UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE INGENIERIA

INGENIERIA
A N E X O
BIBLIOTECA

INGENIERIA
A N E X O
BIBLIOTECA

APUNTES DE
ALGEBRA LINEAL

INGENIERIA
A N E X O
BIBLIOTECA

INGENIERIA
A N E X O
BIBLIOTECA

EDUARDO SOLAR GONZALE
LEDA SPEZIALE DE GUZMA

INGENIERIA
A N E X O
BIBLIOTECA

BIBLIOTECA DE LA UNIVERSIDAD NACIONAL
FACULTAD DE INGENIERIA

2 A
ALG. LINEAL
G.-903246
903246

2-A FACULTAD DE INGENIERIA UNAM.
903246
G1.- 903246

CAPITULO V SISTEMAS DE ECUACIONES LINEALES

INGENIERIA
A N E X O
BIBLIOTECA

INTRODUCCION

El estudio de los sistemas de ecuaciones lineales puede emprenderse desde diversos puntos de vista; el que adoptamos en este capítulo es, posiblemente, el más concreto.

Su propósito fundamental es el de establecer un método para obtener soluciones y, en consecuencia, se presentan únicamente los conceptos necesarios para desarrollar y aplicar el método.

Posteriormente, en el capítulo de Espacios Vectoriales se emplearán las herramientas que proporciona el Algebra Lineal para estudiar los sistemas de ecuaciones lineales desde un punto de vista más general.

G. 903246

V.1 ECUACIONES LINEALES

Supongamos que en una fábrica se producen tres tipos de artículos a los que llamamos A, B y C y que en ella trabajan cincuenta obreros durante ocho horas diarias; es decir, que se dispone de cua -

trocientas "horas-hombre" al día.

Para producir un artículo del tipo A se requieren 20 horas hombre, para uno del tipo B se requieren 100 y para uno del tipo C se requieren 40.

Si en condiciones normales no existen restricciones de materia prima ni de maquinaria y los obreros están capacitados para trabajar en la elaboración de cualquiera de los tres tipos de artículos ¿Cuántos artículos A, B y C pueden producirse diariamente empleando todas las horas-hombre disponibles?

Para responder a esta pregunta podemos plantear el siguiente modelo matemático del problema:

Si x_1, x_2 y x_3 representan el número de productos A, B y C, respectivamente, que se producen por día, entonces $20x_1$, $100x_2$ y $40x_3$ representarán el número de horas-hombre que se requieren para producirlos. Por tanto, si se desea emplear las 400 horas-hombre disponibles, x_1, x_2 y x_3 deben ser tales que

$$20x_1 + 100x_2 + 40x_3 = 400 \qquad --- (1)$$

Expresiones como ésta reciben el nombre de ecuaciones lineales.

Así, una respuesta a la pregunta sobre el número de artículos a producirse diariamente podría ser la siguiente.

"Producir 4 artículos del tipo A, 2 del tipo B y 3 del tipo C", ya que al sustituir los valores

$$x_1 = 4, \; x_2 = 2 \quad y \quad x_3 = 3$$

en la ecuación (1) se verifica la igualdad; esto es

$$20(4) + 100(2) + 40(3) = 80 + 200 + 120 = 400$$

Se dice entonces que el conjunto de valores $x_1 = 4$, $x_2 = 2$ y $x_3 = 3$ es una solución de la ecuación (1), o que la terna ordenada (4, 2, 3) es una solución de dicha ecuación.

Daremos a continuación una definición formal para estos conceptos

V.1.1 DEFINICION

Una ecuación lineal sobre C es una expresión de la forma

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

donde $a_1, a_2, \ldots, a_n, b \in C$

INGENIERIA
A N E X O
BIBLIOTECA

A los símbolos $x_1, x_2, \ldots, x_n$ se les conoce como "incógnitas" de la ecuación, a los números a_i como "coeficientes" de las x_i y a b como el "término independiente".

V.1.2 DEFINICION

Una solución de la ecuación lineal

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

es un conjunto ordenado de n valores $k_1, k_2, \ldots, k_n$ tales que

$$a_1k_1 + a_2k_2 + \ldots + a_nk_n = b$$

- Resolución de una ecuación lineal

En la búsqueda de soluciones para la ecuación

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

pueden distinguirse tres casos.

Caso i) Al menos uno de los coeficientes es diferente de cero.

Si $a_k \neq 0$ la ecuación puede escribirse como

$$a_kx_k = b - a_1x_1 - \dots - a_{k-1}x_{k-1} - a_{k+1}x_{k+1} - \dots - a_nx_n$$

o bien, como

$$x_k = \frac{1}{a_k}\,(b - a_1x_1 - \dots - a_{k-1}x_{k-1} - a_{k+1}x_{k+1} - \dots - a_nx_n)$$

Podemos entonces asignar valores a las incógnitas $x_1, \dots, x_{k-1}$ $x_{k+1}, \dots, x_n$ (arbitrariamente), y de la expresión anterior se obtendrá el valor de x_k que con los valores asignados constituye una solución de la ecuación.

Por ejemplo, la ecuación

$$20x_1 + 100x_2 + 40x_3 = 400 \qquad --- (1)$$

puede escribirse como

$$20x_1 = 400 - 100x_2 - 40x_3$$
$$x_1 = 20 - 5x_2 - 2x_3$$

de donde, haciendo $x_2 = 2$ y $x_3 = 3$ se obtiene

$$x_1 = 20 - 5(2) - 2(3) = 20 - 10 - 6 = 4$$

con lo que se forma la solución

$$x_1 = 4,\ x_2 = 2 \quad y \quad x_3 = 3$$

la cual presentamos al inicio de esta sección.

Si queremos obtener otra solución podemos asignar otros valores a las incógnitas x_2 y x_3; por ejemplo $x_2 = 0$ y $x_3 = 5$, con lo que se obtiene

$$x_1 = 20 - 5(0) - 2(5) = 10$$

En consecuencia, la terna (10, 0, 5) es otra solución de la ecuación (1).

En general, cualquier terna ordenada de la forma

$$(\ 20-5a-2b\ ,\ a\ ,\ b\ )$$

donde a y b son dos números cualesquiera, es una solución de la ecuación (1).

Caso ii) Todos los coeficientes son nulos y el término independiente también lo es.

INGENIERIA
A N E X O
BIBLIOTECA

Entonces la ecuación es de la forma

$$0x_1 + 0x_2 + \dots + 0x_n = 0$$

y es claro que cualquier conjunto de n valores es una solución de la ecuación.

Caso iii) Todos los coeficientes son nulos y el término independiente no lo es.

Entonces la ecuación es de la forma

$$0x_1 + 0x_2 + \dots + 0x_n = b, \text{ con } b \neq 0$$

y es claro que ningún conjunto de n valores podrá ser una solución de la ecuación; es decir, la ecuación no tiene solución.

V.2 SISTEMAS DE ECUACIONES LINEALES

Volvamos al ejemplo de la fábrica y supongamos que los artículos B y C deben producirse en cantidades iguales. Tenemos entonces la restricción adicional

$$x_2 = x_3$$

que, expresada en la forma que establece la definición V.1.1, queda como

$$0x_1 + x_2 - x_3 = 0 \qquad --- (2)$$

Ahora el problema consiste en encontrar una solución que satisfaga "simultáneamente" a las ecuaciones (1) y (2). En consecuencia, las dos soluciones obtenidas anteriormente ya no son útiles, puesto que (4, 2, 3) y (10, 0, 5) no son soluciones de la ecuación (2); esto es

$$0(4) + 1(2) - 1(3) = 0 + 2 - 3 = -1 \neq 0$$

y

$$0(10) + 1(0) - 1(5) = 0 + 0 - 5 = -5 \neq 0$$

A diferencia de éstas, si se producen 6 artículos del tipo A, 2 del tipo B y 2 del tipo C se tiene una solución que satisface ambas restricciones ya que

$$20(6) + 100(2) + 40(2) = 120 + 200 + 80 = 400$$

y

$$0(6) + 1(2) - 1(2) = 0 + 2 - 2 = 0$$

Se dice entonces que la terna ordenada (6, 2, 2) es una solución del sistema

$$\begin{aligned} 20x_1 + 100x_2 + 40x_3 &= 400 \\ 0x_1 + x_2 - x_3 &= 0 \end{aligned}$$

el cual consta de dos ecuaciones lineales con tres incógnitas.

En general, un sistema es un conjunto de ecuaciones lineales que tienen las mismas incógnitas, como lo establece la siguiente definición.

V.2.1 DEFINICION

Un sistema de m ecuaciones lineales con n incógnitas sobre C es una expresión de la forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\;\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

donde $a_{11}, a_{12}, \ldots, a_{mn}, b_1, \ldots, b_m \in C$

Paso

Puesto que un sistema de ecuaciones lineales es un conjunto de ecuaciones con las mismas incógnitas, resulta natural considerar como una solución del sistema a un conjunto de valores que satisface a todas las ecuaciones del sistema, por lo que se establece la siguiente definición.

INGENIERIA
ANEXO
BIBLIOTECA

V.2.2 DEFINICION

Una solución del sistema de ecuaciones lineales

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\dots \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

es un conjunto ordenado de n valores $k_1, k_2, \dots, k_n$ tales que

$$\begin{aligned} a_{11}k_1 + a_{12}k_2 + \dots + a_{1n}k_n &= b_1 \\ a_{21}k_1 + a_{22}k_2 + \dots + a_{2n}k_n &= b_2 \\ &\dots \\ &\dots \\ a_{m1}k_1 + a_{m2}k_2 + \dots + a_{mn}k_n &= b_m \end{aligned}$$

La definición anterior establece claramente lo que deberá entenderse por solución de un sistema de ecuaciones lineales; sin embargo, no nos dice que cualquier sistema de ecuaciones lineales habrá de tener solución.

Hay sistemas de ecuaciones que no admiten solución. Por ejemplo, es claro que el sistema

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 3 \end{aligned}$$

no tiene solución, puesto que no existen dos números cuya suma

ANEXO BIBLIOTECA

sea igual a 1 y también a 3. A este tipo de sistemas les llamaremos "incompatibles".(1)

Si, por el contrario, un sistema de ecuaciones lineales tiene solución diremos que es "compatible".(2)

Los sistemas compatibles pueden tener una sola solución, en cuyo caso diremos que son "determinados"; o más de una solución, en cuyo caso diremos que son "indeterminados".

De acuerdo con esto, los sistemas de ecuaciones lineales pueden clasificarse de la siguiente manera

SISTEMAS DE ECUACIONES LINEALES
- INCOMPATIBLES (no tienen solución)
- COMPATIBLES (tienen solución)
  - DETERMINADOS (una sola solución)
  - INDETERMINADOS (más de una solución)

- Transformaciones elementales

Cuando dos sistemas de ecuaciones lineales tienen las mismas soluciones se dice que son "equivalentes".

El método que emplearemos en este capítulo para obtener las soluciones de un sistema de ecuaciones lineales se basa en el empleo de ciertas transformaciones, llamadas transformaciones elementales, que no alteran las soluciones del sistema; es decir, transformaciones que al aplicarse a un sistema dan como resultado un sistema equivalente.

(1) *En algunos textos se emplea el término "inconsistente" para referirse a este concepto.*

(2) *"consistente".*

Las transformaciones elementales pueden ser de tres tipos y con - sisten en:

I) Intercambiar dos ecuaciones.

II) Multiplicar una ecuación por un número diferente de cero.

III) Multiplicar una ecuación por un número y sumarla a otra ecuación, reemplazando esta última por el resultado obtenido.

Para ilustrar el empleo de estas transformaciones consideremos, por ejemplo, el sistema

$$\begin{aligned} 3x - 2y + z &= -1 \\ x - 2y + 3z &= 1 \\ 6y - 2z &= 4 \end{aligned} \qquad (S_0)$$

Si intercambiamos en él las dos primeras ecuaciones estamos aplicando a S_0 una transformación del tipo I que conduce al sistema

$$\begin{aligned} x - 2y + 3z &= 1 \\ 3x - 2y + z &= -1 \\ 6y - 2z &= 4 \end{aligned} \qquad (S_1)$$

que, evidentemente, tiene las mismas soluciones que S_0.

Si ahora multiplicamos la tercera ecuación de S_1 por $\frac{1}{2}$ estamos aplicando a S_1 una transformación del tipo II que conduce al sistema

$$\begin{aligned} x - 2y + 3z &= 1 \\ 3x - 2y + z &= -1 \\ 3y - z &= 2 \end{aligned} \qquad (S_2)$$

Si ahora multiplicamos la primera ecuación de S_2 por -3 y la sumamos a la segunda ecuación, reemplazando esta última por el resultado obtenido, estamos aplicando una transformación del tipo

III que conduce al sistema

$$\begin{aligned} x - 2y + 3z &= 1 \\ 4y - 8z &= -4 \\ 3y - z &= 2 \end{aligned} \qquad (S_3)$$

Los sistemas S_0, S_1, S_2 y S_3 son, según hemos dicho, equivalentes; esto es, tienen las mismas soluciones.

Es obvio que las transformaciones del tipo I y del tipo II conducen a sistemas equivalentes. El caso de las transformaciones del tipo III no es tan evidente por lo que se demostrará a continuación.

Sea el sistema de m ecuaciones lineales con n incógnitas

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ &\vdots \\ &a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n = b_p \\ &\vdots \\ &a_{q1}x_1 + a_{q2}x_2 + \dots + a_{qn}x_n = b_q \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{aligned} \qquad (S)$$

donde $1 \leq p < q \leq m$; y sea S' el sistema que se obtiene al multiplicar por c la ecuación p y sumarla a la ecuación q; esto es

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \dots\dots\dots\dots + a_{1n}x_n = b_1 \\ &\vdots \\ &a_{p1}x_1 + a_{p2}x_2 + \dots\dots\dots\dots + a_{pn}x_n = b_p \\ &\vdots \\ &(ca_{p1} + a_{q1})x_1 + \dots + (ca_{pn} + a_{qn})x_n = cb_p + b_q \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + \dots\dots\dots\dots + a_{mn}x_n = b_m \end{aligned} \qquad (S')$$

Si $(k_1, k_2, \dots, k_n)$ es una solución de S entonces satisface to das las ecuaciones de S' con excepción, posiblemente, de la ecuación

$$(ca_{p1} + a_{q1})x_1 + \dots + (ca_{pn} + a_{qn})x_n = cb_p + b_q \qquad \text{---}(q')$$

Sin embargo, como $(k_1, k_2, \dots, k_n)$ es solución de S se tiene que

$$a_{p1}k_1 + a_{p2}k_2 + \dots + a_{pn}k_n = b_p \qquad \text{---}(1)$$

y

$$a_{q1}k_1 + a_{q2}k_2 + \dots + a_{qn}k_n = b_q \qquad \text{---}(2)$$

por lo que

$$ca_{p1}k_1 + ca_{p2}k_2 + \dots + ca_{pn}k_n = cb_p \qquad \text{---}(3)$$

En consecuencia, sumando (3) y (2)

$$(ca_{p1} + a_{q1})k_1 + \dots + (ca_{pn} + a_{qn})k_n = cb_p + b_q$$

por lo que $(k_1, k_2, \dots, k_n)$ satisface también la ecuación (q').

De manera recíproca, sea ahora $(\ell_1, \ell_2, \dots, \ell_n)$ una solución de S', entonces satisface todas las ecuaciones de S con excepción, posiblemente, de la ecuación

$$a_{q1}x_1 + a_{q2}x_2 + \dots + a_{qn}x_n = b_q \qquad \text{---}(q)$$

Sin embargo, como $(\ell_1, \ell_2, \dots, \ell_n)$ es solución de S' satisface la ecuación (q'); esto es

$$(ca_{p1} + a_{q1})\ell_1 + \dots + (ca_{pn} + a_{qn})\ell_n = cb_p + b_q \qquad \text{---}(1)$$

además, como también satisface la ecuación p de (S')

$$a_{p1}\ell_1 + a_{p2}\ell_2 + \dots + a_{pn}\ell_n = b_p$$

INGENIERIA ANEXO BIBLIOTECA

se tiene que

$$ca_{p1}\ell_1 + ca_{p2}\ell_2 + \dots + ca_{pn}\ell_n = cb_p \qquad \text{---}(2)$$

Entonces, restando (2) de (1)

$$a_{q1}\ell_1 + a_{q2}\ell_2 + \dots + a_{qn}\ell_n = b_q$$

con lo que $(\ell_1, \ell_2, \dots, \ell_n)$ satisface también la ecuación (q) y los sistemas S y S' son equivalentes. □

- El método de Gauss

El procedimiento más cómodo para obtener las soluciones de un sistema de ecuaciones lineales es, tal vez, el conocido como método de Gauss.

Este método consiste en la eliminación consecutiva de las incógnitas con el propósito de llegar a un sistema que tenga forma "escalonada". Para llevar a cabo dicha eliminación sin alterar las soluciones del sistema, se recurre a las transformaciones elementales que hemos descrito.

Para ilustrar la idea central del método, consideremos el problema de resolver el siguiente sistema de ecuaciones lineales

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ 3x_1 + 4x_2 + x_3 &= -10 \qquad (S_0) \\ -2x_1 - 4x_2 - x_3 &= 0 \end{aligned}$$

Para eliminar la incógnita x_1 de la segunda y de la tercera ecuación, podemos emplear dos tranformaciones del tipo III. Así, multiplicando la primera ecuación por -3 y sumando el resultado a la segunda se obtiene el sistema

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ x_2 - 5x_3 &= -10 \qquad (S_1) \\ -2x_1 - 4x_2 - x_3 &= 0 \end{aligned}$$

y multiplicando ahora la primera ecuación por 2 y sumando el re - sultado a la tercera se obtiene

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ x_2 - 5x_3 &= -10 \qquad (S_2) \\ -2x_2 + 3x_3 &= 6 \end{aligned}$$

con lo que hemos conseguido eliminar x_1 de la segunda y tercera ecuaciones.

Para eliminar x_2 de la tercera ecuación podemos emplear nuevamente una transformación del tipo III, pero tomando ahora la segunda ecuación como "pivote". Así, multiplicando la segunda ecuación por 2 y sumando el resultado a la tercera se obtiene el sistema

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ x_2 - 5x_3 &= -10 \qquad (S_3) \\ -7x_3 &= -14 \end{aligned}$$

donde se observa de inmediato que

$$x_3 = \frac{-14}{-7} = 2$$

Para obtener el valor de x_2 sustituimos el valor obtenido de x_3 en la segunda ecuación de S_3

$$x_2 - 5(2) = -10$$

quedando así una sola incógnita cuyo valor es

$$x_2 = -10 + 10 = 0$$

Por último, para obtener el valor de x_1 sustituimos en la primera ecuación de S_3 los valores obtenidos de x_2 y x_3.

$$x_1 + 1(0) + 2(2) = 3$$

de donde

$$x_1 = 3 - 4 = -1$$

En consecuencia, $x_1 = -1$, $x_2 = 0$ y $x_3 = 2$ es la solución del sistema S_3; y como éste es equivalente a S_0, la terna (-1, 0, 2) es la solución del sistema inicial, con lo que queda resuelto el problema.

Cabe hacer notar que en el párrafo anterior hemos dicho "la" solución del sistema S_3, lo cual lleva implícito que dicho sistema es determinado. Explicaremos ahora el por qué de tal aseveración.

Es evidente que el valor $x_3 = 2$ es el único que satisface la tercera ecuación de S_3; en consecuencia, los únicos valores que sa - tisfacen "simultáneamente" a la segunda y a la tercera ecuación de S_3 son $x_2 = 0$ y $x_3 = 2$. Continuando con este razonamiento concluimos que $x_1 = -1$, $x_2 = 0$ y $x_3 = 2$ es la única solución del sistema S_3.

Como el lector habrá sospechado, éste no es el único caso que pue de presentarse ya que, como hemos visto, existen sistemas que son indeterminados y otros que son incompatibles. Veremos posteriormente algunos ejemplos correspondientes a estos dos casos haciendo notar bajo qué condiciones se presentan; sin embargo, introduciremos primero una herramienta que nos permitirá ahorrarnos al -

gún trabajo y ver con mayor claridad lo que sucede en cada paso cuando utilizamos el método de Gauss.

Si analizamos con cierto cuidado el proceso seguido en el ejemplo anterior, podemos darnos cuenta que no era necesario escribir los símbolos correspondientes a las incógnitas una y otra vez, puesto que todas las operaciones se efectuaron sobre los coeficientes y términos independientes.

El sistema

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ 3x_1 + 4x_2 + x_3 &= -1 \qquad (S_0) \\ -2x_1 - 4x_2 - x_3 &= 0 \end{aligned}$$

queda completamente definido por el valor de sus coeficientes y términos independientes, los cuales pueden presentarse convenientemente en el siguiente arreglo tabular

$$M_0 = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 3 & 4 & 1 & -1 \\ -2 & -4 & -1 & 0 \end{bmatrix}$$

al que se conoce con el nombre de "matriz". Esta matriz, en particular, contiene doce elementos dispuestos en tres renglones y cuatro columnas por lo que se dice que es de orden 3×4.

De la misma manera, los sistemas S_1, S_2 y S_3 pueden ser representados, respectivamente, por las matrices

$$M_1 = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -5 & -10 \\ -2 & -4 & -1 & 0 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -5 & -10 \\ 0 & -2 & 3 & 6 \end{bmatrix}$$

INGENIERIA ANEXO BIBLIOTECA

$$M_3 = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & -7 & -14 \end{bmatrix}$$

las cuales pueden obtenerse a partir de M_0 efectuando, con los renglones, transformaciones análogas a las descritas con las ecuaciones. Estas transformaciones, conocidas como "transformaciones elementales por renglón", consisten en:

I) Intercambiar dos renglones.

II) Multiplicar un renglón por un número diferente de cero.

III) Multiplicar un renglón por un número y sumarlo a otro renglón, reemplazando este último por el resultado obtenido.

La última de las matrices anteriores (M_3) se dice que está en "forma escalonada" o que es una matriz escalonada. En general, se dice que una matriz está en forma escalonada si el número de ceros anteriores al primer elemento no nulo de cada renglón aumenta al pasar de un renglón al siguiente, hasta llegar eventualmente a renglones cuyos elementos son todos nulos.

Por ejemplo, las siguientes matrices también son escalonadas

$$\begin{bmatrix} 1 & -4 & 0 & 2 \\ 0 & 3 & -1 & 5 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 5 & -1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 3 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Regresando al método de Gauss, vemos que es conveniente representar al sistema mediante una matriz y efectuar en ella las transformaciones necesarias para llevarla a la forma escalonada. Haremos esto para obtener las soluciones del siguiente sistema de ecuaciones lineales.

$$\begin{aligned} 3x_1 + 3x_2 - x_3 + x_4 + 4x_5 &= 4 \\ x_1 + x_2 + x_3 - 2x_4 - x_5 &= 1 \\ -2x_1 - 2x_2 + 2x_3 - 3x_4 - 5x_5 &= -3 \end{aligned} \qquad (S_0)$$

Primero representamos al sistema por medio de la matriz

$$M_0 = \begin{bmatrix} 3 & 3 & -1 & 1 & 4 & 4 \\ 1 & 1 & 1 & -2 & -1 & 1 \\ -2 & -2 & 2 & -3 & -5 & -3 \end{bmatrix}$$

La cual trataremos de llevar hasta la forma escalonada mediante transformaciones elementales por renglón.

Por lo general, conviene que el primer elemento no nulo de cada renglón sea un uno (o un menos uno) para eliminar fácilmente los coeficientes que se encuentran por debajo de él, multiplicando simplemente por los simétricos respectivos. Entonces, intercambiando el primero y segundo renglones de M_0 obtenemos la matriz

$$M_1 = \begin{bmatrix} 1 & 1 & 1 & -2 & -1 & 1 \\ 3 & 3 & -1 & 1 & 4 & 4 \\ -2 & -2 & 2 & -3 & -5 & -3 \end{bmatrix}$$

la cual tiene un uno en la primera posición del primer renglón. Ahora, multiplicando dicho primer renglón por -3 y sumando al segundo y, a continuación, multiplicando el mismo primer renglón por 2 y sumando al tercero obtenemos la matriz

INGENIERIA
A N E X O
BIBLIOTECA

$$M_2 = \begin{bmatrix} 1 & 1 & 1 & -2 & -1 & 1 \\ 0 & 0 & -4 & 7 & 7 & 1 \\ 0 & 0 & 4 & -7 & -7 & -1 \end{bmatrix}$$

la cual puede transformarse en una matriz escalonada sumando el segundo renglón al tercero, con lo que se obtiene

$$M_3 = \begin{bmatrix} 1 & 1 & 1 & -2 & -1 & 1 \\ 0 & 0 & -4 & 7 & 7 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

El tercer renglón de esta matriz representa a una ecuación de la forma

$$0x_1 + 0x_2 + \ldots + 0x_n = 0$$

que, como vimos, es satisfecha por cualquier conjunto de n valores. En consecuencia, la matriz M_3 representa al siguiente sistema de dos ecuaciones

$$\begin{aligned} x_1 + x_2 + x_3 - 2x_4 - x_5 &= 1 \\ -4x_3 + 7x_4 + 7x_5 &= 1 \end{aligned} \qquad (S_1)$$

continuando con la idea del ejemplo anterior, de la segunda ecuación de S_1 podemos obtener el valor de x_3; sólo que ahora este valor no es único, sino que está en función de los valores que tomen x_4 y x_5. Así

$$-4x_3 = 1 - 7x_4 - 7x_5$$

por lo que

$$x_3 = -\frac{1}{4} + \frac{7}{4}x_4 + \frac{7}{4}x_5 \qquad \text{--- (1)}$$

Llevando este valor a la primera ecuación de S_1 se obtiene

$$x_1 + x_2 + (-\frac{1}{4} + \frac{7}{4}x_4 + \frac{7}{4}x_5) - 2x_4 - x_5 = 1$$

$$x_1 + x_2 - \frac{1}{4}x_4 + \frac{3}{4}x_5 = \frac{5}{4}$$

$$x_1 = \frac{5}{4} - x_2 + \frac{1}{4}x_4 - \frac{3}{4}x_5 \qquad \text{--- (2)}$$

podemos entonces dar cualquier valor a las incógnitas x_2, x_4 y x_5 y calcular, a partir de (1) y (2), los valores correspondientes de x_1 y x_3. Se dice por ello que el conjunto de expresiones

$$\begin{aligned} x_1 &= \frac{5}{4} - x_2 + \frac{1}{4}x_4 - \frac{3}{4}x_5 \\ x_2 &= x_2 \\ x_3 &= -\frac{1}{4} + \frac{7}{4}x_4 + \frac{7}{4}x_5 \qquad (3) \\ x_4 &= x_4 \\ x_5 &= x_5 \end{aligned}$$

constituye la "solución general" del sistema S_0 que, como se ve, es indeterminado.

Si queremos obtener una "solución particular" del sistema S_0; es decir, una solución en el sentido de la definición V.2.2, bastará con elegir un conjunto de tres valores para x_2, x_4 y x_5; por ejemplo

$$x_2 = 3$$
$$x_4 = 4$$
$$x_5 = -1$$

y calcular, a partir de la solución general, los correspondientes valores de x_1 y x_3. Para los valores elegidos se tiene

$$x_1 = \frac{5}{4} - 3 + \frac{1}{4}(4) - \frac{3}{4}(-1) = 0$$

y

$$x_3 = -\frac{1}{4} + \frac{7}{4}(4) + \frac{7}{4}(-1) = 5$$

por lo que (0, 3, 5, 4, -1) es una solución de S_0.

Si hacemos ahora $x_2 = 1$, $x_4 = 0$ y $x_5 = 1$, de (3) se obtiene

$$x_1 = \frac{5}{4} - 1 - \frac{3}{4} = -\frac{1}{2}$$

y

$$x_3 = -\frac{1}{4} + \frac{7}{4} = \frac{3}{2}$$

por lo que $(-\frac{1}{2}, 1, \frac{3}{2}, 0, 1)$ es otra solución del sistema S_0.

En ocasiones los símbolos correspondientes a las "variables libres" suelen reemplazarse por otras literales, las cuales se convierten en parámetros de la solución general. Así por ejemplo, para el caso anterior podemos expresar la solución general (3) como

$$\begin{aligned} x_1 &= \frac{5}{4} - a + \frac{1}{4}b - \frac{3}{4}c \\ x_2 &= a \\ x_3 &= -\frac{1}{4} + \frac{7}{4}b + \frac{7}{4}c \\ x_4 &= b \\ x_5 &= c \end{aligned}$$

donde a, b y c pueden tomar cualquier valor.

Consideremos ahora el sistema

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

al cual podemos representar con la matriz

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

Efectuando en ella transformaciones elementales por renglón la llevamos hasta la forma escalonada siguiente

$$\begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 0 & 3 & -2 \\ -1 & 2 & -4 & 4 \\ 3 & 2 & 2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -4 & 5 & -4 \\ 0 & 4 & -5 & 5 \\ 0 & -4 & 5 & -4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -4 & 5 & -4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

En la última matriz, el tercer renglón representa a una ecuación de la forma

$$0x_1 + 0x_2 + \ldots + 0x_n = b, \text{ con } b \neq 0$$

que, como vimos, no tiene solución. En consecuencia, el sistema en cuestión es incompatible.

A través de los ejemplos anteriores hemos mostrado lo que sucede al emplear el método de Gauss en cada uno de los tres casos correspondientes a la clasificación de los sistemas de ecuaciones lineales.

En resumen podemos decir lo siguiente:

El método de Gauss consiste en aplicar a un sistema de m ecuaciones con n incógnitas (o a la matriz que lo representa) una sucesión de transformaciones elementales hasta llevarlo a la forma escalonada.

Si durante el proceso se obtiene una ecuación de la forma

$$0x_1 + 0x_2 + \ldots + 0x_n = 0$$

a la que se llama ecuación nula, ésta se desecha puesto que cualquier conjunto de n valores es una solución de la misma.

Si durante el proceso se obtiene una ecuación de la forma

$$0x_1 + 0x_2 + \ldots + 0x_n = b; \text{ con } b \neq 0$$

el sistema es incompatible, puesto que dicha ecuación no tiene solución; de otra manera el sistema es compatible.

Si el sistema es compatible y al reducirlo a la forma escalonada se obtienen n ecuaciones no nulas, entonces el sistema es determinado y su solución se obtiene por sustitución sucesiva de los valores de las incógnitas, a partir de la última cuyo valor es inmediato.

Si el sistema es compatible y al reducirlo a la forma escalonada se obtienen $r < n$ ecuaciones no nulas, entonces el sistema es indeterminado y su solución general se obtiene dejando $n - r$ incógnitas libres (es decir como parámetros) y expresando a las otras r incógnitas en función de éstas.

V.2.3 EJERCICIOS

1.- Para la ecuación lineal

$$x_1 - 4x_2 + 2x_3 + \frac{1}{2}x_4 = 5$$

determinar cuáles de los siguientes conjuntos ordenados son soluciones

a) $(-3, -1, 2)$ b) $(1, -4, \frac{1}{2}, 2)$ c) $(-3, -1, 2, 0)$

d) $(\frac{3}{2}, -1, -1, 3)$ e) $(3, 2, -1, -1, 3)$

2.- Para cada una de las siguientes ecuaciones lineales obtener todas sus soluciones

a) $3x_1 - 2x_2 + x_3 = 5$ b) $3x_1 - 2x_2 + x_3 = 0$

c) $0x_1 + 0x_2 + 0x_3 = 5$ d) $0x_1 + 0x_2 + 0x_3 = 0$

e) $ax = b$; con $a \neq 0$

3.- Resolver los siguientes sistemas de ecuaciones lineales

a) $$\begin{aligned} x + 2y - 3z &= -5 \\ -y + 2z &= 5 \\ -2x + y &= -11 \\ 3x + z &= 13 \end{aligned}$$

b) $$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 &= -1 \\ -x_1 - 2x_2 + 2x_3 - 5x_4 &= 1 \\ 2x_1 + 4x_2 - 5x_3 + 7x_4 &= -2 \end{aligned}$$

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

d) $$\begin{aligned} x_2 - x_3 &= 3 \\ 3x_1 + x_2 + 2x_3 + x_4 &= 6 \\ x_1 + x_3 &= 0 \\ 2x_1 + x_2 + x_3 + x_4 &= 6 \end{aligned}$$

4.- Para cada uno de los siguientes sistemas de ecuaciones lineales

a) $$\begin{aligned} 2x - y - kz &= 0 \\ x - y - 2z &= 1 \\ -x + 2y + 0z &= k \end{aligned}$$

b) $$\begin{aligned} kx + y + z &= 1 \\ x + ky + z &= 1 \\ x + y + kz &= 1 \end{aligned}$$

Determinar para qué valores de k el sistema es:

i) Incompatible

ii) Compatible determinado

iii) Compatible indeterminado

5.- Determinar para qué condiciones de a y b tiene solución el siguiente sistema. Si tales condiciones se cumplen ¿Cuál es la solución del sistema?

$$\begin{aligned} x_1 - x_2 - x_3 &= a \\ -x_1 + x_2 + x_3 &= b \\ x_1 - x_2 + x_3 &= 2a \end{aligned}$$

6.- Un sistema de ecuaciones lineales en el que todos los términos independientes son nulos se dice que es "homogéneo". Un sistema homogéneo siempre es compatible puesto que admite la solución $x_1 = x_2 = \dots = x_n = 0$, llamada solución trivial.

Para cada uno de los siguientes sistemas homogéneos, determinar si el sistema admite soluciones no triviales y en caso afirmativo obtenerlas

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

b) $$\begin{aligned} 2x + 6y + z &= 0 \\ x + 3y &= 0 \\ -x - 3y + 4z &= 0 \end{aligned}$$

*CAPITULO VI MATRICES

INTRODUCCION

El estudio de los sistemas de ecuaciones lineales es un tema que de manera natural nos lleva al concepto de matriz. Así, en el capítulo V se introdujeron las matrices como una ayuda para representar, en forma tabular, un sistema de ecuaciones lineales, y facilitar con ello el empleo de las transformaciones elementales.

A diferencia del capítulo anterior, en éste nos ocuparemos de las matrices como entes matemáticos con existencia propia, independiente de los sistemas de ecuaciones lineales; aunque encuentran en éstos sus principales aplicaciones.

Definiremos la manera como las matrices pueden sumarse, multiplicarse y multiplicarse por escalares; analizando las principales consecuencias de dichas definiciones. Estudiaremos además algunos tópicos y tipos especiales de matrices que son importantes en el campo de las aplicaciones.

Desde un punto de vista algebraico, las matrices rompen con la monotonía establecida por los diversos sistemas numéricos, ya que la multiplicación viola una de las leyes que tradicionalmente se habían cumplido en dichos sistemas: la ley conmutativa. Esto trae como consecuencia que, en algunos aspectos, las matrices se separen del conocido comportamiento algebraico de los números.

INGENIERIA
ANEXO
BIBLIOTECA

VI.1 CONCEPTOS GENERALES

- Matriz

Podemos decir que una matriz es una "tabla" o "arreglo rectangular" de elementos que, usualmente, son números reales o complejos.

El concepto de matriz, sin embargo, puede generalizarse al caso en que los elementos sean polinomios, funciones, operadores o cualquier otro tipo de "entes matemáticos"; conservando su validez la mayoría de los conceptos y propiedades presentados en este capítulo, en el cual se considera a la matriz como un arreglo de números.

VI.1.1 DEFINICION

Una matriz de $m \times n$ con elementos en C es un arreglo de la forma

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

donde $a_{11}, a_{12}, \ldots, a_{mn} \in C$ y $m, n \in Z$.

Una matriz de m×n (léase "m por n") se dice también que es de "orden" m×n.

En forma abreviada, la matriz de la definición anterior puede expresarse como

$$\begin{bmatrix} a_{ij} \end{bmatrix}$$

donde i = 1, 2,..., m   y   j = 1, 2,..., n.

- Renglones y columnas

Al arreglo horizontal

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}$$

se le conoce como el primer renglón de la matriz, al arreglo

$$\begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix}$$

como el segundo renglón, y en general al arreglo horizontal

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

se le conoce como el i-ésimo renglón de la matriz.

En forma análoga, al arreglo vertical

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

se le conoce como la j-ésima columna.

Así, en una matriz de m×n pueden distinguirse m renglones - - - (i = 1, 2,..., m) y n columnas (j = 1, 2,..., n). En particular

INGENIERIA
A N E X O
BIBLIOTECA

si m = n se dice que la matriz es "cuadrada" de orden n.

Comúnmente se representa a las matrices con letras mayúsculas y a sus elementos con letras minúsculas.

Como ejemplos de matrices tenemos las siguientes

$$A = \begin{bmatrix} 1+i & 2 & -3i \\ 0 & 4i & 7 \\ -1 & 1-2i & 3 \\ -2 & 0 & 5 \end{bmatrix} \quad B = \begin{bmatrix} \frac{1}{4} & 0 & -\frac{1}{3} \end{bmatrix} \quad C = \begin{bmatrix} \pi i \\ 1-3i \\ \sqrt{2} \\ 0 \end{bmatrix} \quad D = \begin{bmatrix} -1 & 2 & 4 \\ 3 & 1 & 1 \\ 2 & -1 & 0 \end{bmatrix}$$

donde A es una matriz de 4×3, B es una matriz de 1×3 (conocida como "matriz renglón" o "vector renglón"), C es una matriz de 4×1 (conocida como "matriz columna" o "vector columna"), y D es una matriz cuadrada de orden tres.

- La igualdad de matrices

Se dice que dos matrices son iguales cuando tienen los mismos elementos y éstos se encuentran dispuestos de la misma manera en ambos arreglos.

Esta idea puede expresarse en términos más precisos con ayuda del símbolo a_{ij}, que representa al elemento que se encuentra en la posición correspondiente al renglón i y a la columna j de la matriz A. Así, por ejemplo, para las matrices A, B, C y D citadas anteriormente se tiene que

$a_{23} = 7$

$a_{32} = 1-2i$

$b_{13} = -\frac{1}{3}$

c_{33} no existe

$d_{33} = 0$, etc.

En consecuencia, la igualdad de matrices se define formalmente como sigue

VI.1.2 DEFINICION

Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices de m×n con elementos en C. Diremos que A y B son iguales, lo que representaremos con A = B, si:

$$a_{ij} = b_{ij} \quad ; \text{ para } i = 1, 2, \ldots, m \text{ y } j = 1, 2, \ldots, n$$

Así, por ejemplo, las matrices

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

no son iguales, a pesar de que son del mismo orden y tienen los mismos elementos; ya que, aunque se cumplen las igualdades

$$a_{11} = b_{11}$$
$$a_{12} = b_{12}$$
$$a_{13} = b_{13}$$
$$a_{21} = b_{21}$$

se tiene además que

$$a_{22} \neq b_{22}$$
y
$$a_{23} \neq b_{23}$$

por lo que A y B no satisfacen la condición de igualdad establecida por la definición VI.1.2.

También de VI.1.2 se sigue que, para las matrices

$$M = \begin{bmatrix} x & 2 & 1 \\ 3 & 0 & -5 \\ 0 & -4 & z \end{bmatrix} \quad \text{y} \quad N = \begin{bmatrix} -1 & 2 & 1 \\ 3 & y & -5 \\ 0 & -4 & w \end{bmatrix}$$

la igualdad M = N se cumple si y sólo si $x = -1$, $y = 0$ y $z = w$.

VI.2 ADICION DE MATRICES Y MULTIPLICACION POR UN ESCALAR

La adición de matrices

La primera de las operaciones con matrices que estudiaremos, y también la más sencilla, es la adición. Esta operación puede efectuarse cuando las matrices son del mismo orden y el resultado se obtiene sumando los elementos correspondientes de ambas matrices, de acuerdo con la siguiente definición.

VI.2.1 DEFINICION

Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices de m×n con elementos en C. La suma A + B es una matriz $S = [s_{ij}]$, de m×n, definida por

$$s_{ij} = a_{ij} + b_{ij} \quad ; \text{ para } i = 1,2,\ldots,m \text{ y } j = 1,2,\ldots n.$$

Así, por ejemplo, para las matrices

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

se tiene que

$$A + B = \begin{bmatrix} 3+1 & -5+2 \\ 0+(-2) & 1+i+(-i) \\ -2i+3 & 4+(-4) \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -2 & 1 \\ 3-2i & 0 \end{bmatrix}$$

mientras que la adición de A y C no puede efectuarse, ya que las matrices no son del mismo orden. Se dice por ello que A y C "no son conformables" para la adición y, en consecuencia, la suma $A + C$ no existe.

La adición de matrices, definida por VI.2.1, satisface las propiedades que se enuncian a continuación.

VI.2.2 TEOREMA

Si A, B y C son matrices de m×n cuyos elementos son números complejos, entonces:

i)	$A + (B + C) = (A + B) + C$	asociatividad
ii)	$A + B = B + A$	conmutatividad
iii)	Existe una matriz O de m×n tal que $A + O = A$	elemento idéntico
iv)	Existe una matriz -A de m×n tal que $A + (-A) = O$	elementos inversos

DEMOSTRACION

Se demostrarán a continuación las propiedades ii), iii) y iv).

ii) Sean $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices de m×n con elementos en C.

INGENIERIA
ANEXO
BIBLIOTECA

Por VI.2.1 se tiene que

$$A + B = \left[s_{ij}\right] = \left[a_{ij} + b_{ij}\right]$$

y

$$B + A = \left[t_{ij}\right] = \left[b_{ij} + a_{ij}\right]$$

Como a_{ij} y b_{ij} son números complejos $\forall$ i, j; por iii) de II.1.4

$$t_{ij} = b_{ij} + a_{ij} = a_{ij} + b_{ij} = s_{ij}; \ \forall\, i, j$$

por lo que, de VI.1.2

$$A + B = B + A$$

iii) Sea $A = \left[a_{ij}\right]$ una matriz de m×n con elementos en C.

Si definimos la matriz $O = \left[o_{ij}\right]$ como $o_{ij} = 0$ (cero) para $i = 1,2,\ldots, m$ y $j = 1,2,\ldots, n$; entonces

$$\begin{aligned} A + O &= \left[a_{ij} + o_{ij}\right] && \text{por VI.2.1} \\ &= \left[a_{ij} + 0\right] && \text{por definición de O} \\ &= \left[a_{ij}\right] && \text{por iv) de II.1.4} \\ A + O &= A && \text{como se quería.} \end{aligned}$$

A la matriz O, que es una matriz de m×n cuyos elementos son todos nulos, se le conoce como "matriz nula" o "matriz cero" de m×n.

iv) Sea $A = \left[a_{ij}\right]$ una matriz de m×n con elementos en C.

Si definimos la matriz $-A = \left[v_{ij}\right]$ como $v_{ij} = -a_{ij} \ \forall$ i, j; entonces

$$\begin{aligned} A+(-A) &= \left[a_{ij} + (v_{ij})\right] && \text{por VI.2.1} \\ &= \left[a_{ij} + (-a_{ij})\right] && \text{por definición de -A} \\ &= \left[0\right], \ \forall\, i, j && \text{por v) de II.1.4} \\ A+(-A) &= O && \text{por definición de O} \end{aligned}$$

y la prueba termina.

A la matriz -A, que es una matriz de m×n cuyos elementos son los simétricos de los elementos de A, se le conoce como la "simétrica de A" o la "negativa de A". □

La sustracción de matrices

La resta o sustracción de matrices puede definirse ahora, a partir de la adición y de iv) de VI.2.2, como sigue

> VI.2.3 DEFINICION
>
> Sean $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices de m×n con elementos en C. La diferencia A - B se define como
>
> $$A - B = A + (-B)$$

De acuerdo con esta definición, para obtener la diferencia A - B bastará con restar a los elementos de la matriz A los elementos correspondientes de la matriz B, puesto que

$$A - B = A + (-B) = \left[a_{ij} + (-b_{ij})\right] = \left[a_{ij} - b_{ij}\right]$$

Así, por ejemplo, para las matrices

$$A = \begin{bmatrix} 3 & -5 \\ 0 & 1+i \\ -2i & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ -2 & -i \\ 3 & -4 \end{bmatrix} \qquad y \qquad C = \begin{bmatrix} -1 & 5 & 2i \\ 7+i & 0 & 3 \end{bmatrix}$$

que vimos anteriormente, se tiene

$$A - B = \begin{bmatrix} 3-1 & -5-2 \\ 0-(-2) & 1+i-(-i) \\ -2i-3 & 4-(-4) \end{bmatrix} = \begin{bmatrix} 2 & -7 \\ 2 & 1+2i \\ -3-2i & 8 \end{bmatrix}$$

mientras que la diferencia A - C no existe.

De la definición VI.2.3 se sigue que dos matrices son conformables para la resta si y sólo si son del mismo orden.

La multiplicación por un escalar

En ocasiones, y particularmente desde el punto de vista de las aplicaciones, se requiere multiplicar una matriz por un número, al que genéricamente se le conoce como "escalar". Esta operación, denominada "multiplicación por un escalar", se define formalmente como sigue

> VI.2.4 DEFINICION
>
> Sean $A = \left[a_{ij}\right]$ una matriz de m×n con elementos en C y $\alpha \in C$. El producto αA es una matriz $E = \left[e_{ij}\right]$ de m×n, definida por
>
> $$e_{ij} = \alpha a_{ij} \quad ; \text{ para } i = 1, \ldots, m \quad y \quad j = 1, \ldots, n.$$

Así, por ejemplo, el producto del escalar $\alpha = 2i$ por la matriz

$$A = \begin{bmatrix} -i & 0 & 1 \\ i & -3 & 1+i \end{bmatrix}$$

es la matriz

$$\alpha A = (2i) \begin{bmatrix} -i & 0 & 1 \\ i & -3 & 1+i \end{bmatrix} = \begin{bmatrix} (2i)(-i) & (2i)(0) & (2i)(1) \\ (2i)(i) & (2i)(-3) & (2i)(1+i) \end{bmatrix} = \begin{bmatrix} 2 & 0 & 2i \\ -2 & -6i & -2+2 \end{bmatrix}$$

La multiplicación por un escalar satisface las siguientes propiedades.

VI.2.5 TEOREMA

Si A y B son matrices de m×n con elementos en C y α, $\beta \in C$, entonces:

i) $\alpha(A + B) = \alpha A + \alpha B$

ii) $(\alpha + \beta)A = \alpha A + \beta A$

iii) $\alpha(\beta A) = (\alpha\beta)A$

DEMOSTRACION

Se demostrará a continuación la propiedad i), dejando al lector como ejercicio la demostración de las restantes.

i) Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices de m×n con elementos en C y α un escalar de C, entonces

$$
\begin{aligned}
A + B &= [a_{ij} + b_{ij}] && \text{por VI.2.1} \\
\alpha(A + B) &= [\alpha(a_{ij} + b_{ij})] && \text{por VI.2.4} \\
&= [\alpha a_{ij} + \alpha b_{ij}] && \text{por vi) de II.1.4} \\
&= [\alpha a_{ij}] + [\alpha b_{ij}] && \text{por VI.2.1} \\
\alpha(A + B) &= \alpha A + \alpha B && \text{por VI.2.4}
\end{aligned}
$$

como se quería.

□

VI.2.6 EJERCICIOS

1.- Para las siguientes matrices

$$A = \begin{bmatrix} 5 & -2 & 7 \\ 0 & 3 & a_{23} \\ 2 & -1 & -3 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 2 & -3 \\ -2 & -1 & 0 \\ b_{31} & 1 & 3 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 & -1 \\ -3 & 0 & c_{23} \\ 5 & 1 & 3 \end{bmatrix}$$

determinar los valores de a_{23}, b_{31} y c_{23} que verifican la igualdad $A + 3B = 2C$

2.- Para las siguientes matrices

$$A = \begin{bmatrix} 2i & 1+i \\ 2 & i \\ -1 & 2-i \end{bmatrix} \quad B = \begin{bmatrix} 3 & -1 \\ 3i & i \\ 2 & 1-2i \end{bmatrix} \quad C = \begin{bmatrix} i & 2i \\ 2-i & 1 \\ -1 & 3i \end{bmatrix}$$

calcular A + B, A - B, B - A, 2A - C y 3B + 2C.

3.- Demostrar que si A, B y C son matrices de m×n cuyos elementos son números complejos, entonces:

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

4.- Demostrar que si A es una matriz de m×n con elementos en C y α, $\beta \in C$, entonces:

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

b) $\alpha(\beta A) = (\alpha\beta)A$

5.- Demostrar que si A y B son matrices de m×n cuyos elementos son números complejos, entonces:

a) $A - B = A + (-1)B$

b) $A - B = -(B - A)$

c) $0A = O$

VI.3.- MULTIPLICACION DE MATRICES

Consideremos nuevamente el sistema de ecuaciones lineales

$$\left.\begin{array}{r} 20x_1 + 100x_2 + 40x_3 = 400 \\ 0x_1 + x_2 - x_3 = 0 \end{array}\right\} \quad \text{---(1)}$$

visto al inicio de la sección V.2; y formemos ahora una matriz con los coeficientes de las ecuaciones, a la que llamaremos A; otra con las incógnitas, a la que llamaremos X, y una tercera con los términos independientes, a la que llamaremos B. Esto es

$$A = \begin{bmatrix} 20 & 100 & 40 \\ 0 & 1 & -1 \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad B = \begin{bmatrix} 400 \\ 0 \end{bmatrix}$$

Con ayuda de estas matrices podemos representar al sistema de ecuaciones (1) mediante la expresión

$$AX = B \qquad \text{---(2)}$$

siempre y cuando tengamos una definición adecuada para el producto AX.

Las condiciones que establece el sistema (1) son equivalentes, por VI.1.2, a la siguiente igualdad entre matrices

$$\begin{bmatrix} 20x_1 + 100x_2 + 40x_3 \\ 0x_1 + x_2 - x_3 \end{bmatrix} = \begin{bmatrix} 400 \\ 0 \end{bmatrix}$$

de donde se sigue que la expresión (2) representará al sistema (1) si y sólo si

$$AX = \begin{bmatrix} 20x_1 + 100x_2 + 40x_3 \\ 0x_1 + x_2 - x_3 \end{bmatrix}$$

Veamos ahora cómo puede obtenerse la matriz AX a partir de las matrices A y X.

El primer elemento de AX; es decir, el que se encuentra en el primer renglón y primera columna de dicha matriz, se obtiene sumando los productos de los elementos del primer renglón de A por sus elementos correspondientes en la primera columna de X. En forma esquemática:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \longrightarrow 20x_1 + 100x_2 + 40x_3$$

$$\begin{bmatrix} 20 & 100 & 40 \end{bmatrix}$$

Análogamente, el elemento que se encuentra en el segundo renglón y primera columna de AX se obtiene sumando los productos de los elementos del segundo renglón de A por los de la primera columna de X. Así

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0x_1 + 1x_2 + (-1)x_3$$

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

En general, si A y B son dos matrices tales que el número de columnas de A coincide con el número de renglones de B, el elemento que se encuentra en la posición correspondiente al renglón i y la columna j de la matriz producto AB, se obtiene sumando los productos de los elementos del renglón i de la matriz A por sus elementos correspondientes en la columna j de la matriz B.

Así, si A y B son las matrices

$$A = \begin{array}{cc} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} & \begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \end{matrix} \end{array} \qquad B = \begin{array}{c} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2q} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{nq} \end{bmatrix} \\ \begin{matrix} & & & \uparrow j & & \end{matrix} \end{array}$$

de $m\times n$ y $n\times q$ respectivamente, el elemento ubicado en el renglón i y columna j de la matriz producto AB, al que representaremos con p_{ij}, será

$$p_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

que, en forma compacta, puede expresarse como

$$p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

La multiplicación de matrices

Formalmente, se tiene la siguiente definición para la multiplicación de matrices.

> **VI.3.1 DEFINICION**
>
> Sean $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices con elementos en C, de $m\times n$ y $n\times q$ respectivamente. El producto AB es una matriz $P = \left[p_{ij}\right]$, de $m\times q$, definida por
>
> $$p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \; ; \text{ para } i = 1,\ldots,m \text{ y } j = 1,\ldots,q.$$

A manera de ejemplo, para las matrices

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

se tiene que $AB = \left[p_{ij}\right]$ es una matriz de 4×2, donde

$$p_{11} = \sum_{k=1}^{3} a_{1k}b_{k1} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} =$$

$$= 5(2) + (3)(-3) + (-1)(-1) = 10 - 9 + 1 = 2$$

$$p_{12} = \sum_{k=1}^{3} a_{1k}b_{k2} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} =$$

$$= (5)(0) + (3)(4) + (-1)(3) = 0 + 12 - 3 = 9$$

y de manera similar se calculan

$$p_{21} = (0)(2) + (1)(-3) + (-3)(-1) = 0 - 3 + 3 = 0$$
$$p_{22} = (0)(0) + (1)(4) + (-3)(3) = 0 + 4 - 9 = -5$$
$$p_{31} = (-2)(2) + (0)(-3) + (1)(-1) = -4 + 0 - 1 = -5$$
$$p_{32} = (-2)(0) + (0)(4) + (1)(3) = 0 + 0 + 3 = 3$$
$$p_{41} = (1)(2) + (-1)(-3) + (3)(-1) = 2 + 3 - 3 = 2$$
$$p_{42} = (1)(0) + (-1)(4) + (3)(3) = 0 - 4 + 9 = 5$$

por lo que

$$AB = \begin{bmatrix} 2 & 9 \\ 0 & -5 \\ -5 & 3 \\ 2 & 5 \end{bmatrix}$$

El producto AC no puede obtenerse, puesto que el número de columnas de A no es igual al número de renglones de C. Se dice entonces que las matrices A y C "no son conformables para el producto AC".

INGENIERIA ANEXO BIBLIOTECA

Curiosamente, estas mismas matrices sí resultan conformables para el producto CA.

En efecto, como puede verificarse fácilmente

$$CA = \begin{bmatrix} -2 & -4 & 4 \\ -6 & 1 & -2 \end{bmatrix}$$

De lo anterior se sigue que la multiplicación de matrices no es conmutativa; es decir, no puede establecerse que para dos matrices A y B (conformables para el producto AB) se tenga que AB = BA.

Puesto que AB y BA representan en general matrices diferentes, es importante hacer énfasis en el orden en que se multiplican. Así, en el producto AB se dice que la matriz A "premultiplica" a la matriz B; mientras que en el producto BA se dice que A "postmultiplica a B.

En algunos casos, como el del ejemplo anterior, la multiplicación puede efectuarse en un sentido, digamos AB, pero no en el otro, es decir BA. En otros casos la multiplicación puede efectuarse tanto en un sentido como en el otro, pero los resultados pueden ser diferentes o iguales según las matrices de que se trate.

Cuando dos matrices A y B son tales que AB = BA se dice que son "permutables" (también suele decirse que "conmutan").

Por ejemplo, para las matrices

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

se tiene que

$$AB = \begin{bmatrix} -3 & -4 \\ 0 & 2 \end{bmatrix} \quad y \quad BA = \begin{bmatrix} 6 & -3 \\ 12 & -7 \end{bmatrix}$$

por lo que A y B no son permutables; mientras que para

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

se tiene que

$$AC = \begin{bmatrix} 3 & 2 \\ -6 & 5 \end{bmatrix} \quad y \quad CA = \begin{bmatrix} 3 & 2 \\ -6 & 5 \end{bmatrix}$$

por lo que A y C son permutables.

La multiplicación de matrices satisface la ley asociativa que establece el siguiente enunciado.

VI.3.2 TEOREMA

Sean A, B y C matrices de $m\times n$, $n\times p$ y $p\times q$, respectivamente, cuyos elementos son números complejos, entonces:

$$A(BC) = (AB)C$$

DEMOSTRACION

Sean $A = \left[a_{ij}\right]$, $B = \left[b_{ij}\right]$ y $C = \left[c_{ij}\right]$ matrices de $m\times n$, $n\times p$ y $p\times q$, respectivamente. Entonces, por VI.3.1

$$BC = \left[\sum_{k=1}^{p} b_{ik}c_{kj}\right]$$

donde BC es una matriz de n×q. Entonces

$$A(BC) = \left[\sum_{h=1}^{n} a_{ih} \left(\sum_{k=1}^{p} b_{hk}c_{kj}\right)\right] \qquad \text{por VI.3.1}$$

$$= \left[\sum_{h=1}^{n} \left(\sum_{k=1}^{p} a_{ih}b_{hk}c_{kj}\right)\right] \qquad \text{por vi) de II.1.4}$$

$$= \left[\sum_{k=1}^{p} \left(\sum_{h=1}^{n} a_{ih}b_{hk}c_{kj}\right)\right] \qquad \text{puesto que podemos sumar en cualquier orden.}$$

$$= \left[\sum_{k=1}^{p} \left(\sum_{h=1}^{n} a_{ih}b_{hk}\right)c_{kj}\right] \qquad \text{por vi) de II.1.4}$$

$$A(BC) = (AB)C \qquad \text{por VI.3.1}$$

y la prueba termina. □

Para verificar el teorema anterior en un caso particular, consideremos las matrices

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

Obtengamos primero el producto

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

2x3 3x2 2x2

y, posteriormente, premultipliquemos éste por la matriz A, con lo que se obtiene

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

Por otra parte, obtengamos primero el producto

$$AB = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 3 \\ 1 & -2 & -1 \end{bmatrix}$$

y, a continuación, postmultipliquémoslo por C, con lo que se obtiene

$$(AB)C = \begin{bmatrix} 3 & 8 & 3 \\ 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} -4 & 2 \\ 2 & 0 \end{bmatrix}$$

y hemos llegado al mismo resultado, como cabía esperar del teorema VI.3.2

Con fundamento en dicho teorema podemos escribir simplemente

ABC

ya que no importa cual de los productos (AB o BC) se efectúe primero.

Consideradas simultáneamente, la adición y la multiplicación de matrices tienen las propiedades que se enuncian a continuación, conocidas como leyes distributivas de la multiplicación sobre la adición.

VI.3.3 TEOREMA

Sean A, B y C matrices de m×n, n×p y n×p, respectivamente, y D, E y F matrices de m×n, m×n y n×p, respectivamente, cuyos elementos son números complejos; entonces:

i) $A(B + C) = AB + AC$

ii) $(D + E)F = DF + EF$

DEMOSTRACION

Se demostrará a continuación la distributividad por la izquierda (propiedad i), dejando al lector como ejercicio la demostración de la distributividad por la derecha (propiedad ii).

Sean $A = \left[a_{ij}\right]$, $B = \left[b_{ij}\right]$ y $C = \left[c_{ij}\right]$ matrices de m×n, n×p y n×p, respectivamente; entonces

$$\begin{aligned}
B + C &= \left[b_{ij} + c_{ij}\right] && \text{por VI.2.1}\\
A(B + C) &= \left[\sum_{k=1}^{n} a_{ik}\,(b_{kj} + c_{kj})\right] && \text{por VI.3.1}\\
&= \left[\sum_{k=1}^{n} (a_{ik}b_{kj} + a_{ik}c_{kj})\right] && \text{por vi) de II.1.4}\\
&= \left[\sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}\right] && \text{por ii) y iii) de II.1.4}\\
&= \left[\sum_{k=1}^{n} a_{ik}b_{kj}\right] + \left[\sum_{k=1}^{n} a_{ik}c_{kj}\right] && \text{por VI.2.1}\\
A(B + C) &= AB + AC && \text{por VI.3.1}
\end{aligned}$$

y la prueba termina.

□

Matriz identidad

Se conoce como "matriz identidad" de orden n a una matriz cuadrada de orden n que es de la forma

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Como puede verse, esta matriz está formada con unos y ceros únicamente. Los elementos iguales a uno son aquellos en que coinciden el número del renglón y el de la columna donde se encuentran, y todos los demás elementos son iguales a cero.

Lo anterior permite establecer la siguiente definición para la matriz identidad.

VI.3.4 DEFINICION

Se llama matriz identidad de orden n a la matriz cuadrada de orden n $I_n = \left[\delta_{ij}\right]$, tal que

$$\delta_{ij} = 1, \quad \text{si } i = j$$

y

$$\delta_{ij} = 0, \quad \text{si } i \neq j$$

Al símbolo δ_{ij} de la definición anterior se le conoce como "delta de Kronecker".

La matriz identidad juega un papel muy importante en el álgebra de matrices, ya que constituye un elemento idéntico para la multiplicación.

Por ejemplo, si premultiplicamos la matriz

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

por la matriz identidad de orden tres se tendrá

$$I_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -2i & 4 \\ 7 & 0 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -2i & 4 \\ 7 & 0 \end{bmatrix} = A$$

Si ahora postmultiplicamos dicha matriz por I_2 se tendrá también

$$AI_2 = \begin{bmatrix} 3 & -1 \\ -2i & 4 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -2i & 4 \\ 7 & 0 \end{bmatrix} = A$$

En general, se tiene el siguiente teorema

VI.3.5 TEOREMA

Si A es una matriz de $m \times n$ con elementos en C, entonces:

i) $I_mA = A$

ii) $AI_n = A$

DEMOSTRACION

Se demuestra a continuación la parte i) dejando como ejercicio al lector la demostración de ii).

i) Sea $A = \left[a_{ij}\right]$ una matriz de $m \times n$ con elementos en C y sea $I_m = \left[\delta_{ij}\right]$

$$I_mA = \left[\sum_{k=1}^{m} \delta_{ik}a_{kj}\right] \qquad \text{por VI.3.1}$$

$$= \left[\delta_{ii}a_{ij}\right] \qquad \text{por VI.3.4}$$

$$= \left[1 \cdot a_{ij}\right] \qquad \text{por VI.3.4}$$

$$= \left[a_{ij}\right] \qquad \text{por iv) de II.1.4}$$

I_mA - como se quería.

□

INGENIERIA
ANEXO
BIBLIOTECA

VI.3.6 EJERCICIOS

1.- Para las siguientes matrices

3×3 3×2 2×3

$$A = \begin{bmatrix} 2 & 1-i & 1 \\ i & 0 & 3 \\ 0 & 1+i & -i \end{bmatrix} \qquad B = \begin{bmatrix} 2i & 1+i \\ 1 & 0 \\ 0 & 2-i \end{bmatrix} \qquad C = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 2 & -1 \end{bmatrix}$$

calcular, de ser posible, AB, BA, BC, CB, ABC, CBA y BCA.

2.- Demostrar que si A, B y C son matrices de $m \times n$, $n \times p$ y $n \times p$, respectivamente, y D, E y F son matrices de $m \times n$, $m \times n$ y $n \times p$, respectivamente, cuyos elementos son números complejos, entonces:

a) $(D + E)F = DF + EF$

b) $A(B - C) = AB - AC$

c) $(D - E)F = DF - EF$

3.- Si A y B son dos matrices de $m \times n$ y $n \times p$, respectivamente, y α es un número complejo cualquiera, entonces:

$$\alpha(AB) = (\alpha A)B = A(\alpha B)$$

a) Ilustrar el enunciado anterior mediante un ejemplo.

b) Demostrar dicho enunciado.

4.- Demostrar que si A es una matriz de $m \times n$ con elementos en C, entonces:

$$AI_n = A$$

5.- Para las siguientes matrices

$$A = \begin{bmatrix} a_{11} & 3 & -2 \\ -1 & 2 & -1 \\ 1 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & 0 & 1 \\ -1 & 2 & b_{23} \\ -2 & 3 & b_{33} \end{bmatrix}$$

determinar los valores de a_{11}, b_{11}, b_{23} y b_{33} que satisfacen la igualdad

$$AB = I$$

VI.4 INVERSA DE UNA MATRIZ

En ciertos casos, para una matriz A es posible hallar una matriz X tal que XA = I = AX.

Por ejemplo, para la matriz

$$A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$$

se tiene que la matriz

$$X = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$$

es tal que

$$XA = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$AX = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Se dice entonces que X es "inversa" de la matriz A y se representa con A^{-1}.

VI.4.1 DEFINICION

Sea A una matriz de n×n con elementos en C. Una matriz X se dice que es inversa de A si

$$XA = I_n = AX$$

y se representa con A^{-1}.

Cabe hacer notar que la igualdad XA = AX sólo es posible cuando A y X son matrices cuadradas del mismo orden; en consecuencia, para que una matriz A tenga inversa es condición necesaria que sea cuadrada. Además, la inversa deberá ser también cuadrada y del mismo orden que A.

La definición VI.4.1 establece lo que deberá entenderse por inversa de una matriz cuadrada, pero no dice que toda matriz cuadrada tenga inversa, ni que dicha inversa (en caso de existir) sea única.

En lo que se refiere al primer punto, se puede demostrar, mediante un ejemplo, que no todas las matrices cuadradas tienen inversa.

En efecto, para la matriz

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

una matriz

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

tal que XA = I deberá cumplir con

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

INGENIERIA
ANEXO
BIBLIOTECA

esto es

$$\begin{bmatrix} 3x_{11} & 0 \\ 3x_{21} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

igualdad que, como puede verse, no se satisface para ningún valor de los elementos x_{11}, x_{12}, x_{21}, x_{22}. Luego, no existe inversa para la matriz propuesta.

A las matrices que tienen inversa les llamamos "no singulares"* y a las que no tienen inversa "singulares".

> **VI.4.2 DEFINICION**
>
> Sea A una matriz de n×n con elementos en C. Se dice que A es no singular si existe A^{-1}, en caso contrario se dice que A es singular.

En lo que se refiere a la unicidad, se puede demostrar que la inversa de una matriz cuadrada (si existe) es única, como lo establece el siguiente teorema, en el que se enuncian además otras propiedades importantes de la inversa.

* *Algunos autores emplean el término "regular" en vez de "no singular".*

VI.4.3 TEOREMA

Si A y B son dos matrices no singulares del mismo orden y $\lambda \in C$, entonces:

i) A^{-1} es única

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

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

iv) $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$, si $\lambda \neq 0$

DEMOSTRACION

Se demuestran a continuación i) y iii) dejando al lector como ejercicio la demostración de ii) y iv).

i) Sea A una matriz de $n \times n$ no singular, y sean X, Y dos inversas de A; entonces, por VI.4.1

$$XA = I_n = AX \qquad \text{y} \qquad YA = I_n = AY$$

Por otra parte

$X = XI_n$	por ii) de VI.3.5
$= X(AY)$	por hipótesis
$= (XA)Y$	por VI.3.2
$= I_n Y$	por hipótesis
$X = Y$	por i) de VI.3.5

y en consecuencia la inversa es única.

iii) Sean A y B dos matrices de $n \times n$ no singulares. Por VI.4.2

existen A^{-1} y B^{-1} y puede formarse el producto

$$B^{-1}A^{-1}$$

para el cual se tiene que

$(B^{-1}A^{-1})(AB) = (B^{-1}A^{-1})[(A)(B)]$	
$= [(B^{-1}A^{-1})A]B$	por VI.3.2
$= [B^{-1}(A^{-1}A)]B$	por VI.3.2
$= (B^{-1}I_n)B$	por VI.4.1
$= B^{-1}B$	por VI.3.5
$(B^{-1}A^{-1})(AB) = I_n$	por VI.4.1

En forma análoga puede demostrarse que

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

y, en consecuencia, de VI.4.1 se tiene que $B^{-1}A^{-1}$ es la inversa de AB; esto es

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

como se quería. □

Cabe hacer notar que de la expresión anterior se sigue que el producto de dos matrices no singulares es una matriz no singular; resultado importante del que haremos uso más adelante.

- Cálculo de la inversa por transformaciones elementales.

Como hemos visto, hay matrices cuadradas que tienen inversa y hay otras que no la tienen; por tanto, cabe ahora preguntarse cómo podemos saber si una matriz dada A tiene inversa o no la tiene y, en caso de que la tenga, cómo podemos obtenerla.

Un primer procedimiento que podría ocurrirse consiste en plantear una matriz desconocida X, cuyos elementos x_{ij} queremos determinar. Multiplicar dicha matriz por A y obtener los valores de x_{ij} que hacen posible las igualdades

$$XA = I = AX$$

Este procedimiento, que se fundamenta directamente en la definición de inversa, nos conduciría sin embargo a un sistema de n^2 ecuaciones con n^2 incógnitas, que para valores grandes de n resulta muy arduo resolver.

En su lugar se propone a continuación un método más práctico que se basa en el empleo de las transformaciones elementales por renglón, las cuales se manejaron en el capítulo anterior.

El método consiste en aplicar una sucesión de transformaciones elementales a la matriz A hasta obtener la matriz identidad, y aplicar esta misma sucesión de transformaciones a la matriz I_n con lo que se obtiene A^{-1}. Si no es posible transformar la matriz A en la matriz identidad entonces no existe A^{-1}.

Con el propósito de fundamentar teóricamente este método introduciremos a continuación el concepto de matriz elemental y estableceremos algunos resultados que nos permitirán concluir la validez del método.

— Matrices elementales

INGENIERIA
ANEXO
BIBLIOTECA

Consideremos la matriz

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

y apliquémosle la transformación elemental (T_1) que consiste en intercambiar los renglones segundo y tercero; se obtiene entonces la matriz

$$A_1 = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 4 & -2 \\ 3 & 5 & 6 & 7 \end{bmatrix}$$

Esta matriz puede obtenerse también como resultado de una multiplicación.

En efecto, si premultiplicamos A por la matriz

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

se tendrá

$$E_1A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 & 3 \\ 3 & 5 & 6 & 7 \\ 0 & 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 4 & -2 \\ 3 & 5 & 6 & 7 \end{bmatrix} = A_1$$

La matriz E_1 recibe el nombre de "matriz elemental" y, como puede verse, se obtiene a partir de la matriz identidad efectuando en ella la transformación correspondiente (en este caso el intercambio de los renglones 2 y 3).

Se obtiene así el equivalente algebraico de "aplicar una transformación elemental" que es "premultiplicar por una matriz elemental".

Es claro que existen tres tipos de matrices elementales, correspondientes a los tres tipos de transformaciones elementales.

VI.4.4 DEFINICION

Una matriz elemental es aquella que se obtiene aplicando a I_n una transformación elemental y se representa con:

$I_n^{(i,j)}$ si se obtiene intercambiando los renglones i y j de I_n.

$I_n^{k(i)}$ si se obtiene multiplicando por un número $k \neq 0$ el renglón i de I_n.

$I_n^{k(i,j)}$ si se obtiene multiplicando por k el renglón i de I_n y sumando el resultado al renglón j.

De acuerdo con esta notación, a la matriz E_1 del ejemplo anterior le corresponde el símbolo $I_3^{(2,3)}$

VI.4.5 TEOREMA

Si A es una matriz de m×n con elementos en C, entonces:

i) $I_m^{(i,j)}A$ es la matriz que se obtiene intercambiando los renglones i y j de la matriz A.

ii) $I_m^{k(i)}A$ es la matriz que se obtiene multiplicando por k el renglón i de la matriz A.

iii) $I_m^{k(i,j)}A$ es la matriz que se obtiene sumando al renglón j de la matriz A el renglón i multiplicado por k.

DEMOSTRACION

Se demuestra a continuación la proposición i), las proposiciones ii) y iii) se pueden demostrar de manera similar.

Puesto que $I_m^{(i,j)} = \left[e_{rc}\right]$ es una matriz identidad con los renglones i y j intercambiados, se tiene que

$$\text{para } r \neq i, j;\ e_{rc} = \delta_{rc}$$

$$\text{para } r = i;\ e_{ic} = \begin{cases} 1, & \text{si } c = j \\ 0, & \text{si } c \neq j \end{cases}$$

$$\text{para } r = j;\ e_{jc} = \begin{cases} 1, & \text{si } c = i \\ 0, & \text{si } c \neq i \end{cases}$$

Sea $I_m^{(i,j)}A = B = \left[b_{rc}\right]$

(1) Para $r \neq i$, j se tiene que

$$b_{rc} = \sum_{k=1}^{m} e_{rk}a_{kc} = \sum_{k=1}^{m} \delta_{rk}a_{kc} = \delta_{rr}a_{rc} = 1 \cdot a_{rc} = a_{rc} \quad \forall\ c$$

por lo que el renglón r de B es igual al renglón r de A.

(2) Para r = i se tiene que

$$b_{rc} = b_{ic} = \sum_{k=1}^{m} e_{ik}a_{kc} = e_{ij}a_{jc} = 1 \cdot a_{jc} = a_{jc} \quad \forall\ c$$

por lo que el renglón i de B es igual al renglón j de A.

(3) para r = j se tiene que

$$b_{rc} = b_{jc} = \sum_{k=1}^{m} e_{jk}a_{kc} = e_{ji}a_{ic} = 1 \cdot a_{ic} = a_{ic} \quad \forall\ c$$

por lo que el renglón j de B es igual al renglón i de A.

En consecuencia, de (1), (2) y (3) la matriz B se obtiene intercambiando los renglones i y j de la matriz A, como se quería. □

De acuerdo con el teorema anterior, cuando una matriz se premultiplica por $I_n^{(i,j)}$ se intercambian sus renglones i y j. En particular, si es la misma $I_n^{(i,j)}$ la que se premultiplica por dicha matriz, tomando en cuenta que $I_n^{(i,j)}$ se obtiene intercambiando los renglones i y j de I_n, se tendrá que

$$I_n^{(i,j)} I_n^{(i,j)} = I_n$$

por lo que $I_n^{(i,j)}$ tiene inversa, que es la misma $I_n^{(i,j)}$.

Razonando de manera similar podemos concluir que la inversa de $I_n^{k(i)}$ es $I_n^{\frac{1}{k}(i)}$, y que la inversa de $I_n^{k(i,j)}$ es $I_n^{-k(i,j)}$.

En consecuencia, se puede establecer que

> **VI.4.6 TEOREMA**
>
> Las matrices elementales son no singulares.

y, tomando en cuenta el teorema VI.4.3, se tiene que

> **VI.4.7 TEOREMA**
>
> El producto de matrices elementales es una matriz no singular

- Justificación del método.

Estamos ahora en condiciones de fundamentar el método descrito para obtener la inversa de una matriz mediante transformaciones elementales.

En efecto, sea A una matriz de n×n con elementos en C y

i) Supongamos que existe una sucesión (finita) de transformaciones elementales

$$T_1, T_2, \ldots, T_k$$

que aplicada a la matriz A la transforma en la matriz identidad de orden n; esquemáticamente:

$$A \xrightarrow{T_1} A_1 \xrightarrow{T_2} \cdots \quad A_{k-1} \xrightarrow{T_k} I_n$$

Entonces, existe una sucesión (finita) de matrices elementales

$$E_1, E_2, \ldots, E_k$$

tales que

$$E_k(\ldots(E_2(E_1A))\ldots) = I_n$$

por lo que

$$(E_k \ldots E_2 E_1)A = I_n$$

Si llamamos P al producto $E_k \ldots E_2 E_1$, se tendrá que

$$PA = I_n$$

Por otra parte, como P es un producto de matrices elementales, de VI.4.7 se sigue que P es no singular y existe P^{-1}; por tanto

$$P^{-1}(PA) = P^{-1} I_n$$

$$(P^{-1} P)A = P^{-1} I_n$$

$$I_n A = P^{-1} I_n$$

$$A = P^{-1}$$

y postmultiplicando ahora por P

$$AP = P^{-1} P$$

$$AP = I_n$$

En consecuencia

$$PA = I_n = AP$$

y P es la inversa de A.

El desarrollo anterior indica que la inversa de A (la matriz P) puede calcularse como el producto de k matrices elementales, las cuales deben obtenerse previamente; sin embargo, la matriz P puede calcularse directamente a partir de I_n como se muestra a continuación.

En efecto, se tiene que

$$P = E_k \dots E_2 E_1$$

$$P = (E_k \dots E_2 E_1) I_n$$

$$P = E_k(\dots(E_2(E_1 I_n))\dots)$$

de donde podemos concluir que P se obtiene aplicando a I_n la sucesión de transformaciones elementales $T_1, T_2, \dots, T_k$.

Lo anterior sugiere, para propósitos de cálculo, el empleo de un arreglo formado por dos matrices de n×n.

Inicialmente el arreglo tiene del lado izquierdo a la matriz A y del lado derecho a la matriz identidad I_n. Se efectúan entonces (en ambas matrices simultáneamente) las transformaciones necesarias para obtener en el lado izquierdo la matriz I_n, y al finalizar el proceso se obtiene en el lado derecho la matriz A^{-1}.

En forma esquemática

$$\left[A \mid I_n\right] \xrightarrow{T_1} \dots \xrightarrow{T_k} \left[I_n \mid A^{-1}\right]$$

Para ilustrar lo anterior mediante un ejemplo consideremos la matriz

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

cuya inversa deseamos obtener.

Formemos primero el arreglo $\left[A \mid I_3\right]$ y efectuemos a continuación las transformaciones necesarias para obtener en el lado izquierdo una matriz escalonada (como en el método de Gauss).

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 1 & 0 & 0 \\ 2 & 6 & 1 & 0 & 1 & 0 \\ -1 & -4 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow{T_1} \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & -1 & 2 & 1 & 0 & 1 \end{array}\right] \xrightarrow{T_2} \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 0 \end{array}\right]$$

Y una vez que se ha obtenido ésta continuamos con el proceso hasta obtener en el lado izquierdo la matriz identidad

$$\xrightarrow{T_3}\left[\begin{array}{rrr|rrr}1&3&0&1&0&0\\0&-1&0&5&-2&1\\0&0&1&-2&1&0\end{array}\right]\xrightarrow{T_4}\left[\begin{array}{rrr|rrr}1&0&0&16&-6&3\\0&-1&0&5&-2&1\\0&0&1&-2&1&0\end{array}\right]\xrightarrow{T_5}\left[\begin{array}{rrr|rrr}1&0&0&16&-6&3\\0&1&0&-5&2&-1\\0&0&1&-2&1&0\end{array}\right]$$

con lo que se llega al arreglo $\left[I_3 \middle| A^{-1}\right]$ y, en consecuencia, para la matriz A en cuestión se tiene que

$$A^{-1} = \begin{bmatrix}16&-6&3\\-5&2&-1\\-2&1&0\end{bmatrix}$$

ii) Supongamos ahora que la matriz A no puede ser transformada en la matriz identidad mediante una sucesión de transformaciones elementales.

Se tiene entonces una sucesión de transformaciones elementales

$$T_1,\ T_2,\ \ldots,\ T_r$$

que aplicada a la matriz A la transforma en una matriz C que tiene un renglón de ceros; y existe por tanto una sucesión de matrices elementales

$$E_1,\ E_2,\ \ldots,\ E_r$$

tales que

$$(E_r \ldots E_2\ E_1)A = C$$

Si llamamos Q al producto $E_r \ldots E_2\ E_1$, se tendrá que

$$QA = C$$

Por VI.4.7 Q es una matriz no singular, y si A fuese también no singular por iii) de VI.4.3 se tendría que C es no singular; sin embargo, C es singular puesto que tiene un renglón de ceros y para cualquier matriz M el producto MC tiene un renglón de ceros, es decir que no existe M tal que MC = I.

ANEXO BIBLIOTECA

En consecuencia la matriz A es singular y no existe A^{-1}.

Para ilustrar este caso consideremos la matriz

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

Formemos el arreglo $\left[A \middle| I_3\right]$ y tratemos de obtener en el lado izquierdo la matriz identidad

$$\left[\begin{array}{rrr|rrr}1&3&0&1&0&0\\2&6&1&0&1&0\\-1&-3&2&0&0&1\end{array}\right]\xrightarrow{T_1}\left[\begin{array}{rrr|rrr}1&3&0&1&0&0\\0&0&1&-2&1&0\\0&0&2&1&0&1\end{array}\right]\xrightarrow{T_2}\left[\begin{array}{rrr|rrr}1&3&0&1&0&0\\0&0&1&-2&1&0\\0&0&0&5&-2&1\end{array}\right]$$

Como se ve, en el lado izquierdo del último arreglo se ha obtenido una matriz con un renglón de ceros, por lo que la matriz A es singular y no tiene inversa.

VI.4.8 EJERCICIOS

1.- Para las matrices

$$A = \begin{bmatrix}5&-1&-1&1\\-2&1&0&2\\-2&0&1&-1\end{bmatrix}\ 3\times4 \qquad y \qquad B = \begin{bmatrix}1&1&1\\2&3&2\\2&2&3\\0&0&0\end{bmatrix}$$

4x3

obtener el producto AB

¿Puede decirse que A es inversa de B? ¿Por qué?

2.- Demostrar que si A es una matriz no singular con elementos en C y $\lambda \in C$, entonces:

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

b) $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$ si $\lambda \neq 0$

3.- Para cada una de las siguientes matrices, obtener una matriz P tal que PA sea una matriz escalonada:

i) $A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -2 & 4 & 0 \\ 3 & 1 & 7 & 5 \end{bmatrix}$ ii) $A = \begin{bmatrix} i & 0 & 1 \\ 0 & 1 & i \\ -i & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$

4.- Obtener la inversa, si existe, de cada una de las siguientes matrices

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

y $C = \begin{bmatrix} i & -1 & 2 \\ 0 & 2-i & 1+3i \\ 0 & 1 & i \end{bmatrix}$

5.- Para la matriz

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

determinar el conjunto de valores para los cuales A^{-1} existe y obtenerla.

VI.5 ECUACIONES CON MATRICES

Consideremos ahora las matrices

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

y preguntémonos si es posible hallar una matriz X que satisfaga la siguiente relación .

$$AX + B = 3X$$

Hemos planteado con ello una ecuación entre matrices, donde la matriz X es la incógnita.

En ciertos casos estas ecuaciones, conocidas como ecuaciones matriciales, pueden resolverse siguiendo el mismo procedimiento que se emplea para resolver ecuaciones planteadas con números; esto es, tratando de "despejar" la incógnita en términos de los otros elementos que intervienen en la ecuación. Sin embargo, las propiedades de las operaciones con matrices presentan, como hemos visto, algunas diferencias respecto a las propiedades de las operaciones con números, por lo que debemos tener especial cuidado en que los "pasos" efectuados en el despeje sean válidos en el álgebra de matrices.

Volviendo al ejemplo que nos ocupa, para "pasar" la matriz B al miembro derecho de la ecuación podemos proceder de la siguiente manera:

Por iv) de VI.2.2 existe -B, por lo que, de la expresión original

$$(AX + B) + (-B) = 3X + (-B)$$

en consecuencia

$$AX + \left[B + (-B) \right] = 3X + (-B) \qquad \text{por i) de VI.2.2}$$

$$AX + O = 3X + (-B) \qquad \text{por iv) de VI.2.2}$$

$$AX = 3X + (-B) \qquad \text{por iii) de VI.2.2}$$

Para "pasar" ahora la matriz 3X al miembro izquierdo de la ecua - ción:

por iv) de VI.2.2 existe $-(3X)$, y de la expresión anterior

$$-(3X) + AX = -(3X) + \left[3X + (-B) \right]$$

de donde

$$-(3X) + AX = \left[-(3X) + 3X \right] + (-B) \qquad \text{por i) de VI.2.2}$$

$$-(3X) + AX = O + (-B) \qquad \text{por iv) de VI.2.2}$$

$$-(3X) + AX = -B \qquad \text{por iii) de VI.2.2}$$

Ahora, para "factorizar" a X procedemos como sigue:

Probamos primero que

$$-(\alpha X) = (-\alpha)X$$

por lo que podemos escribir simplemente $-\alpha X$.

En efecto, si α es un escalar de C y X una matriz de m×n con ele mentos en C:

$$\alpha X + \left[(-\alpha)X \right] = \left[\alpha + (-\alpha) \right] X \qquad \text{por ii) de VI.2.5}$$

$$= 0 \cdot X \qquad \text{por v) de II.1.4}$$

$$\alpha X + (-\alpha)X = 0 \qquad \text{por 5.c) de VI.2.6}$$

de donde

$$(-\alpha)X = -(\alpha X) \qquad \text{por iv) de VI.2.2}$$

Llevando este resultado al desarrollo anterior podemos escribir

$$(-3)X + AX = -B$$

de donde se sigue que

$$(-3)(IX) + AX = -B \qquad \text{por i) de VI.3.5}$$

$$\left[(-3)I\right] X + AX = -B \qquad \text{por 3 de VI.3.6}$$

$$(-3I)X + AX = -B \qquad \text{por lo que acabamos de demostrar}$$

$$(-3I + A)X = -B \qquad \text{por ii) de VI.3.3}$$

Finalmente, para despejar X premultiplicamos por la inversa de $(-3I + A)$, lo cual es válido sólo si dicha matriz es no singular. Así:

Si $\exists\,(-3I + A)^{-1}$ se tiene que

$$(-3I + A)^{-1} \left[(-3I + A)X \right] = (-3I + A)^{-1}(-B)$$

y en consecuencia

$$\left[(-3I + A)^{-1}(-3I + A)\right] X = (-3I + A)^{-1}(-B) \quad \text{por VI.3.2}$$

$$IX = (-3I + A)^{-1}(-B) \quad \text{por VI.4.1}$$

$$X = (-3I + A)^{-1}(-B) \quad \text{por i) de VI.3.5}$$

con lo que hemos conseguido expresar a X en términos de las matrices A y B y del escalar 3 que aparecen en la ecuación.

En el desarrollo anterior hemos efectuado uno a uno todos los pasos necesarios para resolver la ecuación, y los hemos justificado formalmente con el propósito de ilustrar cómo puede despejarse la incógnita en una ecuación matricial empleando las propiedades del álgebra de matrices; sin embargo, en la práctica es aconsejable suprimir los pasos que resultan obvios y sólo especificar detalladamente aquellas partes del proceso donde existan dudas. Por otra parte, la justificación formal de los mismos suele dejarse para las demostraciones únicamente.

Regresando al ejemplo, para obtener los elementos de la matriz X bastará con efectuar las operaciones indicadas en la última expresión obtenida. Así

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

para calcular la inversa de esta matriz procedemos como sigue

$$\left[\begin{array}{cc|cc} 2 & 4 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 2 & \frac{1}{2} & 0 \\ 3 & 5 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 2 & \frac{1}{2} & 0 \\ 0 & -1 & -\frac{3}{2} & 1 \end{array}\right]$$

$$\left[\begin{array}{cc|cc} 1 & 2 & \frac{1}{2} & 0 \\ 0 & 1 & \frac{3}{2} & -1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -\frac{5}{2} & 2 \\ 0 & 1 & \frac{3}{2} & -1 \end{array}\right]$$

por lo que

$$X = (-3I + A)^{-1}(-B) = \begin{bmatrix} -\frac{5}{2} & 2 \\ \frac{3}{2} & -1 \end{bmatrix}\begin{bmatrix} -4 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 6 & -3 \\ -5 & -3 & 2 \end{bmatrix}$$

es la matriz que satisface la ecuación propuesta.

— Representación matricial de un sistema de ecuaciones lineales.

Otro ejemplo de ecuación matricial, de uso frecuente en las aplicaciones, lo constituye la llamada representación matricial de un sistema de ecuaciones.

Como se sugirió al inicio de la sección VI.3, con base en las definiciones de igualdad y de multiplicación de matrices, un sistema de m ecuaciones lineales con n incógnitas puede quedar representado por la expresión

$$AX = B$$

donde A es una matriz de m×n que se conoce como "matriz de coeficientes" del sistema, X es una matriz de n×1 conocida como "vector de incógnitas" y B es una matriz de m×1 conocida como "vector de términos independientes".

Esta ecuación puede resolverse premultiplicando por A^{-1} cuando A sea una matriz no singular.

En efecto, si $\exists\ A^{-1}$ se tiene que

$$A^{-1}(AX) = A^{-1}B$$

$$(A^{-1}A)X = A^{-1}B$$

$$IX = A^{-1}B$$

$$X = A^{-1}B$$

Así, por ejemplo, el sistema de ecuaciones lineales

$$\begin{aligned} x_1 + \quad\ 3x_3 &= 2 \\ x_2 - 2x_3 &= -1 \\ x_1 + x_2 + 2x_3 &= 3 \end{aligned}$$

puede expresarse en forma matricial como $AX = B$, donde

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 1 & 1 & 2 \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad y \qquad B = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

Para determinar si existe A^{-1} y obtenerla procedemos como sigue

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

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

por lo que

INGENIERIA
A N E X O
BIBLIOTECA

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

y en consecuencia

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

es la solución del sistema; es decir

$$x_1 = -4, \; x_2 = 3, \; x_3 = 2$$

Diferencias entre el álgebra de números y el álgebra de matrices.

Con objeto de prevenir al lector sobre errores que pueden cometerse al aplicar descuidadamente a las matrices las reglas usuales en el manejo de los números, se presentan a continuación algunas diferencias importantes entre el álgebra de los números y el álgebra de las matrices.

1) La diferencia más general consiste en que podemos sumar o multiplicar dos números cualesquiera, mientras que no siempre podemos hacerlo con las matrices, puesto que éstas deben ser conformables para la operación a efectuar.

   Como consecuencia de ello podemos encontrarnos con ecuaciones matriciales "mal planteadas", en el sentido de que no puedan efectuarse las operaciones propuestas. Por ejemplo, si para las matrices A y B del inicio de esta sección planteamos la ecuación.

$$XA + B = 3X$$

se tendrá que, como A es de 2×2, la matriz X deberá ser de m×2 para que exista el producto XA, y en tales circunstancias XA será también de m×2 por lo que no podrá sumarse con B. Luego, no existe matriz X alguna que permita efectuar las operaciones propuestas en el miembro izquierdo de la ecuación.

Las diferencias más significativas, sin embargo, son las relacionadas con la multiplicación; entre las cuales se cuentan las siguientes.

2) La multiplicación de números es conmutativa, mientras que la multiplicación de matrices no lo es.

Como consecuencia de ello se tiene que, para los números

$$b = c \implies ab = ac$$

y también

$$b = c \implies ab = ca$$

mientras que para las matrices

$$B = C \implies AB = AC$$

pero

$$B = C \not\Rightarrow AB = CA$$

Así, por ejemplo, al despejar la incógnita X de una ecuación matricial

$$AX = B$$

se premultiplican ambos miembros por A^{-1} con lo que se obtiene

$$X = A^{-1} B$$

resultado que, en general, difiere de

$$B A^{-1}$$

que se obtendría premultiplicando por A^{-1} el miembro izquierdo y postmultiplicando por dicha matriz el miembro derecho.

3) El producto de dos números diferentes de cero es diferente de cero, mientras que el producto de dos matrices diferentes de la matriz cero puede ser igual a la matriz cero.

Por ejemplo, para las matrices

$$A = \begin{bmatrix} -3 & -1 \\ 9 & 3 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & -2 \\ -3 & 6 \end{bmatrix}$$

se tiene que $A \neq O$, $B \neq O$ y $AB = O$.

4) La ley cancelativa para la multiplicación tiene una aplicación más restringida en el caso de las matrices.

En efecto, para los números se tiene que

$$\text{si } a \neq 0 \text{ entonces } ab = ac \implies b = c$$

lo cual no es válido para las matrices ya que, por ejemplo, para las matrices A y B citadas anteriormente se tiene que $A \neq O$ y

$$AB = AO$$

sin embargo, esto no implica que $B = O$; es decir, no podemos "cancelar" la matriz A en la expresión anterior.

Para las matrices, la ley cancelativa puede enunciarse de la siguiente manera

Si A es no singular entonces $AB = AC \Longrightarrow B = C$

como el lector podrá demostrar fácilmente.

Antes de concluir esta sección conviene señalar que hay ecuaciones matriciales, del tipo que hemos planteado aquí, las cuales no pueden resolverse empleando el procedimiento que hemos descrito y que, sin embargo, tienen solución. Para estos casos queda el recurso de plantear un sistema de ecuaciones lineales equivalente y resolverlo empleando el método de Gauss.

VI.5.1 EJERCICIOS

1.- Si definimos $A^2 = A\,A$, considere el siguiente desarrollo

$$\begin{aligned}(A+B)^2 - (2A+B)B &= (A+B)^2 - (2AB+B^2)\\ &= (A+B)^2 - 2AB - B^2\\ &= A^2 + 2AB + B^2 - 2AB - B^2\end{aligned}$$

$$(A+B)^2 - (2A+B)B = A^2$$

y compruebe la validez de la última expresión para las matrices

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

¿Hay algún error? Explique en que consiste.

2.- Obtener la matriz X, si existe, tal que:

a) $XAB = C + X$

si $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\qquad B = \begin{bmatrix} 1 & -1 \end{bmatrix}$ $\qquad C = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$

b) $XA + B = XC$

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

c) $AX + C = B$

si $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \\ -1 & -2 \end{bmatrix}$ $\qquad B = \begin{bmatrix} 2 & 4 \\ 3 & 0 \\ 5 & -1 \end{bmatrix}$ $\qquad C = \begin{bmatrix} -5 & 1 \\ 2 & -1 \\ 2 & 0 \end{bmatrix}$

d) $A + XB = XC$

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

3.- Para las matrices

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

y la ecuación $B(XA + B) = C - 3XA$

a) Obtener la expresión de X en términos de A, B y C

b) Obtener los elementos de la matriz X que resuelve la ecuación.

4.- Demostrar que si A es no singular, entonces:

i) $AB = AC \Longrightarrow B = C$
ii) $BA = CA \Longrightarrow B = C$
iii) $AB = CA \not\Longrightarrow B = C$

VI.6 TIPOS ESPECIALES DE MATRICES CUADRADAS

Las matrices cuadradas desempeñan un papel muy importante en la teoría de matrices, especialmente en lo que se refiere a sus aplicaciones. Es por ello que se establece cierta terminología especial para este tipo de matrices, de la cual nos ocuparemos en esta sección.

- Diagonal principal, triángulo superior y triángulo inferior.

En una matriz cuadrada pueden distinguirse tres "regiones":

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

triángulo superior

diagonal principal

triángulo inferior

i) La "diagonal principal", constituida por los elementos a_{ij} tales que $i = j$; es decir por los elementos de la forma a_{ii}.

Dichos elementos se encuentran ubicados en lo que geométricamente sería una de las diagonales del cuadrado formado por la matriz (la diagonal que va de izquierda a derecha y de arriba hacia abajo)

ii) El "triángulo superior", constituido por los elementos a_{ij} tales que $i < j$.

Estos elementos se encuentran situados "por arriba" de la diagonal principal.

iii) El "triángulo inferior", constituido por los elementos a_{ij} tales que $i > j$.

Estos elementos se encuentran situados "por debajo" de la diagonal principal.

Los tipos especiales de matrices cuadradas que veremos en esta sección se refieren a la naturaleza y disposición de los elementos de acuerdo con estas tres "regiones".

- Traza

Se conoce como traza de una matriz cuadrada al número que se obtiene sumando los elementos de su diagonal principal, como lo establece la siguiente definición

VI.6.1 DEFINICION

Sea $A = [a_{ij}]$ una matriz de $n\times n$ con elementos en C.

Se llama traza de A, y se representa con tr A, al número

$$\sum_{i=1}^{n} a_{ii}$$

Así, por ejemplo, para la matriz

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

se tiene que

$$\operatorname{tr} A = a_{11} + a_{22} + a_{33} + a_{44} = 2 + (-4i) + 0 + 5i = 2+i$$

De acuerdo con VI.6.1, la traza define una función del conjunto de matrices cuadradas con elementos en C en el conjunto de los núme - ros complejos. Dicha función tiene las propiedades que se enun - cian a continuación

> VI.6.2 TEOREMA
>
> Si A y B son dos matrices de n×n con elementos en C y $\alpha \in C$:
>
> i) $\operatorname{tr}(A+B) = (\operatorname{tr} A) + (\operatorname{tr} B)$
>
> ii) $\operatorname{tr}(\alpha A) = \alpha(\operatorname{tr} A)$
>
> iii) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

DEMOSTRACION

Se demuestran a continuación i) y ii) dejando al lector como ejer- cicio la demostración de iii).

Sean $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ y $B = \begin{bmatrix} b_{ij} \end{bmatrix}$ dos matrices de n×n con elementos en C y sea $\alpha \in C$:

i) $\operatorname{tr}(A+B) = \operatorname{tr}\begin{bmatrix} a_{ij} + b_{ij} \end{bmatrix}$ por VI.2.1

$= \sum_{i=1}^{n} (a_{ii} + b_{ii})$ por VI.6.1

$= \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii}$ por ii) y iii) de II.1.4

$\operatorname{tr}(A+B) = (\operatorname{tr} A) + (\operatorname{tr} B)$ por VI.6.1

ANEXO
BIBLIOTECA

ii) $\operatorname{tr}(\alpha A) = \operatorname{tr}\begin{bmatrix} \alpha a_{ij} \end{bmatrix}$ por VI.2.4

$= \sum_{i=1}^{n} (\alpha a_{ii})$ por VI.6.1

$= \alpha \sum_{i=1}^{n} a_{ii}$ por vi) de II.1.4

$\operatorname{tr}(\alpha A) = \alpha(\operatorname{tr} A)$ por VI.6.1 □

Matrices triangulares

> VI.6.3 DEFINICION
>
> Sea $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ una matriz de n×n con elementos en C. Se dice que:
>
> i) A es triangular superior si $a_{ij} = 0$ para $i > j$
>
> ii) A es triangular inferior si $a_{ij} = 0$ para $i < j$

Obsérvese que, de acuerdo con esta definición, en una matriz trian gular superior los elementos correspondientes al triángulo inferior son todos nulos. En consecuencia, en una matriz de este tipo sólo pueden hallarse elementos distintos de cero en el triángulo supe - rior y en la diagonal principal. Por ejemplo, las siguientes ma - trices son triangulares superiores

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{21} & a_{22} \\ 0 & 0 & a_{33} \end{bmatrix} \quad \begin{bmatrix} 2 & -3 & 0 & 1+i \\ 0 & -4i & 2 & 4 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 5i \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Por el contrario, en una matriz triangular inferior los elementos del triángulo superior deben ser nulos, como es el caso de las siguientes matrices

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \end{bmatrix} \quad \begin{bmatrix} -i & 0 & 0 & 0 \\ 2 & 5i & 0 & 0 \\ i & -3 & 7 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Con relación a las matrices triangulares, superiores e inferiores, se tiene el siguiente teorema

VI.6.4	TEOREMA
	Si A y B son dos matrices triangulares superiores (inferiores) del mismo orden y $\alpha \in C$, entonces:
	i) A+B es triangular superior (inferior)
	ii) αA es triangular superior (inferior)
	iii) AB es triangular superior (inferior)

DEMOSTRACION

Las propiedades i) y ii) son evidentes, por lo que omitiremos su demostración.

iii) Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices triangulares superiores de orden n. De VI.3.1 se sigue que

$$AB = [p_{ij}] \quad \text{donde} \quad p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Veamos que pasa con los sumandos de la expresión anterior cuando $i > j$, para todos los valores de $k = 1, \ldots, n$:

1°) si $k < i$, de i) de VI.6.3 $a_{ik} = 0$, por lo que $a_{ik}b_{kj} = 0$.

2°) si $k \geq i$, entonces $k > j$ y de i) de VI.6.3 $b_{kj} = 0$, por lo que también $a_{ik}b_{kj} = 0$.

En consecuencia, cuando $i > j$ $p_{ij} = 0$ y, de i) de VI.6.3, AB es triangular superior.

Si A y B son triangulares inferiores la prueba es similar. □

Matriz diagonal y matriz escalar

Una matriz que es triangular superior e inferior a la vez; esto es, una matriz cuyos elementos situados fuera de la diagonal principal son todos nulos, recibe el nombre de matriz diagonal.

Debido a su peculiar estructura, para este tipo de matrices suele emplearse una notación especial en la que se especifican sólo los elementos que integran la diagonal principal, ya que los demás son iguales a cero.

VI.6.5	DEFINICION
	Sea $A = [a_{ij}]$ una matriz de n×n con elementos en C. Se dice que A es una matriz diagonal si $a_{ij} = 0$ para $i \neq j$, y se representa con
	$\mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$

Así, por ejemplo, la matriz

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

es una matriz diagonal y se representa con

$$\mathrm{diag}(2,\ -4i,\ 0,\ 5i)$$

Los cálculos para efectuar operaciones con matrices se simplifican notablemente cuando se trata de matrices diagonales, especialmente la multiplicación y el cálculo de la inversa, como lo establece el siguiente teorema

> VI.6.6 TEOREMA
>
> Si A y B son dos matrices diagonales tales que
>
> $A = \mathrm{diag}(a_{11},\ a_{22}, \ldots,\ a_{nn})$ y $B = \mathrm{diag}(b_{11},\ b_{22}, \ldots,\ b_{nn})$ y $\alpha \in C$, entonces:
>
> i) $A+B = \mathrm{diag}(a_{11}+b_{11},\ a_{22}+b_{22}, \ldots,\ a_{nn}+b_{nn})$
>
> ii) $\alpha A = \mathrm{diag}(\alpha a_{11},\ \alpha a_{22}, \ldots,\ \alpha a_{nn})$
>
> iii) $AB = \mathrm{diag}(a_{11}b_{11},\ a_{22}b_{22}, \ldots,\ a_{nn}b_{nn})$
>
> iv) $A^{-1} = \mathrm{diag}(\frac{1}{a_{11}},\ \frac{1}{a_{22}}, \ldots,\ \frac{1}{a_{nn}})$, si A es no singular.

DEMOSTRACION

Se demuestran a continuación iii) y iv)

iii) Sean A y B dos matrices diagonales de orden n. De VI.3.1 se sigue que

$$AB = \left[\, p_{ij} \,\right], \text{ donde } p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

pero de VI.6.5 se tiene que $a_{ik} = 0$ si $i \neq k$ y $b_{kj} = 0$ si $k \neq j$, por lo que

$$p_{ij} = 0, \quad \text{si} \quad i \neq j$$

y

$$p_{ii} = a_{ii}b_{ii}$$

en consecuencia

$$AB = \mathrm{diag}(a_{11}b_{11},\ a_{22}b_{22}, \ldots,\ a_{nn}b_{nn})$$

iv) Sea $A = \left[a_{ij}\right]$ una matriz de $n\times n$ no singular, y sea $A^{-1} = \left[x_{ij}\right]$ la inversa de A. Entonces

$$A^{-1}A = I_n \qquad \text{por VI.4.1}$$

$$\left[\sum_{k=1}^{n} x_{ik}a_{kj}\right] = I_n \qquad \text{por VI.3.1}$$

$$\left[\sum_{k=1}^{n} x_{ik}a_{kj}\right] = \left[\delta_{ij}\right] \qquad \text{por VI.3.4}$$

$$\sum_{k=1}^{n} x_{ik}a_{kj} = \delta_{ij}; \quad \forall\ i,\ j \qquad \text{por VI.1.2}$$

$$x_{ij}a_{jj} = \delta_{ij}; \quad \forall\ i,\ j \qquad \text{por VI.6.5}$$

Entonces, si $i = j$ de VI.3.4 se tiene que

$$x_{ii}a_{ii} = 1$$

de donde

$$x_{ii} = \frac{1}{a_{ii}}, \text{ si } \quad a_{ii} \neq 0, \text{ lo cual se cumple puesto que } \exists\ A^{-1}$$

Además, si $i \neq j$ de VI.3.4 se tiene que

$$x_{ij}a_{jj} = 0$$

de donde

$$x_{ij} = 0, \text{ puesto que } a_{jj} \neq 0.$$

En consecuencia

$$A^{-1} = \operatorname{diag}\left(\frac{1}{a_{11}}, \frac{1}{a_{22}}, \ldots, \frac{1}{a_{nn}}\right)$$

y la prueba termina. □

Como consecuencia de la propiedad iii) del teorema anterior y de iii) de II.1.4, las matrices diagonales del mismo orden son permutables.

Un caso particular de matriz diagonal es aquel en que todos los elementos de la diagonal principal son iguales. A una matriz de este tipo se le conoce como "matriz escalar"; es decir, una matriz $A = [a_{ij}]$ de $n\times n$ con elementos en C se dice que es una matriz escalar si $a_{ij} = 0$ para $i \neq j$ y $a_{ii} = \alpha \ \forall i$, donde $\alpha \in C$.

Así pues, una matriz escalar es de la forma

$$\begin{bmatrix} \alpha & 0 & 0 & \cdots & 0 \\ 0 & \alpha & 0 & \cdots & 0 \\ 0 & 0 & \alpha & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \alpha \end{bmatrix}$$

y es claro que puede expresarse como αI, donde I es la matriz identidad de orden n; en consecuencia, premultiplicar una matriz M por una matriz escalar αI es equivalente a multiplicar por el escalar α. A esta propiedad, de la cual ya hicimos uso en la sección anterior, se debe el nombre de matriz escalar.

VI.6.7 EJERCICIOS

1.- a) Para las matrices

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

verificar que $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

b) Demostrar que si A y B son dos matrices de $n\times n$ con elementos en C, entonces

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

2.- a) Hallar dos matrices A y B tales que TAREA

$$\operatorname{tr}(AB) \neq (\operatorname{tr} A)(\operatorname{tr} B)$$

b) Para las matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ i & -1 & 0 \\ 2 & -1 & a_{33} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & 0 & 0 \\ 1 & 2 & 0 \\ 0 & -1 & -i \end{bmatrix}$$

obtener una pareja de valores (a_{33}, b_{11}) tales que

$$\operatorname{tr}(AB) = (\operatorname{tr} A)(\operatorname{tr} B)$$

3.- a) Si A es una matriz de 2x2 con elementos en C, demostrar que existen dos matrices, una triangular inferior (X) y otra triangular superior (Y), tales que $A = XY$

b) Para
$$A = \begin{bmatrix} 3 & -1 \\ 2 & 4 \end{bmatrix}$$
obtener dos matrices X (triangular inferior) y Y (triangular superior) tales que
$$A = XY$$

4.- Para la matriz triangular superior
$$T = \begin{bmatrix} 3 & -1 & 4 \\ 0 & 2 & -3 \\ 0 & 0 & -\frac{1}{3} \end{bmatrix}$$
obtener T^{-1}

¿En general, si T es una matriz triangular superior de nxn, su inversa es triangular superior? ¿Por qué?

5.- Para las matrices
$$A = \begin{bmatrix} 1 & 1 & -1 \\ 3 & -1 & -4 \\ -1 & 0 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 2 & 5 \\ 2 & -1 & -1 \\ 1 & 1 & 4 \end{bmatrix}$$
$$C = \begin{bmatrix} 2 & k & -1 \\ 3 & -1 & 2 \\ 1 & -1 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 1 & 1-2k \\ 1 & 1 & 7 \\ -1 & k+2 & 2+3k \end{bmatrix}$$

a) Obtener AB

b) Determinar qué condiciones debe cumplir k para que exista $(CD)^{-1}$

c) Obtener $(AB)(CD)^{-1}$

BIBLIOTECA NACIONAL UNIVERSIDAD DE MEXICO
FACULTAD DE INGENIERIA

VI.7 OPERACIONES SOBRE UNA MATRIZ

Además de las operaciones como la adición y la multiplicación existen "operaciones" de otro tipo, las cuales se efectúan sobre una sola matriz transformándola, generalmente, en otra matriz diferente.

De estas operaciones, que hemos agrupado bajo el título de "operaciones sobre una matriz", nos ocuparemos en esta sección; así como de algunos tipos especiales de matrices definidos en términos de dichas operaciones.

Transposición

La transposición es una operación que transforma una matriz en otra, llamada su transpuesta, cuyos renglones son las columnas de la matriz original y cuyas columnas son los renglones de la matriz original. Al respecto se tiene la siguiente definición

VI.7.1 DEFINICION

Sea $A = \left[a_{ij} \right]$ una matriz de mxn con elementos en C. Se llama transpuesta de A a la matriz de nxm $A^T = \left[c_{ij} \right]$ tal que

$$c_{ij} = a_{ji}$$

G. 903246

De acuerdo con esta definición, el elemento correspondiente al renglón i y columna j de A^T es el que se encuentra en el renglón j y columna i de la matriz A. Así, los renglones de A^T son las columnas de A y las columnas de A^T son los renglones de A.

Por ejemplo, para la matriz

$$A = \begin{bmatrix} 2i & 0 & -i \\ 5 & 1 & 1-3i \end{bmatrix}$$

se tiene que

$$A^T = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = \begin{bmatrix} 2i & 5 \\ 0 & 1 \\ -i & 1-3i \end{bmatrix}$$

Las principales propiedades de la transposición se presentan en el siguiente teorema

> VI.7.2 TEOREMA
>
> Si A y B son dos matrices con elementos en C y $\alpha \in C$, entonces:
>
> i) $(A^T)^T = A$
>
> ii) $(\alpha A)^T = \alpha A^T$
>
> iii) $(A+B)^T = A^T + B^T$, si A + B puede obtenerse
>
> iv) $(AB)^T = B^T A^T$, si AB puede obtenerse

DEMOSTRACION

Las propiedades i), ii) y iii) son evidentes, por lo que omitimos su demostración sugiriéndola al lector como ejercicio.

iv) Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices con elementos en C, de m×n y n×q respectivamente; y sean $A^T = [c_{ij}]$ y $B^T = [d_{ij}]$ sus respectivas transpuestas. Entonces, de

INGENIERIA
ANEXO
BIBLIOTECA

VI.3.1

$$AB = [p_{ij}], \quad \text{donde} \quad p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

y en consecuencia

$$(AB)^T = [p_{ji}] = \left[\sum_{k=1}^{n} a_{jk}b_{ki}\right] \qquad \text{por VI.7.1}$$

$$= \left[\sum_{k=1}^{n} b_{ki}a_{jk}\right] \qquad \text{por iii) de II.1.4}$$

$$= \left[\sum_{k=1}^{n} d_{ik}c_{kj}\right] \qquad \text{por VI.7.1}$$

$$(AB)^T = B^T A^T \qquad \text{por VI.3.1}$$

Matrices simétricas y antisimétricas.

La transposición da lugar a la definición de dos tipos especiales de matrices cuadradas, como se establece a continuación.

> VI.7.3 DEFINICION
>
> Sea A una matriz de n×n con elementos en C. Se dice que:
>
> i) A es simétrica si $A^T = A$
>
> ii) A es antisimétrica si $A^T = -A$

Veamos ahora que características tienen los elementos de una matriz simétrica y de una antisimétrica.

De VI.7.3 A es simétrica si

$$A = A^T$$

en consecuencia, de VI.7.1 A es simétrica si

$$\left[a_{ij}\right] = \left[a_{ji}\right]$$

esto es, si

$$a_{ij} = a_{ji} \quad \forall\, i,j$$

Es decir que los elementos "simétricos" con respecto a la diagonal principal son iguales.

Por ejemplo, la siguiente matriz es simétrica

$$A = \begin{bmatrix} -1 & 5 & 2-i \\ 5 & 3i & -i \\ 2-i & -i & 0 \end{bmatrix} \quad \text{ya que} \quad \begin{cases} a_{12} = a_{21} = 5 \\ a_{13} = a_{31} = 2-i \\ a_{23} = a_{32} = -i \end{cases}$$

De manera similar, para las matrices antisimétricas se tiene

$$A = -A^T$$

$$\left[a_{ij}\right] = \left[-a_{ji}\right]$$

$$\therefore \quad a_{ij} = -a_{ji} \quad \forall\, i,j$$

Es decir que los elementos "simétricos" con respecto a la diagonal principal deben ser uno el negativo del otro. Además, de la expresión anterior se tiene, para $i = j$, que

$$a_{ii} = -a_{ii} \Rightarrow a_{ii} = 0$$

por lo que los elementos de la diagonal principal deben ser nulos.

La siguiente matriz, por ejemplo, es una matriz antisimétrica

$$A = \begin{bmatrix} 0 & -5 & -2+i \\ 5 & 0 & i \\ 2-i & -i & 0 \end{bmatrix} \quad \text{ya que} \quad \begin{cases} a_{12} = -a_{21} = -5 \\ a_{13} = -a_{31} = -2+i \\ a_{23} = -a_{32} = i \\ a_{11} = a_{22} = a_{33} = 0 \end{cases}$$

Las matrices simétricas y antisimétricas tienen, entre otras, las propiedades que se enuncian en los dos siguientes teoremas.

VI.7.4 TEOREMA

Si A y B son dos matrices simétricas (antisimétricas) de $n \times n$ y $\alpha \in C$, entonces:

i) A+B es simétrica (antisimétrica)

ii) αA es simétrica (antisimétrica)

DEMOSTRACION

Se demuestra a continuación únicamente la parte i) para el caso de matrices simétricas.

Sean A y B dos matrices del mismo orden. Entonces, por iii) de VI.7.2

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

Si A y B son simétricas, de VI.7.3 $A^T = A$ y $B^T = B$, por lo que

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

En consecuencia, de VI.7.3 A+B es simétrica y la prueba termina.

□

VI.7.5 TEOREMA

Si A es una matriz de n×n con elementos en C, entonces:

i) $A+A^T$ es simétrica

ii) $A-A^T$ es antisimétrica

DEMOSTRACION

i) Si A es una matriz de n×n, por VI.7.1 A^T es también de n×n y por iii) de VI.7.2

$$(A+A^T)^T = A^T+(A^T)^T$$

en consecuencia, por i) de VI.7.2 y por ii) de VI.2.2

$$(A+A^T)^T = A^T+A = A+A^T$$

por lo que, de VI.7.3, $A+A^T$ es simétrica.

La prueba de ii) es similar. □

Conjugación

La conjugación transforma una matriz en otra, llamada su conjugada, cuyos elementos son los conjugados de los elementos correspondientes en la matriz original, como lo establece la siguiente definición.

VI.7.6 DEFINICION

Sea $A = \left[a_{ij}\right]$ una matriz de m×n con elementos en C. Se llama conjugada de A a la matriz de m×n $\bar{A} = \left[c_{ij}\right]$ tal que

$$c_{ij} = \bar{a}_{ij}$$

Así, por ejemplo para la matriz

$$A = \begin{bmatrix} 2i & 0 & -i \\ 5 & 1 & 1-3i \end{bmatrix}$$

se tiene que

$$\bar{A} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} \\ \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} \end{bmatrix} = \begin{bmatrix} -2i & 0 & i \\ 5 & 1 & 1+3i \end{bmatrix}$$

Las principales propiedades de la conjugación son las siguientes

VI.7.7 TEOREMA

Si A y B son dos matrices con elementos en C y $\alpha \in C$, entonces:

i) $(\bar{\bar{A}}) = A$

ii) $\overline{\alpha A} = \bar{\alpha}\,\bar{A}$

iii) $\overline{A+B} = \bar{A} + \bar{B}$, si A+B puede obtenerse

iv) $\overline{AB} = \bar{A}\,\bar{B}$, si AB puede obtenerse

DEMOSTRACION

Se demuestra únicamente la propiedad iv).

Sean $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices con elementos en C, de m×n y n×q respectivamente; y sean $\bar{A} = \left[c_{ij}\right]$ y $\bar{B} = \left[d_{ij}\right]$ sus respectivas conjugadas. Entonces de VI.3.1

$$AB = \left[\sum_{k=1}^{n} a_{ik}b_{kj}\right]$$

y en consecuencia

$$\begin{aligned}
\overline{AB} &= \left[\overline{\sum_{k=1}^{n} a_{ik}b_{kj}}\right] && \text{por VI.7.6} \\
&= \left[\sum_{k=1}^{n} \overline{a_{ik}b_{kj}}\right] && \text{por v) de II.1.6} \\
&= \left[\sum_{k=1}^{n} \bar{a}_{ik}\bar{b}_{kj}\right] && \text{por vi) de II.1.6} \\
&= \left[\sum_{k=1}^{n} c_{ik}d_{kj}\right] && \text{por VI.7.6} \\
\overline{AB} &= \bar{A}\,\bar{B} && \text{por VI.3.1}
\end{aligned}$$

□

Matrices reales e imaginarias

La conjugación también da lugar a dos tipos especiales de matrices, de acuerdo con la siguiente definición

> VI.7.8 DEFINICION
>
> Sea A una matriz de m×n con elementos en C. Se dice que:
>
> i) A es real si $\bar{A} = A$
>
> ii) A es imaginaria si $\bar{A} = -A$

ANEXO BIBLIOTECA

Los elementos de una matriz real (en el sentido que establece VI.7.8) son, en efecto, números reales; ya que

$$a_{ij} = \bar{a}_{ij} \implies I(a_{ij}) = 0$$

Para una matriz imaginaria se tiene, de acuerdo con VI.7.8, que

$$a_{ij} = -\bar{a}_{ij} \implies R(a_{ij}) = 0$$

por lo que sus elementos son números imaginarios.

Las matrices reales e imaginarias tienen las propiedades enunciadas en los dos siguientes teoremas, cuya demostración se deja al lector.

> VI.7.9 TEOREMA
>
> Si A y B son dos matrices reales (imaginarias), entonces:
>
> i) A+B es real (imaginaria), si A+B puede obtenerse
>
> ii) AB es real (real), si AB puede obtenerse

> VI.7.10 TEOREMA
>
> Si A es una matriz de m×n con elementos en C, entonces:
>
> i) $A+\bar{A}$ es real
>
> ii) $A-\bar{A}$ es imaginaria

- Conjugación-transposición.

Se conoce como conjugación-transposición a la aplicación sucesiva de las dos operaciones definidas anteriormente. A la matriz que se obtiene se le llama conjugada-transpuesta de la matriz original, como lo indica la siguiente definición

VI.7.11 DEFINICION

Sea A una matriz de m×n con elementos en C. Se llama conjugada-transpuesta de A, y se representa con A*, a la matriz de n×m definida por

$$A^* = (\overline{A})^T$$

El orden en que se efectúen las operaciones de transposición y conjugación es indiferente, como lo señala el siguiente teorema

VI.7.12 TEOREMA

Si A es una matriz de m×n con elementos en C, entonces:

$$A^* = (\overline{A})^T = \overline{(A^T)}$$

cuya demostración se deja al lector.

A manera de ejemplo, consideremos nuevamente la matriz

$$A = \begin{bmatrix} 2i & 0 & -i \\ 5 & 1 & 1-3i \end{bmatrix}$$

Al efectuar la conjugación se obtiene

$$\overline{A} = \begin{bmatrix} -2i & 0 & i \\ 5 & 1 & 1+3i \end{bmatrix}$$

y al transponer esta última matriz se tiene

$$A^* = (\overline{A})^T = \begin{bmatrix} -2i & 5 \\ 0 & 1 \\ i & 1+3i \end{bmatrix}$$

Esta misma matriz se habría obtenido transponiendo primero A y conjugando después A^T.

La conjugación-transposición satisface las siguientes propiedades

VI.7.13 TEOREMA

Si A y B son dos matrices con elementos en C y $\alpha \in C$, entonces:

i) $(A^*)^* = A$

ii) $(\alpha A)^* = \overline{\alpha}\, A^*$

iii) $(A+B)^* = A^* + B^*$, si A+B puede obtenerse

iv) $(AB)^* = B^* A^*$, si AB puede obtenerse

DEMOSTRACION

Se demuestran a continuación ii) y iv) únicamente:

ii) $(\alpha A)^* = (\overline{\alpha A})^T$ por VI.7.11

$= (\overline{\alpha}\ \overline{A})^T$ por ii) de VI.7.7

$= \overline{\alpha}\ (\overline{A})^T$ por ii) de VI.7.2

$(\alpha A)^* = \overline{\alpha}\ A^*$ por VI.7.11

iv) $(AB)^* = (\overline{AB})^T$ por VI.7.11

$(AB)^* = (\overline{A}\ \overline{B})^T$ por iv) de VI.7.7

$(AB)^* = (\overline{B})^T (\overline{A})^T$ por iv) de VI.7.2

$(AB)^* = B^* A^*$ por VI.7.11

Matrices hermitianas y antihermitianas

A partir de la conjugación-transposición se definen otros dos tipos especiales de matrices cuadradas, como se establece a continuación.

> VI.7.14 DEFINICION
>
> Sea A una matriz de n×n con elementos en C. Se dice que:
>
> i) A es hermitiana si $A^* = A$
>
> ii) A es antihermitiana si $A^* = -A$

De la definición anterior se sigue que los elementos de una matriz hermitiana deben ser tales que

$$a_{ij} = \overline{a}_{ji} \qquad \forall\ i,j$$

Es decir que los elementos "simétricos" con respecto a la diagonal principal deben ser conjugados. Además, para i = j se tiene que

$$a_{ii} = \overline{a}_{ii} \quad \Longrightarrow \quad I(a_{ii}) = 0$$

por lo que los elementos de la diagonal principal deben ser números reales.

Así, por ejemplo, la siguiente matriz es hermitiana

$$A = \begin{bmatrix} -1 & 5 & 2+i \\ 5 & 3 & i \\ 2-i & -i & 0 \end{bmatrix} \qquad \text{ya que} \quad \begin{cases} a_{12} = \overline{a}_{21} = 5 \\ a_{13} = \overline{a}_{31} = 2+i \\ a_{23} = \overline{a}_{32} = i \\ I(a_{11}) = I(a_{22}) = I(a_{33}) = 0 \end{cases}$$

Para los elementos de una matriz antihermitiana se tiene que

$$a_{ij} = -\overline{a}_{ji} \qquad \forall\ i,j$$

Es decir que los elementos "simétricos" con respecto a la diagonal principal deben ser tales que sus partes reales sólo difieran en el signo y su partes imaginarias sean iguales. Además

$$a_{ii} = -\overline{a}_{ii} \quad \Longrightarrow \quad R(a_{ii}) = 0$$

por lo que los elementos de la diagonal principal deben ser números imaginarios, como en el caso de la siguiente matriz que es antihermitiana

$$A = \begin{bmatrix} -i & -5 & -2-i \\ 5 & 3i & -i \\ 2-i & -i & 0 \end{bmatrix} \qquad \text{ya que} \quad \begin{cases} a_{12} = -\overline{a}_{21} = -5 \\ a_{13} = -\overline{a}_{31} = -2-i \\ a_{23} = -\overline{a}_{32} = -i \\ R(a_{11}) = R(a_{22}) = R(a_{33}) = 0 \end{cases}$$

A continuación se enuncian algunas propiedades relacionadas con las matrices hermitianas y antihermitianas

> VI.7.15 TEOREMA
>
> Si A y B son dos matrices hermitianas (antihermitianas) de n×n, entonces A+B es hermitiana (antihermitiana)

DEMOSTRACION

Se demuestra a continuación el enunciado para el caso de matrices antihermitianas.

Sean A y B dos matrices del mismo orden. Entonces, por iii) de VI.7.13

$$(A+B)^* = A^* + B^*$$

Si A y B son antihermitianas, de VI.7.14 $A^* = -A$ y $B^* = -B$, por lo que

$$(A+B)^* = -A + (-B) = -(A+B)$$

En consecuencia, de VI.7.14 A+B es antihermitiana. □

VI.7.16 TEOREMA

Si A es una matriz de m×n con elementos en C, entonces:

i) $A\,A^*$ es hermitiana

ii) $A^*\,A$ es hermitiana

iii) $A+A^*$ es hermitiana, si A es cuadrada

iv) $A-A^*$ es antihermitiana, si A es cuadrada

La demostración de este teorema se deja al lector como ejercicio.

- Potencia enésima

De manera similar al caso de los números, se conoce como potencia enésima de una matriz cuadrada al producto

$$\underbrace{A\,A\,\ldots\,A}_{n\ \text{factores}}$$

cuya definición formal es la siguiente

VI.7.17 DEFINICION

Sea A una matriz de m×m con elementos en C y sea n ε N. Se llama potencia enésima de A, y se representa con A^n, a la matriz definida por

$$A^0 = I_m$$

$$A^n = A\,A^{n-1}, \text{ para } n \geqslant 1$$

Como se ve, la definición anterior es recurrente; por ejemplo, para la tercera potencia de A se tiene, según VI.7.17, que

$$A^3 = A\,A^2$$

Aplicando nuevamente VI.7.17 se tiene que $A^2 = A\,A^1$, por lo que

$$A^3 = A(A\,A^1)$$

y aplicando una vez más VI.7.17 se tiene que $A^1 = A\,A^0 = AI = A$; por lo que, finalmente

$$A^3 = A(A\,A)$$

Este resultado puede también expresarse como

$$A^3 = A\,A\,A$$

debido a la asociatividad de la multiplicación de matrices.

La potencia enésima de una matriz, así definida, satisface las siguientes propiedades

VI.7.18 TEOREMA

Si A es una matriz cuadrada con elementos en C y $m,n \in N$, entonces:

i) $A^m A^n = A^{m+n}$

ii) $(A^m)^n = A^{mn}$

DEMOSTRACION (por inducción)

Demostraremos primero que

$$A^m A = A A^m, \ \forall\, m \in N \qquad ---(1)$$

En efecto, para m = 1 la proposición establece que

$$A^1 A = A A^1$$

pero, por VI.7.17 $A^1 = A A^0 = AI = A$, y se tiene la expresión

$$A A = A A$$

por lo que la proposición es válida para m = 1.

Suponemos entonces que para algún k

$$A^k A = A A^k$$

y premultiplicando por A se tiene

$$A(A^k A) = A(A A^k)$$

en consecuencia

$(A A^k)A = A(A A^k)$ por VI.3.2

$A^{k+1} A = A A^{k+1}$ por VI.7.17

con lo que se demuestra el enunciado.

i) Sea ahora m un número natural arbitrario. Por (1) y por VI.7.17 se tiene que

$$A^m A^1 = A^1 A^m = A^{m+1}$$

y la proposición se verifica para n = 1.

Supongamos ahora que

$$A^m A^k = A^{m+k}$$

INGENIERIA
ANEXO
BIBLIOTECA

premultiplicando por A se tiene

$$A(A^m A^k) = A A^{m+k}$$

y en consecuencia

$(A A^m)A^k = A A^{m+k}$ por VI.3.2

$(A^m A)A^k = A A^{m+k}$ por (1)

$A^m(A A^k) = A A^{m+k}$ por VI.3.2

$A^m A^{k+1} = A^{m+k+1}$ por VI.7.17

Con lo que se demuestra la propiedad i). La demostración de ii) es similar y se deja al lector. $\square$

A partir de la definición de potencia enésima se establecen los siguientes tipos especiales de matrices cuadradas.

VI.7.19 DEFINICION

Sea A una matriz de m×m con elementos en C. Se dice que A es:

i) Idempotente si $A^2 = A$

ii) Involutoria si $A^2 = I$

iii) Nilpotente (de índice n) si n es el menor número natural tal que $A^n = 0$

iv) Periódica (de período n) si n es el menor número natural distinto de uno tal que $A^n = A$

Obsérvese que una matriz idempotente es un caso particular de matriz periódica (de período dos).

VI.7.20 TEOREMA

Sea A una matriz de m×m con elementos en C:

i) Si A es idempotente entonces

$$A^n = A, \quad \forall\, n \in N$$

ii) Si A es involutoria entonces

$$A^{2n} = I$$

$$A^{2n+1} = A, \quad \forall\, n \in N$$

Se deja al lector la demostración de este teorema.

VI.7.21 EJERCICIOS TAREA

1.- Para las matrices

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

verificar que se cumplen las siguientes propiedades

a) $(A+B)^T = A^T + B^T$

b) $(AB)^T = B^T A^T$

2.- Demostrar que:

a) Si A y B son matrices de m×n con elementos en C entonces:

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

b) Si A es no singular entonces

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

3.- a) Construir dos matrices A y B de 3×3 que sean antisimétricas y verificar que A+B es antisimétrica.

b) Construir una matriz A de 4×4 que sea simétrica y verificar que αA es simétrica para cualquier $\alpha \in C$.

4.- Demostrar que si A y B son matrices simétricas del mismo orden, AB es simétrica si y sólo si A y B son permutables.

5.- a) Demostrar que toda matriz cuadrada M con elementos en C puede expresarse como

$$M = S+A$$

donde S es una matriz simétrica y A es una matriz antisimétrica.

b) Ilustrar el enunciado anterior con

$$M = \begin{bmatrix} i & 1+i & 3 \\ -1 & 0 & 2i \\ 5 & -2 & -3+i \end{bmatrix}$$

6.- Para las matrices

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

verificar que:

a) $\overline{AB} = \overline{A}\ \overline{B}$

b) $\overline{(A^{-1})} = (\overline{A})^{-1}$

7.- Demostrar que si A y B son dos matrices imaginarias de orden $m\times n$ y $n\times q$, respectivamente, entonces AB es real.

8.- Demostrar que si A es una matriz de $m\times n$ con elementos en C, entonces:

$$(\overline{A})^T = \overline{(A^T)}$$

9.- Para las matrices

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

verificar que:

a) $(\alpha A)^* = \overline{\alpha}\, A^* \quad \forall\, \alpha \in C$

b) $(AB)^* = B^*\, A^*$

10.- a) Construir dos matrices hermitianas de orden 3 y verificar que su suma es hermitiana.

b) Construir una matriz A antihermitiana y verificar que αA es antihermitiana si α es real, y que αA es hermitiana si α es imaginario.

11.- Demostrar que si A y B son matrices antihermitianas, AB es hermitiana si y sólo si A y B son permutables.

12.- a) Demostrar que si M es una matriz hermitiana con elementos en C, entonces puede expresarse como

$$M = S + iA$$

donde S es real simétrica y A es real antisimétrica.

b) Ilustrar el enunciado anterior para

$$M = \begin{bmatrix} 1 & 2 & 1-i \\ 2 & -3 & 2i \\ 1+i & -2i & 5 \end{bmatrix}$$

13.- Una matriz A no singular se dice que:

i) es ortogonal si $A^T = A^{-1}$

ii) es unitaria si $A^* = A^{-1}$

a) Determinar bajo qué condiciones el producto de dos matrices ortogonales es ortogonal.

b) Demostrar que la conjugada de una matriz unitaria es unitaria.

14.- Demostrar que si A y B son dos matrices con elementos en C y $m,n \in N$, entonces:

a) $(A^m)^n = A^{mn}$

b) $(AB)^n = A^n B^n$ si y sólo si A y B son permutables

15.- Para las matrices

$$B = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$$

a) Verificar que B es idempotente y C es involutoria

b) Tomando en cuenta el resultado anterior obtener una matriz X, si existe, tal que

$$X = B^2 B^{-1} + C^2 B$$

y verificar el resultado.

c) En general, ¿qué podemos decir de A^{-1} si A es idempotente? Demostrarlo.

16.- Demostrar que una matriz triangular superior de orden 3 tal que $a_{ii} = 0 \ \forall i$, es nilpotente de índice 3.

VI.8 PARTICION DE MATRICES

En ciertos casos puede ser útil "subdividir" las matrices con objeto de simplificar algunos cálculos o para cambiar la presentación de un problema. Surge así el concepto de partición de matrices, el cual presentamos en esta sección; sin embargo, es conveniente introducir antes los conceptos de submatriz e hipermatriz para facilitar la comprensión del concepto de partición.

Submatriz e hipermatriz

Si A es una matriz de m×n con elementos en C, se llama "submatriz de A" a cualquier matriz que pueda obtenerse a partir de A suprimiendo en ésta algunos renglones o columnas.

Por ejemplo, si en la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

suprimimos el tercer renglón y la segunda y cuarta columnas se obtiene la matriz

$$\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}$$

que es una submatriz de A.

A continuación se presentan, a manera de ejemplo, algunas otras submatrices de A

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} \quad \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \quad \begin{bmatrix} a_{24} \end{bmatrix}$$

Se conoce como "hipermatriz" a un arreglo rectangular de matrices; es decir, a una especie de matriz cuyos elementos son matrices.

Por ejemplo, las matrices

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \qquad C = \begin{bmatrix} c_{11} & c_{12} \end{bmatrix} \quad y \quad D = \begin{bmatrix} d_{11} \\ d_{21} \end{bmatrix}.$$

pueden ser presentadas en un arreglo, de la siguiente manera

$$H = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

con lo que se obtiene una hipermatriz. Esta hipermatriz también puede expresarse en términos de los elementos de A, B, C y D como sigue

$$H = \left[\begin{array}{cc|c} a_{11} & a_{12} & b_{11} \\ a_{21} & a_{22} & b_{21} \\ \cline{3-3} a_{31} & a_{32} & d_{11} \\ \hline c_{11} & c_{12} & d_{21} \end{array}\right]$$

- Partición

Consideremos ahora la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{bmatrix}$$

y expresémosla como una hipermatriz, agrupando sus elementos en submatrices de la siguiente manera

$$\left[\begin{array}{cc|c|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{array}\right]$$

Hemos formado con ello la hipermatriz

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

donde

$$A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad A_{12} = \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \qquad A_{13} = \begin{bmatrix} a_{14} \\ a_{24} \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \\ a_{51} & a_{52} \end{bmatrix} \qquad A_{22} = \begin{bmatrix} a_{33} \\ a_{43} \\ a_{53} \end{bmatrix} \qquad A_{23} = \begin{bmatrix} a_{34} \\ a_{44} \\ a_{54} \end{bmatrix}$$

Dicha hipermatriz se dice que es una "partición" de la matriz A.

Otra partición de la misma matriz A puede ser la siguiente.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ \hdashline a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{bmatrix} = \begin{bmatrix} A_{11} \\ \\ A_{21} \end{bmatrix}$$

donde se tienen ahora sólo dos submatrices.

En términos generales, puede decirse que una partición de una ma - triz es la expresión de ésta como una hipermatriz, mediante la agrupación de sus elementos en submatrices; sin embargo, no cual - quier clase de hipermatriz se considera como una partición. Sólo se aceptan como tales aquellas en que:

1) Todas las submatrices que integran un mismo renglón (de la hipermatriz) tienen el mismo número de renglones, y

2) Todas las submatrices que integran una misma columna tie nen el mismo número de columnas.

Así, una hipermatriz como

$$H = \left[\begin{array}{cc|c} a_{11} & a_{12} & b_{11} \\ a_{21} & a_{22} & b_{21} \\ \cline{3-3} a_{31} & a_{32} & d_{11} \\ \cline{1-2} c_{11} & c_{12} & d_{21} \end{array}\right] = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

no es una partición de la matriz

$$M = \begin{bmatrix} a_{11} & a_{12} & b_{11} \\ a_{21} & a_{22} & b_{21} \\ a_{31} & a_{32} & d_{11} \\ c_{11} & c_{12} & d_{21} \end{bmatrix}$$

puesto que las submatrices A y B no tienen el mismo número de renglones (como tampoco lo tienen las submatrices C y D).

En consecuencia, toda partición de una matriz de m×n deberá ser de la forma

$$\begin{array}{r} \\ m_1 \text{ renglones } \Big\{ \\ m_2 \text{ renglones } \Big\{ \\ \\ m_s \text{ renglones } \Big\{ \end{array} \begin{array}{c} \begin{array}{cccc} n_1 \text{ cols.} & n_2 \text{ cols.} & & n_t \text{ cols.} \end{array} \\ \left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1t} \\ \hdashline A_{21} & A_{22} & \cdots & A_{2t} \\ \hdashline \vdots & \vdots & & \vdots \\ \hdashline A_{s1} & A_{s2} & \cdots & A_{st} \end{array}\right] \end{array}$$

donde:

$$m_1+m_2+\ldots+m_s = m \quad \text{y} \quad n_1+n_2+\ldots+n_t = n$$

De acuerdo con esto, se tiene la siguiente definición formal para el concepto de partición.

VI.8.1 DEFINICION

Sea $A = \left[a_{ij} \right]$ una matriz de m×n con elementos en C, y sean $(m_1, m_2, \ldots, m_s)$ números naturales tales que $m_1 + m_2 + \ldots + m_s = m$, y $(n_1, n_2, \ldots, n_t)$ números naturales tales que $n_1 + n_2 + \ldots + n_t = n$.

Se conoce como partición de A inducida por

$(m_1, m_2, \ldots, m_s)$ y $(n_1, n_2, \ldots, n_t)$ al arreglo

$$\begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1t} \\ A_{21} & A_{22} & \ldots & A_{2t} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ A_{s1} & A_{s2} & \ldots & A_{st} \end{bmatrix}$$

donde A_{rc} $(r = 1, 2, \ldots, s$ y $c = 1, 2, \ldots, t)$ es una matriz de $m_r \times n_c$ definida por

$$A_{rc} = \left[a_{ij} \right] \text{ con } \begin{cases} i = \left[m_1 + \ldots + (m_{r-1} + 1) \right], \ldots, \left[m_1 + \ldots + m_r \right] \\ j = \left[n_1 + \ldots + (n_{c-1} + 1) \right], \ldots, \left[n_1 + \ldots + n_c \right] \end{cases}$$

Así, para la primera partición que presentamos:

$$\left[\begin{array}{cc|c|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

se tiene que

$$m_1 = 2, \; m_2 = 3 \quad \text{y} \quad n_1 = 2, \; n_2 = 1, \; n_3 = 1$$

por lo que se trata de la partición de A inducida por los números (2,3) y (2,1,1). En dicha partición la matriz A_{21}, por ejemplo, es una matriz de $m_2 \times n_1$ (es decir de 3×2) definida por

$$A_{21} = \left[a_{ij} \right] \text{ con } \begin{cases} i = 3, 4, 5 \\ j = 1, 2 \end{cases}$$

A N E X O

Operaciones con matrices por partición

Pasemos ahora a ocuparnos de cómo efectuar operaciones con matrices empleando el concepto de partición.

Una vez que se han determinado las particiones, con las hipermatrices obtenidas pueden efectuarse operaciones como si se tratara de matrices cuyos elementos son las submatrices correspondientes; siempre que éstas últimas sean conformables para todas las operaciones requeridas.

Veamos primero unos ejemplos relativos a la adición y a la multiplicación por un escalar, y para ello consideremos las matrices

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

Si establecemos las siguientes particiones para A y B

$$\left[\begin{array}{ccc|c} 3 & 1 & -5 & -2 \\ -1 & 4 & 0 & 2 \\ \hline 0 & -2 & 0 & 7 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 2 & -1 & 3 & 4 \\ 4 & -2 & 1 & 3 \\ \hline -3 & 1 & 6 & -3 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

y sumamos las hipermatrices como si éstas fueran matrices, se obtiene

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix}$$

donde

$$A_{11} + B_{11} = \begin{bmatrix} 3 & 1 & -5 \\ -1 & 4 & 0 \end{bmatrix} + \begin{bmatrix} 2 & -1 & 3 \\ 4 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 & -2 \\ 3 & 2 & 1 \end{bmatrix}$$

$$A_{12} + B_{12} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$A_{21} + B_{21} = \begin{bmatrix} 0 & -2 & 0 \end{bmatrix} + \begin{bmatrix} -3 & 1 & 6 \end{bmatrix} = \begin{bmatrix} -3 & -1 & 6 \end{bmatrix}$$

$$A_{22} + B_{22} = \begin{bmatrix} 7 \end{bmatrix} + \begin{bmatrix} -3 \end{bmatrix} = \begin{bmatrix} 4 \end{bmatrix}$$

por lo que se llega a la hipermatriz

$$\left[\begin{array}{ccc|c} 5 & 0 & -2 & 2 \\ 3 & 2 & 1 & 5 \\ \hline -3 & -1 & 6 & 4 \end{array}\right]$$

la cual constituye una partición de A + B. Es decir

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

como pude verificarse fácilmente sumando las matrices A y B.

Consideremos ahora el escalar $-i$. Si lo multiplicamos por la hipermatriz

$$\left[\begin{array}{ccc|c} 3 & 1 & -5 & -2 \\ -1 & 4 & 0 & 2 \\ \hline 0 & -2 & 0 & 7 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

como si ésta fuera una matriz, se obtiene

$$-i \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} -i\,A_{11} & -i\,A_{12} \\ -i\,A_{21} & -i\,A_{22} \end{bmatrix} = \begin{bmatrix} -3i & -i & 5i & 2i \\ i & -4i & 0 & -2i \\ 0 & 2i & 0 & -7i \end{bmatrix}$$

que es una partición de $-i$ A.

Estos resultados pueden generalizarse mediante el siguiente teorema, cuya demostración se deja al lector.

VI.8.2 TEOREMA

Sean A y B dos matrices de m×n con elementos en C y sea $\alpha \in C$:

i) Si $\left[A_{ij}\right]$ y $\left[B_{ij}\right]$ son las particiones de A y B, respectivamente, inducidas por

$(m_1, m_2, \ldots, m_s)$ y $(n_1, n_2, \ldots, n_t)$, entonces

$$\left[S_{ij}\right] \quad \text{con} \quad S_{ij} = A_{ij} + B_{ij}$$

es la partición de A + B inducida por

$$(m_1, m_2, \ldots, m_s) \;\; \text{y} \;\; (n_1, n_2, \ldots, n_t)$$

ii) Si $\left[A_{ij}\right]$ es la partición de A inducida por

$(m_1, m_2, \ldots, m_s)$ y $(n_1, n_2, \ldots, n_t)$, entonces

$$\left[E_{ij}\right] \quad \text{con} \quad E_{ij} = \alpha A_{ij}$$

es la partición de αA inducida por

$$(m_1, m_2, \ldots, m_s) \;\; \text{y} \;\; (n_1, n_2, \ldots, n_t)$$

Menos evidente que los dos casos anteriores, pero de mayor utilidad práctica, resulta el caso de la multiplicación de matrices por partición.

De manera similar a como sucede con la suma y el producto por un escalar, el producto puede obtenerse multiplicando las hipermatrices como si éstas fueran matrices, siempre que las correspondientes submatrices sean conformables para las operaciones a efectuar con ellas.

Para ilustrar esto mediante un ejemplo consideremos las matrices

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

si establecemos las siguientes particiones para A y B

$$\left[\begin{array}{ccc|c} 3 & 1 & -5 & -2 \\ -1 & 4 & 0 & 2 \\ \hline 0 & -2 & 0 & 7 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \left[\begin{array}{c|cc} -1 & 4 & 1 \\ 0 & 3 & -2 \\ -2 & 3 & 1 \\ \hline 1 & 0 & 0 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

y multiplicamos las hipermatrices como si fueran matrices, se obtiene

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

donde

$$A_{11}B_{11} + A_{12}B_{21} = \begin{bmatrix} 3 & 1 & -5 \\ -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$A_{11}B_{12} + A_{12}B_{22} = \begin{bmatrix} 3 & 1 & -5 \\ -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 3 & -2 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & -4 \\ 8 & -9 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -4 \\ 8 & -9 \end{bmatrix}$$

$$A_{21}B_{11} + A_{22}B_{21} = \begin{bmatrix} 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} 7 \end{bmatrix}\begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} + \begin{bmatrix} 7 \end{bmatrix} = \begin{bmatrix} 7 \end{bmatrix}$$

$$A_{21}B_{12} + A_{22}B_{22} = \begin{bmatrix} 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 3 & -2 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 7 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} -6 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \end{bmatrix} = \begin{bmatrix} -6 & 4 \end{bmatrix}$$

por lo que se llega a la hipermatriz

$$\left[\begin{array}{c|cc} 5 & 0 & -4 \\ 3 & 8 & -9 \\ \hline 7 & -6 & 4 \end{array}\right]$$

que es una partición de AB; esto es

$$AB = \begin{bmatrix} 5 & 0 & -4 \\ 3 & 8 & -9 \\ 7 & -6 & 4 \end{bmatrix}$$

En general, se tiene el siguiente teorema

VI.8.3 TEOREMA

Sean A y B dos matrices con elementos en C, de $m \times n$ y $n \times q$ respectivamente:

Si $\left[A_{ij}\right]$ es la partición de A inducida por $(m_1, m_2, \ldots, m_s)$ y $(n_1, n_2, \ldots, n_t)$, y $\left[B_{ij}\right]$ es la partición de B inducida por $(n_1, n_2, \ldots, n_t)$ y $(q_1, q_2, \ldots, q_u)$; entonces

$$\left[P_{ij}\right] \quad \text{con} \quad P_{ij} = \sum_{k=1}^{t} A_{ik}B_{kj}$$

es la partición de AB inducida por $(m_1, m_2, \ldots, m_s)$ y $(q_1, q_2, \ldots, q_u)$.

DEMOSTRACION

Sean $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ matrices de $m \times n$ y $n \times p$, y sean: $\left[A_{rc}\right]$ la partición de A inducida por $(m_1, m_2, \ldots, m_s)$ y $(n_1, n_2, \ldots, n_t)$, y $\left[B_{rc}\right]$ la partición de B inducida por $(n_1, n_2, \ldots, n_t)$ y $(q_1, q_2, \ldots, q_u)$.

Hagamos

$$P_{rc} = \sum_{k=1}^{t} A_{rk}B_{kc}$$

De la definición VI.8.1. se tiene que A_{rk} es una matriz de $m_r \times n_k$ definida por

$$A_{rk} = \left[a_{ij}\right] \quad \text{con} \quad \begin{cases} i = \left[m_1 + \ldots + (m_{r-1}+1)\right], \ldots, \left[m_1 + \ldots + m_r\right] \\ j = \left[n_1 + \ldots + (n_{k-1}+1)\right], \ldots, \left[n_1 + \ldots + n_k\right] \end{cases}$$

y que B_{kc} es una matriz de $n_k \times q_c$ definida por

$$B_{kc} = \left[b_{ij} \right] \text{ con } \begin{cases} i = \left[n_1 + \ldots + (n_{k-1}+1) \right], \ldots, \left[n_1 + \ldots + n_k \right] \\ j = \left[q_1 + \ldots + (q_{c-1}+1) \right], \ldots, \left[q_1 + \ldots + q_c \right] \end{cases}$$

entonces, de VI.3.1 se tiene que $A_{rk}B_{kc}$ es una matriz de $m_r \times q_c$ definida por

$$A_{rk}B_{kc} = \left[\sum_{\ell = n_1 + \ldots + (n_{k-1}+1)}^{n_1 + \ldots + n_k} a_{i\ell} b_{\ell j} \right]$$

$$\text{con } \begin{cases} i = \left[m_1 + \ldots + (m_{r-1}+1) \right], \ldots, \left[m_1 + \ldots + m_r \right] \\ j = \left[q_1 + \ldots + (q_{c-1}+1) \right], \ldots, \left[q_1 + \ldots + q_c \right] \end{cases}$$

En consecuencia

$$P_{rc} = \sum_{k=1}^{t} \left[\sum_{\ell = n_1 + \ldots + (n_{k-1}+1)}^{n_1 + \ldots + n_k} a_{i\ell} b_{\ell k} \right] = \left[\sum_{\ell=1}^{n} a_{i\ell} b_{\ell j} \right]$$

$$\text{con } \begin{cases} i = \left[m_1 + \ldots + (m_{r-1}+1) \right], \ldots, \left[m_1 + \ldots + m_r \right] \\ j = \left[q_1 + \ldots + (q_{c-1}+1) \right], \ldots, \left[q_1 + \ldots + q_c \right] \end{cases}$$

Por otra parte, de VI.3.1

$$AB = \left[\sum_{\ell=1}^{n} a_{i\ell} b_{\ell j} \right] \quad \text{con } \begin{cases} i = 1, \ldots, m \\ j = 1, \ldots, q \end{cases}$$

si ahora efectuamos en la matriz AB la partición inducida por $(m_1, m_2, \ldots, m_s)$ y $(q_1, q_2, \ldots, q_u)$, según VI.8.1 se obtiene el arreglo $\left[H_{rc} \right]$, donde

$$H_{rc} = \left[\sum_{\ell=1}^{n} a_{i\ell} b_{\ell j} \right] \quad \text{con } \begin{cases} i = \left[m_1 + \ldots + (m_{r-1}+1) \right], \ldots, \left[m_1 + \ldots + m_r \right] \\ j = \left[q_1 + \ldots + (q_{c-1}+1) \right], \ldots, \left[q_1 + \ldots + q_c \right] \end{cases}$$

se tiene entonces que $H_{rc} = P_{rc}$ y la prueba termina.

$\square$

Como hemos visto, la partición de una matriz no es única. Esto nos permite seleccionar aquella que más nos convenga de acuerdo con las condiciones del problema.

Así, el concepto de partición puede ser útil para simplificar los cálculos al efectuar operaciones con matrices, cuando éstas presentan características especiales en su estructura.

A manera de ejemplo, consideremos las matrices

$$X = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \end{bmatrix} \qquad Y = \begin{bmatrix} 2 & -1 \\ 1 & 0 \\ 3 & 4 \\ 6 & -2 \end{bmatrix}$$

para las cuales queremos obtener el producto XY.

Nos conviene efectuar primero las particiones siguientes*

$$X = \left[\begin{array}{cc|cc} 1 & -2 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ \hline 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \end{array}\right] = \begin{bmatrix} A & O \\ O & \frac{1}{2}\,I \end{bmatrix} \qquad Y = \left[\begin{array}{cc} 2 & -1 \\ 1 & 0 \\ \hline 3 & 4 \\ 6 & -2 \end{array}\right] = \begin{bmatrix} B \\ C \end{bmatrix}$$

con lo que

$$XY = \begin{bmatrix} A & O \\ O & \frac{1}{2}\,I \end{bmatrix} \begin{bmatrix} B \\ C \end{bmatrix} = \begin{bmatrix} AB + OC \\ OB + \frac{1}{2}\,IC \end{bmatrix} = \begin{bmatrix} AB \\ \frac{1}{2}C \end{bmatrix}$$

Así, efectuamos únicamente los productos

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

y

$$\frac{1}{2}C = \frac{1}{2}\begin{bmatrix} 3 & 4 \\ 6 & -2 \end{bmatrix} = \begin{bmatrix} 3/2 & 2 \\ 3 & -1 \end{bmatrix}$$

con lo que se obtiene la matriz

$$XY = \begin{bmatrix} 0 & -1 \\ 7 & -3 \\ 3 & -2 \\ 3/2 & 2 \\ 3 & -1 \end{bmatrix}$$

*En estas expresiones la igualdad no está empleada en un sentido estricto, ya que X es una matriz mientras que la partición propuesta es una hipermatriz; sin embargo, es frecuente hallar este abuso de notación en diversos libros y no produce confusión si se le interpreta correctamente.

VI.8.4 EJERCICIOS

1.- Para la matriz

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

Obtener la partición inducida por

a) (2,3,2) y (3,1,2)

b) (3,3,1) y (2,4)

c) (4,3) y (1,2,3)

y en cada caso determinar la submatriz A_{21} correspondiente.

2.- Demostrar el teorema VI.8.2

3.- Para las matrices

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

a) Hallar el producto AB empleando las particiones inducidas por (2,3) y (2,1,2), y por (2,1,2) y (2,2) para A y B, respectivamente.

b) ¿Qué sucede si para A se usa la partición anterior y para

B la inducida por (3,2) y (1,2,1)?

c) Hallar una partición para A que permita obtener el producto AB con la partición de B correspondiente al inciso b), y obtener dicho producto.

4.- Para las matrices

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

Obtener el producto AB haciendo las particiones más convenientes para simplificar los cálculos.

5.- Si A y B son matrices conocidas de orden $m \times n$ y $m \times 1$, respectivamente, y U y V son matrices incógnitas de orden $n \times 1$ y $m \times 1$, respectivamente; transformar la igualdad

$$AU + V = B$$

en un sistema de ecuaciones lineales, determinando la matriz de coeficientes, la de incógnitas y la de términos independientes.

CAPITULO VII DETERMINANTES

INGENIERIA
ANEXO
BIBLIOTECA

INTRODUCCION

Seguramente el lector habrá tenido ya algún contacto con los determinantes; especialmente con los de segundo y tercer orden.

La idea de determinante es, en realidad, más antigua que la de matriz. Descubierta por Cramer durante sus trabajos orientados a la resolución de sistemas de ecuaciones lineales, fue expuesta por primera vez en 1750, cien años antes de que Sylvester y Cayley empezaran a hablar de matrices.

En la actualidad, sin embargo, el concepto de determinante suele presentarse como consecuencia de la teoría de matrices, y ha sufrido, incluso, el proceso de "axiomatización", del cual surge una definición integrada por cuatro postulados.

En este capítulo estableceremos una definición para el concepto de determinante a partir de un razonamiento similar al que históricamente le dio origen, aunque emplearemos en ella el concepto de matriz. En la deducción de las propiedades y en la presentación de los métodos para el cálculo de determinantes se manejarán también algunos elementos de la teoría de matrices.

VII.1 CONCEPTOS BASICOS

Con el propósito de motivar la definición de determinante, consideremos el problema de resolver el siguiente sistema de ecuaciones lineales

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \qquad (1)$$

Para obtener el valor de x_1 podemos multiplicar por a_{22} la primera ecuación y por a_{12} la segunda, con lo que se obtiene el sistema equivalente

$$a_{22}a_{11}x_1 + a_{22}a_{12}x_2 = a_{22}b_1$$

$$a_{12}a_{21}x_1 + a_{12}a_{22}x_2 = a_{12}b_2$$

Restando la segunda ecuación de la primera, y ordenando convenientemente los términos, se llega a

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} \qquad (2)$$

De manera semejante se obtiene también que

$$x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \qquad (3)$$

Sustituyendo estos valores de x_1 y x_2 en las ecuaciones del sistema (1) podemos comprobar que constituyen una solución del mismo.

Las expresiones (2) y (3) tienen el mismo denominador, el cual está expresado en términos de los elementos de la matriz

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

que es la matriz de coeficientes del sistema. Al número correspondiente a dicho denominador se le conoce como el determinante de la matriz A y se le representa con "det A"; esto es:

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

Otra forma de representar al determinante de una matriz consiste en escribir sus elementos tal y como aparecen en el arreglo, pero reemplazando los paréntesis rectangulares por barras verticales para indicar que se trata de un determinante.

Así, para la matriz anterior se tiene

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \qquad (4)$$

Expresión que puede considerarse como la definición del determinante de orden dos.

Es importante resaltar que una matriz es un arreglo de números mientras que su determinante es un número. Por ejemplo, para la matriz

$$A = \begin{bmatrix} 5 & 4 \\ -2 & -3 \end{bmatrix}$$

se tiene que

$$\begin{vmatrix} 5 & 4 \\ -2 & -3 \end{vmatrix} = (5)(-3) - (4)(-2) = -15 + 8 = -7$$

por lo que

$$\det A = -7$$

Regresando a las expresiones (2) y (3), vemos que los numeradores de éstas también pueden ser considerados como determinantes, ya que

$$b_1a_{22} - a_{12}b_2 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}$$

y

$$a_{11}b_2 - b_1a_{21} = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

En consecuencia, los valores de las incógnitas pueden obtenerse como el cociente de dos determinantes; es decir

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \quad \text{y} \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

Al método sugerido por estas dos últimas expresiones se le conoce como "regla de Cramer".

Consideremos ahora un sistema de tres ecuaciones lineales con tres incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \qquad (5)$$

Siguiendo un proceso similar al anterior se encuentra que los valores de las incógnitas que satisfacen al sistema son

$$x_1 = \frac{b_1a_{22}a_{33}-b_1a_{23}a_{32}-a_{12}b_2a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}-a_{13}a_{22}b_3}{a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}} \qquad (6)$$

$$x_2 = \frac{a_{11}b_2a_{33}-a_{11}a_{23}b_3-b_1a_{21}a_{33}+b_1a_{23}a_{31}+a_{13}a_{21}b_3-a_{13}b_2a_{31}}{a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}} \qquad (7)$$

$$x_3 = \frac{a_{11}a_{22}b_3-a_{11}b_2a_{32}-a_{12}a_{21}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-b_1a_{22}a_{31}}{a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}} \qquad (8)$$

Al común denominador de las expresiones (6), (7) y (8) se le conoce como el determinante de la matriz

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

Así, la expresión que define al determinante de tercer orden es

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \qquad (9)$$

Por ejemplo

$$\begin{vmatrix} 1 & 0 & -3 \\ -4 & 5 & 2 \\ 1 & -2 & 0 \end{vmatrix} = (1)(5)(0) - (1)(2)(-2) - (0)(-4)(0) + (0)(2)(1) + (-3)(-4)(-2) - (-3)(5)(1) =$$

$$= 0 + 4 - 0 + 0 - 24 + 15 = -5$$

De acuerdo con (9) podemos escribir la solución del sistema (5) como sigue

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \qquad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

Así pues, resulta natural tratar de generalizar la regla de Cramer al caso de n ecuaciones con n incógnitas, y para ello se requiere una definición general para el determinante de orden n. Sin embargo, tal definición no puede obtenerse de la misma manera que en los casos de segundo y tercer orden (es decir, resolviendo un sistema de n ecuaciones lineales con n incógnitas), pues a medida que n aumenta los cálculos se hacen más complicados y para n arbitrario son irrealizables.

En consecuencia, buscaremos establecer una ley general a partir del análisis de las expresiones (4) y (9), que definen a los determinantes de segundo y tercer orden. Para dicho análisis, empero, se requieren algunos conceptos que no hemos manejado aún y que estudiaremos a continuación.

- Permutaciones

Las permutaciones de los elementos de un conjunto (finito) son las diferentes maneras en que éstos pueden ser arreglados.

Por ejemplo, las permutaciones del conjunto

$$S = \{1, 2, 3\}$$

son los arreglos

$$p_1 = (1, 2, 3)$$
$$p_2 = (1, 3, 2)$$
$$p_3 = (2, 1, 3)$$
$$p_4 = (2, 3, 1)$$
$$p_5 = (3, 1, 2)$$
$$p_6 = (3, 2, 1)$$

Para los propósitos de este capítulo nos interesan únicamente las permutaciones de conjuntos formados por números naturales. Se tiene entonces la siguiente definición

VII.1.1 DEFINICION

Una permutación del conjunto $S = \{1, 2, \ldots, n\}$ es un arreglo de la forma

$$(\alpha_1 \;\; \alpha_2, \ldots, \alpha_n)$$

donde $\alpha_i \in S \;\; \forall\, i \quad y \quad \alpha_i \neq \alpha_j \quad si \quad i \neq j$

Así, para la permutación p_1 del ejemplo anterior se tiene que

$$\alpha_1 = 1, \quad \alpha_2 = 2 \quad y \quad \alpha_3 = 3$$

mientras que para p_4 se tiene que

$$\alpha_1 = 2, \quad \alpha_2 = 3 \quad y \quad \alpha_3 = 1$$

Cuando en una permutación todos los números aparecen en el orden natural, como en p_1, se dice que ésta es la "permutación principal" del conjunto.

En general, $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ es la permutación principal del conjunto $S = \{1, 2, \ldots, n\}$ si $\alpha_i = i, \forall i$.

Respecto al número de permutaciones de un conjunto se tiene el siguiente teorema

> VII.1.2 TEOREMA
>
> El conjunto $S = \{1, 2, \ldots, n\}$ tiene
> $n!$ permutaciones diferentes

DEMOSTRACION (por inducción)

Para $n = 1$ se tiene una sola permutación $(1!)$.

Para $n = k$ se supone que el conjunto tiene $k!$ permutaciones diferentes.

Sea ahora $S = \{1, 2, \ldots, k+1\}$, y consideremos una permutación arbitraria de S

$$(\alpha_1, \alpha_2, \ldots, \alpha_{k+1})$$

El primer elemento (α_1) puede ser seleccionado de $k+1$ maneras diferentes, y para cada una de ellas los k elementos restantes pueden arreglarse, por hipótesis, de $k!$ maneras diferentes. En consecuencia, se tienen

$$(k+1)(k!) = (k+1)!$$

permutaciones diferentes y la prueba termina.

$\square$

En una permutación se dice que hay una "inversión" por cada dos nú-

meros que se encuentren en un orden diferente al natural. Así, por ejemplo, en la permutación

$$(3, 1, 4, 2)$$

se tienen tres inversiones; es decir, podemos encontrar tres parejas

$$(3, 1) \quad (3, 2) \quad y \quad (4, 2)$$

donde el primer elemento es mayor que el segundo y aparece antes que éste en la permutación.

Se tiene al respecto la siguiente definición

INGENIERIA
ANEXO
BIBLIOTECA

> VII.1.3 DEFINICION
>
> i) Una permutación $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ tiene m inversiones si existen m parejas (α_i, α_j) tales que $i < j$ y $\alpha_i > \alpha_j$
>
> ii) Una permutación es de clase par si tiene un número par de inversiones; en caso contrario se dice que es de clase impar.

Así, por ejemplo, para las permutaciones del conjunto $S = \{1, 2, 3\}$ se tiene que

$p_1 = (1, 2, 3)$	es de clase par	(cero inversiones)
$p_2 = (1, 3, 2)$	es de clase impar	(una inversión)
$p_3 = (2, 1, 3)$	es de clase impar	(una inversión)
$p_4 = (2, 3, 1)$	es de clase par	(dos inversiones)
$p_5 = (3, 1, 2)$	es de clase par	(dos inversiones)
$p_6 = (3, 2, 1)$	es de clase impar	(tres inversiones)

En general, de las n! permutaciones de un conjunto - - - - - - - - $S = \{1, 2, \ldots, n\}$ la mitad son de clase par y la mitad de clase impar.

Definición de determinante.

Recordemos ahora las expresiones que definen a los determinantes de segundo y tercer orden.

Para $n = 2$ se tiene

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

y para $n = 3$ se tiene

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

Podemos observar que en ambos casos:

1) El determinante es la suma de n! productos, la mitad de ellos con singo + y la mitad con signo -.
2) Cada uno de los productos consta de n factores.
3) En cada producto hay un elemento de cada renglón y un elemento de cada columna.
4) Si los factores se ordenan de tal manera que los primeros índices formen una permutación principal, corresponde el signo + a los productos cuyos segundos índices forman una permutación de clase par y corresponde el signo - a los productos cuyos segundos índices forman una permutación de clase impar.

Ahora, para establecer una ley general consideremos la siguiente matriz cuadrada de orden n

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

y un producto arbitrario de n de sus elementos, en el cual haya un elemento de cada renglón y un elemento de cada columna; esto es

$$a_{1\alpha_1} \; a_{2\alpha_2} \; \ldots \; a_{n\alpha_n}$$

donde $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ es una permutación de los índices - - - - - $\{1, 2, \ldots, n\}$.

Como el conjunto $S = \{1, 2, \ldots, n\}$ tiene n! permutaciones diferentes podemos formar, para cada una de ellas, un producto de este tipo; esto es

$$a_{1\alpha_{k1}} \; a_{2\alpha_{k2}} \; \ldots \; a_{n\alpha_{kn}}, \quad \text{donde } k = 1, 2, \ldots, n!$$

En cada uno de dichos productos los elementos están ordenados de manera que los primeros índices forman una permutación principal y los segundos índices forman una permutación - - - - - - - - - - $P_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn})$. Para dicha permutación hacemos

$$\varepsilon(P_k) = \begin{cases} +, & \text{si } P_k \text{ es de clase par} \\ -, & \text{si } P_k \text{ es de clase impar} \end{cases}$$

Y formamos con ello un producto provisto de signo, esto es

$$\varepsilon(P_k) \; a_{1\alpha_{k1}} \; a_{2\alpha_{k2}} \; \ldots \; a_{n\alpha_{kn}}$$

al que podemos representar, de manera compacta, con

$$\varepsilon(p_k) \prod_{i=1}^{n} a_{i\alpha_{ki}}$$

A la suma de los n! productos de este tipo se le conoce como el determinante de A. Se tiene así la siguiente definición

VII.1.4 DEFINICION

Sea $A = \left[a_{ij} \right]$ una matriz de n×n con elementos en C, y sea $p_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn})$ una permutación del conjunto {1, 2,..., n}. Se llama determinante de A al número

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k) \prod_{i=1}^{n} a_{i\alpha_{ki}}$$

donde $\varepsilon(p_k) = \begin{cases} +, & \text{si } p_k \text{ es de clase par} \\ -, & \text{si } p_k \text{ es de clase impar} \end{cases}$

Para calcular el valor de un determinante a partir de la definición anterior, podemos proceder de la siguiente manera:

Se obtiene el "término principal", que es el producto de los elementos de la diagonal principal con los factores ordenados de tal manera que tanto los primeros como los segundos índices formen una permutación principal.

A partir del término principal se obtienen los demás términos dejando fijos los primeros índices y permutando los segundos de todas las maneras posibles. Como de los segundos índices hay n! permutaciones diferentes se obtendrán n! términos, correspondiéndoles

el signo + o el signo - según sea la permutación de clase par o de clase impar.

Así, por ejemplo, para obtener el desarrollo del determinante de tercer orden

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

escribimos el término principal

$$a_{11} \quad a_{22} \quad a_{33}$$

y permutamos los segundos índices de todas las maneras posibles

p_k	clase de p_k	$\varepsilon(p_k)$
$p_1 = (1, 2, 3)$	par	+
$p_2 = (1, 3, 2)$	impar	-
$p_3 = (2, 1, 3)$	impar	-
$p_4 = (2, 3, 1)$	par	+
$p_5 = (3, 1, 2)$	par	+
$p_6 = (3, 2, 1)$	impar	-

con lo que se obtienen los seis términos del desarrollo; esto es

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

Esta expresión coincide con la definición que teníamos para el determinante de tercer orden.

VII.1.5 EJERCICIOS

1.- Determinar cuántas inversiones tiene cada una de las siguientes permutaciones

a) (1, 5, 3, 2, 4)

b) (3, 2, 1, 4)

2.- a) Escribir la permutación de los cinco primeros números naturales que tiene el mayor número de inversiones.

b) Demostrar que el mayor número de inversiones que puede tener una permutación de n números es

$$\frac{n(n-1)}{2}$$

3.- Determinar la condición que debe cumplir n para que la permutación

$$(n-1,\ n,\ n-2,\ n-3,\ n-4, \ldots,\ 1)$$

sea de clase impar.

4.- Calcular, a partir de la definición VII.1.4, el valor del determinante de la matriz

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

VII.2 PROPIEDADES DE LOS DETERMINANTES

Las principales propiedades de los determinantes pueden ser consideradas en dos grupos: Las primeras se refieren a las condiciones bajo las cuales se puede concluir que un determinante es nulo mediante la simple inspección de las líneas de la matriz (renglones y columnas), así como a los efectos producidos en el determinante al efectuar transformaciones elementales con las líneas de la matriz. El segundo grupo se refiere a las propiedades del determinante en relación con las operaciones definidas para las matrices.

Para demostrar algunas de estas propiedades se requieren ciertos resultados relativos a intercambios de números en una permutación, los cuales presentaremos a continuación a manera de lemas.

Al intercambiar dos números adyacentes en una permutación de clase par ésta se transforma en una permutación de clase impar y viceversa. Se dice por ello que se obtiene una permutación de "paridad" diferente.

En efecto, si los números a intercambiar se encuentran inicialmente en el orden natural, después del intercambio la permutación tendrá una inversión más ya que el resto de las inversiones no se alteran por ser adyacentes los números que se intercambian; por otra parte, si los números no se encuentran inicialmente en el orden natural, después del intercambio la permutación tendrá una inversión menos. En ambos casos se obtiene una permutación de paridad diferente.

Generalizando el resultado anterior a un número impar de intercambios se obtiene el siguiente resultado

LEMA 1

Al efectuar en una permutación un número impar de intercambios de números adyacentes se obtiene una permutación de paridad diferente.

Con ayuda de este lema podemos probar a su vez el siguiente enunciado

LEMA 2

Al intercambiar en una permutación dos números cualesquiera se obtiene una permutación de paridad diferente.

DEMOSTRACION

Sea

$$(\alpha_1, \alpha_2, \ldots, \alpha_{r-1}, \alpha_r, \alpha_{r+1}, \ldots, \alpha_{s-1}, \alpha_s, \alpha_{s+1}, \ldots, \alpha_n)$$

una permutación del conjunto $S = \{1, 2, \ldots, n\}$.

Para intercambiar los números α_r y α_s, colocamos primero a α_s entre α_{r-1} y α_r

$$(\alpha_1, \alpha_2, \ldots, \alpha_{r-1}, \alpha_r, \alpha_{r+1}, \ldots, \alpha_{s-1}, \alpha_s, \alpha_{s+1}, \ldots, \alpha_n)$$

para lo cual es necesario efectuar $(s-1)-(r-1) = s-r$ intercambios de números adyacentes.

A continuación se requiere colocar a α_r entre α_{s-1} y α_{s+1}

$$(\alpha_1, \alpha_2, \ldots, \alpha_{r-1}, \alpha_s, \alpha_r, \alpha_{r+1}, \ldots, \alpha_{s-1}, \alpha_{s+1}, \ldots, \alpha_n)$$

para lo cual debemos efectuar $(s-1)-r = s-r-1$ intercambios de números adyacentes.

En consecuencia, el total de intercambios de números adyacentes será

$$(s-r) + (s-r-1) = 2(s-r)-1$$

que es un número impar; por lo que, del Lema 1, se obtiene una permutación de paridad diferente. □

Respecto al número de intercambios necesario para llevar una permutación cualquiera a la permutación principal se tiene el siguiente enunciado

LEMA 3

Si una permutación tiene m inversiones, se requieren m intercambios de números adyacentes para transformarla en la permutación principal.

DEMOSTRACION

Consideremos una permutación del conjunto $S = \{1, 2, \ldots, n\}$ con un total de m inversiones; entonces:

el número n estará antes de j_n números menores que él,
el número $n-1$ estará antes de j_{n-1} números menores que él,
⋮
el número 2 estará antes de j_2 números menores que él,

de tal forma que

$$\sum_{r=2}^{n} j_r = m$$

Esto significa que para colocar al númerc n en la posición que le corresponde en la permutación principal, deberán efectuarse j_n intercambios de números adyacentes; para colocar al número n-1 deberán efectuarse j_{n-1} intercambios, y así hasta colocar al número 2 en su posición correspondiente. Al final del proceso se habrán efectuado un total de

$$j_n + j_{n-1} + \ldots + j_2 = m$$

intercambios de números adyacentes.

$\square$

Una primera propiedad de los determinantes consiste en "descompo - ner" el determinante de una matriz A en la suma de dos determinantes, de acuerdo con el siguiente enunciado.

> VII.2.1 TEOREMA
>
> Si los elementos del renglón r de una matriz $A = \left[a_{ij}\right]$ pueden expresarse como
>
> $$a_{rj} = b_{rj} + c_{rj}$$
>
> entonces
>
> $$\det A = \det B + \det C$$
>
> donde $B = \left[b_{ij}\right]$ y $C = \left[c_{ij}\right]$ son tales que
>
> $$b_{ij} = \begin{cases} a_{ij}, & \text{para } i \neq r \\ b_{rj}, & \text{para } i = r \end{cases} \quad \text{y} \quad c_{ij} = \begin{cases} a_{ij}, & \text{para } i \neq r \\ c_{rj}, & \text{para } i = r \end{cases}$$

Podemos "visualizar" esta propiedad con ayuda de la siguiente ex - presión

$$\begin{array}{l} \\ \\ \\ \text{renglón} \rightarrow \\ r \\ \\ \end{array} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{r1}+c_{r1} & b_{r2}+c_{r2} & \cdots & b_{rn}+c_{rn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{r1} & c_{r2} & \cdots & c_{rn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

DEMOSTRACION

De VII.1.4 se tiene que

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

pero $a_{r\alpha_{kr}} = b_{r\alpha_{kr}} + c_{r\alpha_{kr}}$, por lo que

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots (b_{r\alpha_{kr}} + c_{r\alpha_{kr}}) \cdots a_{n\alpha_{kn}}$$

y en consecuencia

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}} a_{2\alpha_{k2}} \cdots b_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}} +$$

$$+ \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots c_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

Entonces, si $B = \left[b_{ij} \right]$ y $C = \left[c_{ij} \right]$ son tales que

$$b_{ij} = \begin{cases} a_{ij}, & \text{para } i \neq r \\ b_{rj}, & \text{para } i = r \end{cases} \quad \text{y} \quad c_{ij} = \begin{cases} a_{ij}, & \text{para } i \neq r \\ c_{rj}, & \text{para } i = r \end{cases}$$

se tendrá

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k)\, b_{1\alpha_{k1}}\, b_{2\alpha_{k2}} \cdots b_{r\alpha_{kr}} \cdots b_{n\alpha_{kn}} +$$

$$+ \sum_{k=1}^{n!} \varepsilon(p_k)\, c_{1\alpha_{k1}}\, c_{2\alpha_{k2}} \cdots c_{r\alpha_{kr}} \cdots c_{n\alpha_{kn}}$$

por lo que, de VII.1.4 nuevamente

$$\det A = \det B + \det C$$

□

A continuación se presentan, en el teorema VII.2.2, las propiedades que hemos considerado del primer grupo

VII.2.2 TEOREMA

Sea $A = \left[a_{ij} \right]$ una matriz de $n \times n$ con elementos en C:

i) Si los elementos de una línea de A (renglón o columna) son todos nulos, entonces $\det A = 0$.

ii) Si B se obtiene de A multiplicando los elementos de una de sus líneas por un número $\lambda \in C$, entonces $\det B = \lambda \det A$.

iii) Si B se obtiene de A intercambiando dos líneas paralelas (dos renglones o dos columnas), entonces $\det B = -\det A$.

iv) Si dos líneas paralelas de A son proporcionales, entonces $\det A = 0$.

v) Si B se obtiene de A sumando a los elementos de una línea los elementos de una línea paralela multiplicados por un número $\lambda \in C$, entonces $\det B = \det A$.

DEMOSTRACION

Se demuestran a continuación las propiedades para el caso de los renglones; en el caso de las columnas la demostración es similar.

i) De VII.1.4 se tiene que

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

En consecuencia, si todos los elementos del renglón r son iguales a cero, cada uno de los términos del desarrollo anterior tendrá al menos un factor igual a cero (el factor $a_{r\alpha_{kr}}$), por lo que

$\det A = 0$

ii) Sea $B = \left[b_{ij}\right]$ una matriz obtenida a partir de A multiplicando los elementos del renglón r por un número λ; esto es

$$b_{ij} = \begin{cases} a_{ij}, & \text{para } i \neq r \\ \lambda a_{ij}, & \text{para } i = r \end{cases}$$

De VII.1.4 se sigue que

$$\det B = \sum_{k=1}^{n!} \varepsilon(p_k)\, b_{1\alpha_{k1}}\, b_{2\alpha_{k2}} \cdots b_{r\alpha_{kr}} \cdots b_{n\alpha_{kn}}$$

por lo que

$$\det B = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots \lambda a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

$$= \lambda \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

$$\det B = \lambda \det A$$

iii) Sea $B = \left[b_{ij}\right]$ una matriz obtenida a partir de A intercambiando los renglones r y s, donde $r < s$; es decir

$$b_{ij} = \begin{cases} a_{ij}, & \text{para } i \neq r, s \\ a_{sj}, & \text{para } i = r \\ a_{rj}, & \text{para } i = s \end{cases}$$

y consideremos un término cualquiera del desarrollo de det B según la definición VII.1.4; esto es

$$\varepsilon(p_k)\, b_{1\alpha_{k1}}\, b_{2\alpha_{k2}} \cdots b_{r\alpha_{kr}} \cdots b_{s\alpha_{ks}} \cdots b_{n\alpha_{kn}}$$

donde el signo $\varepsilon(p_k)$ depende de la permutación

$$p_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kr}, \ldots, \alpha_{ks}, \ldots, \alpha_{kn})$$

sustituyendo los elementos de B por su expresión en términos de los de A se obtiene

$$\varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{s\alpha_{kr}} \cdots a_{r\alpha_{ks}} \cdots a_{n\alpha_{kn}}$$

que puede escribirse como

$$\varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{r\alpha_{ks}} \cdots a_{s\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

Esta expresión constituye también un término en el desarrollo de det A, salvo el signo $\varepsilon(p_k)$; puesto que en dicho desarrollo al producto

$$a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{r\alpha_{ks}} \cdots a_{s\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

le corresponde el signo $\varepsilon(q_k)$ que depende de la permutación

$$q_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{ks}, \ldots, \alpha_{kr}, \ldots, \alpha_{kn})$$

la cual puede obtenerse a partir de p_k intercambiando los números α_{kr} y α_{ks}. En consecuencia, por el Lema 2 p_k y q_k son de paridad diferente, por lo que

$$\varepsilon(p_k) = -\varepsilon(q_k)$$

se tiene por tanto que

$$\sum_{k=1}^{n!} \varepsilon(p_k)\, b_{1\alpha_{k1}}\, b_{2\alpha_{k2}} \cdots b_{n\alpha_{kn}} =$$

$$= \sum_{k=1}^{n!} -\varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{n\alpha_{kn}}$$

y en consecuencia

$$\det B = -\det A$$

iv) Sea $A = \left[a_{ij}\right]$ una matriz cuyo renglón s es igual al renglón r multiplicado por un número λ; esto es

$$a_{sj} = \lambda a_{rj}$$

y formemos una matriz B cuyos elementos sean iguales a los de A salvo los del renglón s, que serán iguales a los del renglón r; esto es

$$b_{ij} = \begin{cases} a_{ij}, & \text{para } i \neq s \\ a_{rj}, & \text{para } i = s \end{cases}$$

Entonces, por la propiedad ii)

$$\det A = \lambda \det B$$

Ahora, si formamos una matriz C intercambiando en B los renglones r y s, se tendrá que B = C y por tanto

$$\det B = \det C$$

pero por la propiedad iii) se tendrá también que

$$\det B = -\det C$$

por lo cual

$$\det B = 0$$

y en consecuencia

$$\det A = \lambda \times 0 = 0$$

v) Sea $B = \left[b_{ij}\right]$ una matriz obtenida a partir de A sumando a los elementos del renglón r los elementos del renglón s multiplicados por un número λ; esto es:

$$b_{ij} = \begin{cases} a_{ij} & , \text{ para } i \neq r \\ a_{rj} + \lambda a_{sj}, & \text{para } i = r \end{cases}$$

Entonces por VII.2.1, podemos expresar al determinante de B como

$$\det B = \det A + \det C$$

donde $C = \left[c_{ij}\right]$ es tal que

$$c_{ij} = \begin{cases} a_{ij} & , \text{ para } i \neq r \\ \lambda a_{sj}, & \text{para } i = r \end{cases}$$

y sus renglones r y s son proporcionales; en consecuencia, de la propiedad iv)

$$\det C = 0$$

y por tanto

$$\det B = \det A$$

$\square$

Veremos ahora las propiedades respecto a las operaciones con matrices

VII.2.3 TEOREMA

Si $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ son dos matrices de $n\times n$ con elementos en C, entonces:

i) $\det A = \det A^T$

ii) $\det(\lambda A) = \lambda^n \det A$

iii) $\det(AB) = (\det A)(\det B)$

DEMOSTRACION

i) Sea $A^T = \left[c_{ij}\right]$, donde $c_{ij} = a_{ji}$ por VI.7.7; y consideremos un término cualquiera del desarrollo de det A^T según la definición VII.1.4:

$$\varepsilon(p_k)\, c_{1\alpha_{k1}}\, c_{2\alpha_{k2}} \cdots c_{n\alpha_{kn}}$$

donde $\varepsilon(p_k)$ depende de la permutación

$$p_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn})$$

sustituyendo c_{ij} por a_{ji} se obtiene

$$\varepsilon(p_k)\, a_{\alpha_{k1}1}\, a_{\alpha_{k2}2} \cdots a_{\alpha_{kn}n}$$

ordenando los factores de tal manera que los primeros índices (que forman la permutación p_k) formen una permutación principal, se obtendrá la expresión

$$\varepsilon(p_k)\, a_{1\beta_{k1}}\, a_{2\beta_{k2}} \cdots a_{n\beta_{kn}}$$

donde ahora los segundos índices forman una permutación

$$q_k = (\beta_{k1}, \beta_{k2}, \ldots, \beta_{kn})$$

Como se ve, dicha expresión constituye también un término en el desarrollo de det A salvo, posiblemente, el signo $\varepsilon(p_k)$; puesto que a este producto le corresponde el signo $\varepsilon(q_k)$ en el desarrollo de det A.

Ahora bien; si p_k es de clase impar, por el Lema 3 se requiere un número impar de intercambios de números adyacentes para llevarla a la permutación principal, mismo número de intercambios que se efectúan en la permutación de los segundos índices, que es la permutación principal, para obtener q_k, por lo que q_k es también de clase impar. Mediante un razonamiento análogo se concluye que si p_k es de clase par entonces q_k es también de clase par; por lo que

$$\varepsilon(p_k) = \varepsilon(q_k)$$

se tiene por tanto que

$$\sum_{k=1}^{n!} \varepsilon(p_k)\, c_{1\alpha_{k1}}\, c_{2\alpha_{k2}} \cdots c_{n\alpha_{kn}} = \sum_{k=1}^{n!} \varepsilon(q_k)\, a_{1\beta_{k1}}\, a_{2\beta_{k2}} \cdots a_{n\beta_{kn}}$$

y en consecuencia

$$\det A^T = \det A$$

ii) La prueba es inmediata si consideramos que λA es una matriz que se obtiene de A multiplicando por λ cada uno de sus n renglones; entonces, aplicando reiteradamente ii) de VII.2.2 se concluye que

$$\det(\lambda A) = \lambda^n \det A$$

iii) La prueba de esta propiedad requiere de la generalización, a n sumandos, del teorema VII.2.1. El lector podrá aceptar fácilmente dicha generalización o, si lo prefiere, realizar la prueba por inducción.

Sean ahora dos matrices cualesquiera de n×n

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad y \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}$$

y consideremos una partición de B escribiéndola como

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} \quad \text{donde } B_i = (b_{i1}, b_{i2}, \dots, b_{in})$$

El producto AB también puede escribirse como

$$AB = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix} \quad \text{donde } C_i = a_{i1}B_1 + a_{i2}B_2 + \dots + a_{in}B_n$$

Entonces, considerando los n sumandos que constituyen el 1er. renglón de AB, por la generalización del teorema VII.2.1 podemos escribir el determinante de AB como la suma de n determinantes; esto es

$$\det(AB) = \begin{vmatrix} a_{11}B_1 \\ C_2 \\ \vdots \\ C_n \end{vmatrix} + \begin{vmatrix} a_{12}B_2 \\ C_2 \\ \vdots \\ C_n \end{vmatrix} + \dots + \begin{vmatrix} a_{1n}B_n \\ C_2 \\ \vdots \\ C_n \end{vmatrix}$$

Si llamamos $D_1, D_2, \dots, D_n$ a tales determinantes se tendrá

$$\det(AB) = D_1 + D_2 + \dots + D_n$$

Considerando ahora los n sumandos que constituyen el segundo renglón de cada una de las matrices correspondientes a dichos determinantes, podemos escribir

$$D_1 = \begin{vmatrix} a_{11}B_1 \\ a_{21}B_1 \\ \vdots \\ C_n \end{vmatrix} + \begin{vmatrix} a_{11}B_1 \\ a_{22}B_2 \\ \vdots \\ C_n \end{vmatrix} + \dots + \begin{vmatrix} a_{11}B_1 \\ a_{2n}B_n \\ \vdots \\ C_n \end{vmatrix} = D_{11} + D_{12} + \dots + D_{1n}$$

$$D_2 = \begin{vmatrix} a_{12}B_2 \\ a_{21}B_1 \\ \vdots \\ C_n \end{vmatrix} + \begin{vmatrix} a_{12}B_2 \\ a_{22}B_2 \\ \vdots \\ C_n \end{vmatrix} + \dots + \begin{vmatrix} a_{12}B_2 \\ a_{2n}B_n \\ \vdots \\ C_n \end{vmatrix} = D_{21} + D_{22} + \dots + D_{2n}$$

$$\vdots$$

$$D_n = \begin{vmatrix} a_{1n}B_n \\ a_{21}B_1 \\ \vdots \\ C_n \end{vmatrix} + \begin{vmatrix} a_{1n}B_n \\ a_{22}B_2 \\ \vdots \\ C_n \end{vmatrix} + \dots + \begin{vmatrix} a_{1n}B_n \\ a_{2n}B_n \\ \vdots \\ C_n \end{vmatrix} = D_{n1} + D_{n2} + \dots + D_{nn}$$

y podemos continuar este proceso hasta considerar los n sumandos del último renglón. Al final de dicho proceso se obtendrá que det(AB) es la suma de n^n determinantes

$$D_{j_1 j_2 \dots j_n} \quad ; \quad \text{donde} \begin{cases} j_1 = 1, 2, \dots, n \\ j_2 = 1, 2, \dots, n \\ \vdots \\ j_n = 1, 2, \dots, n \end{cases}$$

de los cuales la mayor parte es igual a cero por tener renglones que son proporcionales (considérese, por ejemplo, D_{11} cuyos dos primeros renglones son proporcionales). De hecho, sólo pueden llegar a ser diferentes de cero los n! determi -

nantes para los cuales $j_1, j_2, \ldots, j_n$ constituyen una permutación de los números $\{1, 2, .., n\}$.

Entonces

$$\det(AB) = \sum_{k=1}^{n!} D_{\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kn}}$$

donde $(\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kn})$ es una permutación de $\{1, 2, \ldots, n\}$

Tomando en cuenta la estructura de $D_{\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kn}}$ podemos escribir

$$\det(AB) = \sum_{k=1}^{n!} \begin{vmatrix} a_{1\alpha_{k_1}} & B_{\alpha_{k_1}} \\ a_{2\alpha_{k_2}} & B_{\alpha_{k_2}} \\ \vdots & \\ a_{n\alpha_{kn}} & B_{\alpha_{kn}} \end{vmatrix}$$

por lo que, aplicando reiteradamente ii) de VII.2.2, se tiene que

$$\det(AB) = \sum_{k=1}^{n!} a_{1\alpha_{k_1}} a_{2\alpha_{k_2}} \cdots a_{n\alpha_{kn}} \begin{vmatrix} B_{\alpha_{k_1}} \\ B_{\alpha_{k_2}} \\ \vdots \\ B_{\alpha_{kn}} \end{vmatrix}$$

Obsérvese en esta expresión que el determinante de la derecha puede transformarse en det B mediante el intercambio de renglones hasta lograr que la permutación

$$p_k = (\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kn})$$

se transforme en una permutación principal. Así, si p_k es de clase par dicho determinante es igual a det B y si p_k es de

clase impar es igual a − det B.

En consecuencia, podemos escribir

$$\det(AB) = \sum_{k=1}^{n!} a_{1\alpha_{k_1}} a_{2\alpha_{k_2}} \cdots a_{n\alpha_{kn}} \, \varepsilon(p_k) \det B$$

esto es

$$\det(AB) = \det B \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k_1}} a_{2\alpha_{k_2}} \cdots a_{n\alpha_{kn}}$$

$$= (\det B)(\det A)$$

$$\det(AB) = (\det A)(\det B)$$

como se quería. $\square$

INGENIERIA
ANEXO
BIBLIOTECA

VII.2.4 EJERCICIOS

1.- Determinar cuántos intercambios de números adyacentes son necesarios para transformar en la permutación principal cada una de las siguientes permutaciones

a) (3, 1, 5, 4, 2)

b) (4, 2, 1, 3)

2.- A partir de VII.2.2 y tomando en cuenta que

$$\begin{vmatrix} 2 & 1 & 3 & -2 \\ 1 & 4 & 0 & -1 \\ 0 & 2 & 1 & 0 \\ a & b & 2a & 3b \end{vmatrix} = 13a + 39b$$

Obtener el valor de los siguientes determinantes

a) $\begin{vmatrix} 2 & -1 & 3 & -2 \\ 1 & -4 & 0 & -1 \\ 0 & -2 & 1 & 0 \\ a & -b & 2a & 3b \end{vmatrix}$ b) $\begin{vmatrix} 2+ai & 1+bi & 3+2ai & -2+3bi \\ 1 & 4 & 0 & -1 \\ 0 & 2 & 1 & 0 \\ a & b & 2a & 3b \end{vmatrix}$

c) $\begin{vmatrix} 2 & 1 & 3 & -2 \\ 3a & 3b & 6a & 9b \\ 0 & 2 & 1 & 0 \\ a & b & 2a & 3b \end{vmatrix}$ d) $\begin{vmatrix} 2 & 1 & 3 & -2 \\ 1 & 4 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ a & b & 2a & 3b \end{vmatrix}$

e) $\begin{vmatrix} 3 & 1 & 2 & -2 \\ 0 & 4 & 1 & -1 \\ 1 & 2 & 0 & 0 \\ 2a & b & a & 3b \end{vmatrix}$ f) $\begin{vmatrix} 2 & 1 & 3 & -2 \\ i & 4i & 0 & -i \\ 0 & 2i & i & 0 \\ ai & bi & 2ai & 3bi \end{vmatrix}$

3.- Demostrar que

$$\begin{vmatrix} x-a & d & x \\ y-b & d & y \\ z-c & d & z \end{vmatrix} = \begin{vmatrix} x & d & a \\ y & d & b \\ z & d & c \end{vmatrix}$$

4.- Proponer dos matrices A y B de 2×2 tales que

$$\det(A + B) \neq \det A + \det B$$

y verificar que para dichas matrices

$$\det(AB) = (\det A)(\det B)$$

VII.3 CALCULO DE DETERMINANTES

Como el lector se habrá dado cuenta a lo largo de la sección VII.1, el empleo de la definición VII.1.4 para calcular el valor de un determinante resulta un proceso demasiado laborioso; especialmente en los casos de ejemplos numéricos, en los cuales se requiere obtener primero la expresión del desarrollo correspondiente al orden de la matriz en cuestión, para posteriormente identificar los números que corresponden a cada uno de los elementos en el desarrollo y efectuar las operaciones indicadas.

Es por ello que no se acostumbra en la práctica el empleo de la definición para el cálculo de determinantes; a cambio se han desarrollado otros métodos cuya aplicación resulta más simple y que conducen a los mismos resultados.

- Regla de Sarrus

El más sencillo de tales métodos es el que se conoce como "regla de Sarrus", el cual posiblemente habrá sido utilizado ya por el lector. Este método se emplea para calcular determinantes de segundo y de tercer orden.

Para calcular el valor de un determinante de segundo orden empleando la regla de Sarrus, se efectúa el producto de los elementos de la diagonal principal y a éste se resta el producto de los elementos de la "diagonal secundaria". En forma esquemática:

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

como se ve, el resultado que arroja la regla de Sarrus coincide con la definición de determinante de segundo orden.

Así, por ejemplo, para calcular el determinante de la matriz

$$A = \begin{bmatrix} 5 & 4 \\ -2 & -3 \end{bmatrix}$$

empleando la regla de Sarrus se procede como sigue

$$\det A = \begin{vmatrix} 5 & 4 \\ -2 & -3 \end{vmatrix} = (5)(-3) - (-2)(4) = -15 + 8 = -7$$

con lo que se llega al mismo resultado que se obtuvo en la sec - ción VII.1.

Para calcular el valor de un determinante de tercer orden empleando la regla de Sarrus, se efectúa el producto de los elementos de la diagonal principal y de las dos "diagonales paralelas" a ella; el término "diagonales paralelas" se debe a que, cuando se emplea el artificio que consiste en volver a escribir los dos primeros renglones a continuación del tercero, los elementos en cuestión aparecen formando "diagonales" paralelas a la principal. A la suma de dichos productos se restan los productos de los elementos de la diagonal secundaria y de las dos "paralelas" a ella. En forma esquemática:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{matrix} \\ \\ \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix} \quad \text{ó} \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+$$

$$+ a_{31}a_{23}a_{12}-a_{31}a_{22}a_{13}-$$

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

El desarrollo anterior coincide, salvo en el orden de algunos factores y términos, con la definición de determinante de tercer or - den.

A manera de ejemplo, calculemos el determinante de la matriz

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

empleando la regla de Sarrus

$$\det A = \begin{vmatrix} 1 & 0 & -3 \\ -4 & 5 & 2 \\ 1 & -2 & 0 \end{vmatrix} = (1)(5)(0)+(-4)(-2)(-3)+(1)(2)(0)$$

$$- (1)(5)(-3)-(-2)(2)(1)-(0)(0)(-4) =$$

$$= 0 -24 + 0 + 15 + 4 -0 = -5$$

se obtiene así el mismo resultado de la sección VII.1.

Es importante subrayar que la regla de Sarrus sólo se aplica a determinantes de segundo y de tercer orden. En ocasiones se pretende erróneamente "generalizar" esta regla para calcular determinantes de orden mayor; sin embargo, se puede comprobar fácilmente que al aplicar la supuesta "regla de Sarrus" a un determinante de or- den superior al tercero se obtiene un desarrollo que no coincide con el de la definición.

El método que veremos a continuación, conocido como "desarrollo por cofactores", no sólo es aplicable al cálculo de determinantes de cualquier orden, sino que constituye el fundamento de todos los métodos de aplicación práctica.

— Desarrollo por cofactores

Con el propósito de familiarizar al lector con las ideas fundamentales del desarrollo por cofactores, consideremos nuevamente el desarrollo del determinante de tercer orden obtenido al final de la sección VII.1:

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

Como en cada término hay un elemento de cada renglón y de cada columna, podemos seleccionar una línea cualquiera y factorizar los elementos de ésta. Por ejemplo, eligiendo el primer renglón podemos factorizar sus elementos y escribir

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

Cada uno de los factores que multiplican a los elementos del primer renglón en la expresión anterior constituye el desarrollo de un determinante de segundo orden. Así:

para a_{11} tenemos que

$$(a_{22}\ a_{33} - a_{23}\ a_{32}) = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

para a_{12} tenemos que

$$(-a_{21}\ a_{33} + a_{23}\ a_{31}) = \begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} = (-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

para a_{13} tenemos que

$$(a_{21}\ a_{32} - a_{22}\ a_{31}) = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Cada uno de estos determinantes puede ser obtenido de la matriz original suprimiendo el renglón y la columna en que se encuentra el elemento correspondiente. Tales determinantes reciben el nombre de "menores".

Así, por ejemplo, el menor de a_{12} se puede obtener de la siguiente manera

$$\begin{bmatrix} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ a_{21} & \cancel{a_{22}} & a_{23} \\ a_{31} & \cancel{a_{32}} & a_{33} \end{bmatrix} \rightarrow \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

Regresando a la expresión anterior para det A, vemos que los factores que multiplican a los elementos del primer renglón no son, en todos los casos, los menores correspondientes.

En el caso de a_{12} dicho factor es igual al menor con el signo cambiado. Esto se identifica con el hecho de que tal elemento es de "característica impar"; es decir que la suma del número del renglón y de la columna en que se encuentra es un número impar - - - $(1 + 2 = 3)$.

Sólo tenemos un elemento de característica impar en el desarrollo anterior por haber elegido el primer renglón para factorizar sus elementos. Si hubiésemos elegido el segundo renglón tendríamos dos elementos de característica impar (a_{21} y a_{23}) y, en tal caso, los factores que multiplican a éstos en el desarrollo del determinante serían iguales a sus correspondientes menores con el signo cambiado.

Surge así el concepto de "cofactor", como el factor que multiplica al elemento en el desarrollo del determinante. Dicho cofactor es igual al menor, o al negativo de éste, según sea par o impar la característica del elemento.

Finalmente, el determinante de A puede expresarse como

$$\det A = a_{11}(-1)^{1+1}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^{1+2}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} +$$

$$+ a_{13}(-1)^{1+3}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Expresión que se conoce como el "desarrollo por cofactores según el primer renglón". Es claro que pueden obtenerse desarrollos similares para cada uno de los otros renglones y columnas de la matriz A.

VII.3.1 DEFINICION

Sea $A = \left[a_{ij}\right]$ una matriz de $n \times n$ con elementos en C.

i) Se llama menor del elemento a_{ij}, y se representa con M_{ij}, al determinante de la matriz que se obtiene suprimiendo en A el renglón i y la columna j.

ii) Se llama cofactor del elemento a_{ij}, y se representa con C_{ij}, al producto $(-1)^{i+j}M_{ij}$.

Así, por ejemplo, para la matriz

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

el menor del elemento a_{23} es el determinante

$$M_{23} = \begin{vmatrix} 1 & -3 & 0 \\ 0 & -1 & 1 \\ 3i & 0 & -i \end{vmatrix} = i - 9i = -8i$$

mientras que su cofactor es el producto

$$C_{23} = (-1)^{2+3}\begin{vmatrix} 1 & -3 & 0 \\ 0 & -1 & 1 \\ 3i & 0 & -i \end{vmatrix} = (-1)(-8i) = 8i$$

Para el elemento a_{24} se tiene

$$M_{24} = \begin{vmatrix} 1 & -3 & 2i \\ 0 & -1 & i \\ 3i & 0 & 2 \end{vmatrix} = -2 + 9 - 6 = 1$$

y

$$C_{24} = (-1)^6 (1) = (1)(1) = 1$$

etc.

VII.3.2 TEOREMA

Si $A = \left[a_{ij}\right]$ es una matriz de $n \times n$ con elementos en C y r es un número entero tal que $1 \leqslant r \leqslant n$, entonces:

i) $\det A = \sum\limits_{j=1}^{n} a_{rj}\, C_{rj}$

ii) $\det A = \sum\limits_{i=1}^{n} a_{ir}\, C_{ir}$

DEMOSTRACION

Se demuestra primero la parte i)

De VII.1.4 se tiene que

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}} \qquad (1)$$

y en consecuencia, cada término del desarrollo de det A tiene uno y sólo un factor $a_{r\alpha_{kr}}$ correspondiente al renglón r, cuyo segundo índice α_{kr} recorre todos los valores 1, 2,..., n a lo largo de los n! términos. Podemos en consecuencia escribir

$$\det A = a_{r1}\, F_{r1} + a_{r2}\, F_{r2} + \ldots + a_{rn}\, F_{rn} \qquad (2)$$

y debemos probar que los F_{rj} $(j = 1, 2, \ldots, n)$ son los cofactores de los elementos a_{rj} en el sentido de la definición VII.3.1.

Como el factor F_{rj} se obtiene sumando los términos del desarrollo (1) en los cuales aparece el elemento a_{rj} y factorizando éste, se tendrá que F_{rj} es una suma de productos formados por n-1 factores entre los cuales no hay elementos del renglón r ni de la columna j; esto es

$$F_{rj} = \sum_{k=1}^{(n-1)!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots$$

$$\cdots\, a_{(r-1)\alpha_{k(r-1)}}\, a_{(r+1)\alpha_{k(r+1)}} \cdots a_{n\alpha_{kn}} \qquad (3)$$

Donde $q_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{k(r-1)}, \alpha_{k(r+1)}, \ldots, \alpha_{kn})$ es una permutación del conjunto $\{1, 2, \ldots, j-1, j+1, \ldots, n\}$ que se obtiene suprimiendo en $p_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kr}, \ldots, \alpha_{kn})$ el número α_{kr} (que es igual a j en los términos correspondientes a F_{rj}).

En consecuencia, con la posible excepción del signo $\varepsilon(p_k)$ la expresión (3) corresponde al desarrollo del menor M_{rj}, de acuerdo con la definición VII.3.1

Para mostrar la relación que existe entre $\varepsilon(p_k)$ y $\varepsilon(q_k)$, consideremos la permutación

$$p_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kr}, \ldots, \alpha_{kn})$$

y traslademos a α_{kr} hasta el último lugar por medio de n-r intercambios de números adyacentes, se obtiene así la permutación

$$p'_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn}, \alpha_{kr})$$

cuyo signo, respecto al de p_k, está dado por

$$\varepsilon(p'_k) = (-1)^{n-r}\, \varepsilon(p_k)$$

Por otra parte, los primeros n-1 números de p'_k coinciden con la permutación

$$q_k = (\alpha_{k_1}, \alpha_{k_2}, \dots, \alpha_{kn})$$

y como $\alpha_{kr} = j$, en p_k' habrá n-j números mayores que α_{kr} y que se encuentran antes que éste en la permutación, por lo que p_k' tiene n-j inversiones más que q_k y, en consecuencia

$$\varepsilon(q_k) = (-1)^{n-j} \varepsilon(p_k')$$

por lo que, respecto al signo de p_k tendremos

$$\varepsilon(q_k) = (-1)^{n-j} (-1)^{n-r} \varepsilon(p_k) = (-1)^{2n-r-j} \varepsilon(p_k) = (-1)^{-(r+j)} \varepsilon(p_k)$$

o lo que es equivalente

$$\varepsilon(p_k) = (-1)^{r+j} \varepsilon(q_k)$$

Finalmente, llevando este resultado al desarrollo (3) se obtiene

$$F_{rj} = (-1)^{r+j} M_{rj}$$

por lo que, de VII.3.1

$$F_{rj} = C_{rj}$$

como se quería.

La parte ii) del teorema se sigue inmediatamente de la parte i) del mismo y de la propiedad i) del teorema VII.2.3.

□

De acuerdo con el teorema VII.3.2, el valor de un determinante puede obtenerse a partir de los elementos de una cualquiera de sus líneas, sumando los productos de éstos por sus respectivos cofactores.

Por ejemplo, para calcular el determinante de la matriz

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

Podemos elegir cualquier renglón o columna para desarrollar por cofactores. En este caso, un somero análisis de la matriz dada nos sugiere la elección del tercer renglón en virtud de que dos de sus elementos son nulos y el producto de éstos por sus respectivos cofactores será igual a cero, sea cual fuere el valor de los cofactores, por lo que no será necesario calcularlos.

Así, se tendrá que

$$\det A = 0\times C_{31} + 2\times C_{32} + (-1)\times C_{33} + 0\times C_{34} = 2C_{32} - C_{33}$$

Para calcular los menores correspondientes podemos emplear la regla de Sarrus por tratarse de determinantes de tercer orden.

$$M_{32} = \begin{vmatrix} 2 & -5 & 2 \\ 4 & 0 & 1 \\ -1 & -7 & 1 \end{vmatrix} = 0 - 56 + 5 - 0 + 14 + 20 = -17$$

$$M_{33} = \begin{vmatrix} 2 & 1 & 2 \\ 4 & -6 & 1 \\ -1 & 6 & 1 \end{vmatrix} = -12 + 48 - 1 - 12 - 12 - 4 = 7$$

Los cofactores serán entonces

$$C_{32} = 17 \quad \text{y} \quad C_{33} = 7$$

por lo que

$$\det A = 2\times 17 - 7 = 27$$

En el caso general, el desarrollo por cofactores transforma el problema de calcular un determinante de orden n en el de calcular n determinantes de orden n-1. Cada uno de estos determinantes puede desarrollarse a su vez por cofactores, obteniéndose menores de orden n-2, y así sucesivamente. Se acostumbra continuar el proceso hasta obtener menores de orden 3 o de orden 2, cuyo valor puede obtenerse empleando la regla de Sarrus.

Condensación

El desarrollo por cofactores, aunque de aplicación general, no es un método eficiente dada la gran cantidad de determinantes de orden menor que se requiere calcular. Por ejemplo, para un determinante de quinto orden se tendrán 5 menores de cuarto orden, y para cada uno de ellos 4 menores de tercer orden; por lo que se requiere calcular un total de 20 determinantes de tercer orden.

Como se vio en el ejemplo anterior, en ciertos casos especiales puede evitarse el cálculo de algunos cofactores cuando los elementos correspondientes son nulos; en especial, si alguna de las líneas tuviese todos los elementos excepto uno iguales a cero, sin duda escogeríamos dicha línea para efectuar el desarrollo por cofactores, ya que sólo requeriríamos calcular uno de ellos (el correspondiente al elemento distinto de cero). Esta situación, claramente deseable, puede lograrse con ayuda de las propiedades que establece el teorema VII.2.2; en particular, mediante la aplicación reiterada de la propiedad v).

El método que se propone a continuación, conocido como "método de condensación", se basa precisamente en esta idea y consiste en lo siguiente:

1) Elegir una línea que contenga el mayor número de ceros posible.

2) Elegir un elemento no nulo de dicha línea (de preferencia un 1 o un -1) y aplicar reiteradamente la propiedad v) de VII.2.2 hasta reducir a cero todos los demás elementos de la línea.

3) Desarrollar por cofactores según dicha línea.

4) Repetir los tres pasos anteriores hasta obtener un determinante de tercer orden (o de segundo si se prefiere) y obtener su valor mediante la regla de Sarrus.

A manera de ejemplo, usaremos a continuación el método de condensación para calcular el determinante de la matriz

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

Elegimos la cuarta columna para efectuar el desarrollo por tener ésta dos elementos nulos, y seleccionamos como "pivote" al tercer elemento de dicha columna por tratarse de un uno. Entonces, multiplicando por 2 y por -3 el tercer renglón y sumándolo al primero y al cuarto renglones, respectivamente, se obtiene

$$
\det A = \begin{array}{c|ccccc|c}
 & & & & \downarrow & & \\
 & -1 & 1 & -5 & -2 & 3 & (2) \\
 & 3 & 2 & 1 & 0 & -1 & \\
 & 1 & -1 & 2 & 1 & 0 & * \\
 & 0 & 2 & 1 & 3 & -1 & (-3) \\
 & 1 & 2 & 4 & 0 & 1 &
\end{array}
= \begin{vmatrix} 1 & -1 & -1 & 0 & 3 \\ 3 & 2 & 1 & 0 & -1 \\ 1 & -1 & 2 & 1 & 0 \\ -3 & 5 & -5 & 0 & -1 \\ 1 & 2 & 4 & 0 & 1 \end{vmatrix}
$$

En consecuencia, desarrollando por cofactores según la cuarta columna

$$
\det A = (1)(-1)^7 \begin{vmatrix} 1 & -1 & -1 & 3 \\ 3 & 2 & 1 & -1 \\ -3 & 5 & -5 & -1 \\ 1 & 2 & 4 & 1 \end{vmatrix}
$$

Para calcular el valor de éste determinante de cuarto orden, elegimos ahora el primer renglón para el desarrollo y la primera columna como pivote, por lo que sumando ésta a la segunda y tercera columnas, multiplicándola por -3 y sumándola a la cuarta se obtiene

$$
\det A = (-1) \begin{array}{c|cccc|}
\rightarrow & 1 & -1 & -1 & 3 \\
 & 3 & 2 & 1 & -1 \\
 & -3 & 5 & -5 & -1 \\
 & 1 & 2 & 4 & 1 \\
\multicolumn{1}{c}{} & \multicolumn{1}{c}{*} & (1) & (1) & \multicolumn{1}{c}{(-3)}
\end{array}
= (-1) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 3 & 5 & 4 & -10 \\ -3 & 2 & -8 & 8 \\ 1 & 3 & 5 & -2 \end{vmatrix}
$$

y desarrollando por cofactores según el primer renglón

$$
\det A = (-1)(1)(-1)^2 \begin{vmatrix} 5 & 4 & -10 \\ 2 & -8 & 8 \\ 3 & 5 & -2 \end{vmatrix}
$$

Para continuar con el proceso, seleccionamos ahora el segundo renglón para el desarrollo y la primera columna como pivote, con lo que

$$
\det A = (-1) \begin{array}{c|ccc|}
 & 5 & 4 & -10 \\
\rightarrow & 2 & -8 & 8 \\
 & 3 & 5 & -2 \\
\multicolumn{1}{c}{} & \multicolumn{1}{c}{*} & (4) & \multicolumn{1}{c}{(-4)}
\end{array}
= (-1) \begin{vmatrix} 5 & 24 & -30 \\ 2 & 0 & 0 \\ 3 & 17 & -14 \end{vmatrix}
$$

Finalmente, desarrollando por cofactores y empleando la regla de Sarrus se obtiene

$$
\det A = (-1)(2)(-1)^3 \begin{vmatrix} 24 & -30 \\ 17 & -14 \end{vmatrix} = 2(-336 + 510) = 348
$$

que es el valor del determinante.

Como puede verse, el método de condensación ofrece en cada ciclo un gran número de posibilidades para la selección de la línea y del elemento pivote. Una selección adecuada en cada caso puede contribuir notablemente a simplificar los cálculos correspondientes.

- Determinante de una matriz triangular.

El cálculo del determinante de una matriz triangular resulta particularmente sencillo, ya que su valor es igual al producto de los elementos de su diagonal principal, como lo establece el siguiente teorema.

VII.3.3 TEOREMA

Si $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ es una matriz triangular superior (inferior), entonces

$$\det A = \prod_{i=1}^{n} a_{ii}$$

DEMOSTRACION

Haremos una prueba por inducción en el caso de una matriz triangular superior.

Para n = 2 se tiene

$$A_2 = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}$$

y por la regla de Sarrus

$$\det A_2 = a_{11}a_{22} - 0 = a_{11}a_{22}$$

Sea ahora

$$A_n = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Desarrollando por cofactores según el último renglón se tiene que

$$\det A_n = a_{nn}(-1)^{n+n} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ 0 & a_{22} & \cdots & a_{2,n-1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & a_{n-1,n-1} \end{vmatrix}$$

pero, por hipótesis de inducción

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ 0 & a_{22} & \cdots & a_{2,n-1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & a_{n-1,n-1} \end{vmatrix} = a_{11}a_{22} \cdots a_{n-1,n-1}$$

y en consecuencia

$$\det A_n = a_{nn}(a_{11}\ a_{22} \cdots a_{n-1,n-1}) = \prod_{i=1}^{n} a_{ii}$$

Por otra parte, si A es triangular inferior entonces A^T es triangular superior; en consecuencia por i) de VII.2.3 y el resultado anterior se tendrá que

$$\det A = \det A^T = \prod_{i=1}^{n} a_{ii}$$

Esto completa la demostración. □

La facilidad con que se calcula el determinante de una matriz triangular sugiere otro método general para el cálculo de determinantes, el cual consiste en transformar la matriz dada en una matriz triangular empleando las propiedades del teorema VII.2.2, y calcular el determinante de ésta multiplicando los elementos de su diagonal principal. Debido a que una matriz triangular es una matriz escalonada, el procedimiento resulta similar al método de Gauss para la resolución de sistemas; sin embargo, se debe tener presente que en este caso las transformaciones del tipo I y del tipo II pueden alterar el valor del determinante.

A manera de ejemplo, calcularemos a partir de este método el valor del determinante que empleamos para ilustrar el método de conden -

sación.

Multiplicando por 3 el primer renglón y sumándolo al segundo, y sumando el primer renglón al tercero y al quinto se obtiene

$$\det A = \begin{vmatrix} -1 & 1 & -5 & -2 & 3 \\ 3 & 2 & 1 & 0 & -1 \\ 1 & -1 & 2 & 1 & 0 \\ 0 & 2 & 1 & 3 & -1 \\ 1 & 2 & 4 & 0 & 1 \end{vmatrix} = \begin{vmatrix} -1 & 1 & -5 & -2 & 3 \\ 0 & 5 & -14 & -6 & 8 \\ 0 & 0 & -3 & -1 & 3 \\ 0 & 2 & 1 & 3 & -1 \\ 0 & 3 & -1 & -2 & 4 \end{vmatrix}$$

sumando al segundo renglón el quinto multiplicado por -2 y, a continuación, multiplicando por 2 y por 3 el segundo renglón y sumando al cuarto y quinto renglones, respectivamente, se obtiene

$$\det A = \begin{vmatrix} -1 & 1 & -5 & -2 & 3 \\ 0 & -1 & -12 & -2 & 0 \\ 0 & 0 & -3 & -1 & 3 \\ 0 & 2 & 1 & 3 & -1 \\ 0 & 3 & -1 & -2 & 4 \end{vmatrix} = \begin{vmatrix} -1 & 1 & -5 & -2 & 3 \\ 0 & -1 & -12 & -2 & 0 \\ 0 & 0 & -3 & -1 & 3 \\ 0 & 0 & -23 & -1 & -1 \\ 0 & 0 & -37 & -8 & 4 \end{vmatrix}$$

Multiplicando el tercer renglón por $-\frac{23}{3}$ y por $-\frac{37}{3}$, y sumándolo al cuarto y al quinto renglones y, a continuación, multiplicando por $-\frac{13}{20}$ el cuarto renglón y sumándolo al quinto se obtiene

$$\det A = \begin{vmatrix} -1 & 1 & -5 & -2 & 3 \\ 0 & -1 & -12 & -2 & 0 \\ 0 & 0 & -3 & -1 & 3 \\ 0 & 0 & 0 & \frac{20}{3} & -24 \\ 0 & 0 & 0 & \frac{13}{3} & -33 \end{vmatrix} = \begin{vmatrix} -1 & 1 & -5 & -2 & 3 \\ 0 & -1 & -12 & -2 & 0 \\ 0 & 0 & -3 & -1 & 3 \\ 0 & 0 & 0 & \frac{20}{3} & -24 \\ 0 & 0 & 0 & 0 & -\frac{87}{5} \end{vmatrix}$$

por lo que, finalmente

$$\det A = (-1)(-1)(-3)\left(\frac{20}{3}\right)\left(\frac{-87}{5}\right) = 348.$$

que es el valor buscado.

VII.3.4 EJERCICIOS

1.- Aplicar la regla de Sarrus para calcular el valor de los siguientes determinantes

a) $\begin{vmatrix} a & -3a \\ 2 & 1 \end{vmatrix}$ b) $\begin{vmatrix} 3 & b & -c \\ -2 & 1 & -1 \\ -9 & c & 3b \end{vmatrix}$

2.- Obtener el menor y el cofactor de los elementos a_{22} y a_{23} de la matriz

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

3.- Obtener el valor del determinante

$$\begin{vmatrix} 4 & -6 & 0 & 1 \\ 0 & 2 & -1 & 0 \\ 2 & 1 & -5 & 2 \\ 1 & 6 & -7 & 1 \end{vmatrix}$$

a) desarrollando por cofactores según la tercera columna

b) aplicando lo que sería la regla de Sarrus para un determinante de 4o. orden.

Y comparar estos resultados con el obtenido en el problema 4 de VII.1.5

4.- Calcular el determinante del problema anterior, desarrollando por cofactores según el segundo renglón.

5.- Calcular el valor del determinante

$$\begin{vmatrix} 2 & 1 & 0 & 2 & -1 \\ 1 & 0 & -2 & 0 & 3 \\ -4 & -1 & 3 & 1 & -5 \\ 1 & 2 & 0 & 4 & 3 \\ -2 & 1 & 4 & 0 & 1 \end{vmatrix}$$

a) por el método de condensación

b) transformando a una matriz triangular.

6.- Demostrar que el valor de

$$\begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ b_1 & b_2 & a_{33} & 0 \\ c_1 & c_2 & c_3 & a_{44} \end{vmatrix}$$

es independiente de los valores de b_1, b_2, c_1, c_2 y c_3

VII.4 ALGUNAS APLICACIONES

Como complemento a las secciones anteriores, en las cuales se presentaron los principales aspectos relativos al concepto de determinante, en ésta veremos dos de sus principales aplicaciones: al cálculo de la inversa de una matriz y a la resolución de sistemas de ecuaciones lineales. Como consecuencia del desarrollo de estos tópicos se obtendrán además algunas propiedades adicionales que son de gran utilidad en el empleo de los determinantes.

- Cálculo de la inversa por medio de la adjunta

Se conoce como adjunta de una matriz cuadrada A a la transpuesta de la matriz que se obtiene reemplazando los elementos de A por sus respectivos cofactores, como lo establece la siguiente definición

VII.4.1 DEFINICION

Sea $A = \left[a_{ij}\right]$ una matriz de $n \times n$ con elementos en C, y sea C_{ij} el cofactor del elemento a_{ij}. Se llama Adjunta de A a la matriz

$$\text{Adj } A = \left[b_{ij}\right] \text{, donde } b_{ij} = C_{ji}$$

Así, por ejemplo, para obtener la adjunta de la matriz

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

se calculan los cofactores de todos sus elementos

$$C_{11} = \begin{vmatrix} -1 & 4 \\ 1 & 0 \end{vmatrix} = -4, \quad C_{12} = (-1)\begin{vmatrix} 3 & 4 \\ 2 & 0 \end{vmatrix} = 8, \quad C_{13} = \begin{vmatrix} 3 & -1 \\ 2 & 1 \end{vmatrix} = 5,$$

$$C_{21} = (-1)\begin{vmatrix} 0 & 2 \\ 1 & 0 \end{vmatrix} = 2, \quad C_{22} = \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} = -4, \quad C_{23} = (-1)\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} = -1,$$

$$C_{31} = \begin{vmatrix} 0 & 2 \\ -1 & 4 \end{vmatrix} = 2, \quad C_{32} = (-1)\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 2, \quad C_{33} = \begin{vmatrix} 1 & 0 \\ 3 & -1 \end{vmatrix} = -1,$$

y se ordenan de la siguiente manera

$$\text{Adj } A = \begin{bmatrix} C_{ji} \end{bmatrix} = \begin{bmatrix} -4 & 2 & 2 \\ 8 & -4 & 2 \\ 5 & -1 & -1 \end{bmatrix}$$

La adjunta tiene la siguiente propiedad importante. Si multiplicamos la matriz A del ejemplo anterior por su adjunta obtenemos

$$A(\text{Adj } A) = \begin{bmatrix} 1 & 0 & 2 \\ 3 & -1 & 4 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} -4 & 2 & 2 \\ 8 & -4 & 2 \\ 5 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} = 6\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y cabe preguntarse que relación tiene el número 6 con la matriz dada. Si calculamos su determinante encontramos que

$$\det A = \begin{vmatrix} 1 & 0 & 2 \\ 3 & -1 & 4 \\ 2 & 1 & 0 \end{vmatrix} = 6$$

Es decir que

$$A(\text{Adj } A) = (\det A) I_3$$

Esta expresión no sólo es válida para la matriz del ejemplo anterior sino que constituye un resultado general, como lo establece el siguiente teorema

VII.4.2 TEOREMA

Si A es una matriz de n×n con elementos en C, entonces

$$A(\text{Adj } A) = (\text{Adj } A)A = (\det A)I_n$$

DEMOSTRACION

Sea $A = \begin{bmatrix} a_{ij} \end{bmatrix}$.

De VII.4.1

$$\text{Adj } A = \begin{bmatrix} b_{ij} \end{bmatrix}, \text{ donde } b_{ij} = C_{ji}$$

por lo que, de VI.3.1

$$A(\text{Adj } A) = \begin{bmatrix} p_{ij} \end{bmatrix}$$

donde $p_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} a_{ik} C_{jk}$

En consecuencia, para $i = j$ se tiene que

$$p_{ij} = p_{ii} = \sum_{k=1}^{n} a_{ik} C_{ik}$$

que es el desarrollo por cofactores, según el i-ésimo renglón, del determinante de A; por lo que, de i) de VII.3.2

$$p_{ii} = \det A \qquad (1)$$

Probaremos ahora que $p_{ij} = 0$ cuando $i \neq j$

Por i) de VII.3.2 se tiene que

$$\begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{j1}C_{j1} + a_{j2}C_{j2} + \cdots + a_{jn}C_{jn}$$

por lo que, haciendo $a_{jk} = a_{ik}$ tenemos

$$\begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}$$

Pero este determinante tiene dos renglones iguales por lo que, de iv) de VII.2.2, su valor es cero y

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn} = 0$$

Esta expresión nos indica que la suma de los productos de los elementos de un renglón por los cofactores de los elementos de otro renglón es igual a cero.

En consecuencia

$$p_{ij} = \sum_{k=1}^{n} a_{ik}C_{jk} = 0, \quad \text{si} \quad i \neq j \qquad (2)$$

Así que, de (1) y (2)

$$p_{ij} = \begin{cases} \det A, & \text{si} \quad i = j \\ 0 \quad, & \text{si} \quad i \neq j \end{cases}$$

por lo que, de VI.3.4 y VI.2.4

$$A(\text{Adj } A) = (\det A)I_n$$

La prueba de

$$(\text{Adj } A)A = (\det A)I_n$$

es similar y se sugiere al lector intentarla como ejercicio. □

Con ayuda de este resultado puede demostrarse el importante teorema que se enuncia a continuación

VII.4.3 TEOREMA

Sea A una matriz de n×n con elementos en C:

A^{-1} existe si y sólo si $\det A \neq 0$

DEMOSTRACION

a) De VII.4.2 se tiene que

$$A(\text{Adj } A) = (\det A)I_n$$

por lo que, si $\det A \neq 0$ de la expresión anterior se sigue que

$$\frac{1}{\det A}\left[A(\text{Adj } A)\right] = I_n$$

esto es

$$A \left[\frac{1}{\det A}(\text{Adj } A) \right] = I_n$$

En consecuencia, de VI.4.1

$$A^{-1} = \frac{1}{\det A}(\text{Adj } A) \qquad (1)$$

y A^{-1} existe, ya que la adjunta de A existe para toda matriz A.

Entonces

$$\det A \neq 0 \Rightarrow \exists A^{-1} \qquad (2)$$

b) Si A^{-1} existe, entonces

$$A A^{-1} = I_n \qquad \text{por VI.4.1}$$

$$\det(A A^{-1}) = \det I_n$$

$$(\det A)(\det A^{-1}) = \det I_n \qquad \text{por iii) de VII.2.3}$$

$$(\det A)(\det A^{-1}) = 1 \qquad \text{por VII.3.3} \qquad (3)$$

por lo que $\det A \neq 0$; esto es

$$\exists A^{-1} \Rightarrow \det A \neq 0 \qquad (4)$$

Finalmente, de (2) y (4) se sigue que

$$\exists A^{-1} \Longleftrightarrow \det A \neq 0$$

como se quería demostrar.

□

La expresión (1) de la demostración anterior sugiere un método para calcular la inversa de una matriz, el cual consiste en multiplicar el recíproco del determinante por la adjunta. Puntualizamos dicho resultado en el siguiente corolario

VII.4.4 COROLARIO

si $\det A \neq 0$, entonces $A^{-1} = \frac{1}{\det A}(\text{Adj } A)$

Así, por ejemplo, la inversa de la matriz

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

es

$$A^{-1} = \frac{1}{6} \begin{bmatrix} -4 & 2 & 2 \\ 8 & -4 & 2 \\ 5 & -1 & -1 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{4}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \end{bmatrix}$$

Otro resultado importante que puede derivarse de la demostración del teorema VII.4.3 es el siguiente.

VII.4.5 COROLARIO

si $\exists A^{-1}$ entonces $\det A^{-1} = \frac{1}{\det A}$

Esta propiedad es consecuencia inmediata de la expresión (3).

- Regla de Cramer

Para terminar este capítulo enunciaremos formalmente el método para resolver sistemas de ecuaciones lineales sugerido al principio de la sección VII.1, el cual nos sirvió como motivación para la definición general de determinante.

VII.4.6 TEOREMA (REGLA DE CRAMER)

Sea

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots&\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

un sistema de n ecuaciones lineales con n incógnitas, y sea $A = \left[a_{ij}\right]$ su matriz de coeficientes.

Si $\det A \neq 0$ entonces $x_k = \dfrac{\det A_k}{\det A}$, $(k = 1, 2, \ldots, n)$

donde $A_k = \left[c_{ij}\right]$ es tal que $c_{ij} = \begin{cases} a_{ij}, & \text{para } j \neq k \\ b_i, & \text{para } j = k \end{cases}$

Es decir que, si $\det A \neq 0$, entonces el valor de la k-ésima incógnita en la solución del sistema puede calcularse como el cociente de los determinantes de las matrices A_k y A, donde A_k se obtiene reemplazando en A la k-ésima columna por el vector de términos independientes.

DEMOSTRACION

El sistema es equivalente a la expresión

$$AX = B$$

donde

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad y \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Entonces, si $\det A \neq 0$ por VII.4.3 $\exists\ A^{-1}$ y se tendrá que

$$X = A^{-1}B$$

por lo que, de VII.4.4

$$X = \left[\frac{1}{\det A}(\text{Adj } A)\right] B$$

Esto es

$$X = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \ldots & C_{n1} \\ C_{12} & C_{22} & \ldots & C_{n2} \\ \vdots & & & \vdots \\ C_{1n} & C_{2n} & \ldots & C_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$X = \frac{1}{\det A}\begin{bmatrix} b_1C_{11} + b_2C_{21} + \ldots + b_nC_{n1} \\ b_1C_{12} + b_2C_{22} + \ldots + b_nC_{n2} \\ \vdots \\ b_1C_{1n} + b_2C_{2n} + \ldots + b_nC_{nn} \end{bmatrix}$$

Así que

$$x_k = \frac{1}{\det A}(b_1C_{1k} + b_2C_{2k} + \ldots + b_nC_{nk}), \quad (k = 1, 2, \ldots, n)$$

pero, por VII.3.2, la expresión

$$b_1C_{1k} + b_2C_{2k} + \ldots + b_nC_{nk}$$

corresponde al desarrollo por cofactores, según la k-ésima columna, de una matriz que se obtiene a partir de A reemplazando los elementos de la k-ésima columna por el vector de términos independientes; es decir que

$$x_k = \frac{1}{\det A}(\det A_k), \text{ donde } A_k = \left[c_{ij} \right] \text{ con } c_{ij} = \begin{cases} a_{ij}, & \text{para } j \neq k \\ b_i, & \text{para } j = k \end{cases}$$

lo que demuestra el teorema. $\square$

Así, por ejemplo, para resolver el siguiente sistema de ecuaciones lineales empleando la regla de Cramer

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

calculamos primero

$$\det A = \begin{vmatrix} 2 & 3 & -1 \\ -1 & -1 & 2 \\ 1 & 2 & -1 \end{vmatrix} = -2$$

Como $\det A \neq 0$ calculamos ahora

$$\det A_1 = \begin{vmatrix} 1 & 3 & -1 \\ 3 & -1 & 2 \\ -2 & 2 & -1 \end{vmatrix} = -10$$

$$\det A_2 = \begin{vmatrix} 2 & 1 & -1 \\ -1 & 3 & 2 \\ 1 & -2 & -1 \end{vmatrix} = 4$$

$$\det A_3 = \begin{vmatrix} 2 & 3 & 1 \\ -1 & -1 & 3 \\ 1 & 2 & -2 \end{vmatrix} = -6$$

por lo que la solución es

$$x_1 = \frac{-10}{-2} = 5, \qquad x_2 = \frac{4}{-2} = -2, \qquad x_3 = \frac{-6}{-2} = 3$$

VII.4.7 EJERCICIOS

1.- Obtener la matriz A para la cual

$$\text{Adj } A = \begin{bmatrix} 1 & -2 & 7 \\ -3 & 0 & -3 \\ 1 & -2 & 1 \end{bmatrix} \qquad \text{y} \qquad A^T = \begin{bmatrix} 1 & a_{21} & a_{31} \\ 2 & a_{22} & a_{32} \\ -1 & 3 & 1 \end{bmatrix}$$

2.- Para cada una de las siguientes matrices determinar si existe su inversa, en caso afirmativo calcularla por el método de la adjunta y verificar que se cumple el corolario VII.4.5

a)
$$A = \begin{bmatrix} 1 & -1 & 0 \\ 3 & 2 & 1 \\ 0 & -2 & 1 \end{bmatrix} \quad \begin{matrix} 4/7 & 1/7 & -1/7 \\ -3/7 & 1/7 & -1/7 \\ -6/7 & 2/7 & 5/7 \end{matrix}$$

b)
$$B = \begin{bmatrix} -2 & 0 & 1 \\ 3 & 5 & 2 \\ -1 & 5 & 4 \end{bmatrix}$$

3.- Demostrar que si A es una matriz no singular de 3×3, entonces

$$\det A = \sqrt{\det(\text{Adj } A)}$$

6.- Una compañía recibió de cuatro proveedores los siguientes presupuestos:

	Prov. A	Prov. B	Prov. C	Prov. D
No. compresoras	2	0	1	1
No. medidores	0	4	2	2
No. válvulas	4	4	0	4
No. reguladores	1	2	2	0
Costo total en millones de pesos	6	5	5	4

Sabiendo que los proveedores tienen los mismos precios para cada artículo; ¿Cuántos millones de pesos cuesta cada impresora? ¿Qué ventajas presenta en este caso el empleo de la regla de Cramer sobre el método de Gauss?

CAPITULO VIII ESTRUCTURAS ALGEBRAICAS

INTRODUCCION

En capítulos anteriores hemos tratado con diferentes tipos de entes matemáticos tales como números complejos, polinomios y matrices, y hemos efectuado con ellos ciertas operaciones; sin embargo, no todas las operaciones se comportaron de la misma manera: la multiplicación de polinomios, por ejemplo, resultó ser una operación conmutativa mientras que la multiplicación de matrices no lo es. En este capítulo analizaremos los aspectos más relevantes en el comportamiento de las llamadas "operaciones binarias".

Cuando un conjunto esta provisto de una o varias operaciones binarias se tiene un "sistema algebraico". Dicho sistema posee cierta "estructura" que está determinada por las propiedades de las operaciones definidas en el conjunto.

Es posible que dos conjuntos formados por elementos de diferente naturaleza y provistos de operaciones distintas tengan, sin embargo, el mismo "comportamiento algebraico"; es decir, que las opera ciones obedezcan a las mismas leyes. Se dice en tal caso que ambos sistemas poseen la misma "estructura algebraica".

A ciertas estructuras fundamentales se les han asignado nombres específicos como el de "grupo", "anillo" o "campo".

No se pretende realizar aquí un estudio exhaustivo de las diversas estructuras algebraicas existentes, sino más bien presentar los conceptos básicos que nos permitan identificar y comprender la estructura algebraica de los sistemas más comunes en matemáticas.

VIII.1 OPERACIONES BINARIAS Y SUS PROPIEDADES

El concepto de operación binaria es fundamental para el estudio de las estructuras algebraicas; en consecuencia, necesitamos empezar por definir formalmente lo que entenderemos por una operación binaria.

No se trata ya de una operación en particular, como la adición de números complejos o la multiplicación de matrices, sino del concepto mismo de operación binaria; es decir, de aquello que es común a todas las operaciones de este tipo que conocemos.

¿Qué tienen en común operaciones como la adición de números racionales, la sustracción de polinomios y la multiplicación de matrices, por ejemplo?

Fundamentalmente lo siguiente:

- Se aplican a dos elementos de la misma especie (de ahí el término "binaria").
- Asignan a dichos elementos un único "resultado", que es otro elemento de la misma especie, por medio de un criterio determinado. En general, el resultado asignado depende no sólo de quiénes sean los elementos sino también del orden en el que éstos sean considerados.

Podemos decir entonces que una operación binaria es una regla que asigna a cada par ordenado de elementos de un conjunto un único elemento de dicho conjunto. Con ayuda del concepto de función, la definición de operación binaria puede enunciarse de la siguiente manera

VIII.1.1 DEFINICION

Una operación binaria $*$ definida en un conjunto S es una función de $S\times S$ en S. La imagen del par ordenado (a,b) bajo la operación $*$ se representa con $a*b$.

La adición de números racionales, por ejemplo, es una función de $Q\times Q$ en Q, denotada por el símbolo $+$, que asigna a cada par ordenado de números racionales $(\frac{a}{b}, \frac{c}{d})$ un único número racional representado por $\frac{a}{b} + \frac{c}{d}$, al que se conoce como "la suma" de $\frac{a}{b}$ y $\frac{c}{d}$.

La expresión

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

INGENIERIA ANEXO BIBLIOTECA

de la definición I.4.4 especifica como obtener el número $\frac{a}{b} + \frac{c}{d}$ a partir de los números $\frac{a}{b}$ y $\frac{c}{d}$; es decir, especifica la regla o criterio de asignación.

Al aplicar dicha regla al par ordenado $(\frac{1}{2}, \frac{4}{3})$, por ejemplo, se obtiene el número racional $\frac{11}{6}$; es decir

$$\frac{1}{2} + \frac{4}{3} = \frac{11}{6}$$

Entre otros ejemplos de operaciones binarias conocidas tenemos los siguientes:

La adición y la multiplicación en el conjunto de los números naturales

La sustracción en el conjunto de los números enteros

La división en el conjunto de los números complejos diferentes de cero

La adición y la sustracción de polinomios

La adición y la multiplicación en el conjunto de matrices cuadradas de orden n

La unión y la intersección de conjuntos, etc.

Es claro que el concepto de operación binaria establecido por VIII.1.1 no se restringe a las operaciones usuales, como las que acabamos de mencionar, sino que admite la existencia de "nuevas" operaciones binarias.

Para definir una operación binaria en un conjunto S bastará con especificar una regla que asigne a cada par ordenado de elementos de S un único elemento de S. Por ejemplo, podemos enunciar la siguiente regla para un par ordenado cualquiera de números naturales:

"Al primer elemento agregarle el doble del segundo"

con lo que hemos definido una operación binaria en el conjunto N.

Para representar dicha operación podemos utilizar cualquier símbolo, aunque es recomendable emplear símbolos distintos a los de las operaciones usuales para evitar confusiones.

Por ejemplo, si elegimos el símbolo Δ para representar la operación anterior, ésta quedará definida por la expresión

$$m \,\Delta\, n = m + 2n; \quad \forall\, m, n \in N \qquad (1)$$

Si aplicamos dicha operación al par ordenado (1,3), por ejemplo, se obtendrá como resultado el número 7; lo que se expresa de la siguiente manera

$$1 \,\Delta\, 3 = 7$$

si la aplicamos ahora al par ordenado (3,1) se obtendrá como resultado 5; esto es

$$3 \,\Delta\, 1 = 5$$

Aunque la manera más usual de definir una operación binaria es mediante una expresión, digamos "matemática", como la expresión (1), en ciertos casos suele hacerse también mediante una "tabla".

Dichas tablas son particularmente útiles cuando el conjunto sobre el que se define la operación es finito y tiene pocos elementos. Por ejemplo, en el conjunto

$$G = \{\alpha, \beta, \gamma\}$$

podemos definir una operación $\square$ mediante la siguiente tabla

$\square$	α	β	γ
α	α	β	γ
β	β	β	α
γ	α	β	γ

(2)

Para buscar en la tabla el resultado asignado a un par ordenado en particular, se busca al primer elemento en la columna de la izquierda y al segundo en el renglón superior; el resultado se encuentra en la intersección del renglón y la columna correspondientes.

Por ejemplo, el resultado de aplicar la operación $\square$ al par ordenado (β,γ) es, de acuerdo con la tabla (2):

$$\beta \square \gamma = \alpha$$

Para (γ,β) se tendrá en cambio

$$\gamma \square \beta = \beta$$

Antes de pasar a ocuparnos de las propiedades cabe resaltar que, de acuerdo con la definición VIII.1.1, una operación binaria * definida en un conjunto S asigna siempre como resultado un elemento de S; es decir que

$$\forall\ a, b \in S:\ a*b \in S$$

Algunos autores se refieren a esta situación diciendo que el conjunto S es "cerrado" respecto a la operación *.

De acuerdo con esto, la sustracción no es una operación binaria en el conjunto de los números naturales ya que, para los números naturales 1 y 3, por ejemplo, la diferencia 1-3 no pertenece al

conjunto N.

- Cerradura

VIII.1.2 DEFINICION

Sea * una operación binaria definida en un conjunto S, y sea T un subconjunto de S. Se dice que T es cerrado respecto a la operación * si

$$\forall\ a, b \in T:\ a*b \in T$$

Es decir que el subconjunto T es cerrado respecto a la operación * si al aplicar dicha operación a dos elementos cualesquiera de T se obtiene como resultado otro elemento de T.

Así, por ejemplo, para la multiplicación de números racionales los siguientes subconjuntos de Q son cerrados

$\{x \mid x \in Q,\ x > 0\}$

$\{0\}$

$\{0,\ 1,\ -1\}$

así como también lo es el propio Q; mientras que los siguientes subconjuntos no son cerrados respecto a dicha operación

$\{x \mid x \in Q,\ x < 0\}$

$\{0,\ 1,\ 2\}$

Respecto a la operación Δ definida por la expresión (1), el conjunto N no tiene subconjuntos cerrados salvo el propio N.

Para la operación $\square$ definida por la tabla (2), el conjunto

$\{\beta, \gamma\}$ es el único subconjunto de G que no es cerrado respecto a dicha operación.

- Elementos idénticos

VIII.1.3 DEFINICION

Sea $*$ una operación binaria definida en un conjunto S:

i) Un elemento $e \in S$ es un idéntico izquierdo para $*$ si

$$e*a = a, \quad \forall\, a \in S$$

ii) Un elemento $e \in S$ es un idéntico derecho para $*$ si

$$a*e = a, \quad \forall\, a \in S$$

iii) Un elemento $e \in S$ es un idéntico para $*$ si es idéntico izquierdo e idéntico derecho.

Así, por ejemplo, para la multiplicación definida en el conjunto M de todas las matrices de m×n con elementos en C, la matriz I_m es un idéntico izquierdo ya que

$$I_m\, A = A, \quad \forall\, A \in M$$

y la matriz I_n es un idéntico derecho puesto que

$$A\, I_n = A, \ \forall\, A \in M$$

El número cero es un elemento idéntico para la adición definida en Q, ya que

$$0 + x = x$$

y

$$x + 0 = x, \quad \forall\, x \in Q$$

y el número uno lo es para la multiplicación definida en dicho

conjunto, puesto que

$$1 \cdot x = x$$

y

$$x \cdot 1 = x, \ \forall\, x \in Q$$

El conjunto N no tiene idéntico izquierdo para la operación Δ definida por la expresión (1), ya que la condición

$$e \,\Delta\, n = n$$

es equivalente, por definición de la operación Δ, a la expresión

$$e + 2n = n$$

o sea

$$e = -n$$

Entonces, si $n \in N$ se tendrá que $-n \notin N$, por lo que Δ no tiene idéntico izquierdo en N.

Dicha operación tampoco tiene idéntico derecho en N, ya que

$$n \,\Delta\, e = n, \ \forall\, n$$

es equivalente a

$$n + 2e = n$$
$$2e = 0$$
$$e = 0$$

y el cero no es un elemento de N.

Como otro ejemplo consideremos la misma regla de correspondencia para la operación Δ, pero definida ahora en el conjunto de los números enteros; esto es

$$m \,\Delta\, n = m + 2n; \ \forall\, m, n \in Z$$

Al buscar un idéntico izquierdo para Δ en el conjunto Z partimos de la condición

$$e \Delta n = n$$

y llegamos nuevamente a la expresión

$$e = -n$$

donde ahora, si $n \varepsilon Z$ se tiene que $-n \varepsilon Z$.

Sin embargo, de acuerdo con dicha expresión cada elemento tendría su "propio" idéntico izquierdo, y el "idéntico izquierdo" de un elemento no lo sería para otro diferente.

Por ejemplo, un "idéntico izquierdo" para 1 sería -1 ya que

$$-1 \Delta 1 = 1$$

pero para 2 se tendría que

$$-1 \Delta 2 = 3$$

En consecuencia, la operación Δ no tiene idéntico izquierdo en Z.

Por otra parte, el número cero es un idéntico derecho para la operación Δ definida en Z ya que

$$n \Delta 0 = n, \forall n \varepsilon Z$$

y además $0 \varepsilon Z$.

Finalmente, para la operación $\square$ definida por la tabla (2), el conjunto G tiene dos idénticos izquierdos (que son α y γ) y no tiene idéntico derecho.

- Elementos inversos

VIII.1.4 DEFINICION

Sea $*$ una operación binaria definida en un conjunto S, y:

i) Sea e un idéntico izquierdo para $*$. Un elemento $\hat{a} \varepsilon S$ es un inverso izquierdo del elemento $a \varepsilon S$ si

$$\hat{a} * a = e$$

ii) Sea e un idéntico derecho para $*$. Un elemento $\hat{a} \varepsilon S$ es un inverso derecho del elemento $a \varepsilon S$ si

$$a * \hat{a} = e$$

iii) Sea e un idéntico para $*$. Un elemento $\hat{a} \varepsilon S$ es un inverso del elemento $a \varepsilon S$ si

$$\hat{a} * a = e \quad y \quad a * \hat{a} = e$$

Como se sigue de la definición anterior, para poder hablar de elementos inversos se requiere que existan elementos idénticos.

Así pues, como la operación Δ definida en N por la expresión (1) no tiene idéntico izquierdo, no puede haber inversos izquierdos para dicha operación; y como tampoco existe idéntico derecho, ningún elemento de N podrá tener inverso derecho para la operación Δ.

Para la operación $\square$ definida en G por la tabla (2) se tiene lo siguiente:

Como α y γ son idénticos izquierdos, entonces α tiene dos inversos izquierdos, que son α y γ, ya que

$\alpha \;\square\; \alpha = \alpha$

$\gamma \;\square\; \alpha = \alpha$

β no tiene inverso izquierdo

γ tiene dos inversos izquierdos, que son α y γ, ya que

$\alpha \;\square\; \gamma = \gamma$

$\gamma \;\square\; \gamma = \gamma$

Como no hay idéntico derecho ningún elemento de G tiene inverso derecho.

Para la adición en Q el cero es un elemento idéntico por lo que, para dicha operación, un inverso del número $x \in Q$ es el número $-x \in Q$ (llamado su simétrico), ya que

$$-x + x = 0 \quad y \quad x + (-x) = 0$$

Para la multiplicación en Q, como el número uno es un elemento idéntico, un inverso del número racional $x \neq 0$ es el número $\frac{1}{x} \in Q$ (llamado su recíproco), puesto que

$$\frac{1}{x} \cdot x = 1 \qquad y \qquad x \cdot \frac{1}{x} = 1$$

- Asociatividad

VIII.1.5 DEFINICION

Sea $*$ una operación binaria definida en un conjunto S

Se dice que $*$ es asociativa si

$$\forall\, a, b, c \in S: \quad a*(b*c) = (a*b)*c$$

Consideremos, por ejemplo, la operación $\square$ definida en G por la tabla (2). Esta operación no es asociativa ya que, para los elementos β, β y γ se tiene que

$$\beta \;\square\; (\beta \;\square\; \gamma) = \beta \;\square\; \alpha = \beta$$

y

$$(\beta \;\square\; \beta) \;\square\; \gamma = \beta \;\square\; \gamma = \alpha$$

por lo que

$$\beta \;\square\; (\beta \;\square\; \gamma) \neq (\beta \;\square\; \beta) \;\square\; \gamma$$

y, en consecuencia, existe al menos un caso para el cual no se cumple la igualdad de la definición VIII.1.5, igualdad que debe cumplirse en todos los casos para que la operación sea asociativa.

La adición y la multiplicación en Q son operaciones asociativas, ya que

$$x + (y + z) = (x + y) + z$$

y

$$x\,(yz) = (xy)\,z$$

para todos los valores de $x, y, z \in Q$, como lo establecen los teoremas I.4.5 y I.4.8; mientras que la sustracción en Q no lo es, ya que la igualdad

$$x - (y - z) = (x - y) - z$$

no se cumple en todos los casos.

Para determinar si la operación Δ definida por la expresión (1) es asociativa o no lo es, debemos analizar para que valores de m, n y p se satisface la igualdad

$$m \Delta (n \Delta p) = (m \Delta n) \Delta p$$

Por una parte tenemos que

$$m \Delta (n \Delta p) = m \Delta (n + 2p) = m + 2(n + 2p) = m + 2n + 4p$$

y por otra

$$(m \Delta n) \Delta p = (m + 2n) \Delta p = m + 2n + 2p$$

resultados que nunca serán iguales puesto que p no puede valer cero.

Cabe hacer notar que cuando una operación * es asociativa pode - mos escribir

$$a*b*c$$

sin que exista ambigüedad respecto al significado de la expresión, puesto que en cualquier orden que coloquemos los paréntesis se obtendrá el mismo resultado.

- Conmutatividad

VIII.1.6 DEFINICION

Sea $\star$ una operación binaria definida en un conjunto S.

Se dice que * es conmutativa si

$$\forall\ a, b \in S: \quad a*b = b*a$$

La operación $\square$ definida por la tabla (2) tampoco es conmutativa ya que, como vimos

$$\beta \square \gamma = \alpha$$

y

$$\gamma \square \beta = \beta$$

por lo que

$$\beta \square \gamma \neq \gamma \square \beta$$

y existe al menos un caso para el cual no se cumple la igualdad de la definición VIII.1.6.

La adición y la multiplicación en Q son operaciones conmutativas ya que

$$x + y = y + x$$

y

$$xy = yx$$

para todos los valores de $x, y \in Q$, como lo establecen los teore mas I.4.5 y I.4.8, mientras que la división en Q no lo es, ya que la igualdad

$$x \div y = y \div x$$

no se cumple en todos los casos.

Para determinar si la operación Δ definida por la expresión (1). es conmutativa o no lo es, debemos analizar para qué valores de m y n se satisface la igualdad

$$m \Delta n = n \Delta m$$

Por una parte tenemos que

$$m \Delta n = m + 2n$$

y por otra

$$n \, \Delta \, m = n + 2m$$

resultados que serán iguales sólo cuando m y n sean iguales y no para todos los valores de m y n, por lo que la operación Δ no es conmutativa.

VIII.1.7 EJERCICIOS

1.- Para cada uno de los siguientes conjuntos determinar si $*$ es una operación binaria definida en S:

a) $S = \{x \mid x \in Z,\ x < 0\}$, y $*$ es la multiplicación en Z.

b) $S = \{-1, 0, 1\}$, y

$*$	-1	0	1
-1	1	0	-1
0	0	0	0
1	-1	0	1

c) S es el conjunto de matrices de 3×2 con elementos en R, y $*$ es la multiplicación de matrices.

2.- Para el conjunto $L = \{a, b, c\}$ y la operación Δ definida por la tabla, obtener todos los subconjuntos de L que son cerrados para Δ.

Δ	a	b	c
a	a	c	a
b	a	b	b
c	c	b	c

3.- Sea □ la operación definida en Z como

$$a \,\square\, b = a + 3b + 1; \quad \forall\, a, b \in Z.$$

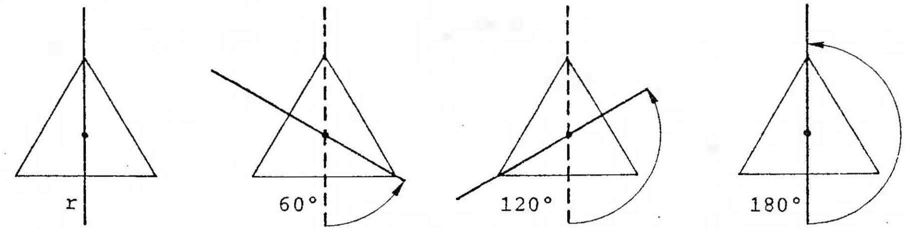
Determinar si dicha operación:

a) Es asociativa

b) Tiene idéntico derecho

c) Tiene inverso derecho para todo elemento de Z.

4.- Sea S el conjunto de los ángulos que puede girar la recta r de la primera figura conservando la simetría del triángulo equilátero



Considérese la suma de dos de tales ángulos como el giro final que experimenta la recta a partir de su posición original. Construir una tabla que defina a dicha operación, y determinar si ésta:

a) Es conmutativa

b) Es asociativa (Se sugiere verificar únicamente tres o cuatro de los veintisiete casos posibles que deberían verificarse en una prueba formal)

c) Tiene idéntico

d) Tiene inverso para todo elemento de S.

VIII.2 ESTRUCTURA DE GRUPO

La estructura algebraica más simple que consideraremos en este capítulo es la de grupo.

Se emplea el nombre de grupo para designar la estructura que poseen los sistemas formados por un conjunto y una operación binaria cuando dicha operación es asociativa, está dotada de elemento idéntico y todo elemento del conjunto tiene inverso para la operación.

- Definición de grupo

La definición que se presenta a continuación está constituida por tres postulados independientes que, en conjunto, son suficientes para deducir todas las propiedades características de la estructura de grupo.

VIII.2.1 DEFINICION
Sea G un conjunto no vacío y sea $*$ una operación binaria definida en G. El sistema $(G,*)$ tiene estructura de grupo si:

i) $\forall\ a, b, c \in G \quad a*(b*c) = (a*b)*c$

ii) $\exists\ e \in G$ tal que $e*a = a,\ \forall\ a \in G$

iii) $\forall\ a \in G, \exists\ \hat{a} \in G$ tal que $\hat{a}*a = e$

El postulado i) de la definición nos dice que la operación $*$ debe ser asociativa.

El postulado ii) nos indica que debe existir un elemento de G que sea idéntico izquierdo para la operación $*$. Conviene, al respecto, hacer un par de observaciones:

1) Aunque el símbolo $\exists\ e$ se lee "existe un e",† debe interpretarse como "existe al menos un e". Para indicar que existe exactamente un elemento e se dice "existe uno y sólo un e", para lo cual no se tiene un símbolo adoptado universalmente.

2) Aunque el postulado ii) sólo exige la existencia de (al menos) un idéntico izquierdo, dicho elemento es también un idéntico derecho y, además, es único. La demostración se verá más adelante en las "propiedades elementales de los grupos".

El postulado iii) de la definición VIII.2.1 nos indica que todo elemento de G debe tener (al menos) un inverso izquierdo. Nuevamente, como se demostrará más adelante, dicho inverso izquierdo es también inverso derecho y, además, es único.

- Ejemplos de grupos

Los ejemplos más conocidos de grupos los encontramos entre los diversos sistemas numéricos que ya hemos manejado.

El conjunto de los números enteros y la operación de adición constituyen un sistema con estructura de grupo ya que, como sabemos

i) $\forall\ a, b, c \in Z: \quad a + (b + c) = (a + b) + c$

ii) $0 \in Z$ y es tal que $0 + a = a,\ \forall\ a \in Z$

iii) $\forall\ a \in Z, \exists -a \in Z$ tal que $-a + a = 0$

por lo que $(Z, +)$ es un grupo.

Otros sistemas conocidos que también tienen estructura de grupo

† *a dicho símbolo se le conoce como cuantificador existencial*

son los siguientes:

- Los números racionales con la adición
- Los números complejos con la adición
- Los números complejos diferentes de cero con la multiplicación
- Los polinomios con la adición
- Las matrices de m×n con la adición
- Las matrices no singulares de orden n con la multiplicación

etc.

Otro ejemplo de grupo, que a diferencia de los anteriores consta de un conjunto finito, lo constituye el sistema $(S,\cdot)$ donde $S = \{1, -1, i, -i\}$ y $\cdot$ es la multiplicación usual de números complejos.

Para demostrar que dicho sistema tiene estructura de grupo debemos comprobar, primero, que la multiplicación de números complejos es una operación definida en S; para lo cual bastará con verificar que S es cerrado respecto a dicha operación. Podemos lograr esto construyendo la siguiente tabla

$\cdot$	1	-1	i	-i
1	1	-1	i	-i
-1	-1	1	-i	i
i	i	-i	-1	1
-i	-i	i	1	-1

donde se aprecia claramente que todos los resultados posibles caen dentro del conjunto S.

Una vez comprobado que la multiplicación es una operación definida en S, debemos verificar que ésta satisface las propiedades que

exigen los postulados i), ii) y iii) de la definición de grupo:

El postulado i) se cumple como consecuencia de la propiedad asociativa de la multiplicación en C (Teorema II.1.4), puesto que $S \subset C$.

El postulado ii) también se cumple ya que $1 \in S$ y es tal que $1 \cdot a = a, \forall\, a \in S$, como puede observarse en la misma tabla.

Respecto al postulado iii), todo elemento de S tiene inverso izquierdo en S para la multiplicación, ya que

$$1 \cdot 1 = 1 \quad \text{por lo que} \quad \hat{1} = 1$$
$$-1 \cdot (-1) = 1 \quad \text{por lo que} \quad -\hat{1} = -1$$
$$-i \cdot i = 1 \quad \text{por lo que} \quad \hat{i} = -i$$
$$i \cdot (-i) = 1 \quad \text{por lo que} \quad -\hat{i} = i$$

En consecuencia, hemos demostrado que el sistema $(S,\cdot)$ tiene estructura de grupo.

Como un último ejemplo relativo a la definición de grupo, investiguemos si el sistema (R, §) tiene tal estructura; donde R es el conjunto de los números reales y § la operación definida por

$$x \; § \; y = x + y - 2xy$$

Primero determinamos si § es una operación definida en R:

Sean $x, y \in R$; entonces, por i) de I.5.3 se tiene que $x + y \in R$ y además que $2xy \in R$; por lo que, de I.5.5 y I.5.3, se sigue que

$$x + y - 2xy \in R$$

En consecuencia, $\forall\, x, y \in R$ se tiene que

$x \S y \in R$

y la operación § está definida en R.

i) Para determinar si § es asociativa, por una parte calculamos

$$x \S (y \S z) = x \S (y + z - 2yz)$$
$$= x + (y + z - 2yz) - 2x(y + z - 2yz)$$
$$x \S (y \S z) = x + y + z - 2yz - 2xy - 2xz + 4xyz.$$

y por la otra

$$(x \S y) \S z = (x + y - 2xy) \S z$$
$$= (x + y - 2xy) + z - 2(x + y - 2xy)z$$
$$(x \S y) \S z = x + y + z - 2yz - 2xy - 2xz + 4xyz.$$

En consecuencia, podemos afirmar que

$$x \S (y \S z) = (x \S y) \S z; \ \forall\, x, y, z \in R$$

ii) Sea e un idéntico izquierdo para §, entonces

$$e \S x = x$$

es decir

$$e + x - 2ex = x$$
$$e - 2ex = 0$$
$$e(1 - 2x) = 0$$

donde se observa que $e = 0$ es un idéntico izquierdo para § en R.

En efecto, vemos que $0 \in R$ y además

$$0 \S x = 0 + x - 2 \cdot 0 \cdot x = 0 + x - 0 = x, \ \forall\, x \in R.$$

iii) Sea $\hat{x}$ un inverso izquierdo de $x \in R$, entonces

$$\hat{x} \S x = 0$$

esto es

$$\hat{x} + x - 2\hat{x}x = 0$$
$$\hat{x} - 2\hat{x}x = -x$$
$$\hat{x}(1 - 2x) = -x$$

por lo que, si $x \neq \frac{1}{2}$ se tendrá que

$$\hat{x} = \frac{-x}{1 - 2x} \in R$$

Sin embargo, si $x = \frac{1}{2} \not\exists\, \hat{x}$ ya que la expresión

$$\hat{x} \S \frac{1}{2} = 0$$

es equivalente a

$$\hat{x} + \frac{1}{2} - 2\hat{x}\,\frac{1}{2} = 0$$
$$\hat{x} + \frac{1}{2} - \hat{x} = 0$$

y no se cumple para valor alguno de $\hat{x} \in R$.

En consecuencia, como el número real $\frac{1}{2}$ no tiene inverso para la operación § , no se satisface el postulado iii) de la definición VIII.2.1 y el sistema (R, §) no tiene estructura de grupo.

- Propiedades elementales de los grupos

Como consecuencia de los postulados que establece la definición de grupo se deducen una serie de propiedades, las cuales son comunes a todos los sistemas que tienen dicha estructura. En este apartado presentamos, a manera de teoremas, algunas de estas propiedades.

Como tales propiedades son comunes a todos los grupos, éstas deben ser demostradas directamente a partir de los postulados que integran la definición VIII.2.1; sin embargo, es posible organizar el trabajo mediante una secuencia de resultados que permita aprovechar propiedades ya demostradas para probar otras subsecuentes, como lo haremos a continuación.

VIII.2.2 TEOREMA (Ley de cancelación izquierda)

Si $(G,*)$ es un grupo, entonces $\forall$ a, b, c ε G:

$$a*b = a*c \Rightarrow b = c$$

DEMOSTRACION

Sea

$$a*b = a*c$$

Por iii) de VIII.2.1 $\exists$ $\hat{a}$ ε G y entonces

$$\hat{a}*(a*b) = \hat{a}*(a*c)$$

Ahora, por i) de VIII.2.1 podemos escribir

$$(\hat{a}*a)*b = (\hat{a}*a)*c$$

En consecuencia, por iii) de VIII.2.1

$$e*b = e*c$$

por lo que, de ii) de VIII.2.1 se sigue que

$$b = c$$

y la prueba termina. □

VIII.2.3 TEOREMA

Si $(G,*)$ es un grupo y e es un idéntico izquierdo para $*$, entonces e es un idéntico para dicha operación.

DEMOSTRACION

Sean e un idéntico izquierdo para $*$ y a un elemento cualquiera de G:

Por iii) de VIII.2.1 $\exists$ $\hat{a}$ ε G y por i) de VIII.2.1 podemos escribir

$$\hat{a}*(a*e) = (\hat{a}*a)*e$$

En consecuencia, por iii) de VIII.2.1 se tendrá que

$$\hat{a}*(a*e) = e*e$$

pero como e es un idéntico izquierdo, por ii) de VIII.2.1 podemos escribir

$$\hat{a}*(a*e) = e$$

Nuevamente, por iii) de VIII.2.1 se tendrá

$$\hat{a}*(a*e) = \hat{a}*a$$

y de VIII.2.2 se sigue que

$$a*e = a$$

por lo que e es también un idéntico derecho para $*$.

En consecuencia, de la definición VIII.1.3, e es un idéntico para $*$. □

VIII.2.4 TEOREMA

Si $(G,*)$ es un grupo entonces el idéntico para $*$ es único

DEMOSTRACION

Sea e un idéntico para $*$ y sea z otro idéntico para dicha operación; esto es

$$z*a = a \quad y \quad a*z = a, \; \forall\, a \in G$$

Entonces, haciendo $a = e$ en la segunda expresión se tendrá que

$$e*z = e$$

pero como e es un idéntico para $*$ podemos escribir

$$e*z = e*e$$

En consecuencia, de VIII.2.2 se sigue que

$$z = e$$

por lo que no puede haber dos idénticos diferentes y la prueba termina. □

Como consecuencia de los dos últimos teoremas, y del postulado ii) de la definición VIII.2.1, podemos concluir que en un grupo existe uno y sólo un elemento idéntico para la operación.

VIII.2.5 TEOREMA

Si $(G,*)$ es un grupo y $\hat{a}$ es un inverso izquierdo del elemento $a \in G$, entonces $\hat{a}$ es un inverso de a.

DEMOSTRACION

Sean a un elemento cualquiera de G, $\hat{a}$ un inverso izquierdo de a y e el idéntico para $*$:

Por i) de VIII.2.1 podemos escribir

$$\hat{a}*(a*\hat{a}) = (\hat{a}*a)*\hat{a}$$

como $\hat{a}$ es inverso izquierdo de a se tiene que

$$\hat{a}*(a*\hat{a}) = e*\hat{a}$$

y como e es idéntico para $*$ se sigue que

$$\hat{a}*(a*\hat{a}) = \hat{a}$$

$$\hat{a}*(a*\hat{a}) = \hat{a}*e$$

por lo que, de VIII.2.2

$$a*\hat{a} = e$$

y $\hat{a}$ es un inverso derecho de a.

En consecuencia, de VIII.2.4 se sigue que $\hat{a}$ es un inverso de a. □

VIII.2.6 TEOREMA

Si $(G,*)$ es un grupo entonces el inverso de $a \in G$ para la operación $*$ es único.

VIII.2.7 TEOREMA (Ley de cancelación derecha)

Si $(G,*)$ es un grupo, entonces $\forall$ a, b, c $\in$ G:

$$b*a = c*a \implies b = c$$

Se deja al lector como ejercicio la demostración de los dos teoremas anteriores.

VIII.2.8 TEOREMA

Si $(G,*)$ es un grupo y $a, b \in G$ entonces las ecuaciones

$$a*x = b$$
$$y*a = b$$

tienen, respectivamente, soluciones únicas $x, y \in G$.

DEMOSTRACION

Para encontrar una solución de la ecuación

$$a*x = b \qquad (1)$$

podemos utilizar las propiedades que establece la definición de grupo en la siguiente forma:

Como $a \in G$, por iii) de VIII.2.1, $\exists\ \hat{a} \in G$ y de la expresión (1) se sigue que

$$\hat{a}*(a*x) = \hat{a}*b$$

y en consecuencia

$$(\hat{a}*a)*x = \hat{a}*b \quad \text{por i) de VIII.2.1}$$
$$e*x = \hat{a}*b \quad \text{por iii) de VIII.2.1}$$
$$x = \hat{a}*b \quad \text{por ii) de VIII.2.1}$$

Así, el elemento $\hat{a}*b \in G$ es una solución de la ecuación (1) ya que

$$a*(\hat{a}*b) = (a*\hat{a})*b = e*b = b$$

Para demostrar que dicha solución es única, supongamos dos soluciones x y x' de la ecuación (1); entonces

$$a*x = b \quad \text{y} \quad a*x' = b$$

por lo que

$$a*x = a*x'$$

y de VIII.2.2 se sigue que

$$x = x'$$

por lo que ambas soluciones son iguales.

Esto completa la demostración del teorema para el caso correspondiente a la primera ecuación. La prueba del segundo caso es similar y se sugiere al lector hacerla como ejercicio. □

A continuación se enuncian otras tres propiedades de los grupos, cuya demostración también se deja al lector, con las cuales concluye la lista de propiedades que aquí presentamos. Dicha lista no es exhaustiva; sin embargo, contiene las propiedades elementales de uso más frecuente.

VIII.2.9 TEOREMA

Si $(G,*)$ es un grupo y $\hat{a}$ representa el inverso de $a \in G$ para $*$, entonces:

i) $\hat{(\hat{a})} = a; \quad \forall\ a \in G$

ii) $\widehat{(a*b)} = \hat{b}*\hat{a}; \quad \forall\ a, b \in G$

iii) $\widehat{(a_1 * a_2 * \dots * a_n)} = \hat{a}_n * \dots * \hat{a}_2 * \hat{a}_1; \ \forall\ a_1, a_2, \dots, a_n \in G$

Antes de concluir este apartado conviene hacer notar que los tres postulados que integran la definición VIII.2.1 no son el único conjunto de postulados que puede emplearse para definir la estructura de grupo.

Los postulados ii) y iii), por ejemplo, pueden ser reemplazados por los dos siguientes:

II) $\exists\ e \in G$ tal que $a*e = a,\ \forall\ a \in G$

III) $\forall\ a \in G, \exists\ \hat{a} \in G$ tal que $a*\hat{a} = e$

los cuales pueden considerarse como las "versiones derechas" de los postulados ii) y iii).

Otra definición de grupo podría quedar integrada por el postulado i) de VIII.2.1 y por la siguiente condición:

$\forall\ a, b \in G$ las ecuaciones

$$a*x = b$$

$$y*a = b$$

tienen soluciones $x, y \in G$.

- Subgrupos

Cuando un sistema $(G,*)$ tiene estructura de grupo es posible que algunos subconjuntos de G con la operación $*$ tengan, por sí mismos, estructura de grupo. En tal caso se dice que éstos son subgrupos de G, como lo establece la siguiente definición.

VIII.2.10 DEFINICION

Sea $(G,*)$ un grupo y sea $S \subset G$, se dice que S es un subgrupo de G para la operación $*$ si $(S,*)$ es un grupo.

Por ejemplo, vimos anteriormente que el sistema $(S,\cdot)$, donde $S = \{1, -1, i, -i\}$ y $\cdot$ es la multiplicación, tiene estructura de grupo; como S es un subconjunto del conjunto de números complejos diferentes de cero y éste forma un grupo con la multiplicación, entonces $(S,\cdot)$ es un subgrupo de $(C-\{0\},\cdot)$.

El sistema $(S,\cdot)$, a su vez, tiene también subgrupos, como el formado por el conjunto $\{1, -1\}$ y la multiplicación o el formado por el conjunto $\{1\}$ con dicha operación.

En general, si $(G,*)$ es un grupo y $e \in G$ es el idéntico para $*$, entonces el sistema $(\{e\}, *)$ es un subgrupo de G. Dicho subgrupo, así como el mismo $(G,*)$, reciben el nombre de subgrupos impropios; cualquier otro subgrupo de G, si lo hay, se considera propio.

El siguiente teorema nos permite determinar cuándo un subconjunto es un subgrupo, sin tener que verificar todas las condiciones de la definición VIII.2.1.

VIII.2.11 TEOREMA

Sea $(G,*)$ un grupo y sea $S \subset G$, S es un subgrupo de G para la operación $*$ si y sólo si:

i) $\forall\ a, b \in S$: $a*b \in S$

ii) $\forall\ a \in S$: $\hat{a} \in S$

INGENIERIA
A N E X O
BIBLIOTECA

DEMOSTRACION

Probaremos primero que un conjunto $S \subset G$ que satisface las condiciones i) y ii) del teorema cumple con los postulados establecidos por la definición VIII.2.1:

1) Como $S \subset G$ y $*$ es una operación definida en G, asocia a cada par ordenado de elementos de S uno y sólo un elemento de G; además, por i) de VIII.2.11 el conjunto S es cerrado para $*$, por lo que ésta es una operación definida en S.

2) Como $(G,*)$ es un grupo, $\forall$ a, b, c ε G se cumple que

$$a*(b*c) = (a*b)*c$$

por lo que, en particular, dicha expresión se cumplirá $\forall$ a, b, c ε S ya que $S \subset G$.

3) Por ii) de VIII.2.11 todo elemento de S tiene inverso para $*$ en el conjunto S.

4) Para mostrar que el idéntico también pertenece al conjunto S, tomemos un elemento a ε S; por ii) se tendrá que $\hat{a}$ ε S y por i) $a*\hat{a}$ ε S, y por lo tanto e ε S.

En consecuencia, de 1) a 4) se sigue que $(S,*)$ es un grupo, y por VIII.2.10 es un subgrupo de G.

Por otra parte, si alguna de las dos condiciones del teorema VIII.2.11 no se cumple, de VIII.2.1 se sigue que el sistema $(S,*)$ no es un grupo y, por tanto, tampoco un subgrupo de G. Esto completa la demostración

$\square$

Para ilustrar la aplicación del teorema anterior a un caso particular, consideremos nuevamente el sistema formado por el conjunto de los números enteros y la operación de adición, el cual tiene estructura de grupo, y busquemos determinar si los siguientes subconjuntos de Z son subgrupos para la adición

$$S = \{\ldots -6, -4, -2, 0, 2, 4, 6, \ldots\} = \{2m \mid m \in Z\}$$

$$T = \{1, 2, 3, 4, 5, \ldots\} = \{m \mid m \in Z, m > 0\}$$

Para el primer caso:

i) Sean $a = 2m$ y $b = 2n$ dos elementos cualesquiera de S, con m, n ε Z; entonces

$$a + b = 2m + 2n = 2(m + n) \in S,$$

ya que $m + n \in Z$.

ii) $\hat{a} = -2m = 2(-m) \in S$, ya que $-m \in Z$.

En consecuencia, $(S,+)$ es un subgrupo de $(Z,+)$.

Para el segundo caso:

i) Sean $a = m$ y $b = n$ dos elementos cualesquiera de T, con m, n ε Z y m, n > 0; entonces

$$a + b = m + n \in T,$$

ya que $m + n \in Z$ y $m + n > 0$.

ii) $\hat{a} = -m \notin T$, ya que $-m < 0$.

En consecuencia, $(T,+)$ no es un subgrupo de $(Z,+)$.

- Grupos abelianos

Con respecto a la definición VIII.2.1 podemos preguntarnos que pasaría al agregarle un postulado más, independiente de los anteriores.

La respuesta es que se obtendría una estructura más completa; más "rica" en el sentido que podríamos efectuar en ella ciertos proce sos algebraicos que no serían válidos en estructuras más simples.

Si el postulado que se agrega a la definición VIII.2.1 es la propiedad conmutativa de la operación, la estructura obtenida se conoce como "grupo conmutativo" o "grupo abeliano". El segundo nombre, empleado en honor del matemático noruego Niels Henrik Abel, es el que adoptamos en la siguiente definición.

> VIII.2.12 DEFINICION
>
> Un grupo $(G,*)$ se dice que es abeliano si:
>
> $$\forall\ a,\ b \in G \quad a*b = b*a$$

Los ejemplos de grupos presentados para ilustrar la definición VIII.2.1 constituyen también ejemplos de grupos abelianos, salvo el último de ellos: el formado por las matrices no singulares de orden n con la multiplicación.

Entre las propiedades adicionales que poseen los grupos abelianos, como consecuencia de la conmutatividad, se encuentran las siguientes.

Si $(G,*)$ es un grupo abeliano, entonces:

i) $\forall\ a,\ b,\ c \in G;\quad a*b = c*a \Rightarrow b = c$

ii) $\forall\ a,\ b,\ c \in G;\quad b*a = a*c \Rightarrow b = c$

iii) $\forall\ a,\ b \in G;\quad \widehat{(a*b)} = \hat{a}*\hat{b}$

iv) $\forall\ a,\ b \in G;$ las ecuaciones

$$a*x = b$$

$$y*a = b$$

tienen la misma solución en G.

VIII.2.13 EJERCICIOS

1.- Entre los sistemas conocidos que se citaron como ejemplos de grupos se encuentran:

a) Los números complejos diferentes de cero con la multiplicación.

b) Las matrices de $m \times n$ con la adición.

Demostrar que dichos sistemas poseen estructura de grupo, indicando los teoremas que establecen cada una de las propiedades que exige la definición VIII.2.1.

2.- Para cada uno de los siguientes conjuntos con la operación $*$ definida, indicar por qué no tienen estructura de grupo

a) $S = \{x \mid x \in Q,\ -1 \leq x \leq 1\} \qquad a*b = a + b$

b) $S = \{1,\ 2,\ 3,\ 4\} \qquad a*b = a$

c) $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle|\ a,b,c,d \in R \right\} \qquad A*B = AB^T$

3.- Sea $C = \{x \mid x \in Q,\ 0 \leq x \leq 10\}$ el conjunto de todas las calificaciones posibles en la asignatura de Algebra Lineal; donde se define la operación promedio como

$$a\ P\ b = \frac{a + b}{2}, \quad \forall\ a,\ b \in C$$

Determinar si el sistema (C,P) tiene estructura de grupo.

4.- Demostrar que si $(G,*)$ es un grupo: entonces, $\forall\ a,\ b,\ c \in G$:

a) $b*a = c*a \Rightarrow b = c$

b) el inverso de a para $*$ es único

c) $\widehat{(a*b)} = \hat{b}*\hat{a}$

5.- Sea A un conjunto no vacío, y sea $*$ una operación definida en A que es asociativa y para la cual las ecuaciones

$$a*x = b$$
$$y*a = b$$

tienen soluciones $x, y \in A$. Demostrar que $(A,*)$ es un grupo.

6.- Si $(G,*)$ es un grupo y $a, b, c \in G$, obtener la solución de cada una de las siguientes ecuaciones:

a) $a*x*b = a$

b) $x*a*x*b = x*c$

7.- Sea M el conjunto de matrices cuadradas de orden dos con elementos en R y sea + la adición usual de matrices. Determinar cuáles de los siguientes conjuntos constituyen subgrupos de $(M,+)$:

$$S = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \middle| \; a, b, c \in R \right\}$$

$$T = \left\{ \begin{bmatrix} a & 1 \\ 1 & c \end{bmatrix} \middle| \; a, c \in R \right\}$$

$$U = \left\{ \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix} \middle| \; a, c \in R^+ \right\}$$

8.- Demostrar que si $(G,*)$ es un grupo abeliano y $a, b, c \in G$, entonces:

a) $a*b = c*a \Rightarrow b = c$

b) Las ecuaciones $a*x = b$ *y* $y*a = b$ tienen la misma solución

c) $\widehat{(a*b)} = \hat{a}*\hat{b}$

VIII.3 ESTRUCTURAS DE ANILLO Y DE CAMPO

En esta sección presentaremos diversas estructuras algebraicas, cada vez más completas, relativas a sistemas formados por un conjunto y dos operaciones binarias.

Es claro que pueden deducirse varias propiedades para cada una de dichas estructuras a partir de sus definiciones correspondientes, como se hizo con la estructura de grupo; sin embargo, no se realizarán aquí tales deducciones y solamente se sugieren algunas como ejercicio.

En las definiciones de estas estructuras se emplearán los símbolos $+$ y $\cdot$ para representar dos operaciones definidas en el conjunto, sin que ésto signifique que se trata de la adición y la multiplicación de números.

El empleo de tales símbolos obedece, ciertamente, a que los ejemplos más conocidos de este tipo de estructuras se encuentran entre los conjuntos de números con las operaciones de adición y multiplicación usuales. De esta manera, el empleo de $+$ y $\cdot$ así como de la simbología correspondiente para idénticos e inversos, puede ser útil para comprender y recordar las propiedades de las diversas estructuras asociándolas con los sistemas numéricos conocidos.

- Anillos

VIII.3.1 DEFINICION

Sea A un conjunto no vacío y sean + y $\cdot$ dos operaciones binarias definidas en A. El sistema (A, +, $\cdot$) tiene estructura de anillo si:

i) $\forall\, a, b, c \in A \quad a + (b + c) = (a + b) + c$

ii) $\forall\, a, b \in A \quad a + b = b + a$

iii) $\exists\, 0 \in A$ tal que $0 + a = a, \ \forall\, a \in A$

iv) $\forall\, a \in A \ \exists\, -a \in A$ tal que $-a + a = 0$

v) $\forall\, a, b, c \in A \quad a\cdot(b\cdot c) = (a\cdot b)\cdot c$

vi) $\forall\, a, b, c \in A \quad a\cdot(b + c) = (a\cdot b) + (a\cdot c)$ y $(b + c)\cdot a = (b\cdot a) + (c\cdot a)$

De los cuatro primeros postulados de la definición anterior se sigue que un anillo es un grupo abeliano para la primera operación; en consecuencia, todas las propiedades de los grupos y de los grupos abelianos son válidas en la estructura (A,+) conocida como "la estructura aditiva" del anillo.

Al elemento 0 del postulado iii), que representa al idéntico para la primera operación, se le conoce como "el cero" del anillo. Cabe enfatizar que este elemento no es el número cero necesariamente; incluso, el conjunto A puede estar formado por elementos que no sean números.

El postulado v) de la definición se refiere a la segunda operación, y establece que ésta debe ser asociativa.

Las propiedades a que se refiere el postulado vi) se conocen como propiedades distributivas: Cuando dos operaciones + y $\cdot$, definidas en un conjunto A, son tales que

$$\forall\, a, b, c \in A \quad a\cdot(b + c) = (a\cdot b) + (a\cdot c)$$

se dice que la operación $\cdot$ es distributiva por la izquierda sobre la operación +, y cuando son tales que

$$\forall\, a, b, c \in A \quad (b + c)\cdot a = (b\cdot a) + (c\cdot a)$$

se dice que $\cdot$ es distributiva por la derecha sobre +.

Como ejemplos de sistemas con estructura de anillo tenemos los siguientes:

- Los números enteros con la adición y la multiplicación.
- Los números racionales con la adición y la multiplicación.
- Los números reales con la adición y la multiplicación.
- Los números complejos con la adición y la multiplicación.
- Los polinomios con la adición y la multiplicación.
- Las matrices cuadradas de orden n con la adición y la multiplicación.
- El conjunto $S = \{2m \mid m \in Z\}$ con la adición y la multiplicación definidas en Z.

De manera similar al concepto de subgrupo, un subconjunto de un anillo que es un anillo para las mismas operaciones, se dice que es un subanillo de éste. Así, el anillo del último ejemplo que acabamos de mencionar es un subanillo de (Z, +, $\cdot$); que a su vez lo es de (Q, +, $\cdot$); etc.

- Anillos conmutativos y Anillos con unidad

En la definición de anillo no se indica que la segunda operación deba ser conmutativa, ni que deba tener elemento idéntico; en caso de que tales propiedades se satisfagan los anillos toman los nombres específicos que se indican a continuación.

VIII.3.2 DEFINICION

Sea (A, +, ·) un anillo:

i) Si $\forall$ a, b ε A $\quad a \cdot b = b \cdot a$

se dice que el anillo es conmutativo.

ii) Si $\exists$ 1 ε A tal que $1 \cdot a = a = a \cdot 1, \quad \forall\, a \,\varepsilon\, A$

se dice que el anillo tiene unidad.

Al elemento 1 de la definición anterior, que es idéntico para la segunda operación, se le conoce como "la unidad"† del anillo. Nuevamente, este elemento no es el número uno necesariamente.

De los sistemas que acabamos de citar como ejemplos de anillos, todos son anillos conmutativos con excepción de las matrices y todos son anillos con unidad excepto el del conjunto S.

- Dominios Enteros

Cuando dos elementos a y b de un anillo son tales que

$$a \neq 0, \; b \neq 0 \quad y \quad a \cdot b = 0$$

se dice que son "divisores propios de cero".

La estructura denominada "dominio entero" posee como característica adicional la no existencia de divisores propios de cero, como lo establece la siguiente definición.

† *Se emplea aquí el artículo determinado "la" porque dicho elemento es único, como el lector podrá demostrar fácilmente.*

VIII.3.3 DEFINICION

Sea (A, +, ·) un anillo conmutativo con unidad de por lo menos dos elementos, donde $0 \neq 1$; si

$$a \cdot b = 0 \implies a = 0 \text{ ó } b = 0$$

se dice que (A, +, ·) es un dominio entero.

Los números enteros y los polinomios, con las operaciones usuales de adición y multiplicación, son ejemplos de dominios enteros.

Las matrices cuadradas de orden n, con la adición y la multiplicación usuales, no constituyen un dominio entero ya que no son un anillo conmutativo y, además, contienen divisores propios de cero.

- Campos

Al incorporar los inversos para la segunda operación se obtiene la estructura algebraica más completa que veremos en este capítulo. Dicha estructura recibe el nombre de "campo"† y contiene las propiedades comunes a los sistemas numéricos más completos algebraicamente; entre los que se encuentran los números racionales, los números reales y los números complejos con sus respectivas operaciones de adición y multiplicación.

Un campo es un anillo conmutativo con unidad cuyos elementos distintos del cero tienen inverso para la segunda operación, como se establece en la siguiente definición.

† *Algunos autores le llaman cuerpo*

VIII.3.4 DEFINICION

Sea K un conjunto de por lo menos dos elementos, y sean + y · dos operaciones binarias definidas en K. El sistema (K, +, ·) tiene estructura de campo si:

i) $\forall$ a, b, c ε K a + (b + c) = (a + b) + c

ii) $\forall$ a, b ε K a + b = b + a

iii) $\exists$ *0* ε K tal que *0* + a = a, $\forall$ a ε K

iv) $\forall$ a ε K $\exists$ -a ε K tal que -a + a = *0*

v) $\forall$ a, b, c ε K a·(b·c) = (a·b)·c

vi) $\forall$ a, b ε K a·b = b·a

vii) $\exists$ *1* ε K tal que *1*·a = a, $\forall$ a ε K

viii) $\forall$ a ε K, a $\neq$ *0*, $\exists$ a^{-1} tal que a^{-1}·a = *1*

ix) $\forall$ a, b, c ε K a·(b + c) = (a·b) + (a·c)
y
(b + c)·a = (b·a) + (c·a)

En esta definición se listan, una a una, todas las propiedades que debe satisfacer un sistema para tener estructura de campo. Puede demostrarse fácilmente que dicha definición es equivalente a la que, de manera más compacta, se enuncia a continuación.

VIII.3.4' DEFINICION

Sea K un conjunto de por lo menos dos elementos, y sean + y · dos operaciones binarias definidas en K. El sistema (K, +, ·) es un campo si:

i) (K,+) es un grupo abeliano, cuyo elemento idéntico denotamos con *0*.

ii) (K-{*0*},·) es un grupo abeliano.

iii) · es distributiva por la izquierda y por la derecha sobre +.

Posiblemente el lector se pregunte por qué se llegó a la estructura de campo agregando un postulado a la estructura de anillo conmutativo con unidad, en lugar de hacerlo con la de dominio entero qué es más completa. La razón es que, al incorporar los inversos directamente en la estructura de anillo conmutativo con unidad se obtienen como consecuencia la diferencia entre el cero y la unidad y la no existencia de divisores propios de cero; lo cual se deduce del siguiente teorema.

VIII.3.5 TEOREMA

Todo campo es un dominio entero

DEMOSTRACION

Sea (K, +, ·) un campo. Entonces, por los postulados i) a vii) y ix) de la definición VIII.3.4, (K, +, ·) es un anillo conmutativo con unidad de por lo menos dos elementos.

Para demostrar que es un dominio entero debemos probar únicamente que el cero es diferente de la unidad y que no existen divisores propios de cero.

Sean *0* y *1* el cero y la unidad, respectivamente, de (K, +, ·). Probaremos primero el siguiente lema:

0·a = *0*, $\forall$ a ε K

Demostración:

0·a = (*0* + *0*)·a por iii) de VIII.3.4

0·a = (*0*·a) + (*0*·a) por ix) de VIII.3.4

$- (0 \cdot a) + (0 \cdot a) = - (0 \cdot a) + [(0 \cdot a) + (0 \cdot a)]$ por iv) de VIII.3.4

$- (0 \cdot a) + (0 \cdot a) = [- (0 \cdot a) + (0 \cdot a)] + (0 \cdot a)$ por i) de VIII.3.4

$0 = 0 + (0 \cdot a)$ por iv) de VIII.3.4

$0 = 0 \cdot a$ por iii) de VIII.3.4

con lo que queda demostrado el lema.

1) Probaremos ahora que $0 \neq 1$ por contradicción.

Supongamos que

$$0 = 1$$

y consideremos un elemento $a \neq 0$ de K (que debe haberlo ya que K tiene por lo menos dos elementos); entonces

$$0 \cdot a = 1 \cdot a$$

pero por el lema y por vii) de VIII.3.4 se puede concluir que

$$0 = a$$

con lo que se presenta una contradicción. En consecuencia

$$0 \neq 1$$

2) Finalmente, mostraremos que en K no hay divisores propios de cero:

Sea

$$a \cdot b = 0$$

si $a \neq 0$, entonces por viii) de VIII.3.4 $\exists\ a^{-1}$ y

$$a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0$$

de donde se sigue que

$(a^{-1} \cdot a) \cdot b = a^{-1} \cdot 0$ por v) de VIII.3.4

$1 \cdot b = a^{-1} \cdot 0$ por viii) de VIII.3.4

$b = a^{-1} \cdot 0$ por vii) de VIII.3.4

$b = 0$ por el lema 1

si ahora $a \cdot b = 0$ y $b \neq 0$, por vi) de VIII.3.4 podemos escribir

$$b \cdot a = 0$$

y mediante un razonamiento similar al anterior se concluye que

$$a = 0$$

INGENIERIA
ANEXO
BIBLIOTECA

En consecuencia

$$a \cdot b = 0 \implies a = 0 \text{ o } b = 0$$

Esto completa la demostración.

□

Como ejemplos de campos, además de los sistemas (Q, +, ·), (R, +, ·) y (C, +, ·) que ya hemos mencionado, podemos citar al conjunto

$$\{a + b\sqrt{2} \mid a, b \in Q\}$$

con las operaciones de adición y multiplicación de números reales. Dicho sistema es un subcampo de (R, +, ·).

Otro ejemplo de campo lo encontramos en el conjunto {p, q, r, s} con las operaciones + y · definidas por las siguientes tablas

+	p	q	r	s
p	r	s	p	q
q	s	r	q	p
r	p	q	r	s
s	q	p	s	r

·	p	q	r	s
p	s	p	r	q
q	p	q	r	s
r	r	r	r	r
s	q	s	r	p

¿Cuál es el cero y cuál es la unidad de dicho campo?

Para concluir esta sección presentamos la solución al problema de identificar la estructura algebraica del sistema formado por el conjunto de los números racionales y las operaciones ⊕ y ⊙ definidas a continuación

$$x \oplus y = x + y + 1$$

$$x \odot y = x + y + xy, \quad \forall\, x, y \in Q$$

Solución:

Para la operación $\oplus$ tenemos lo siguiente

0) Como $x, y, 1 \in Q$, por i) de I.5.3 se tiene que $x + y + 1 \in Q$; esto es

$$x \oplus y \in Q, \quad \forall\, x, y \in Q$$

y $\oplus$ es una operación definida en Q.

1) Si $x, y, z \in Q$, entonces

$$x \oplus (y \oplus z) = x \oplus (y + z + 1) = x + y + z + 2$$

y

$$(x \oplus y) \oplus z = (x + y + 1) \oplus z = x + y + z + 2$$

por lo que $x \oplus (y \oplus z) = (x \oplus y) \oplus z$.

2) Si $x, y \in Q$ entonces

$$x \oplus y = x + y + 1$$

y

$$y \oplus x = y + x + 1 = x + y + 1$$

por lo que $x \oplus y = y \oplus x$.

3) Sea $x \in Q$ y sea z un idéntico izquierdo para $\oplus$, entonces:

$$z + x + 1 = x$$

$$z = -1$$

y la operación $\oplus$ tiene idéntico izquierdo en Q.

4) Sea $x \in Q$ y sea $\tilde{x}$ un inverso izquierdo de x para $\oplus$; entonces

$$\tilde{x} + x + 1 = -1$$

$$\tilde{x} = -x - 2$$

y todo elemento $x \in Q$ tiene inverso izquierdo en Q para la operación $\oplus$.

En consecuencia, $(Q, \oplus)$ es un grupo abeliano.

Para la operación $\odot$ se tiene lo siguiente

00) Como $x, y \in Q$, por i) de I.5.3 $xy \in Q$ y $x + y + xy \in Q$; esto es

$$x \odot y \in Q, \quad \forall\, x, y \in Q$$

y $\odot$ es una operación definida en Q.

5) Si $x, y \in Q$; entonces

$$x \odot (y \odot z) = x \odot (y + z + yz) = x + y + z + yz + xy + xz + xyz.$$

y

$$(x \odot y) \odot z = (x + y + xy) \odot z = x + y + z + yz + xy + xz + xyz.$$

por lo que $x \odot (y \odot z) = (x \odot y) \odot z$.

6) Si $x, y \in Q$; entonces

$$x \odot y = x + y + xy$$

y

$$y \odot x = y + x + yx = x + y + xy$$

por lo que $x \odot y = y \odot x$.

7) Sea $x \in Q$ y sea u un idéntico izquierdo para $\odot$; entonces

$$u + x + ux = x$$

$$u(1 + x) = 0$$

por lo que $u = 0$ es un idéntico izquierdo para $\odot$ en Q.

8) Sea $x \in Q$ y sea $x^{\urcorner}$ un inverso izquierdo de x para $\odot$; entonces

$$x^{\urcorner} + x + x^{\urcorner}x = 0$$

$$x^{\urcorner}(1 + x) = -x$$

si $x \neq -1$ se sigue que

$$x' = \frac{-x}{1 + x} \in Q$$

Por otra parte, si $x = -1$ y suponemos que $\exists\ x'$ se tendrá que

$$x' \cdot 0 = 1$$

lo cual es un absurdo y, por lo tanto, $\nexists\ -1'$.

En consecuencia, todo elemento $x \in Q$, excepto $x = -1$, tiene inverso izquierdo para $\odot$ (Recuérdese que -1 es el cero de la estructura).

9) Si $x, y, z \in Q$ entonces

$$\begin{aligned} x \odot (y \oplus z) &= x \odot (y + z + 1) \\ &= x + y + z + 1 + xy + xz + x \\ x \odot (y \oplus z) &= 2x + y + z + xy + xz + 1 \end{aligned}$$

$$\begin{aligned} (x \odot y) \oplus (x \odot z) &= (x + y + xy) \oplus (x + z + xz) \\ &= x + y + xy + x + z + xz + 1 \\ (x \odot y) \oplus (x \odot z) &= 2x + y + z + xy + xz + 1 \end{aligned}$$

por lo que

$$x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$$

Por otra parte

$$\begin{aligned} (y \oplus z) \odot x &= (y + z + 1) \odot x \\ &= y + z + 1 + x + yx + zx + x \\ (y \oplus z) \odot x &= 2x + y + z + yx + zx + 1 \end{aligned}$$

$$\begin{aligned} (y \odot x) \oplus (z \odot x) &= (y + x + yx) \oplus (z + x + zx) \\ &= y + x + yx + z + x + zx + 1 \\ (y \odot x) \oplus (z \odot x) &= 2x + y + z + yx + zx + 1 \end{aligned}$$

por lo que

$$(y \oplus z) \odot x = (y \odot x) \oplus (z \odot x)$$

En conclusión: el sistema $(Q, \oplus, \odot)$ tiene estructura de campo.

VIII.3.6 EJERCICIOS

1.- Sea $(A, +, \cdot)$ un anillo. Demostrar que $\forall\ a, b \in A$:

a) $-(-a) = a$

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

2.- Si $A = \{\circ, \square\}$ y $+, \cdot$ son las operaciones definidas por las siguientes tablas

$+$	$\circ$	$\square$
$\circ$	$\circ$	$\square$
$\square$	$\square$	$\circ$

$\cdot$	$\circ$	$\square$
$\circ$	$\circ$	$\circ$
$\square$	$\circ$	$\square$

a) Demostrar que $(A, +, \cdot)$ es un anillo

b) ¿Es conmutativo?

c) ¿Tiene unidad?

3.- Sea $(A, +, \cdot)$ un anillo y sea $S \subset A$. Si definimos $a - b = a + (-b)$, demostrar que si:

$$a - b \in S \quad \text{y} \quad a \cdot b \in S, \ \forall\ a, b \in S$$

entonces S es un subanillo de A.

4.- En un anillo no conmutativo $(A, +, \cdot)$ se define una operación $*$ como

$$x * y = x \cdot y - y \cdot x$$

Calcular

$$x * (y * z) + y * (z * x) + z * (x * y)$$

5.- Sea el sistema (Z,+) cuya estructura es de grupo abeliano. Si definimos una operación $*$ como

$a*b = kab$; $\forall$ a, b $\in$ Z, donde k es un entero

a) Demostrar que el sistema (Z, +, $*$) tiene estructura de anillo conmutativo.

b) Si k = 1, ¿Qué estructura tiene (Z, +, $*$)?

c) Si k = 0, demostrar que (Z, +, $*$) no es un dominio entero.

6.- Demostrar que el conjunto $\{a + b\sqrt{2} \mid a, b \in Q\}$ forma un campo para las operaciones de adición y multiplicación definidas en R.

5815804

VIII.4 ISOMORFISMOS Y HOMOMORFISMOS

El concepto de isomorfismo es de relevante importancia en las matemáticas; especialmente desde el punto de vista de sus aplicaciones. Dicho concepto se encuentra subyacente en el empleo mismo de modelos matemáticos, ya sea para la resolución de problemas físicos o para la resolución de problemas matemáticos más complicados.

El término "isomorfo", que etimológicamente significa "de igual forma", se emplea en el álgebra para denotar la idea de que dos sistemas son tan parecidos que pueden considerarse, en esencia, como el mismo.

Por ejemplo, al analizar las tablas siguientes

·	1	-1
1	1	-1
-1	-1	1

□	p	q
p	p	q
q	q	p

nos queda la impresión de que podríamos reemplazar los símbolos 1, -1 y · por los símbolos p, q y □, respectivamente. Con ello cambiaríamos un sistema por otro sin alterar esencialmente los resultados.

En general, la sustitución de los elementos de un conjunto A por los elementos de otro conjunto B puede hacerse mediante una función $f: A \rightarrow B$. Cuando dicha función es biyectiva los elementos de A y de B se encuentran en relación "uno a uno", y cada uno de ellos puede considerarse como el "reflejo" de su elemento correspondiente en el otro conjunto.

Si en el conjunto A está definida una operación $*$ y en el conjunto B una operación Δ, es necesario que los resultados obtenidos en el sistema $(A,*)$ se conserven al efectuar las operaciones en el sistema (B,Δ) con las imágenes respectivas; por lo que la función f debe ser tal que

$$f(a*b) = f(a)\ \Delta\ f(b);\quad \forall\ a,\ b \in A$$

Esta propiedad de la función f garantiza que se llega al mismo resultado empleando cualquiera de los dos procedimientos siguientes:

1) Efectuando la operación $*$ en el sistema $(A,*)$ y aplicando después la función f al resultado.

2) Aplicando la función f a cada uno de los elementos y efectuando después la operación Δ con las imágenes en el sistema (B,Δ).

La diferencia entre ambos procedimientos puede apreciarse claramente en el diagrama siguiente

$$\begin{tikzcd}
(a,b) \arrow[r, "*"] \arrow[d, Rightarrow, "f"'] & a*b \arrow[d, "f"] \\
(f(a),f(b)) \arrow[r, Rightarrow, "\Delta"] & \begin{array}{c} f(a*b) \\ f(a)\ \Delta\ f(b) \end{array}
\end{tikzcd}$$

Una función que satisface dicha propiedad se dice que es un "homomorfismo" de $(A,*)$ en (B,Δ). Cuando la función es, además, biyectiva se dice que es un "isomorfismo" entre $(A,*)$ y (B,Δ), y que dichos sistemas son "isomorfos".

El caso particular del isomorfismo ofrece una ventaja adicional:

por ser f una función biyectiva existe su inversa f^{-1} y podemos, mediante esta última, emprender el "regreso" del sistema (B,Δ) al sistema $(A,*)$ una vez que se ha efectuado la operación. Esto último fundamenta el empleo de modelos matemáticos para la resolución de problemas físicos ya que, debido al isomorfismo, es posible emplear un proceso como el que se indica en el siguiente diagrama.

$$\begin{tikzcd}
\text{elementos físicos} \arrow[r, dashed, "\textit{operación}", "\textit{física}"'] \arrow[d, "\begin{array}{c}\textit{modelo}\\ (f)\end{array}"'] & \text{resultado físico} \\
\text{elementos matemáticos} \arrow[r, "\textit{operación}", "\textit{matemática}"'] & \text{resultado matemático} \arrow[u, "\begin{array}{c}\textit{interpretación}\\ \textit{del modelo } (f^{-1})\end{array}"']
\end{tikzcd}$$

A continuación se presentan las definiciones formales de estos conceptos así como algunos ejemplos y teoremas relacionados.

VIII.4.1 DEFINICION Sean $(G,*)$ y (G',Δ) dos grupos. Una función $f: G \to G'$ es un homomorfismo si $f(a*b) = f(a)\ \Delta\ f(b),\quad \forall\ a,\ b \in G$ Si f es, además, biyectiva, se dice que es un isomorfismo y que los grupos $(G,*)$ y (G',Δ) son isomorfos.

Por ejemplo, para los grupos

$(Z,+)$ y $(S,\cdot)$, donde $S = \{1,\ -1,\ i,\ -i\}$, la función

$$f(m) = i^m \ , \quad \forall\, m \in Z$$

constituye un homomorfismo; ya que, $\forall\, m,\ n \in Z$:

$$f(m + n) = i^{m+n} = i^m \cdot i^n = f(m) \cdot f(n)$$

pero no se trata de un isomorfismo puesto que f no es biyectiva. Es claro que no puede existir un isomorfismo entre dichos grupos ya que S es finito (tiene cuatro elementos) y Z es infinito.

Un ejemplo de grupos isomorfos lo encontramos en el conjunto S de matrices simétricas de orden dos con elementos en R y el conjunto R^3 de las ternas ordenadas de números reales; ambos con las operaciones de adición definidas en forma usual.

En efecto, si

$$M = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

representa un elemento arbitrario de S, podemos definir una función $f\colon S \to R^3$ mediante la regla

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \xrightarrow{f} (a,\ b,\ c)$$

Se tendrá entonces, para dos matrices cualesquiera de S

$$M = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad y \quad N = \begin{bmatrix} d & e \\ e & f \end{bmatrix}$$

que

$$f(M + N) = f\left(\begin{bmatrix} a + d & b + e \\ b + e & c + f \end{bmatrix}\right) = (a + d,\ b + e,\ c + f)$$

y que

$$f(M) + f(N) = (a,\ b,\ c) + (d,\ e,\ f) = (a + d,\ b + e,\ c + f)$$

por lo que

$$f(M + N) = f(M) + f(N); \quad \forall\, M,\ N \in S$$

Además, como f es biyectiva se trata de un isomorfismo; en consecuencia, $(S,+)$ y $(R^3,+)$ son isomorfos.

Un conocido ejemplo de isomorfismo entre el grupo multiplicativo $(R^+,\cdot)$ y el grupo aditivo $(R,+)$ lo constituye la función logaritmo; ya que, además de ser biyectiva, cumple con la condición

$$L(x \cdot y) = L(x) + L(y); \quad \forall\, x,\ y \in R^+$$

Como es del conocimiento del lector, la función logaritmo se emplea comúnmente para transformar problemas multiplicativos en problemas aditivos.

VIII.4.2 TEOREMA

Sean $(G,*)$ y (G',Δ) dos grupos, y $f\colon G \to G'$ un homomorfismo:

i) Si e es el idéntico para $*$ y e' es el idéntico para Δ, entonces

$$f(e) = e'$$

ii) Si $\hat{a}$ es el inverso de $a \in G$ para $*$ y $\widehat{f(a)}$ es el inverso de $f(a) \in G'$ para Δ, entonces

$$f(\hat{a}) = \widehat{f(a)}$$

DEMOSTRACION

Se demuestra únicamente la parte i) dejando al lector como ejercicio la demostración de la parte ii).

i) Sea $a \in G$, como e es el idéntico para $*$ en G

$$e*a = a; \quad \forall\, a \in G$$

por lo que

$$f(e*a) = f(a)$$

como f es un homomorfismo

$$f(e*a) = f(e)\ \Delta\ f(a)$$

por lo que

$$f(e)\ \Delta\ f(a) = f(a)$$

y en consecuencia

$$f(e) = e'$$

como queríamos. □

> VIII.4.3 TEOREMA
>
> Sean $(G,*)$ y (G',Δ) dos grupos. Si $f: G \to G'$ es un isomorfismo entonces $f^{-1}: G' \to G$ es un isomorfismo.

DEMOSTRACION

Como $f: G \to G'$ es biyectiva; entonces, de la teoría de funciones, sabemos que existe su inversa $f^{-1}: G' \to G$ y que también es biyectiva.

Para probar que f^{-1} es un homomorfismo consideremos dos elementos $a' = f(a)$ y $b' = f(b)$ de G', donde $a, b \in G$. Entonces

$$f^{-1}(a'\ \Delta\ b') = f^{-1}\left[f(a)\ \Delta\ f(b)\right]$$

Como f es un homomorfismo podemos escribir

$$f^{-1}(a'\ \Delta\ b') = f^{-1}\left[f(a*b)\right]$$

Como $f^{-1}\left[f(x)\right] = x$, se tiene que