la suma vectorial de los momentos de todos los vectores (fuerzas) del sistema, con respecto a un mismo punto.

De esta forma, si consideramos el conjunto de fuerzas $\mathbf{F_1}, \mathbf{F_2}, \ldots, \mathbf{F_n}$, aplicadas en los puntos con vectores de posición dados por $\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_n}$, la resultante de las fuerzas corresponde a

$$\mathbf{F_r} = \sum_{i=1}^{n} \mathbf{F_i},$$

mientras que para los momentos, respecto al origen, el resultado es

$$\mathbf{M_0} = \sum_{i=1}^{n} \mathbf{M_i} = \sum_{i=1}^{n} \mathbf{r_i} \times \mathbf{F_i}.$$

En caso de que los momentos se calculen con respecto a otro punto Q, el momento resultante es

$$\mathbf{M_Q} = \sum_{i=1}^{n} \mathbf{M_{Q_i}} = \sum_{i=1}^{n} (\mathbf{r_i} - \mathbf{r_Q}) \times \mathbf{F_i} = \sum_{i=1}^{n} \mathbf{r_i} \times \mathbf{F_i} - \mathbf{r_Q} \times \sum_{i=1}^{n} \mathbf{F_i},$$

expresión que puede escribirse en la forma

$$\mathbf{M_Q} = \mathbf{M_0} - \mathbf{r_Q} \times \mathbf{F_r},$$

lo cual muestra que el momento resultante depende del punto con respecto al cual se calcula.

De la última relación se concluye que si la resultante del sistema que se está considerando es nula, es decir, si $\mathbf{F_r} = 0$, el momento es invariante. En otras palabras, para cualquier punto Q, se tiene la relación $\mathbf{M_Q} = \mathbf{M_0}$.

Una aplicación importante de este resultado es la determinación del momento resultante de un sistema de vectores opuestos, no coaxiales y de igual módulo, como se muestra en la Figura 5.18.

Dicho sistema recibe el nombre de *par* y, en este caso, para calcular el momento, tenemos que:

$$\mathbf{M_0} = \mathbf{r_2} \times \mathbf{F} - \mathbf{r_1} \times \mathbf{F} = (\mathbf{r_2} - \mathbf{r_1}) \times \mathbf{F}.$$

Al considerar la fuerza y el momento resultantes para un sistema, es posible establecer el siguiente teorema.

Teorema 5.5.2. *La proyección del momento resultante sobre la resultante general del sistema es constante.*

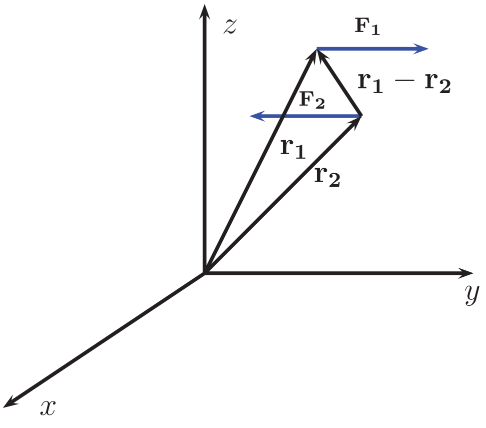


Figura 5.18. Momento producido por un par

Demostración. La proyección vectorial de $\mathbf{M}_c$ sobre la resultante general del sistema es

$$\mathbf{P} = \|\mathbf{M}_c\| \cos\alpha \mathbf{U}_F,$$

donde α es el ángulo entre la resultante y el momento, y $\mathbf{U}_F$ es un vector unitario en la dirección de la fuerza resultante. Esta relación se transforma en

$$\mathbf{P} = \|\mathbf{M}_c\| \left(\frac{\mathbf{M}_c.\mathbf{F}}{\|\mathbf{M}_c\|\|\mathbf{F}\|} \right) \frac{1}{\|\mathbf{F}\|}\mathbf{F},$$

con lo cual

$$\mathbf{P} = \left(\frac{\mathbf{M}_c.\mathbf{F}}{\|\mathbf{F}\|^2} \right) \mathbf{F}.$$

Al tener en cuenta que $\mathbf{M}_c = \mathbf{M}_0 - \mathbf{r} \times \mathbf{F}$, la relación para la proyección del momento se transforma en

$$\mathbf{P} = \left(\frac{\mathbf{M}_c.\mathbf{F}}{\|\mathbf{F}\|^2} \right) \mathbf{F} = \left[\frac{(\mathbf{M}_0 - \mathbf{r} \times \mathbf{F}).\mathbf{F}}{\|\mathbf{F}\|^2} \right] \mathbf{F}.$$

Al equiparar las componentes del vector $\mathbf{F}$ en cada lado de la igualdad y aplicando la ley distributiva del producto escalar

obtenemos la ecuación
$$\frac{\mathbf{M}_c.\mathbf{F}}{\|\mathbf{F}\|^2}=\frac{\mathbf{M}_0.\mathbf{F}-\mathbf{r}\times\mathbf{F}.\mathbf{F}}{\|\mathbf{F}\|^2}.$$

Pero $\mathbf{r}\times\mathbf{F}.\mathbf{F}=0$, puesto que los vectores $\mathbf{r}\times\mathbf{F}$ y $\mathbf{F}$ son perpendiculares. De este modo,
$$\frac{\mathbf{M}_c.\mathbf{F}}{\|\mathbf{F}\|^2}=\frac{\mathbf{M}_0.\mathbf{F}}{\|\mathbf{F}\|^2}.$$

Como el momento resultante de un sistema cambia al modificarse el punto con respecto al cual se calcula, existen puntos sobre los que el momento total y la resultante del sistema son vectores paralelos. Al lugar geométrico de los puntos que poseen esta propiedad se le conoce como *eje central* del sistema. En este caso, si $\mathbf{r_c}$ es el vector de posición de un punto sobre el eje central, se cumple que
$$\mathbf{M_c}=t\mathbf{F_r}=\mathbf{M_0}-\mathbf{r_c}\times\mathbf{F_r}.$$

Si la resultante del sistema y el momento respecto al origen están dados por
$$\mathbf{F}=f_x\mathbf{i}+f_y\mathbf{j}+f_z\mathbf{k}\ \text{ y }\mathbf{M}=m_x\mathbf{i}+m_y\mathbf{j}+m_z\mathbf{k}$$
y el vector de posición de un punto sobre el eje central esta dado por $\mathbf{r}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$, entonces, la *ecuación del eje central* es
$$t(f_x\mathbf{i}+f_y\mathbf{j}+f_z\mathbf{k})=m_x\mathbf{i}+m_y\mathbf{j}+m_z\mathbf{k}-\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\x&y&z\\f_x&f_y&f_z\end{vmatrix},$$
la cual da origen al sistema
$$\begin{aligned}tf_x&=m_x-(yf_z-zf_y)\\tf_y&=m_y-(zf_x-xf_z)\\tf_z&=m_z-(xf_y-yf_x),\end{aligned}$$
que conforman las ecuaciones paramétricas del eje central.

Ejemplo 5.5.3. Hallar la ecuación del eje central para el conjunto de fuerzas que actúan sobre los puntos indicados.

$$\begin{array}{ll} \mathbf{F_1} = 3\mathbf{i} + 2\mathbf{j} - \mathbf{k} & \mathbf{r_1} = \mathbf{i} - \mathbf{j} + \mathbf{k} \\ \mathbf{F_2} = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k} & \mathbf{r_2} = \mathbf{j} + 2\mathbf{k} \\ \mathbf{F_3} = \mathbf{i} + \mathbf{j} + \mathbf{k} & \mathbf{r_3} = \mathbf{i} - \mathbf{j} + 2\mathbf{k} \\ \mathbf{F_4} = -2\mathbf{i} + 2\mathbf{j} - \mathbf{k} & \mathbf{r_4} = \mathbf{i} + \mathbf{j} + \mathbf{k}. \end{array}$$

Solución

En este caso, la fuerza resultante es $\mathbf{F} = 5\mathbf{i} + 4\mathbf{j} + \mathbf{k}$. El momento de cada una de las fuerzas, con respecto al origen, es:

$$\begin{aligned} \mathbf{r_1} \times \mathbf{F_1} &= -\mathbf{i} + 4\mathbf{j} + 5\mathbf{k} \\ \mathbf{r_2} \times \mathbf{F_2} &= 4\mathbf{i} + 6\mathbf{j} - 3\mathbf{k} \\ \mathbf{r_3} \times \mathbf{F_3} &= -3\mathbf{i} + \mathbf{j} + 2\mathbf{k} \\ \mathbf{r_4} \times \mathbf{F_4} &= -3\mathbf{i} - \mathbf{j} + 4\mathbf{k}, \end{aligned}$$

con lo cual el momento resultante respecto al origen es

$$\mathbf{M} = -3\mathbf{i} + 10\mathbf{j} + 8\mathbf{k}.$$

De esta forma, la ecuación vectorial para el eje central es

$$t(5\mathbf{i} + 4\mathbf{j} + \mathbf{k}) = -3\mathbf{i} + 10\mathbf{j} + 8\mathbf{k} - (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) \times (5\mathbf{i} + 4\mathbf{j} + \mathbf{k}),$$

que da origen al sistema de ecuaciones

$$\begin{aligned} 5t &= -y + 4z - 3 \\ 4t &= x - 5z + 10 \\ t &= -4x + 5y + 8. \end{aligned}$$

Este sistema tiene infinitas soluciones, que pueden escribirse así:

$$\begin{aligned} x - 5z &= -\frac{48}{7} \\ y - 4z &= -\frac{97}{14} \\ t &= \frac{11}{14}. \end{aligned}$$

Si tomamos $z = s$ como un parámetro, las ecuaciones anteriores pueden escribirse en la forma

$$\begin{aligned} x &= -\frac{48}{7} + 5s \\ y &= -\frac{97}{14} + 4s \\ z &= s \\ t &= \frac{11}{14}\,. \end{aligned}$$

Las tres primeras relaciones determinan la ecuación de una recta paralela a la fuerza y es la que contiene a los puntos del eje central, mientras que el valor del parámetro t nos indica que el momento con respecto a puntos sobre el eje central es $|t|$ veces el módulo de la fuerza resultante.

Las ecuaciones paramétricas del eje central

$$\begin{aligned} tf_x &= m_x - (yf_z - zf_y) \\ tf_y &= m_y - (zf_x - xf_z) \\ tf_z &= m_z - (xf_y - yf_x) \end{aligned}$$

constituyen un sistema lineal cuya solución puede escribirse así:

$$\begin{aligned} f_z x - f_x z &= -\frac{f_z^2 m_y - f_y f_z m_z - f_y f_x m_x + f_x^2 m_y}{f_x^2 + f_y^2 + f_z^2} \\ f_z y - f_y z &= \frac{f_z^2 m_x - f_x f_z m_z - f_y f_x m_y + f_y^2 m_x}{f_x^2 + f_y^2 + f_z^2} \\ t &= \frac{f_x m_x + f_y m_y + f_z m_z}{f_x^2 + f_y^2 + f_z^2} = \frac{\mathbf{F} \cdot \mathbf{M}}{\|\mathbf{F}\|^2}. \end{aligned}$$

Las ecuaciones anteriores indican que el valor de t puede calcularse directamente a partir de las componentes de la fuerza y el momento resultantes, mientras que las coordenadas de los puntos sobre el eje central satisfacen un sistema de ecuaciones de la forma

$$\begin{aligned} f_z x - f_x z &= A \\ f_z y - f_y z &= B, \end{aligned}$$

donde los coeficientes A y B están dados por

$$\begin{aligned} A &= -\frac{f_z^2 m_y - f_y f_z m_z - f_y f_x m_x + f_x^2 m_y}{f_x^2 + f_y^2 + f_z^2} \\ B &= \frac{f_z^2 m_x - f_x f_z m_z - f_y f_x m_y + f_y^2 m_x}{f_x^2 + f_y^2 + f_z^2}, \end{aligned}$$

el cual corresponde a un sistema con infinitas soluciones que pueden escribirse en forma paramétrica como

$$\begin{aligned} x &= \frac{A}{f_z} + \frac{f_x}{f_z} s \\ y &= \frac{B}{f_z} + \frac{f_y}{f_z} s \\ z &= s, \end{aligned}$$

que son las ecuaciones paramétricas de una recta paralela $\mathbf{F}$.

Ejemplo 5.5.4. Hallar la ecuación del eje central para el conjunto de fuerzas que actúan en el punto indicado:

$$\begin{array}{ll} \mathbf{F_1} = 3\mathbf{i} - 5\mathbf{j} + 4\mathbf{k} & \mathbf{r_1} = 2\mathbf{i} - \mathbf{j} - \mathbf{k} \\ \mathbf{F_2} = 4\mathbf{i} - \mathbf{j} - \mathbf{k} & \mathbf{r_2} = \mathbf{i} - \mathbf{j} + \mathbf{k} \\ \mathbf{F_3} = \mathbf{i} - 2\mathbf{j} + \mathbf{k} & \mathbf{r_3} = \mathbf{i} + \mathbf{j} + 6\mathbf{k} \\ \mathbf{F_4} = \mathbf{i} + 2\mathbf{j} - \mathbf{k} & \mathbf{r_4} = \mathbf{i} - 2\mathbf{j} - 3\mathbf{k} \\ \mathbf{F_5} = \mathbf{i} - 6\mathbf{j} + 2\mathbf{k} & \mathbf{r_5} = \mathbf{i} - 2\mathbf{j} + 3\mathbf{k}. \end{array}$$

Solución

La fuerza resultante para este sistema es $\mathbf{F} = 10\mathbf{i} - 12\mathbf{j} + 5\mathbf{k}$, en tanto que el momento de cada una de las fuerzas, con respecto al origen, está dado por

$$\begin{aligned}
\mathbf{r_1} \times \mathbf{F_1} &= -9\mathbf{i} - 11\mathbf{j} - 7\mathbf{k} \\
\mathbf{r_2} \times \mathbf{F_2} &= 2\mathbf{i} + 5\mathbf{j} + 3\mathbf{k} \\
\mathbf{r_3} \times \mathbf{F_3} &= 13\mathbf{i} + 5\mathbf{j} - 3\mathbf{k} \\
\mathbf{r_4} \times \mathbf{F_4} &= 8\mathbf{i} - 2\mathbf{j} + 4\mathbf{k} \\
\mathbf{r_5} \times \mathbf{F_5} &= 14\mathbf{i} + \mathbf{j} - 4\mathbf{k},
\end{aligned}$$

con lo cual el momento resultante respecto al origen es

$$\mathbf{M} = 28\mathbf{i} - 2\mathbf{j} - 7\mathbf{k}.$$

De esta forma, los valores de A y B son

$$\begin{aligned}
A &= -\frac{f_z^2 m_y - f_y f_z m_z - f_y f_x m_x + f_x^2 m_y}{f_x^2 + f_y^2 + f_z^2} = -10 \\
B &= \frac{f_z^2 m_x - f_x f_z m_z - f_y f_x m_y + f_y^2 m_x}{f_x^2 + f_y^2 + f_z^2} = 18.
\end{aligned}$$

De esta manera, la ecuación del eje central es

$$\begin{aligned}
x &= \frac{A}{f_z} + \frac{f_x}{f_z} s = -2 + 2s \\
y &= \frac{B}{f_z} + \frac{f_y}{f_z} s = \frac{18}{5} - \frac{12}{5} s \\
z &= s.
\end{aligned}$$

Por último, el valor de t es 1, como puede verificarse fácilmente.

Hay otra propiedad importante que caracteriza al eje central, que queda expresada en el siguiente teorema.

Teorema 5.5.3. *Los puntos sobre los cuales se minimiza el módulo del momento están sobre el eje central.*

Demostración. En este caso se trata de minimizar la función

$$H(x, y, z) = \|\mathbf{M}_c\|^2 = M_x^2 + M_y^2 + M_z^2,$$

para lo cual es necesario resolver el sistema de ecuaciones

$$\frac{\partial H}{\partial x} = 0, \quad \frac{\partial H}{\partial y} = 0 \text{ y } \frac{\partial H}{\partial z} = 0,$$

que es equivalente al sistema

$$\begin{aligned}
M_x \frac{\partial M_x}{\partial x} + M_y \frac{\partial M_y}{\partial x} + M_z \frac{\partial M_z}{\partial x} &= 0 \\
M_x \frac{\partial M_x}{\partial y} + M_y \frac{\partial M_y}{\partial y} + M_z \frac{\partial M_z}{\partial y} &= 0 \\
M_x \frac{\partial M_x}{\partial z} + M_y \frac{\partial M_y}{\partial z} + M_z \frac{\partial M_z}{\partial z} &= 0.
\end{aligned} \tag{5.5.1}$$

Si tenemos en cuenta que el momento respecto a un punto C es

$$\mathbf{M_c} = \mathbf{M_0} - \mathbf{r_c} \times \mathbf{F_r},$$

las ecuaciones de las componentes cartesianas del momento son

$$\begin{aligned}
M_x &= M_{0x} - (yF_z - zF_y) \\
M_y &= M_{0y} - (zF_x - xF_z) \\
M_z &= M_{0z} - (xF_y - yF_x),
\end{aligned}$$

con lo cual

$$\frac{\partial M_x}{\partial x} = 0, \quad \frac{\partial M_x}{\partial y} = -F_z \text{ y } \frac{\partial M_x}{\partial z} = F_y,$$

mientras que

$$\frac{\partial M_y}{\partial x} = F_z, \quad \frac{\partial M_y}{\partial y} = 0 \text{ y } \frac{\partial M_y}{\partial z} = -F_x$$

y, finalmente,

$$\frac{\partial M_z}{\partial x} = -F_y, \quad \frac{\partial M_z}{\partial y} = F_x \text{ y } \frac{\partial M_z}{\partial z} = 0.$$

Con estas relaciones, el sistema 5.5.1 se transforma en

$$\begin{aligned} M_yF_z - M_zF_y &= 0 \\ M_zF_x - M_xF_z &= 0 \\ M_xF_y - M_yF_x &= 0, \end{aligned}$$

de donde podemos deducir que

$$\frac{M_y}{F_y} = \frac{M_z}{F_z}, \quad \frac{M_x}{F_x} = \frac{M_z}{F_z} \text{ y } \frac{M_y}{F_y} = \frac{M_x}{F_x}.$$

Esto es:

$$\frac{M_x}{F_x} = \frac{M_y}{F_y} = \frac{M_z}{F_z} = t,$$

con lo cual concluimos que el vector **M** es un múltiplo del vector **F**, como se ha definido previamente el eje central.

5.6 Ejercicios propuestos capítulo 5

1. Dados los vectores $\mathbf{a} = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ y $\mathbf{b} = \mathbf{i} - 2\mathbf{j} - 2\mathbf{k}$, calcular $\mathbf{a} \times \mathbf{b}$, y verificar que es perpendicular a los vectores **a** y **b**.

2. Hallar la ecuación del plano que pasa por el punto $(2, 3, 4)$ y es paralelo a los vectores $\mathbf{a} = \mathbf{i} + \mathbf{j} - 2\mathbf{k}$ y $\mathbf{a} = 3\mathbf{i} - 2\mathbf{j} - 5\mathbf{k}$.

3. Hallar la ecuación del plano que pasa por los puntos $A(2, 3, 4)$, $B(2, -1, 0)$ y $C(-2, -3, 0)$.

4. Hallar la ecuación del plano que pasa por el punto $A(1, 3, 1)$ y es paralelo al plano con ecuación $3x + 2y - 7z = 2$.

5. Hallar la ecuación del plano que pasa por el punto $A(0, 2, 1)$ y es perpendicular al plano con ecuación $x + y - 2z = 3$.

6. Hallar la ecuación del plano que tiene por intersección con los ejes x, y y z, los números a, b y c.

7. Hallar la ecuación del plano que pasa por el punto $A(1,0,1)$ y es perpendicular a la intersección de los planos cuyas ecuaciones son $3x+2y-7z=2$ y $x-2y-2z=3$.

8. Hallar el área de la poligonal de la Figura 5.19.

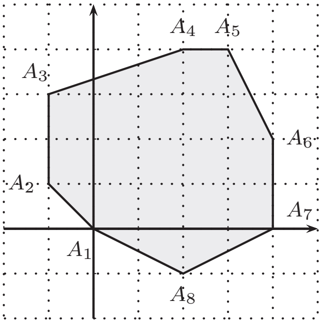


Figura 5.19. Área de la región poligonal convexa

9. Determine un par de planos cuya intersección sea la recta con ecuaciones paramétricas $x=2-3s$, $y=3+s$ y $z=2-4s$.

10. Demuestre la *identidad de Jacobi:*

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}+(\mathbf{b}\times\mathbf{c})\times\mathbf{a}+(\mathbf{c}\times\mathbf{a})\times\mathbf{b}=0.$$

11. Demuestre que

$$\|\mathbf{a}\times\mathbf{b}\|^2+(\mathbf{a}\cdot\mathbf{b})^2=\|\mathbf{a}\|^2\|\mathbf{b}\|^2.$$

12. Determine el área del triángulo con vértices en los puntos $A(2,3,4)$, $B(2,-1,0)$ y $C(-2,-3,0)$.

13. Determine el área del triángulo con vértices en los puntos A, B y C, donde $\mathbf{AB}=\mathbf{i}-2\mathbf{j}+3\mathbf{k}$ y $\mathbf{AC}=2\mathbf{i}-3\mathbf{j}+4\mathbf{k}$.

14. Determine el volumen del paralelepípedo con un vértice en el origen y aristas $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$, $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}$ y $\mathbf{w} = \mathbf{i} + 3\mathbf{j} - 2\mathbf{k}$.

15. Determine el área del paralelogramo con vértices en los puntos A, B y C, donde $\mathbf{AB} = 3\mathbf{i} - \mathbf{j} + 3\mathbf{k}$ y $\mathbf{AC} = -2\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$.

16. Hallar la ecuación del plano que pasa por el punto de corte de las rectas

$$\begin{array}{lcl} x = 2 - 3t & & x = 5 + 2s \\ y = 3 + 2t & \text{y} & y = 1 - 3s \\ z = 4 + 2t & & z = 2 + s, \end{array}$$

y es paralelo al plano con ecuación $x + 2y - z = 2$.

17. Hallar la ecuación del plano que pasa por el punto de corte de las rectas

$$\begin{array}{lcl} x = 2 - 3t & & x = 5 + 2s \\ y = 3 + 2t & \text{y} & y = 1 - 3s \\ z = 4 + 2t & & z = 2 + s, \end{array}$$

y es perpendicular a la recta $x = 2t - 2, \quad y = 5t - 1$ y $z = t$.

18. Hallar la distancia entre las rectas

$$\begin{array}{lcl} x = 2 - 3t & & x = 2 + 2s \\ y = 3 + 2t & \text{y} & y = 1 + s \\ z = 4 + 2t & & z = 3 + 5s. \end{array}$$

19. Halle la intersección del plano con ecuación $x - 3y + 4z = 0$ y el plano que pasa por los puntos $A(3, 3, 4)$, $B(0, -1, 0)$ y $C(-2, -2, 0)$.

20. Halle la intersección del plano con ecuación $2x - y + 4z = 2$ y la recta con ecuación $x = 2t - 2, \quad y = 5t - 1,$ y $z = t$.

21. Muestre que

$$(\mathbf{x} \times \mathbf{y}).\mathbf{z} = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

22. Muestre que dos vectores de $\mathbb{R}^3$, $\mathbf{v_1}$ y $\mathbf{v_2}$, son paralelos si y solo si $\mathbf{v_1} \times \mathbf{v_2} = 0$.

23. Halle la ecuación del eje central para el conjunto de fuerzas que actúan sobre los puntos indicados.

$$\begin{array}{ll} \mathbf{F_1} = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k} & \mathbf{r_1} = 2\mathbf{i} - \mathbf{j} + \mathbf{k} \\ \mathbf{F_2} = \mathbf{i} - \mathbf{j} + 3\mathbf{k} & \mathbf{r_2} = -\mathbf{j} + 2\mathbf{k} \\ \mathbf{F_3} = \mathbf{i} - 2\mathbf{j} + \mathbf{k} & \mathbf{r_3} = \mathbf{i} - \mathbf{j} + 3\mathbf{k} \\ \mathbf{F_4} = -2\mathbf{i} + 2\mathbf{j} - 3\mathbf{k} & \mathbf{r_4} = 4\mathbf{i} + \mathbf{j} + \mathbf{k} \\ \mathbf{F_5} = \mathbf{i} - 4\mathbf{j} + 3\mathbf{k} & \mathbf{r_5} = \mathbf{i} - 6\mathbf{j} - \mathbf{k}. \end{array}$$

Capítulo 6

Transformaciones lineales

6.1 Introducción

En este capítulo nos interesamos por aquellas aplicaciones entre espacios vectoriales que preservan las operaciones de suma y producto por escalar. A este tipo de aplicaciones las denominamos *homomorfismos o transformaciones lineales.*

Al igual que en el caso del espacio de las matrices, acá también es posible sumar transformaciones lineales y multiplicarlas por un escalar. Aun cuando suponemos que el lector está familiarizado con los conceptos y las definiciones de la teoría de funciones, en el desarrollo del capítulo definimos aquellos conceptos más importantes para la comprensión clara del tema.

El origen de la teoría de las transformaciones lineales está en los trabajos del matemático francés Francisco Viète (1540-1603). La teoría inicial sobre las transformaciones lineales fue extendiéndose poco a poco en muchas ramas del álgebra y, posteriormente, con la aparición del concepto de *invariante*, ocupó un lugar especial en toda la matemática. Uno de los aspectos más importantes del concepto de *transformación lineal* radica en el arte de reducir un problema no resuelto a una serie de problemas con solución conocida (Bell, 2002: 128-132).

6.2 Transformaciones lineales

Definición 6.2.1. *Sean V y W espacios vectoriales. Una transformación lineal es una función $T : V \to W$ de un espacio vectorial V en un espacio vectorial W que cumple las siguientes condiciones:*

1. *Para cualquier x y y en V, se tiene que $T(x+y) = T(x)+T(y)$.*

2. *Para cualquier x en V y cualquier escalar α (en $\mathbb{R}$ o en $\mathbb{C}$) se tiene que $T(\alpha x) = \alpha T(x)$.*

El par de condiciones anteriores es equivalente a que para cualquier x, y en V y α, λ en $\mathbb{R}$, se tiene que

$$T(\alpha x + \lambda y) = \alpha T(x) + \lambda T(y).$$

En esta definición debemos observar que la suma $x + y$ se realiza entre elementos del espacio vectorial V, mientras que la suma $T(x) + T(y)$ se hace entre elementos del espacio vectorial W.

Además, si tenemos que $V = W$, a la aplicación lineal que se obtiene, $T : V \to V$, la denominamos *endomorfismo*.

Ejemplo 6.2.1. Determine si la aplicación dada por

$$\begin{aligned} T : \mathbb{R}^2 \to &\quad \mathbb{R}^3 \\ (x, y) \to &\quad T(x, y) = (2x, y, x + y) \end{aligned}$$

es una transformación lineal del espacio $\mathbb{R}^2$ en el espacio $\mathbb{R}^3$.
Solución
Para determinar si esta transformación es lineal, debemos verificar que cumple las dos condiciones de la definición.

1. Sean $u = (x_1, y_1)$ y $v = (x_2, y_2)$ dos vectores de $\mathbb{R}^2$; entonces,

$$\begin{aligned} T(u+v) &= T((x_1, y_1) + (x_2, y_2)) \\ &= T(x_1 + x_2, y_1 + y_2) \\ &= (2(x_1 + x_2), y_1 + y_2, x_1 + x_2 + y_1 + y_2) \\ &= (2x_1, y_1, x_1 + y_1) + (2x_2, y_2, x_2 + y_2) \\ &= T(x_1, y_1) + T(x_2, y_2) \\ &= T(u) + T(v). \end{aligned}$$

2. Sea α en $\mathbb{R}$ y $u = (x_1, y_1)$ en $\mathbb{R}^2$; así:

$$\begin{aligned} T(\alpha u) &= T(\alpha(x_1, y_1)) \\ &= T((\alpha x_1, \alpha y_1)) \\ &= T(\alpha x_1, \alpha y_1) \\ &= (2\alpha x_1, \alpha y_1, \alpha x_1 + \alpha y_1) \\ &= \alpha(2x_1, y_1, x_1 + y_1) \\ &= \alpha T(x_1, y_1) \\ &= \alpha T(u), \end{aligned}$$

lo cual nos muestra que dicha transformación es lineal.

Ejemplo 6.2.2. Sean A una matriz de tamaño $m \times n$; R_m el conjunto de todos los vectores columna de m componentes y R_n el conjunto de todos los vectores columna de n componentes; determine si la aplicación dada por

$$\begin{aligned} T : R_m \to & \quad R_n \\ X \to & \quad T(X) = AX \end{aligned}$$

es una transformación lineal.

Solución

Veamos si las dos propiedades que caracterizan a las transformaciones lineales se cumplen.

1. Sean X y Y dos vectores de R_m:

$$\begin{aligned} T(X+Y) &= A(X+Y) \\ &= AX+AY \\ &= T(X)+T(Y). \end{aligned}$$

2. Si α está en $\mathbb{R}$ y X es un vector de R_m, entonces

$$\begin{aligned} T(\alpha X) &= A(\alpha X) \\ &= \alpha AX \\ &= \alpha T(X), \end{aligned}$$

lo cual nos permite comprobar que esta aplicación es una transformación lineal.

Ejemplo 6.2.3. Consideremos una rotación R_θ alrededor del origen, en sentido contrarreloj, de un ángulo de θ radianes, como se ilustra en la Figura 6.1.

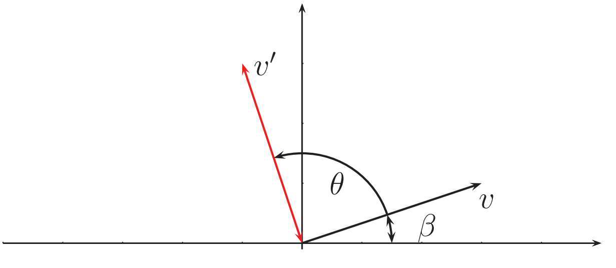


Figura 6.1. Rotación del punto $v=(x,y)$ un ángulo θ

Podemos ver de manera geométrica que la aplicación dada por $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ es una aplicación lineal.

Solución

En las Figuras 6.2 y 6.3 se muestra cómo se transforman los vectores $u+v$ y αu bajo una rotación R_θ alrededor del origen.

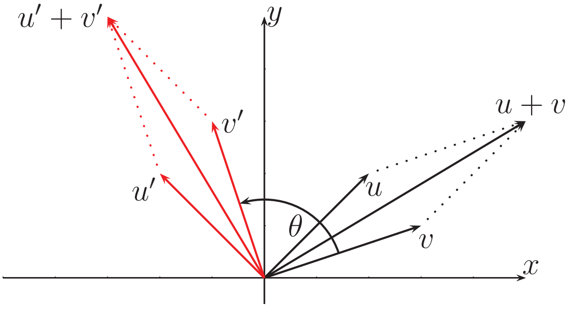


Figura 6.2. Rotación de $u + v$ un ángulo θ

Ahora encontremos $R_\theta(x, y)$ y verifiquemos que es una transformación lineal.

Sea $v = (x, y)$ un vector dado; si $r = \|v\|$, $x = r\cos\beta$ y $y = r \operatorname{sen}\beta$,

$$\begin{aligned} R_\theta(x, y) &= (r\cos(\beta+\theta), r\operatorname{sen}(\beta+\theta)) \\ &= (r\cos\beta\cos\theta - r\operatorname{sen}\beta\operatorname{sen}\theta, r\cos\beta\operatorname{sen}\theta - r\operatorname{sen}\beta\cos\theta) \\ &= (x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta - y\cos\theta). \qquad (6.2.1) \end{aligned}$$

De este modo, si queremos calcular

$$R_\theta(u + v) = R_\theta(x + x', y + y'),$$

obtenemos:

$$\begin{aligned} R_\theta(u+v) &= ((x+x')\cos\theta - (y+y')\operatorname{sen}\theta, (x+x')\operatorname{sen}\theta - (y+y')\cos\theta) \\ &= (x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta - y\cos\theta) \\ &\qquad + (x'\cos\theta - y'\operatorname{sen}\theta, x'\operatorname{sen}\theta - y'\cos\theta) \\ &= R_\theta(u) + R_\theta(v). \end{aligned}$$

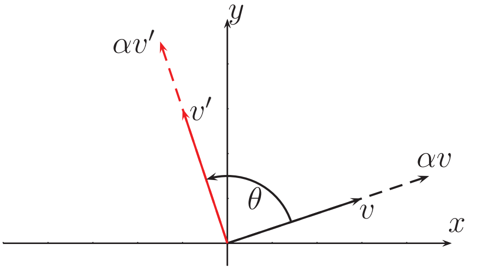


Figura 6.3. Rotación de αv un ángulo θ

Si c es un escalar,

$$\begin{aligned} R_\theta(cu) &= (cx\cos\theta - cy\operatorname{sen}\theta, cx\operatorname{sen}\theta - cy\cos\theta) \\ &= c(x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta - y\cos\theta) \\ &= cR_\theta(u). \end{aligned}$$

Los resultados anteriores nos muestran que la transformación $R_\theta(u)$ es una transformación lineal.

Se puede llegar a la relación 6.2.1 asumiendo que la transformación es lineal y determinando los valores de $R_\theta(1,0)$ y $R_\theta(0,1)$. En la Figura 6.4 vemos que R_θ conserva las magnitudes, es decir, $\|R_\theta(u)\| = \|u\|$ y, en consecuencia,

$$\begin{aligned} R_\theta(1,0) &= (\cos\theta, \operatorname{sen}\theta) \\ R_\theta(0,1) &= \left(\cos\left(\theta + \frac{\pi}{2}\right), \operatorname{sen}\left(\theta + \frac{\pi}{2}\right)\right) = (-\operatorname{sen}\theta, \cos\theta). \end{aligned}$$

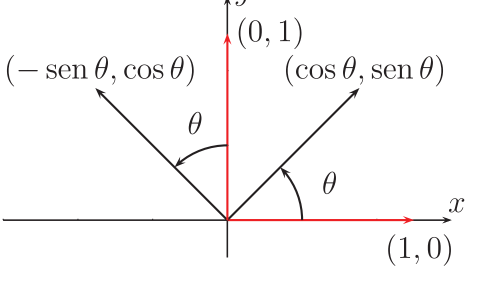


Figura 6.4. Rotación de los vectores $(1,0)$ y $(0,1)$ un ángulo θ

Con las imágenes de los vectores $(1,0)$ y $(0,1)$, y las propiedades de las transformaciones lineales, calculamos $R_\theta(x,y)$ de la siguiente forma:

$$\begin{aligned}
R_\theta(x,y) &= R_\theta(x(1,0)+y(0,1))\\
&= R_\theta(x(1,0))+R_\theta(y(0,1))\\
&= xR_\theta(1,0)+yR_\theta(0,1)\\
&= x(\cos\theta,\operatorname{sen}\theta)+y(-\operatorname{sen}\theta,\cos\theta)\\
&= (x\cos\theta-y\operatorname{sen}\theta, x\operatorname{sen}\theta+y\cos\theta),
\end{aligned}$$

que es la misma relación dada en 6.2.1.

En consecuencia, la rotación R_θ queda definida como

$$\begin{aligned}
R_\theta:\mathbb{R}^2 &\to \mathbb{R}^2\\
(x,y) &\to R_\theta(x,y)=(x\cos\theta-y\operatorname{sen}\theta, x\operatorname{sen}\theta+y\cos\theta).
\end{aligned}$$

Este último resultado podemos escribirlo en forma matricial de la siguiente forma:

$$R_\theta(x,y)=AX=\begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix},\quad \text{donde}\quad X=\begin{bmatrix}x\\ y\end{bmatrix}.$$

La matriz A es conocida como la *matriz de rotaciones*.

Ejemplo 6.2.4. Verifique que la aplicación $T:\mathbb{R}^2\to\mathbb{R}^2$, dada por la expresión

$$T\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}y\\ x\end{bmatrix},$$

es una transformación lineal.

Solución

Como la aplicación T podemos escribirla en la forma $T(X)=AX$, donde A es una matriz de tamaño 2×2 y X es un vector columna, entonces, por las propiedades del producto de matrices, tenemos que:

1. Para cualquier X en $\mathbb{R}^2$ y α en $\mathbb{R}$,

$$T(\alpha X) = A(\alpha X) = \alpha AX = \alpha T(X).$$

2. Para cualesquier X y Y en $\mathbb{R}^2$,

$$T(X+Y) = A(X+Y) = AX + AY = T(X) + T(Y).$$

A esta transformación lineal se le conoce como una *reflexión* alrededor de la recta $y = x$, como puede verse en la Figura 6.5.

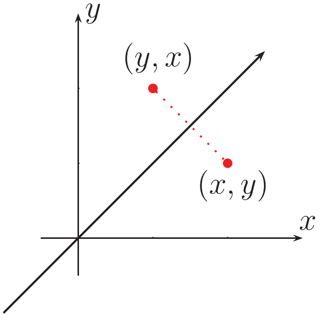


Figura 6.5. Reflexión del vector (x, y) alrededor de $y = x$

Geométricamente, el punto (x, y) se refleja sobre la recta $y = x$, transformándose en el punto (y, x).

Ejemplo 6.2.5. Demuestre que las siguientes aplicaciones no son transformaciones lineales.

1.
$$\begin{aligned} T : \mathbb{R}^2 &\to \quad \mathbb{R}^2 \\ \begin{bmatrix} x \\ y \end{bmatrix} &\to \quad T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+4 \\ 2 \end{bmatrix}. \end{aligned}$$

2.
$$\begin{aligned} T : \mathbb{R} &\to \quad \mathbb{R} \\ x &\to \quad T(x) = \operatorname{sen} x. \end{aligned}$$

3.
$$\begin{aligned} T:\mathbb{R}^2 &\to \mathbb{R}^3 \\ \begin{bmatrix} x \\ y \end{bmatrix} &\to T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^2 \\ x+y \\ 0 \end{bmatrix}. \end{aligned}$$

Solución

Para el primer caso, sólo basta probar que

$$T(u+v) \neq T(u) + T(v).$$

Sea $u = (x_1, y_1)$ y $v = (x_2, y_2)$; entonces,

$$T(u+v) = T\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 4 \\ 2 \end{bmatrix},$$

mientras que

$$T(u) + T(v) = \begin{bmatrix} x_1 + 4 \\ 2 \end{bmatrix} + \begin{bmatrix} x_2 + 4 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 8 \\ 4 \end{bmatrix},$$

lo cual prueba que $T(u+v) \neq T(u) + T(v)$.

Para la segunda aplicación se tiene que

$$T(x+y) = \operatorname{sen}(x+y) \neq \operatorname{sen} x + \operatorname{sen} y.$$

En la tercera aplicación, nuevamente tenemos que

$$T(u+v) = T\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} = \begin{bmatrix} (x_1 + x_2)^2 \\ x_1 + x_2 + y_1 + y_2 \\ 0 \end{bmatrix},$$

mientras que

$$\begin{aligned} T(u) + T(v) &= \begin{bmatrix} x_1^2 \\ x_1 + y_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2^2 \\ x_2 + y_2 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} x_1^2 + x_2^2 \\ x_1 + y_1 + x_2 + y_2 \\ 0 \end{bmatrix}, \end{aligned}$$

lo cual prueba que $T(u+v) \neq T(u) + T(v)$.

6.3 Propiedades de las transformaciones lineales

Una de las más importantes aplicaciones de las transformaciones lineales tiene que ver con el hecho de que una transformación lineal $T : V \to W$ queda completamente determinada por las imágenes de los elementos de una base para el espacio vectorial V. El siguiente teorema nos muestra cómo se aplica una transformación lineal a un vector que es combinación lineal de un conjunto dado de vectores.

Teorema 6.3.1. *Sea $T : V \to W$ una transformación lineal y sean $x_1, x_2, \ldots, x_n$ vectores en V; entonces, dados los escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, se tiene que*

$$T(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n) = \alpha_1 T(x_1) + \alpha_2 T(x_2) + \ldots + \alpha_n T(x_n).$$

Demostración. Usando inducción matemática, tenemos que para $n = 1$ es inmediato que $T\alpha_1(x_1) = \alpha_1 T(x_1)$. Supongamos que la hipótesis es válida hasta $n-1$, es decir,

$$T(\alpha_1 x_1 + \ldots + \alpha_{n-1} x_{n-1}) = \alpha_1 T(x_1) + \ldots + \alpha_{n-1} T(x_{n-1});$$

y sean los vectores $x_1, x_2, \ldots, x_n$ en V. Así,

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$$

es también un vector en V.

Usando la primera propiedad de la definición de transformación lineal y aplicando la hipótesis de inducción matemática, se tiene que

$$\begin{aligned} T(y) &= T(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n) \\ &= T(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_{n-1} x_{n-1}) + T(\alpha_n x_n) \\ &= T(\alpha_1 x_1) + T(\alpha_2 x_2) + \ldots + T(\alpha_{n-1} x_{n-1}) + T(\alpha_n x_n). \end{aligned}$$

Ahora, por la segunda condición de la definición de transformación lineal, $T(\alpha_i x_i) = \alpha_i T(x_i)$ para $i = 1, 2, \ldots, n$; por lo tanto,

$$T(y) = \alpha_1 T(x_1) + \alpha_2 T(x_2) + \ldots + \alpha_n T(x_n),$$

como queríamos probar.

Si en este momento imponemos ciertas condiciones a los vectores $x_1, x_2, \ldots, x_n$ de V, conseguimos el siguiente teorema.

Teorema 6.3.2. *Sean* $T : V \to W$ *una transformación lineal y* $B = \{x_1, x_2, \ldots, x_n\}$ *una base para* V. *Si para* $i = 1, 2, \ldots, n$, $T(x_i) = y_i$, *entonces, para cualquier* x *en* V, $T(x)$ *está completamente determinada y, además,*

$$T(x) = \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_n y_n,$$

donde $\alpha_1, \alpha_2, \ldots, \alpha_n$ *son escalares tales que*

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n.$$

Ejemplo 6.3.1. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal tal que

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad \text{y} \quad T\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}.$$

Halle $T(x, y)$ y $T(4, 3)$.

Solución

En primer lugar, escribimos el vector $(x, y) \in \mathbb{R}^2$ como una combinación lineal de los vectores $(1, 1)$ y $(2, 1)$. Resolviendo la ecuación vectorial,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha + 2\beta \\ \alpha + \beta \end{bmatrix},$$

de donde obtenemos el sistema de ecuaciones

$$\begin{aligned} \alpha + 2\beta &= x \\ \alpha + \beta &= y. \end{aligned}$$

Al resolver este sistema vemos que $\alpha = -x + 2y$ y $\beta = x - y$, con lo cual

$$\begin{bmatrix} x \\ y \end{bmatrix} = (-x + 2y) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (x - y) \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

y, por lo tanto,

$$T\begin{bmatrix} x \\ y \end{bmatrix} = (-x+2y)T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (x-y)T\begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Usando el teorema anterior, concluimos que

$$T\begin{bmatrix} x \\ y \end{bmatrix} = (-x+2y)\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + (x-y)\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} y \\ -x+3y \\ 4x-4y \end{bmatrix}.$$

De aquí obtenemos que

$$T\begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix};$$

este último resultado se consigue sustituyendo (x, y) por $(4, 3)$.

Teorema 6.3.3 (Propiedades de las transformaciones). *Si $T : V \to W$ es una transformación lineal, entonces:*

1. $T(0_v) = 0_w$, *donde* 0_v *y* 0_w *son los ceros de* V *y* W *respectivamente.*

2. $T(-x) = -T(x)$ *para cualquier* x *en* V.

3. *Si* $\{x_1, x_2, \ldots, x_n\}$ *es un conjunto de vectores linealmente dependientes en* V, *entonces* $\{T(x_1), T(x_2), \ldots, T(x_n)\}$ *es un conjunto de vectores linealmente dependientes en* W.

4. *Si* $T : V \to W$ *y* $S : W \to H$ *son transformaciones lineales, entonces la composición* $S \circ T : V \to H$ *es una transformación lineal.*

Demostración. Probemos las tres primeras afirmaciones; la última se deja como ejercicio.

Dadas las transformaciones T y S, entonces,

1. Del teorema 6.3.1 concluimos que

$$T\left(\sum_{i=1}^{n}\alpha_i x_i\right)=\sum_{i=1}^{n}\alpha_i T(x_i),\quad x_i\in V,\quad \alpha_i\in\mathbb{R}, i=1,\ldots,n.$$

Si cada uno de los $\alpha_i=0$, ambos lados de la relación anterior son nulos, es decir, $T(0_v)=0_w$, como queríamos probar.

2. En este caso, si $\alpha_1=-1$ y los demás escalares son nulos, conseguimos que $T(-x_1)=-T(x_1)$ para cualquier $x_1\in V$.

3. En este caso, si

$$\alpha_1x_1+\alpha_2x_2+\ldots+\alpha_nx_n=0_v$$

con alguno de los $\alpha_i\neq 0$, al aplicar el primer numeral del teorema 6.3.1 tenemos que

$$0_w=\alpha_1T(x_1)+\alpha_2T(x_2)+\ldots+\alpha_nT(x_n),$$

de donde se concluye que el conjunto $T(x_1),T(x_2),\ldots,T(x_n)$ es linealmente dependiente.

Es importante aclarar que la independencia lineal no es una propiedad que se conserve en las transformaciones lineales; por ejemplo, si $T:\mathbb{R}^3\to\mathbb{R}^2$ es la transformación lineal tal que

$$T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x\\y\end{bmatrix},$$

entonces, si elegimos los vectores $u=(1,0,0)$ y $v=(0,1,0)$ como dos vectores que son linealmente independientes,

$$T\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}\quad\text{y}\quad T\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix},$$

lo cual muestra que $T(u)$ y $T(v)$ son dos vectores linealmente independientes; mientras que si tomamos los vectores $a = (1, 0, 1)$ y $b = (2, 0, 0)$ como dos vectores linealmente independientes,

$$T\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y } T\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Así, $T(a)$ y $T(b)$ son linealmente dependientes.

Ejemplo 6.3.2. En una empresa se producen artículos que se componen de tres partes: A, B y C. La Tabla 6.1 indica cuántas de cada una de las partes se requieren para hacer un determinado artículo.

	Artículo 1	**Artículo 2**	**Artículo 3**
Parte A	6	8	4
Parte B	3	4	2
Parte C	1	2	1

Tabla 6.1. Datos del ejemplo 6.3.2

Sea x_i el número de artículos de la clase i que se requieren en una orden de ventas para $i = 1, 2$ o 3, y sea y_i el número de partes del tipo i para fabricar los artículos pedidos en la orden. Entonces, la transformación $T : \mathbb{R}^3 \to \mathbb{R}^3$, que envía al vector de orden de venta (x_1, x_2, x_3) al vector de partes (y_1, y_2, y_3), está dada por

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 & 8 & 4 \\ 3 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Si la orden pide 30 artículos del tipo 1, 20 del tipo 2 y 25 del tipo 3, entonces

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = T\begin{bmatrix} 30 \\ 20 \\ 25 \end{bmatrix} = \begin{bmatrix} 6 & 8 & 4 \\ 3 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} 30 \\ 20 \\ 25 \end{bmatrix} = \begin{bmatrix} 440 \\ 220 \\ 95 \end{bmatrix},$$

es decir, para cumplir la orden, se requieren 440 unidades de partes del tipo A, 220 del tipo B y 95 del tipo C.

Supongamos ahora, en este mismo problema, que los artículos son hechos en hierro y madera. Sea z_1 la cantidad de hierro y z_2 la cantidad de madera requeridas para fabricar y_1 unidades A, y_2 unidades B y y_3 unidades C. La Tabla 6.2 muestra la cantidad de material empleado en la producción de las partes A, B y C.

Material	**Parte A**	**Parte B**	**Parte C**
Hierro	0	1	6
Madera	0,5	0	4

Tabla 6.2. Cantidades de material por tipo de parte

La transformación que envía al vector partes Y al vector de materiales Z está dada por $S : \mathbb{R}^3 \to \mathbb{R}^2$, tal que

$$S(Y) = \begin{bmatrix} 0 & 1 & 6 \\ 0,5 & 0 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

Ahora, para encontrar las cantidades de materiales requeridos en términos del vector de la orden de venta X, consideremos la composición $S \circ T$, que se calcula de la siguiente forma:

$$S \circ T(X) = S(T(X)) = \begin{bmatrix} 0 & 1 & 6 \\ 0,5 & 0 & 4 \end{bmatrix} \begin{bmatrix} 6 & 8 & 4 \\ 3 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

la cual, luego de efectuar el producto de matrices, se transforma en

$$S \circ T(X) = S(T(X)) = \begin{bmatrix} 9 & 16 & 8 \\ 7 & 12 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

De esta manera, para completar la orden $(30, 20, 25)$, el material utilizado sería

$$S \circ T(X) = S(T(X)) = \begin{bmatrix} 9 & 16 & 8 \\ 7 & 12 & 6 \end{bmatrix} \begin{bmatrix} 30 \\ 20 \\ 25 \end{bmatrix} = \begin{bmatrix} 790 \\ 600 \end{bmatrix},$$

o sea, 790 unidades de hierro y 600 unidades de madera.

6.4 Rango y núcleo

En esta sección estudiamos dos subespacios asociados a una transformación lineal $T : V \to W$. Se trata del espacio conocido como el *núcleo de la transformación*, el cual es un subespacio de V formado por todos los vectores que son transformados en el vector 0_w mediante la aplicación de la transformación T. A este subespacio lo denotamos con N_T. El otro espacio importante, en el caso de las transformaciones, es el subespacio $T(v)$ de W, constituido por todas las imágenes de T y que representamos con R_T. Formalmente, tenemos las siguientes definiciones:

$$N_T = \{x \in V \,/\, T(x) = 0_w\};$$

$$R_T = \{y \in W \,/\, T(x) = y, \quad \text{para algún}\, x \in V\}.$$

Teorema 6.4.1. *Sea $T : V \to W$ una transformación lineal. Entonces, N_T es subespacio de V y R_T es un subespacio de W.*

Demostración. Es claro que $N_T \neq \Phi$, puesto que $T(0_v) = 0_w$, donde 0_v es el cero de V y 0_w es el cero de W. Ahora, si $x, y \in N_T$, entonces $T(x) = T(y) = 0_w$ y, en consecuencia,

$$T(x + y) = T(x) + T(y) = 0_w + 0_w = 0_w,$$

lo cual prueba que $x + y \in N_T$. Además, si $\alpha \in \mathbb{R}$ y $x \in N_T$,

$$T(\alpha x) = \alpha T(x) = \alpha 0_w = 0_w;$$

esto es, $\alpha x \in N_T$. Se comprueba así que N_T es un subespacio.

De modo similar, para demostrar que R_T es un subespacio, tenemos que $R_T \neq \Phi$, puesto que $T(0_v) = 0_w$, con lo cual $0_v \in R_T$. Además, si $y_1, y_2 \in R_T$, existen $x_1, x_2 \in V$ tales que $T(x_1) = y_1$ y $T(x_2) = y_2$; y como

$$T(x_1 + x_2) = T(x_1) + T(x_2) = y_1 + y_2,$$

y $x_1 + x_2 \in V$, entonces $y_1 + y_2 \in R_T$. Por último, si $\alpha \in \mathbb{R}$ y $y \in R_T$, existe $x \in V$ tal que $T(x) = y$. Como $T(\alpha x) = \alpha T(x) = \alpha y$ y $\alpha x \in V$, $\alpha y \in R_T$, lo cual prueba que R_T es un subespacio de W.

Definición 6.4.1. *Sea $T : V \to W$ una transformación lineal; a la dimensión de R_T se le llama rango de la transformación lineal. A la dimensión de N_T se le denomina nulidad de la transformación lineal.*

Ejemplo 6.4.1. Dada la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$,

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y \\ -2y + z \\ x - y + z \end{bmatrix},$$

calcular N_T y R_T.

Solución

La terna (x, y, z) es un elemento de N_T si y solo si

$$\begin{bmatrix} x + y \\ -2y + z \\ x - y + z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

entonces, obtenemos el sistema homogéneo

$$\begin{aligned} x + y &= 0 \\ -2y + z &= 0 \\ x - y + z &= 0. \end{aligned}$$

La forma escalonada reducida de la matriz de coeficientes de este sistema está dada por

$$\begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix},$$

lo cual indica que este sistema tiene infinitas soluciones, que en forma paramétrica están constituidas por

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\dfrac{t}{2} \\ \dfrac{t}{2} \\ t \end{bmatrix} = \frac{t}{2} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$$

De este modo, $N_T = gen\{(-1, 1, 2)\}$ y $dim\, N_T = 1$; esto es, la nulidad de T es uno.

Vamos ahora a calcular el rango de la transformación. Sea $z = (z_1, z_2, z_3) \in R_T$; entonces, existe $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ tal que $T(x) = z$, es decir,

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ -2x_2 + x_3 \\ x_1 - x_2 + x_3 \end{bmatrix},$$

de lo cual concluimos que

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = x_1 \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}_{v_1} + x_2 \underbrace{\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}}_{v_2} + x_3 \underbrace{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}_{v_3}.$$

Esto último nos dice que el rango de la transformación es una combinación lineal de los vectores v_1, v_2 y v_3; en consecuencia,

$$R_T = gen\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Como este conjunto de vectores es linealmente dependiente y, además,

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix},$$

se tiene que

$$R_T = gen\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} \right\}$$

y, por lo tanto, $dim\, R_T = 2$, es decir, el rango de la transformación es dos.

Debemos notar que para este ejemplo se cumple la relación $dim\, \mathbb{R}^3 = dim\, N_T + dim\, R_T$. Este resultado lo demostraremos en la siguiente sección para una transformación lineal de un espacio V en un espacio W.

Teorema 6.4.2. *Sea $T : V \to W$ una transformación lineal de V en W. Si $\{v_1, v_2, \ldots, v_n\}$ es una base para V, entonces el conjunto de vectores $\{T(v_1), T(v_2), \ldots T(v_n)\}$ es un generador de R_T.*

Demostración. Sea z cualquier elemento de R_T; así, existe un elemento $x \in V$ tal que $T(x) = z$. Como $\{v_1, v_2, \ldots, v_n\}$ es una base para V, entonces existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que $x = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$.

En tal caso,

$$z = T(x) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \ldots + \alpha_n T(v_n);$$

es decir, $R_T = gen\, \{T(v_1), T(v_2), \ldots, T(v_n)\}$.

6.5 Transformaciones lineales inyectivas

Entre las transformaciones lineales, las inyectivas son de especial importancia. Estas dejan invariante la dimensión, tienen por núcleo el espacio nulo y transforman una base del dominio en una base de la imagen. Con el propósito de establecer algunas propiedades de las transformaciones lineales inyectivas, presentamos la siguiente definición.

Definición 6.5.1 (**Aplicación inyectiva**). *Una aplicación* $T : V \to W$ *es inyectiva (o uno a uno) cuando para ella ocurre que elementos diferentes de* V *tienen imágenes distintas en* W*; esto es, si para cualquier* a *y* b *en* V*, si* $a \neq b$*, entonces* $T(a) \neq T(b)$*.*

La definición anterior es equivalente a afirmar que si $T(a) = T(b)$, entonces $a = b$.

Ejemplo 6.5.1. Sea $f : \mathbb{R} \to \mathbb{R}^+$, donde $f(x) = e^x$; esta función es inyectiva. En el caso de las funciones reales de variable real, se puede verificar que una función es inyectiva en un intervalo cualquiera si su derivada allí no cambia de signo; así, si la derivada es positiva en ese intervalo, la función es creciente; mientras que si la derivada es negativa en dicho intervalo, la función será decreciente. Para la función dada, como su derivada es ella misma, la función es siempre creciente.

El siguiente teorema nos proporciona un criterio para determinar cuándo una transformación lineal es inyectiva, o uno a uno.

Teorema 6.5.1. *Sea* $T : V \to W$ *una transformación lineal. Entonces,* T *es uno a uno si y solo si* $N_T = \{0\}$*.*

Demostración. Supongamos que T es uno a uno. Sea $x \in N_T$; entonces, $T(x) = 0_w$. Como, además, $T(0_v) = 0_w$, al ser T uno a uno, $0_v = x$; de esta forma, $N_T \subset \{0\}$ y, así, $N_T = \{0\}$.

Supongamos ahora que $N_T = \{0\}$ y $x, y \in V$ tales que $T(x) = T(y)$; entonces, $T(x) - T(y) = 0$. Pero como T es una transformación lineal,

$$T(x) - T(y) = T(x - y) = 0$$

y, en consecuencia, $x - y \in N_T$; por lo tanto, $x - y = 0$, es decir, $x = y$, lo que implica que T es inyectiva.

El siguiente teorema nos presenta una importante relación entre el rango y la nulidad de una transformación lineal.

Teorema 6.5.2. *Sea $T : V \to W$ una transformación lineal, donde V es un espacio vectorial de dimensión finita n; entonces, se cumple que*

$$dim\, V = dim\, N_T + dim\, R_T.$$

Demostración. Por el teorema 6.4.1, sabemos que N_T es un subespacio de V. Sea $A = \{x_1, x_2, \ldots, x_p\}$ una base para N_T. Vamos a extender la base A hasta que obtengamos una nueva base $B = \{x_1, x_2, \ldots, x_p, y_{p+1}, y_{p+2}, \ldots, y_n\}$ para V. Debemos probar que $C = \{T(y_{p+1}), T(y_{p+2}), \ldots, T(y_n)\}$ es una base para R_T.

Sea $y \in R_T$; entonces, existe $x \in V$ tal que $T(x) = y$ y, por lo tanto,

$$T(x) = T(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p + \alpha_{p+1} y_{p+1} + \alpha_{p+2} y_{p+2} + \ldots + \alpha_n y_n),$$

donde $\alpha_i \in \mathbb{R}$. Como T es una transformación lineal,

$$T(x) = \alpha_1 T(x_1) + \ldots + \alpha_p T(x_p) + \alpha_{p+1} T(y_{p+1}) + \ldots + \alpha_n T(y_n)$$

y al ser

$$T(x_1) = T(x_2) = \ldots = T(x_p) = 0,$$

obtenemos que

$$T(x) = \alpha_{p+1} T(y_{p+1}) + \alpha_{p+2} T(y_{p+2}) + \ldots + \alpha_n T(y_n),$$

lo cual nos indica que $T(y_{p+1}), T(y_{p+2}), \ldots, T(y_n)$ generan a R_T.

Vamos a probar ahora que estos vectores son linealmente independientes. Sean $\beta_{p+1}, \beta_{p+2}, \ldots, \beta_n$ escalares tales que

$$\beta_{p+1}T(y_{p+1}) + \beta_{p+2}T(y_{p+2}) + \ldots + \beta_n T(y_n) = 0;$$

como T es una transformación lineal,

$$T(\beta_{p+1}y_{p+1} + \beta_{p+2}y_{p+2} + \ldots + \beta_n y_n) = 0.$$

Si tomamos

$$x = \beta_{p+1}y_{p+1} + \beta_{p+2}y_{p+2} + \ldots + \beta_n y_n, \tag{6.5.1}$$

entonces $x \in N_T$. Como A es una base para N_T

$$x = \gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_p x_p. \tag{6.5.2}$$

Para ciertos escalares $\gamma_1, \gamma_2, \ldots, \gamma_p$ de las relaciones 6.5.1 y 6.5.2 tenemos que

$$\gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_p x_p - \beta_{p+1}y_{p+1} - \beta_{p+2}y_{p+2} - \ldots - \beta_n y_n = 0;$$

pero como B es una base para V, entonces:

$$\gamma_1 = \gamma_2 = \ldots = \gamma_p = \beta_{p+1} = \beta_{p+2} = \ldots = \beta_n = 0$$

y, en consecuencia, $\{T(y_{p+1}), T(y_{p+2}), \ldots, T(y_n)\}$ son linealmente independientes. Por lo tanto, $C = \{T(y_{p+1}), T(y_{p+2}), \ldots, T(y_n)\}$ es una base para R_T y, además, $dim\, R_T = n - p$, con lo cual

$$dim\, V = n = p + (n - p) = dim\, N_T + dim\, R_T.$$

Aplicando los dos teoremas anteriores se tiene el siguiente corolario.

Corolario 6.5.1. *Sea $T : V \to W$ una transformación lineal en la cual $dim\, V = n$; T es uno a uno si y solo si $dim\, R_T = n$.*

Teorema 6.5.3. *Si $B = \{x_1, x_2, \ldots, x_n\}$ es una base para V, entonces T es inyectiva si y solo si $T(x_1), T(x_2), \ldots, T(x_n)$ es una base para $T(V)$; es decir, si y solo si $T(B)$ es un conjunto linealmente independiente de vectores en W.*

Demostración. Dada una base $B = \{x_1, x_2, \ldots, x_n\}$, como $T(B)$ es un conjunto de generadores que consta de n vectores de $T(V)$, entonces:

$$\begin{aligned} T(B)\,\text{es base de}\,T(V) &\Leftrightarrow T(B)\,\text{es linealmente independiente} \\ &\Leftrightarrow \text{rango}\ T(B) = n \\ &\Leftrightarrow dim\,T(V) = n = dim\,V \end{aligned}$$

y, por lo tanto, T es inyectiva.

Ejemplo 6.5.2. Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ una transformación lineal tal que

$$T\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad T\begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \text{ y } T\begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Halle N_T y R_T, y determine si T es o no una transformación inyectiva.

Solución

Si consideramos el conjunto de vectores

$$B = \left\{ \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} \right\},$$

obtenemos que

$$det\begin{bmatrix} 0 & -4 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 4 \end{bmatrix} = 1,$$

lo cual nos muestra que los tres vectores son linealmente independientes y, por lo tanto, constituyen una base para $\mathbb{R}^3$. Entonces, según el teorema 6.4.2, el conjunto

$$C = \left\{ \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

es un generador para R_T.

Pero claramente, este conjunto es linealmente dependiente y puede verificarse con facilidad que

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 4 \\ 3 \end{bmatrix} + \frac{1}{8} \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

de tal modo que

$$R_T = gen \left\{ \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\}.$$

Como estos dos últimos vectores son linealmente independientes, la base para R_T está dada por estos dos vectores; luego, $dim\, R(T) = 2$.

Para encontrar R_T, debemos hallar primero el valor de $T(x, y, z)$ para un vector cualquiera (x, y, z) en $\mathbb{R}^3$. En efecto, si (x, y, z) es un vector en $\mathbb{R}^3$, entonces

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} + \beta \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix},$$

con $\alpha, \beta, \gamma \in \mathbb{R}$. De esta forma,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\alpha \end{bmatrix} + \begin{bmatrix} -4\beta \\ \beta \\ 0 \end{bmatrix} + \begin{bmatrix} \gamma \\ 0 \\ 4\gamma \end{bmatrix} = \begin{bmatrix} -4\beta + \gamma \\ \beta \\ -\alpha + 4\gamma \end{bmatrix}.$$

Por lo tanto, el sistema de ecuaciones que se obtiene esta dado por

$$
\begin{aligned}
-4\beta+\gamma &= x\\
\beta &= y\\
-\alpha+4\gamma &= z.
\end{aligned}
$$

Al resolver este sistema lineal para α, β y γ, conseguimos

$$
\begin{aligned}
\alpha &= 4x+16y-z\\
\beta &= y\\
\gamma &= x+4y,
\end{aligned}
$$

y de esta forma obtenemos que

$$
\begin{aligned}
T\begin{bmatrix} x\\ y\\ z\end{bmatrix} &= (4x+16y-z)\,T\begin{bmatrix} 0\\ 0\\ -1\end{bmatrix}+y\,T\begin{bmatrix} -4\\ 1\\ 0\end{bmatrix}+(x+4y)\,T\begin{bmatrix} 1\\ 0\\ 4\end{bmatrix}\\
&= (4x+16y-z)\begin{bmatrix} 4\\ 3\end{bmatrix}+y\begin{bmatrix} 0\\ 2\end{bmatrix}+(x+4y)\begin{bmatrix} 1\\ 1\end{bmatrix}\\
&= \begin{bmatrix} 16x+64y-4z+x+4y\\ 12x+48y-3z+2y+x+4y\end{bmatrix}\\
&= \begin{bmatrix} 17x+68y-4z\\ 13x+54y-3z\end{bmatrix}.
\end{aligned}
$$

Ahora, podemos hallar N_T teniendo en cuenta que $(x, y, z) \in N_T$ si se cumple que $T(x, y, z) = (0, 0)$, o sea que

$$
\begin{bmatrix} 17x+68y-4z\\ 13x+54y-3z\end{bmatrix}=\begin{bmatrix} 0\\ 0\end{bmatrix}.
$$

Si utilizamos la eliminación de Gauss-Jordan, obtenemos que la forma escalonada reducida de la matriz de coeficientes de dicho sistema está dada por

$$
\begin{bmatrix} 1 & 0 & -\frac{6}{17}\\ 0 & 0 & \frac{1}{34}\end{bmatrix}.
$$

Entonces,

$$x = \frac{6}{17}z \quad \text{y} \quad y = -\frac{1}{34}z,$$

donde z es un parámetro y, en consecuencia,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{34}\begin{bmatrix} 12z \\ -z \\ 34z \end{bmatrix} = \frac{z}{34}\begin{bmatrix} 12 \\ -1 \\ 34 \end{bmatrix} = t\begin{bmatrix} 12 \\ -1 \\ 34 \end{bmatrix};$$

esto es,

$$N_T = gen\left\{\begin{bmatrix} 12 \\ -1 \\ 34 \end{bmatrix}\right\},$$

o sea que $dim\, N_T = 1$. Los valores calculados comprueban que $dim\, \mathbb{R}^3 = dim\, N_T + dim\, R_T$, y que T no es inyectiva.

6.6 Isomorfismos

Otra de las nociones importantes en el caso de las aplicaciones es la de *sobreyectividad*. Vamos a recordar esta definición.

Definición 6.6.1 (**Aplicación sobreyectiva**). *Dada una aplicación $T : V \to W$, se dice que es sobreyectiva cuando se cumple que, dado un y cualquiera en W, existe un x en V tal que $T(x) = y$.*

Esta definición significa que $T(V) = W$, es decir, el conjunto de imágenes es igual al conjunto de llegada.

Ejemplo 6.6.1. Sea $f : \mathbb{R} \to \mathbb{R}^+$, donde $f(x) = e^x$; esta función es sobreyectiva.

Ejemplo 6.6.2. Sea $f : \mathbb{R} \to \mathbb{R}$, donde $f(x) = e^x$; esta función no es sobreyectiva, puesto que si y es negativo, no existe ningún x tal que $e^x = y$.

Definición 6.6.2 (Aplicación biyectiva). *Si tenemos una aplicación $T : V \to W$, decimos que es biyectiva si cumple que es inyectiva y sobreyectiva.*

Ejemplo 6.6.3. Sea $f : \mathbb{R} \to \mathbb{R}^+$, donde $f(x) = e^x$; esta función es sobreyectiva e inyectiva por lo que, en consecuencia, es biyectiva.

Ejemplo 6.6.4. Sea $f : [a, b] \to [c, d]$, donde

$$f(x) = \frac{d-c}{b-a}(x-a) + c, \qquad a \neq b.$$

Esta es una función inyectiva, puesto que si $x_1 \neq x_2$, entonces $f(x_1) \neq f(x_2)$. Además, dado y en $[c, d]$, existe

$$x = \frac{y(a-b) - ad + bc}{c-d}$$

en $[a, b]$ tal que $f(x) = y$, con lo cual se prueba que es sobreyectiva. Por lo tanto, es una biyección entre $[a, b]$ y $[c, d]$.

Para aplicaciones que sean lineales, tenemos la siguiente definición.

Definición 6.6.3. *Sea $T : V \to W$ una transformación lineal entre dos espacios vectoriales V y W. Se dice que T es un isomorfismo si T es biyectivo; esto es, T es inyectivo y sobreyectivo. En este caso, los espacios V y W son isomorfos.*

Si la transformación va de V sobre sí mismo, hablamos de un *automorfismo.*

El siguiente teorema nos proporciona una serie de resultados importantes acerca de los isomorfismos.

Teorema 6.6.1. *Sean V, W y H espacios vectoriales, entonces, se cumple que:*

1. *Sea $T : V \to W$ una transformación lineal entre dos espacios vectoriales V y W. T es un isomorfismo si y solo si $R_T = W$ y $N_T = \{0\}$.*

2. *Sean $T : V \to W$ una transformación lineal entre dos espacios vectoriales V y W, y la dimensión de V finita; entonces, T es un isomorfismo si y solo si $dim\, V = dim\, W$.*

3. *Si V es un espacio de dimensión finita, la transformación $T : V \to V$ es un automorfismo si y solo si T es inyectiva o si y solo si T es sobreyectiva.*

4. *Si $T : V \to W$ es un isomorfismo, $T^{-1} : W \to V$ es también un isomorfismo.*

5. *Si $T : V \to W$ y $S : W \to H$ son dos isomorfismos, $S \circ T$ es un isomorfismo.*

Demostración. Sean V, W y H espacios vectoriales.

1. Sabemos que T es isomorfismo si y solo si es inyectiva y sobreyectiva. Pero T es sobreyectiva si y solo si $R_T = W$; además, T es uno a uno si y solo si $N_T = \{0\}$. Por lo tanto, T es un isomorfismo si y solo si $R_T = W$ y $N_T = \{0\}$.

2. De nuevo, sabemos que T es un isomorfismo si y solo si es uno a uno y sobreyectiva; pero T es sobreyectiva si y solo si $R_T = W$ y, por lo tanto, $dim\, R_T = dim\, W$. Ahora, T es uno a uno si y solo si $dim\, N_T = 0$ y, en consecuencia, $dim\, V = dim\, R_T$, lo que implica que T es isomorfismo si y solo si $dim\, V = dim\, W$.

3. Del numeral anterior, si $W = V$, entonces T es un isomorfismo si y solo si $dim\, V = R_T$; por lo tanto, T es sobreyectiva. Como $dim\, V = dim\, R_T$, entonces $N_T = \{0\}$ y, por lo tanto, T es uno a uno.

4. Sabemos que si T es una aplicación biyectiva, su inversa existe y también es biyectiva. Solo falta verificar que $T^{-1}; W \to V$ es una transformación lineal, es decir,

$$T^{-1}(\alpha x + \beta y) = \alpha T^{-1}(x) + \beta T^{-1}(y)$$

para vectores $x, y \in W$ y escalares $\alpha, \beta \in \mathbb{R}$. Si aplicamos la definición de la inversa, se cumple la relación

$$T(T^{-1}(\alpha x + \beta y)) = \alpha x + \beta y, \tag{6.6.1}$$

mientras que, por ser T una transformación lineal,

$$\begin{aligned} T(\alpha T^{-1}(x) + \beta T^{-1}(y)) &= \alpha T(T^{-1}(x)) + \beta T(T^{-1}(y)) \\ &= \alpha x + \beta y. \end{aligned} \tag{6.6.2}$$

Si tenemos en cuenta que la transformación es inyectiva, a partir de las ecuaciones 6.6.1 y 6.6.2 concluimos que

$$T^{-1}(\alpha x + \beta y) = \alpha T^{-1}(x) + \beta T^{-1}(y),$$

lo cual significa que T^{-1} es también una transformación lineal.

5. En la Figura 6.6 se puede observar el proceso de composición de isomorfismos.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[dr, "S \circ T"'] & W \arrow[d, "S"] \\ & H \end{tikzcd}$$

Figura 6.6. Composición de isomorfismos

Es claro que la composición de dos transformaciones lineales es también una transformación lineal. Ahora probemos que esta composición es inyectiva y sobreyectiva. Para esto, sean x y y dos elementos en V tales que $x \neq y$. Como T es inyectivo,

$$T(x) \neq T(y)$$

y como S es también inyectivo, entonces

$$S(T(x)) \neq S(T(y)),$$

lo cual significa que

$$S \circ T(x) \neq S \circ T(y).$$

Esto prueba que la composición es inyectiva.

Solo resta demostrar que la transformación $S \circ T$ es sobreyectiva. Sea z en H; como S es una transformación sobreyectiva, existe un elemento y en W tal que $S(y) = z$; pero, al tener y en W y considerando que T es una transformación de V en W, existe un x en V tal que $T(x) = y$. En consecuencia, dado z en H, existe un x en V tal que $S(T(x)) = z$, es decir, $S \circ T(x) = z$, probando así que $S \circ T$ es una aplicación sobreyectiva. Esto último muestra que la composición $S \circ T$ es un isomorfismo.

6.7 Forma matricial de una transformación lineal

En esta sección mostramos que cualquier transformación lineal $T : V \to W$ de un espacio vectorial V de dimensión finita en otro espacio vectorial W de dimensión finita, se puede representar por medio de una matriz. Dicha representación se hace con respecto a dos bases B y B' para los espacios V y W respectivamente. Una vez escogidas dichas bases, como bases ordenadas, la representación matricial de la transformación será única.

También nos referimos al álgebra de las transformaciones lineales, suma, producto y múltiplos escalares; dichas operaciones reflejan las operaciones correspondientes con matrices. Por último fijamos nuestra atención en las transformaciones lineales invertibles.

Para iniciar, necesitamos la siguiente definición.

Definición 6.7.1. *Sea $B = \{v_1, v_2, \ldots, v_n\}$ una base para el espacio vectorial V y sea $v \in V$. Si*

$$v = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n,$$

entonces el vector de coordenadas de v respecto a la base B está dado por el vector columna

$$v_B = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix},$$

donde $\alpha_1, \alpha_2, \ldots, \alpha_n$ son escalares.

Ejemplo 6.7.1. Sea el vector $v = (3, 4) \in \mathbb{R}^2$. Hallar su vector de coordenadas x_B respecto a la base $B = \{(1, 2), (-1, 0)\}$.

Solución

En este caso, debemos calcular escalares α y β tales que

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \end{bmatrix};$$

esto es equivalente al sistema

$$\begin{aligned} \alpha - \beta &= 3 \\ 2\alpha &= 4, \end{aligned}$$

cuya solución es $\alpha = 2$ y $\beta = -1$; por lo tanto, el vector de coordenadas respecto a la base B está dado por

$$v_b = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Ejemplo 6.7.2. Si $B = \{e_1, e_2, \ldots, e_n\}$ es la base natural de $\mathbb{R}^n$ y si $x = (x_1, x_2, \ldots, x_n)^T$ es un vector en $\mathbb{R}^n$, entonces

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

y, por lo tanto,

$$x_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Teorema 6.7.1. *Sea $B = \{v_1, v_2, \ldots, v_n\}$ una base para el espacio vectorial $\mathbb{R}^n$.*

1. *La transformación*

$$\begin{aligned} C_B : \mathbb{R}^n &\to \mathbb{R}^n \\ x &\to x_B \end{aligned}$$

   es tal que:
   a) C_B es una transformación lineal.
   b) C_B es uno a uno y sobreyectiva.

2. *Los vectores $y_1, y_2, \ldots, y_r$ en $\mathbb{R}^n$ son linealmente independientes si y solo si sus vectores coordenados respecto a la base B de $\mathbb{R}^n$ son linealmente independientes.*

Demostración. Sean x y y dos vectores de $\mathbb{R}^n$, con coordenadas

$$x_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{y} \quad y_B = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix};$$

es decir,

$$\begin{aligned} x &= x_1 v_1 + x_2 v_2 + \ldots + x_n v_n \\ y &= y_1 v_1 + y_2 v_2 + \ldots + y_n v_n. \end{aligned}$$

Si α y β son escalares en $\mathbb{R}$, entonces

$$\begin{aligned} \alpha x + \beta y &= \alpha(x_1v_1 + x_2v_2 + \ldots + x_nv_n) + \beta(y_1v_1 + y_2v_2 + \ldots + y_nv_n) \\ &= (\alpha x_1 + \beta y_1)v_1 + (\alpha x_2 + \beta y_2)v_2 + \ldots + (\alpha x_n + \beta y_n)v_n), \end{aligned}$$

de donde concluimos que

$$(\alpha x+\beta y)_B = \begin{bmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \\ \vdots \\ \alpha x_n + \beta y_n \end{bmatrix} = \alpha \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \beta \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \alpha x_B + \beta y_B,$$

lo cual muestra que C_B es una transformación lineal.

Veamos ahora que esta transformación es uno a uno. En efecto, como tenemos que

$$x_B = y_B \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

entonces $x_i = y_i$, con $i = 1, \ldots, n$; por lo tanto,

$$x_1v_1 + x_2v_2 + \ldots + x_nv_n = y_1v_1 + y_2v_2 + \ldots + y_nv_n$$

y, en consecuencia, $x = y$.

Nos falta probar que C_B es una transformación sobreyectiva. Sea

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n,$$

con $x_i \in \mathbb{R}$; entonces, como B es una base para $\mathbb{R}^n$,

$$x = x_1v_1 + x_2v_2 + \ldots + x_nv_n$$

es un elemento de $\mathbb{R}^n$. Así, como

$$x = x_1v_1 + x_2v_2 + \ldots + x_nv_n \in \mathbb{R}^n,$$

entonces

$$C_B(x) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

con lo cual probamos que C_B es una transformación sobreyectiva.

La prueba del segundo numeral la dejamos como ejercicio para el lector.

Observación 6.7.1. Si V es un espacio vectorial de dimensión finita, en V se puede introducir el concepto de *coordenadas* para un vector.

Observación 6.7.2. Sea V un espacio vectorial de dimensión n sobre $\mathbb{R}$ y sea $B = \{v_1, v_2, \ldots, v_n\}$ una base ordenada para el espacio vectorial V. Sabemos que si $x \in V$, entonces existen escalares únicos $x_1, x_2, \ldots, x_n$ tales que

$$x = x_1v_1 + x_2v_2 + \ldots + x_nv_n;$$

esto es,

$$x_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

La función

$$\begin{aligned} C_B : V &\to \mathbb{R}^n \\ x &\to x_B \end{aligned}$$

es una transformación lineal uno a uno y sobre; es decir, C_B es un isomorfismo. Esta última afirmación nos dice que el vector de coordenadas x_B es una representación del vector x en $\mathbb{R}^n$.

6.8 Cambio de coordenadas

Sean $B_1 = \{v_1, v_2, \ldots, v_n\}$ y $B_2 = \{w_1, w_2, \ldots, w_n\}$ dos bases ordenadas para $\mathbb{R}^n$ tales que

$$x = a_1v_1 + a_2v_2 + \ldots + a_nv_n$$
$$x = b_1w_1 + b_2w_2 + \ldots + b_nw_n,$$

es decir, los vectores de coordenadas respecto a estas dos bases son:

$$x_{B_1} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{y} \quad x_{B_2} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Se trata de ver ahora cómo están relacionados los vectores de coordenadas x_{B_1} y x_{B_2}.

Para esto, representemos cada vector w_j en la base B_1. Supongamos que $(w_j)_{B_1} = P_j \in \mathbb{R}^n$; entonces,

$$\begin{aligned} x_{B_1} &= (b_1w_1 + b_2w_2 + \ldots + b_nw_n)_{B_1} \\ &= (b_1w_1)_{B_1} + (b_2w_2)_{B_1} + \ldots + (b_nw_n)_{B_1} \\ &= b_1P_1 + b_2P_2 + \ldots + b_nP_n \\ &= [P_1, P_2, \ldots P_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= Px_{B_2}, \end{aligned}$$

donde P es una matriz de tamaño $n \times n$. De acá podemos afirmar que si $P = [(y_1)_{B_1}, \ldots, (y_n)_{B_1}]$, entonces

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = P \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

A esta matriz P se le conoce como la *matriz de paso* de la base B_2 a la base B_1.

Como los vectores $w_1, w_2, \ldots, w_n$ son linealmente independientes se puede probar fácilmente que los vectores $P_1, P_2, \ldots, P_n$ también lo son y, por lo tanto, podemos concluir que la matriz P es no singular; en consecuencia, obtenemos que

$$x_{B_2} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = P^{-1} x_{B_1} = P^{-1} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

donde la matriz P^{-1} es conocida como la *matriz de paso* de la base B_1 a la base B_2.

Ejemplo 6.8.1. Sean

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

una base de $\mathbb{R}^3$ y

$$B_2 = \left\{ \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

una segunda base para $\mathbb{R}^3$.

Sea el vector $x = (1, 2, 3) \in \mathbb{R}^3$. Se puede ver que

$$x_{B_1} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \quad \text{y} \quad x_{B_2} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix};$$

esto es claro, ya que

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

mientras que

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = -1 \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Ahora, como

$$\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

entonces

$$P_1 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}_{B_1} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}.$$

Además,

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

y también

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix};$$

con ello,

$$P_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}_{B_1} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad P_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{B_1} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}.$$

De esta forma, la matriz $P = [P_1, P_2, P_3]$ es

$$P = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}.$$

Entonces,

$$Px_{B_2} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = x_{B_1};$$

igualmente, puede verificarse que

$$\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = P^{-1} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

6.9 Matriz que representa una transformación

Teniendo en cuenta la noción de *vector coordenado*, encontremos la matriz que representa una transformación lineal. A manera de ejemplo, sea la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$, donde

$$\begin{aligned} B_1 &= \left\{ v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}, \\ B_2 &= \left\{ w_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, w_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} \end{aligned}$$

y la transformación lineal está definida por

$$T(v_1) = -w_1 + 2w_2,\ T(v_2) = w_2 \text{ y } T(v_3) = 3w_1 + w_2,$$

siendo B_1 y B_2 bases para $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente.

Como, de acuerdo con la definición de la transformación, tenemos las siguientes relaciones:

$$\begin{aligned} T(v_1) &= -w_1 + 2w_2, & \text{entonces } (T(v_1))_{B_2} &= \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \\ T(v_2) &= 0w_1 + w_2, & \text{entonces } (T(v_2))_{B_2} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ y} \\ T(v_3) &= 3w_1 + w_2, & \text{entonces } (T(v_3))_{B_2} &= \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \end{aligned}$$

obtenemos la matriz

$$A = [(T(v_1))_{B_2}, (T(v_2))_{B_2}, (T(v_3))_{B_2}]$$

o lo que es equivalente:

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

Veamos que si x es cualquier vector de $\mathbb{R}^3$, se cumple que

$$A[x_{B_1}] = [T(x)]_{B_2}.$$

En efecto, sea $x \in \mathbb{R}^3$; entonces,

$$x = a_1v_1 + a_2v_2 + a_3v_3$$

y, por lo tanto,

$$A[x_{B_1}] = \begin{bmatrix} -1 & 0 & 3 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

Al realizar el producto obtenemos que

$$A[x_{B_1}] = \begin{bmatrix} -a_1 + 3a_3 \\ 2a_1 + a_2 + a_3 \end{bmatrix};$$

pero si ahora tenemos en cuenta que T es una transformación lineal,

$$\begin{aligned} T(X) &= a_1T(v_1) + a_2T(v_2) + a_3T(v_3) \\ &= a_1(-w_1 + 2w_2) + a_2(0w_1 + w_2) + a_3(3w_1 + w_2) \\ &= (-a_1 + 3a_3)w_1 + (2a_1 + a_2 + a_3)w_2; \end{aligned}$$

por lo tanto,

$$[T(x)]_{B_2} = \begin{bmatrix} -a_1 + 3a_3 \\ 2a_1 + a_2 + a_3 \end{bmatrix} = A[x]_{B_2}.$$

En este caso se dice que A es la matriz que representa la transformación con respecto a las bases B_1 y B_2.

Definición 6.9.1. *Sean* $T : V \to W$ *una transformación lineal y* $B_1 = \{v_1, v_2, \ldots, v_n\}$ *y* $B_2 = \{w_1, w_2, \ldots, w_m\}$ *dos bases ordenadas para* V *y* W *respectivamente.*

Si

$$T(v_j) = a_{1j}w_1 + a_{2j}w_2 + \ldots + a_{mj}w_m,$$

con $j = 1, 2, \ldots n$, *entonces la matriz de la transformación con respecto a las bases* B_1 *y* B_2, *denotada por* $[T]_{B_1B_2}$, *se define por*

$$[T]_{B_1B_2} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix};$$

es decir, la matriz de la transformación T *respecto a las bases dadas está dada por la expresión*

$$[T]_{B_1B_2} = [(T(v_1))_{B_2}, (T(v_2))_{B_2}, \ldots, (T(v_n))_{B_2}].$$

Observación 6.9.1. Es claro que si $T : V \to W$ es una transformación lineal en la cual $dim\, V = n$ y $dim\, W = m$, y si B_1 y B_2 son bases ordenadas para V y W respectivamente, entonces $[T]_{B_1B_2}$ es una matriz de tamaño $m \times n$.

Observación 6.9.2. Si en la observación anterior $V = W$, y B_1 es una base de V, entonces la matriz $[T]_{B_1B_2}$ se denotará simplemente como $[T]_{B_1}$.

El siguiente teorema es una generalización del ejercicio anterior.

Teorema 6.9.1. *Sea* $T : V \to W$ *una transformación lineal y sean* $B_1 = \{v_1, v_2, \ldots, v_n\}$ *y* $B_2 = \{w_1, w_2, \ldots, w_m\}$ *dos bases ordenadas para* V *y* W *respectivamente. Si*

$$[T]_{B_1B_2} = [a_{ij}]_{m\times n},$$

entonces para cualquier $x \in V$ *se tiene que*

$$[T(x)]_{B_2} = A[x]_{B_1}.$$

Demostración. Sabemos que

$$[T]_{B_1B_2} = [a_{ij}]_{m\times n}.$$

Además,

$$[T(v_j)]_{B_2} = a_{1j}w_1 + a_{2j}w_2 + \ldots + a_{mj}w_m = A^j,$$

donde A^J es la columna j de la matriz A.

Si $x \in V$, entonces, al ser

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n,$$

tenemos que

$$[x]_{B_1} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

y, en consecuencia,

$$T(x) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \ldots + \alpha_n T(v_n).$$

Al aplicar las propiedades de la transformación lineal, tenemos:

$$\begin{aligned}
T(x) &= \alpha_1 \sum_{i=1}^{m} a_{i1}w_i + \alpha_2 \sum_{i=1}^{m} a_{i2}w_i + \ldots + \alpha_n \sum_{i=1}^{m} a_{in}w_i. \\
&= \sum_{i=1}^{m} \alpha_1 a_{i1}w_i + \sum_{i=1}^{m} \alpha_2 a_{i2}w_i + \ldots + \sum_{i=1}^{m} \alpha_n a_{in}w_i \\
&= \sum_{i=1}^{m} (\alpha_1 a_{i1}w_i + \alpha_2 a_{i2}w_i + \ldots + \alpha_n a_{in}w_i) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_j a_{ij} w_i \right).
\end{aligned}$$

Al intercambiar los índices de las sumatorias, obtenemos

$$\begin{aligned}
T(X) &= \sum_{j=1}^{n}\left(\sum_{i=1}^{m}\alpha_j a_{ij} w_i\right)\\
&= \sum_{j=1}^{n}(\alpha_j a_{1j} w_1 + \alpha_j a_{2j} w_2 + \ldots + \alpha_j a_{mj} w_m)\\
&= \left(\sum_{j=1}^{n}\alpha_j a_{1j}\right) w_1 + \left(\sum_{j=1}^{n}\alpha_j a_{2j}\right) w_2 + \ldots + \left(\sum_{j=1}^{n}\alpha_j a_{mj}\right) w_m,
\end{aligned}$$

de donde

$$[T(x)]_{B_2} = \begin{bmatrix} \sum_{j=1}^{n}\alpha_j a_{1j} \\ \sum_{j=1}^{n}\alpha_j a_{2j} \\ \vdots \\ \sum_{j=1}^{n}\alpha_j a_{mj} \end{bmatrix} = \begin{bmatrix} A_1[x]_{B_1}^t \\ A_2[x]_{B_1}^t \\ \vdots \\ A_m[x]_{B_1}^t \end{bmatrix} = A[x]_{B_1},$$

siendo A_i la fila i de la matriz A, con $i = 1, \ldots, m$.

6.10 Álgebra de las transformaciones lineales

En esta sección suponemos que la suma, la multiplicación y la composición de funciones reales son conocidas, al igual que la multiplicación de una función por un escalar.

Definimos acá las operaciones anteriores para el caso de las transformaciones lineales y analizamos cómo se traducen estas operaciones en el álgebra de las matrices.

Definición 6.10.1. *Sean $T : V \to W$ y $S : V \to W$ dos transformaciones lineales; entonces,*

1. *La suma $(T + S) : V \to W$ es la transformación lineal dada por $(T + S)(x) = T(x) + S(x)$, con $x \in V$.*

2. *Si α es un escalar, el múltiplo escalar $\alpha T : V \to W$ es la transformación lineal dada por $(\alpha T)(x) = \alpha T(x)$ para cualquier $x \in V$.*

Ahora demostremos que si A y B son las representaciones matriciales de las transformaciones T y S respectivamente, entonces las matrices de $T + S$ y αT son $A + B$ y αB respectivamente.

Teorema 6.10.1. *Sean $T : V \to W$ y $S : V \to W$ dos transformaciones lineales, donde V y W son dos espacios vectoriales tales que $\dim V = n$ y $\dim W = m$. Sean, además, $B_1 = \{v_1, v_2, \ldots, v_n\}$ y $B_2 = \{w_1, w_2, \ldots, w_m\}$ bases ordenadas para V y W respectivamente. Si*

$$[T]_{B_1B_2} = A \quad y \ [S]_{B_1B_2} = B$$

y $\alpha \in \mathbb{R}$, entonces

$$[T + S]_{B_1B_2} = A + B \ y \ [\alpha T]_{B_1B_2} = \alpha A.$$

Demostración. Sean

$$\begin{aligned} T(v_j) &= a_{1j}w_1 + a_{2j}w_2 + \ldots, a_{mj}w_m \text{ y} \\ S(v_j) &= b_{1j}w_1 + b_{2j}w_2 + \ldots, b_{mj}w_m, \end{aligned}$$

esto es,

$$[T(v_j)]_{B_2} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \text{ y } [S(v_j)]_{B_2} = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{bmatrix},$$

con $j = 1, 2, \ldots n$. Entonces,

$$\begin{aligned} (T + S)(v_j) &= T(v_j) + S(v_j) \\ &= (a_{1j}w_1 + a_{2j}w_2 + \ldots, a_{mj}w_m) \\ &\quad +(b_{1j}w_1 + b_{2j}w_2 + \ldots, b_{mj}w_m) \\ &= (a_{1j} + b_{1j})w_1 + (a_{2j} + b_{2j})w_2 + \ldots + (a_{1j} + b_{mj})w_m \end{aligned}$$

y, por lo tanto,

$$[(T+S)(v_j)]_{B_2} = \begin{bmatrix} a_{1j}+b_{1j} \\ a_{2j}+b_{2j} \\ \vdots \\ a_{mj}+b_{mj} \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} + \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{bmatrix}.$$

Como esta relación es válida para todas las columnas,

$$[T+S]_{B_1B_2} = [T]_{B_1B_2} + [S]_{B_1B_2}.$$

La segunda parte del teorema la dejamos como ejercicio.

Ejemplo 6.10.1. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y-x \\ x \\ x+y \end{bmatrix}$$

y sea $S : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$S\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ -y \\ x \end{bmatrix}.$$

1. Hallar $[T+S]_{B_1B_2}$;

2. Hallar $[\alpha T]_{B_1B_2}$;

donde α es un escalar y B_1 y B_2 son las bases naturales de $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente.

Solución

1. En el caso de $T+S$ tenemos que

$$(T+S)\begin{bmatrix} x \\ y \end{bmatrix} = T\begin{bmatrix} x \\ y \end{bmatrix} + S\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y-x \\ x \\ x+y \end{bmatrix} + \begin{bmatrix} 0 \\ -y \\ x \end{bmatrix};$$

por lo tanto,

$$(T+S)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y-x \\ x-y \\ x \end{bmatrix}.$$

Luego,

$$(T+S)\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = -1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ y}$$

$$(T+S)\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 1\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

En consecuencia,

$$[T+S]_{B_1B_2} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix}.$$

Ahora

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = -1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ y}$$

$$T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix};$$

entonces,

$$[T]_{B_1B_2} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Además,

$$S\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\1\end{bmatrix} \text{ y } S\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}0\\-1\\0\end{bmatrix},$$

con lo cual obtenemos que

$$[S]_{B_1B_2}=\begin{bmatrix}0&0\\0&-1\\1&0\end{bmatrix}.$$

De esta forma se ve fácilmente que

$$\begin{aligned}[T]_{B_1B_2}+[S]_{B_1B_2} &= \begin{bmatrix}-1&1\\1&0\\1&1\end{bmatrix}+\begin{bmatrix}0&0\\0&-1\\1&0\end{bmatrix}\\ &= \begin{bmatrix}-1&1\\1&-1\\2&1\end{bmatrix}=[T+S]_{B_1B_2}.\end{aligned}$$

2. Sea $\alpha\in\mathbb{R}$; entonces,

$$(\alpha T)\begin{bmatrix}x\\y\end{bmatrix}=\alpha T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\alpha y-\alpha x\\\alpha x\\\alpha x+\alpha y\end{bmatrix}$$

y, por lo tanto,

$$(\alpha T)\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}-\alpha\\\alpha\\\alpha\end{bmatrix} \text{ y } (\alpha T)\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}\alpha\\0\\\alpha\end{bmatrix},$$

de donde

$$[\alpha T]_{B_1B_2}=\begin{bmatrix}-\alpha&\alpha\\\alpha&0\\\alpha&\alpha\end{bmatrix}=\alpha\begin{bmatrix}-1&1\\1&0\\1&1\end{bmatrix}=\alpha[T]_{B_1B_2},$$

como lo afirmamos en el teorema.

Para el siguiente teorema recordemos que si $S : V \to W$ y $T : W \to U$ son transformaciones lineales, entonces la composición $S \circ T : V \to U$ es la transformación lineal definida por

$$(T \circ S)(x) = T(S(x)).$$

Teorema 6.10.2. *Sean $S : V \to W$ y $T : W \to U$ dos transformaciones lineales, y sean las bases ordenadas de V, W y U formadas por $B_1 = \{v_1, v_2, \ldots, v_n\}$, $B_2 = \{w_1, w_2, \ldots, w_m\}$ y $B_3 = \{u_1, u_2, \ldots, u_p\}$ respectivamente. Si $[S]_{B_1 B_2} = B$ y si $[T]_{B_2 B_3} = A$, entonces $[T \circ S]_{B_1 B_3} = AB$.*

Demostración. Sabemos que $[S]_{B_1 B_2} = B = [b_{ij}]$ es una matriz de tamaño $m \times n$; entonces, $(S(v_j))_{B_2}$ es la columna j de B. Así,

$$S(v_j) = b_{1j}w_1 + b_{2j}w_2 + \ldots + b_{mj}w_m = v'_j$$

y, por lo tanto,

$$\begin{aligned}
(T \circ S)(v_j) &= T(S(v_j)) \\
&= T(v'_j) \\
&= T(b_{1j}w_1 + b_{2j}w_2 + \ldots + b_{mj}w_m) \\
&= b_{1j}T(w_1) + b_{2j}T(w_2) + \ldots + b_{mj}T(w_m).
\end{aligned}$$

También se sabe que $[T]_{B_2 B_3} = A$; entonces,

$$[T(w_k)]_{B_3} = A^{(k)},$$

de modo que

$$T(w_k) = a_{1k}u_1 + a_{2k}u_2 + \ldots + a_{pk}u_p = w'_k,$$

para $k = 1, 2, \ldots, m$. De esta manera,

$$\begin{aligned}
(T \circ S)(v_j) &= b_{1j}w'_1 + b_{2j}w'_2 + \ldots + b_{mj}w'_m \\
&= b_{1j}(a_{11}u_1 + \ldots + a_{p1}u_p) + \ldots + b_{mj}(a_{1m}u_1 + \ldots + a_{pm}u_p),
\end{aligned}$$

donde podemos observar que el coeficiente de u_i es

$$b_{1j}a_{11} + b_{2j}a_{12} + \ldots + b_{mj}a_{1m} = A_1^T B^{(j)}.$$

Así, el vector de coordenadas de la transformación $(T \circ S)(v_j)$ respecto a la base B_3 es

$$\begin{bmatrix} A_1^T B^{(j)} \\ A_2^T B^{(j)} \\ \vdots \\ A_p^T B^{(j)} \end{bmatrix},$$

con $j = 1, 2, \ldots, n$. Esto último es la columna j de AB y, por lo tanto, $[T \circ S]_{B_1 B_3} = AB$.

El siguiente ejemplo ilustra la aplicación de este teorema.

Ejemplo 6.10.2. Sean $S : \mathbb{R}^4 \to \mathbb{R}^2$ y $T : \mathbb{R}^2 \to \mathbb{R}^3$ dos transformaciones lineales definidas por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ -y \\ x-y \end{bmatrix} \text{ y } S\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x-y+w \\ 2y-z \end{bmatrix}.$$

Considerando B_1, B_2 y B_3 como las bases naturales para $\mathbb{R}^4$, $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente, se puede verificar con facilidad que

$$[T]_{B_2 B_3} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \\ 1 & -1 \end{bmatrix} \text{ y } [S]_{B_1 B_2} = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Además, $T \circ S : \mathbb{R}^4 \to \mathbb{R}^3$ está definida por

$$
\begin{aligned}
(T \circ S) \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} &= T \left(S \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \right) \\
&= T \begin{bmatrix} x - y + w \\ 2y - z \end{bmatrix} = \begin{bmatrix} 2x - 2y + 2w \\ -2y + z \\ x - 3y + w + z \end{bmatrix}
\end{aligned}
$$

y

$$
[T \circ S]_{B_1 B_2} = \begin{bmatrix} 2 & -2 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 1 & -3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix},
$$

con lo cual verificamos que $[T \circ S]_{B_1 B_3} = [T]_{B_2 B_3} [S]_{B_1 B_2} = AB$.

6.11 Transformaciones lineales invertibles

En esta sección nos referimos a las transformaciones lineales de la forma $T : V \to V$, donde V es un espacio vectorial de dimensión finita. Dichas transformaciones son llamadas también *operadores lineales*. Podemos ver con facilidad que la matriz asociada a la transformación con respecto a una base ordenada para V es una matriz cuadrada y, por lo tanto, es posible pensar en su inversa.

Debido a la correspondencia existente entre las transformaciones lineales y las matrices, tenemos que si $I : V \to V$ es una transformación lineal tal que, dada la base $B_1 = \{v_1, v_2, \ldots, v_n\}$, se cumple que $I(v_j) = v_j$, entonces $[I]_{B_1} = I_n$, esto es, la matriz identidad de tamaño $n \times n$.

A continuación definimos el concepto de *operador lineal invertible o no singular*.

Definición 6.11.1. *Sea* V *un espacio vectorial de dimensión* n. *Sea* $L(V) = L$ *el conjunto de todos los operadores lineales* $T : V \to V$.

1. *El operador lineal* $I : V \to V$ *tal que* $I(v) = v$ *para cualquier* $V \in V$, *es llamado el operador lineal identidad.*

2. *Un operador lineal* $T : V \to V$ *se dice que es invertible o no singular si existe* $S \in L$ *tal que* $S \circ T = T \circ S = I$. *En este caso,* S *es el inverso del operador* T *y se denota con* T^{-1}.

Teorema 6.11.1. *Si el operador lineal* $T : V \to V$ *es no singular, entonces* T^{-1} *es un operador lineal.*

Demostración. Sabemos que $T^{-1} : V \to V$. Si $x, y \in V$, entonces existen $z, w \in V$ tales que $z = T^{-1}(x)$ y $w = T^{-1}(y)$. De esta manera,

$$\begin{aligned} T(z) = T(T^{-1}(x)) &= I(x) = x, \\ T(w) = T(T^{-1}(y)) &= I(y) = y, \end{aligned}$$

de donde concluimos que

$$\begin{aligned} T^{-1}(x + y) &= T^{-1}(T(z) + T(w)) \\ &= T^{-1}(T(z + w)), \quad \text{pues T es lineal} \end{aligned}$$

$$\begin{aligned} T^{-1}(x + y) &= (T^{-1} \circ T)(z + w) \\ &= I(z + w) \\ &= z + w \\ &= T^{-1}(x) + T^{-1}(y). \end{aligned}$$

Ahora, supongamos que $\alpha \in \mathbb{R}$ y sea $x \in V$. Si tomamos $z = T^{-1}(x)$, entonces $T(z) = x$ y, por lo tanto,

$$\begin{aligned}
T^{-1}(\alpha x) &= T^{-1}(\alpha T(z)) \\
&= T^{-1}(T(\alpha z)), \quad \text{pues T es lineal} \\
&= (T^{-1} \circ T)(\alpha z) \\
&= I(\alpha z) \\
&= \alpha z \\
&= \alpha T^{-1}(x),
\end{aligned}$$

con lo cual concluimos que el inverso de un operador lineal es también un operador lineal.

Dada la relación existente entre transformaciones lineales y matrices, debemos esperar que no todo operador lineal posea su inverso. Aquí se hace necesario establecer condiciones que nos permitan decidir cuándo un operador lineal es no singular. El siguiente teorema resuelve esta pregunta en términos de la singularidad o no singularidad de la matriz de la transformación con respecto a una base ordenada.

Teorema 6.11.2. *Sea $T : V \to V$ un operador lineal y sea $B = \{v_1, v_2, \dots, v_n\}$ una base para V; entonces, T es no singular si y solo si $[T]_B$ es una matriz no singular. Además,*

$$[T^{-1}]_B = ([T]_B)^{-1}$$

cuando T^{-1} existe.

Demostración. Sea $A = [T]_B$ y supongamos que A es no singular. Sea $S : V \to V$ una transformación lineal tal que $A^{-1} = [S]_B$. Entonces, por el teorema 6.10.2, se tiene que

$$[S \circ T]_B = [S]_B[T]_B = A^{-1}A = I_n.$$

De manera similar,

$$[T \circ S]_B = [T]_B [S]_B = AA^{-1} = I_n.$$

Por lo tanto, para cualquier $x \in V$,

$$(S \circ T)(x) = x = (T \circ S)(x),$$

de donde

$$S = T^{-1}$$

y, en consecuencia, T es no singular.

Supongamos ahora que T es no singular; entonces, $T \circ T^{-1} = I$. De esta forma,

$$[T \circ T^{-1}]_B = I_n.$$

Pero hemos visto que la matriz de la composición de dos transformaciones lineales es el producto de las matrices que representan a cada transformación, es decir,

$$[T \circ T^{-1}]_B = [T]_B [T^{-1}]_B;$$

esto es, $[T]_B^{-1}$ existe y $[T^{-1}]_B = [T]_B^{-1}$.

Ejemplo 6.11.1. Sea la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ x - 2y \\ y + 3z \end{bmatrix}.$$

Determinar si existe o no la transformación inversa.

Solución

Si consideramos la base natural de $\mathbb{R}^3$, los valores de la transformación para los vectores de esta base están dados por

$$T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \quad T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \text{ y } T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix},$$

a partir de lo cual concluimos que

$$[T]_B = A = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 3 \end{bmatrix}.$$

Veamos ahora si A^{-1} existe. Para ello utilizamos la eliminación de Gauss-Jordan.

$$\left[\begin{array}{rrr|rrr} -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 & 1 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow[F_1 \leftrightarrow F_2]{} \left[\begin{array}{rrr|rrr} 1 & -2 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[F_2+F_1]{} \left[\begin{array}{rrr|rrr} 1 & -2 & 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 & 1 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow[-\frac{1}{2}F_2]{} \left[\begin{array}{rrr|rrr} 1 & -2 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[F_1+2F_2]{F_3-F_2} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 3 & \frac{1}{2} & \frac{1}{2} & 1 \end{array}\right] \xrightarrow[\frac{1}{3}F_3]{} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \end{array}\right].$$

De esta manera concluimos que A es no singular y, además,

$$A^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}.$$

Sea ahora $S : \mathbb{R}^3 \to \mathbb{R}^3$ un operador lineal cuya matriz con respecto a la base natural de $\mathbb{R}^3$ es A^{-1}; entonces,

$$S \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

con lo cual

$$S\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ -\frac{x}{2} - \frac{y}{2} \\ \frac{x}{6} + \frac{y}{6} + \frac{z}{3} \end{bmatrix}.$$

Por el teorema 6.10.2 sabemos que

$$[T \circ S]_B = [T]_B[S]_B = AA^{-1} = I_3;$$

de allí podemos concluir que para cualquier vector $x \in V$,

$$(S \circ T)(x) = x = (T \circ S)(x)$$

y, por lo tanto, $S = T^{-1}$.

Se puede ver fácilmente que

$$\begin{aligned}
(S \circ T)\begin{bmatrix} x \\ y \\ z \end{bmatrix} &= (T \circ S)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = S\left(T\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) \\
&= S\begin{bmatrix} -x \\ x - 2y \\ y + 3z \end{bmatrix} = \begin{bmatrix} -(-x) \\ -\left(-\frac{x}{2}\right) - \left(\frac{x-2y}{2}\right) \\ \left(-\frac{x}{6}\right) + \left(\frac{x-2y}{2}\right) + \left(\frac{y+3z}{3}\right) \end{bmatrix} \\
&= \begin{bmatrix} x \\ y \\ z \end{bmatrix}.
\end{aligned}$$

De modo similar se prueba que

$$(T \circ S)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

como lo afirmaba el teorema 6.11.2.

6.12 Ejercicios resueltos

Ejercicio 6.12.1. Sea T una transformación lineal tal que

a) $T(1,1,-1,1) = (2,0,3,2)$.

b) $T \circ T(1,1,-1,1) = (-3,5,4,2)$.

c) $T^{-1}(0,0,1,0) = \{(1,-1,a,-a)/a \in \mathbb{R}\}$.

d) $N_T = \{(x,y,z,t)/x = y = 0, z + t = 0\}$.

1. Calcule la matriz de la transformación con respecto a la base natural de $\mathbb{R}^4$.
2. Calcule el rango y la nulidad de la transformación y una base para cada uno de estos subespacios.

Solución

1. Las condiciones que nos proporciona el ejercicio nos indican que

a) $T(1,1,-1,1) = (2,0,3,2)$.

b) $T \circ T(1,1,-1,1) = T(2,0,3,2) = (-3,5,4,2)$.

c) $T(1,-1,0,0) = (0,0,1,0)$.

d) $T(0,0,1,-1) = (0,0,0,0)$.

Como $A = \{(1,1,-1,1),(2,0,3,2),(1,-1,0,0),(0,0,1,-1)\}$ es una base para $\mathbb{R}^4$, entonces la matriz de la transformación con respecto a las bases A y $E = \{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}$ está dada por

$$M[T]_{AE} = \begin{bmatrix} 2 & -3 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ 2 & 2 & 0 & 0 \end{bmatrix}.$$

Ahora, de las condiciones a), c) y d) tenemos

$$\begin{aligned} T(1,1,-1,1)+T(1,-1,0,0)+\ T(0,0,1,-1) &= (2,0,3,2) \\ &\quad +(0,0,1,0)+(0,0,0,0), \end{aligned}$$

con lo cual $T(2,0,0,0)=(2,0,4,2)$; o sea que

$$e)\ T(1,0,0,0)=(1,0,2,1).$$

Si ahora hacemos la operación $e)-c)$, entonces,

$$T(1,0,0,0)-T(1,-1,0,0)=(1,0,2,1)-(0,0,1,0);$$

esto es,

$$f)\ T(0,1,0,0)=(1,0,1,1).$$

También se puede ver que

$$(0,0,1,0)=\frac{1}{5}\left[(2,0,3,2)-(1,1,-1,1)+(0,0,1,-1)-(1,-1,0,0)\right]$$

y, por lo tanto,

$$T(0,0,1,0)=\frac{1}{5}\left[(-3,5,4,2)-(2,0,3,2)+(0,0,0,0)-(0,0,1,0)\right],$$

de modo que $T(0,0,1,0)=(-1,1,0,0)$.

De manera similar, tenemos que

$$(0,0,0,1)=\frac{1}{5}\left[(2,0,3,2)-(1,1,-1,1)-4(0,0,1,-1)-(1,-1,0,0)\right],$$

de donde

$$T(0,0,0,1)=\frac{1}{5}\left[(-3,5,4,2)-(2,0,3,2)-4(0,0,0,0)-(0,0,1,0)\right],$$

y concluimos que

$$T(0,0,0,1)=(-1,1,0,0).$$

En síntesis, la matriz de la transformación con respecto a la base natural de $\mathbb{R}^4$ está dada por

$$M[T] = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

2. Según lo establecido en el numeral anterior,

$$R_T = gen\ \{(1,0,2,1),(1,0,1,1),(-1,1,0,0),(-1,1,0,0)\};$$

pero estos cuatro vectores son linealmente dependientes. Como la forma escalonada reducida de la matriz

$$M[T] = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \text{ es } \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

esto nos muestra que los tres primeros vectores forman un conjunto de generadores linealmente independientes para el rango de T. De esta manera, tenemos que

$$R_T = gen\ \{(1,0,2,1),(1,0,1,1),(-1,1,0,0)\}$$

y, en consecuencia, $dim\,(R_T) = 3$.

Por último, el núcleo de la transformación está dado por

$$N_T = \{(x,y,z,t)/x = y = 0, z + t = 0\};$$

luego, el vector $(x,y,z,t) \in N_T$ si

$$(x,y,z,t) = (0,0,z,-z) = z(0,0,1,-1)$$

y, por lo tanto,

$$N_T = gen\ \{(0,0,1,-1)\},$$

con lo cual $dim\,(N_T) = 1$.

Ejercicio 6.12.2. Sea $T : \mathbb{R}^4 \to \mathbb{R}^3$ una transformación lineal que cumple las siguientes condiciones:

a) $T(1,0,1,0) = (1,1,-1)$.

b) $T(-1,-2,0,0) = (1,1,1)$.

c) $(1,1,1,1) \in T^{-1}\{(0,0,1)\}$.

d) $(2,2,2,1) \in N_T$.

1. Calcule la matriz de la transformación con respecto a la base natural de $\mathbb{R}^4$.
2. Determine si la transformación es inyectiva y sobreyectiva.
3. Obtenga el núcleo y el rango de la transformación, al igual que la dimensión de estos.
4. Dadas las bases $B = \{(2,2,0,1),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}$ y $B' = \{(1,1,0),(0,1,0),(0,0,1)\}$ determine, $M[T]_{BB'}$.

Solución

1. Las condiciones que definen la transformación lineal podemos reescribirlas en la siguiente forma:

a) $T(1,0,1,0) = (1,1,-1)$.

b) $T(-1,-2,0,0) = (1,1,1)$.

c) $T(1,1,1,1) = (0,0,1)$.

d) $T(2,2,2,1) = (0,0,0)$.

Sea $A = \{(0,0,1,0),(-1,-2,0,0),(1,1,1,1),(2,2,2,1)\}$. Con estos vectores podemos calcular las imágenes de la transformación para los vectores de la base natural de $\mathbb{R}^4$. De esta forma, al ser

$$(1,0,0,0) = 2(1,0,1,0) - (-1,-2,0,0) - 2(2,2,2,1) + 2(1,1,1,1)$$

y aplicando las propiedades de la transformación lineal, obtenemos

$$T(1,0,0,0) = 2(1,1,-1) - (1,1,1) - 2(0,0,0) + 2(0,0,1);$$

en consecuencia,

$$T(1,0,0,0) = (1,1,-1).$$

De nuevo, al ser

$$(0,1,0,0) = -1(1,0,1,0) + (2,2,2,1) - (1,1,1,1),$$

entonces

$$T(0,1,0,0) = -1(1,1,-1) + (0,0,0) - (0,0,1),$$

con lo cual

$$T(0,1,0,0) = (-1,-1,0).$$

Igualmente, ya que

$$(0,0,1,0) = -(1,0,1,0) + (-1,-2,0,0) + 2(2,2,2,1) - 2(1,1,1,1),$$

se tiene que

$$T(0,0,1,0) = -(1,1,-1) + (1,1,1) + 2(0,0,0) - 2(0,0,1),$$

con lo que concluimos que

$$T(0,0,1,0) = (0,0,0).$$

Por último, como

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

tenemos que

$$T(0,0,0,1) = 2(0,0,1) - (0,0,0),$$

de donde

$$T(0,0,0,1) = (0,0,2).$$

Por lo tanto, la matriz de la transformación con respecto a esta base es

$$M[T]_E = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix}.$$

2. Como $(2,2,2,1) \in N_T$, entonces la transformación T no es inyectiva. También sabemos que la transformación T será sobreyectiva si $dim\,(R_T) = 3$. Para ello vamos a calcular el núcleo N_T: un vector $(x,y,z,t) \in N_T$ si

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

De esta relación concluimos que $x - y = 0$ y $-x + 2t = 0$, es decir, $x = y$ y $y = 2t$, de modo que

$$N_T = \left\{(2t, 2t, z, t)/z, t \in \mathbb{R}\right\},$$

con lo cual

$$N_T = gen\left\{(2,2,0,1), (0,0,1,0)\right\}$$

y, por lo tanto, la dimensión del núcleo es 2.

Como

$$dim\,\mathbb{R}^4 = 4 = dim\,N_T + dim\,R_T,$$

entonces

$$dim\,R_T = 2 < 3$$

y, por lo tanto, la transformación no es sobreyectiva. Esto da respuesta a la tercera pregunta del ejercicio.

4. Se puede probar con facilidad que

$$M[T]_{BB'} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

En efecto, como ya hemos visto que

$$T\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} x-y \\ x-y \\ -x+2t \end{bmatrix},$$

entonces

$$T\begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$T\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = -1\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$$

$$T\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$T\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

De esta forma concluimos que

$$M[T]_{BB'} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

como habíamos afirmado.

Ejercicio 6.12.3. Sea $T : \mathbb{R}^3 \to \mathbb{R}^4$ una transformación lineal que cumple las siguientes condiciones

a) $T(1,0,1) = (1,-1,2,2)$.

b) $T(0,1,0) = (-1,1,1,1)$.

c) $T^{-1}(-1,0,3,1) = \{(0,1,1)\}$.

1. Calcule tanto la matriz de la transformación con respecto a las bases canónicas, como también los subespacios N_T y R_T.
2. Dé bases para estos subespacios y determine si la transformación es inyectiva.
3. Si $G : \mathbb{R}^4 \to \mathbb{R}^3$ es una transformación lineal cuya matriz de representación con respecto a las bases canónicas es

$$A = M[G]_{BB'} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

calcule las matrices que representan a las transformaciones $T \circ G$ y $G \circ T$.

Solución

1. Las condiciones que definen la transformación podemos escribirlas nuevamente en la forma:

a) $T(1,0,1) = (1,-1,2,2)$.

b) $T(0,1,0) = (-1,1,1,1)$.

c) $T(0,1,1) = (-1,0,3,1)$.

Al ser $(1,0,0) = (1,0,1) + (0,1,0) - (0,0,1)$, entonces

$$\begin{aligned} T(1,0,0) &= T(1,0,1) + T(0,1,0) - T(0,0,1) \\ T(1,0,0) &= (1,-1,2,2) + (-1,1,1,1) - (-1,0,3,1) \\ T(1,0,0) &= (1,0,0,2). \end{aligned}$$

De modo similar: ya que $(0,0,1) = 0(1,0,1) - (0,1,0) + (0,0,1)$, luego

$$\begin{aligned}
T(0,0,1) &= 0T(1,0,1) - T(0,1,0) + T(0,0,1)\\
T(0,0,1) &= 0(1,-1,2,2) - (-1,1,1,1) + (-1,0,3,1)\\
T(0,0,1) &= (0,-1,2,0).
\end{aligned}$$

Como el conjunto de vectores $B = \{(1,0,0),(0,1,0),(0,0,1)\}$ es la base canónica de $\mathbb{R}^3$, la matriz de la transformación respecto a estas bases está dada por

$$B = M[T]_{BB'} = \begin{bmatrix} 1 & -1 & 0\\ 0 & 1 & -1\\ 0 & 1 & 2\\ 2 & 1 & 0 \end{bmatrix},$$

siendo B' la base canónica de $\mathbb{R}^4$. De este modo, la transformación queda definida por

$$T\begin{bmatrix} x\\ y\\ z \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0\\ 0 & 1 & -1\\ 0 & 1 & 2\\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} x\\ y\\ z \end{bmatrix} = \begin{bmatrix} x-y\\ y-z\\ y+2z\\ 2x+y \end{bmatrix}.$$

2. Para calcular el N_T debemos resolver el sistema de ecuaciones correspondiente a

$$T\begin{bmatrix} x\\ y\\ z \end{bmatrix} = \begin{bmatrix} 0\\ 0\\ 0 \end{bmatrix},$$

es decir,

$$\begin{aligned}
x-y &= 0\\
y-z &= 0\\
y+2z &= 0\\
2x+y &= 0.
\end{aligned}$$

Resolviendo este sistema lineal homogéneo se tiene que

$$x = y = z = 0$$

y, por lo tanto, $N_T = \{(0,0,0)\}$. De allí que la transformación sea inyectiva.

Con respecto al subespacio R_T, sabemos que

$$R_T = gen\,\{(0,-1,2,0),(-1,1,1,1),(1,0,0,2)\}.$$

Como $A = \{(0,-1,2,0),(-1,1,1,1),(1,0,0,2)\}$ es un conjunto linealmente independiente, estos tres vectores forman una base para R_T y, en consecuencia, $dim\,R_T = 3$. En conclusión, se cumple que $dim\,\mathbb{R}^3 = dim\,N_T + dim\,R_T$.

3. Como $T : \mathbb{R}^3 \to \mathbb{R}^4$ y $G : \mathbb{R}^4 \to \mathbb{R}^3$, entonces $T \circ G : \mathbb{R}^4 \to \mathbb{R}^4$ y $G \circ T : \mathbb{R}^3 \to \mathbb{R}^3$. Por el teorema 6.10.2 sabemos que

$$[G \circ T]_E = [G]_{EE'}[G]_{EE'},$$

siendo E y E' las bases naturales de $\mathbb{R}^3$ y $\mathbb{R}^4$ respectivamente. De esta manera,

$$[G \circ T]_E = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 2 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 3 & 0 & -3 \\ 0 & 2 & 1 \end{bmatrix}$$

y, en consecuencia, $G \circ T$ será una transformación lineal de $\mathbb{R}^3 \to \mathbb{R}^3$, tal que

$$G \circ T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x \\ 3x - 3z \\ 2y + z \end{bmatrix}.$$

De modo similar se calcula $[T \circ G]_{E'}$, donde E' es la base natural del espacio vectorial $\mathbb{R}^4$.

Ejercicio 6.12.4. Sea $T : \mathbb{R}^3 \to \mathbb{R}^4$ una transformación lineal tal que

$$M[T]_{EE'} = \begin{bmatrix} 2 & 0 & 1 \\ a & 3 & -1 \\ -1 & 1 & b \\ 0 & -2 & 1 \end{bmatrix},$$

donde E y E' son las bases naturales de $\mathbb{R}^3$ y $\mathbb{R}^4$ respectivamente.

1. Calcule $T(x, y, z)$.

2. ¿Para qué valores de a y b el vector $(1, -1, -2) \in N_T$?

3. Tomando los valores de a y b calculados en el numeral anterior y dados los siguientes subespacios de $\mathbb{R}^3$, $L = gen\{(0, -2, 4)\}$ y $H = gen\{(1, 1, -1), (0, 1, -2)\}$, halle $T(H) \cap T(L)$.

Solución

1. Para calcular $T(x, y, z)$ usamos la matriz de representación de la transformación y obtenemos

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

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + z \\ ax + 3y - z \\ -x + y + bz \\ -2y + z \end{bmatrix}.$$

2. Si $(1, -1, -2) \in N_T$, entonces

$$T \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ a & 3 & -1 \\ -1 & 1 & b \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

con lo que se obtiene el sistema de ecuaciones

$$\begin{aligned} 2-2 &= 0 \\ a-3+2 &= 0 \\ -1-1-2b &= 0 \\ 2-2 &= 0; \end{aligned}$$

esto implica que $a = 1$ y $b = -1$.

3. Para responder este numeral, sabemos que

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

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

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

y, por lo tanto,

$$T(H) = gen\,\{(1,5,1,-3),(-2,5,3,-4)\},$$

mientras que

$$T(L) = gen\,\{(4,-10,-6,8)\},$$

pues la imagen de un conjunto generador de un subespacio cualquiera A es un conjunto generador de $T(A)$.

Como

$$(4, -10, -6, 8) = -2(1, 5, 1, -3) + 0(-2, 5, 3, -4),$$

entonces $T(L) \subset T(H)$, por lo que $T(L) \cap T(H) = T(L)$.

Ejercicio 6.12.5. Dada una transformación $T : \mathbb{R}^3 \to \mathbb{R}^4$:

1. ¿Puede ser inyectiva si $dim\, R_T = 2$?
2. Si $T : \mathbb{R}^3 \to \mathbb{R}^3$ es una transformación lineal sobreyectiva y $G : \mathbb{R}^3 \to \mathbb{R}^4$ es una transformación lineal inyectiva, ¿es $G \circ T$ inyectiva?

Solución

Sabemos que

$$dim\, \mathbb{R}^3 = dim\, N_T + dim\, R_T;$$

pero como $dim\, R_T = 2$, entonces $dim\, N_T = 1$ y, en consecuencia, la transformación no es inyectiva.

Para la segunda pregunta, nuevamente aprovechamos el hecho de que

$$dim\, \mathbb{R}^3 = dim\, N_T + dim\, R_T;$$

pero como T es sobreyectiva, $dim\, R_T = 3$; entonces, $dim\, N_T = 0$ y, por lo tanto, T es inyectiva.

Veamos ahora si $G \circ T : \mathbb{R}^3 \to \mathbb{R}^3$ es inyectiva. Sea $x \in \mathbb{R}^3$ tal que $(G \circ T)(x) = 0$; de allí resulta que $G(T(x)) = 0$; luego, $T(x) \in N_G$. Pero como G es inyectiva, entonces, por el teorema 6.6.1, $N_G = \{0\}$, con $0 \in \mathbb{R}^3$. Por lo tanto, $T(x) = 0$ y $x \in N_T$, con lo cual probamos que $x = 0$ y, en consecuencia, el núcleo de la transformación $G \circ T$ es $\{0\}$, de donde concluimos, según el mismo teorema, que $G \circ T$ inyectiva.

Ejercicio 6.12.6 (Computador). Dada la transformación lineal $T : \mathbb{R}^4 \to \mathbb{R}^4$ determinada por

$$T(x, y, z, t) = (x - y - t, 2x + 4z - t, x + y, x + y - 2z + t),$$

utilizamos DERIVE para calcular $M[T]_E$, donde E es la base canónica de $\mathbb{R}^4$. Además, se quiere calcular el rango y el núcleo de dicha transformación.

Solución

Iniciamos el problema definiendo la transformación lineal mediante la instrucción AUTOR:

$$\#1.\ \ F(x, y, z, t) := [x - y - t, 2x + 4z - t, x + y, x + y - 2z + t].$$

Ahora, calculemos la imagen de cada vector de la base elegida:

$$\#2.\ \ [F(1, 0, 0, 0), F(0, 1, 0, 0), F(0, 0, 1, 0), F(0, 0, 0, 1)]'$$

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

Si definimos la matriz $A = M[T]_E$, entonces podemos calcular la expresión para la transformación lineal y con ella resolver el sistema

$$A[x, y, z, t]^t = [0, 0, 0, 0]^t$$

para determinar el núcleo de la transformación. Esto podemos verlo en el siguiente grupo de instrucciones:

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

$$\#5.\ \ A.[x, y, z, t]' = [0, 0, 0, 0]$$

$$\#6.\ \ \text{SOLVE}(A.[x, y, z, t]' = [0, 0, 0, 0], [x, y, z, t])$$

$$\#7.\ \ 2x + t = 0\ \wedge\ 2y - t = 0\ \wedge\ 2z - t = 0$$

$$\#8.\ \ \text{SOLVE}(2x + t = 0\ \wedge\ 2y - t = 0\ \wedge\ 2z - t = 0, [y, z, t])$$

$$\#9.\ \ y = -x\ \wedge\ z = -x\ \wedge\ t = -2x$$

De la última instrucción concluimos que el núcleo de la transformación es $N_T = gen\{(1,-1,-1,-2)\}$.

Finalmente, para determinar el rango de la transformación, llevamos la matriz A^t a la forma escalonada reducida, con lo cual obtenemos

$$\#10. \quad Row\_reduce(A')$$

$$\#11. \quad \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

y, por lo tanto, el rango de la transformación está conformado por

$$R_T = gen\{(1,0,0,1/2),(0,1,0-1/2),(0,0,1,3/2)\}.$$

Ejercicio 6.12.7. Si $T:\mathbb{R}^2 \to \mathbb{R}^2$ es una transformación lineal tal que sus valores para la base canónica de $\mathbb{R}^2$ son:

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \text{ y } T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix},$$

halle $T(x,y)$ y T^{-1} en caso de que este exista.
Solución:
Como $E = \{(1,0),(0,1)\}$ es la base natural de $\mathbb{R}^2$,

$$[T]_E = A = \begin{bmatrix} 3 & -1 \\ 4 & 2 \end{bmatrix}.$$

De este modo, la transformación queda definida como

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x-y \\ 4x+2y \end{bmatrix}.$$

Ahora, como A es la matriz de representación de la transformación, veamos si tiene inversa:

$$\left[\begin{array}{cc|cc} 3 & -1 & 1 & 0 \\ 4 & 2 & 0 & 1 \end{array}\right] \xrightarrow[\frac{1}{3}F_1]{} \left[\begin{array}{cc|cc} 1 & -\frac{1}{3} & \frac{1}{3} & 0 \\ 4 & 2 & 0 & 1 \end{array}\right] \xrightarrow[F_2-4F_1]{} \left[\begin{array}{cc|cc} 1 & -\frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{10}{3} & -\frac{4}{3} & 1 \end{array}\right]$$

$$\xrightarrow[\frac{3}{10}F_2]{} \left[\begin{array}{cc|cc} 1 & -\frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 1 & -\frac{4}{10} & \frac{3}{10} \end{array}\right] \xrightarrow[F_1+\frac{1}{3}F_2]{} \left[\begin{array}{cc|cc} 1 & 0 & \frac{2}{10} & \frac{1}{10} \\ 0 & 1 & -\frac{4}{10} & \frac{3}{10} \end{array}\right]$$

y, por lo tanto,

$$A^{-1} = \begin{bmatrix} \frac{2}{10} & \frac{1}{10} \\ -\frac{4}{10} & \frac{3}{10} \end{bmatrix}.$$

De esta manera, la transformación inversa está dada por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 2 & 1 \\ -4 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 2x+y \\ -4x+3y \end{bmatrix}.$$

Podemos verificar con facilidad que $(T(T^{-1}))(x,y) = (x,y)$.

Ejercicio 6.12.8. Sean $S : \mathbb{R}^4 \to \mathbb{R}^3$ y $T : \mathbb{R}^3 \to \mathbb{R}^3$ dos transformaciones lineales definidas por

$$S\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2x-3z+2y \\ x+y+z+w \\ x+2y+2w \end{bmatrix} \text{ y } T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y+2z \\ 2x+3y+z \end{bmatrix},$$

y sean B, B' y B'' las bases naturales para los espacios $\mathbb{R}^4$, $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente.

1. Halle $(T \circ s)(x,y,z,w)$.
2. Halle $A = [S]_{BB'}$, $B = [T]_{B'B''}$ y $[T \circ S]_{BB''}$.
3. Demuestre que $[T \circ S]_{BB''} = AB$.

Solución

Para calcular la composición de transformaciones tenemos:

1.

$$\begin{aligned}(T\circ S)(X) &= T\begin{bmatrix} 2x-3z+2y \\ x+y+z+w \\ x+2y+2w \end{bmatrix}\\ &= \begin{bmatrix} 2x-3z+2y-(x+y+z+w)+2(x+2y+2w) \\ 2(2x-3z+2y)+3(x+y+z+w)+x+2y+2w \end{bmatrix}\\ &= \begin{bmatrix} 3x+5y-4z+3w \\ 8x+9y-3z+5w \end{bmatrix}.\end{aligned}$$

2. Las imágenes de los vectores de la base natural de $\mathbb{R}^4$ son:

$$S\begin{bmatrix}1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}2\\1\\1\end{bmatrix};\ S\begin{bmatrix}0\\1\\0\\0\end{bmatrix}=\begin{bmatrix}2\\1\\2\end{bmatrix};$$

$$S\begin{bmatrix}0\\0\\1\\0\end{bmatrix}=\begin{bmatrix}-3\\1\\0\end{bmatrix};\ S\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\1\\2\end{bmatrix}.$$

En consecuencia, la matriz de la transformación es

$$B=[S]_{BB'}=\begin{bmatrix}2&2&-3&0\\1&1&1&1\\1&2&0&2\end{bmatrix}.$$

De modo similar, para los vectores de la base natural de $\mathbb{R}^3$ se tiene que las imágenes bajo la transformación T son

$$T\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}1\\2\end{bmatrix},\ T\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}-1\\3\end{bmatrix}\ \text{y}\ T\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}$$

y, por lo tanto, la matriz de esta transformación es

$$A=[T]_{B'B''}=\begin{bmatrix}1&-1&2\\2&3&1\end{bmatrix}.$$

Por último, para $H = T \circ S$, las imágenes de los vectores de la base natural de $\mathbb{R}^4$ son

$$H(e_1) = \begin{bmatrix} 3 \\ 8 \end{bmatrix},\ H(e_2) = \begin{bmatrix} 5 \\ 9 \end{bmatrix},\ H(e_3) = \begin{bmatrix} -4 \\ -3 \end{bmatrix} \text{ y } H(e_4) = \begin{bmatrix} 3 \\ 5 \end{bmatrix},$$

con lo cual vemos que la matriz de $T \circ S$ es

$$[T \circ S]_{BB''} = \begin{bmatrix} 3 & 5 & -4 & 3 \\ 8 & 9 & -3 & 5 \end{bmatrix}.$$

3. Dadas las matrices $A = [T]_{B'B''}$ y $B = [S]_{BB'}$, entonces,

$$\begin{aligned} AB &= \begin{bmatrix} 1 & -1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & -3 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 5 & -4 & 3 \\ 8 & 9 & -3 & 5 \end{bmatrix} \\ &= [T \circ S]_{BB''}, \end{aligned}$$

como queríamos verificar.

6.13 Ejercicios propuestos capítulo 6

Determine si las siguientes funciones son transformaciones lineales.

1. $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ x - y \\ y + z \end{bmatrix}.$$

2. $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ yz \end{bmatrix}.$$

3. $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x + 2y \\ z \\ |y| \end{bmatrix}.$$

4. $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

5. $T : M_{nn} \to M_{nn}$, definida por $T\,[A] = A + A^t$.

6. $T : M_{nn} \to M_{nn}$, definida por $T\,[A] = AA^t$.

7. $T : M_{nn} \to M_{nn}$, definida por $T\,[A] = AB$, donde B es una matriz fija de $n \times n$.

8. Encuentre núcleo, imagen, nulidad y rango para las siguientes transformaciones lineales.

   a) $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ x - y \\ y + z \end{bmatrix}.$$

   b) $T : \mathbb{R}^3 \to \mathbb{R}^2$, definida por

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} z \\ y \end{bmatrix}.$$

c) $T:\mathbb{R}^4\to\mathbb{R}^3$, definida por

$$T\begin{bmatrix} x\\ y\\ z\\ w\end{bmatrix}=\begin{bmatrix} x+z\\ y+w\\ w\end{bmatrix}.$$

9. Encuentre la representación matricial de la transformación lineal dada. Obtenga también el núcleo e imagen de dicha transformación.

a) $T:\mathbb{R}^2\to\mathbb{R}^3$, definida por

$$T\begin{bmatrix} x\\ y\end{bmatrix}=\begin{bmatrix} x+y\\ x-y\\ 2x+3y\end{bmatrix}.$$

b) $T:\mathbb{R}^3\to\mathbb{R}^3$, definida por

$$T\begin{bmatrix} x\\ y\\ z\end{bmatrix}=\begin{bmatrix} -x+2y+z\\ 2x-4y-2z\\ -3x+6y+3z\end{bmatrix}.$$

c) $T:\mathbb{R}^3\to\mathbb{R}^3$, definida por

$$T\begin{bmatrix} x\\ y\\ z\end{bmatrix}=\begin{bmatrix} 2y+z\\ 2x-2z\\ -3x+6y+3z\end{bmatrix}.$$

d) $T:\mathbb{R}^4\to\mathbb{R}^3$, definida por

$$T\begin{bmatrix} x\\ y\\ z\\ w\end{bmatrix}=\begin{bmatrix} -x+2y+z+w\\ -4y-2z+3w\\ -3x+6y+3w\end{bmatrix}.$$

e) $T : M_{nn} \to M_{nn}$, definida por

$$T\,[A] = A + A^t.$$

f) $T : M_{nn} \to M_{nn}$, definida por

$$T\,[A] = AA^t.$$

g) $T : \mathbb{R}^5 \to \mathbb{R}^4$, definida por

$$T\begin{bmatrix} t \\ x \\ y \\ x \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 3 & -1 \\ 1 & 0 & 0 & 2 & -1 \\ 2 & 0 & -1 & 5 & -1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} t \\ x \\ y \\ x \\ w \end{bmatrix}.$$

10. Sean las transformaciones lineales

$$T_1 : \mathbb{R}^3 \to \mathbb{R}^2, \qquad T_1(x_1, x_2, x_3) = (x_1, x_1 + x_2).$$

$$T_2 : \mathbb{R}^3 \to \mathbb{R}^2, \qquad T_2(x_1, x_2, x_3) = (x_1 - x_3, x_1 + x_2).$$

$$T_3 : \mathbb{R}^1 \to \mathbb{R}^3, \qquad T_3(x_1) = (x_1, 2x_1, -x_1).$$

Halle cada una de las siguientes transformaciones lineales:

a) $T_1 + T_2$.

b) $T_1 \circ T_3$.

c) $T_2 \circ T_3$.

d) $3T_1 - 2T_2$.

11. Sean las transformaciones lineales

$$T_1 : \mathbb{R}^3 \to \mathbb{R}^1, \qquad T_1(x_1, x_2, x_3) = x_1 + x_1 + x_2.$$

$$T_2 : \mathbb{R}^3 \to \mathbb{R}^1, \qquad T_2(x_1, x_2, x_3) = x_1 - x_3.$$

$$T_3 : \mathbb{R}^1 \to \mathbb{R}^2, \qquad T_3(x_1) = (2x_1, -x_1).$$

$$T_4 : \mathbb{R}^2 \to \mathbb{R}^2, \qquad T_4(x_1, x_2) = (x_1 + x_2, x_1).$$

Halle la matriz de cada una de las siguientes transformaciones lineales:

a) $2T_1 - 2T_2$.

b) $T_3 \circ T_1$.

c) $T_3 \circ T_2$.

d) T_4^2.

e) $T_3 \circ (T_1 - 3T_2)$.

Capítulo 7

Valores y vectores propios

7.1 Introducción

En el capítulo anterior vimos que para un determinado par de bases, la transformación quedaba caracterizada de manera única por una matriz. Es claro que si escogemos las bases de modo adecuado, la matriz de representación de la transformación puede tener una forma sencilla para realizar los cálculos.

Existe gran cantidad de problemas matemáticos en los cuales para su solución se requiere resolver sistemas de ecuaciones lineales; en estos problemas, la elección de las bases adecuadas es importante, pues no solamente agiliza los cálculos, sino que también garantiza mayor precisión en la solución del problema. En este capítulo desarrollamos los elementos básicos de la teoría de los valores y vectores propios, y el problema de la diagonalización. Así mismo, presentamos algunos ejemplos que nos darán una clara idea del uso de estos conceptos.

En el caso de las transformaciones lineales, nos ocupamos solo del problema de los endomorfismos, es decir, transformaciones lineales de un espacio vectorial sobre sí mismo. En este caso, únicamente necesitamos encontrar una base con respecto a la cual la matriz de representación de la transformación sea una matriz diagonal.

7.2 Valores y vectores propios de una matriz

Definición 7.2.1. *Sea A una matriz de orden $n \times n$ con componentes reales o complejas. El número λ (real o complejo) se llama valor propio de A si existe un vector no nulo v en $\mathbb{R}^n$ o $\mathbb{C}^n$ tal que*

$$Av = \lambda\, v. \tag{7.2.1}$$

El vector $v \neq 0$ se denomina vector propio de la matriz A asociado al valor propio λ.

Ejemplo 7.2.1. Para la matriz

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix},$$

$\lambda = 7$ es un valor propio y el vector $(1, 3)$ es un vector propio asociado a $\lambda = 7$, puesto que

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 21 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

como lo afirma la definición.

Es obvio que si v es un vector propio asociado al valor propio λ, entonces cv, con $c \in \mathbb{R}$, también es un vector propio. La ecuación 7.2.1 se puede expresar en la forma

$$(A - \lambda I)v = 0, \tag{7.2.2}$$

donde I es la matriz identidad de orden n. La ecuación 7.2.2 representa un sistema homogéneo con solución no trivial (diferente de cero), lo cual ocurre si y solo si

$$\det(A - \lambda I) = 0. \tag{7.2.3}$$

La ecuación 7.2.3 es la *ecuación característica* de la matriz A.

Si consideramos la matriz A de tamaño $n \times n$, obtenemos que

$$A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}; \tag{7.2.4}$$

entonces, por la definición de determinante, el $det(A - \lambda I)$ está formado por una suma de factores que corresponden a elementos de filas y columnas diferentes de la matriz $A - \lambda I$.

Es claro que el producto de todos los elementos de la diagonal de dicha matriz está conformado por n factores de la forma $a_{ii} - \lambda$; por lo tanto, el producto de los elementos de la diagonal tiene la forma

$$(a_{11} - \lambda)(a_{22} - \lambda) \ldots (a_{nn} - \lambda).$$

Si tomamos a λ como una variable, en el producto anterior el término de mayor exponente es $(-\lambda)^n$, lo que nos muestra que la ecuación característica es un polinomio de grado n, el cual es conocido como el *polinomio característico* de la matriz A y lo representamos como

$$P(\lambda) = \det(A - \lambda I).$$

De este modo, la ecuación 7.2.3 es una ecuación polinomial en λ de grado n y se puede escribir como

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0. \tag{7.2.5}$$

De acuerdo con el teorema fundamental del álgebra, la matriz A de tamaño $n \times n$ tiene exactamente n valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$, reales o complejos, algunos de los cuales se pueden repetir. Si un valor propio dado aparece k veces como una raíz de 7.2.3, se dice que ese valor propio tiene multiplicidad k.

Nota 7.2.1. Los valores propios de una matriz triangular superior o inferior son los elementos de la diagonal principal. En efecto, para el caso de una matriz triangular superior A, la ecuación característica toma la forma

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{nn}-\lambda \end{vmatrix} = 0.$$

Al expandir el determinante obtenemos

$$(a_{11}-\lambda)\cdots(a_{nn}-\lambda) = 0.$$

Teorema 7.2.1. *Los vectores propios $v_1, v_2, \ldots, v_k$ de la matriz A, asociados a valores propios diferentes $\lambda_1, \lambda_2, \ldots, \lambda_k$, son linealmente independientes.*

Demostración. Supongamos que los vectores propios $v_1, v_2, \ldots, v_k$ de la matriz A son linealmente dependientes, y tomemos $g, 2 \leq g \leq k$, tal que v_g sea el primer vector que se puede escribir como combinación lineal de los vectores $v_1, v_2, \ldots, v_{g-1}$, es decir,

$$v_g = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_{g-1} v_{g-1} \tag{7.2.6}$$

para ciertos escalares $\alpha_1, \alpha_1, \ldots, \alpha_{g-1}$.

Si multiplicamos ambos lados de la ecuación 7.2.6 por A, se tiene

$$\begin{aligned} Av_g &= \alpha_1 Av_1 + \alpha_2 Av_2 + \ldots + \alpha_{g-1} Av_{g-1} \\ \lambda_g v_g &= \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \ldots + \lambda_{g-1} \alpha_{g-1} v_{g-1}; \end{aligned} \tag{7.2.7}$$

multiplicando la ecuación 7.2.6 por λ_g y restándole la ecuación 7.2.7, obtenemos

$$0 = \alpha_1(\lambda_g - \lambda_1)v_1 + \alpha_2(\lambda_g - \lambda_2)v_2 + \ldots + \alpha_{g-1}(\lambda_g - \lambda_{g-1})v_{g-1}.$$

Pero como este conjunto de vectores es linealmente independiente, se tiene que

$$\alpha_1(\lambda_g - \lambda_1) = \alpha_2(\lambda_g - \lambda_2) = \ldots = \alpha_{g-1}(\lambda_g - \lambda_{g-1}) = 0.$$

Como los valores propios son diferentes,

$$\alpha_1 = \alpha_2 = \ldots = \alpha_{g-1} = 0$$

y por la ecuación 7.2.6,

$$v_g = 0v_1 + 0v_2 + \ldots + 0v_{g-1} = 0,$$

lo cual es absurdo, pues v_g es un vector propio y por definición, los vectores propios siempre son diferentes del vector 0. Por lo tanto, los vectores propios $\{v_1, v_2, \ldots, v_k\}$ constituyen un conjunto linealmente independiente.

El conjunto de vectores propios correspondientes al valor propio λ es el conjunto

$$E_\lambda = \{v \in \mathbb{R}^n : \ Av = \lambda\, v\}.$$

Otro aspecto importante con respecto a los vectores propios de una matriz lo constituye el hecho de que el conjunto de vectores propios asociados a un misma valor propio, E_λ, es un subespacio vectorial de $\mathbb{R}^n$ o $\mathbb{C}^n$. Esto podemos verlo con facilidad, ya que si v_1 y v_2 son dos vectores propios de la matriz A asociados a λ_1, entonces

$$\begin{aligned} A(v_1+v_2) &= A(v_1)+A(v_2)=\lambda_1 v_1+\lambda_1 v_2=\lambda_1(v_1+v_2)\\ A(\alpha v_1) &= \alpha(Av_1)=\alpha(\lambda_1 v_1)=\lambda_1(\alpha v_1); \end{aligned}$$

por lo tanto, λ_1 sigue siendo un valor propio para v_1+v_2 y αv_1. Este subespacio es conocido como el *subespacio característico* de A, asociado al valor propio λ.

Si estamos interesados en calcular los valores propios y los vectores propios correspondientes a cada valor propio de una matriz, teniendo en cuenta las definiciones dadas hasta aquí, podemos resumir este procedimiento en los siguientes pasos:

1. Se calcula el polinomio característico de la matriz A, es decir, se calcula $p(\lambda)=\det(A-\lambda I)$.

2. Se encuentran las raíces del polinomio característico, es decir, se resuelve la ecuación $p(\lambda)=0$.

3. Por último, se resuelve el sistema homogéneo $(A-\lambda I)\,v=0$, correspondiente a cada valor propio λ_i, para $i=1,\,2,\,\ldots,\,n$.

El último paso, es equivalente a calcular el espacio característico o propio correspondiente al valor propio λ, que es el conjunto

$$E_\lambda=\{v:\ Av=\lambda\, v\}\,.$$

Vamos a ilustrar este procedimiento con algunos ejemplos.

Ejemplo 7.2.2. Dada la matriz

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

halle sus valores y espacios propios.

Solución

Encontremos el polinomio característico de la matriz A.

$$\begin{aligned}
\det(A-\lambda I) &= \begin{vmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} \\
&= (1-\lambda)\left[(2-\lambda)(1-\lambda)-1\right]-(-1)\left[-(1-\lambda)+0\right] \\
&= (1-\lambda)\left[(2-\lambda)(1-\lambda)-1-1\right] \\
&= (1-\lambda)\left[2-3\lambda+\lambda^2-2\right] \\
&= \lambda(1-\lambda)(\lambda-3)=0.
\end{aligned}$$

De este modo, los valores propios de la matriz A corresponden a $\lambda = 0$, $\lambda = 1$ y $\lambda = 3$.

Ahora hallemos los vectores propios asociados con estos valores propios. Para ello, resolvemos el sistema homogéneo $(A-\lambda I)\,v = 0$, para cada uno de los valores de λ.

Comencemos con $\lambda = 0$. Esto es, resolvamos el sistema $A\,v = 0$:

$$(A-0I)\,v_1 = Av_1 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

donde $v_1 = [x_1, x_2, x_3]^t$. La matriz aumentada para este sistema es

$$\left[\begin{array}{rrr|r} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right];$$

al reducir por filas obtenemos

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

De acá tenemos $x_1 = x_3$ y $x_2 = x_3$, de manera que los vectores propios asociados con $\lambda = 0$ son de la forma:

$$v_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Por tanto, el espacio característico para $\lambda = 0$ es

$$E_0 = \left\{ x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : x_3 \in \mathbb{R} \right\} = gen \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Para $\lambda = 1$:

$$(A - 1I)\, v_2 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

donde $v_2 = [x_1, x_2, x_3]^t$. Con ello obtenemos la matriz

$$\left[\begin{array}{rrr|r} 0 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right].$$

De nuevo, al aplicar operaciones elementales por fila se llega finalmente a

$$\left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

En este caso, las soluciones son: $x_1 = -x_3$ y $x_2 = 0$; luego, los vectores propios asociados con $\lambda = 1$ son de la forma:

$$v_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

de modo que el espacio característico para $\lambda = 1$ es

$$E_1 = \left\{ x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} : x_3 \in \mathbb{R} \right\} = gen \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Para $\lambda = 3$ tenemos:

$$(A - 3I)\, v_3 = \begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

donde $v_3 = [x_1, x_2, x_3]^t$. En consecuencia, la matriz aumentada es

$$\left[\begin{array}{rrr|r} -2 & -1 & 0 & 0 \\ -1 & -1 & -1 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right]$$

y al aplicar operaciones elementales por filas se obtiene

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Las soluciones son: $x_1 = x_3$ y $x_2 = -2x_3$. Los vectores propios asociados con $\lambda = 3$ son de la forma:

$$v_3 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

de modo que el espacio característico para $\lambda = 3$ es

$$E_3 = \left\{ x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} : x_3 \in \mathbb{R} \right\} = gen \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

Aquí podemos ver que la matriz posee tres valores propios diferentes y, en consecuencia, nos encontramos con tres vectores propios distintos que, como se mostró, son linealmente independientes.

Ejemplo 7.2.3. Dada la matriz

$$A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix},$$

halle los valores y espacios propios.
Solución
Encontremos el polinomio característico de la matriz A.

$$\det(A - \lambda I) = \begin{vmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix}$$

$$\begin{aligned} \det(A-\lambda I) &= (5-\lambda)\left[(5-\lambda)(2-\lambda)-4\right] \\ &\quad -4\left[4(2-\lambda)-4\right]+2\left[8-2(5-\lambda)\right] \\ &= (5-\lambda)\left(10-7\lambda+\lambda^2-4\right) \\ &\quad -4(8-4\lambda-4)+2(8-10+2\lambda) \\ &= (5-\lambda)(\lambda-1)(\lambda-6)+16(\lambda-1)+4(\lambda-1) \\ &= (\lambda-1)\left[(5-\lambda)(\lambda-6)+20\right] \\ &= (\lambda-1)\left(-\lambda^2+11\lambda-10\right) \\ &= -(\lambda-1)(\lambda-1)(\lambda-10) \\ &= -(\lambda-1)^2(\lambda-10). \end{aligned}$$

Luego, los valores propios son: $\lambda = 1$, con multiplicidad algebraica 2 (la raíz se repite 2 veces) y $\lambda = 10$.

Encontremos los vectores propios asociados con estos valores propios. De nuevo resolvemos el sistema homogéneo $(A - \lambda I)\, v = 0$, para $\lambda = 1$ y $\lambda = 10$.

Para $\lambda = 10$, tenemos con $u_1 = [x_1, x_2, x_3]^t$.

$$(A - 10I)\, u_1 = \begin{bmatrix} -5 & 4 & 2 \\ 4 & -5 & 2 \\ 2 & 2 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

y, así, la matriz aumentada del sistema es

$$\left[\begin{array}{ccc|c} -5 & 4 & 2 & 0 \\ 4 & -5 & 2 & 0 \\ 2 & 2 & -8 & 0 \end{array}\right].$$

Al aplicar operaciones elementales por filas se obtiene

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

Las soluciones del sistema son: $x_1 = 2x_3$ y $x_2 = 2x_3$; por lo tanto, los vectores propios asociados con $\lambda = 10$ son de la forma

$$u_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

El espacio característico o propio para $\lambda = 10$ es:

$$E_{10} = \left\{ x_3 \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} : \ x_3 \in \mathbb{R} \right\} = gen \left\{ \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Resolvamos ahora el sistema para $\lambda = 1$, es decir,

$$(A - 1I)\ u = \begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

donde $u = [x_1, x_2, x_3]^t$. La matriz aumentada de dicho sistema es

$$\left[\begin{array}{ccc|c} 4 & 4 & 2 & 0 \\ 4 & 4 & 2 & 0 \\ 2 & 2 & 1 & 0 \end{array}\right].$$

Al aplicar operaciones elementales por filas se obtiene

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

Las soluciones al sistema homogéneo están dadas por el vector

$$u = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{bmatrix},$$

donde x_2 y x_3 son parámetros reales. El espacio propio asociado al valor $\lambda = 1$ es:

$$\begin{aligned} E_1 &= \left\{ x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{bmatrix} :\ x_2 \in \mathbb{R},\ x_3 \in \mathbb{R} \right\} \\ &= gen \left\{ \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}. \end{aligned}$$

Observe que $dimE_1 = 2$, igual a la multiplicidad algebraica de la ecuación característica. De este modo, los vectores propios linealmente independientes de A son

$$\left\{ \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Este es un hecho importante en la teoría de diagonalización de matrices, como lo veremos en la sección 7.3.

Ejemplo 7.2.4. Dada la matriz

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

encuentre sus valores propios y los espacios propios correspondientes a cada valor propio.

Solución

Primero, calculamos el polinomio característico de la matriz A.

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & -3 & 3-\lambda \end{vmatrix}$$

$$\begin{aligned} \det(A - \lambda I) &= (-\lambda)\left[(-\lambda)(3-\lambda)+3\right] - [-1] \\ &= (-\lambda)\left(\lambda^2 - 3\lambda + 3\right) + 1 \\ &= -\lambda^3 + 3\lambda^2 - 3\lambda + 1 \\ &= -\left(\lambda^3 - 3\lambda^2 + 3\lambda - 1\right) \\ &= -(\lambda - 1)^3 . \end{aligned}$$

Al igualar a cero el polinomio característico, obtenemos que

$$-(\lambda - 1)^3 = 0,$$

con lo cual el valor propio es $\lambda = 1$ y tiene multiplicidad algebraica 3 (la raíz se repite 3 veces). Al resolver el sistema homogéneo

$$(A - \lambda I)\, v = 0, \text{con}\, \lambda = 1 \text{ y } v = [x_1, x_2, x_3]^t,$$

se tiene la forma de los vectores propios asociados a dicho valor. Esto es,

$$\left[\begin{array}{rrr|r} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & -3 & 2 & 0 \end{array}\right].$$

Al aplicar operaciones elementales por filas se obtiene

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

de modo que

$$v = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix};$$

por tanto, el espacio propio es:

$$E_1 = \left\{ x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \,:\, x_3 \in \mathbb{R} \right\} = gen \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

En este caso, la multiplicidad algebraica es 3, diferente de la dimensión de E_1 que es 1.

Ejemplo 7.2.5. Pruebe que si A es una matriz invertible de orden $n \times n$ y λ es un valor propio no nulo, entonces λ^{-1} es un valor propio de A^{-1}.

Solución

Sabemos que si λ es un valor propio de A, entonces $Av = \lambda v$, para $v \neq 0$. Como A es invertible y $\lambda \neq 0$, se tiene que

$$\begin{aligned} A^{-1} A\, v &= A^{-1} \lambda\, v \\ v &= \lambda\, A^{-1} v \\ \lambda^{-1} v &= A^{-1} v. \end{aligned}$$

De modo que λ^{-1} es un valor propio de A^{-1}.

Ejemplo 7.2.6. Pruebe que si λ es un valor propio de la matriz A, entonces λ^2 es un valor propio de A^2.

Solución

Como λ es un valor propio de A, entonces existe un vector $v \neq 0$ tal que

$$Av = \lambda v. \tag{7.2.8}$$

Ahora, multiplicamos por A la ecuación 7.2.8 y obtenemos

$$A^2 v = A(\lambda v) = \lambda(Av) = \lambda(\lambda v) = \lambda^2 v;$$

luego, $A^2 v = \lambda^2 v$, o sea, λ^2 es un valor propio de A^2.

Ejemplo 7.2.7. Pruebe que si $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores propios de la matriz A, entonces $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$.

Solución

El lado izquierdo de la ecuación $\det(A - \lambda I) = 0$ se puede expresar en la forma

$$\det(A - \lambda I) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n),$$

donde $\lambda_1, \ldots, \lambda_n$ son los ceros de $P(\lambda)$. Si hacemos $\lambda = 0$ en esta última ecuación, obtenemos

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n,$$

como se quería probar.

7.3 Diagonalización

Definición 7.3.1. *Dos matrices A y B de $n \times n$ son semejantes o equivalentes si existe una matriz invertible P de orden $n \times n$ tal que*

$$B = P^{-1} A P.$$

Teorema 7.3.1. *Si A y B son matrices semejantes de tamaño $n \times n$, entonces A y B tienen los mismos valores propios.*

Demostración. Para probar que las matrices A y B tienen los mismos valores propios, basta demostrar que poseen la misma ecuación característica. En efecto, como A y B son semejantes, entonces $B = P^{-1}AP$ y

$$\begin{aligned}
\det(B - \lambda I) &= \det\big(P^{-1}AP - \lambda I\big) \\
&= \det\big(P^{-1}AP - P^{-1}(\lambda I)P\big) \\
&= \det\big(P^{-1}(A - \lambda I)P\big) \\
&= \det\big(P^{-1}\big)\det\big(A - \lambda I\big)\det P \\
&= \det\big(A - \lambda I\big),
\end{aligned}$$

ya que $\det\big(P^{-1}\big)\det P = 1$. En consecuencia, A y B tienen los mismos valores propios.

Definición 7.3.2. *Una matriz A de orden $n \times n$ es diagonalizable si existe una matriz invertible P de orden $n \times n$ tal que $D = P^{-1}AP$, donde D es una matriz diagonal cuyas componentes son los valores propios de A, o sea,*

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

La matriz P se construye poniendo como columnas los vectores propios que generan los espacios característicos.

Teorema 7.3.2. *Una matriz A de $n \times n$ es diagonalizable si y solo si tiene n vectores propios linealmente independientes.*

Demostración. Supongamos que A es diagonalizable, es decir, existe una matriz invertible P tal que $D = P^{-1}AP$ es diagonal.

Sean $v_1, v_2, \ldots, v_n$ las columnas de P. Así,

$$\begin{aligned} PD &= \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{bmatrix}. \end{aligned}$$

Como $AP = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_n \end{bmatrix}$ y $P^{-1}AP = D$, entonces $AP = PD$, lo cual implica que

$$\begin{bmatrix} Av_1 & Av_2 & \cdots & Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{bmatrix}.$$

En otras palabras, $Av_i = \lambda_i p_i$ para cada vector columna v_i. Esto significa que los vectores columna v_i de P son los vectores propios de A. Además, como P es invertible, sus vectores columna son linealmente independientes. Por tanto, A tiene n vectores propios que son linealmente independientes.

Recíprocamente, suponga que A tiene n vectores propios linealmente independientes, $v_1, v_2, \ldots, v_n$, con valores propios asociados $\lambda_1, \lambda_2, \ldots, \lambda_n$. Sea

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix},$$

donde

$$v_1 = \begin{bmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \end{bmatrix}, \quad v_2 = \begin{bmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{n2} \end{bmatrix}, \cdots, v_n = \begin{bmatrix} p_{1n} \\ p_{2n} \\ \vdots \\ p_{nn} \end{bmatrix}.$$

Entonces, P es invertible, puesto que sus columnas son linealmente independientes. Ahora,

$$AP = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}.$$

Note que la columna i de AP es

$$A \begin{bmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{bmatrix} = Av_i = \lambda_i v_i.$$

Luego, AP es la matriz cuya columna i es $\lambda_i v_i$, es decir,

$$AP = \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{bmatrix}.$$

Por otro lado,

$$\begin{aligned} PD &= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{bmatrix} = AP. \end{aligned}$$

En consecuencia, $AP = PD$ o, de modo equivalente, $P^{-1}AP = D$, ya que P es invertible, con lo cual A es diagonalizable.

Corolario 7.3.1. *Si la matriz A de orden $n \times n$ tiene n valores propios diferentes, entonces A es diagonalizable.*

Demostración. La conclusión es una consecuencia inmediata del teorema 7.2.1, pues en este caso tenemos n vectores propios $v_1, v_2, \ldots, v_n$ de la matriz A, los cuales son linealmente independientes, ya que cada uno está asociado a un valores propio diferente $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Este corolario nos proporciona un criterio para diagonalizar matrices: una matriz A de orden $n \times n$ es diagonalizable si y solo si tiene n vectores propios linealmente independientes, de modo que si hay n valores propios diferentes, la matriz es diagonalizable. También podemos afirmar que la matriz A es diagonalizable si la multiplicidad algebraica de cada valor propio λ es igual a la dimensión del espacio propio, esto es, $dimE_\lambda$, la cual se conoce como la *multiplicidad geométrica* de λ.

Ejemplo 7.3.1. Dada la matriz

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

determine si es diagonalizable; en caso de serlo, halle una matriz P tal que $P^{-1}AP = D$.

Solución

Esta es la matriz del ejemplo 7.2.2. Los vectores propios son linealmente independientes, de manera que la matriz es invertible y, además,

$$P = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix},$$

de modo que

$$D = \frac{1}{6}\begin{bmatrix} 2 & 2 & 2 \\ -3 & 0 & 3 \\ 1 & -2 & 2 \end{bmatrix}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Es claro que la diagonal está conformada por los valores propios de la matriz original.

Ejemplo 7.3.2. Dada la matriz

$$A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix},$$

determine si es diagonalizable; en caso de serlo, halle una matriz P tal que $P^{-1}AP = D$.

Solución

Los vectores propios asociados a los valores propios $\lambda = 1$ y $\lambda = 10$ son, respectivamente,

$$\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

Estos vectores son linealmente independientes (verificarlo). Por tanto, la matriz invertible P y su inversa P^{-1}, en este caso, están dadas por:

$$P = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \\ -2 & 0 & 1 \end{bmatrix} \quad \text{y} \quad P^{-1} = \frac{1}{9}\begin{bmatrix} 1 & 1 & -4 \\ -4 & 5 & -2 \\ 2 & 2 & 1 \end{bmatrix}.$$

Además,

$$
\begin{aligned}
D &= P^{-1}AP\\
&= \frac{1}{9}\begin{bmatrix} 1 & 1 & -4\\ -4 & 5 & -2\\ 2 & 2 & 1\end{bmatrix}\begin{bmatrix} 5 & 4 & 2\\ 4 & 5 & 2\\ 2 & 2 & 2\end{bmatrix}\begin{bmatrix} 1 & -1 & 2\\ 0 & 1 & 2\\ -2 & 0 & 1\end{bmatrix}\\
&= \begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 10\end{bmatrix}.
\end{aligned}
$$

Debemos aclarar que si cambiamos el orden de las columnas de la matriz P, también se modifica el orden de los valores propios que conforman la matriz diagonal.

Ejemplo 7.3.3. Determine si la matriz

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

es diagonalizable.

Solución

Esta matriz solo tiene un único vector propio linealmente independiente, de modo que no es diagonalizable, pues no se puede construir la matriz invertible P.

Ejemplo 7.3.4. Pruebe que si $D = P^{-1}AP$, entonces se cumple que $A^n = PD^nP^{-1}$, para todo n entero positivo.

Solución

$$
\begin{aligned}
D^n &= (P^{-1}AP)^n = (P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)\\
&= P^{-1}A(PP^{-1})\cdots P^{-1}A(PP^{-1})P\\
&= P^{-1}\underbrace{A\cdots A}_{n\,veces}P = P^{-1}A^nP.
\end{aligned}
$$

Al despejar A^n se tiene:

$$PD^nP^{-1} = PP^{-1}A^nPP^{-1} = A^n.$$

Este ejemplo presenta un método fácil para calcular potencias de matrices diagonalizables; el siguiente nos muestra cómo proceder en estos casos.

Ejemplo 7.3.5. Calcular A^{100} y A^{155}, donde

$$A = \begin{bmatrix} 41 & -30 \\ 56 & -41 \end{bmatrix}.$$

Solución

Calculemos los valores y vectores propios:

$$\det(A - \lambda I) = \begin{vmatrix} 41 - \lambda & -30 \\ 56 & -41 - \lambda \end{vmatrix} = 0 \Longrightarrow \lambda^2 - 1 = 0,$$

de lo cual se deduce que

$$\lambda_1 = 1 \Longrightarrow v_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \text{y} \quad \lambda_2 = -1 \Longrightarrow v_2 = \begin{bmatrix} 5 \\ 7 \end{bmatrix}.$$

Como $A = PDP^{-1}$, donde

$$D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad P = \begin{bmatrix} 3 & 5 \\ 4 & 7 \end{bmatrix} \quad \text{y} \quad P^{-1} = \begin{bmatrix} 7 & -5 \\ -4 & 3 \end{bmatrix},$$

entonces, se tiene que:

$$A^{100} = \begin{bmatrix} 3 & 5 \\ 4 & 7 \end{bmatrix} \begin{bmatrix} 1^{100} & 0 \\ 0 & (-1)^{100} \end{bmatrix} \begin{bmatrix} 7 & -5 \\ -4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Ahora,

$$A^{155} = \begin{bmatrix} 3 & 5 \\ 4 & 7 \end{bmatrix} \begin{bmatrix} 1^{155} & 0 \\ 0 & (-1)^{155} \end{bmatrix} \begin{bmatrix} 7 & -5 \\ -4 & 3 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ 56 & -41 \end{bmatrix}.$$

De esta manera, $A^{155} = A$.

7.4 Diagonalización de matrices simétricas

Como lo hemos mencionado antes, para poder diagonalizar una matriz de tamaño $n\times n$ se requiere que esta tenga n vectores propios. Vamos a probar que en el caso de las matrices simétricas se puede garantizar que esos n vectores propios existen.

Iniciemos esta sección presentando una importante característica de los valores propios de una matriz simétrica.

Teorema 7.4.1. *Los valores propios de una matriz simétrica de entradas reales son reales.*

Demostración. Sea A una matriz simétrica de tamaño $n \times n$ con entradas reales y sea $P(\lambda) = \det(A-\lambda I)$ el polinomio característico de A.

Supongamos que $P(\lambda)$ tiene una raíz compleja $\lambda = \alpha + \beta i$; por lo tanto, su conjugada $\bar{\lambda} = \alpha-\beta i$ es también una raíz del polinomio característico.

Consideremos la matriz

$$\begin{aligned}
B &= (A-(\alpha+\beta i)I)(A-(\alpha-\beta i)I)\\
&= AA-(\alpha-\beta i)A-(\alpha+\beta i)A+(\alpha+\beta i)(\alpha-\beta i)I\\
&= A^2-2\alpha A+(\alpha^2+\beta^2)I\\
&= (A-\alpha I)^2+\beta^2 I.
\end{aligned}$$

Como el determinante de un producto de matrices es el producto de los determinantes de cada matriz, es claro que

$$\det B = \det(A-(\alpha+\beta i)I)\det(A-(\alpha-\beta i)I) = 0$$

y, en consecuencia, B es una matriz singular. Como A es simétrica, entonces

$$(A-(\alpha+\beta i)I)^t = (A^t-(\alpha+\beta i)I^t) = (A-(\alpha+\beta i)I),$$

mientras que

$$(A - (\alpha - \beta i)I)^t = (A^t - (\alpha - \beta i)I^t) = (A - (\alpha - \beta i)I).$$

Por tanto

$$\begin{aligned} B^t &= (A - (\alpha - \beta i)I)^t(A - (\alpha + \beta i)I)^t \\ &= (A - (\alpha - \beta i)I)(A - (\alpha + \beta i)I) \\ &= AA - (\alpha - \beta i)A - (\alpha + \beta i)A + (\alpha + \beta i)(\alpha - \beta i)I \\ &= A^2 - 2\alpha A + (\alpha^2 + \beta^2)I \\ &= B, \end{aligned}$$

lo cual muestra que B es una matriz simétrica.

Al ser B singular, existe un vector $v \neq 0$ tal que $Bv = 0$ y, por lo tanto,

$$\begin{aligned} v^tBv &= 0 \\ &= v^t((A - \alpha I)^2 + \beta^2 I)v \\ &= v^t(A - \alpha I)^2 v + v^t\beta^2 I v \\ &= v^t(A - \alpha I)^2 v + \beta^2 v^t v \\ &= (v^t(A - \alpha I)^t)((A - \alpha I)v) + \beta^2 v^t v \\ &= ((A - \alpha I)v)^t((A - \alpha I)v) + \beta^2 v^t v. \end{aligned}$$

Como ambos sumandos son no negativos, entonces concluimos que $\beta = 0$, ya que $v \neq 0$ y, por lo tanto, los valores propios son reales.

Vamos ahora a probar una propiedad especial de los vectores propios de una matriz simétrica.

Teorema 7.4.2. *Los vectores propios de una matriz simétrica, asociados a valores propios diferentes, son ortogonales.*

Demostración. Sea A una matriz simétrica y $\lambda_1 \neq \lambda_2$ valores propios con vectores propios v_1 y v_2 respectivamente, es decir:

$$\begin{aligned} Av_1 &= \lambda_1 v_1 \\ Av_2 &= \lambda_2 v_2. \end{aligned}$$

Si consideramos el producto interno $v_1 \cdot v_2 = v_1^t v_2$, entonces

$$\begin{aligned} \lambda_1(v_1 \cdot v_2) &= (\lambda_1 v_1)^t v_2 \\ &= (Av_1)^t v_2 \\ &= v_1^t A^t v_2 \\ &= v_1^t A v_2 \\ &= v_1^t \lambda_2 v_2 \\ &= \lambda_2 v_1^t v_2 \\ &= \lambda_2(v_1 \cdot v_2). \end{aligned}$$

Por lo tanto,

$$\lambda_1(v_1 \cdot v_2) - \lambda_2(v_1 \cdot v_2) = (\lambda_1 - \lambda_2) v_1 \cdot v_2 = 0;$$

pero, al ser $\lambda_1 - \lambda_2 \neq 0$, se tiene que $v_1 \cdot v_2 = 0$ y, así, los vectores son ortogonales.

Recordemos que en el caso en el cual una matriz es diagonalizable, las columnas de la matriz P, tal que $D = P^{-1}AP$, son los n vectores propios linealmente independientes de la matriz A. Al considerar el caso en el cual la matriz A es simétrica, hemos visto que sus vectores propios, correspondientes a valores propios distintos, son ortogonales. Esto nos motiva para dar la siguiente definición.

Definición 7.4.1. *Una matriz C de tamaño $n \times n$ es ortogonal si y solo si $C^{-1} = C^t$.*

Lo anterior significa que la inversa es igual a su transpuesta. De la definición 7.4.1 podemos obtener el siguiente teorema.

Teorema 7.4.3. *Si C es una matriz ortogonal, entonces:*

1. $\det C = \pm 1$.
2. *Las columnas de C son vectores ortogonales.*
3. *Si C es ortogonal y A es simétrica, entonces $C^{-1}AC$ es simétrica.*

Demostración. Si C es ortogonal, entonces:

1. Al ser C una matriz ortogonal, obtenemos que

$$CC^{-1} = CC^t = I$$

y, por lo tanto,

$$\det(CC^{-1}) = \det(CC^t) = 1.$$

En consecuencia,

$$\det(C)\det(C^t) = (\det C)^2 = 1$$

y así,

$$\det(C) = \pm 1.$$

2. Sea $C = [v_1, v_2, \ldots, v_n]$ una matriz ortogonal, donde v_j representa la columna j de C; entonces

$$C^tC = C^{-1}C = I_n = \begin{bmatrix} v_1^t \\ v_2^t \\ \vdots \\ v_n^t \end{bmatrix} \begin{bmatrix} v_1, & v_2, & \ldots, & v_n \end{bmatrix}$$

y, por lo tanto,

$$I_n = \begin{bmatrix} v_1^t v_1 & v_1^t v_2 & \ldots & v_1^t v_n \\ v_2^t v_1 & v_2^t v_2 & \ldots & v_2^t v_n \\ \vdots & \vdots & \ldots & \vdots \\ v_n^t v_1 & v_n^t v_2 & \ldots & v_n^t v_n \end{bmatrix},$$

de lo cual concluimos que

$$v_i^t v_j = \begin{cases} 1, & \text{si } i = j; \\ 0, & \text{si } i \neq j. \end{cases}$$

Pero como

$$v_i^t v_j = v_i \cdot v_j,$$

se tiene que estos vectores son ortogonales.

3. Para probar la última propiedad, vamos a calcular $(C^{-1}AC)^t$.

$$\begin{aligned}
(C^{-1}AC)^t &= C^t(C^{-1}A)^t \\
&= C^t(A^t(C^{-1})^t) \\
&= C^t(A(C^t)^t) \\
&= C^t(AC) \\
&= (C^tAC) \\
&= (C^{-1}AC),
\end{aligned}$$

lo cual prueba que esta matriz es simétrica.

En la definición 6.3.1 del concepto de base en los espacios vectoriales establecimos que un espacio vectorial de dimensión finita puede tener diferentes bases. Si tenemos las coordenadas de un vector x cualquiera con respecto a una base ordenada, surge el problema de determinar las coordenadas de este mismo vector con respecto a una nueva base ordenada; esto es lo que hemos estudiado como el problema del cambio de bases. En esta sección regresamos a dicho problema, pero le damos un enfoque un poco diferente, para aplicarlo posteriormente al problema de la diagonalización.

Sean $B_1 = \{u_1, u_2, \ldots, u_n\}$ y $B_2 = \{v_1, v_2, \ldots, v_n\}$ dos bases ordenadas y sea x un vector cuyas coordenadas en B_1 y B_2 están dadas por $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ y $(\beta_1, \beta_2, \ldots, \beta_n)$ respectivamente. Entonces,

$$\sum_{i=1}^{n} \alpha_i u_i = \sum_{i=1}^{n} \beta_i v_i.$$

Esta ecuación podemos escribirla en la forma

$$\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ u_1 & u_2 & \cdots & u_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

De esta manera, si llamamos B_1 a la matriz cuyas columnas son los vectores de la base B_1 y B_2 a la matriz cuyas columnas son los vectores de la base B_2, entonces la relación anterior puede escribirse como

$$B_1 \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = B_2 \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Como B_1 y B_2 son bases, las matrices B_1 y B_2 son invertibles y, por lo tanto, podemos escribir las relaciones

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = B_1^{-1} B_2 \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \qquad \text{o} \qquad \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = B_2^{-1} B_1 \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Así, con la matriz $P = B_1^{-1}B_2$ pasamos las coordenadas de la segunda base a la primera y con la matriz $P^{-1} = B_2^{-1}B_1$ pasamos las coordenadas de la primera base a la segunda.

Hay otro problema que vamos a necesitar en esta sección. Hemos visto que si tenemos una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$, respecto a las bases B_1 de $\mathbb{R}^n$ y B_1' de $\mathbb{R}^m$, la transformación queda determinada de manera única por una matriz A_1; ahora nos planteamos el problema de establecer cómo se modifica esta matriz cuando hacemos un cambio a bases B_2 de $\mathbb{R}^n$ y B_2' de $\mathbb{R}^m$.

Sean X y X' los vectores columna que dan las coordenadas de un vector $x \in \mathbb{R}^n$ respecto a las bases B_1 y B_2, respectivamente, y sean Y y Y' los vectores columna que dan las coordenadas de un vector $y = T(x) \in \mathbb{R}^m$ respecto a las bases B_1' y B_2', respectivamente. Entonces, se tiene que

$$\begin{aligned}
X' &= P_1 X, \text{ siendo } P_1 \text{ la matriz de cambio de base de } B_1 \text{ a } B_2;\\
Y' &= P_2 Y, \text{ siendo } P_2 \text{ la matriz de cambio de base de } B_1' \text{ a } B_2';\\
Y &= A_1 X \ \text{ es la ecuación de } T(x) \text{ referida a las bases de } B_1 \text{ y } B_1';\\
Y' &= A_2 X' \ \text{ es la ecuación de } T(x) \text{ referida a las bases de } B_2 \text{ y } B_2',
\end{aligned}$$

donde

$$P_1 = B_2^{-1} B_1 \text{ y } P_2 = (B_2')^{-1} B_1',$$

como se ve en la Figura 7.1.

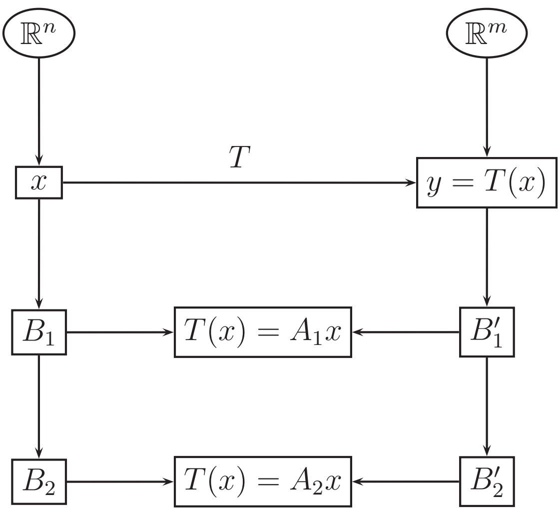


Figura 7.1. Efecto del cambio de base en una transformación

Sustituyendo las tres primeras ecuaciones en la última tenemos:

$$P_2 A_1 X = A_2 P_1 X.$$

Se deduce entonces que

$$P_2 A_1 = A_2 P_1,$$

y de este modo concluimos que

$$A_1 = P_2^{-1} A_2 P_1 \ \text{ o } \ A_2 = P_2 A_1 P_1^{-1}.$$

Cuando la transformación es un endomorfismo, es decir, una transformación $T : \mathbb{R}^n \to \mathbb{R}^n$, la situación es como se observa en la Figura 7.2. En este caso, las matrices A_1 y A_2 representan al endomorfismo en las bases B_1 y B_2 respectivamente.

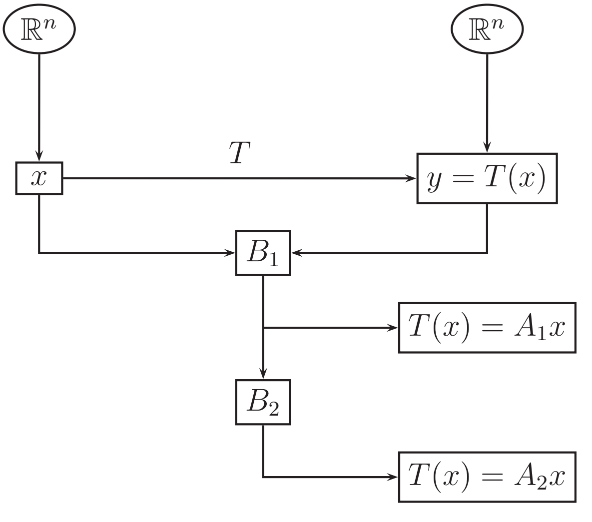


Figura 7.2. Efecto del cambio de base en un endomorfismo

Sean X y X' los vectores columna que dan las coordenadas de un vector $x \in \mathbb{R}^n$ respecto a las bases B_1 y B_2, respectivamente, y sean Y y Y' los vectores columna que dan las coordenadas de un vector $y = T(x) \in \mathbb{R}^n$ respecto a las bases B_1 y B_2, respectivamente.

Entonces,

$$\begin{aligned}
X' &= PX, \text{ siendo } P \text{ la matriz de cambio de base de } B_1 \text{ a } B_2;\\
Y' &= PY, \text{ siendo } P \text{ la matriz de cambio de base de } B_1 \text{ a } B_2;\\
Y &= A_1X \ \text{ es la ecuación de } T(x) \text{ referida a la base } B_1;\\
Y' &= A_2X' \ \text{ es la ecuación de } T(x) \text{ referida a la base } B_2,
\end{aligned}$$

donde

$$P = B_2^{-1}B_1.$$

Al sustituir los valores de X, X' y Y' en la última igualdad, obtenemos que

$$PA_1X = A_2PX,$$

a partir de lo cual concluimos que

$$PA_1 = A_2P;$$

despejando, podemos escribirlo de las formas

$$A_1 = P^{-1}A_2P \text{ o } A_2 = PA_1P^{-1}.$$

Se ve aquí que las matrices A_1 y A_2 son semejantes.

Con estos elementos formulemos ahora el principal resultado de esta sección.

Teorema 7.4.4. *Para toda matriz simétrica real A, existe una matriz ortogonal C, tal que la matriz $D = CAC^{-1}$, semejante a A, es una matriz diagonal.*

Este teorema lo que hace es afirmar que las matrices simétricas son diagonalizables y que, además, la matriz de semejanza es una matriz ortogonal.

Demostración. La prueba de este teorema la realizamos mediante inducción matemática aplicada sobre el tamaño de la matriz.

Veamos lo que ocurre en el caso de matrices simétricas de tamaño 2×2. Sea $A = [A_1, A_2]$ una matriz simétrica real, con columnas A_1 y A_2, y sea $T(x) : \mathbb{R}^2 \to \mathbb{R}^2$ el endomorfismo representado por A con respecto a la base canónica $B_1 = I$, como se ve en la Figura 7.3.

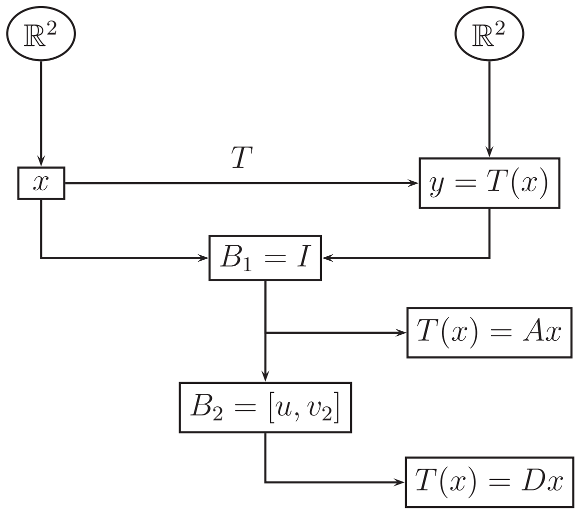


Figura 7.3. Cambio de base en un endomorfismo en $\mathbb{R}^2$

Sea D la matriz que representa a dicho endomorfismo respecto a la nueva base ortogonal ordenada $B_2 = \{u, v_2\}$, donde u es el vector propio normalizado asociado al valor propio λ_1. Recordemos que las matrices que representan al endomorfismo en las distintas bases son matrices semejantes y su relación está dada por $CA = DC$, siendo C la matriz no singular que corresponde al cambio de la primera a la segunda base, esto es,

$$C = B_2^{-1} B_1 = B_2^{-1} I = B_2^{-1}.$$

De esta forma, $C^{-1} = B_2$ y como B_2 es ortogonal, entonces

$$B_2^{-1} = B_2^t,$$

con lo cual obtenemos que la matriz C^t cumple la relación

$$C^t = (B_2^{-1})^t = (B_2^t)^t = B_2 = C^{-1},$$

es decir, la matriz C también es ortogonal.

Al ser C ortogonal, y teniendo en cuenta la relación entre A y D, tenemos que

$$D = CAC^{-1} = CAC^t$$

y, en consecuencia,

$$\begin{aligned} D^t &= (AC^t)^t C^t \\ &= (C^t)^t A^t C^t \\ &= CAC^{-1} \\ &= D, \end{aligned}$$

donde la penúltima igualdad se obtiene por la simetría de A y por ser C ortogonal. Esta relación lo que nos garantiza es que la matriz D es simétrica.

Veamos finalmente que D es diagonal.

$$\begin{aligned} D &= CAC^t = B_2^t A B_2 \\ &= \begin{bmatrix} u^t \\ v_2^t \end{bmatrix} [A_1, A_2][u, v_2] \\ &= \begin{bmatrix} u^t \\ v_2^t \end{bmatrix} \begin{bmatrix} A_1^t \\ A_2^t \end{bmatrix} [u, v_2] \\ &= \begin{bmatrix} u^t \\ v_2^t \end{bmatrix} \begin{bmatrix} A_1^t u & A_1^t v_2 \\ A_2^t u & A_2^t v_2 \end{bmatrix}. \end{aligned}$$

Al ser u un vector propio asociado al valor propio λ_1 y teniendo en cuenta que la matriz A es simétrica,

$$T(u) = \lambda_1 u = Au = A^t u$$

y, de este modo,

$$\lambda_1 u = \begin{bmatrix} \lambda_1 u_{11} \\ \lambda_1 u_{21} \end{bmatrix} = \begin{bmatrix} A_1^t \\ A_2^t \end{bmatrix} u = \begin{bmatrix} A_1^t u \\ A_2^t u \end{bmatrix}.$$

Al sustituir esta relación en la expresión D, concluimos que

$$D = \begin{bmatrix} u^t \\ v_2^t \end{bmatrix} \begin{bmatrix} \lambda_1 u_{11} & A_1^t v_2 \\ \lambda_1 u_{21} & A_2^t v_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 u^t u & b_{21} \\ \lambda_1 v_2^t u & b_{22} \end{bmatrix}.$$

Como la base B_2 es ortogonal y u está normalizado,

$$u^t u = 1 \text{ y } v_2^t u = 0,$$

con lo cual

$$D = \begin{bmatrix} \lambda_1 & b_{12} \\ 0 & b_{22} \end{bmatrix}$$

y por la simetría de D, entonces $b_{12} = 0$, con lo cual

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & b_{22} \end{bmatrix}.$$

Por último, como A y D tienen los mismos valores propios (son semejantes), $b_{22} = \lambda_2$, es decir,

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Supongamos ahora que las matrices simétricas de orden n son diagonalizables. Veamos que las de orden $n+1$ también lo son.

Sea $A = [A_1, A_2, \ldots, A_{n+1}]$ una matriz simétrica de orden $n+1$, con columnas $A_1, A_2, \ldots, A_{n+1}$ y $T : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ el endomorfismo representado por A con respecto a la base canónica $B_1 = I$. Sea D la matriz que representa a dicho endomorfismo respecto a la nueva base ortogonal ordenada $B_2 = \{u, v_2, v_3, \ldots v_{n+1}\}$, donde u es el vector propio normalizado asociado al valor propio λ_1. Recordemos que las matrices que representan al endomorfismo en las distintas bases son matrices semejantes y su relación está dada por $CA = DC$, siendo C la matriz no singular que corresponde al cambio de la primera a la segunda base, esto es,

$$C = B_2^{-1} B_1 = B_2^{-1} I = B_2^{-1};$$

de esta forma, $C^{-1} = B_2$ y como B_2 es ortogonal, entonces

$$B_2^{-1} = B_2^t,$$

con lo cual obtenemos que

$$C^t = (B_2^{-1})^t = (B_2^t)^t = B_2 = C^{-1},$$

es decir, la matriz C también es ortogonal.

De igual modo que en el caso de orden 2, se cumple que como C es ortogonal, y teniendo en cuenta la relación entre A y D, entonces

$$D = CAC^{-1} = CAC^t$$

y, en consecuencia,

$$\begin{aligned} D^t &= (AC^t)^t C^t \\ &= (C^t)^t A^t C^t \\ &= CAC^{-1} \\ &= D, \end{aligned}$$

donde la penúltima igualdad se obtiene por la simetría de A y por ser C ortogonal.

De este modo vemos que D es simétrica; solo nos falta ver que es diagonal.

$$\begin{aligned} D &= CAC^t = B_2^t A B_2 \\ D &= \begin{bmatrix} u^t \\ v_2^t \\ v_3^t \\ \vdots \\ v_{n+1}^t \end{bmatrix} [A_1, A_2, A_3, \ldots, A_{n+1}][u, v_2, v_3, \ldots, v_{n+1}] \end{aligned}$$

$$D = \begin{bmatrix} u^t \\ v_2^t \\ v_3^t \\ \vdots \\ v_{n+1}^t \end{bmatrix} \begin{bmatrix} A_1^t \\ A_2^t \\ A_3^t \\ \vdots \\ A_{n+1}^t \end{bmatrix} [u, v_2, v_3, \ldots, v_{n+1}]$$

$$D = \begin{bmatrix} u^t \\ v_2^t \\ v_3^t \\ \vdots \\ v_{n+1}^t \end{bmatrix} \begin{bmatrix} A_1^t u & A_1^t v_2 & A_1^t v_3 & \ldots & A_1^t v_{n+1} \\ A_2^t u & A_2^t v_2 & A_2^t v_3 & \ldots & A_2^t v_{n+1} \\ A_3^t u & A_3^t v_2 & A_3^t v_3 & \ldots & A_3^t v_{n+1} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ A_{n+1}^t u & A_{n+1}^t v_2 & A_{n+1}^t v_3 & \ldots & A_{n+1}^t v_{n+1} \end{bmatrix}.$$

Como u es un vector propio asociado al valor propio λ_1 y teniendo en cuenta que la matriz A es simétrica, obtenemos que

$$T(u) = \lambda_1 u = Au = A^t u$$

y, de este modo,

$$\lambda_1 u = \begin{bmatrix} \lambda_1 u_{11} \\ \lambda_1 u_{21} \\ \lambda_1 u_{31} \\ \vdots \\ \lambda_1 u_{(n+1)1} \end{bmatrix} = \begin{bmatrix} A_1^t \\ A_2^t \\ A_3^t \\ \vdots \\ A_{n+1}^t \end{bmatrix} u = \begin{bmatrix} A_1^t u \\ A_2^t u \\ A_3^t u \\ \vdots \\ A_{n+1}^t u \end{bmatrix}.$$

Sustituyendo esta relación en la última relación para D, podemos escribir la matriz D en la siguiente forma:

$$D = \begin{bmatrix} u^t \\ v_2^t \\ v_3^t \\ \vdots \\ v_{n+1}^t \end{bmatrix} \begin{bmatrix} \lambda_1 u_{11} & A_1^t v_2 & A_1^t v_3 & \ldots & A_1^t v_{n+1} \\ \lambda_1 u_{21} & A_2^t v_2 & A_2^t v_3 & \ldots & A_2^t v_{n+1} \\ \lambda_1 u_{31} & A_3^t v_2 & A_3^t v_3 & \ldots & A_3^t v_{n+1} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ \lambda_1 u_{(n+1)1} & A_{n+1}^t v_2 & A_{n+1}^t v_3 & \ldots & A_{n+1}^t v_{n+1} \end{bmatrix}.$$

Efectuando el producto, obtenemos finalmente la matriz

$$D = \left[\begin{array}{c|ccc} \lambda_1 u^t u & b_{12} & \dots & b_{1(n+1)} \\ \hline \lambda_1 v_2^t u & & & \\ \vdots & & Q_n & \\ \lambda_1 v_{n+1}^t u & & & \end{array}\right],$$

donde Q_n es una matriz de tamaño $n \times n$. Además, como la base $B_2 = \{u, v_2, v_3, \dots v_{n+1}\}$ es ortogonal y el vector u está normalizado,

$$u^t u = 1 \text{ y } v_j^t u = 0, \text{ con } j = 1, \dots, n+1,$$

obtenemos que

$$D = \left[\begin{array}{c|ccc} \lambda_1 & b_{12} & \dots & b_{1(n+1)} \\ \hline 0 & & & \\ \vdots & & Q_n & \\ 0 & & & \end{array}\right]$$

y por la simetría de la matriz D, $b_{1j} = 0$, con $j = 1, \dots, n+1$. De este modo, D tiene la forma

$$D = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q_n & \\ 0 & & & \end{array}\right].$$

Como la matriz D es simétrica y semejante a la matriz A, tiene los mismos valores propios de A; en consecuencia, la matriz Q_n es simétrica, de orden n y sus valores propios son $\lambda_2, \lambda_3, \dots, \lambda_{n+1}$. Por lo tanto, si aplicamos la hipótesis de inducción, Q_n es diagonalizable, es decir, existe una matriz R ortogonal, $n \times n$, tal que

$$R^t Q_n R = \begin{bmatrix} \lambda_2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_{n+1} \end{bmatrix}.$$

Sea ahora la matriz P de tamaño $(n+1)\times(n+1)$ dada por

$$P = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & R & \\ 0 & & & \end{array}\right].$$

P es ortogonal y se cumple que

$$P^tDP = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & R^t & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q_n & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & R & \\ 0 & & & \end{array}\right]$$

y, por lo tanto, obtenemos que

$$P^tDP = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & R^tQ_nR & \\ 0 & & & \end{array}\right] = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{n+1} \end{bmatrix}.$$

En consecuencia, la matriz D es diagonalizable y como

$$\begin{aligned} D &= B_2^tAB_2, \\ P^tDP &= R^tB_2^tAB_2R \\ &= (B_2R)^tA(B_2R) \\ &= \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{n+1} \end{bmatrix}. \end{aligned}$$

Así, la matriz ortogonal B_2R diagonaliza a la matriz A.

Ejemplo 7.4.1. Dada la matriz simétrica

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

hallar una matriz ortogonal que diagonalice a A.

Solución

El teorema anterior nos garantiza que la matriz A es diagonalizable. En primer lugar, calculamos los valores propios de la matriz, y luego los vectores propios asociados a cada uno de estos valores propios; una vez normalizado cada uno de los vectores propios, obtenemos la matriz que diagonaliza a A.

El polinomio característico de A está dado por

$$P(\lambda) = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{vmatrix} = -\lambda^3 + 4\lambda^2 - 3\lambda = \lambda(1-\lambda)(\lambda-3).$$

De esta forma, los valores propios de la matriz A son $\lambda = 0$, $\lambda = 1$ y $\lambda = 3$.

Para el valor propio $\lambda = 0$, el vector propio se obtiene al resolver el sistema homogéneo

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

La forma escalonada reducida del sistema está dada por

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; \quad \text{luego,} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -z \\ z \\ z \end{bmatrix}$$

y, por lo tanto, el vector propio normalizado para $\lambda = 0$ es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

Para el valor propio $\lambda = 1$, el vector propio se obtiene al resolver el sistema homogéneo

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

La forma escalonada reducida del sistema está dada por

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; \quad \text{luego,} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -z \\ z \end{bmatrix}$$

y, por lo tanto, el vector propio normalizado para $\lambda = 1$ es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

Para el valor propio $\lambda = 3$, el vector propio se obtiene al resolver el sistema homogéneo

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

La forma escalonada reducida del sistema está dada por

$$\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; \quad \text{luego,} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2z \\ z \\ z \end{bmatrix}$$

y, por lo tanto, el vector propio normalizado para $\lambda = 3$ es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

Con los vectores propios normalizados, obtenemos entonces la matriz ortogonal C que diagonaliza a A. Esta es

$$C = \begin{bmatrix} -\frac{\sqrt{3}}{3} & 0 & \frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \end{bmatrix}.$$

A modo de prueba, basta verificar que C^tAC es la matriz diagonal con los valores propios en la diagonal:

$$C^tAC = \begin{bmatrix} -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{3}}{3} & 0 & \frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \end{bmatrix}$$

$$C^tAC = \begin{bmatrix} -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \end{bmatrix} \begin{bmatrix} 0 & 0 & \sqrt{6} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Es fácil verificar también que C es ortogonal.

7.5 Formas cuadráticas

Las formas cuadráticas constituyen una de las tantas aplicaciones de la diagonalización y una poderosa herramienta para el estudio de las cónicas y la solución de problemas de optimización. Iniciamos esta sección con la siguiente definición.

Definición 7.5.1. *Sea V un espacio vectorial real de dimensión n. Una forma cuadrática $Q : V \to \mathbb{R}$ es una aplicación de V en $\mathbb{R}$ que transforma a un vector $x \in V$ mediante la expresión $Q(x) = x^tAx$,* *donde A es una matriz simétrica de orden n y x es el vector de coordenadas de x respecto a una base de V. De esta forma, $Q(x)$ está dada por*

$$Q(x) = [x_1, \ldots, x_n] \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A la última ecuación se le denomina *expresión matricial de la forma cuadrática.*

Si se efectúan las operaciones, obtenemos la relación

$$Q(x) = a_{11}x_1^2 + \ldots a_{nn}x_n^2 + 2a_{12}x_1x_2 + \ldots + 2a_{n(n-1)n}x_{n-1}x_n$$

o lo que es equivalente:

$$Q(x) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}x_ix_j,$$

donde debe tenerse en cuenta que $a_{ij} = a_{ji}$. Es claro que $Q(x)$ es un polinomio cuadrático en las variables x_ix_j.

Pasar de la representación matricial a la representación polinómica de la forma cuadrática o viceversa es bien sencillo. Para ello solo basta observar que los coeficientes de la diagonal, a_{ii}, son los coeficientes de x_i^2, mientras que las mitades de los coeficientes con variables diferentes, x_ix_j, son los elementos de la posición a_{ij} y a_{ji} de la matriz simétrica A.

Ejemplo 7.5.1. Dada la forma cuadrática

$$Q(x) = 3x_1^2 + 2x_2^2 + 4x_1x_2 - 3x_1x_3 - 5x_2x_3,$$

halle su forma matricial.

Solución

Según el comentario anterior, los elementos de la diagonal de la matriz son 3, 2 y 0 y, además, $a_{12} = a_{21} = 4/2$, $a_{13} = a_{31} = -3/2$, $a_{23} = a_{32} = -5/2$ y, en consecuencia, la forma matricial está dada por

$$Q(x) = [x_1, x_2, x_3] \begin{bmatrix} 3 & 2 & -\frac{3}{2} \\ 2 & 2 & -\frac{5}{2} \\ -\frac{3}{2} & -\frac{5}{2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Es claro que la dimensión del espacio vectorial V es 3.

Uno de los aspectos importantes a estudiar en una forma cuadrática es el signo que esta toma para distintos valores del vector x. Teniendo en cuenta su signo, se acostumbra a clasificar las formas cuadráticas en cinco grupos:

- $Q(x)$ es definida positiva si $Q(x) > 0 \ \forall x \neq 0$.
- $Q(x)$ es definida negativa si $Q(x) < 0 \ \forall x \neq 0$.
- $Q(x)$ es semidefinida positiva si $Q(x) \geq 0 \ \forall x \neq 0$.
- $Q(x)$ es semidefinida negativa si $Q(x) \leq 0 \ \forall x \neq 0$.
- $Q(x)$ es indefinida si $Q(x) > 0$ para algún $x \neq 0 \in V$ y $Q(x) < 0$ para algún $x \neq 0 \in V$.

El estudio del signo de una forma cuadrática y, por ende, el tipo de forma cuadrática, puede hacerse de modo sencillo si la matriz simétrica que representa a la forma cuadrática tiene una estructura particular. Para esto debemos primero justificar que hay ciertos invariantes en la representación matricial de la forma cuadrática. Siguiendo este camino, probamos el siguiente teorema.

Teorema 7.5.1. *El cambio de base dado por la ecuación $x = Cx'$ transforma a la forma cuadrática $Q(x) = x^t Ax$ en $Q(x) = (x')^t Bx'$, siendo B la matriz simétrica $B = C^t AC$.*

Demostración. Al sustituir la ecuación $x = Cx'$ en la expresión de la forma cuadrática obtenemos que

$$Q(x) = x^t A x = (Cx')^t A(Cx') = (x')^t (C^t AC)x' = (x')^t Bx',$$

con $B = C^t AC$. Además, es claro que

$$B^t = (C^t AC)^t = (AC)^t (C^t)^t = C^t A^t C = C^t AC = B,$$

con lo cual se muestra que B es una matriz simétrica y, por lo tanto, $(x')^t Bx'$ es una forma cuadrática.

El teorema anterior lo que nos permite concluir es que una forma cuadrática puede representarse mediante distintas matrices simétricas, dependiendo de la base que elijamos para el espacio vectorial.

Definición 7.5.2. *Dos matrices simétricas A y B son congruentes si existe una matriz no singular C tal que $B = C^t AC$.*

Teorema 7.5.2. *Si A y B son matrices congruentes, entonces sus determinantes tienen el mismo signo.*

Demostración. Si A y B son congruentes, entonces $B = C^t AC$ y, por lo tanto,

$$\det B = \det C^t \det A \det C = (\det C)^2 \det A,$$

puesto que $\det C^t = \det C$. En consecuencia, los determinantes de A y B tienen el mismo signo.

En la sección previa hemos probado que toda matriz simétrica se puede diagonalizar mediante una matriz ortogonal; esto es equivalente a afirmar que toda matriz simétrica es congruente con una matriz diagonal. En este caso la forma cuadrática solo tendrá elementos de la forma x_i^2 en su representación polinómica.

Es importante aclarar que una matriz simétrica puede ser congruente a varias matrices diagonales; de hecho, al conjunto de expresiones

$$Q(x) = (x')^t Bx' = b_1(x'_1)^2 + b_2(x'_2)^2 + \ldots + b_n(x'_n)^2$$

se le conoce como *expresiones diagonales* de la forma cuadrática.

Al tener una forma cuadrática diferentes expresiones diagonales, los coeficientes b_i de cada expresión dependen de la base que hayamos escogido para la representación. No obstante, como lo probaremos en el teorema 7.5.3, el número de coeficientes positivos y el número de coeficientes negativos en todas las expresiones diagonales es el mismo. Esto da origen a la siguiente definición.

Definición 7.5.3. *Dada una expresión diagonal de una forma cuadrática, si p es el número de coeficientes positivos y q es el número de coeficientes negativos de la expresión diagonal, al par (p, q) se le denomina signatura de la forma cuadrática.*

Es claro que la expresión $p + q$ corresponde al número de filas no nulas de la matriz diagonal congruente con A y, por lo tanto, representa el rango de las matrices diagonales congruentes con A.

Teorema 7.5.3 (Ley de inercia de Sylvester). *Todas las expresiones diagonales de una forma cuadrática tienen la misma signatura.*

Demostración. Supongamos que hay dos expresiones diagonales que difieren en el número de coeficientes positivos y negativos. En este caso la forma cuadrática estaría dada por

$$\begin{aligned} Q(x) &= b_1(x'_1)^2 + \ldots + b_r(x'_r)^2 - b_{r+1}(x'_{r+1})^2 - \ldots - b_t(x'_t)^2 \\ Q(x) &= c_1(x''_1)^2 + \ldots + c_s(x''_s)^2 - c_{s+1}(x''_{s+1})^2 - \ldots - c_t(x''_t)^2, \end{aligned}$$

donde $b_i > 0$, con $i = 1, \ldots t$; $c_i > 0$, con $i = 1, \ldots t$ y $r < s$. Como ambas son representaciones de la misma forma cuadrática, al igualarlas obtenemos la expresión

$$\sum_{i=1}^{r} b_i(x_i')^2 + \sum_{i=s+1}^{t} c_i(x_i'')^2 = \sum_{i=1}^{s} c_i(x_i'')^2 + \sum_{i=r+1}^{t} b_i(x_i')^2.$$

Como las variables $\{x_i'\}$ y $\{x_i''\}$ están relacionadas mediante la ecuación de cambio de base $x' = Cx''$, tomemos $x_i' = 0$ para $i = 1, \ldots r$ y $x_i'' = 0$ para $i = s+1, \ldots t$. Entonces, en la ecuación para el cambio de base se obtiene un sistema homogéneo de r ecuaciones con s incógnitas. Al ser $s > r$, el sistema tiene soluciones no triviales.

Reemplazando estas condiciones obtenemos la ecuación

$$0 = \sum_{i=1}^{s} c_i(x_i'')^2 + \sum_{i=r+1}^{t} b_i(x_i')^2.$$

Si tenemos en cuenta que $c_i > 0$ y $b_i > 0$, entonces $x_i'' = 0$ para $i = 1, \ldots s$, con lo cual $x_i'' = 0$ para $i = 1, \ldots s, s+1, \ldots, t$ y, por lo tanto, solo tendríamos la solución trivial, en contradicción con el hecho de que el sistema tiene infinitas soluciones, luego, r y s deben ser iguales.

La importancia del teorema anterior radica en que si conocemos la signatura de una expresión diagonal de la forma cuadrática, entonces podemos saber qué tipo de forma cuadrática tenemos.

Teorema 7.5.4. *Sea la forma cuadrática $Q(x) = x^t Ax$ en n variables y sea*

$$Q(x) = b_1(x_1')^2 + b_2(x_2')^2 + \ldots + b_n(x_n')^2$$

una expresión diagonal de la forma cuadrática con signatura (p, q); entonces:

- *$Q(x)$ es definida positiva si y solo si $p = n$ y $q = 0$.*
- *$Q(x)$ es definida negativa si y solo si $p = 0$ y $q = n$.*
- *$Q(x)$ es semidefinida positiva si y solo si $q = 0$.*
- *$Q(x)$ es semidefinida negativa si y solo si $p = 0$.*
- *$Q(x)$ es indefinida si y solo si $p \neq 0$ y $q \neq 0$.*

Demostración. Vamos a probar solo la segunda propiedad, las demás quedan propuestas como ejercicios. Supongamos que $Q(x)$ es definida negativa; entonces, $\forall x \in \mathbb{R}^n$, $Q(x) < 0$. Sea x un vector de la forma $x = (0, \ldots, 1, \ldots, 0)$, con el 1 en la posición i para $i = 1, \ldots, n$; así, $Q(x) = b_i < 0$. Como esto es válido para $i = 1, \ldots, n$, el número de coeficientes negativos en la expresión diagonal de la forma cuadrática es n, es decir, $p = 0$ y $q = n$.

Supongamos ahora que $p = 0$ y $q = n$; entonces,

$$Q(x) = b_1(x_1')^2 + b_2(x_2')^2 + \ldots + b_n(x_n')^2$$

solo se anula para el vector nulo; en los demás casos será siempre negativa y, en consecuencia, $Q(x)$ es definida negativa.

Una de las maneras de hallar una expresión diagonal para una forma cuadrática $Q(x) = x^tAx$ es aprovechar el hecho de que toda matriz simétrica es diagonalizable mediante una matriz ortogonal y que la matriz diagonal está formada por los valores propios de A. En este caso, la signatura de la expresión diagonal queda determinada por el signo de los valores propios de A, y el teorema anterior puede formularse nuevamente en términos del signo de los valores propios de la matriz A.

Teorema 7.5.5. *Dada la forma cuadrática $Q(x) = x^tAx$, esta se transforma mediante el cambio de base $x = Cx'$ con la matriz C ortogonal, en la expresión diagonal $Q(x) = (x')^tDx'$, donde D es* *la matriz diagonal* $D = (\lambda_1, \lambda_2, \ldots \lambda_n)$, *es decir, la matriz diagonal con los valores propios de* A.

Demostración. Al ser A una matriz simétrica, esta es diagonalizable y la matriz de cambio de base es una matriz C ortogonal; por lo tanto, si $x = Cx'$, entonces

$$Q(x) = x^t Ax = (Cx')^t A(Cx') = (x')^t (C^t AC)x' = (x')^t Dx'.$$

De esta forma,

$$Q(x) = \lambda_1(x'_1)^2 + \lambda_2(x'_2)^2 + \ldots + \lambda_n(x'_n)^2.$$

El teorema anterior permite estudiar la signatura de una forma cuadrática cuando se tiene su expresión diagonal a partir del análisis de los signos de los valores propios. Así, el teorema 7.5.4 puede enunciarse nuevamente de la siguiente manera.

Teorema 7.5.6. *Sea la forma cuadrática* $Q(x) = x^t Ax$ *en* n *variables y sea*

$$Q(x) = \lambda_1(x'_1)^2 + \lambda_2(x'_2)^2 + \ldots + \lambda_n(x'_n)^2$$

la expresión diagonal formada con los valores propios de A *de la forma cuadrática con signatura* (p, q); *entonces:*

- $Q(x)$ *es definida positiva si y solo si todos sus valores propios son positivos* ($p = n$ *y* $q = 0$).
- $Q(x)$ *es definida negativa si y solo si todos sus valores propios son negativos* ($p = 0$ *y* $q = n$).
- $Q(x)$ *es semidefinida positiva si y solo si sus valores propios son nulos o positivos* ($q = 0$).

- *$Q(x)$ es semidefinida negativa si y solo si sus valores propios son nulos o negativos ($p = 0$).*

- *$Q(x)$ es indefinida si y solo si tiene valores propios positivos y negativos ($p \neq 0$ y $q \neq 0$).*

Ejemplo 7.5.2. Sea

$$Q(x) = 3x_1 + 2x_2 - 2\sqrt{6}x_1x_2$$

una forma cuadrática en dos variables; determine:

1. Su forma matricial.

2. Su signatura y tipo.

3. La matriz C de cambio de base.

4. Calcular $x = Cx'$.

5. Calcular C^tAC.

Solución
La forma matricial de $Q(x)$ está dada por

$$Q(x) = [x_1, x_2]\begin{bmatrix} 3 & -\sqrt{6} \\ -\sqrt{6} & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Los valores propios de la matriz de la forma cuadrática corresponden a las soluciones de la ecuación

$$\det\begin{bmatrix} 3-\lambda & -\sqrt{6} \\ -\sqrt{6} & 2-\lambda \end{bmatrix} = (3-\lambda)(2-\lambda) - 6 = \lambda(\lambda - 5) = 0.$$

Por lo tanto, los valores propios son $\lambda = 0$ y $\lambda = 5$, de lo cual concluimos que la forma lineal es semidefinida positiva (sus valores propios son nulos o positivos).

Para hallar la matriz de transición C debemos calcular cada uno de los vectores propios. El correspondiente a $\lambda = 0$ se obtiene a partir de la matriz

$$\begin{bmatrix} 3 & -\sqrt{6} \\ -\sqrt{6} & 2 \end{bmatrix},$$

cuya forma escalonada reducida es

$$\begin{bmatrix} 1 & -\frac{\sqrt{6}}{3} \\ 0 & 0 \end{bmatrix},$$

con lo cual obtenemos la solución

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ \sqrt{6} \end{bmatrix} t.$$

Al normalizar este vector tenemos

$$u_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 2 \\ \sqrt{6} \end{bmatrix}.$$

Para $\lambda = 5$ el vector propio se consigue a partir de la matriz

$$\begin{bmatrix} -2 & -\sqrt{6} \\ -\sqrt{6} & -3 \end{bmatrix},$$

cuya forma escalonada reducida es

$$\begin{bmatrix} 1 & \frac{\sqrt{6}}{2} \\ 0 & 0 \end{bmatrix},$$

con lo cual obtenemos la solución

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -3 \\ \sqrt{6} \end{bmatrix} t.$$

Al normalizar este vector tenemos

$$u_2 = \frac{1}{\sqrt{15}} \begin{bmatrix} -3 \\ \sqrt{6} \end{bmatrix}.$$

Con estos vectores, la matriz C queda determinada por

$$C = \begin{bmatrix} \frac{\sqrt{10}}{5} & -\frac{\sqrt{15}}{5} \\ \frac{\sqrt{15}}{5} & \frac{\sqrt{10}}{5} \end{bmatrix}$$

y, en consecuencia, la ecuación para cambio a las nuevas coordenadas $x' = C^t x$ es

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{10}}{5} & \frac{\sqrt{15}}{5} \\ -\frac{\sqrt{15}}{5} & \frac{\sqrt{10}}{5} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

De esta manera, la expresión diagonal de la forma cuadrática será $Q(x) = 5(x_2')^2$.

Finalmente, verificamos que $C^tAC = D$.

En este caso tenemos

$$\begin{aligned} C^tAC &= \begin{bmatrix} \frac{\sqrt{10}}{5} & \frac{\sqrt{15}}{5} \\ -\frac{\sqrt{15}}{5} & \frac{\sqrt{10}}{5} \end{bmatrix} \begin{bmatrix} 3 & -\sqrt{6} \\ -\sqrt{6} & 2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{10}}{5} & -\frac{\sqrt{15}}{5} \\ \frac{\sqrt{15}}{5} & \frac{\sqrt{10}}{5} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ -\sqrt{15} & \sqrt{10} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{10}}{5} & -\frac{\sqrt{15}}{5} \\ \frac{\sqrt{15}}{5} & \frac{\sqrt{10}}{5} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}, \end{aligned}$$

que es la matriz diagonal con los valores propios de A.

7.5.1 Ecuaciones cuadráticas en $\mathbb{R}^2$ y $\mathbb{R}^3$

Una aplicación de las ideas presentadas sobre diagonalización de matrices simétricas, es el estudio de las ecuaciones cuadráticas en $\mathbb{R}^2$ y $\mathbb{R}^3$ y con centro en el origen, las cuales pueden escribirse en la forma $Q(x) = d$, donde d es cualquier número real y $Q(x)$ es una forma cuadrática. Recordemos que las gráficas de las ecuaciones cuadráticas es lo que conocemos como *secciones cónicas* o simplemente *cónicas*. El problema de identificar y graficar una cónica es sencillo cuando conocemos los ejes de simetría de esta; no obstante, si la ecuación cuadrática posee un término de la forma xy o término cruzado, lo que tenemos es una sección cónica cuyos ejes de simetría están rotados un determinado ángulo con respecto a los ejes usuales xy.

El problema de identificación se hace fácil si logramos identificar los ejes de simetría de la sección cónica. En este caso, al diagonalizar la matriz de la forma cuadrática asociada a la ecuación cuadrática obtenemos la ecuación de la cónica escrita en su forma estándar. Con el fin de identificar y graficar estas secciones cónicas, iniciamos con la siguiente definición.

Definición 7.5.4. *Una ecuación cuadrática en* $\mathbb{R}^2$ *con términos lineales es una expresión de la forma*

$$a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2 + b_1x_1 + b_2x_2 = c, \tag{7.5.1}$$

donde $c \in \mathbb{R}$, $c \geq 0$.

Esta ecuación representa una sección cónica sobre la cual se han realizado dos transformaciones geométricas, una rotación y una traslación de los ejes de simetría de la cónica. La forma cuadrática

$$Q(x) = a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2$$

es la que determina la rotación, mientras que

$$b_1x_1 + b_2x_2$$

es la expresión responsable de la traslación de ejes.

Los comentarios anteriores nos sugieren que si logramos identificar los parámetros que determinan estas trasformaciones, podemos identificar la sección cónica correspondiente.

La ecuación 7.5.1 puede escribirse, usando matrices, en la forma

$$[x_1, x_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [b_1, b_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = c,$$

lo cual es equivalente a la ecuación

$$x^t Ax + Bx = c, \tag{7.5.2}$$

donde A es la matriz simétrica dada por

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}, \quad B = [b_1, b_2] \quad \text{y} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Por los resultados de la sección anterior sabemos que la forma cuadrática $x^t Ax$ es diagonalizable; por lo tanto, podemos aplicar un cambio de base de la forma $x = Cx'$, donde C es la matriz ortogonal que se obtiene con los vectores propios de A normalizados. Así, la ecuación 7.5.2 se transforma en la ecuación

$$\begin{aligned} (x')^t C^t AC(x') + BC(x') &= c \\ (x')^t D(x') + BC(x') &= c, \end{aligned} \tag{7.5.3}$$

donde D es la matriz diagonal formada con los valores propios λ_1 y λ_2 de A, y como se mencionó previamente, las columnas de C son los vectores propios de A normalizados.

Después de realizar los productos de la ecuación 7.5.3 podemos ver que no aparecen términos de la forma x_1x_2 y esta puede escribirse como

$$\lambda_1(x_1')^2 + \lambda_2(x_2')^2 + d_1x_1' + d_2x_2' = c.$$

En este punto vamos a introducir un concepto similar al concepto de *signatura* de una forma cuadrática, que será de gran utilidad para la identificación de la ecuación cuadrática.

Definición 7.5.5. *Dada la forma cuadrática x^tAx, la inercia de la forma cuadrática $In(Q(x))$ es la terna $In(Q(x)) = (p, q, n)$, donde p es el número de valores propios positivos de A, q es el número de valores propios negativos de A y n es el número de valores propios nulos de la matriz A.*

Si tenemos en cuenta la definición de la inercia de una forma cuadrática, la Tabla 7.1 muestra la clasificación de las cónicas en dos variables, dependiendo de los valores de su inercia.

Sección cónica	**Inercia**
Elipse	$(2, 0, 0)$
Hipérbola	$(1, 1, 0)$
Parábola	$(1, 0, 1)$
Parábola	$(0, 1, 1)$

Tabla 7.1. Clasificación de las cónicas según su inercia

Por último, si escribimos la ecuación cuadrática en la forma

$$\lambda_1\left(x_1' + \frac{d_1}{2\lambda_1}\right)^2 + \lambda_2\left(x_2' + \frac{d_2}{2\lambda_2}\right)^2 = c + \frac{d_1^2}{4\lambda_1} + \frac{d_2^2}{4\lambda_2},$$

para $\lambda_1 \neq 0$ y $\lambda_2 \neq 0$, y hacemos la traslación de ejes

$$x_1'' = x_1' + \frac{d_1}{2\lambda_1} \quad \text{y} \quad x_2'' = x_2' + \frac{d_2}{2\lambda_2},$$

obtenemos la forma canónica de la sección cónica correspondiente:

$$\lambda_1\left(x_1''\right)^2 + \lambda_2\left(x_2''\right)^2 = c + \frac{d_1^2}{4\lambda_1} + \frac{d_2^2}{4\lambda_2}.$$

Ejemplo 7.5.3. Describa la cónica dada por la ecuación

$$x_1^2 + x_2^2 - 6x_1x_2 + 2\sqrt{2}x_1 - 4\sqrt{2}x_2 = 8$$

y realice su gráfica.

Solución

La forma matricial de esta ecuación está dada por

$$[x_1, x_2]\begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [2\sqrt{2}, -4\sqrt{2}]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 8.$$

Como la matriz de la forma cuadrática es

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

sus valores propios corresponden a la solución de la ecuación

$$\det\begin{bmatrix} 1-\lambda & -3 \\ -3 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 - 9 = 0,$$

lo cual es equivalente a $(\lambda+2)(\lambda-4) = 0$ y, por lo tanto, los valores propios son $\lambda = 4$ y $\lambda = -2$, de tal modo que $In(Q(x)) = (1, 1, 0)$. En consecuencia, la cónica es una hipérbola, como se muestra en la Figura 7.4.

Para averiguar cuáles son los ejes principales de la hipérbola, calculamos los vectores propios correspondientes a cada valor propio.

Para $\lambda = -2$,

$$\begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \Rightarrow x_1 = x_2 \Rightarrow u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

mientras que para $\lambda = 4$,

$$\begin{bmatrix} -3 & -3 \\ -3 & -3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow x_1 = -x_2 \Rightarrow u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

y, por lo tanto, la matriz de cambio de base es

$$C = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$[x_1', x_2'] \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} + \sqrt{2}[2, -4] \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = 8.$$

Luego de realizar los productos y simplificar, llegamos a la expresión

$$-2(x_1')^2 + 4(x_2')^2 - 2x_1' - 6x_2' = 8.$$

Estos nuevos ejes están orientados según los vectores unitarios

$$u_1 = \frac{1}{\sqrt{2}}(1, 1) \text{ y } u_2 = \frac{1}{\sqrt{2}}(-1, 1),$$

y la ecuación resultante puede escribirse en la forma

$$-(x_1')^2 - x_1' + 2(x_2')^2 - 3x_2' = 4,$$

que es lo mismo que la ecuación

$$-\left(x_1' + \frac{1}{2}\right)^2 + 2\left(x_2' - \frac{3}{4}\right)^2 = 4 - \frac{1}{4} + \frac{9}{8} = \frac{39}{8}\,.$$

Si ahora hacemos el cambio de ejes

$$x_1'' = x_1' + \frac{1}{2} \quad \text{y} \quad x_2'' = x_2' - \frac{3}{4},$$

llegamos a la ecuación

$$-\left(x_1''\right)^2 + 2\left(x_2''\right)^2 = \frac{39}{8}$$

o, de manera equivalente,

$$-\frac{\left(x_1''\right)^2}{\frac{39}{8}} + \frac{\left(x_2''\right)^2}{\frac{39}{16}} = 1,$$

que es la forma canónica de una hipérbola cortando al eje x_2'' en $\sqrt{39/16} \approx 1,5612$. La gráfica de la cónica y sus nuevos ejes pueden verse en la Figura 7.4.

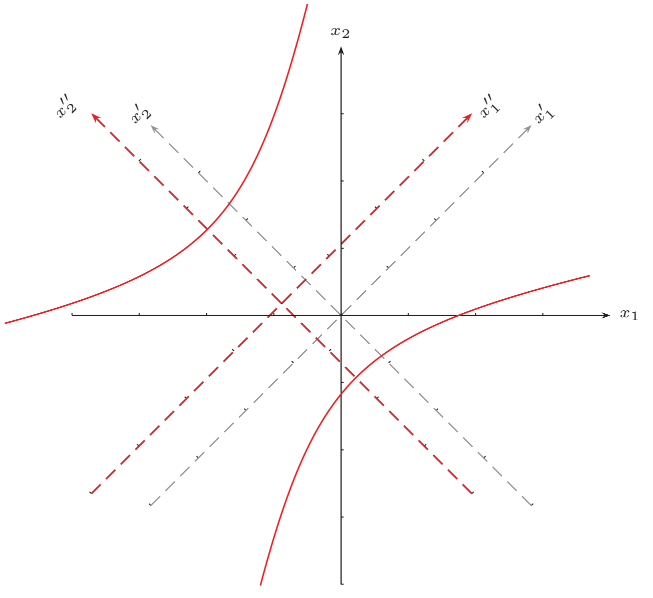


Figura 7.4. Gráfica de la sección cónica del ejercicio 7.5.3

Ejemplo 7.5.4. Describa la cónica dada por la ecuación

$$2x_1^2 + 10x_2^2 + 6x_1x_2 + \sqrt{10}x_1 + 5\sqrt{10}x_2 = 9$$

y realice su gráfica.

Solución

La forma matricial de la ecuación cuadrática está dada por

$$[x_1, x_2]\begin{bmatrix} 2 & 3 \\ 3 & 10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [\sqrt{10}, 5\sqrt{10}]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 9.$$

Como la matriz de la forma cuadrática es

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 10 \end{bmatrix},$$

los valores propios corresponden a la solución de la ecuación

$$\det\begin{bmatrix} 2-\lambda & 3 \\ 3 & 10-\lambda \end{bmatrix} = (\lambda - 2)(\lambda - 10) - 9 = 0,$$

lo cual es equivalente a $(\lambda-11)(\lambda-1)=0$ y, por lo tanto, los valores propios son $\lambda=11$ y $\lambda=1$, de tal modo que $In(Q(x))=(2,0,0)$. En consecuencia, la cónica correspondiente es una elipse.

Para averiguar cuáles son los ejes principales de esta elipse, calculamos los vectores propios correspondientes a cada valor propio.

Para $\lambda=1$,

$$\begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} \Rightarrow x_1 = -3x_2 \Rightarrow u_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \end{bmatrix},$$

mientras que para $\lambda=11$,

$$\begin{bmatrix} -9 & 3 \\ 3 & -1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -\frac{1}{3} \\ 0 & 0 \end{bmatrix} \Rightarrow x_1 = \frac{1}{3}x_2 \Rightarrow u_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Por lo tanto, la matriz de cambio de base es

$$C = \frac{1}{\sqrt{10}} \begin{bmatrix} -3 & 1 \\ 1 & 3 \end{bmatrix}$$

$$[x_1', x_2'] \begin{bmatrix} 1 & 0 \\ 0 & 11 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} + \sqrt{10}[1,5]\frac{1}{\sqrt{10}} \begin{bmatrix} -3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = 9.$$

Tras de realizar los productos y simplificar, llegamos a la expresión

$$(x_1')^2 + 11(x_2')^2 + 2x_1' + 16x_2' = 9.$$

Estos nuevos ejes están orientados según los vectores unitarios

$$u_1 = \frac{1}{\sqrt{10}}(-3,1) \ \text{ y } u_2 = \frac{1}{\sqrt{10}}(1,3),$$

y la ecuación resultante puede escribirse en la forma

$$(x_1'+1)^2 + 11\left(x_2' + \frac{8}{11}\right)^2 = 9 + 1 + \frac{64}{11} = \frac{174}{11}\,.$$

Si ahora hacemos el cambio de ejes

$$x_1'' = x_1' + 1 \ \text{ y } \ x_2'' = x_2' + \frac{8}{11},$$

llegamos a la ecuación

$$(x_1'')^2 + 11\,(x_2'')^2 = \frac{174}{11}$$

o, de manera equivalente,

$$\frac{(x_1'')^2}{\frac{174}{11}} + \frac{(x_2'')^2}{\frac{174}{121}} = 1,$$

que es la forma canónica de una elipse con semiejes

$$\sqrt{\frac{174}{11}} \approx 3{,}977 \ \text{ y } \ \frac{\sqrt{174}}{11} \approx 1{,}199$$

sobre los ejes x_1'' y x_2''. La gráfica de dicha elipse, junto con sus nuevos ejes, puede verse en la Figura 7.5.

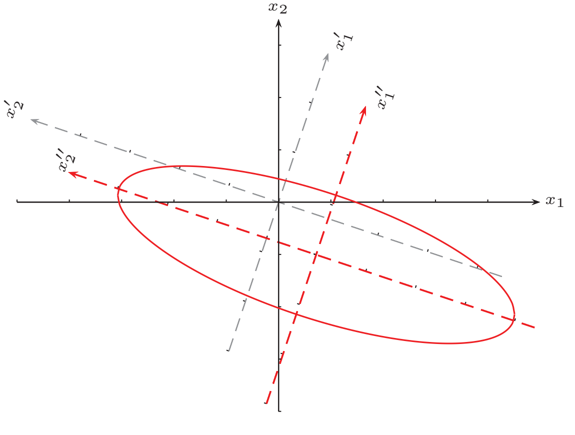


Figura 7.5. Gráfica de la sección cónica del ejercicio 7.5.4

Ejemplo 7.5.5. Describa la cónica dada por la ecuación

$$9x_1^2 + 16x_2^2 + 24x_1x_2 - 20x_1 + 15x_2 = 9$$

y realice su gráfica.

Solución

La forma matricial de la ecuación cuadrática está dada por

$$[x_1, x_2] \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [-20, 15] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 9.$$

Como la matriz de la forma cuadrática es

$$A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix},$$

los valores propios corresponden a la solución de la ecuación

$$\det \begin{bmatrix} 9-\lambda & 12 \\ 12 & 16-\lambda \end{bmatrix} = (\lambda - 9)(\lambda - 16) - 144 = 0,$$

lo cual es equivalente a $(\lambda - 25)\lambda = 0$ y, por lo tanto, los valores propios son $\lambda = 25$ y $\lambda = 0$, de tal modo que $In(Q(x)) = (1, 0, 1)$. En consecuencia, la cónica correspondiente es una parábola, como se observa en la Figura 7.6.

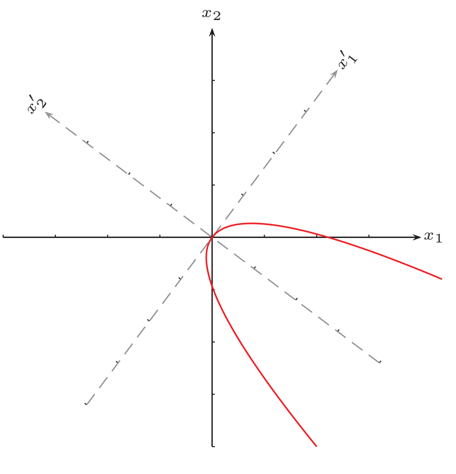


Figura 7.6. Gráfica de la sección cónica del ejercicio 7.5.5

Para averiguar cuáles son los ejes principales de esta parábola, calculamos los vectores propios correspondientes a cada valor propio.

Para $\lambda = 25$,

$$\begin{bmatrix} -16 & 12 \\ 12 & -9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -3/4 \\ 0 & 0 \end{bmatrix} \Rightarrow x_1 = \frac{3}{4}x_2 \Rightarrow u_1 = \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix},$$

mientras que para $\lambda = 0$,

$$\begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \frac{4}{3} \\ 0 & 0 \end{bmatrix} \Rightarrow x_1 = -\frac{4}{3}x_2 \Rightarrow u_1 = \frac{1}{5}\begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

y, por lo tanto, la matriz de cambio de base es

$$C = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}.$$

La ecuación de la cónica toma ahora la forma

$$[x_1', x_2']\begin{bmatrix} 25 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} + [-20, 15]\frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = 9.$$

Luego de realizar los productos y simplificar, llegamos a la expresión

$$25(x_1')^2 + 25x_2' = 0.$$

Estos nuevos ejes están orientados según los vectores unitarios

$$u_1 = \frac{1}{5}(3, 4) \;\text{ y }\; u_2 = \frac{1}{5}(-4, 3),$$

y la ecuación resultante puede escribirse en la forma

$$x_2' = -(x_1')^2,$$

que es la forma canónica de una parábola. La gráfica de dicha parábola y sus ejes pueden verse en la Figura 7.6.

Definición 7.5.6. *Una ecuación cuadrática en* $\mathbb{R}^3$ *con términos lineales es una expresión de la forma*

$$a_{11}x^2+a_{22}y^2+a_{33}z+2a_{12}xy+2a_{13}xz+2a_{23}yz+b_1x+b_2y+b_3z=c, \tag{7.5.4}$$

donde $c \in \mathbb{R}$, $c \geq 0$.

Esta ecuación cuadrática representa una superficie cuadrática sobre la cual se han realizado dos transformaciones geométricas, una rotación y una traslación de los ejes de simetría de la superficie cuadrática. La forma cuadrática

$$Q(x) = a_{11}x^2 + a_{22}y^2 + a_{33}z + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz$$

es la que determina la rotación, mientras que $b_1x + b_2y + b_3z$ son los términos responsables de la traslación de ejes.

El problema de la identificación de la superficie cuadrática, así como el de la identificación de los ejes principales de esta, se realizan de manera similar a como se hizo en el caso de las cónicas. Para ello vamos primero a escribir la forma matricial de una ecuación cuadrática. En este caso, la ecuación 7.5.4 toma la forma

$$[x, y, z] \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + [b_1, b_2, b_3] \begin{bmatrix} x \\ y \\ z \end{bmatrix} = c.$$

Como lo hemos expuesto previamente, si tenemos en cuenta los teoremas sobre formas cuadráticas, podemos identificar el tipo de superficie cuadrática que la ecuación cuadrática representa. Además, si también tenemos en cuenta que la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

es una matriz simétrica, sabemos que es diagonalizable y que si C es la matriz de cambio de base tal que $X = CX'$, donde X y X' representan los vectores

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{y} \quad X' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix},$$

entonces la ecuación cuadrática puede escribirse en la forma

$$\begin{aligned} (CX')^t A(CX') + B(CX') &= c \\ (X')^t(C^t AC)X' + BCX' &= c \\ (X')^t DX' + BCX' &= c, \end{aligned}$$

donde D es la matriz diagonal formada por los vectores propios de la matriz A. En este caso, el tipo de superficie cuadrática queda determinado por la inercia de la forma cuadrática.

En la tabla 7.2 se identifica el tipo de superficie cuadrática conociendo el signo de los valores propios, esto es, identificando la inercia de la forma cuadrática asociada a la ecuación cuadrática en tres variables.

Tipo de superficie	**Inercia**
Elipsoide	$(3,0,0)$
Paraboloide hiperbólico	$(1,1,1)$
Paraboloide elíptico	$(2,0,1)$
Hiperboloide de una hoja	$(2,1,0)$
Hiperboloide de dos hojas	$(1,2,0)$
Cilindro parabólico	$(1,0,2)$

Tabla 7.2. Clasificación de las superficies cuadráticas por su inercia

La nueva ecuación de la forma cuadrática está ahora dada por

$$\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2 + e_1x' + e_2y' + e_3z' = c. \qquad (7.5.5)$$

En caso de que los valores propios sean no nulos, esta ecuación podemos escribirla nuevamente como

$$\lambda_1\left(x'+\frac{e_1}{2\lambda_1}\right)^2+\lambda_2\left(y'+\frac{e_2}{2\lambda_2}\right)^2+\lambda_3\left(z'+\frac{e_3}{2\lambda_3}\right)^2=c+f,$$

donde f está dado por

$$f=\frac{e_1^2}{4\lambda_1}+\frac{e_2^2}{4\lambda_2}+\frac{e_3^2}{4\lambda_3}\,.$$

Si finalmente hacemos las transformaciones

$$x''=x'+\frac{e_1}{2\lambda_1};\quad y''=y'+\frac{e_2}{2\lambda_2};\quad z''=z'+\frac{e_3}{2\lambda_3}\,,$$

llegamos a la expresión

$$\lambda_1\left(x''\right)^2+\lambda_2\left(y''\right)^2+\lambda_3\left(z''\right)^2=\delta,$$

donde δ está dado por

$$\delta=c+\frac{e_1^2}{4\lambda_1}+\frac{e_2^2}{4\lambda_2}+\frac{e_3^2}{4\lambda_3}\,.$$

En esta expresión, los nuevos ejes están orientados según la dirección de los vectores propios normalizados.

En la tabla 7.3 se muestran las ecuaciones canónicas de las superficies cuadráticas. En estas ecuaciones, todas las superficies están centradas en el origen del sistema de coordenadas.

Es importante aclarar que cuando hacemos la identificación de la superficie, en algunos casos se puede llegar a lo que se conoce como una *superficie cuadrática degenerada.* Tal es el caso del hiperboloide de una hoja, el cual, al simplificarlo, puede tomar la forma de alguna de las siguientes ecuaciones:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0,\quad \frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=0,$$

$$-\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=0,$$

las cuales representan conos.

Elipsoide
$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$
Hiperboloide de una hoja
$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1;\quad \frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1;\quad -\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$
Hiperboloide de dos hojas
$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1;\quad -\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1;\quad -\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$
Paraboloide elíptico
$z=\frac{x^2}{a^2}+\frac{y^2}{b^2};\quad y=\frac{x^2}{a^2}+\frac{z^2}{c^2};\quad x=\frac{y^2}{a^2}+\frac{z^2}{c^2}$
Paraboloide hiperbólico
$\pm z=\frac{x^2}{a^2}-\frac{y^2}{b^2};\quad \pm y=\frac{x^2}{a^2}-\frac{z^2}{c^2};\quad \pm x=\frac{y^2}{a^2}-\frac{z^2}{c^2}$
Cilindro parabólico
$x^2=ay+bz;\quad y^2=ax+bz;\quad z^2=ax+by$

Tabla 7.3. Ecuaciones canónicas de las superficies cuadráticas

Ejemplo 7.5.6. Identificar la superficie cuadrática correspondiente a la ecuación cuadrática $2x^2+2y^2+4z^2-4xy-8xz-8yz+8x=15$.
Solución

La forma matricial de esta ecuación cuadrática está dada por

$$[x,y,z]\begin{bmatrix} 2 & -2 & -4\\ -2 & 2 & -4\\ -4 & -4 & 4\end{bmatrix}\begin{bmatrix} x\\ y\\ z\end{bmatrix}+[8,0,0]\begin{bmatrix} x\\ y\\ z\end{bmatrix}=15.$$

Los valores propios de la forma cuadrática asociada a esta ecuación son las soluciones de la ecuación

$$\begin{vmatrix} 2-\lambda & -2 & -4 \\ -2 & 2-\lambda & -4 \\ -4 & -4 & 4-\lambda \end{vmatrix} = -\lambda^3 + 8\lambda^2 + 16\lambda - 128 = 0.$$

Si tenemos en cuenta que

$$-\lambda^3 + 8\lambda^2 + 16\lambda - 128 = (\lambda+4)(4-\lambda)(\lambda-8) = 0,$$

los valores propios de la matriz A son $\lambda = 8$, $\lambda = 4$ y $\lambda = -4$; en consecuencia, la inercia de la forma cuadrática está dada por $In(Q(x)) = (2, 1, 0)$, que corresponde a un hiperboloide de una hoja.

Para calcular la matriz de cambio de base es necesario que determinemos los vectores propios correspondientes a cada uno de los valores propios.

Si $\lambda = 8$, la matriz del sistema homogéneo correspondiente es

$$\begin{bmatrix} -6 & -2 & -4 \\ -2 & -6 & -4 \\ -4 & -4 & -4 \end{bmatrix},$$

que cuando la llevamos a la forma escalonada reducida es

$$\begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix},$$

lo cual nos indica que la solución del sistema está dada por

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \quad \Rightarrow \quad u_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}.$$

Si $\lambda = 4$, la matriz del sistema homogéneo correspondiente es

$$\begin{bmatrix} -2 & -2 & -4 \\ -2 & -2 & -4 \\ -4 & -4 & 0 \end{bmatrix},$$

que cuando la llevamos a la forma escalonada reducida es

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

lo cual nos indica que la solución del sistema está dada por

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \Rightarrow \quad u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Si $\lambda = -4$, la matriz del sistema homogéneo correspondiente es

$$\begin{bmatrix} 6 & -2 & -4 \\ -2 & 6 & -4 \\ -4 & -4 & 8 \end{bmatrix},$$

que cuando la llevamos a la forma escalonada reducida es

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

lo cual nos indica que la solución del sistema está dada por

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad u_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

En consecuencia, la matriz de cambio de base está dada por

$$C = \frac{\sqrt{6}}{6} \begin{bmatrix} -1 & -\sqrt{3} & \sqrt{2} \\ -1 & \sqrt{3} & \sqrt{2} \\ 2 & 0 & \sqrt{2} \end{bmatrix}.$$

Haciendo el cambio de variables $X = CX'$ en la ecuación 7.5.5 y teniendo en cuenta que

$$BCX' = \frac{8\sqrt{3}}{3}\,z' - 4\sqrt{2}\,y' - \frac{4\sqrt{6}}{3}\,x',$$

tras realizar las operaciones llegamos a la ecuación

$$8(x')^2 + 4(y')^2 - 4(z')^2 + \frac{8\sqrt{3}}{3}\,z' - 4\sqrt{2}\,y' - \frac{4\sqrt{6}}{3}\,x' = 15,$$

la cual puede escribirse en la forma

$$8\left(x' - \frac{\sqrt{6}}{12}\right)^2 + 4\left(y' - \frac{\sqrt{2}}{2}\right)^2 - 4\left(z' - \frac{\sqrt{3}}{3}\right)^2 = 16.$$

Por último, haciendo el cambio de coordenadas

$$x'' = x' - \frac{\sqrt{6}}{12}, \quad y'' = y' - \frac{\sqrt{2}}{2}, \quad z'' = z' - \frac{\sqrt{3}}{3},$$

obtenemos la ecuación

$$8\,(x'')^2 + 4\,(y'')^2 - 4\,(z'')^2 = 16,$$

o en forma equivalente,

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

que corresponde a un hiperboloide de una hoja y cuyos ejes principales están dados por los vectores propios normalizados.

7.6 Ejercicios propuestos capítulo 7

1. Halle el polinomio característico de las siguientes matrices

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix};\ B = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix};\ C = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix};\ D = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 4 \\ 2 & 1 & 2 \end{bmatrix}; \; F = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 2 & 1 \\ -1 & 5 & 1 \end{bmatrix}; \; G = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 1 & 5 & -1 \end{bmatrix}.$$

2. Calcule los valores propios de las siguientes matrices

$$A = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix}; \; B = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}; \; C = \begin{bmatrix} 5 & 1 \\ 0 & 4 \end{bmatrix}$$

$$D = \begin{bmatrix} 2 & 0 & -2 \\ -3 & -1 & 2 \\ 2 & 0 & -2 \end{bmatrix}; \; E = \begin{bmatrix} 2 & 4 & 3 \\ -2 & 2 & 0 \\ 3 & 0 & 2 \end{bmatrix}.$$

3. Dada la matriz

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

   a) Pruebe que es diagonalizable.

   b) Halle los vectores propios de la matriz A.

   c) Pruebe que $A = CDC^{-1}$, donde C es la matriz formada con los vectores propios de A.

4. Dada la matriz

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

   a) Pruebe que es diagonalizable.

   b) Halle los vectores propios de la matriz A.

   c) Pruebe que $A = CDC^{-1}$, donde C es la matriz formada con los vectores propios de A.

5. Dada la matriz

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

   a) Pruebe que es diagonalizable.
   b) Halle los vectores propios de la matriz A.
   c) Pruebe que $A = CDC^{-1}$, donde C es la matriz formada con los vectores propios de A.

6. Pruebe que si A y B son matrices no singulares, entonces AB y BA tienen los mismos valores propios.

7. Pruebe que si A es idempotente ($A^2 = A$), los únicos valores que pueden tomar sus valores propios son 0 y 1.

8. Calcular A^n para cada una de las matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}; \quad A = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & -2 \\ 01 & 0 & 3 \end{bmatrix}.$$

9. Halle los vectores propios de la matriz simétrica

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

   y verifique que son ortogonales.

10. Encuentre la matriz ortogonal que diagonaliza a la matriz simétrica

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

   y compruebe que $D = C^t AC$.

11. Muestre que A y A^t tienen los mismos valores propios.

12. Si A es una matriz $n \times n$, muestre que $\det(A)$ es el producto de todas las raíces del polinomio característico de A.

13. Muestre que si A es una matriz $n \times n$, entonces A es singular si 0 es un valor propio de A.

14. Muestre que si A es una matriz $n \times n$ diagonalizable, entonces A^t es diagonalizable.

15. Muestre que si A es una matriz ortogonal, entonces cA es ortogonal si y solo si $c = \pm 1$.

16. En los siguientes ejercicios, identifique la gráfica de la ecuación, localice los ejes principales, escriba la ecuación en su forma canónica y realice la gráfica.

    a) $5x^2 + 5y^2 - 6xy - 30\sqrt{2}x + 18\sqrt{2}y + 82 = 0$.
    b) $5x^2 + 12xy - 12\sqrt{13}x = 360$.
    c) $8x^2 + 8y^2 - 16xy - 33\sqrt{2}x - 31\sqrt{2}y + 70 = 0$.
    d) $6x^2 + 9y^2 - 4xy - 4\sqrt{5}x + 18\sqrt{5}y = 5$.
    e) $-x^2 + y^2 + 2\sqrt{3}xy + 6x = 0$.

17. En los siguientes ejercicios, identifique la gráfica de la ecuación, localice los ejes principales y escriba la ecuación en su forma canónica.

    a) $4x^2 + 4z^2 + 4yz + 8\sqrt{2}x + 4 = 0$.
    b) $4x^2 + 4y^2 + 8z^2 + 4xy - 4xz - 4yz + 6x - 10y + 2z = 9/4$.
    c) $-x^2 - y^2 - z^2 + 4xy + 4xz + 4yz + 3\sqrt{2}/2x - 3\sqrt{2}2y = 6$.
    d) $x^2 + y^2 + 2z^2 - 2xy + 4xz + 4yz = 16$.
    e) $x^2 + y^2 - z^2 + 4xy - 4xz - 4yz + 6x - 10y + z = 1$.
    f) $-x^2 - y^2 - z^2 + 4xy - 4xz - 4yz = 3$.
    g) $x^2 + y^2 + 2z^2 + 2yz = 1$.

Apéndice

Respuesta de los ejercicios seleccionados

A.1 Ejercicios propuestos capítulo 1, página 91

1. *a*) $CAB = \begin{bmatrix} -3 & -10 & -22 & -7 \\ 41 & 51 & 71 & 43 \end{bmatrix}$.

   b) $A^{-1}B = \frac{1}{25}\begin{bmatrix} 12 & 13 & 27 & 12 \\ -22 & 22 & 13 & 3 \\ 11 & 14 & 6 & 11 \end{bmatrix}$.

   c) $(B + D)A$ no se puede realizar.

   d) CBA no se puede realizar.

   e) $A + B$ no se puede realizar.

   f) $A^{-1}C^t = \frac{1}{25}\begin{bmatrix} 6 & -1 \\ 14 & 6 \\ -7 & 22 \end{bmatrix}$.

   g) $B^tC^t = \begin{bmatrix} -7 & 13 \\ 1 & 10 \\ 2 & 9 \\ -2 & 10 \end{bmatrix}$.

h) $(A^{-1}C^t)^t = \frac{1}{25}\begin{bmatrix} 6 & 14 & -7 \\ -1 & 6 & 22 \end{bmatrix}$.

i) $(B - D)A$ no se puede realizar.

j) $(B - A)D$ no se puede realizar.

3. *a*) $A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & -2 & -2 \\ 1 & -1 & 1 \end{bmatrix}$.

b) $A|B = \left[\begin{array}{rrr|r} 1 & 3 & 1 & -2 \\ 1 & -2 & -2 & 4 \\ 1 & -1 & 1 & 0 \end{array}\right]$.

c) $A^{-1} = \frac{1}{12}\begin{bmatrix} 4 & 4 & 4 \\ 3 & 0 & -3 \\ -1 & -4 & 5 \end{bmatrix}$.

d) La solución del sistema es $x_1 = 2/3$, $x_2 = -1/2$ y $x_3 = -7/6$.

5. *a*) La solución del sistema es: $x = 3$, $y = 1$ y $z = 1$.

b) La solución del sistema es: $x = 0$, $y = 0$ y $z = 0$.

c) E sistema tiene infinitas soluciones de la forma $x_4 = t$, $x_3 = -2 - 2t$, $x_2 = 4$ y $x_1 = -12 - 4t$.

d) La solución del sistema es: $x_1 = 1/2$, $x_2 = -55/4$, $x_3 = -17/4$ y $x_4 = 6$ y $x_5 = -1$.

7. Se deben emplear $x_1 = 13/3$, $x_2 = 4/3$, $x_3 = 1$ galones de las marcas A, B y C respectivamente.

9. Haciendo el producto de las dos matrices, se obtiene la matriz

$$\begin{bmatrix} 11550 & 12600 \\ 8500 & 9350 \end{bmatrix}.$$

La primera fila representa el costo de la lista de Ana en cada mercado, y la segunda, el costo de la lista de Sebastián en cada mercado.

11. A es singular si $\lambda = 0$, $\lambda = -4$ o $\lambda = -1$.

13. Como AB es no singular, entonces

$$(AB)(AB)^{-1} = I;$$

pero $(AB)(AB)^{-1} = A(BB^{-1})A^{-1} = AA^{-1} = I$ y, en consecuencia, A es no singular.

De manera semejante, como AB es no singular,

$$(AB)^{-1}(AB) = I;$$

pero tenemos que

$$(AB)^{-1}(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I$$

y, en consecuencia, B es no singular.

15. Tenemos que

$$(AB)^t = B^tA^t = (-B)(-A) = BA;$$

pero (AB) es simétrica si y solo si $(AB)^t = AB$, si y solo si $B^tA^t = AB$, y si y solo si $BA = AB$.

17. Como

$$(A + B)^t = A^t + B^t,$$

al ser ambas simétricas tenemos que

$$(A + B)^t = A^t + B^t = A + B$$

y, por lo tanto, $A + B$ es simétrica.

19. *a*) Rango 1: para ningún valor.

    b) Rango 2, si $a = 2$.

    c) Rango 3, para cualquier valor de $\mathbb{R} \neq 2$.

21. *a*) No hay solución si $a = -1$.

    b) Hay solución única si $a \neq -1$ y $a \neq 1$.

    c) Hay infinitas soluciones si $a = 1$.

23. El polinomio interpolador es $P(x) = \frac{5}{6}x^2 - \frac{9}{2}x + \frac{23}{3}$.

25. *a*) $P(A) = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$.

    b) $Q(A) = \begin{bmatrix} 8 & 8 \\ 12 & 20 \end{bmatrix}$.

27. $P(A) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

29. Los valores de k son $\pm\sqrt{2/5}$.

31. *a*) $A^2 = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$, $A^3 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $A^4 = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$, $A^5 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$, $A^6 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $A^7 = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

    b) $A^{2001} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$.

33. $X = \frac{1}{21}\begin{bmatrix} 91 & 35 & -21 & -7 \\ 64 & 14 & -9 & -19 \\ -66 & -21 & 6 & 15 \end{bmatrix}$.

A.2 Ejercicios propuestos capítulo 2, página 152

1. Como A es antisimétrica, $A^t = -A$ y, por lo tanto,

$$\begin{aligned} \det A^t &= \det(-A) \\ \Rightarrow \det A &= (-1)^n \det A \\ \det A(1-(-1)^n) &= 0 \end{aligned}$$

   si n es par, entonces, $2\det A = 0$, luego $\det A = 0$.

3. $\det(A^t(-3B)^{-1}(4C^{-1})) = -128/81$.

5.
$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -a & 1 & 0 & \dots & 0 \\ 0 & -a & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ & & \dots & -a & 1 \end{bmatrix}.$$

7. $\det A = 55$.

9.
$$A^{-1} = -\frac{1}{52}\begin{bmatrix} 2 & -10 & -4 \\ -14 & -8 & 2 \\ -6 & 4 & -14 \end{bmatrix}.$$

11. No tiene inversa para todo $\alpha \in \mathbb{R}$.

13. Se sabe que $adj(A)A = \det AI_n$; entonces, se tiene que

$$\begin{aligned} \det(adj(A)A) &= \det((\det A)I_n) \\ \det(adj(A))\det A &= (\det A)^n \\ \det(adj(A)) &= (\det A)^{n-1}; \end{aligned}$$

   esto último bajo el supuesto de que $\det A \neq 0$.

15. $\det A = -1$ y $\det B = 8$.

17. $\det A = 120$.

19. Supongamos que $\det A \neq 0$, como

$$adj(A)A = A(adj(A)) = \det AI_n$$

entonces

$$\begin{aligned}
&\Rightarrow \quad [Aadj(A)]^t = \det AI_n^t = \det AI_n \\
&\Rightarrow \quad (adj(A))^t A^t = \det AI_n \\
&\Rightarrow \quad (adj(A))^t A^t (adj A^t) = \det AI_n (adj A^t) \\
&\Rightarrow \quad (adj(A))^t A^t (adj A^t) = \det A (adj A^t) \\
&\Rightarrow \quad (adj(A))^t \det AI_n = \det A (adj A^t) \\
&\Rightarrow \quad \det A (adj(A))^t I_n = \det A (adj A^t) \\
&\Rightarrow \quad \det A (adj(A))^t = \det A (adj A^t) \\
&\Rightarrow \quad (adj(A))^t = (adj A^t).
\end{aligned}$$

21. Como $AA^t = I_n$, entonces,

$$\begin{aligned}
&\Rightarrow \quad \det(AA^t)) = 1 \\
&\Rightarrow \quad \det A(\det A^t) = 1 \\
&\Rightarrow \quad (\det A)^2 = 1.
\end{aligned}$$

por lo tanto, $\det A = \pm 1$.

23.

$$\det \begin{bmatrix} 3a & 3b & 3c \\ 4a & 4b & 4c \\ c & 2c & 3c \end{bmatrix} = 0.$$

A.3 Ejercicios propuestos capítulo 3, página 221

1. a. Sí es espacio vectorial.

   c. No es cerrado para la suma.

   e. No es distributivo para la suma de escalares.

   g. Sí es espacio vectorial.

   i. No es distributivo y el 1 no es elemento neutro.

3. *a*) No es subespacio.

   b) No es subespacio.

   c) No es subespacio.

   d) No es subespacio.

   e) Sí es subespacio.

   f) Sí es subespacio.

5. Son linealmente independientes y

$$(1,2) = \frac{2+\sqrt{3}}{\sqrt{2}+\sqrt{3}}(1,\sqrt{2}) + \frac{2-\sqrt{2}}{\sqrt{2}+\sqrt{3}}(-1,\sqrt{3}).$$

7. Sean v_1 y v_2 vectores en $W+U$; entonces,

$$\begin{aligned} \alpha v_1 &= \alpha(w_1+u_1) \\ &= \alpha w_1 + \alpha u_1, \quad \text{con } \alpha w_1 \in W, \alpha u_1 \in U; \end{aligned}$$

luego $\alpha v_1 \in W+U$.

Para la suma se tiene:

$$\begin{aligned} v_1+v_2 &= (w_1+u_1)+(w_2+u_2) \\ &= (w_1+w_2)+(u_1+u_2), \quad \text{con } (w_1+w_2) \in W, (u_1+u_2) \in U; \end{aligned}$$

en consecuencia, $v_1+v_2 \in W+U$ y, por lo tanto, es subespacio.

9. *a*) $(1,0) = (1,1) - (0,1)$.

   b) $(2,1) = 3/2(1,1) - 1/2(-1,1)$.

   c) $(4,3) = 2(2,1) - (0,-1)$.

   d) $(1,1) = 2/3(2,1) + 1/3(-1,1)$.

11. Es un conjunto linealmente dependiente.

13. *a*) Es linealmente independiente.

    b) Es linealmente dependiente.

15. Sea
$$\alpha_1(a_1 + z) + \ldots + \alpha_n(a_n + z) = 0;$$
entonces, se tiene que
$$\alpha_1 a_1 + \ldots + \alpha_n a_n + (\alpha_1 + \alpha_2 + \ldots + \alpha_n)z = 0.$$
Como z no está en el espacio generado por los vectores a_i, entonces se tiene que $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$; pero al ser estos vectores linealmente independientes, se cumple que
$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$$
y, en consecuencia, $\{a_1 + z, a_2 + z, \ldots, a_n + z\}$ es un conjunto linealmente independiente.

17. Los tres vectores son linealmente dependientes; pero $v_1 = (1,1,2)$ y $v_2 = (2,-1,3)$ son linealmente independientes y, por lo tanto, estos dos vectores forman un base para este subespacio. La dimensión del subespacio es 2.

19. Solo los dos primeros polinomios son linealmente independientes; por lo tanto, la base es $\{t^3 + t^2 - t + 1, t^2 + 1\}$ y la dimensión es 2.

A.4 Ejercicios propuestos capítulo 4, página 290

1. Las diagonales del paralelogramo están dadas por $\|x+y\|$ y $\|x-y\|$; por lo tanto,

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2\\ \|x-y\|^2 &= \langle x-y, x-y\rangle = \|x\|^2 - 2\langle x,y\rangle + \|y\|^2.\end{aligned}$$

   Al sumar estas dos relaciones conseguimos

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2,$$

   donde $2\|x\|^2 + 2\|y\|^2$, son los cuadrados de los lados del paralelogramo.

3. Sean α y β los ángulos que el vector

$$\vec{v} = \|\vec{a}\|\vec{b} + \|\vec{b}\|\vec{a}$$

   forma con los vectores $\vec{a}$ y $\vec{b}$ respectivamente; entonces,

$$\begin{aligned}\cos\alpha &= \frac{\vec{a}\cdot\vec{v}}{\|\vec{a}\|\|\vec{v}\|} = \frac{\|\vec{a}\|\vec{b}\cdot\vec{a} + \|\vec{b}\|\vec{a}\cdot\vec{a}}{\|\vec{a}\|\|\vec{v}\|} = \frac{\vec{b}\cdot\vec{a} + \|\vec{b}\|\|\vec{a}\|}{\|\vec{v}\|}\\ \cos\beta &= \frac{\vec{b}\cdot\vec{v}}{\|\vec{b}\|\|\vec{v}\|} = \frac{\|\vec{a}\|\vec{b}\cdot\vec{b} + \|\vec{b}\|\vec{b}\cdot\vec{a}}{\|\vec{b}\|\|\vec{v}\|} = \frac{\|\vec{b}\|\|\vec{a}\| + \vec{b}\cdot\vec{a}}{\|\vec{v}\|}.\end{aligned}$$

   En consecuencia, los ángulos entre $\vec{a}$ y $\vec{v}$ y $\vec{b}$ y $\vec{v}$ son iguales; por lo tanto, $\vec{v}$ está sobre la bisectriz del ángulo formado por los vectores $\vec{a}$ y $\vec{b}$.

5. El producto escalar de estos vectores es:

$$\begin{aligned}\vec{c}\cdot\vec{d} &= (\|\vec{a}\|\vec{b} + \|\vec{b}\|\vec{a})\cdot(\|\vec{a}\|\vec{b} - \|\vec{b}\|\vec{a})\\ &= \|\vec{a}\|^2\|\vec{b}\|^2 - \|\vec{a}\|^2\|\vec{b}\|^2 = 0;\end{aligned}$$

   por lo tanto, son ortogonales.

7. Dado que

$$
\begin{aligned}
\|a+b\|^2 &= \langle a+b, a+b\rangle = \|a\|^2 + 2\langle a,b\rangle + \|b\|^2 \\
\|a-b\|^2 &= \langle a-b, a-b\rangle = \|a\|^2 - 2\langle a,b\rangle + \|b\|^2,
\end{aligned}
$$

al restar estas expresiones se tiene que

$$\|a+b\|^2 - \|a-b\|^2 = 4\langle a,b\rangle,$$

que es la expresión que se pide.

9. Puesto que

$$
\begin{aligned}
\cos\alpha &= \frac{\overrightarrow{AB}.\overrightarrow{AC}}{\|\overrightarrow{AB}\|\|\overrightarrow{AC}\|} \Rightarrow \|\overrightarrow{AC}\|\cos\alpha = \frac{\overrightarrow{AB}.\overrightarrow{AC}}{\|\overrightarrow{AB}\|} \\
\cos\beta &= \frac{\overrightarrow{BA}.\overrightarrow{BC}}{\|\overrightarrow{BA}\|\|\overrightarrow{BC}\|} \Rightarrow \|\overrightarrow{BC}\|\cos\beta = \frac{\overrightarrow{BA}.\overrightarrow{BC}}{\|\overrightarrow{AB}\|}
\end{aligned}
$$

al sumar estas dos expresiones se tiene

$$
\begin{aligned}
\|\overrightarrow{AC}\|\cos\alpha + \|\overrightarrow{BC}\|\cos\beta &= \frac{\overrightarrow{AB}.\overrightarrow{AC}}{\|\overrightarrow{AB}\|} + \frac{\overrightarrow{BA}.\overrightarrow{BC}}{\|\overrightarrow{AB}\|} \\
&= \frac{\overrightarrow{AB}.\overrightarrow{AC} + \overrightarrow{BA}.\overrightarrow{BC}}{\|\overrightarrow{AB}\|} \\
&= \frac{\overrightarrow{AB}.(\overrightarrow{AC} + \overrightarrow{CB})}{\|\overrightarrow{AB}\|} \\
&= \frac{\overrightarrow{AB}.\overrightarrow{AB}}{\|\overrightarrow{AB}\|} = \|\overrightarrow{AB}\|.
\end{aligned}
$$

11. El ángulo entre f y g es 0. El ángulo entre f y h es $\pi/3$. El ángulo entre h y g es $\pi/6$.

13. La base ortonormal es:

$$u_1 = \frac{\sqrt{3}}{6}(1,1,1), \quad u_2 = \frac{\sqrt{6}}{6}(1,1,-2), \quad u_3 = \frac{\sqrt{2}}{2}(-1,1,0).$$

La base ortonormal es:

$$u_1 = \frac{\sqrt{6}}{6}(2,1,-1), \quad u_2 = \frac{\sqrt{2}}{2}(0,1,1), \quad u_3 = \frac{\sqrt{3}}{3}(1,-1,1).$$

15. El ángulo entre v_1 y v_2 es $35,26$. Entre v_1 y v_3 es $75,04$. Entre v_3 y v_2 es $50,79$.
El ángulo entre u_1 y u_2 es 30. Entre u_1 y u_3 es 180. Entre u_3 y u_2 es 150.

17. La proyección escalar es $14/\sqrt{30}$. La proyección vectorial es $7/15(-2,-5,1)$.

19. Cualquier vector de la forma $t(2,-3,7)$.

21. La ecuación del plano es $-2x + y - 4z = -35$.

23. El valor de a es $16/7$.

25. La distancia es $26/\sqrt{41}$.

A.5 Ejercicios propuestos capítulo 5, página 341

1. $\mathbf{a} \times \mathbf{b} = -6\mathbf{i} + 4\mathbf{j} - 7\mathbf{k}$.

3. La ecuación del plano es $x - 2y + 2z = 4$.

5. La ecuación del plano es $x + y + z = 3$.

7. La ecuación del plano es $18x + y + 8z = 26$.

9. Los planos pedidos son $x + 3y = 11$ y $z + 4y = 14$.

11.

$$\begin{aligned}
\|\mathbf{a}\times\mathbf{b}\|^2 &= (\mathbf{a}\times\mathbf{b})\cdot(\mathbf{a}\times\mathbf{b})\\
&= \mathbf{a}\cdot(\mathbf{b}\times(\mathbf{a}\times\mathbf{b}))\\
&= \mathbf{a}\cdot((\mathbf{b}\cdot\mathbf{b})\mathbf{a}-(\mathbf{b}\cdot\mathbf{a})\mathbf{b})\\
&= (\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b})-(\mathbf{b}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{a})\\
&= \|\mathbf{a}\|^2\|\mathbf{b}\|^2-(\mathbf{b}\cdot\mathbf{a})^2.
\end{aligned}$$

13. Área del triángulo: $\sqrt{6}/2$.

15. Área del paralelogramo: $\sqrt{314}$.

17. La ecuación del plano pedido es $2x+5y+z=17$.

19. La recta de intersección de los planos es

$$\frac{4(x-6)}{17}=\frac{4(y-2)}{11}=z.$$

21.

$$\mathbf{X}\times\mathbf{Y}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\x_1&x_2&x_3\\y_1&y_2&y_3\end{vmatrix}=\mathbf{i}\begin{vmatrix}x_2&x_3\\y_2&y_3\end{vmatrix}-\mathbf{j}\begin{vmatrix}x_1&x_3\\y_1&y_3\end{vmatrix}+\mathbf{k}\begin{vmatrix}x_1&x_2\\y_1&y_2\end{vmatrix};$$

luego,

$$(\mathbf{X}\times\mathbf{Y})\cdot\mathbf{Z}=z_1\begin{vmatrix}x_2&x_3\\y_2&y_3\end{vmatrix}-z_2\begin{vmatrix}x_1&x_3\\y_1&y_3\end{vmatrix}+z_3\begin{vmatrix}x_1&x_2\\y_1&y_2\end{vmatrix};$$

por lo tanto,

$$(\mathbf{X}\times\mathbf{Y})\cdot\mathbf{Z}=\begin{vmatrix}z_1&z_2&z_3\\x_1&x_2&x_3\\y_1&y_2&y_3\end{vmatrix}=\begin{vmatrix}x_1&x_2&x_3\\y_1&y_2&y_3\\z_1&z_2&z_3\end{vmatrix}.$$

23. La ecuación del eje central es

$$x=\frac{11s-37}{11};\quad y=-\frac{44s+421}{66};\quad z=s;\quad t=-\frac{43}{22}.$$

A.6 Ejercicios propuestos capítulo 6, página 416

1. Sí es transformación lineal.

3. No es transformación lineal.

5. Sí es transformación lineal.

7. Sí es transformación lineal.

8\. b) El núcleo es $N_t = gen\{(1,0,0)\}$ y la nulidad es 1. El rango de la transformación es $gen\{(0,1),(1,0)\}$.

9\. a) El núcleo es $N_t = \{(0,0)\}$ y la nulidad es 0. El rango de la transformación es $gen\{(1,1,2),(1,-1,3)\}$. Además,

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

9\. c) El núcleo es $N_t = \{(0,0,0)\}$ y la nulidad es 0. El rango de la transformación es $gen\{(0,2-3),(2,0,6),(1,-2,3)\}$. Además,

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & -2 \\ -3 & 6 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

9\. e) El núcleo es el conjunto de las matrices antisimétricas. El rango de la transformación es el conjunto de las matrices simétricas.

9\. g) El núcleo es $N_t = gen\{(-2,0,1,1,0),(0,1,0,0,0)\}$ y la nulidad es 2. El rango de la transformación es $gen\{(1,1,2,0),(-1,0,-1,-1,),(-1,-1,-1,0)\}$.

10\. b) $T_1 \circ T_3 = (x_1, 3x_1)$.

10. d) $3T_1 - 2T_2 = (x_1 + 2x_3, x_1 + x_2)$.

11. b)

$$T_3 \circ T_1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

11. d)

$$T_4 \circ T_4 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

A.7 Ejercicios propuestos capítulo 7, página 488

1. *a*) El polinomio característico es $p(\lambda) = \lambda^2 - 6\lambda - 1$.
   b) El polinomio característico es $p(\lambda) = \lambda^2 - 2\lambda - 3$.
   c) El polinomio característico es $p(\lambda) = \lambda^2 - 9\lambda + 18$.
   d) El polinomio característico es $p(\lambda) = \lambda^2 - 4\lambda + 3$.
   e) El polinomio característico es $p(\lambda) = -\lambda^3 + 6\lambda^2 - 7\lambda + 2$.
   f) El polinomio característico es $p(\lambda) = -\lambda^3 + 4\lambda^2 - 3\lambda + 2$.
   g) El polinomio característico es $p(\lambda) = -\lambda^3 + \lambda^2 + 9\lambda - 9$.

3. Los valores propios son $\lambda = -1$, $\lambda = -2$ y $\lambda = 2$; por lo tanto, es diagonalizable. Además:

$$v_{-1} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \; v_{-2} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}; \; v_2 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}; \quad C = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix}.$$

5. Los valores propios son $\lambda = -1$ y $\lambda = 1$; los vectores propios son:

$$v_{-1} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}; \; v_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \; v_1 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}; \quad C = \begin{bmatrix} -1 & 1 & -3 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

7. Si λ es un valor propio de la matriz A, entonces se cumple que $Av = \lambda v$ y, por lo tanto, al multiplicar en ambos lados por A se tiene que $A^2v = \lambda Av$. Si A es nilpotente, es decir, $A^2 = A$, entonces $Av = \lambda^2 v$; esto significa que λ^2 es también un valor propio de A. Multiplicando nuevamente por A se tiene $A^2v = Av = \lambda^3 v$; de esta manera, λ^3 es igualmente un valor propio de A. Continuando de esta forma se tiene que $\{\lambda, \lambda^2 \ldots \lambda^n \ldots\}$ son valores propios de A. Como el número de valores propios de A es a lo sumo n, se tiene que $\lambda^n = 0$ o $\lambda^n = 1$.

9. Los valores propios son: $\lambda = 1$, $\lambda = \sqrt{7}$ y $\lambda = -\sqrt{7}$, y los vectores propios son:

$$v_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad v_{\sqrt{7}} = \begin{bmatrix} \sqrt{21} + 2\sqrt{3} \\ 0 \\ 3 \end{bmatrix}; \quad v_{-\sqrt{7}} = \begin{bmatrix} 2\sqrt{3} - \sqrt{21} \\ 0 \\ 3 \end{bmatrix}.$$

11. Como

$$\begin{aligned} \det(A^t - \lambda I) &= \det(A^t - \lambda I^t) \\ &= \det(A - \lambda I)^t \\ &= \det(A - \lambda I) \end{aligned}$$

y, en consecuencia, A y A^t tienen el mismo polinomio característico y, por lo tanto, los mismos valores propios.

13. Si el valor propio es $\lambda = 0$, entonces $\det(A^t - \lambda I) = \det A = 0$, y de esta forma A es singular. Si ahora afirmamos que A es singular, $\det(A^t - 0I) = \det A = 0$ y, por lo tanto, 0 es una valor propio de A.

15. Como A es ortogonal, $AA^t = I$. Sea ahora $cA(cA)^t = c^2AAt = c^2I$; para ser ortogonal se requiere que $c^2 = 1$ o simplemente $c = \pm 1$.

16. b) Los valores propios de la matriz asociada a la forma cuadrática son $\lambda = 9$ y $\lambda = -4$; por lo tanto, es una hipérbola.

16. d) Los valores propios de la matriz asociada a la forma cuadrática son $\lambda = 10$ y $\lambda = 5$; por lo tanto, es una elipse.

17. a) Los valores propios de la matriz asociada a la forma cuadrática son $\lambda = 2 - 2\sqrt{2}$, $\lambda = 2 + 2\sqrt{2}$ y $\lambda = 4$; por lo tanto, es un hiperboloide de una hoja.

17. c) Los valores propios de la matriz asociada a la forma cuadrática son $\lambda = -3$, $\lambda = -3$ y $\lambda = 3$; por lo tanto, es un hiperboloide de dos hojas.

17. e) Los valores propios de la matriz asociada a la forma cuadrática son $\lambda = 1 - 2\sqrt{3}$, $\lambda = 1 + 2\sqrt{3}$, y $\lambda = -1$; por lo tanto, es un hiperboloide de dos hojas.

17. g) Los valores propios de la matriz asociada a la forma cuadrática son $\lambda = \frac{3-\sqrt{5}}{2}$, $\lambda = \frac{3+\sqrt{5}}{2}$ y $\lambda = 1$; por lo tanto, es un hiperboloide de una hoja.

Bibliografía

Álvarez, R., F. Marcellán y J. Arvesú, *Álgebra lineal y aplicaciones,* Madrid, Síntesis, 1999.

Bell, E. T, *Historia de las matemáticas,* México, Fondo de Cultura Económica, 1985.

Boyer, Carl B., *Historia de la matemática,* Madrid, Alianza, 2001.

Burgos, Juan de, *Álgebra lineal y geometría cartesiana,* México, Mc Graw Hill, 2006.

Gerber, Harvey, *Álgebra lineal,* México, Grupo Editorial Iberoamérica, 1992.

Grossman, Stanley, *Álgebra lineal,* 6.ª ed., México, Mc Graw Hill, 2008.

Kolman, Bernard y David Hill, *Álgebra lineal,* 8.ª ed., México, Prentice Hall, 2006.

Lay, David C., *Álgebra lineal y sus aplicaciones,* México, Pearson, 2010.

Lipschutz, S., *Álgebra Lineal,* 2.ª ed., Madrid, McGraw Hill Interamericana, 1991.

Poole, David, *Álgebra lineal. Una introducción moderna,* México, Cengage Learning, 2007.

Proskuriakov, I. V., *2000 Problemas de álgebra lineal,* Barcelona, Reverté, 2008.

Soler, Francisco, *Álgebra lineal y programación lineal,* 6.^{a} ed., Bogotá, Ecoe Ediciones.

Strang, Gilbert, *Álgebra lineal y sus aplicaciones,* México, Cengage Learning, 2007.

Takahashi, Alonso, *Álgebra lineal,* Bogotá, Ed. Universidad Nacional de Colombia, 2002.

Índice alfabético