APUNTES DE ÁLGEBRA LINEAL

Eduardo Solar González
Leda Speziale de Guzmán

DIVISIÓN DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS BÁSICAS

1985

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE INGENIERIA

APUNTES DE
ALGEBRA LINEAL

EDUARDO SOLAR GONZALEZ
LEDA SPEZIALE DE GUZMAN

DIVISION DE CIENCIAS BASICAS
DEPARTAMENTO DE MATEMATICAS BASICAS

FI/DCB/81-041

P R O L O G O

La presente obra, cuyo contenido comprende los conceptos básicos del Algebra Lineal, es la continuación de nuestro trabajo publicado bajo el título Algebra (primera parte), aunque es prácticamente independiente del mismo.

Sin embargo, para un buen aprovechamiento del material que comprende este volumen, es necesario que el lector conozca las propiedades de los números reales, de los números complejos y de los polinomios, incluyendo el cálculo de raíces. Es conveniente también que tenga conocimientos de álgebra vectorial y de cálculo con funciones de una variable.

En el capítulo V, primero de este volumen, se estudian los sistemas de ecuaciones lineales desde un punto de vista práctico, proporcionándose un método general para obtener soluciones.

Se aborda después de manera amplia el tema de las matrices, fundamentalmente desde el punto de vista algebraico. Se estudian las operaciones de adición, multiplicación por un escalar y multiplicación de matrices, así como sus propiedades, incluyendo el problema de la inversa; se presentan además algunos tipos especiales de matrices de uso frecuente en las aplicaciones.

La definición de determinante se plantea aquí desde el punto de vista tradicional; se estudian sus propiedades fundamentales y se desarrollan método generales para el cálculo de determinantes.
Se incluyen además un par de aplicaciones al cálculo de la inversa

y a la resolución de sistemas de ecuaciones lineales.

En el capítulo VIII se estudian los principales tipos de estructuras algebraicas, introduciéndose además el importante concepto de isomorfismo.

Los últimos dos capítulos, que son los de mayor extensión, comprenden los temas centrales del álgebra lineal desde el punto de vista abstracto.

En el capítulo IX se define matemáticamente la estructura de espacio vectorial, se estudian sus propiedades y conceptos fundamentales inherentes, ilustrándolos con ayuda de ejemplos. Estos conceptos se aplican a la construcción de una teoría para los sistemas de ecuaciones lineales y al tratamiento algebraico de las funciones.

Se estudian también en este capítulo los espacios vectoriales con producto interno y los conceptos métricos correspondientes, concluyendo con un resultado de gran importancia para las aplicaciones.

En el capítulo X se estudian las transformaciones entre espacios vectoriales, sus conceptos fundamentales, el álgebra de las transformaciones lineales y la relación de éstas con las matrices; haciendo énfasis en los conceptos de valor y vector característico y su aplicación al problema de la diagonalización.

Concluye el capítulo con el tratamiento de tipos especiales de operadores en espacios con producto interno, abordando el problema de la diagonalización en estos espacios, así como la descomposición espectral.

En el presente trabajo hemos tratato de conservar la idea del ante rior, en el sentido de buscar una presentación para los conceptos fundamentales que sea accesible al estudiante; sin renunciar a la formalidad mínima que debe tener un libro sobre el tema.

Queremos reiterar aquí nuestro reconocimiento a los profesores con quienes tuvimos el agrado de trabajar en la preparación de las notas que, bajo el título de Apuntes de Algebra, publicó la Facultad de Ingeniería de la UNAM durante seis años y que sirvieron de orientación para la elaboración de este material. Asimismo, deseamos expresar nuevamente nuestro agradecimiento a la Sra. Rosa María Arenas por su excelente trabajo en la mecanografía del manuscrito.

Nuestra gratitud también para la Facultad de Ingeniería de la UNAM, cuyos profesores y autoridades nos brindaron el apoyo que hizo posible la elaboración de este trabajo, que esperamos le sea de utilidad a la institución.

Finalmente, conscientes de que a pesar de nuestro empeño habrán de aparecer errores, suplicamos la comprensión de los lectores y solicitamos su colaboración para hacernos llegar sus críticas y sugerencias, las que seguramente contribuirán a mejorar futuras ediciones.

EDUARDO SOLAR GONZALEZ

LEDA SPEZIALE DE GUZMAN

C O N T E N I D O

CAPITULO IX: ESPACIOS VECTORIALES

CAPITULO X: TRANSFORMACIONES LINEALES

CAPITULO V SISTEMAS DE ECUACIONES LINEALES

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.

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

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

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

A los símbolos $x_1, x_2, \dots, x_n$ 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

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

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

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

- 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.

Entonces la ecuación es de la forma

$$0x_1 + 0x_2 + .. + 0x_n = 0$$

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

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

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

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

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

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

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

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.

V.2.2 DEFINICION

Una solución del sistema de ecuaciones lineales

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

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

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

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 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 consisten 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

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

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

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

Si $(k_1, k_2, \ldots, k_n)$ es una solución de S entonces satisface todas 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')

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

se tiene que

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

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

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

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

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

y multiplicando ahora la primera ecuación por 2 y sumando el resultado a la tercera se obtiene

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

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

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

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 satisfacen "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 puede 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

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

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

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

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

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

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 trans - formaciones necesarias para llevarla a la forma escalonada. Hare mos 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, intercam - biando 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 se gundo y, a continuación, multiplicando el mismo primer renglón ɔr 2 y sumando al tercero obtenemos la matriz

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

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

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

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

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

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

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

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

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

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

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

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

$$x_1 = \frac{5}{4} - a + \frac{1}{4}b - \frac{3}{4}c$$

$$x_2 = a$$

$$x_3 = -\frac{1}{4} + \frac{7}{4}b + \frac{7}{4}c$$

$$x_4 = b$$

$$x_5 = c$$

donde a, b y c pueden tomar cualquier valor.

Consideremos ahora el sistema

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

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

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

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

$$a)\quad \begin{aligned} 2x_1 + x_2 - x_3 &= 0 \\ x_1 + 2x_2 + 4x_3 &= 0 \\ 3x_1 + 2x_2 + 3x_3 &= 0 \end{aligned} \qquad\qquad b)\quad \begin{aligned} 2x + 6y + z &= 0 \\ x + 3y &= 0 \\ -x - 3y + 4z &= 0 \end{aligned}$$

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.

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×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

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

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

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

como el segundo renglón, y en general al arreglo horizontal

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

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 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 $m \times n$ con elementos en C. Diremos que A y B son iguales, lo que representaremos con $A = B$, si:

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

Así, por ejemplo, las matrices

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

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

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

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 $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices de $m\times n$ con elementos en C. La suma $A + B$ es una matriz $S = \left[s_{ij}\right]$, de $m\times n$, 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

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

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.

Por VI.2.1 se tiene que

$$A + B = \begin{bmatrix} s_{ij} \end{bmatrix} = \begin{bmatrix} a_{ij} + b_{ij} \end{bmatrix}$$

y

$$B + A = \begin{bmatrix} t_{ij} \end{bmatrix} = \begin{bmatrix} b_{ij} + a_{ij} \end{bmatrix}$$

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 $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ una matriz de m×n con elementos en $\mathbb{C}$.

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

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

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 $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ una matriz de m×n con elementos en $\mathbb{C}$.

Si definimos la matriz $-A = \begin{bmatrix} v_{ij} \end{bmatrix}$ como $v_{ij} = -a_{ij}\ \forall$ i, j; entonces

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

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 $A = [a_{ij}]$ y $B = [b_{ij}]$ 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

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

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

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

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 $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices de m×n con ele - mentos en C y α un escalar de C, entonces

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

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

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

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

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

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

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

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

$$AX = B \qquad ---(2)$$

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

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

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

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

Veamos ahora cómo puede obtenerse la matriz AX a partir de las ma trices 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:

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

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

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

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

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

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

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

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

$$p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \quad ; \text{ para } i = 1,\ldots,m \text{ y } j = 1,\ldots,q.$$

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".

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 $n \times q$. 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

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

Obtengamos primero el producto

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

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 $m\times n$, $n\times p$ y $n\times p$, respectivamente; entonces

$$B + C = \left[b_{ij} + c_{ij}\right] \qquad \text{por VI.2.1}$$

$$A(B + C) = \left[\sum_{k=1}^{n} a_{ik}\,(b_{kj} + c_{kj})\right] \qquad \text{por VI.3.1}$$

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

$$= \left[\sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}\right] \qquad \text{por ii) y iii) de II.1.4}$$

$$= \left[\sum_{k=1}^{n} a_{ik}b_{kj}\right] + \left[\sum_{k=1}^{n} a_{ik}c_{kj}\right] \qquad \text{por VI.2.1}$$

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

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 $I_m = \left[\delta_{ij}\right]$

$I_mA = \left[\sum_{k=1}^{m} \delta_{ik}a_{kj}\right]$	por VI.3.1
$= \left[\delta_{ii}a_{ij}\right]$	por VI.3.4
$= \left[1 \cdot a_{ij}\right]$	por VI.3.4
$= \left[a_{ij}\right]$	por iv) de II.1.4
I_mA -	como se quería.

$\square$

VI.3.6 EJERCICIOS

1.- Para las siguientes matrices

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

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 $n \times n$ 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

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

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×n no singular, y sean X, Y dos inversas de A; entonces, por VI.4.1

$$XA = I_n = AX \qquad y \qquad YA = I_n = AY$$

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×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

$$\begin{aligned}
(B^{-1} A^{-1})(AB) &= (B^{-1} A^{-1})\left[(A)(B)\right] & \\
&= \left[(B^{-1} A^{-1})A\right]B & \text{por VI.3.2} \\
&= \left[B^{-1} (A^{-1} A)\right]B & \text{por VI.3.2} \\
&= (B^{-1} I_n)B & \text{por VI.4.1} \\
&= B^{-1} B & \text{por VI.3.5} \\
(B^{-1} A^{-1})(AB) &= I_n & \text{por VI.4.1}
\end{aligned}$$

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

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 trans - formación elemental" que es "premultiplicar por una matriz elemental".

Es claro que existen tres tipos de matrices elementales, corres - pondientes 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 ele mentales.

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:

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

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

$$P = E_k \ldots E_2 \, E_1$$

$$P = (E_k \ldots E_2 \, E_1) \, I_n$$

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

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

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

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

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

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

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 singu -

lar; 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.

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 $\left[A \mid I_3\right]$ y tratemos de obtener en el lado izquierdo la matriz identidad

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

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

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

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

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

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

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

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

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

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

$$AX + 0 = 3X + (-B) \qquad \text{por iv) de VI.2.2}$$

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

Para "pasar" ahora la matriz 3X al miembro izquierdo de la ecuación:

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

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

de donde

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

$$-(3X) + AX = 0 + (-B) \qquad \text{por iv) de VI.2.2}$$

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

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 $m\times n$ con elementos en C:

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

$$= 0 \cdot X$$ por v) de II.1.4

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

de donde

$$(-\alpha) X = -(\alpha X)$$ por iv) de VI.2.2

Llevando este resultado al desarrollo anterior podemos escribir

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

de donde se sigue que

$$(-3)(IX) + AX = -B$$ por i) de VI.3.5

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

$$(-3I) X + AX = -B$$ por lo que acabamos de demostrar

$$(-3I + A) X = -B$$ 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

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

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

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

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

$$B = C \;\not\Longrightarrow\; AB = CA$$

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 $AB = AC \Rightarrow B = C$

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

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

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

b) $XA + B = XC$

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

c) $AX + C = B$

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

d) $A + XB = XC$

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

3.- Para las matrices

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

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:

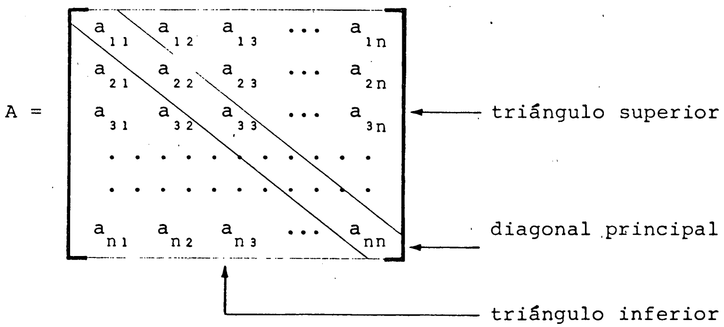
i) $AB = AC \implies B = C$
ii) $BA = CA \implies B = C$
iii) $AB = CA \not\Rightarrow B = C$

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":



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

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

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 $n \times n$ con elementos en C y $\alpha \in C$:
>
> i) $\text{tr}(A+B) = (\text{tr } A) + (\text{tr } B)$
>
> ii) $\text{tr}(\alpha A) = \alpha(\text{tr } A)$
>
> iii) $\text{tr}(AB) = \text{tr}(BA)$

DEMOSTRACION

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

Sean $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices de $n \times n$ con elementos en C y sea $\alpha \in C$:

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

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

ii) $\operatorname{tr}(\alpha A) = \operatorname{tr}\left[\alpha a_{ij}\right]$ 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 $A = \left[a_{ij}\right]$ una matriz de $n\times n$ 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 triangular 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 superior y en la diagonal principal. Por ejemplo, las siguientes matrices 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 $A = \left[a_{ij}\right]$ y $B = \left[b_{ij}\right]$ dos matrices triangulares superiores de orden n. De VI.3.1 se sigue que

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

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 $k \geqslant i$, 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 $A = \left[a_{ij}\right]$ una matriz de $n \times n$ 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

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

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

$A = \text{diag}(a_{11},\ a_{22}, \ldots,\ a_{nn})$ y $B = \text{diag}(b_{11},\ b_{22}, \ldots,\ b_{nn})$ y $\alpha \in C$, entonces:

i) $A+B = \text{diag}(a_{11}+b_{11},\ a_{22}+b_{22}, \ldots,\ a_{nn}+b_{nn})$

ii) $\alpha A = \text{diag}(\alpha a_{11},\ \alpha a_{22}, \ldots,\ \alpha a_{nn})$

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

iv) $A^{-1} = \text{diag}(\frac{1}{a_{11}},\ \frac{1}{a_{22}}, \ldots,\ \frac{1}{a_{nn}})$, 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

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

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

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

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

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

y la prueba termina. $\square$

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

$$\begin{bmatrix} \alpha & 0 & 0 & \ldots & 0 \\ 0 & \alpha & 0 & \ldots & 0 \\ 0 & 0 & \alpha & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot\;\cdot\;\cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\;\cdot\;\cdot & \cdot \\ 0 & 0 & 0 & \ldots & \alpha \end{bmatrix}$$

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 $\quad tr(AB) = tr(BA)$

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

$$tr(AB) = tr(BA)$$

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

$$tr(AB) \neq (tr\ A)(tr\ B)$$

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

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

3.- a) Si A es una matriz de 2×2 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 $n\times n$, 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 $(AB)(CD)^{-1}$

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 m×n con elementos en C.
> Se llama transpuesta de A a la matriz de n×m
> $A^T = \left[c_{ij} \right]$ tal que
>
> $$c_{ij} = a_{ji}$$

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

VI.3.1

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

y en consecuencia

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

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

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

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

- 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

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

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

Por ejemplo, la siguiente matriz es simétrica

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

De manera similar, para las matrices antisimétricas se tiene

$$A = -A^T$$

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

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

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

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

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 $n\times n$ 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 $n\times n$, por VI.7.1 A^T es también de $n\times n$ 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 $\overline{A} = \left[c_{ij}\right]$ tal que

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

Así, por ejemplo para la matriz

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

se tiene que

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

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) $(\overline{\overline{A}}) = A$

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

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

iv) $\overline{AB} = \overline{A}\,\overline{B}$, 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 $m\times n$ y $n\times q$ respectivamente; y sean $A = \left[c_{ij}\right]$ y $B = \left[d_{ij}\right]$ sus respectivas conjugadas. Entonces de VI.3.1

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

y en consecuencia

$$\overline{AB} = \left[\overline{\sum_{k=1}^{n} a_{ik}b_{kj}}\right] \qquad \text{por VI.7.6}$$

$$= \left[\sum_{k=1}^{n} \overline{a_{ik}b_{kj}}\right] \qquad \text{por v) de II.1.6}$$

$$= \left[\sum_{k=1}^{n} \overline{a}_{ik}\overline{b}_{kj}\right] \qquad \text{por vi) de II.1.6}$$

$$= \left[\sum_{k=1}^{n} c_{ik}d_{kj}\right] \qquad \text{por VI.7.6}$$

$$\overline{AB} = \overline{A}\ \overline{B} \qquad \text{por VI.3.1}$$

□

- Matrices reales e imaginarias

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

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

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

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

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

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

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 $m \times n$ con elementos en $\mathbb{C}$. Se llama conjugada-transpuesta de A, y se representa con A^*, a la matriz de $n \times m$ 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 $m \times n$ con elementos en $\mathbb{C}$, 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 nxn 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

$$a_{ii} = \overline{a}_{ii} \quad \Rightarrow \quad I(a_{ii}) = 0$$

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

Así, por ejemplo, la siguiente matriz es hermitiana

$$A = \begin{bmatrix} -1 & 5 & 2+i \\ 5 & 3 & i \\ 2-i & -i & 0 \end{bmatrix} \quad \text{ya que} \left\{ \begin{array}{l} a_{12} = \bar{a}_{21} = 5 \\ a_{13} = \bar{a}_{31} = 2+i \\ a_{23} = \bar{a}_{32} = i \\ I(a_{11}) = I(a_{22}) = I(a_{33}) = 0 \end{array} \right.$$

Para los elementos de una matriz antihermitiana se tiene que

$$a_{ij} = -\bar{a}_{ji} \qquad \forall\ i,j$$

Es decir que los elementos "simêtricos" 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

$$a_{ii} = -\bar{a}_{ii} \quad \Longrightarrow \quad R(a_{ii}) = 0$$

por lo que los elementos de la diagonal principal deben ser núme - ros imaginarios, como en el caso de la siguiente matriz que es an- tihermitiana

$$A = \begin{bmatrix} -i & -5 & -2-i \\ 5 & 3i & -i \\ 2-i & -i & 0 \end{bmatrix} \quad \text{ya que} \left\{ \begin{array}{l} a_{12} = -\bar{a}_{21} = -5 \\ a_{13} = -\bar{a}_{31} = -2-i \\ a_{23} = -\bar{a}_{32} = -i \\ R(a_{11}) = R(a_{22}) = R(a_{33}) = 0 \end{array} \right.$$

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 $m\times n$ 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 mxm con elementos en C y sea $n \in N$. Se llama potencia enésima de A, y se representa con A^n, a la matriz definida por

$$A^0 = I_m$$

$$A^n = A\ A^{n-1}, \text{ para } n \geq 1$$

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

$$A^m\ A^k = A^{m+k}$$

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.

□

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

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 $m\times n$ 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×n y n×q, respectivamente, entonces AB es real.

8.- Demostrar que si A es una matriz de m×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) $(\alpha A)^* = \overline{\alpha}\ A^* \qquad \forall\ \alpha \in C$

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

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \quad C = \begin{bmatrix} c_{11} & c_{12} \end{bmatrix} \quad y \quad D = \begin{bmatrix} d_{11} \\ d_{21} \end{bmatrix}$$

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 exprésemosla como una hipermatriz, agrupando sus elementos en submatrices de la siguiente manera

$$\left[\begin{array}{cc:c:c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{array}\right]$$

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.

$$\left[\begin{array}{cccc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{array}\right] = \begin{bmatrix} A_{11} \\ \\ A_{21} \end{bmatrix}$$

donde se tienen ahora sólo dos submatrices.

En términos generales, puede decirse que una partición de una matriz es la expresión de ésta como una hipermatriz, mediante la agrupación de sus elementos en submatrices; sin embargo, no cualquier 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 tienen 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

$$
\begin{array}{ll}
 & \begin{array}{cccc} \overbrace{\text{cols.}}^{n_1} & \overbrace{\text{cols.}}^{n_2} & & \overbrace{\text{cols.}}^{n_t} \end{array} \\
\begin{array}{l} m_1 \text{ renglones } \} \\ m_2 \text{ renglones } \} \\ \\ m_s \text{ renglones } \} \end{array} &
\left[\begin{array}{c|c|c|c}
A_{11} & A_{12} & \cdots & A_{1t} \\ \hline
A_{21} & A_{22} & \cdots & A_{2t} \\ \hline
\vdots & \vdots & & \vdots \\ \hline
A_{s1} & A_{s2} & \cdots & A_{st}
\end{array}\right]
\end{array}
$$

donde:

$$m_1+m_2+\ldots+m_s = m \quad y \quad n_1+n_2+\ldots+n_t = n$$

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

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ A_{s1} & A_{s2} & \cdots & A_{st} \end{bmatrix}$$

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

$$A_{21} = \left[a_{ij} \right] \text{ con } \begin{cases} i = 3,4,5 \\ j = 1,2 \end{cases}$$

- 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

$(m_1, m_2, \ldots, m_s)$ 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

$(m_1, m_2, \ldots, m_s)$ 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

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

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

$$A_{rk} = \left[a_{ij}\right] \text{ con } \begin{cases} i = \left[m_1+\ldots+(m_{r-1}+1)\right], \ldots, \left[m_1+\ldots+m_r\right] \\ j = \left[n_1+\ldots+(n_{k-1}+1)\right], \ldots, \left[n_1+\ldots+n_k\right] \end{cases}$$

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

$$H_{rc} = \left[\sum_{\ell=1}^{n} a_{i\ell} b_{\ell j} \right] \quad \text{con} \quad \begin{cases} i = \left[m_1 + \ldots + (m_{r-1}+1) \right], \ldots, \left[m_1 + \ldots + m_r \right. \\ j = \left[q_1 + \ldots + (q_{c-1}+1) \right], \ldots, \left[q_1 + \ldots + q_c \right. \end{cases}$$

se tiene entonces que $H_{rc} = P_{rc}$ y la prueba termina.

□

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

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, considere̲mos el problema de resolver el siguiente sistema de ecuaciones linea̲les

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \qquad (1)$$

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 sis̲tema (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

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{aligned} & a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + \\ & + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{aligned} \qquad (9)$$

Por ejemplo

$$\begin{vmatrix} 1 & 0 & -3 \\ -4 & 5 & 2 \\ 1 & -2 & 0 \end{vmatrix} = \begin{aligned} & (1)(5)(0) - (1)(2)(-2) - (0)(-4)(0) + \\ & + (0)(2)(1) + (-3)(-4)(-2) - (-3)(5)(1) = \\ = \; & 0 + 4 - 0 + 0 - 24 + 15 = -5 \end{aligned}$$

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
>
> $$(\alpha_1 \quad \alpha_2, \ldots, \alpha_n)$$
>
> donde $\alpha_i \in S \ \forall i$ y $\alpha_i \neq \alpha_j$ si $i \neq j$

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

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,..., 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

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

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

$$a_{1\alpha_1} \quad a_{2\alpha_2} \quad \cdots \quad a_{n\alpha_n}$$

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

$$a_{1\alpha_{k1}} \quad a_{2\alpha_{k2}} \quad \cdots \quad a_{n\alpha_{kn}}, \quad \text{donde } k = 1, 2, \ldots, n!$$

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

$$\varepsilon(p_k)\ a_{1\alpha_{k1}} \quad a_{2\alpha_{k2}} \quad \cdots \quad a_{n\alpha_{kn}}$$

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 $\{1, 2, \ldots, n\}$. Se llama determinante de A al número

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k) \prod_{i=1}^{n} a_{i\alpha_{ki}}$$

$$\text{donde} \quad \varepsilon(p_k) = \begin{cases} +, & \text{si } p_k \text{ es de clase par} \\ -, & \text{si } p_k \text{ es de clase impar} \end{cases}$$

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.

$\square$

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 número 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.

□

Una primera propiedad de los determinantes consiste en "descomponer" 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 expresión

$$
\begin{array}{l}
\\ \\ \\ \text{renglón } r \rightarrow \\ \\ \\
\end{array}
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\cdots & \cdots & \cdots & \cdots \\
b_{r1}+c_{r1} & b_{r2}+c_{r2} & \cdots & b_{rn}+c_{rn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix} =
$$

$$
= \begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\cdots & \cdots & \cdots & \cdots \\
b_{r1} & b_{r2} & \cdots & b_{rn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
+
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\cdots & \cdots & \cdots & \cdots \\
c_{r1} & c_{r2} & \cdots & c_{rn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
$$

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 = λ 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

$$\det B = \sum_{k=1}^{n!} \varepsilon(p_k)\, b_{1\alpha_{k_1}}\, b_{2\alpha_{k_2}} \cdots b_{r\alpha_{kr}} \cdots b_{n\alpha_{kn}}$$

por lo que

$$\det B = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k_1}}\, a_{2\alpha_{k_2}} \cdots \lambda a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

$$= \lambda \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k_1}}\, a_{2\alpha_{k_2}} \cdots a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}}$$

$\det B = \lambda \det A$

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

$$\varepsilon(p_k)\, b_{1\alpha_{k_1}}\, b_{2\alpha_{k_2}} \cdots b_{r\alpha_{kr}} \cdots b_{s\alpha_{ks}} \cdots b_{n\alpha_{kn}}$$

donde el signo $\varepsilon(p_k)$ depende de la permutación

$$p_k = (\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kr}, \ldots, \alpha_{ks}, \ldots, \alpha_{kn})$$

sustituyendo los elementos de B por su expresión en términos de los de A se obtiene

$$\varepsilon(p_k)\ a_{1\alpha_{k_1}}\ a_{2\alpha_{k_2}}\ \dots\ a_{s\alpha_{kr}}\ \dots\ a_{r\alpha_{ks}}\ \dots\ a_{n\alpha_{kn}}$$

que puede escribirse como

$$\varepsilon(p_k)\ a_{1\alpha_{k_1}}\ a_{2\alpha_{k_2}}\ \dots\ a_{r\alpha_{ks}}\ \dots\ a_{s\alpha_{kr}}\ \dots\ a_{n\alpha_{kn}}$$

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

$$a_{1\alpha_{k_1}}\ a_{2\alpha_{k_2}}\ \dots\ a_{r\alpha_{ks}}\ \dots\ a_{s\alpha_{kr}}\ \dots\ a_{n\alpha_{kn}}$$

le corresponde el signo $\varepsilon(q_k)$ que depende de la permutación

$$q_k = (\alpha_{k_1}, \alpha_{k_2}, \dots, \alpha_{ks}, \dots, \alpha_{kr}, \dots, \alpha_{kn})$$

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

$$\sum_{k=1}^{n!} \varepsilon(p_k)\ b_{1\alpha_{k_1}}\ b_{2\alpha_{k_2}}\ \dots\ b_{n\alpha_{kn}} =$$

$$= \sum_{k=1}^{n!} -\varepsilon(p_k)\ a_{1\alpha_{k_1}}\ a_{2\alpha_{k_2}}\ \dots\ a_{n\alpha_{kn}}$$

y en consecuencia

$$\det B = -\det A$$

iv) Sea $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ 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 $B = \begin{bmatrix} b_{ij} \end{bmatrix}$ 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

$$\det B = \det A$$

□

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×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 índi -

ces, 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 $n \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad y \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}$$

y consideremos una partición de B escribiéndola como

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} \qquad \text{donde } B_i = (b_{i1}, b_{i2}, \dots, b_{in})$$

El producto AB también puede escribirse como.

$$AB = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix} \qquad \text{donde } C_i = a_{i1}B_1 + a_{i2}B_2 + \dots + a_{in}B_n$$

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

$$D_1 = \begin{vmatrix} a_{11}B_1 \\ a_{21}B_1 \\ \vdots \\ C_n \end{vmatrix} + \begin{vmatrix} a_{11}B_1 \\ a_{22}B_2 \\ \vdots \\ C_n \end{vmatrix} + \ldots + \begin{vmatrix} a_{11}B_1 \\ a_{2n}B_n \\ \vdots \\ C_n \end{vmatrix} = D_{11} + D_{12} + \ldots + D_{1n}$$

$$D_2 = \begin{vmatrix} a_{12}B_2 \\ a_{21}B_1 \\ \vdots \\ C_n \end{vmatrix} + \begin{vmatrix} a_{12}B_2 \\ a_{22}B_2 \\ \vdots \\ C_n \end{vmatrix} + \ldots + \begin{vmatrix} a_{12}B_2 \\ a_{2n}B_n \\ \vdots \\ C_n \end{vmatrix} = D_{21} + D_{22} + \ldots + D_{2n}$$

$$\vdots$$

$$D_n = \begin{vmatrix} a_{1n}B_n \\ a_{21}B_1 \\ \vdots \\ C_n \end{vmatrix} + \begin{vmatrix} a_{1n}B_n \\ a_{22}B_2 \\ \vdots \\ C_n \end{vmatrix} + \ldots + \begin{vmatrix} a_{1n}B_n \\ a_{2n}B_n \\ \vdots \\ C_n \end{vmatrix} = D_{n1} + D_{n2} + \ldots + D_{nn}$$

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

$$D_{j_1 j_2 \ldots j_n} \quad ; \qquad \text{donde} \left\{ \begin{array}{l} j_1 = 1, 2, \ldots, n \\ j_2 = 1, 2, \ldots, n \\ \vdots \\ j_n = 1, 2, \ldots, n \end{array} \right.$$

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 $\{1, 2, \ldots, n\}$.

Entonces

$$\det(AB) = \sum_{k=1}^{n!} D_{\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn}}$$

donde $(\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn})$ es una permutación de $\{1, 2, \ldots, n\}$

Tomando en cuenta la estructura de $D_{\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn}}$ podemos escribir

$$\det(AB) = \sum_{k=1}^{n!} \begin{vmatrix} a_{1\alpha_{k1}} & B_{\alpha_{k1}} \\ a_{2\alpha_{k2}} & B_{\alpha_{k2}} \\ \vdots & \\ a_{n\alpha_{kn}} & B_{\alpha_{kn}} \end{vmatrix}$$

por lo que, aplicando reiteradamente ii) de VII.2.2, se tiene que

$$\det(AB) = \sum_{k=1}^{n!} a_{1\alpha_{k1}} a_{2\alpha_{k2}} \cdots a_{n\alpha_{kn}} \begin{vmatrix} B_{\alpha_{k1}} \\ B_{\alpha_{k2}} \\ \vdots \\ B_{\alpha_{kn}} \end{vmatrix}$$

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

$$p_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn})$$

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

$$\det(AB) = \sum_{k=1}^{n!} a_{1\alpha_{k1}} a_{2\alpha_{k2}} \cdots a_{n\alpha_{kn}} \varepsilon(p_k) \det B$$

esto es

$$\det(AB) = \det B \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}} a_{2\alpha_{k2}} \cdots a_{n\alpha_{kn}}$$

$$= (\det B)(\det A)$$

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

como se quería.

$\square$

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 sección 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:

$$\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array} \quad \text{ó} \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{array}{l} = a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+ \\ + a_{31}a_{23}a_{12}-a_{31}a_{22}a_{13}- \\ - a_{32}a_{23}a_{11}-a_{33}a_{12}a_{21} \end{array}$$

El desarrollo anterior coincide, salvo en el orden de algunos factores y términos, con la definición de determinante de tercer orden.

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

$$\det A = \begin{vmatrix} 1 & 0 & -3 \\ -4 & 5 & 2 \\ 1 & -2 & 0 \end{vmatrix} \begin{array}{l} = (1)(5)(0)+(-4)(-2)(-3)+(1)(2)(0) \\ - (1)(5)(-3)-(-2)(2)(1)-(0)(0)(-4) = \\ = 0 -24 + 0 + 15 + 4 -0 = -5 \end{array}$$

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 orden 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

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \rightarrow \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

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×n con elementos en C y r es un número entero tal que $1 \leqslant r \leqslant n$, entonces:

i) $\det A = \sum_{j=1}^{n} a_{rj} C_{rj}$

ii) $\det A = \sum_{i=1}^{n} a_{ir} C_{ir}$

DEMOSTRACION

Se demuestra primero la parte i)

De VII.1.4 se tiene que

$$\det A = \sum_{k=1}^{n!} \varepsilon(p_k)\, a_{1\alpha_{k1}}\, a_{2\alpha_{k2}} \cdots a_{r\alpha_{kr}} \cdots a_{n\alpha_{kn}} \tag{1}$$

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

$$\det A = a_{r1} F_{r1} + a_{r2} F_{r2} + \ldots + a_{rn} F_{rn} \tag{2}$$

y debemos probar que los F_{rj} (j = 1, 2,..., 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

$$F_{rj} = \sum_{k=1}^{(n-1)!} \varepsilon(p_k)\, a_{1\alpha_{k_1}}\, a_{2\alpha_{k_2}} \cdots$$

$$\cdots\, a_{(r-1)\alpha_{k(r-1)}}\, a_{(r+1)\alpha_{k(r+1)}} \cdots a_{n\alpha_{kn}} \tag{3}$$

Donde $q_k = (\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{k(r-1)}, \alpha_{k(r+1)}, \ldots, \alpha_{kn})$ es una permutación del conjunto $\{1, 2, \ldots, j-1, j+1, \ldots, n\}$ que se obtiene suprimiendo en $p_k = (\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kr}, \ldots, \alpha_{kn})$ 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

$$p_k = (\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kr}, \ldots, \alpha_{kn})$$

y traslademos a α_{kr} hasta el último lugar por medio de n-r intercambios de números adyacentes, se obtiene así la permutación

$$p'_k = (\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{kn}, \alpha_{kr})$$

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

$$q_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kn})$$

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.

$\square$

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

$$C_{32} = 17 \qquad \text{y} \qquad C_{33} = 7$$

por lo que

det A = 2×17 - 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

$$\det A = \begin{array}{c} \begin{array}{ccccc} & & & \downarrow & \end{array} \\ \left|\begin{array}{ccccc} -1 & 1 & -5 & -2 & 3 \\ 3 & 2 & 1 & 0 & -1 \\ 1 & -1 & 2 & 1 & 0 \\ 0 & 2 & 1 & 3 & -1 \\ 1 & 2 & 4 & 0 & 1 \end{array}\right| \end{array} \begin{array}{c} \\ (2) \\ \\ * \\ (-3) \\ \\ \end{array} = \left|\begin{array}{ccccc} 1 & -1 & -1 & 0 & 3 \\ 3 & 2 & 1 & 0 & -1 \\ 1 & -1 & 2 & 1 & 0 \\ -3 & 5 & -5 & 0 & -1 \\ 1 & 2 & 4 & 0 & 1 \end{array}\right|$$

En consecuencia, desarrollando por cofactores según la cuarta columna

$$\det A = (1)(-1)^7 \left|\begin{array}{cccc} 1 & -1 & -1 & 3 \\ 3 & 2 & 1 & -1 \\ -3 & 5 & -5 & -1 \\ 1 & 2 & 4 & 1 \end{array}\right|$$

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

$$\det A = (-1) \begin{array}{c} \rightarrow \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \left|\begin{array}{cccc} 1 & -1 & -1 & 3 \\ 3 & 2 & 1 & -1 \\ -3 & 5 & -5 & -1 \\ 1 & 2 & 4 & 1 \end{array}\right| \\ \begin{array}{cccc} * & (1) & (1) & (-3) \end{array} \end{array} = (-1) \left|\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 3 & 5 & 4 & -10 \\ -3 & 2 & -8 & 8 \\ 1 & 3 & 5 & -2 \end{array}\right|$$

y desarrollando por cofactores según el primer renglón

$$\det A = (-1)(1)(-1)^2 \left|\begin{array}{ccc} 5 & 4 & -10 \\ 2 & -8 & 8 \\ 3 & 5 & -2 \end{array}\right|$$

Para continuar con el proceso, seleccionamos ahora el segundo renglón para el desarrollo y la primera columna como pivote, con lo que

$$\det A = (-1) \rightarrow \begin{vmatrix} 5 & 4 & -10 \\ 2 & -8 & 8 \\ 3 & 5 & -2 \end{vmatrix} = (-1) \begin{vmatrix} 5 & 24 & -30 \\ 2 & 0 & 0 \\ 3 & 17 & -14 \end{vmatrix}$$

$$\begin{matrix} * & (4) & (-4) \end{matrix}$$

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

$$A_n = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Desarrollando por cofactores según el último renglón se tiene que

$$\det A_n = a_{nn}(-1)^{n+n} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ 0 & a_{22} & \cdots & a_{2,n-1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & a_{n-1,n-1} \end{vmatrix}$$

pero, por hipótesis de inducción

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ 0 & a_{22} & \cdots & a_{2,n-1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & a_{n-1,n-1} \end{vmatrix} = a_{11}\, a_{22} \;\cdots\; a_{n-1,n-1}$$

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

$$\text{Adj } A = \left[b_{ij} \right], \text{ donde } \quad b_{ij} = C_{ji}$$

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

$$\text{Adj } A = \left[C_{ji}\right] = \begin{bmatrix} -4 & 2 & 2 \\ 8 & -4 & 2 \\ 5 & -1 & -1 \end{bmatrix}$$

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 $A = \left[a_{ij}\right]$.

De VII.4.1

$$\text{Adj } A = \left[b_{ij}\right] \text{ , donde } b_{ij} = C_{ji}$$

por lo que, de VI.3.1

$$A(\text{Adj } A) = \left[p_{ij}\right]$$

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

$$\begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array}\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & & & \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{j1}C_{j1} + a_{j2}C_{j2} + \dots + a_{jn}C_{jn}.$$

por lo que, haciendo $a_{jk} = a_{ik}$ tenemos

$$\begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array}\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & & & \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{i1}C_{j1} + a_{i2}C_{j2} + \dots + a_{in}C_{jn}$$

Pero este determinante tiene dos renglones iguales por lo que, de iv) de VII.2.2, su valor es cero y

$$a_{i1}\,C_{j1} + a_{i2}\,C_{j2} + \dots + a_{in}\,C_{jn} = 0 .$$

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)

$$p_{ij} = \begin{cases} \det A, & \text{si} \quad i = j \\ 0 \;, & \text{si} \quad i \neq j \end{cases}$$

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

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

$\det(A A^{-1}) = \det I_n$

$(\det A)(\det A^{-1}) = \det I_n$ por iii) de VII.2.3

$(\det A)(\det A^{-1}) = 1$ por VII.3.3 (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

$$\exists A^{-1} \Leftrightarrow \det A \neq 0$$

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

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

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

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 ≠ 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

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad y \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

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

$$X = \left[\frac{1}{\det A}(\operatorname{Adj} A) \right] B$$

Esto es

$$X = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$X = \frac{1}{\det A} \begin{bmatrix} b_1C_{11} + b_2C_{21} + \dots + b_nC_{n1} \\ b_1C_{12} + b_2C_{22} + \dots + b_nC_{n2} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ b_1C_{1n} + b_2C_{2n} + \dots + b_nC_{nn} \end{bmatrix}$$

Así que

$$x_k = \frac{1}{\det A}(b_1C_{1k} + b_2C_{2k} + \dots + b_nC_{nk}), \quad (k = 1, 2, \dots, n)$$

pero, por VII.3.2, la expresión

$$b_1C_{1k} + b_2C_{2k} + \dots + b_nC_{nk}$$

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 indepen dientes; 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 ecuacio - nes 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 cuál

$$\text{Adj } A = \begin{bmatrix} 1 & -2 & 7 \\ -3 & 0 & -3 \\ 1 & -2 & 1 \end{bmatrix} \qquad \text{y} \qquad A^T = \begin{bmatrix} 1 & a_{21} & a_{31} \\ 2 & a_{22} & a_{32} \\ -1 & 3 & 1 \end{bmatrix}$$

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)

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

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×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×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

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

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 símbo lo, 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

□	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 □ 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

$$\forall\ a,\ b \in S:\quad a*b \in S$$

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
>
> $$\forall\, a, b \in T: \quad a*b \in T$$

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 $m \times n$ 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 $n \in Z$ se tiene que $-n \in Z$.

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

$$n \; \Delta \; 0 = n, \; \forall \, n \in Z$$

y además $0 \in Z$.

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 $\hat{a} \in S$ es un inverso izquierdo del elemento $a \in S$ si

$$\hat{a} * a = e$$

ii) Sea e un idéntico derecho para $*$. Un elemento $\hat{a} \in S$ es un inverso derecho del elemento $a \in S$ si

$$a * \hat{a} = e$$

iii) Sea e un idéntico para $*$. Un elemento $\hat{a} \in S$ es un inverso del elemento $a \in S$ 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

$$-x + x = 0 \quad \text{y} \quad x + (-x) = 0$$

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

$$\frac{1}{x} \cdot x = 1 \quad \text{y} \quad x \cdot \frac{1}{x} = 1$$

- Asociatividad

VIII.1.5 DEFINICION

Sea $*$ una operación binaria definida en un conjunto S

Se dice que $*$ es asociativa si

$$\forall\, a, b, c \in S: \; a*(b*c) = (a*b)*c$$

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 podemos 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 * 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 ε Q, como lo establecen los teoremas 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

$$a \,\square\, b = a + 3b + 1; \quad \forall\ a, b \in Z.$$

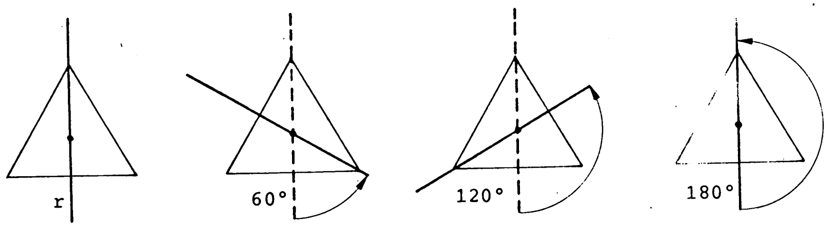
Determinar si dicha operación:

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 inter - pretarse como "existe al menos un e". Para indicar que existe exactamente un elemento e se dice "existe uno y só lo 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 di versos sistemas numéricos que ya hemos manejado.

El conjunto de los números enteros y la operación de adición cons tituyen un sistema con estructura de grupo ya que, como sabemos

i) $\forall\, a, b, c \in Z$: $a + (b + c) = (a + b) + c$

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 $m \times n$ 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

$1 \cdot 1 = 1$ por lo que $\hat{1} = 1$

$-1 \cdot (-1) = 1$ por lo que $-\hat{1} = -1$

$-i \cdot i = 1$ por lo que $\hat{i} = -i$

$i \cdot (-i) = 1$ por lo que $-\hat{i} = i$

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

$$x \S y = x + y - 2xy$$

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. $\square$

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 $\exists\ \hat{a} \in G$ 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

$$a*x = b \quad y \quad a*x' = b$$

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) $\widehat{(\hat{a})} = a; \quad \forall\, a \in G$

ii) $\widehat{(a*b)} = \hat{b}*\hat{a}; \quad \forall\, a, b \in G$

iii) $\widehat{(a_1 * a_2 * \ldots * a_n)} = \hat{a}_n * \ldots * \hat{a}_2 * \hat{a}_1; \; \forall\, a_1, a_2, \ldots, a_n \in G$

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) $\forall\, a \in S:\ \hat{a} \in S$

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 $m, n \in Z$; 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 $m, n \in Z$ y $m, n > 0$; entonces

$$a + b = m + n \in T,$$

ya que $m + n \in Z$ y $m + n > 0$.

ii) $a = -m \not\in T$, 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) $\forall\ a,\ b,\ c \in G;\quad a*b = c*a \implies b = c$

ii) $\forall\ a,\ b,\ c \in G;\quad b*a = a*c \implies b = c$

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

$$a \mathrel{P} b = \frac{a + b}{2}, \quad \forall\ a,\ b \in C$$

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) $b*a = c*a \implies b = c$

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) $a*b = c*a \implies b = c$

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 · 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 · 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 • dos operaciones binarias definidas en A. El sistema (A, +, •) 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 $(A, +, \cdot)$ un anillo:

i) Si $\forall\ a, b \in A \quad a \cdot b = b \cdot a$

se dice que el anillo es conmutativo.

ii) Si $\exists\ 1 \in A$ tal que $1 \cdot a = a = a \cdot 1, \quad \forall\ a \in A$

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 $\cdot$ dos operaciones binarias definidas en K. El sistema (K, +, $\cdot$) 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) $\exists$ $\mathit{0}$ ε K tal que $\mathit{0} + a = a$, $\forall$ a ε K

iv) $\forall$ a ε K $\exists$ -a ε K tal que $-a + a = \mathit{0}$

v) $\forall$ a, b, c ε K $a\cdot(b\cdot c) = (a\cdot b)\cdot c$

vi) $\forall$ a, b ε K $a\cdot b = b\cdot a$

vii) $\exists$ $\mathit{1}$ ε K tal que $\mathit{1}\cdot a = a$, $\forall$ a ε K

viii) $\forall$ a ε K, $a \neq \mathit{0}$, $\exists$ a^{-1} tal que $a^{-1}\cdot a = \mathit{1}$

ix) $\forall$ a, b, c ε K $a\cdot(b + c) = (a\cdot b) + (a\cdot c)$
y
$(b + c)\cdot a = (b\cdot a) + (c\cdot a)$

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 $\cdot$ dos operaciones binarias definidas en K. El sistema (K, +, $\cdot$) es un campo si:

i) (K,+) es un grupo abeliano, cuyo elemento idéntico denotamos con $\mathit{0}$.

ii) (K-{$\mathit{0}$},$\cdot$) es un grupo abeliano.

iii) $\cdot$ 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 que 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 $(K, +, \cdot)$ un campo. Entonces, por los postulados i) a vii) y ix) de la definición VIII.3.4, $(K, +, \cdot)$ 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 $(K, +, \cdot)$. Probaremos primero el siguiente lema:

$$0 \cdot a = 0, \quad \forall\, a \in K$$

Demostración:

$0 \cdot a = (0 + 0) \cdot a$ por iii) de VIII.3.4

$0 \cdot a = (0 \cdot a) + (0 \cdot a)$ 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

$$a = 0$$

En consecuencia

$$a \cdot b = 0 \implies a = 0 \text{ o } b = 0$$

Esto completa la demostración.

$\square$

Como ejemplos de campos, además de los sistemas $(Q, +, \cdot)$, $(R, +, \cdot)$ y $(C, +, \cdot)$ 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 $(R, +, \cdot)$.

Otro ejemplo de campo lo encontramos en el conjunto $\{p, q, r, s\}$ con las operaciones $+$ y $\cdot$ 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 $\oplus$ y $\odot$ 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

$$x \odot y \in Q, \forall x, y \in Q$$

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

$$x^{\urcorner} = \frac{-x}{1 + x} \in Q$$

Por otra parte, si $x = -1$ y suponemos que $\exists\ x^{\urcorner}$ se tendrá que

$$x^{\urcorner} \cdot 0 = 1$$

lo cual es un absurdo y, por lo tanto, $\nexists\ -1^{\urcorner}$.

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

$$x \odot (y \oplus z) = x \odot (y + z + 1)$$

$$= x + y + z + 1 + xy + xz + x$$

$$x \odot (y \oplus z) = 2x + y + z + xy + xz + 1$$

$$(x \odot y) \oplus (x \odot z) = (x + y + xy) \oplus (x + z + xz)$$

$$= x + y + xy + x + z + xz + 1$$

$$(x \odot y) \oplus (x \odot z) = 2x + y + z + xy + xz + 1$$

por lo que

$$x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$$

Por otra parte

$$(y \oplus z) \odot x = (y + z + 1) \odot x$$

$$= y + z + 1 + x + yx + zx + x$$

$$(y \oplus z) \odot x = 2x + y + z + yx + zx + 1$$

$$(y \odot x) \oplus (z \odot x) = (y + x + yx) \oplus (z + x + zx)$$

$$= y + x + yx + z + x + zx + 1$$

$$(y \odot x) \oplus (z \odot x) = 2x + y + z + yx + zx + 1$$

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:

$a - b \in S$ y $a \cdot b \in S$, $\forall$ 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.

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 $f: A \to B$. 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

$$\begin{tikzcd}
(a,b) \arrow[r, "*"] \arrow[d, Rightarrow, "f"'] & a*b \arrow[d, "f"] \\
(f(a),f(b)) \arrow[r, Rightarrow, "\Delta"] & \begin{matrix} f(a*b) \\ f(a)\ \Delta\ f(b) \end{matrix}
\end{tikzcd}$$

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.

$$\begin{tikzcd}[column sep=huge, row sep=huge]
\text{elementos físicos} \arrow[r, dashed, "\textit{operación}", "\textit{física}"'] \arrow[d, "\textit{modelo}\ (f)"'] & \text{resultado físico} \\
\text{elementos matemáticos} \arrow[r, "\textit{operación}", "\textit{matemática}"'] & \text{resultado matemático} \arrow[u, "\textit{interpretación}\ \textit{del modelo}\ (f^{-1})"']
\end{tikzcd}$$

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 $f\colon G \to G'$ 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 $f: S \to R^3$ 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 $f: G \to G'$ 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. $\square$

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

$$f^{-1}(a' \, \Delta \, b') = a*b$$

Finalmente, como $a = f^{-1}(a')$ y $b = f^{-1}(b')$

$$f^{-1}(a' \, \Delta \, b') = f^{-1}(a') * f^{-1}(b')$$

por lo que f^{-1} es un homomorfismo y la prueba termina. $\square$

Para concluir esta sección cabe mencionar que los conceptos de isomorfismo y homomorfismo entre estructuras algebraicas no se limitan al caso de la estructura de grupo.

Por ejemplo, en el caso de dos anillos $(A, +, \cdot)$ y $(A', \oplus, \odot)$, una función $f: A \rightarrow A'$ tal que

$$f(a + b) = f(a) \oplus f(b)$$

y

$$f(a \cdot b) = f(a) \odot f(b); \quad \forall \, a, b \in A.$$

es un homomorfismo de A en A'. Si f es biyectiva se trata de un isomorfismo y se dice que los anillos son isomorfos.

De esta manera, se habla de isomorfismos entre anillos, entre campos y entre espacios vectoriales.

VIII.4.4 EJERCICIOS

1.- Para los grupos $(R,*)$ y (R^+,Δ), donde

$$x * y = x + y - 1; \quad \forall\, x, y \in R$$
$$x \Delta y = \frac{1}{e} xy; \quad \forall\, x, y \in R^+$$

a) Demostrar que la función $f: R \rightarrow R^+$ definida por

$$f(x) = e^x \qquad \text{es un isomorfismo}$$

b) Verificar que, para dicha función, se cumplen las conclusiones que establece el teorema VIII.4.2

2.- Un isomorfismo de un grupo en sí mismo se conoce, frecuentemente, con el nombre de "automorfismo".
Determinar si la función "valor absoluto" es un automorfismo en cada uno de los siguientes grupos: $(R^+, +)$ y $(R^+,\cdot)$.

3.- Sea f un isomorfismo entre los grupos $(G,*)$ y (G',Δ) y sean $\hat{a}$ el inverso de $a \in G$ para $*$ y $\widehat{f(a)}$ el inverso de $f(a)$ para Δ.
Demostrar que

$$f(\hat{a}) = \widehat{f(a)}$$

4.- Considérense los anillos formados por los conjuntos

$$A = \left\{ \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \;\middle|\; x, y \in R \right\} \quad y \quad A' = \{a+bi \mid a,b \in R,\ i^2 = -1\}$$

y las operaciones de adición y multiplicación definidas en forma usual para las matrices y para los números complejos, respectivamente. Determinar si las funciones H y S son isomorfismos entre $(A, +, \cdot)$ y $(A', +, \cdot)$

a) $$\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \xrightarrow{H} x + y\, i$$

b) $\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \xrightarrow{S} x + xy\, i$

5.- Sea f un isomorfismo entre los anilllos (A, +, ·) y (B, +, ·). Si 1_A es el idéntico en A para · y 1_B es el idéntico en B para ·, demostrar que

$$f(1_A) = 1_B$$

CAPITULO IX   ESPACIOS VECTORIALES

INTRODUCCION

Comúnmente, la primera noción que se tiene de vector es la de una "cantidad con magnitud, dirección y sentido". Esta noción surge a partir de conceptos físicos, tales como fuerza y velocidad, cuya cabal descripción requiere algo más que un simple número.

Dichos conceptos han encontrado una adecuada representación geométrica en los llamados "segmentos dirigidos" (flechas), los cuales pueden ser caracterizados y manejados analíticamente mediante parejas o ternas ordenadas de números.

Seguramente el lector ha tenido ya una amplia experiencia con este tipo de vectores. En particular, los ha sumado y los ha multiplicado por escalares (números).

Con la adición y la multiplicación por un escalar pueden formarse expresiones como

$$\alpha\bar{a} + \beta\bar{b} + \gamma\bar{c}$$

donde $\bar{a}$, $\bar{b}$, $\bar{c}$ son vectores y α, β, γ son escalares; expresiones a las que se denomina "combinaciones lineales".

En muchas ramas de la matemática se presentan conjuntos, con elementos de muy diversa naturaleza, donde se emplea con frecuencia el concepto de combinación lineal. Quizá el ejemplo más conocido sea el de las ternas ordenadas de números reales; sin embargo, se tienen también los polinomios, las matrices y las funciones como otros ejemplos de conjuntos para los cuales tiene sentido el concepto de combinación lineal.

Tales conjuntos, con las leyes de adición y multiplicación por un escalar definidas de la manera usual, tienen en común gran número de propiedades algebraicas; tienen, por tanto, una estructura común. A dicha estructura se le conoce con el nombre de "Espacio Vectorial".†

En forma genérica, a los elementos de un espacio vectorial se les llama "vectores", por lo que, en este contexto, la palabra vector adquiere un significado más amplio que el de "cantidad con magnitud, dirección y sentido".

Este capítulo está destinado al estudio de la estructura de Espacio Vectorial, incluyendo algunas aplicaciones a la teoría de sistemas de ecuaciones lineales y al comportamiento algebraico de las funciones.

Con el propósito de ilustrar las definiciones y teoremas fundamentales se presentan diversos ejemplos relativos a espacios vectoriales conocidos, especialmente al de las ternas ordenadas de nú-

† *Algunos autores emplean el término "Espacio Lineal" para referirse a este mismo concepto.*

meros reales; sin embargo, no debe olvidarse que con tales ejem - plos se pretende ayudar a comprender los conceptos básicos inhe - rentes a la estructura y no los detalles propios de los casos par ticulares (espacios vectoriales específicos), los cuales deben ser ya del dominio del lector.

IX.1 LA ESTRUCTURA DE ESPACIO VECTORIAL

- El espacio vectorial de las ternas.

En los cursos de geometría analítica y de álgebra vectorial se trabaja frecuentemente con el conjunto

$$R^3 = \{(x, y, z) \mid x, y, z \in R\}$$

constituido por todas las ternas ordenadas de números reales. En dicho conjunto, la adición y la multiplicación por un escalar se definen, usualmente, de la siguiente manera:

Si $\bar{a} = (a_1, a_2, a_3)$ y $\bar{b} = (b_1, b_2, b_3)$ son dos elementos de R^3 y λ es un número real, entonces

$$\bar{a} + \bar{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3) \qquad \text{(I)}$$

y

$$\lambda\bar{a} = (\lambda a_1, \lambda a_2, \lambda a_3) \qquad \text{(II)}$$

Así, por ejemplo, para los vectores $\bar{a} = (2, -1, 4)$ y $\bar{b} = (3, 0, -1)$ y el escalar $\lambda = -\frac{1}{2}$, se tiene que

$$\bar{a} + \bar{b} = (5, -1, 3) \quad \text{y} \quad \lambda\bar{a} = (-1, \tfrac{1}{2}, -2)$$

Como consecuencia de tales definiciones se tiene que, para todas las ternas $\bar{a}$, $\bar{b}$ y $\bar{c}$, y todos los números reales α y β:

1) $\overline{a} + \overline{b}$ es único y $\overline{a} + \overline{b} \in R^3$

2) $\overline{a} + (\overline{b} + \overline{c}) = (\overline{a} + \overline{b}) + \overline{c}$

3) $\overline{0} + \overline{a} = \overline{a}$, donde $\overline{0} = (0, 0, 0)$

4) $-\overline{a} + \overline{a} = \overline{0}$, donde $-\overline{a} = (-a_1, -a_2, -a_3)$

5) $\overline{a} + \overline{b} = \overline{b} + \overline{a}$

6) $\alpha\overline{a}$ es único y $\alpha\overline{a} \in R^3$

7) $\alpha(\overline{a} + \overline{b}) = \alpha\overline{a} + \alpha\overline{b}$

8) $(\alpha + \beta)\overline{a} = \alpha\overline{a} + \beta\overline{a}$

9) $\alpha(\beta\overline{a}) = (\alpha\beta)\overline{a}$

10) $1 \cdot \overline{a} = \overline{a}$

Estas diez propiedades, que se demuestran fácilmente a partir de las expresiones (I) y (II), no son las únicas que pueden establecerse como consecuencia de las definiciones usuales de adición y multiplicación por un escalar en R^3; sin embargo, son suficientes para deducir, a partir de ellas, todas las demás reglas necesarias para el manejo algebraico de ternas y números.

Existen muchos otros sistemas que satisfacen también dichas propiedades. Todos ellos son ejemplos de espacios vectoriales, de acuerdo con la siguiente definición.

- Definición de espacio vectorial

IX.1.1 DEFINICION

Sea V un conjunto no vacío y sea (K, +, •) un campo. Se dice que V es un espacio vectorial sobre K si están definidas dos leyes de composición, llamadas adición y multiplicación por un escalar, tales que

i) La adición asigna a cada pareja ordenada $(\bar{u}, \bar{v})$ de elementos de V un único elemento $\bar{u} + \bar{v} \in V$, llamado la suma de $\bar{u}$ y $\bar{v}$.

ii) $\forall\ \bar{u}, \bar{v}, \bar{w} \in V: \quad \bar{u} + (\bar{v} + \bar{w}) = (\bar{u} + \bar{v}) + \bar{w}$

iii) $\exists\ \bar{0} \in V$ tal que $\bar{0} + \bar{v} = \bar{v}, \forall\ \bar{v} \in V$

iv) $\forall\ \bar{v} \in V\ \exists\ -\bar{v} \in V$ tal que $-\bar{v} + \bar{v} = \bar{0}$

v) $\forall\ \bar{u}, \bar{v} \in V: \quad \bar{u} + \bar{v} = \bar{v} + \bar{u}$

vi) La multiplicación por un escalar asigna a cada pareja ordenada $(\alpha, \bar{v})$ de elementos $\alpha \in K$ y $\bar{v} \in V$ un único elemento $\alpha\bar{v} \in V$, llamado el producto de α por $\bar{v}$.

vii) $\forall\ \alpha \in K;\ \bar{u}, \bar{v} \in V: \quad \alpha(\bar{u} + \bar{v}) = \alpha\bar{u} + \alpha\bar{v}$

viii) $\forall\ \alpha, \beta \in K;\ \bar{v} \in V: \quad (\alpha + \beta)\bar{v} = \alpha\bar{v} + \beta\bar{v}$

ix) $\forall\ \alpha, \beta \in K;\ \bar{v} \in V: \quad \alpha(\beta\bar{v}) = (\alpha\beta)\bar{v}$

x) Si 1 es la unidad de K: $1\bar{v} = \bar{v}, \forall\ \bar{v} \in V$

A los elementos de V se les llama vectores y a los de K escalares.

Cabe hacer notar que la adición es una función de $V \times V$ en V y, por lo tanto, es una operación binaria. A diferencia de ésta, la multiplicación por un escalar no lo es ya que es una función de $K \times V$ en V; sin embargo, es frecuente referirse a las dos como "operaciones" en un sentido menos estricto que el de la definición VIII.1.1.

Si precisamos los elementos que integran los conjuntos V y K, e indicamos cómo se suman los elementos de V y cómo se multiplican por escalares de K, podemos obtener diversos ejemplos de espacios vectoriales. El lector puede verificar fácilmente que cada uno de los siguientes es un ejemplo de espacio vectorial.

Ejemplo 1. V es el conjunto de todas las matrices de $m \times n$ con elementos en C, K es el campo de los números complejos, y la adición y la multiplicación por un escalar son las usuales (Definiciones VI.2.1 y VI.2.4).

Ejemplo 2. V es el conjunto de todos los polinomios con coeficientes en C, K es el campo de los números complejos, y la adición y la multiplicación por un escalar son las usuales (Definición III.1.4 y expresión previa a los ejercicios III.1.10).

Ejemplo 3. $V = R^n$, $K = R$, y la adición y la multiplicación por un escalar son las usuales; esto es:

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

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1,\ \lambda x_2, \ldots,\ \lambda x_n)$$

Ejemplo 4. V es el conjunto de todas las funciones reales de variable real, K es el campo de los números reales, y la adición y la multiplicación por un escalar son las usuales; esto es:

$$(f + g)(x) = f(x) + g(x)$$

$$(\lambda f)(x) = \lambda f(x)$$

Ejemplo 5. V es el conjunto de todas las soluciones de una ecuación diferencial de la forma

$$y'' + ay' + by = 0$$

donde a y b son constantes, K es el campo de los números reales, y la adición y multiplicación por un escalar son las usuales.

- Propiedades algebraicas fundamentales.

De los diez postulados que integran la definición IX.1.1 los primeros cinco se refieren únicamente a la adición, y establecen que el sistema (V, +) es un grupo abeliano; por lo tanto, con base en los resultados del capítulo VIII podemos enunciar las siguientes propiedades, las cuales son comunes a todos los espacios vectoriales.

IX.1.2 TEOREMA

Si V es un espacio vectorial sobre K, entonces

i) $\forall\ \bar{u}, \bar{v}, \bar{w} \in V$: $\bar{u} + \bar{v} = \bar{u} + \bar{w} \Rightarrow \bar{v} = \bar{w}$

ii) El vector $\bar{0}$ es único y es tal que
$$\bar{v} + \bar{0} = \bar{v},\ \forall\ \bar{v} \in V$$

iii) El vector $-\bar{v}$ es único y es tal que
$$\bar{v} + (-\bar{v}) = \bar{0}$$

iv) La ecuación $\bar{u} + \bar{x} = \bar{v}$ tiene solución única en V.

v) $\forall\ \bar{v} \in V$: $-(-\bar{v}) = \bar{v}$

vi) $\forall\ \bar{u}, \bar{v} \in V$: $-(\bar{u} + \bar{v}) = -\bar{u} + (-\bar{v})$

Tomando en cuenta, además, los postulados vi) a x) de la defini -

ción, los cuales involucran a la multiplicación por un escalar, podemos establecer otras importantes propiedades que, junto con las anteriores, rigen los procedimientos algebraicos en un espacio vectorial.

IX.1.3 TEOREMA

Sea V un espacio vectorial sobre K:

i) $\forall\ \alpha \in K$: $\qquad \alpha\bar{0} = \bar{0}$

ii) $\forall\ \bar{v} \in V$: $\qquad 0\bar{v} = \bar{0}$, donde 0 es el cero de K.

iii) $\forall\ \alpha \in K,\ \bar{v} \in V$: $\qquad (-\alpha)\bar{v} = -(\alpha\bar{v}) = \alpha(-\bar{v})$

iv) $\forall\ \alpha \in K,\ \bar{v} \in V$: $\qquad \alpha\bar{v} = \bar{0} \Longrightarrow \alpha = 0$ *o* $\bar{v} = \bar{0}$

v) $\forall\ \alpha \in K,\ \bar{u},\ \bar{v} \in V$: $\qquad \alpha\bar{u} = \alpha\bar{v}$ *y* $\alpha \neq 0 \Longrightarrow \bar{u} = \bar{v}$

vi) $\forall\ \alpha,\ \beta \in K,\ \bar{v} \in V$: $\qquad \alpha\bar{v} = \beta\bar{v}$ *y* $\bar{v} \neq 0 \Longrightarrow \alpha = \beta$

DEMOSTRACION

Se demuestran a continuación únicamente i), iii) y iv)

i) $\alpha\bar{0} = \alpha(\bar{0} + \bar{0})$ por iii) de IX.1.1

$\alpha\bar{0} = \alpha\bar{0} + \alpha\bar{0}$ por vii) de IX.1.1

$\alpha\bar{0} + \bar{0} = \alpha\bar{0} + \alpha\bar{0}$ por ii) de IX.1.2

$\bar{0} = \alpha\bar{0}$ por i) de IX.1.2

Con lo que se ha demostrado la propiedad i). Se sugiere al lector como ejercicio demostrar este mismo enunciado sin hacer uso del teorema IX.1.2; es decir, demostrarlo directamente a partir de la definición IX.1.1.

iii) $(-\alpha)\bar{v} + \alpha\bar{v} = (-\alpha + \alpha)\bar{v}$ por viii) de IX.1.1

$(-\alpha)\bar{v} + \alpha\bar{v} = 0\bar{v}$ por iv) de VIII.3.4

$(-\alpha)\bar{v} + \alpha\bar{v} = \bar{0}$ por ii) de IX.1.3

En consecuencia, $(-\alpha)\overline{v}$ es el inverso aditivo de $\alpha\overline{v}$; esto es

$$(-\alpha)\overline{v} = -(\alpha\overline{v})$$

La prueba de la segunda igualdad se deja como ejercicio.

iv) sea $\alpha\overline{v} = \overline{0}$

y supongamos que $\alpha \neq 0$ y $\overline{v} \neq \overline{0}$; entonces, por viii) de VIII.3.4 $\exists\ \alpha^{-1} \in K$ tal que $\alpha^{-1}\alpha = 1$, y en consecuencia

$$\alpha^{-1}(\alpha\overline{v}) = \alpha^{-1}\,\overline{0}$$

por lo que:

$$\begin{aligned} \alpha^{-1}(\alpha\overline{v}) &= \overline{0} && \text{por i) de IX.1.3} \\ (\alpha^{-1}\alpha)\overline{v} &= \overline{0} && \text{por ix) de IX.1.1} \\ 1\,\overline{v} &= \overline{0} && \text{por viii) de VIII.3.4} \\ \overline{v} &= \overline{0} && \text{por x) de IX.1.1} \end{aligned}$$

lo cual contradice la hipótesis; por lo tanto

$$\alpha = 0 \quad o \quad \overline{v} = \overline{0}$$

como se quería. $\square$

En un espacio vectorial, la sustracción se define a partir de la adición de la siguiente manera.

IX.1.4 DEFINICION

Si V es un espacio vectorial sobre K, entonces

$$\overline{u} - \overline{v} = \overline{u} + (-\overline{v}) \ ; \quad \forall\ \overline{u}, \overline{v} \in V$$

Al vector $\overline{u} - \overline{v}$ se le llama la diferencia $\overline{u}$ menos $\overline{v}$.

- Subespacios

Es posible que un espacio vectorial tenga subconjuntos que sean, por si mismos, espacios vectoriales.

Consideremos, por ejemplo, el conjunto de puntos sobre una recta en el espacio, que pasa por el origen y que tiene como números directores 1, 2, 3. Dicho conjunto puede considerarse como un subconjunto de R^3 ya que la ecuación de la recta, en forma simétrica, es

$$\frac{x}{1} = \frac{y}{2} = \frac{z}{3}$$

y, por lo tanto, las coordenadas y, z de los puntos guardan la siguiente relación con la coordenada x

$$y = 2x, \quad z = 3x$$

Así, el conjunto de todos los puntos sobre la recta puede ser expresado como

$$L = \{(x,\ 2x,\ 3x) \mid x \in R\}$$

el cual es un subconjunto de R^3.

El lector puede verificar fácilmente que el conjunto L, con la adición y la multiplicación por un escalar usuales en R^3, satisface los postulados de la definición IX.1.1 y, por lo tanto, es un espacio vectorial sobre R. Se dice por ello que es un subespacio de R^3.

En general, se tiene la siguiente definición

IX.1.5 DEFINICION

Sea V un espacio vectorial sobre K y sea S un subconjunto de V. S es un subespacio de V si es un espacio vectorial sobre K respecto a la adición y la multiplicación por un escalar definidas en V.

De acuerdo con esta definición, para concluir que un conjunto dado S es un subespacio de V debemos verificar que se cumplen los diez postulados de la definición IX.1.1 para los elementos de S; por ello, cabe preguntarse si podemos reducir el número de condiciones a verificar tomando en cuenta que S está contenido en V. La respuesta es afirmativa, como se desprende del siguiente enunciado.

IX.1.6 TEOREMA

Sea V un espacio vectorial sobre K y sea S un subconjunto de V. S es un subespacio de V si y sólo si

i) $\forall\ \bar{u}, \bar{v} \in S: \quad \bar{u} + \bar{v} \in S$

ii) $\forall\ \alpha \in K, \bar{v} \in S: \quad \alpha\bar{v} \in S$

Es decir que bastará con verificar la "cerradura" de S respecto a la adición y a la multiplicación por un escalar definidas en V para concluir que S es un subespacio de V.

DEMOSTRACION

Primeramente, supongamos que se verifican las condiciones i) y

ii) de IX.1.6, esto implica que se cumplen los postulados i) y vi) de la definición IX.1.1 para los elementos de S.

Además, como todos los elementos de S pertenecen a V, es inmediato que se satisfacen también los postulados ii), v), vii), viii), ix) y x) de la definición, para los elementos de S.

Para demostrar que se satisfacen también los postulados iii) y iv) consideremos un vector $\bar{v} \in S$, y sean 0 y -1 el cero de K y el inverso aditivo de la unidad de K, respectivamente; entonces

$0\bar{v} \in S$	por ii) de IX.1.6
$\bar{0} \in S$	por ii) de IX.1.3

Además:

$(-1)\bar{v} \in S$	por ii) de IX.1.6
$-(1\bar{v}) \in S$	por iii) de IX.1.3
$-\bar{v} \in S$	por x) de IX.1.1

En consecuencia, S es un espacio vectorial sobre K y, por IX.1.5, es un subespacio de V.

Supongamos ahora que no se cumple alguna de las condiciones de IX.1.6; entonces, por i) o por vi) de IX.1.1 S no es un espacio vectorial sobre K y, por IX.1.5, S no es un subespacio de V. □

Con ayuda de este teorema podemos ahora demostrar que el conjunto

$$L = \{(x, 2x, 3x) \mid x \in R\}$$

es un subespacio de R^3, verificando únicamente dos condiciones.

En efecto, sean $\bar{u} = (x, 2x, 3x)$ y $\bar{v} = (y, 2y, 3y)$ dos elementos cualesquiera de L y sea $\alpha \in R$, entonces:

$$\bar{u} + \bar{v} = (x+y,\ 2x+2y,\ 3x+3y) = (x+y,\ 2[x+y],\ 3[x+y]) \in L$$

$$\alpha\bar{v} = (\alpha y,\ \alpha[2y],\ \alpha[3y]) = (\alpha y,\ 2[\alpha y],\ 3[\alpha y]) \in L$$

por lo que, de IX.1.6, L es un subespacio de R^3.

De la misma manera puede verificarse que los siguientes conjuntos son también subespacios de R^3.

$$A = \{(x, y, 0) \mid x, y \in R\}$$

$$B = \{(x, y, z) \mid x + 3y - z = 0;\ x, y, z \in R\}$$

$$C = \{(0, 0, 0)\}$$

etc.

A diferencia de los anteriores, el conjunto

$$D = \{(x, y, z) \mid x - z = -1;\ x, y, z \in R\}$$

que también es un subconjunto de R^3, no es un subespacio.

Para probarlo podemos proceder de la siguiente forma: como la condición $x - z = -1$ es equivalente a $z = x + 1$, es posible expresar al conjunto D como

$$D = \{(x, y, x + 1) \mid x, y \in R\}$$

Así, si $\bar{u} = (x, y, x + 1)$ y $\bar{v} = (z, t, z + 1)$ son dos vectores cualesquiera de D, se tiene que

$$\bar{u} + \bar{v} = (x + z,\ y + t,\ [x + z] + 2) \notin D$$

por lo que, de IX.1.6, D no es un subespacio de R^3.

Como ejemplos adicionales mostraremos que:

a) El conjunto de matrices triangulares superiores, con elementos en C, es un espacio vectorial sobre C.

b) El conjunto de polinomios de tercer grado, con coeficientes en C, no es un espacio vectorial sobre C.

a) Sea $T = \{ [t_{ij}] \mid t_{ij} \in C,\ t_{ij} = 0 \text{ para } i > j\}$

y sean $A = [a_{ij}]$, $B = [b_{ij}]$ dos matrices cualesquiera de T; esto es:

$$a_{ij} = 0, \text{ para } i > j$$
y
$$b_{ij} = 0, \text{ para } i > j$$

como

$$A + B = [s_{ij}], \text{ donde } s_{ij} = a_{ij} + b_{ij}$$

se tiene que, para $i > j$

$$s_{ij} = 0 + 0 = 0$$

por lo que $A + B \in T$

Por otra parte, si $\alpha \in C$ se tiene que

$$\alpha A = [e_{ij}], \text{ donde } e_{ij} = \alpha a_{ij}$$

entonces, para $i > j$

$$e_{ij} = \alpha a_{ij} = \alpha \cdot 0 = 0$$

por lo que $\alpha A \in T$.

En consecuencia, por IX.1.6, T es un subespacio del conjunto de matrices de m×n con elementos en C (espacio vectorial del ejem plo 1) y es, por tanto, un espacio vectorial sobre C.

b) Sea $S = \{ax^3 + bx^2 + cx + d \mid a, b, c, d \in C \text{ y } a \neq 0\}$

y sean $p_1(x) = ax^3 + bx^2 + cx + d$ y $p_2(x) = ex^3 + fx^2 + gx + h$

dos polinomios cualesquiera de S.

La suma $p_1(x) + p_2(x) = (a+e)x^3 + (b+f)x^2 + (c+g)x + d + h$, no siempre es un elemento de S: por ejemplo - - - - - - - - - si $a = -e$ se tiene que $p_1(x) + p_2(x) \notin S$. En consecuencia, por IX.1.6, S no es un subespacio del conjunto de todos los polinomios con coeficientes en C (espacio vectorial del ejemplo 2) y, por tanto, no es un espacio vectorial.

En general, si V es un espacio vectorial sobre K, los subespacios constituidos por el mismo V y por el conjunto $\{\overline{0}\}$ (el conjunto cuyo único elemento es el vector cero de V y al cual se le conoce como "espacio cero") se denominan subespacios impropios. Cualquier otro subespacio de V, si existe, se considera propio.

IX.1.7 EJERCICIOS

1.- Sea $V = \{(a,b) \mid a, b \in R\}$

Mostrar, en cada caso, que V no es un espacio vectorial sobre R para la adición y la multiplicación por un escalar definidas por

a) $(a,b) + (c,d) = (a + d, b + c)$ y $k(a,b) = (ka,kb)$

b) $(a,b) + (c,d) = (a + c, b + d)$ y $k(a,b) = (a,b)$

c) $(a,b) + (c,d) = (ac,bd)$ y $k(a,b) = (ka,kb)$

Indicar, en cada caso, cuáles axiomas de la definición IX.1.1 no se cumplen.

2.- Sean $C^n = \{(z_1, z_2, \ldots, z_n) \mid z_i \in C,\ i = 1, 2, \ldots, n\}$ donde C es el campo de los números complejos. La adición y la multiplicación por un escalar se definen usualmente en C^n de manera semejante al caso de R^n; esto es

$$(z_1, z_2, \ldots, z_n) + (w_1, w_2, \ldots, w_n) =$$
$$= (z_1 + w_1, z_2 + w_2, \ldots, z_n + w_n)$$

$$\alpha(z_1, z_2, \ldots, z_n) = (\alpha z_1, \alpha z_2, \ldots, \alpha z_n)$$

Demostrar que con las operaciones usuales:

a) C^n es un espacio vectorial sobre R.

b) R^n no es un espacio vectorial sobre C.

3.- Demostrar a partir de la definición IX.1.1 que en un espacio vectorial V sobre K:

$$\forall\ \alpha \in K\ ,\ \alpha\overline{0} = \overline{0}$$

4.- Demostrar que en un espacio vectorial V sobre K, $\forall\ \overline{v} \in V$, $\alpha \in K$

a) $0\overline{v} = \overline{0}$, donde 0 es el cero de K

b) $-(\alpha\overline{v}) = \alpha(-\overline{v})$

5.- Verificar que

a) $A = \{f \mid f(5) = 0\}$ y $B = \{g \mid g(1) = g(3)\}$ son subespacios del espacio de funciones reales de varia - ble real sobre R.

b) $A = \{(a, b, c) \mid a, b, c \in R,\ b \leqslant 0\}$ no es subespacio de R^3.

c) $A = \left\{ \begin{bmatrix} a & b \\ \frac{a}{2} & -b \end{bmatrix} \middle|\ a, b \in R \right\}$ es subespacio del espacio de matrices de orden m×n con elementos en R.

6.- Determinar cuáles de los siguientes subconjuntos del conjunto de polinomios con coeficientes en R, son subespacios:

a) $A = \{a_0+a_1x^2+a_2x^4+\ldots+a_nx^{2n} | a_i \varepsilon R,\ i = 0,1,2,\ldots,n\}$

b) $B = \{ax^2 + x + 2a \mid a \varepsilon R\}$

c) $C = \{ax^2 + bx + 2a + b \mid a, b \varepsilon R\}$

7.- Sean S y T dos subespacios de un espacio vectorial V sobre un campo K, y considérense los conjuntos

$$S \cup T = \{\bar{x} \mid \bar{x} \varepsilon S \text{ o } \bar{x} \varepsilon T\}$$

$$S + T = \{\bar{x} + \bar{y} \mid \bar{x} \varepsilon S \text{ y } \bar{y} \varepsilon T\}$$

Demostrar que S + T es un espacio vectorial sobre K; mientras que, en general, S U T no lo es. ¿En qué casos S U T es un espacio vectorial sobre K?

IX.2 DEPENDENCIA LINEAL, BASE Y DIMENSION.

- Combinación lineal

Para los conceptos que se presentan en esta sección es de fundamental importancia la idea de combinación lineal, la cual se introduce a continuación partiendo de un ejemplo

Consideremos los vectores de R^3

$$\bar{a} = (3, 0, -2) \quad y \quad \bar{b} = (4, 1, -1)$$

y obtengamos un vector $\bar{c}$ tal que

$$\bar{c} = -\bar{a} + 2\bar{b}$$

Efectuando las operaciones indicadas, con las leyes usuales de adición y multiplicación por un escalar, se llega al resultado

$$\bar{c} = (5, 2, 0)$$

Decimos entonces que el vector $\bar{c}$ se obtuvo mediante una combinación lineal de los vectores $\bar{a}$ y $\bar{b}$ o, más brevemente, que el vector $\bar{c}$ es una combinación lineal de los vectores $\bar{a}$ y $\bar{b}$. En general se tiene la siguiente definición.

IX.2.1 DEFINICION

Un vector $\bar{w}$ es una combinación lineal de los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ si puede ser expresado en la forma

$$\bar{w} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \ldots + \alpha_n\bar{v}_n$$

donde $\alpha_1, \alpha_2, \ldots, \alpha_n$ son escalares.

Cabe señalar que lo esencial del concepto de combinación lineal está en la expresión

$$\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n$$

la cual tiene sentido debido a la existencia de las leyes de adición y multiplicación por un escalar en un espacio vectorial, así como a la propiedad asociativa de la adición.

IX.2.2 TEOREMA

Sea $S = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ un conjunto no vacío de vectores de un espacio vectorial V. El conjunto de todas las combinaciones lineales de los vectores de S, denotado con L(S), es un subespacio de V.

DEMOSTRACION

Sean $\overline{w}_1$, $\overline{w}_2$ dos vectores cualesquiera de L(S); esto es

$$\overline{w}_1 = \alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n$$

$$\overline{w}_2 = \beta_1\overline{v}_1 + \beta_2\overline{v}_2 + \ldots + \beta_n\overline{v}_n$$

Entonces, la suma

$$\overline{w}_1 + \overline{w}_2 = (\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n) + (\beta_1\overline{v}_1 + \beta_2\overline{v}_2 + \ldots + \beta_n\overline{v}_n)$$

puede expresarse, aplicando reiteradamente las propiedades ii), v) y viii) de IX.1.1, como

$$\overline{w}_1 + \overline{w}_2 = (\alpha_1 + \beta_1)\overline{v}_1 + (\alpha_2 + \beta_2)\overline{v}_2 + \ldots + (\alpha_n + \beta_n)\overline{v}_n$$

por lo que $\overline{w}_1 + \overline{w}_2 \in L(S)$.

Además, si λ es un escalar

$$\lambda\overline{w}_1 = \lambda(\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n)$$

entonces, por vii) y ix) de IX.1.1

$$\lambda\overline{w}_1 = (\lambda\alpha_1)\overline{v}_1 + (\lambda\alpha_2)\overline{v}_2 + \ldots + (\lambda\alpha_n)\overline{v}_n$$

en consecuencia $\lambda\overline{w}_1 \in L(S)$ y, por IX.1.6, L(S) es un subespacio de V.

$\square$

Así pues, L(S) es un espacio vectorial.

En particular, si $S = \emptyset$ se define a L(S) como $\{\overline{0}\}$; es decir, el espacio cero. Más adelante podrá apreciarse la conveniencia de esta definición.

- Dependencia lineal.

Consideremos nuevamente los vectores $\overline{a} = (3, 0, -2)$, $\overline{b} = (4, 1, -1)$ y $\overline{c} = (5, 2, 0)$ del ejemplo anterior, para los cuales

$$\overline{c} = -\overline{a} + 2\overline{b} \qquad (1)$$

Esta expresión nos indica que el vector $\overline{c}$ es una combinación lineal de los vectores $\overline{a}$ y $\overline{b}$; sin embargo, indica también la existencia de cierta relación entre los tres vectores. Cuando esto sucede se dice que los vectores son linealmente dependientes.

Observe que de la expresión (1) se sigue que

$$\overline{a} = 2\overline{b} - \overline{c} \qquad (2)$$

y también que

$$\bar{b} = \frac{1}{2}\,\bar{a} + \frac{1}{2}\,\bar{c} \qquad (3)$$

Es decir: $\bar{a}$ es una combinación lineal de $\bar{b}$ y $\bar{c}$, y $\bar{b}$ lo es, a su vez, de $\bar{a}$ y $\bar{c}$.

Así, la dependencia lineal es una característica del conjunto formado por los tres vectores, y la relación entre ellos puede indicarse también de la siguiente manera

$$\bar{a} - 2\bar{b} + \bar{c} = \bar{0} \qquad (4)$$

En esta última expresión es el vector cero el que aparece como una combinación lineal de los vectores $\bar{a}$, $\bar{b}$ y $\bar{c}$.

El lector puede advertir, sin embargo, que siempre es posible expresar al vector cero como combinación lineal de un conjunto de vectores, aunque no exista relación alguna entre ellos, simplemente haciendo todos los escalares iguales a cero; por ello, la dependencia lineal se presenta cuando el vector cero es una combinación lineal de un conjunto de vectores, pero con escalares no todos nulos.

IX.2.3 DEFINICION

Sea $S = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ un conjunto de vectores:

i) S es linealmente dependiente si existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, no todos iguales a cero, tales que

$$\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n = \overline{0}$$

ii) S es linealmente independiente si la igualdad

$$\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n = \overline{0}$$

sólo se satisface con $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

A la expresión

$$\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n = \overline{0}$$

se le conoce como ecuación de dependencia lineal.

Así, por ejemplo, para determinar si el conjunto de vectores de R^3

$$A = \{(-1, 0, 2), (0, -4, 2), (2, 0 -4)\}$$

es linealmente dependiente o independiente planteamos la ecuación

$$\alpha(-1, 0, 2) + \beta(0, -4, 2) + \gamma(2, 0, -4) = \overline{0}$$

la cual, para las leyes usuales de adición y multiplicación por un escalar, es equivalente a

$$(-\alpha + 2\gamma, -4\beta, 2\alpha + 2\beta -4\gamma) = (0, 0, 0)$$

Entonces, por definición de igualdad en R^3 dicha expresión es equi_ valente, a su vez, al siguiente sistema homogéneo

$$\begin{aligned} -\alpha + 2\gamma &= 0 \\ -4\beta &= 0 \\ 2\alpha + 2\beta - 4\gamma &= 0 \end{aligned}$$

Escalonando la matriz de coeficientes del sistema se obtiene

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

por lo que dicho sistema admite soluciones no triviales y, en consecuencia, los vectores son linealmente dependientes.†

Si queremos hallar escalares (no todos nulos) que expresen al vector cero como combinación lineal de los vectores de A, obtenemos la solución general del sistema, que es

$$\begin{aligned} \alpha &= 2k \\ \beta &= 0 \\ \gamma &= k \end{aligned}$$

Entonces, para $k = 1$ se tienen los escalares $\alpha = 2$, $\beta = 0$ y $\gamma = 1$; esto es:

$$2(-1, 0, 2) + 0(0, -4, 2) + 1(2, 0, -4) = (0, 0, 0)$$

como puede comprobarse fácilmente.

A diferencia del conjunto A, el conjunto

$$B = \{(1, 0, -2), (-4, 2, 0), (0, 2, -4)\}$$

† *Es frecuente decir que "los vectores" son linealmente dependientes (o independientes) en vez de decir que el conjunto es linealmente dependiente (o independiente).*

es linealmente independiente, ya que la ecuación

$$\alpha(1, 0, -2) + \beta(-4, 2, 0) + \gamma(0, 2, -4) = (0, 0, 0)$$

es equivalente al sistema

$$\begin{aligned} \alpha - 4\beta &= 0 \\ 2\beta + 2\gamma &= 0 \\ -2\alpha - 4\gamma &= 0 \end{aligned}$$

cuya matriz de coeficientes se transforma en

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

por lo que el sistema sólo admite la solución trivial:

$$\alpha = 0, \ \beta = 0 \ \ y \ \ \gamma = 0$$

Observe que, como consecuencia de la definición IX.2.3, un conjunto formado por dos o más vectores es linealmente dependiente cuando al menos uno de ellos es una combinación lineal de los otros vectores del conjunto. En caso contrario; esto es, cuando ninguno de los vectores es combinación lineal de los restantes, el conjunto es linealmente independiente.

Para el ejemplo anterior, tanto el primero como el tercer vector del conjunto A son combinaciones lineales de los otros dos vectores, ya que

$$(-1, 0, 2) = 0(0, -4, 2) - \tfrac{1}{2}(2, 0, -4)$$

$$(2, 0, -4) = -2(-1, 0, 2) + 0(0, -4, 2)$$

mientras que para el conjunto B ninguno de los vectores es combinación lineal de los otros elementos del conjunto.

A continuación se presentan dos importantes teoremas, relativos al concepto de dependencia lineal.

IX.2.4 TEOREMA

Todo conjunto que contiene al vector $\overline{0}$ es linealmente dependiente.

DEMOSTRACION

Sea $S = \{\overline{v}_1, \ldots, \overline{0}, \ldots, \overline{v}_n\}$

Entonces, por i) de IX.1.3 $0\overline{v}_1 + \ldots + \alpha_i\overline{0} + \ldots + 0\overline{v}_n = \overline{0}$, para cualquier $\alpha_i \neq 0$ y, en consecuencia, por i) de IX.2.3, S es linealmente dependiente.

□

IX.2.5 TEOREMA

Si S es un conjunto linealmente independiente entonces cualquier subconjunto de S es linealmente independiente.

DEMOSTRACION

Si $S = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ es linealmente independiente, se tiene que

$$\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n = \overline{0} \iff \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0 \qquad (1)$$

Sea $S' = \{\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_m\}$ un subconjunto de S; entonces, la ecuación

$$\beta_1\overline{w}_1 + \beta_2\overline{w}_2 + \ldots + \beta_m\overline{w}_m = \overline{0}$$

es equivalente a

$$\gamma_1\overline{v}_1 + \gamma_2\overline{v}_2 + \ldots + \gamma_n\overline{v}_n = \overline{0}$$

donde $\gamma_i = \beta_j$ si $\overline{v}_i = \overline{w}_j$ y $\gamma_i = 0$ si $\overline{v}_i \not\in S'$

En consecuencia, de (1) se tiene que $\gamma_1 = \gamma_2 = \ldots = \gamma_n = 0$ por lo que $\beta_1 = \beta_2 = \ldots = \beta_m = 0$ y S' es linealmente independiente.

$\square$

- Conjunto generador

Cuando todos los vectores de un espacio pueden obtenerse mediante combinaciones lineales de un conjunto finito de vectores, se dice que tal conjunto es un generador del espacio.

Por ejemplo, los vectores de R^3 pueden ser expresados como combinaciones lineales de los cuatro vectores del conjunto

$$G = \{(-2, 0, 0), (0, 1, 2), (0, 0, -1), (0, 1, -1)\}$$

como se demuestra a continuación.

Sea (x, y, z) un vector cualquiera de R^3. Dicho vector será una combinación lineal de los vectores de G si existen escalares α, β, γ, δ tales que

$$(x, y, z) = \alpha(-2, 0, 0) + \beta(0, 1, 2) + \gamma(0, 0, -1) + \delta(0, 1, -1) \quad (1)$$

Para la adición, la multiplicación por un escalar y la igualdad usuales en R^3, la expresión (1) es equivalente al sistema

$$\begin{aligned} -2\alpha &= x \\ \beta + \delta &= y \\ 2\beta - \gamma - \delta &= z \end{aligned}$$

para el cual, aplicando el método de Gauss se obtiene

$$\left[\begin{array}{cccc|c} -2 & 0 & 0 & 0 & x \\ 0 & 1 & 0 & 1 & y \\ 0 & 2 & -1 & -1 & z \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -\frac{x}{2} \\ 0 & 1 & 0 & 1 & y \\ 0 & 0 & -1 & -3 & -2y+z \end{array}\right] \rightarrow$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -\frac{x}{2} \\ 0 & 1 & 0 & 1 & y \\ 0 & 0 & 1 & 3 & 2y-z \end{array}\right]$$

por lo que dicho sistema es compatible y, en consecuencia, existen escalares α, β, γ y δ que satisfacen la expresión (1) para cualquier valor de x, y, z; como se quería demostrar.

Por otra parte, como el sistema es indeterminado, para cada vector (x, y, z) de R^3 existe una infinidad de valores para los escalares α, β, γ y δ. De acuerdo a la solución general del sistema, dichos valores son

$$(2) \quad \left\{ \begin{array}{l} \alpha = -\frac{x}{2} \\ \beta = y - k \\ \gamma = 2y - z - 3k \\ \delta = k \quad ; \quad \forall\, k \in R \end{array} \right.$$

Así, para k = 1 se obtiene

$$\begin{array}{l} \alpha = -\frac{x}{2} \\ \beta = y - 1 \\ \gamma = 2y - z - 3 \\ \delta = 1 \end{array}$$

por lo que

$$(x,y,z) = -\frac{x}{2}(-2,0,0)+(y-1)(0,1,2)+(2y-z-3)(0,0,-1)+(0,1,-1) \tag{3}$$

como puede comprobarse fácilmente.

De la misma manera, para $k = -\frac{z}{3}$ se obtiene

$$(x,y,z) = -\frac{x}{2}(-2,0,0)+\frac{3y+z}{3}(0,1,2)+2y(0,0,-1) - \frac{z}{3}(0,1,-1) \tag{4}$$

y así podríamos continuar asignando valores al parámetro k y obteniendo escalares α, β, γ, δ que satisfacen la expresión (1); sin embargo, para mostrar que cualquier vector de R es una combinación lineal de los vectores de G es suficiente cualquiera de las expresiones (3) o (4). Hemos mostrado así que G es un generador de R^3, conforme a la siguiente definición.

> IX.2.6 DEFINICION
>
> Sea V un espacio vectorial sobre K, y sea
>
> $$G = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m\}$$
>
> un conjunto de vectores de V. Se dice que G es un generador de V si para todo vector $\bar{x} \in V$ existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_m$ tales que
>
> $$\bar{x} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \ldots + \alpha_m\bar{v}_m$$

De acuerdo con esta definición, todos los vectores de V son combinaciones lineales de los elementos de G, por lo que

$$V \subseteq L(G)$$

Por otra parte, como G está formado por vectores de V, debido a la cerradura de éste para la adición y la multiplicación por un escalar, toda combinación lineal de vectores de G es un elemento de V; esto es

$$L(G) \subseteq V$$

En consecuencia

$$V = L(G)$$

con lo que hemos demostrado el siguiente teorema

IX.2.7	TEOREMA
	Sea V un espacio vectorial sobre K y sea G un subconjunto de V, G es un generador de V si y sólo si $V = L(G)$.

Como hemos visto, un conjunto generador es un conjunto finito; sin embargo, no todos los espacios vectoriales pueden ser generados por un conjunto finito de valores: el conjunto de todos los polinomios con coeficientes en C no puede ser generado por un conjunto finito, como tampoco el conjunto de todas las funciones reales de variable real. Este tipo de espacios se dice que son de "dimensión infinita".

- Base

Consideremos nuevamente el conjunto generador del ejemplo anterior

$$G = \{(-2, 0, 0), (0, 1, 2), (0, 0, -1), (0, 1, -1)\}$$

Si en la expresión (2) hacemos $k = 0$ se obtiene

$$\alpha = -\frac{x}{2}, \quad \beta = y, \quad \gamma = 2y - z, \quad \delta = 0$$

por lo que

$$(x, y, z) = -\frac{x}{2}(-2, 0, 0) + y(0, 1, 2) + (2y - z)(0, 0, -1) \qquad (5)$$

En esta última expresión se advierte que cualquier vector de R^3 es una combinación lineal de los tres primeros vectores del conjunto G, por lo que el cuarto vector de G no es indispensable para generar a los vectores de R^3.

Así, el conjunto

$$G' = \{(-2, 0, 0), (0, 1, 2), (0, 0, -1)\}$$

es también un generador de R^3, pero con un vector menos que G. Esto se debe a que el conjunto G es linealmente dependiente, como puede verificarse fácilmente.

En general, cuando un conjunto generador es linealmente dependiente podemos suprimir alguno de sus vectores y obtener con ello otro conjunto generador del mismo espacio.

En efecto, sea $G = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m\}$ (con $m > 2$) un generador de V.

Si G es linealmente dependiente, alguno de sus vectores es una combinación lineal de los demás elementos de V. Podemos suponer, sin pérdida de generalidad, que $\bar{v}_m$ satisface tal condición; es decir

$$\bar{v}_m = \beta_1\bar{v}_1 + \beta_2\bar{v}_2 + \ldots + \beta_{m-1}\,\bar{v}_{m-1}$$

Entonces, llevando este resultado a la expresión

$$\bar{x} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \ldots + \alpha_m\bar{v}_m \,, \quad \forall\, \bar{x} \in V$$

se tiene que

$$\overline{x} = \alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_m (\beta_1\overline{v}_1 + \beta_2\overline{v}_2 + \ldots + \beta_{m-1}\overline{v}_{m-1})$$

y finalmente

$$\overline{x} = (\alpha_1 + \alpha_m\beta_1)\ \overline{v}_1 + (\alpha_2 + \alpha_m\beta_2)\ \overline{v}_2 + \ldots + (\alpha_{m-1} + \alpha_m\beta_{m-1})\ \overline{v}_{m-1}$$

por lo que

$$G' = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_{m-1}\}$$

es también un generador de V. □

El resultado anterior puede aplicarse reiteradamente a un conjunto generador linealmente dependiente hasta obtener un conjunto que sea linealmente independiente.

En este punto del procedimiento no es posible ya suprimir un vector y obtener un nuevo conjunto generador, puesto que cualquier vector que se suprima no podrá ser expresado como combinación lineal de los vectores restantes.

Podemos entonces concluir que hemos llegado a un conjunto generador de "tamaño mínimo", que recibe el nombre de base conforme a la siguiente definición.

IX.2.8 DEFINICION

Se llama base de un espacio vectorial V a un conjunto generador de V que es linealmente independiente.

En el ejemplo anterior vimos que el conjunto

$$G' = \{(-2, 0, 0), (0, 1, 2), (0, 0, -1)\}$$

es un generador de R^3; además, como puede comprobarse fácilmente, G' es linealmente independiente. En consecuencia, por la definición IX.2.8 dicho conjunto es una base de R^3.

Como ejemplo adicional, consideremos el conjunto M de matrices triangulares superiores de orden dos, con elementos en R, cuya traza es igual a cero.

Una matriz de dicho conjunto es de la forma

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \text{ con } a + c = 0$$

por lo que

$$M = \left\{ \begin{bmatrix} a & b \\ 0 & -a \end{bmatrix} \middle| \; a, b \in R \right\}$$

El lector puede comprobar que este conjunto es un espacio vectorial sobre R, para las leyes usuales de adición y multiplicación por un escalar.

Si deseamos obtener una base de dicho espacio debemos buscar un conjunto de elementos de M que pueda generar a todos los vectores del espacio, y verificar que dicho conjunto es linealmente independiente.

Se propone así el conjunto

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$$

para el cual vemos que

$$\begin{bmatrix} a & b \\ 0 & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

por lo que B es un generador de M.

Verifiquemos ahora que B es linealmente independiente.

La ecuación

$$\alpha \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \beta \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

es equivalente al sistema

$$\begin{aligned} \alpha &= 0 \\ \beta &= 0 \\ -\alpha &= 0 \end{aligned}$$

cuya única solución es la trivial, por lo que B es linealmente independiente y, en consecuencia, es una base de M.

De la misma manera, puede comprobarse que los siguientes conjuntos son también bases de M

$$C = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \right\}$$

$$D = \left\{ \begin{bmatrix} \frac{1}{5} & 1 \\ 0 & -\frac{1}{5} \end{bmatrix}, \begin{bmatrix} 3 & \sqrt{2} \\ 0 & -3 \end{bmatrix} \right\} \text{ etc.}$$

aunque, sin duda, B es la más "simple" de todas ellas desde diver sos puntos de vista.

- Dimensión.

Con relación al ejemplo anterior, seguramente el lector habrá sospechado que cualquier otra base del espacio M deberá estar formada por dos matrices. Y así es.

En general, todas las bases de un espacio vectorial tienen el mismo número de elementos, como demostraremos más adelante.

Para ello nos apoyaremos en el siguiente teorema

IX.2.9 TEOREMA

Sea V un espacio vectorial sobre K. Si $B = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ es una base de V, entonces cualquier conjunto de vectores de V con más de n elementos es linealmente dependiente.

DEMOSTRACION

Sea $S = \{\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_m\}$ un conjunto de m vectores de V, con $m > n$.

Como $B = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ es una base de V, por IX.2.8 es un generador de V y por IX.2.6 existen escalares α_{ij} tales que

$$\begin{aligned}
\overline{w}_1 &= \alpha_{11}\overline{v}_1 + \alpha_{12}\overline{v}_2 + \ldots + \alpha_{1n}\overline{v}_n \\
\overline{w}_2 &= \alpha_{21}\overline{v}_1 + \alpha_{22}\overline{v}_2 + \ldots + \alpha_{2n}\overline{v}_n \\
&\vdots \\
\overline{w}_m &= \alpha_{m1}\overline{v}_1 + \alpha_{m2}\overline{v}_2 + \ldots + \alpha_{mn}\overline{v}_n
\end{aligned}$$

llevando este resultado a la ecuación

$$\lambda_1\overline{w}_1 + \lambda_2\overline{w}_2 + \ldots + \lambda_m\overline{w}_m = \overline{0}$$

se obtiene

$$\lambda_1(\alpha_{11}\overline{v}_1+\ldots+\alpha_{1n}\overline{v}_n)+\lambda_2(\alpha_{21}\overline{v}_1+\ldots+\alpha_{2n}\overline{v}_n)+\ldots+\lambda_m(\alpha_{m1}\overline{v}_1+\ldots+\alpha_{mn}\overline{v}_n) = \overline{0}$$

esto es

$$(\lambda_1\alpha_{11}+\lambda_2\alpha_{21}+\ldots+\lambda_m\alpha_{m1})\overline{v}_1+\ldots+(\lambda_1\alpha_{1n}+\lambda_2\alpha_{2n}+\ldots+\lambda_m\alpha_{mn})\overline{v}_n = \overline{0}$$

Como B es una base, por IX.2.8 es linealmente independiente y, en consecuencia

$$\begin{aligned}
\lambda_1\alpha_{11} + \lambda_2\alpha_{21} + \ldots + \lambda_m\alpha_{m1} &= 0\\
\lambda_1\alpha_{12} + \lambda_2\alpha_{22} + \ldots + \lambda_m\alpha_{m2} &= 0\\
&\vdots\\
\lambda_1\alpha_{1n} + \lambda_2\alpha_{2n} + \ldots + \lambda_m\alpha_{mn} &= 0
\end{aligned}$$

Como $m > n$ este sistema admite soluciones no triviales para $\lambda_1, \lambda_2, \ldots, \lambda_m$; por lo que el conjunto S es linealmente dependiente.

Podemos ahora demostrar el siguiente enunciado

IX.2.10 TEOREMA

Sea V un espacio vectorial sobre K. Si $B = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ es una base de V, entonces cualquier otra base de dicho espacio está formada por n vectores.

En efecto, sea $B' = \{\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_m\}$ otra base de V, como B' es linealmente independiente y B es una base de V, entonces $m \leqslant n$.

Por otra parte, como B' es una base de V y B es linealmente independiente: $n \leq m$.

En consecuencia

$$m = n$$

como se quería. $\square$

Este resultado nos permite establecer una definición apropiada para el concepto de "dimensión", la cual va más allá del contexto geométrico en el que usualmente se interpreta dicha palabra.

IX.2.11 DEFINICION

Sea V un espacio vectorial sobre K. Si $B = \{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_n\}$ es una base de V se dice que V es de dimensión n, lo cual se denota con

$$\dim V = n$$

En particular, si $V = \{\overline{0}\}$, $\dim V = 0$

Entonces, salvo en el caso del espacio cero (para el cual no existe base), la dimensión de un espacio vectorial queda definida como el número de elementos que tiene cualquier base de dicho espacio.

Así, para el espacio

$$M = \left\{ \begin{bmatrix} a & b \\ 0 & -a \end{bmatrix} \middle| \; a, b \in R \right\}$$

del ejemplo anterior, se tiene que

$$\dim M = 2$$

Por la misma razón podemos concluir que

$$\dim R^3 = 3$$

y para el subespacio de R^3

$$L = \{(x,\ 2x,\ 3x) \mid x \in R\}$$

tenemos que dim L = 1, puesto que el conjunto

$$\{(1,\ 2,\ 3)\}$$

es una base de L.

El concepto de dimensión tiene otras interpretaciones interesantes. Por ejemplo, de la definición IX.2.11 y del teorema IX.2.9 podemos concluir que la dimensión de un espacio es el número máximo de vectores que puede contener un conjunto linealmente independiente en dicho espacio.

A continuación se presentan dos importantes teoremas relacionados con el concepto de dimensión.

IX.2.12 TEOREMA

Si V es un espacio vectorial de dimensión n, cualquier conjunto linealmente independiente formado por n vectores de V es una base de dicho espacio.

DEMOSTRACION

Sea $S = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n\}$ un conjunto linealmente independiente de vectores de V, y supongamos que S no es una base de V.

Entonces, por IX.2.8, S no es un generador de V; esto es

$$\exists \ \bar{x} \in V \quad \text{tal que} \quad \bar{x} \notin L(S)$$

en consecuencia, $\bar{x}$ no es una combinación lineal de los vectores de S por lo que el conjunto

$$S' = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n, \bar{x}\}$$

es linealmente independiente, lo cual contradice el teorema IX.2.9.

Por lo tanto, la hipótesis de que S no es una base de V es falsa y queda demostrado el teorema.

$\square$

IX.2.13 TEOREMA

Si V es un espacio vectorial de dimensión n y W es un subespacio de V, entonces

$$\dim W \leq n$$

En particular, si $\dim W = n$ entonces $W = V$.

DEMOSTRACION

Sea $B = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k\}$ una base de W.

Como $W \subseteq V$, B es también un conjunto de vectores de V y, como es linealmente independiente, por IX.2.9 no puede contener más de n elementos; es decir

$$k \leqslant n$$

como se quería.

En particular, si $k = n$ por IX.2.12, B también es una base de V; por lo que $W = V$.

$\square$

- Coordenadas

Consideremos nuevamente el espacio

$$M = \left\{ \begin{bmatrix} a & b \\ 0 & -a \end{bmatrix} \middle| \; a, b \in R \right\}$$

para el cual el conjunto

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$$

es una base.

Como B es un generador, cualquier matriz de M puede expresarse como combinación lineal de los elementos de B. Por ejemplo, la matriz

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

puede expresarse como

$$A = 7\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + (-3)\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

A los escalares 7 y -3, que expresan a la matriz A como combinación lineal de los elementos de B, se les llama "coordenadas de A en la base B", y a la pareja ordenada $(7, -3)^T$ se le conoce como "vector de coordenadas" de A en la base B y se representa con

$$(A)_B = \begin{bmatrix} 7 \\ -3 \end{bmatrix}$$

El orden en que aparecen los escalares en el vector de coordenadas corresponde al orden que tienen los vectores en la base, por lo que las bases son conjuntos ordenados.

Así, la base

$$F = \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\}$$

se considera distinta a la base B, y se tiene que

$$(A)_F = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$$

IX.2.14 DEFINICION

Sea $B = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ una base de un espacio vectorial V sobre K, y sea $\overline{x} \in V$. Si

$$\overline{x} = \alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n$$

los escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ se llaman coordenadas de $\overline{x}$ en la base B; y el vector de K^n

$$(\overline{x})_B = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$$

se llama vector de coordenadas de $\overline{x}$ en la base B.

Cabe hacer notar que, mediante el concepto de vector de coordenadas, es posible representar a un vector de cualquier espacio vectorial de dimensión finita mediante un conjunto ordenado de escalares; es decir, mediante un elemento de K^n.

Como veremos a continuación, dicha representación es única. Más aún, la función $f: V \to K^n$ definida por

$$f(\overline{x}) = (\overline{x})_B$$

para una base B fija, es un isomorfismo, como se puede demostrar.

IX.2.15 TEOREMA

Sea $B = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ una base de un espacio vectorial V sobre K. Para cualquier $\overline{x} \in V$ el vector $(\overline{x})_B$ es único.

DEMOSTRACION

Sea

$$(\overline{x})_B = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$$

esto es

$$\overline{x} = \alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + x_n\overline{v}_n \qquad (1)$$

y sea

$$(\overline{x})'_B = (\beta_1, \beta_2, \ldots, \beta_n)^T$$

otro vector de coordenadas de $\overline{x}$ en B; es decir

$$\overline{x} = \beta_1\overline{v}_1 + \beta_2\overline{v}_2 + \ldots + \beta_n\overline{v}_n \qquad (2)$$

Restando (2) de (1) se obtiene

$$\overline{0} = (\alpha_1 - \beta_1)\overline{v}_1 + (\alpha_2 - \beta_2)\overline{v}_2 + \ldots + (\alpha_n - \beta_n)\overline{v}_n$$

por lo que, como B es linealmente independiente

$$\begin{aligned} \alpha_1 - \beta_1 &= 0 \\ \alpha_2 - \beta_2 &= 0 \\ &\vdots \\ \alpha_n - \beta_n &= 0 \end{aligned}$$

esto es

$$\begin{aligned} \alpha_1 &= \beta_1 \\ \alpha_2 &= \beta_2 \\ &\vdots \\ \alpha_n &= \beta_n \end{aligned}$$

y en consecuencia $(\overline{x})_B = (\overline{x})'_B$, lo que demuestra el teorema. $\square$

Consideremos nuevamente el conjunto

$$G' = \{(-2, 0, 0), (0, 1, 2), (0, 0, -1)\}$$

que, como vimos, es una base de R^3.

Para dicho conjunto se obtuvo que

$$(x, y, z) = -\frac{x}{2}(-2, 0, 0) + y(0, 1, 2) + (2y - z)(0, 0, -1)$$

por lo que las coordenadas de un vector arbitrario (x, y, z) en la base G' son

$$\left[(x, y, z)\right] = (-\frac{x}{2}, y, 2y - z)^T$$

Si en lugar de G' tomamos la base

$$E = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

para un vector arbitrario (x, y, z) se tendrá que

$$(x, y, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1)$$

por lo que sus coordenadas en la base E son

$$\left[(x, y, z)\right]_E = (x, y, z)^T$$

esto es, coinciden con sus propias componentes.

Como se puede demostrar, E es la única base de R^3 que tiene tales características y recibe el nombre de "base canónica" de R^3.

Para R^2 la base canónica es el conjunto

$$\{(1, 0), (0, 1)\}$$

Y, en general, la base canónica de R^n es el conjunto

$\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$

cuyos vectores

$$\bar{e}_i = (e_{i1}, e_{i2}, \ldots, e_{in})$$

son tales que

$$e_{ij} = \begin{cases} 1 & \text{para} \quad i = j \\ 0 & \text{para} \quad i \neq j \end{cases}$$

IX.2.16 EJERCICIOS

1.- Sean $\bar{u}$ y $\bar{v}$ los siguientes vectores de R^3

$\bar{u} = (1, 2, 2)$ y $\bar{v} = (-1, 3, 1)$

a) Expresar a cada uno de los vectores $(3, 1, 3)$, $(2, -6, -2)$ y $(-1, 8, 4)$ como una combinación lineal de $\bar{u}$ y $\bar{v}$.

b) ¿Para qué valores de k es el vector $(5, k, 1)$ una combinación lineal de $\bar{u}$ y $\bar{v}$?

2.- Determinar si los siguientes conjuntos de vectores son linealmente dependientes o independientes, en sus respectivos espacios vectoriales reales:

a) $\{(0,1), (0,-3)\}$

b) $\left\{ \begin{bmatrix} 1 & 5 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 5 & 3 \\ 2 & -1 \end{bmatrix} \right\}$

c) $\{x^2 - x + 2, x^2 + 1, -x^2 + x - 3\}$

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

3.- Sean $\overline{u}$, $\overline{v}$ y $\overline{w}$ tres vectores linealmente independientes en un espacio vectorial real. Verificar que los vectores $\overline{r}$, $\overline{s}$, $\overline{t}$:

a) Son linealmente dependientes si

$\overline{r} = -\overline{v}$, $\overline{s} = -2\overline{u} + 3\overline{v} + 2\overline{w}$ y $\overline{t} = \overline{u} - \overline{v} - \overline{w}$

b) Son linealmente independientes si

$\overline{r} = \overline{u} + \overline{v} + 2\overline{w}$, $\overline{s} = \overline{w}$ y $\overline{t} = \overline{u} + 2\overline{v} - 5\overline{w}$

4.- Como puede verificarse, el conjunto C de los números complejos es un espacio vectorial tanto sobre R como sobre C.

Determinar si el conjunto $S = \{1 + 3i \ , \ 1 - 2i\}$ es linealmente dependiente o independiente cuando:

a) El campo es R.

b) El campo es C.

Como se desprende de este ejemplo, la dependencia o independencia lineal de un conjunto de vectores puede variar al cambiar el campo en el que se consideran los escalares.

5.- Sea $S = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$

Demostrar que si

$\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n = \beta_1\overline{v}_1 + \beta_2\overline{v}_2 + \ldots + \beta_n\overline{v}_n$

y $\exists\, i$ tal que $\alpha_i \neq \beta_i$, donde $i = 1, 2, \ldots, n$,

entonces S es linealmente dependiente.

6.- Determinar, en cada caso, si el conjunto A es generador del espacio V indicado:

a)

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

$$V = \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \;\middle|\; a, b, c \in R \right\}$$

b) $A = \{x^2 -2x + 1,\ 3x + 2,\ -x^2 + x -1\}$

$$V = \{ax^2 + bx + c \mid a, b, c \in R\}$$

c) $A = \{(1, 0, 1), (0, 3, -1), (0, 4, 1), (2, 1, 0)\}$

$$V = \{(a, b, c) \mid a, b, c \in R\}$$

7.- Obtener una base del espacio generado por

$A = \{x^2 -3x + 2,\ 4x + 1,\ x^2 + x + 3,\ -x^2 + 7x -1\}$

y decir cuál es la dimensión de dicho espacio.

8.- Sea $A = \{\bar{s}, \bar{t}, \bar{u}, \bar{v}\}$ un conjunto generador de un espacio vectorial real V, donde

$$\bar{s} - 2\bar{t} = 2\bar{u}$$
$$\bar{s} - \bar{u} = \bar{v}$$
$$2\bar{v} - 2\bar{t} = \bar{s}$$

Obtener una base de V y la dimensión de éste.

9.- Sean los espacios

$$V = \{ax^2 + bx + 2a + b \mid a, b \in R\} \quad \text{sobre } R$$

y

$$W = \left\{ \begin{bmatrix} a & a+b \\ c & 2b-c \end{bmatrix} \middle| \ a, b, c \in C \right\} \quad \text{sobre } C$$

a) Obtener dos bases de V y dos bases de W, comprobando en cada caso que se satisface la definición IX.2.8.

b) Decir cuál es la dimensión de V y cuál la de W.

c) Hallar el vector de coordenadas de $\bar{v}_1 = x^2 - 2x$ en cada una de las bases de V obtenidas en a) y hallar el vector de coordenadas de

$$\bar{v}_2 = \begin{bmatrix} 1 & 2-i \\ 3-2i & -1 \end{bmatrix}$$ en cada una de las bases de W obtenidas en a)

10.- Sean $A = \{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ y $B = \{\bar{w}_1, \bar{w}_2, \bar{w}_3\}$ dos bases de un espacio vectorial V sobre R, para las cuales

$$\bar{w}_1 = \bar{v}_1 + 2\bar{v}_2$$

$$\bar{w}_2 = \bar{v}_1 + \bar{v}_2 + \bar{v}_3$$

$$\bar{w}_3 = \bar{v}_2 - 2\bar{v}_3$$

Determinar las coordenadas de $\bar{x} \in V$ en la base B si

$$(\bar{x})_A = (1, 0, 2)^T$$

11.- Demostrar que la función $f: V \to K^n$ definida por

$$f(\bar{x}) = (\bar{x})_B$$

donde V es un espacio vectorial sobre K con dim V = n y B es una base de V, es un isomorfismo.

Esto es, demostrar que:

i) $(\bar{x})_B = (\bar{y})_B \iff \bar{x} = \bar{y}$

ii) $[\bar{x} + \bar{y}]_B = [\bar{x}]_B + [\bar{y}]_B; \quad \forall\ \bar{x}, \bar{y} \in V$

iii) $[\alpha\bar{x}]_B = \alpha[\bar{x}]_B; \quad \forall\ \bar{x} \in V, \alpha \in K$

IX.3 ESPACIOS VECTORIALES ASOCIADOS A UNA MATRIZ. TEORIA DE LOS SISTEMAS DE ECUACIONES LINEALES.

A partir de los elementos que integran una matriz pueden definirse diversos espacios vectoriales. Dos de ellos, conocidos como espacio renglón y espacio columna, son de especial importancia en las aplicaciones del Algebra Lineal.

Al estudiar el primero de éstos veremos, además, un procedimiento sencillo para obtener bases de espacios vectoriales. La definición del segundo nos permitirá, a su vez, establecer resultados teóricos importantes acerca de las soluciones de un sistema de ecuaciones lineales.

- Espacio renglón.

Sea A una matriz de m×n con elementos en R

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

sus renglones pueden ser considerados como vectores de R^n; esto es:

$$\bar{r}_1 = (a_{11}, a_{12}, \ldots, a_{1n})$$
$$\bar{r}_2 = (a_{21}, a_{22}, \ldots, a_{2n})$$
$$\vdots$$
$$\bar{r}_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

El conjunto de todas las combinaciones lineales de dichos vectores es un subespacio de R^n al que se conoce como "espacio generado por los renglones de A" o, más brevemente, como "espacio renglón de A".

Claramente, este concepto puede generalizarse al caso en que los elementos de la matriz sean números complejos o elementos de un campo K arbitrario, con el consecuente reemplazo de R^n por C^n o K^n. De esta manera, las definiciones y teoremas se enuncian para el caso general de un campo K arbitrario; aunque, por sencillez, los ejemplos se presentan sobre el campo de los números reales.

IX.3.1 DEFINICION

Sea $A = \left[a_{ij}\right]$ una matriz de m×n con elementos en un campo K, y sea $\overline{r}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ el i-ésimo renglón de A. Si $A_r = \{\overline{r}_1, \overline{r}_2, \ldots, \overline{r}_m\}$, el conjunto $L(A_r)$ se llama espacio renglón de A.

Observe que $\overline{r}_1, \overline{r}_2, \ldots, \overline{r}_m$ son elementos de K^n por lo que, de IX.2.2, L(A) es un subespacio de K^n.

Es posible que los renglones de una matriz en particular no sean linealmente independientes; es decir, que alguno de ellos sea una combinación lineal de los otros renglones.

Para determinar si éste es el caso, lo cual pocas veces puede advertirse a simple vista, se emplea un procedimiento que consiste en aplicar a la matriz una sucesión de transformaciones elementales hasta obtener lo que se conoce como su forma canónica escalonada.

Como resultado de ello se sabe cuál es el número máximo de renglones linealmente independientes de la matriz y, además, se obtiene una base muy simple de su espacio renglón, la cual permite definir claramente dicho espacio.

Este procedimiento se basa en el hecho de que el espacio renglón de una matriz no cambia al aplicarle una sucesión finita de transformaciones elementales, como justificaremos formalmente a continuación.

IX.3.2 DEFINICION

Dos matrices A y B son equivalentes (por renglones), lo cual se denota mediante $A \sim B$, si alguna de ellas puede obtenerse a partir de la otra mediante una sucesión finita de transformaciones elementales (por renglón).

Es claro que si A puede obtenerse como resultado de aplicar a B una sucesión finita de transformaciones elementales, también B puede obtenerse a partir de A aplicando a ésta las transformaciones elementales inversas de las que integran la sucesión mencionada.

IX.3.3 TEOREMA

Dos matrices equivalentes tienen el mismo espacio renglón

DEMOSTRACION

Sea $A_r = \{\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_i, \ldots, \bar{r}_j, \ldots, \bar{r}_m\}$ el conjunto de los renglones de A:

i) Si B se obtiene de A intercambiando los renglones i, j; entonces

$$B_r = \{\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_j, \ldots, \bar{r}_i, \ldots, \bar{r}_m\}$$

por lo que

$$L(B_r) = \{\bar{x} | \bar{x} = \alpha_1\bar{r}_1 + \alpha_2\bar{r}_2 + \ldots + \alpha_j\bar{r}_j + \ldots + \alpha_i\bar{r}_i + \ldots + \alpha_m\bar{r}_m\}$$
$$= \{\bar{x} | \bar{x} = \alpha_1\bar{r}_1 + \alpha_2\bar{r}_2 + \ldots + \alpha_i\bar{r}_i + \ldots + \alpha_j\bar{r}_j + \ldots + \alpha_m\bar{r}_m\}$$

$$L(B_r) = L(A_r)$$

ii) Si B se obtiene de A multiplicando el renglón i por un escalar $\lambda \neq 0$, entonces

$$B_r = \{\bar{r}_1, \bar{r}_2, \ldots, \lambda\bar{r}_i, \ldots, \bar{r}_j, \ldots, \bar{r}_m\}$$

por lo que

$$L(B_r) = \{\bar{x} | \bar{x} = \alpha_1\bar{r}_1 + \alpha_2\bar{r}_2 + \ldots + \alpha_i(\lambda\bar{r}_i) + \ldots + \alpha_j\bar{r}_j + \ldots + \alpha_m\bar{r}_m\}$$
$$= \{\bar{x} | \bar{x} = \alpha_1\bar{r}_1 + \alpha_2\bar{r}_2 + \ldots + (\lambda\alpha_i)\bar{r}_i + \ldots + \alpha_j\bar{r}_j + \ldots + \alpha_m\bar{r}_m\}$$

$$L(B_r) = L(A_r)$$

iii) Si B se obtiene de A multiplicando el renglón i por λ y sumándolo al renglón j, entonces

$$B_r = \{\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_i, \ldots, \lambda\bar{r}_i + \bar{r}_j, \ldots, \bar{r}_m\}$$

por lo que

$$L(B_r) = \{\bar{x} | \bar{x} = \alpha_1\bar{r}_1 + \alpha_2\bar{r}_2 + \ldots + \alpha_i\bar{r}_i + \ldots$$
$$\ldots + \alpha_j(\lambda\bar{r}_i + \bar{r}_j) + \ldots + \alpha_m\bar{r}_m\}$$
$$= \{\bar{x} | \bar{x} = \alpha_1\bar{r}_1 + \alpha_2\bar{r}_2 + \ldots + (\alpha_i + \lambda\alpha_j)\bar{r}_i + \ldots$$
$$\ldots + \alpha_j\bar{r}_j + \ldots + \alpha_m\bar{r}_m\}$$

$$L(B_r) = L(A_r)$$

De i), ii) y iii), el espacio renglón de una matriz no cambia al efectuar en ella una transformación elemental (de cualquiera de los tres tipos).

Aplicando reiteradamente este resultado, el espacio renglón de una matriz no cambia si efectuamos en ella una sucesión finita de transformaciones elementales; por lo que matrices equivalentes tienen el mismo espacio renglón.

□

En el procedimiento que mencionamos, la aplicación sucesiva de transformaciones elementales se efectúa hasta obtener una "forma canónica escalonada".

Se dice que una matriz es una forma canónica escalonada cuando además de ser una matriz escalonada, el primer elemento distinto de cero en cada renglón es un uno y dicho elemento es el único diferente de cero en la columna en que se encuentra.

Por ejemplo, de las siguientes matrices escalonadas sólo la última es una forma canónica escalonada

$$\begin{bmatrix} 1 & -3 & 0 & 2 & 4 \\ 0 & 2 & 1 & -5 & 3 \\ 0 & 0 & 0 & -4 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 0 & 1 & 0 & 5 & -1 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & -3 & 0 & 1 \\ 0 & 1 & 4 & 0 & -2 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Dos propiedades importantes, relativas a las formas canónicas escalonadas, son las siguientes:

- Para una matriz dada A existe una y sólo una forma canónica escalonada que es equivalente a la matriz A.

- Los renglones no nulos de una forma canónica escalonada constituyen una base de su espacio renglón.

La última propiedad es válida también para una matriz escalonada cualquiera, pero en el caso de una forma canónica es aún más evidente.

A manera de ejemplo utilizamos a continuación el procedimiento mencionado para obtener el espacio renglón de la matriz

$$A = \begin{bmatrix} 3 & -1 & -5 & 9 \\ 1 & -1 & -3 & 3 \\ 2 & 0 & -2 & 6 \end{bmatrix}$$

Aplicando a dicha matriz transformaciones elementales hasta obtener su forma canónica equivalente se obtiene:

$$A = \begin{bmatrix} 3 & -1 & -5 & 9 \\ 1 & -1 & -3 & 3 \\ 2 & 0 & -2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -3 & 3 \\ 3 & -1 & -5 & 9 \\ 2 & 0 & -2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -3 & 3 \\ 0 & 2 & 4 & 0 \\ 0 & 2 & 4 & 0 \end{bmatrix} \sim$$

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

Por ser esta última matriz una forma canónica escalonada, los vectores (1, 0, -1, 3) y (0, 1, 2, 0), que son sus renglones no nulos, constituyen un conjunto linealmente independiente y, por tanto, una base de su espacio renglón.

Por el teorema IX.3.3, dicho espacio coincide con el espacio renglón de A, por lo que el conjunto

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

también es una base de $L(A_r)$.

Entonces, cualquier vector de dicho espacio será de la forma

$$a(1, 0, -1, 3) + b(0, 1, 2, 0) = (a, b, -a + 2b, 3a)$$

y en consecuencia

$$L(A_r) = \{(a, b, -a + 2b, 3a) \mid a, b \in R\}$$

El espacio renglón de A es, en este caso, un subespacio de R^4 de dimensión dos.

- Espacio columna, rango.

De manera similar al espacio renglón, se define el "espacio columna" de una matriz como el conjunto de todas las combinaciones lineales de las columnas de dicha matriz; esto es

> IX.3.4 DEFINICION
>
> Sea $A = \left[a_{ij}\right]$ una matriz de m×n con elementos en un campo K, y sea $\bar{c}_i = (a_{1i}, a_{2i}, \ldots, a_{mi})^T$ la i-ésima columna de A. Si $A_c = \{\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_n\}$, el conjunto $L(A_c)$ se llama espacio columna de A.

Estrictamente hablando, hemos definido las columnas de una matriz como arreglos "verticales", a diferencia de los renglones a los que hemos considerado "horizontales"; empero, no hay razón para que los vectores columna no puedan ser escritos y manejados como arreglos horizontales cuando así sea conveniente.

En consecuencia, como las columnas de una matriz son los renglones de su transpuesta, el espacio columna de una matriz A coincide con el espacio renglón de A^T (salvo en el "aspecto" vertical y horizontal de los arreglos). De esta manera, podemos obtener el espacio columna de A aplicando a A^T el procedimiento anterior.

Así, por ejemplo, para obtener el espacio columna de la matriz

$$A = \begin{bmatrix} 3 & -1 & -5 & 9 \\ 1 & -1 & -3 & 3 \\ 2 & 0 & -2 & 6 \end{bmatrix}$$

del ejemplo anterior, efectuamos transformaciones elementales sobre su transpuesta hasta obtener una forma canónica escalonada

$$A^T = \begin{bmatrix} 3 & 1 & 2 \\ -1 & -1 & 0 \\ -5 & -3 & -2 \\ 9 & 3 & 6 \end{bmatrix} \sim \begin{bmatrix} -1 & -1 & 0 \\ 3 & 1 & 2 \\ -5 & -3 & -2 \\ 9 & 3 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & -2 & 2 \\ 0 & 2 & -2 \\ 0 & -6 & 6 \end{bmatrix} \sim$$

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

De donde

$$L(A_r^T) = \{(a,b,a-b) \mid a,\ b \in R\}$$

por lo que

$$L(A_c) = \{(a,b,a-b) \mid a,\ b \in R\}$$

o bien, si se prefiere en forma vertical

$$L(A_c) = \left\{ \begin{bmatrix} a \\ b \\ a-b \end{bmatrix} \middle|\ a,\ b \in R \right\}$$

Pasando a otro aspecto, con referencia a los espacios renglón y columna de la matriz A del ejemplo vimos que

$$L(A_r) \subset R^4 \qquad y \qquad L(A_c) \subset R^3$$

por lo que, evidentemente

$$L(A_r) \neq L(A_c)$$

Sin embargo, también se obtuvo que

$$\dim L(A_r) = \dim L(A_c)$$

Esto no es una casualidad debida a la matriz que se eligió para el ejemplo, sino una propiedad general como se establece a continuación

IX.3.5 TEOREMA

Para cualquier matriz A se tiene que

$$\dim L(A_r) = \dim L(A_c)$$

DEMOSTRACION

Sea $A = \left[a_{ij}\right]$ una matriz de m×n y supongamos que su espacio renglón es de dimensión $k \leq m$; es decir

$$\dim L(A_r) = k$$

y que $B = \{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k\}$ es una base de $L(A_r)$.

Formemos una matriz X con los vectores de B como renglones; esto es

$$X = \begin{bmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \vdots \\ \overline{x}_k \end{bmatrix}$$

Como B es una base de $L(A_r)$, cada uno de los renglones de A es una combinación lineal de $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$; por lo que

$$\bar{r}_1 = y_{11}\bar{x}_1 + y_{12}\bar{x}_2 + \ldots + y_{1k}\bar{x}_k$$
$$\bar{r}_2 = y_{21}\bar{x}_1 + y_{22}\bar{x}_2 + \ldots + y_{2k}\bar{x}_k$$
$$\vdots$$
$$\bar{r}_m = y_{m1}\bar{x}_1 + y_{m2}\bar{x}_2 + \ldots + y_{mk}\bar{x}_k$$

o bien, en forma matricial

$$A = YX$$

donde

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1k} \\ y_{21} & y_{22} & \cdots & y_{2k} \\ \vdots & \vdots & & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mk} \end{bmatrix}$$

Si expresamos ahora a las matrices A, Y, X como

$$A = \begin{bmatrix} \bar{c}_1, & \bar{c}_2, \ldots, & \bar{c}_n \end{bmatrix}, \quad Y = \begin{bmatrix} \bar{y}_1, & \bar{y}_2, \ldots, & \bar{y}_k \end{bmatrix} \text{ y } X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{bmatrix}$$

La igualdad A = YX corresponde a

$$\begin{bmatrix} \bar{c}_1, & \bar{c}_2, \ldots, & \bar{c}_n \end{bmatrix} = \begin{bmatrix} \bar{y}_1, & \bar{y}_2, \ldots, & \bar{y}_k \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{bmatrix}$$

por lo que

$$\bar{c}_1 = x_{11}\bar{y}_1 + x_{21}\bar{y}_2 + \ldots + x_{k1}\bar{y}_k$$
$$\bar{c}_2 = x_{12}\bar{y}_1 + x_{22}\bar{y}_2 + \ldots + x_{k2}\bar{y}_k$$
$$\vdots$$
$$\bar{c}_n = x_{1n}\bar{y}_1 + x_{2n}\bar{y}_2 + \ldots + x_{kn}\bar{y}_k$$

y las columnas de A son combinaciones lineales de las k columnas que forman la matriz Y; en consecuencia

$$\dim L(A_c) \leqslant k$$

esto es

$$\dim L(A_c) \leqslant \dim L(A_r).$$

Mediante un procedimiento similar se demuestra que también

$$\dim L(A_r) \leqslant \dim L(A_c)$$

por lo que

$$\dim L(A_r) = \dim L(A_c)$$

como se quería.

□

El teorema anterior nos permite definir el concepto de rango como se enuncia a continuación

IX.3.6 DEFINICION

Se llama rango de una matriz A, y se denota con R(A), al número

$$R(A) = \dim L(A_r) = \dim L(A_c)$$

De esta manera, para la matriz

$$A = \begin{bmatrix} 3 & -1 & -5 & 9 \\ 1 & -1 & -3 & 3 \\ 2 & 0 & -2 & 6 \end{bmatrix}$$

que hemos empleado en los ejemplos anteriores, se tiene que

$$R(A) = 2.$$

De la definición IX.3.6, es claro que el rango de una matriz representa el número máximo de renglones (y de columnas) linealmente independientes que contiene la matriz.

En particular, para una matriz cuadrada el concepto de rango tiene relación con otras características importantes como se indica en el siguiente teorema, cuya demostración se deja al lector.

> IX.3.7 TEOREMA
>
> Si A es una matriz de n×n, los siguientes enunciados son equivalentes:
>
> i) $R(A) = n$
>
> ii) $A \sim I_n$
>
> iii) $\exists\ A^{-1}$
>
> iv) $\det A \neq 0$
>
> v) Los renglones de A son linealmente independientes
>
> vi) Las columnas de A son linealmente independientes

- Condiciones para la existencia y unicidad de soluciones de un sistema

En este apartado se establece una condición necesaria y suficiente para la existencia de soluciones de un sistema de ecuaciones lineales, utilizando el concepto de rango. Dicho concepto se emplea también para deducir las condiciones de unicidad y multiplicidad de soluciones.

Consideremos para ello un sistema arbitrario de m ecuaciones lineales con n incógnitas

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array} \qquad \text{--- (I)}$$

Como se vio en el capítulo VI, dicho sistema es equivalente a la ecuación matricial

$$A\overline{x} = \overline{b} \qquad \text{--- (II)}$$

donde

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \qquad \overline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \overline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Estas matrices, como sabemos, reciben los nombres de "matriz de coeficientes", "vector de incógnitas" y "vector de términos independientes", respectivamente.

Definamos ahora una nueva matriz, a la que llamaremos "matriz ampliada" del sistema y representaremos con $(A,\overline{b})$, de la siguiente manera

$$(A,\bar{b}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

Para el análisis que se desarrolla a continuación nos apoyaremos en la siguiente expresión vectorial planteada en R^n, o en K^n si se prefiere, la cual es también equivalente al sistema (I):

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \qquad \text{---(III)}$$

De esta expresión se sigue que el sistema tiene solución si, y solamente si, existen escalares $x_1, x_2, \ldots, x_n$ que expresen a $\bar{b}$ como combinación lineal de las columnas de A; esto es, si y sólo si $\bar{b} \in L(A_c)$.

Para enunciar esta condición en términos de los rangos de A y $(A,\bar{b})$ bastará con observar que la matriz $(A,\bar{b})$ está formada por las mismas columnas de A y por una columna adicional, que es $\bar{b}$. Entonces, cuando $\bar{b} \in L(A_c)$ los rangos de A y $(A,\bar{b})$ son iguales; y cuando $\bar{b} \notin L(A_c)$ el rango de $(A,\bar{b})$ es mayor, en una unidad, que el rango de A.

En consecuencia, podemos concluir que

IX.3.8 TEOREMA

El sistema de ecuaciones lineales $A\bar{x} = \bar{b}$ es compatible si y sólo si $R(A) = R(A,\bar{b})$

Supongamos ahora que el sistema (I) es compatible y veamos bajo qué condiciones es determinado, comparando el rango de A con el número de incógnitas del sistema:

Si R(A) = n, el conjunto formado por las n columnas de A es linealmente independiente y, por tanto, constituye una base de $L(A_c)$; entonces los escalares $x_1, x_2, \ldots, x_n$ son las coordenadas del vector $\overline{b}$ en dicha base y, de IX.2.15, el sistema tiene solución única.

Supongamos ahora que R(A) = r < n; entonces cualquier base de $L(A_c)$ contiene r columnas de A.

Consideremos, sin pérdida de generalidad, que las primeras r columnas de A constituyen una base de dicho espacio†; de esta manera la expresión (III) puede escribirse como

$$\overline{b} = x_1\overline{c}_1 + \ldots + x_r\overline{c}_r + x_{r+1}\overline{c}_{r+1} + \ldots + x_n\overline{c}_n$$

donde $\{\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_r\}$ es una base de $L(A_c)$.

Debido a la cerradura de un espacio vectorial para la adición y la multiplicación por un escalar se tiene que, para cualquier valor de los escalares $x_{r+1}, \ldots, x_n$, el vector

$$x_{r+1}\overline{c}_{r+1} + \ldots + x_n\overline{c}_n$$

es un elemento de $L(A_c)$; como también lo es el vector

$$\overline{b} - (x_{r+1}\overline{c}_{r+1} + \ldots + x_n\overline{c}_n)$$

† *En caso de que las primeras r columnas no sean independientes podemos reescribir la expresión (III), debido a la conmutatividad de la adición, hasta que las primeras r columnas de A que aparezcan sean linealmente independientes.*

debido a la cerradura para la sustracción.

En consecuencia, para cualquier valor de $x_{r+1},\ldots, x_n$ la ecuación

$$\bar{b} - (x_{r+1}\bar{c}_{r+1} + \ldots + x_n\bar{c}_n) = x_1\bar{c}_1 + \ldots + x_r\bar{c}_r$$

tiene solución única para las variables $x_1, x_2,\ldots, x_r$.

Esto implica que el sistema (I) es indeterminado, puesto que podemos asignar valores arbitrarios a las incógnitas $x_{r+1},\ldots, x_n$ y obtener los correspondientes valores de $x_1, x_2,\ldots, x_r$ que satisfacen la expresión (III).

En resumen puede concluirse lo siguiente

> IX.3.9 TEOREMA
>
> Sea $A\bar{x} = \bar{b}$ un sistema compatible de m ecuaciones lineales con n incógnitas: si $R(A) = n$ el sistema es determinado y si $R(A) < n$ el sistema es indeterminado.

- Estructura del conjunto solución.

Consideremos nuevamente un sistema arbitrario de m ecuaciones lineales con n incógnitas, cuya representación matricial es

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

Una solución de dicho sistema es cualquier vector $\bar{x}_0 \in R^n$ tal que $A\bar{x}_0 = \bar{b}$. Al conjunto formado por todas las soluciones de un sistema se le conoce como "conjunto solución" del mismo, conforme a la siguiente definición

IX.3.10 DEFINICION

Sean A y $\bar{b}$ dos matrices de m×n y m×1 respectivamente, al conjunto $\{\bar{x} \mid A\bar{x} = \bar{b}\}$ se le llama conjunto solución del sistema $A\bar{x} = \bar{b}$.

Es claro que cada sistema en particular tiene su propio conjunto solución. Dicho conjunto puede ser vacío (en el caso de un sistema incompatible), puede tener un solo elemento (sistema compatible determinado) o puede estar constituido por un número infinito de elementos (sistema compatible indeterminado), dependiendo de los coeficientes y términos independientes que tenga el sistema.

Para obtener el conjunto solución de un sistema dado no se requiere introducir un nuevo procedimiento. Bastará con hallar sus soluciones mediante los procedimientos estudiados en el capítulo V y expresarlas como un conjunto de vectores.

Por ejemplo, para el sistema

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

en el capítulo V se concluyó que tiene como solución única la siguiente

$$x_1 = -1, \; x_2 = 0, \; x_3 = 2$$

por lo que su conjunto solución es

$$\{(-1, 0, 2)\}$$

el cual está formado por un solo vector.

De manera similar, para hallar el conjunto solución del sistema

$$\begin{aligned} -x_1 + x_2 - x_3 + x_4 &= 4 \\ -2x_1 + x_2 - x_3 + 3x_4 &= 5 \\ 3x_1 + x_2 - x_3 - 7x_4 &= 0 \\ 2x_2 - 2x_3 - 2x_4 &= 6 \end{aligned}$$

aplicamos el método de Gauss y obtenemos

$$\left[\begin{array}{cccc|c} -1 & 1 & -1 & 1 & 4 \\ -2 & 1 & -1 & 3 & 5 \\ 3 & 1 & -1 & -7 & 0 \\ 0 & 2 & -2 & -2 & 6 \end{array}\right] \quad \left[\begin{array}{cccc|c} -1 & 1 & -1 & 1 & 4 \\ 0 & -1 & 1 & 1 & -3 \\ 0 & 4 & -4 & -4 & 12 \\ 0 & 2 & -2 & -2 & 6 \end{array}\right] \quad \left[\begin{array}{cccc|c} 1 & -1 & 1 & -1 & -4 \\ 0 & 1 & -1 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{aligned} x_1 &= 2x_4 - 1 \\ x_2 &= x_3 + x_4 + 3 \\ x_3 &= x_3 \\ x_4 &= x_4 \end{aligned}$$

por lo que su conjunto solución es

$$\{(2b-1,\ a+b+3,\ a,\ b) \mid a,\ b \in R\}$$

el cual contiene un número infinito de vectores.

En este apartado estudiaremos la estructura que tiene el conjunto solución de un sistema arbitrario de ecuaciones lineales.

Primeramente observamos que el conjunto solución de un sistema de m ecuaciones lineales con n incógnitas es un subconjunto de R^n (en general de K^n), por lo que cabe preguntarse si será un subespacio de éste.

La respuesta depende únicamente de que el sistema sea homogéneo o no lo sea. En el primer caso la respuesta es afirmativa, como lo establece el siguiente teorema, mientras que en el segundo es negativa como veremos más adelante.

> IX.3.11 TEOREMA
>
> El conjunto solución del sistema $A\overline{x} = \overline{0}$ es un espacio vectorial

DEMOSTRACION

Sea A una matriz de m×n y sean $\overline{x}_1$, $\overline{x}_2$ ε R^n dos soluciones del sistema $A\overline{x} = \overline{0}$; esto es

$$A\overline{x}_1 = \overline{0} \quad \text{y} \quad A\overline{x}_2 = \overline{0}$$

Entonces, por i) de VI.3.3, por hipótesis y por iii) de IX.1.1, respectivamente

$$A(\overline{x}_1 + \overline{x}_2) = A\overline{x}_1 + A\overline{x}_2 = \overline{0} + \overline{0} = \overline{0}$$

y en consecuencia $\overline{x}_1 + \overline{x}_2$ es solución del sistema.

Además, si α es un escalar, por el enunciado del ejercicio 3 de VI.3.6, por hipótesis y por i) de IX.1.3 se tiene que

$$A(\alpha\overline{x}_1) = \alpha(A\overline{x}_1) = \alpha\overline{0} = \overline{0}$$

por lo que $\alpha\overline{x}_1$ es solución del sistema.

Finalmente, por IX.1.6, el conjunto de soluciones de $A\overline{x} = \overline{0}$ es un subespacio de R^n y, por IX.1.5, es un espacio vectorial.

□

Debido a esto se acostumbra hablar del "espacio solución" de un sistema homogéneo.

Así, por ejemplo, para el sistema

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

su espacio solución es

$$\{(-\frac{z}{2}, 0, z) \mid z \in R\}$$

Puede verificarse fácilmente que éste es un subespacio de R^3.

Regresemos ahora al caso general del sistema no homogéneo

$$A\bar{x} = \bar{b}, \text{ con } \bar{b} \neq \bar{0}$$

y sean $\bar{x}_1$, $\bar{x}_2$ dos soluciones de dicho sistema.

Para la suma se tiene

$$A(\bar{x}_1 + \bar{x}_2) = A\bar{x}_1 + A\bar{x}_2 = \bar{b} + \bar{b} \neq \bar{b}$$

por lo que $\bar{x}_1 + \bar{x}_2$ no es solución del sistema y, en consecuencia, su conjunto solución no es un espacio vectorial.

Sin embargo, si en lugar de la suma consideramos la diferencia de dos soluciones tenemos

$$A(\bar{x}_1 - \bar{x}_2) = A\bar{x}_1 - A\bar{x}_2 = \bar{b} - \bar{b} = \bar{0}$$

por lo que

IX.3.12 TEOREMA

Si $\bar{x}_1$, $\bar{x}_2$ son soluciones del sistema $A\bar{x} = \bar{b}$, entonces $\bar{x}_1 - \bar{x}_2$ es solución del sistema $A\bar{x} = \bar{0}$

Al sistema $A\bar{x} = \bar{0}$ se le conoce como "sistema homogéneo asociado" al sistema $A\bar{x} = \bar{b}$. Es evidente que para cualquier sistema no homogéneo existe uno y sólo un sistema homogéneo asociado.

De acuerdo con el teorema IX.3.12, la diferencia de dos soluciones cualesquiera del sistema $A\bar{x} = \bar{b}$ es una solución del homogéneo asociado. En particular, si $\bar{x}_0$ representa una solución fija y $\bar{x}$ una solución cualquiera de $A\bar{x} = \bar{b}$, la diferencia $\bar{x} - \bar{x}_0$ es una solución de su sistema homogéneo asociado; por lo que

$$\bar{x} - \bar{x}_0 = \bar{u}$$

para algún vector $\bar{u}$ del espacio solución de $A\bar{x} = \bar{0}$.

De lo anterior se deduce que

LEMA

Si U es el espacio solución del sistema homogéneo $A\bar{x} = \bar{0}$ y $\bar{x}_0$ es una solución fija del sistema $A\bar{x} = \bar{b}$, entonces cualquier solución $\bar{x}$ del sistema $A\bar{x} = \bar{b}$ es tal que

$$\bar{x} = \bar{u} + \bar{x}_0$$

para algún $\bar{u} \in U$.

Con relación a este enunciado se puede concluir, además, que para cada $\bar{x}$ el vector $\bar{u}$ es único.

En efecto, si $\bar{u}_1$, $\bar{u}_2$ son tales que

$$\bar{x} = \bar{u}_1 + \bar{x}_0 \qquad y \qquad \bar{x} = \bar{u}_2 + \bar{x}_0$$

entonces, de $\bar{x} - \bar{x} = \bar{0}$ se sigue que

$$(\bar{u}_1 + \bar{x}_0) - (\bar{u}_2 + \bar{x}_0) = \bar{0}$$

por lo que

$$\bar{u}_1 + \bar{x}_0 = \bar{u}_2 + \bar{x}_0$$

y en consecuencia

$$\bar{u}_1 = \bar{u}_2$$

$\square$

Así pues, podemos afirmar que para cada solución $\bar{x}$ del sistema $A\bar{x} = \bar{b}$ existe una y sólo una solución $\bar{u}$ del homogéneo $A\bar{x} = \bar{0}$ tal que $\bar{x} = \bar{u} + \bar{x}_0$.

Sin embargo, no podemos todavía asegurar que para cada $\bar{u}$ exista una y sólo una $\bar{x}$ que satisfaga dicha relación.

Probaremos esto a continuación mostrando, primero, que para cada solución $\bar{u}$ de $A\bar{x} = \bar{0}$ el vector $\bar{x} = \bar{u} + \bar{x}_0$ es solución de $A\bar{x} = \bar{b}$. En efecto

$$A(\bar{u} + \bar{x}_0) = A\bar{u} + A\bar{x}_0 = \bar{0} + \bar{b} = \bar{b}$$

Además, la unicidad de $\bar{x}$ es inmediata ya que la adición asigna a cada pareja $(\bar{u}, \bar{x}_0)$ uno y sólo un vector $\bar{u} + \bar{x}_0$.

Finalmente podemos concluir que

IX.3.13 TEOREMA

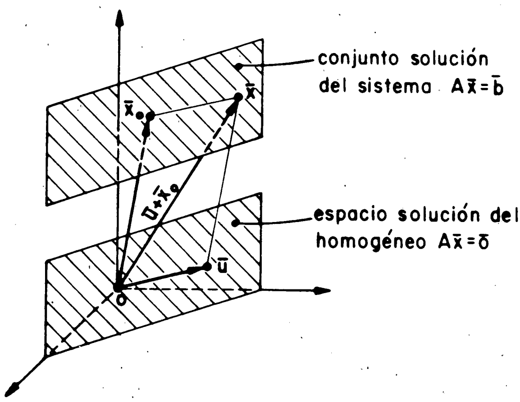
Si $\overline{x}$ es una solución cualquiera del sistema $A\overline{x} = \overline{b}$ y $\overline{u}$ es una solución cualquiera del sistema $A\overline{x} = \overline{0}$, entonces para cada solución $\overline{x}_0$ de $A\overline{x} = \overline{b}$ la relación

$$\overline{x} = \overline{u} + \overline{x}_0$$

establece una biyección entre el conjunto solución del sistema $A\overline{x} = \overline{b}$ y el espacio solución del sistema $A\overline{x} = \overline{0}$

En otras palabras: cada solución $\overline{x}$ del sistema $A\overline{x} = \overline{b}$ tiene su correspondiente "reflejo" $\overline{u}$ en el espacio solución del sistema homogéneo $A\overline{x} = \overline{0}$, determinado unívocamente por la expresión $\overline{x} = \overline{u} + \overline{x}_0$; y viceversa.

Una interpretación geométrica de este resultado podría ser la siguiente



En consecuencia, para el conjunto solución del sistema $A\overline{x} = \overline{b}$ se tiene que

IX.3.14 TEOREMA

Si U es el espacio solución del sistema homogéneo $A\overline{x} = \overline{0}$ y $\overline{x}_0$ es una solución del sistema $A\overline{x} = \overline{b}$, entonces

$$\{\overline{u} + \overline{x}_0 \mid \overline{u} \in U\}$$

es el conjunto solución del sistema $A\overline{x} = \overline{b}$.

Así, por ejemplo, para el siguiente sistema de ecuaciones lineales

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

aplicando el método de Gauss se obtiene

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

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

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

por lo que su conjunto solución es

$$S = \{(1 - \frac{z}{2}, -1, z) \mid z \in R\}$$

Dicho conjunto puede ser expresado como

$$S = \{(-\frac{z}{2}, 0, z) + (1, -1, 0) \mid z \in R\} \qquad --- (1)$$

donde

$$U = \{(-\frac{z}{2}, 0, z) \mid z \in R\}$$

es el espacio solución del sistema homogéneo asociado y

$$\bar{x}_0 = (1, -1, 0)$$

es una solución particular del sistema.

Ahora bien, del teorema IX.3.14 se sigue que la expresión (1) no es única, sino que existe una expresión de este tipo para cada solución $\bar{x}_0$ del sistema.

En la expresión (1) la solución $\bar{x}_0 = (1, -1, 0)$ corresponde al caso $z = 0$ de la solución general obtenida; sin embargo, podemos seleccionar cualquier otra solución particular para efectuar la descomposición a que se refiere el teorema IX.3.14.

Por ejemplo, para $z = 6$ se tiene la solución

$$\bar{x}_0 = (-2, -1, 6)$$

y el vector $(1 - \frac{z}{2}, -1, z)$ puede expresarse como

$$(1 - \frac{z}{2}, -1, z) = (1 - \frac{z}{2} + 2, 0, z - 6) + (-2, -1, 6)$$

$$(1 - \frac{z}{2}, -1, z) = (\frac{6 - z}{2}, 0, z - 6) + (-2, -1, 6)$$

Entonces, haciendo $z - 6 = k$ se tiene la siguiente expresión para el conjunto solución del sistema

$$S = \{(-\frac{k}{2}, 0, k) + (-2, -1, 6) \mid k \in R\} \qquad --- (2)$$

donde

$$\{(-\frac{k}{2}, 0, k) \mid k \in R\}$$

en el mismo espacio solución U del sistema homogéneo asociado.

- Variedad lineal

Como ya vimos, el conjunto solución del sistema $A\bar{x} = \bar{b}$, con $\bar{b} \neq \bar{0}$, no tiene estructura de espacio vectorial; sin embargo, los teoremas IX.3.13 y IX.3.14 sugieren que dicho conjunto (cuando no es vacío) es una especie de "espacio vectorial trasladado".

Este tipo de estructura recibe, en general, el nombre de variedad lineal.

IX.3.15 DEFINICION

Sea V un espacio vectorial sobre K. Un subconjunto L de V se dice que es una variedad lineal de V si

$$L = \{\bar{w} + \bar{v}_0 \mid \bar{w} \in W\}$$

para algún subespacio W de V y algún vector $\bar{v}_0 \in V$. Al subespacio W se le llama espacio asociado y al vector $\bar{v}_0$ apoyo, de la variedad lineal.

Algunas propiedades importantes de este tipo de estructura son las siguientes

IX.3.16 TEOREMA

i) El espacio asociado a una variedad lineal es único.

ii) Cualquier vector de una variedad lineal es un apoyo de ésta.

iii) Dos variedades lineales con el mismo espacio asociado son iguales si y sólo si tienen algún vector en común.

iv) Una variedad lineal es un subespacio si y sólo si contiene el vector cero.

v) Cualquier conjunto constituido por un solo vector es una variedad lineal.

DEMOSTRACION

i) Sean W, U dos espacios asociados a una variedad lineal L con apoyo $\bar{v}_0$, y sea $\bar{x}$ un elemento cualquiera de L; entonces

$$\bar{x} = \bar{w} + \bar{v}_0 \text{ , para uno y sólo un } \bar{w} \in W$$

y

$$\bar{x} = \bar{u} + \bar{v}_0 \text{ , para uno y sólo un } \bar{u} \in U$$

en consecuencia

$$\bar{w} + \bar{v}_0 = \bar{u} + \bar{v}_0$$

por lo que, $\forall\ \bar{x} \in L$:

$$\bar{w} = \bar{u}$$

de donde

$$W = U.$$

ii) Sea L una variedad lineal con espacio asociado W y apoyo $\bar{v}_0$, y sea $\bar{x}$ un elemento cualquiera de L; entonces

$$\bar{x} = \bar{w} + \bar{v}_0\text{, para uno y sólo un } \bar{w} \in W$$

si $\overline{y}_0 \in L$ se tendrá que

$$\overline{y}_0 = \overline{w}_0 + \overline{v}_0 \ , \text{ donde } \overline{w}_0 \in W$$

y en consecuencia

$$- \overline{w}_0 + \overline{y}_0 = \overline{v}_0$$

entonces

$$\overline{x} = \overline{w} + \overline{v}_0 = \overline{w} + (- \overline{w}_0 + \overline{y}_0) = (\overline{w} - \overline{w}_0) + \overline{y}_0$$

Como $\overline{w}_0 \in W$ y W es un subespacio: $\overline{u} = \overline{w} - \overline{w}_0 \in W$; de donde

$$\overline{x} = \overline{u} + \overline{y}_0, \text{ para uno y sólo un } \overline{u} \in W.$$

Es decir

$$L = \{\overline{u} + \overline{y}_0 \mid \overline{u} \in W\}$$

por lo que $\overline{y}_0$ es un apoyo de L.

La demostración de las propiedades restantes se deja al lector como ejercicio

$\square$

IX.3.17 EJERCICIOS

1.- Para la matriz

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

Obtener:

a) Su forma canónica escalonada.

b) Su espacio renglón, dando de él una base y la dimensión.

c) Su espacio columna, dando de él una base y la dimensión.

2.- Determinar cuáles de las siguientes matrices tienen el mismo espacio renglón.

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

¿Qué conclusión puede obtenerse del resultado anterior respecto al teorema IX.3.13?

3.- Para la matriz

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

Determinar si cada uno de los vectores

$$\bar{x} = (-1,\ 2,\ 3)^T$$
$$\bar{y} = (3,\ 1,\ -2)^T$$

pertenece al espacio generado por sus columnas y en caso afirmativo expresarlo como una combinación lineal de ellas.

4.- Para la matriz

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

a) Verificar que su espacio renglón y su espacio columna tienen la misma dimensión ¿Son iguales dichos espacios?

b) ¿Cuál es el rango de A?

5.- Demostrar que si A es una matriz de $n \times n$, los siguientes enunciados son equivalentes

a) $R(A) = n$

b) $\det A \neq 0$

c) Los renglones de A son linealmente independientes.

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

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

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

Determinar

i) Su espacio solución

ii) El espacio renglón de su matriz de coeficientes

iii) El espacio columna de su matriz de coeficientes

7.- Para cada uno de los sistemas del ejercicio anterior determinar el rango de la matriz de coeficientes y la dimensión del conjunto solución.

Generalizar los resultados obtenidos para un sistema de m ecuaciones con n incógnitas con $R(A) = r$ y demostrarlo.

8.- Demostrar que:

a) Dos variedades lineales con el mismo espacio asociado son iguales si y sólo si tienen algún vector en común.

b) Una variedad lineal es un subespacio si y sólo si contiene al vector cero.

c) Cualquier conjunto constituido por un solo vector es una variedad lineal.

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

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

b)
$$\begin{aligned} -2x_1 + x_2 - x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 1 \\ -x_1 + x_2 - 3x_3 &= -5 \end{aligned}$$

Hallar su conjunto solución y mostrar que tiene estructura de variedad lineal.

IX.4 ESPACIO DE FUNCIONES

Seguramente el lector ha tenido ya amplia experiencia en el manejo de funciones. En sus cursos de cálculo se habrá familiarizado con conceptos tales como dominio, gráfica, derivada, etc; los cuales forman parte del estudio de las funciones como modelos matemáticos de fenómenos físicos.

En esta sección nos ocuparemos de las funciones desde un punto de vista más bien algebraico.

Las funciones constituyen un espacio vectorial para las operaciones de adición y multiplicación por un escalar; y el estudio de dicho espacio tiene un interés especial por tratarse de un espacio de dimensión infinita.

El análisis que aquí presentamos se limita al caso de las funciones reales de variable real, aunque la mayoría de las conclusiones pueden generalizarse para funciones de un conjunto S en un campo K.

Iniciamos nuestro estudio puntualizando aquellos conceptos básicos que, aunque ya debe conocer el lector, adquieren una importancia relevante desde el punto de vista algebraico.

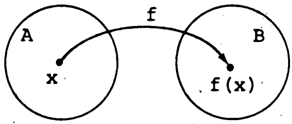
- El espacio vectorial de las funciones reales de variable real.

Recordemos que una función de un conjunto A en un conjunto B es una regla que asocia a cada elemento de A uno y sólo un elemento de B.

Para indicar que f es una función de A en B escribimos

$$f: A \rightarrow B$$

Si $x \in A$, el símbolo $f(x)$ representa al elemento de B asociado a x por la función f, al cual se conoce como "imagen de x" bajo la función f. En la siguiente figura se encuentran representados estos conceptos.



$x \in A$ $\quad$ $f(x) \in B$

En particular, si A y B son ambos el conjunto de los números reales f será una función del tipo

$$f: R \to R$$

llamada "función real de variable real".

Como ejemplo de este tipo de funciones podemos citar a las funciones f, g y h definidas mediante las expresiones siguientes

$$f(x) = x^2 \quad , \quad \forall x \in R$$
$$g(x) = -\sqrt{2} \text{ sen } x \quad , \quad \forall x \in R$$
$$h(x) = 2x^3 + e^{-2x} \quad , \quad \forall x \in R$$

La función f, así definida, asocia a cada número real x el número (real) que se obtiene elevando éste al cuadrado; la función g, por su parte, asigna a cada número real x el que se obtiene multiplicando el valor del seno de éste por el número $-\sqrt{2}$; etc.

La igualdad de funciones se define usualmente de la siguiente manera

IX.4.1 DEFINICION

Sea F el conjunto de todas las funciones reales de variable real, y sean f, g ε F. Se dice que f y g son iguales, lo cual se denota mediante f = g, cuando

$$f(x) = g(x), \forall x \varepsilon R$$

Es muy importante hacer notar que la igualdad de dos funciones equivale a un número infinito de igualdades entre números reales, una para cada valor de x, lo cual constituye una primera diferencia fundamental con los espacios vectoriales que hemos manejado hasta ahora.

A manera de ejemplo, consideremos las funciones f, g ε F definidas por

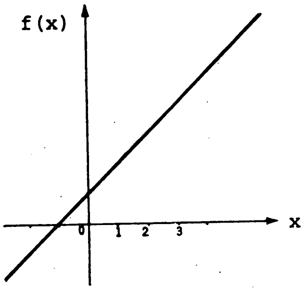
$$f(x) = x + 1 \qquad , \quad \forall x \varepsilon R$$

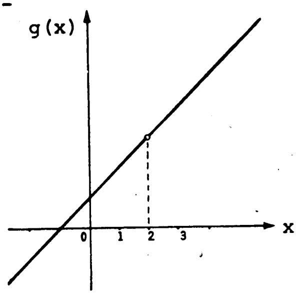
$$g(x) = \frac{x^2 - x - 2}{x - 2} \quad , \quad \forall x \varepsilon R$$

Estas funciones no son iguales ya que la condición f(x) = g(x), aunque se cumple para muchos valores de x, no se satisface para x = 2 (puesto que g(2) no está definida). Escribimos entonces

$$f \neq g$$

Las funciones f y g de este ejemplo están representadas gráficamente por las líneas rectas que se muestran a continuación





En algunas ocasiones, y fundamentalmente con propósitos de aplicación, suele hablarse de "igualdad por intervalos", reemplazando la condición

$$f(x) = g(x), \ \forall x \in R$$

por

$$f(x) = g(x), \ \forall x \in I$$

donde I es un intervalo dado.

Así, por ejemplo, para las funciones f y g definidas anteriormente se cumple que

$$f(x) = g(x), \ \forall x \in (0,2)$$

por lo que se dice que f es igual a g en el intervalo (0,2).

Para estas mismas funciones se tiene

$f = g$ en el intervalo $(2,\infty)$

$f = g$ en el intervalo $[-1,1]$

$f \neq g$ en el intervalo $(0,3)$; etc.

En el conjunto F, la adición y la multiplicación por un escalar se definen como sigue

IX.4.2 DEFINICION

Sea $\mathcal{F}$ el conjunto de funciones reales de variable real, y sean $f, g \in \mathcal{F}$; $\alpha \in R$:

i) La suma de f y g es una función $f + g$ definida por

$$(f + g)(x) = f(x) + g(x), \ \forall\, x \in R.$$

ii) El producto de α y f es una función αf definida por

$$(\alpha f)(x) = \alpha \cdot f(x), \ \forall\, x \in R.$$

El lector puede verificar fácilmente que el conjunto $\mathcal{F}$, con las operaciones de adición y multiplicación por un escalar así definidas, es un espacio vectorial sobre R.

El vector cero de dicho espacio es una función, frecuentemente denotada por θ, tal que

$$\theta(x) = 0, \ \forall\, x \in R$$

En efecto

$$
\begin{aligned}
(f + \theta)(x) &= f(x) + \theta(x), & \forall\, x \in R & \quad \text{por i) de IX.4.2}\\
&= f(x) + 0, & \forall\, x \in R & \quad \text{por definición de } \theta\\
(f + \theta)(x) &= f(x), & \forall\, x \in R & \quad \text{por iv) de I.5.3}
\end{aligned}
$$

entonces

$$f + \theta = f \qquad \text{por IX.4.1}$$

y en consecuencia θ es el vector cero de $\mathcal{F}$.

□

- Dependencia lineal de funciones.

Puesto que F es un espacio vectorial podemos hablar de combinaciones lineales, dependencia e independencia lineal de funciones.

Por ejemplo, la función f definida por

$$f(x) = \operatorname{sen}^2 x, \ \forall x \in R$$

es una combinación lineal de las funciones g y h tales que

$$g(x) = \cos^2 x \quad y \quad h(x) = 5, \ \forall x \in R$$

En efecto, de la identidad† trigonométrica

$$\operatorname{sen}^2 x + \cos^2 x = 1, \ \forall x \in R$$

se sigue que

$$\operatorname{sen}^2 x = -\cos^2 x + 1 \quad , \ \forall x \in R$$

$$\operatorname{sen}^2 x = -\cos^2 x + \frac{1}{5}(5), \ \forall x \in R$$

Entonces, por la definición de f, g y h en este ejemplo

$$f(x) = -[g(x)] + \frac{1}{5}[h(x)], \ \forall x \in R$$

y como $-[g(x)] = (-g)(x)$ (Demuéstrelo), se tiene que

$$f(x) = (-g)(x) + \frac{1}{5}[h(x)], \ \forall x \in R$$

de donde, por ii) y i) de IX.4.2

$$f(x) = (-g)(x) + (\tfrac{1}{5}h)(x), \ \forall x \in R$$

$$f(x) = (-g + \tfrac{1}{5}h)(x) \quad , \ \forall x \in R$$

† *igualdad que se cumple para cualquier valor de la variable.*

Finalmente, por IX.4.1

$$f = -g + \frac{1}{5} h$$

y en consecuencia f es una combinación lineal de g y h.

Cabe hacer notar que la relación anterior también puede ser expresada como

$$f + g - \frac{1}{5} h = \theta$$

donde la función cero aparece como combinación lineal de las funciones f, g y h, con escalares no todos nulos. Se confirma con esto que el conjunto de funciones {f, g, h} es linealmente dependiente, de acuerdo con la definición IX.2.3.

En este ejemplo, la dependencia lineal de las funciones f, g y h pudo determinarse fácilmente gracias a una conocida identidad trigonométrica; sin embargo, no siempre resulta sencillo establecer si un conjunto dado de funciones es linealmente dependiente o independiente.

A continuación estudiaremos este problema para el caso general de un conjunto constituido por n funciones cualesquiera $f_1, f_2, \ldots, f_n$

En este caso la ecuación de dependencia lineal de la definición IX.2.3 es

$$\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n = \theta \qquad --- \text{(I)}$$

Por IX.4.1, IX.4.2 y la definición de la función cero, dicha ecuación es equivalente a

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) + \ldots + \alpha_n f_n(x) = 0, \; \forall \, x \in R \qquad --- \text{(II)}$$

Esta expresión representa un número infinito de ecuaciones: una para cada número real x.

Así, cualquier solución $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ de la ecuación (I) deberá satisfacer la igualdad de la expresión (II) para todos los valores reales de x.

Para determinar si existen soluciones no triviales de la ecuación (I), quizás lo primero que se ocurre es asignar diversos valores a x en la expresión (II) para llegar a un sistema de ecuaciones lineales, del cual trataremos de obtener un conjunto de escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ que satisfagan dicha ecuación.

Si asignamos a x los valores $x_1, x_2, \ldots, x_n$, se llega al siguiente sistema homogéneo de n ecuaciones lineales con n incógnitas.

$$\begin{array}{ll}
\alpha_1 f_1(x_1) + \alpha_2 f_2(x_1) + \ldots + \alpha_n f_n(x_1) = 0 & \quad (\text{para } x = x_1) \\
\alpha_1 f_1(x_2) + \alpha_2 f_2(x_2) + \ldots + \alpha_n f_n(x_2) = 0 & \quad (\text{para } x = x_2) \\
\cdots\cdots\cdots\cdots\cdots\cdots & \quad \vdots \\
\alpha_1 f_1(x_n) + \alpha_2 f_2(x_n) + \ldots + \alpha_n f_n(x_n) = 0 & \quad (\text{para } x = x_n)
\end{array}$$

Dicho sistema puede tener solución distinta de la trivial o puede no tenerla.

En el primer caso, la obtención de una solución distinta de la trivial no permite llegar a conclusión alguna respecto a la dependencia lineal de las funciones, ya que sólo puede asegurarse que la solución obtenida satisface la igualdad de la expresión (II) para los n valores de x que hemos elegido; en consecuencia, no sabemos si dicha solución satisface la igualdad para los demás valores reales de x.

En el segundo caso, cuando el sistema sólo admite la solución trivial, sí puede concluirse que el conjunto de funciones es lineal - mente independiente.

En efecto, si no existe un conjunto de escalares (no todos nulos) que satisfaga la igualdad de la expresión (II) para n valores da - dos de x, tampoco existirá un conjunto de escalares que la satisfa ga para todos los valores reales de x.

De lo anterior se concluye el siguiente enunciado

IX.4.3 TEOREMA

Sea $\{f_1, f_2, \ldots, f_n\}$ un conjunto de n funciones reales de varia - ble real. Si existen n valores $x_1, x_2, \ldots, x_n \in R$ tales que el sistema

$$\begin{array}{l} \alpha_1 f_1(x_1) + \alpha_2 f_2(x_1) + \ldots + \alpha_n f_n(x_1) = 0 \\ \alpha_1 f_1(x_2) + \alpha_2 f_2(x_2) + \ldots + \alpha_n f_n(x_2) = 0 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \alpha_1 f_1(x_n) + \alpha_2 f_2(x_n) + \ldots + \alpha_n f_n(x_n) = 0 \end{array}$$

sólo admite la solución trivial, entonces el conjunto de funciones es linealmente independiente.

A manera de ejemplo apliquemos el método que sugiere este teorema al conjunto constituido por las funciones f_1, f_2 tales que

$$f_1(x) = x \text{ sen } x \quad \text{y} \quad f_2(x) = \cos x \ , \ \forall\, x \in R$$

La ecuación de dependencia lineal es

$$\alpha_1 f_1 + \alpha_2 f_2 = \theta$$

la cual es equivalente a la expresión

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) = 0, \forall x \in R$$

que para las funciones del ejemplo es la siguiente

$$\alpha_1 x \text{ sen } x + \alpha_2 \cos x = 0, \forall x \in R$$

Si asignamos a x los valores $x_1 = 0$ y $x_2 = \pi$, de la expresión anterior se obtienen las ecuaciones

$$\alpha_1 \cdot 0 + \alpha_2 \cdot 1 = 0 \qquad \text{(para } x = 0\text{)}$$
$$\alpha_1 \cdot 0 + \alpha_2(-1) = 0 \qquad \text{(para } x = \pi\text{)}$$

las cuales integran un sistema homogéneo que admite soluciones no triviales de la forma

$$\alpha_1 = k$$
$$\alpha_2 = 0, \text{ con } k \neq 0$$

Como vimos, esto no permite concluir respecto a la dependencia lineal de las funciones f_1 y f_2, puesto que la igualdad

$$k\, x \text{ sen } x + 0 \cdot \cos x = 0, \text{ con } k \neq 0$$

sólo puede asegurarse que se cumple para los valores $x = 0$ y $x = \pi$.

Si ahora asignamos a x los valores 0 y $\frac{\pi}{2}$, se llega al sistema

$$\alpha_1 \cdot 0 + \alpha_2 \cdot 1 = 0$$
$$\alpha_1 \cdot \frac{\pi}{2} + \alpha_2 \cdot 0 = 0$$

cuya única solución es la trivial; esto es $\alpha_1 = \alpha_2 = 0$.

En consecuencia, por IX.4.3, el conjunto $\{f_1, f_2\}$ es linealmente independiente.

Como se deduce del ejemplo, llegar a una conclusión respecto a la independencia lineal empleando el procedimiento anterior puede depender de la elección que hagamos para los valores $x_1, x_2, \ldots, x_n$.

Hay otro procedimiento, basado también en una condición suficiente para la independencia lineal, en el cual no se requiere elegir conjuntos de valores. Dicho procedimiento es aplicable a conjuntos de funciones derivables y se refiere a la independencia lineal "por intervalos", consecuencia de la "igualdad por intervalos" que vimos en seguida de la definición IX.4.1.

- El Wronskiano.

Consideremos un conjunto de funciones $f_1, f_2, \ldots, f_n$, derivables al menos n-1 veces en el intervalo (a,b).

La ecuación de dependencia lineal para el conjunto $\{f_1, f_2, \ldots, f_n\}$ en el intervalo (a,b) es

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) + \ldots + \alpha_n f_n(x) = 0, \quad \forall\, x \in (a,b) \qquad (1)$$

Derivando n-1 veces, la expresión (1) implica que

$$\begin{array}{lr} \alpha_1 f_1'(x) + \alpha_2 f_2'(x) + \ldots + \alpha_n f_n'(x) = 0, \quad \forall\, x \in (a,b) & (2) \\ \alpha_1 f_1''(x) + \alpha_2 f_2''(x) + \ldots + \alpha_n f_n''(x) = 0, \quad \forall\, x \in (a,b) & (3) \\ \vdots \qquad\qquad \vdots \quad \vdots \qquad\qquad \vdots & \vdots \\ \alpha_1 f_1^{(n-1)}(x) + \alpha_2 f_2^{(n-1)}(x) + \ldots + \alpha_n f_n^{(n-1)}(x) = 0, \quad \forall\, x \in (a,b) & (n) \end{array}$$

Ahora bien, para cada valor de x en el intervalo (a,b), las igualdades de las expresiones (1) a (n) constituyen un sistema homogéneo de n ecuaciones lineales con n incógnitas.

Definamos una función $W(x)$ cuyo valor sea igual al determinante de la matriz de coeficientes de dicho sistema; esto es

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \dots & f_n(x) \\ f_1'(x) & f_2'(x) & \dots & f_n'(x) \\ \vdots & & & \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \dots & f_n^{(n-1)}(x) \end{vmatrix}$$

Esta función se conoce como el Wronskiano† de las funciones $f_1, f_2, \dots, f_n$, en el intervalo (a,b).

Supongamos ahora que existe un valor x_0, en dicho intervalo, para el cual

$$W(x_0) \neq 0$$

Entonces, para dicho valor el sistema homogéneo sólo admite la solución trivial.

Luego, como las ecuaciones (2) a (n) son consecuencia de la ecuación (1), ésta no puede tener solución distinta de la trivial para $x = x_0$; por lo que la única solución de (1) $\forall\ x \in (a,b)$ es $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$, y el conjunto $\{f_1, f_2, \dots, f_n\}$ es linealmente independiente.

Hemos demostrado con ello el siguiente enunciado.

†*En recuerdo del matemático polaco J.M. Hoene Wronsky, quien la introdujo.*

IX.4.4 TEOREMA

Sea $\{f_1, f_2, \ldots, f_n\}$ un conjunto de n funciones reales de variable real, derivables al menos n-1 veces en el intervalo (a,b); y sea

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \ldots & f_n^{(n-1)}(x) \end{vmatrix}$$

Si $W(x_0) \neq 0$ para algún $x_0 \in (a,b)$, entonces el conjunto de funciones es linealmente independiente en dicho intervalo.

Utilicemos ahora este teorema para demostrar que las funciones del ejemplo anterior son linealmente independientes en el intervalo $(-\infty, \infty)$.

En efecto, el Wronskiano de las funciones f_1, f_2, definidas por

$$f_1(x) = x \operatorname{sen} x \quad y \quad f_2(x) = \cos x, \quad \forall\, x \in R$$

es

$$W(x) = \begin{vmatrix} x \operatorname{sen} x & \cos x \\ x \cos x + \operatorname{sen} x & -\operatorname{sen} x \end{vmatrix} = -x \operatorname{sen}^2 x - x \cos^2 x - \operatorname{sen} x \cos x$$

$$W(x) = -x(\operatorname{sen}^2 x + \cos^2 x) - \operatorname{sen} x \cos x = -x - \operatorname{sen} x \cos x.$$

y es claro que existen valores reales para los cuales $W(x)$ es diferente de cero. Por ejemplo, para $x_0 = \pi$

$$W(x_0) = W(\pi) = -\pi - \operatorname{sen}\pi\cos\pi = -\pi - (0)(-1) = -\pi \neq 0$$

por lo que las funciones son linealmente independientes en cualquier intervalo que contenga al punto $x_0 = \pi$.

Cabe hacer notar que el teorema IX.4.4 establece una condición suficiente para la independencia lineal de funciones en un intervalo; sin embargo, dicha condición no es necesaria ya que el Wronskiano puede ser nulo en todo el intervalo para un conjunto linealmente independiente, como se muestra en el siguiente ejemplo.

Sean f_1 y f_2 tales que

$$f_1(x) = x^2 \qquad y \qquad f_2(x) = x\cdot|x|; \quad \forall\, x \in R$$

Entonces, para $x \geqslant 0$

$$W(x) = \begin{vmatrix} x^2 & x^2 \\ 2x & 2x \end{vmatrix} = 2x^3 - 2x^3 = 0$$

y para $x < 0$

$$W(x) = \begin{vmatrix} x^2 & -x^2 \\ 2x & -2x \end{vmatrix} = -2x^3 + 2x^3 = 0$$

por lo que

$$W(x) = 0, \quad \forall\, x \in R$$

Sin embargo, el conjunto $\{f_1, f_2\}$ es linealmente independiente en el intervalo $(-\infty, \infty)$ ya que, para $x \geqslant 0$ la ecuación

$$\alpha_1 f_1 + \alpha_2 f_2 = \theta$$

es equivalente a

$$\alpha_1 x^2 + \alpha_2 x^2 = 0, \quad \forall\, x \geqslant 0$$

cuyas soluciones no triviales son de la forma

$$\alpha_1 = k \quad y \quad \alpha_2 = -k, \text{ con } k \neq 0$$

Dichas soluciones, sin embargo, no satisfacen la condición

$$\alpha_1 x^2 + \alpha_2(-x^2) = 0, \quad \forall\, x < 0$$

por lo que la única solución de

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) = 0, \quad \forall\, x \in R$$

es $\alpha_1 = \alpha_2 = 0$ y las funciones son linealmente independientes en el intervalo $(-\infty, \infty)$.

- Subespacios de funciones.

A diferencia de los otros espacios vectoriales que hemos visto, el espacio F de las funciones reales de variable real no puede ser generado por un conjunto finito de vectores y se dice por ello que es de dimensión infinita.

A los diversos subespacios de F se les llama "espacios de funciones" o "espacios funcionales". Como ejemplos de estos podemos citar los siguientes:

1) El conjunto de los polinomios.
2) El conjunto de los polinomios de grado menor o igual que n.
3) El conjunto de las funciones definidas en un intervalo.
4) El conjunto de las funciones continuas en un intervalo.
5) El conjunto de las funciones derivables en un punto.
6) El conjunto de las funciones integrables en un intervalo.
7) El conjunto de las soluciones de la ecuación diferencial $y'' + ay' + by = 0$, etc.

Algunos de estos subespacios son de dimensión finita, como es el caso de los ejemplos 2 y 7.

El espacio del ejemplo 2 está constituido por todos los polinomios de la forma

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

donde $a_1, a_2, \ldots, a_n$ son constantes.

Es evidente que, para un n fijo, el conjunto

$$\{1, x, x^2, \ldots, x^n\}$$

es un generador de dicho espacio.

Además, se puede demostrar (se sugiere al lector hacerlo) que dicho conjunto es linealmente independiente para cualquier valor de n, por lo que es una base del espacio en cuestión y su dimensión es n+1.

En lo que se refiere al espacio del ejemplo 7, con ayuda de otros conceptos (algunos de los cuales están fuera de los temas de este libro), se puede demostrar que dicho espacio es de dimensión dos y establecer un procedimiento para hallar una base del mismo.

IX.4.5 EJERCICIOS

1.- Hallar $(f + g)(x)$ y $(\alpha f)(x)$ para

a) $$f(x) = \begin{cases} 2x^2 & \forall\ x \leq 2 \\ \text{sen}\ x & \forall\ x > 2 \end{cases} \qquad g(x) = \begin{cases} -x^2 & \forall\ x \in (0,4) \\ 0 & \forall\ x \leq 0\ y\ \forall\ x \geq 4 \end{cases}$$

b)
$$f(x) = \begin{cases} \operatorname{sen}^2 x & \forall\ x < 0 \\ 1 & \forall\ x \geqslant 0 \end{cases} \qquad g(x) = \begin{cases} \cos^2 x & \forall\ x \leqslant 5 \\ 0 & \forall\ x > 5 \end{cases}$$

2.- Demostrar que en el conjunto F de funciones reales de variable real, se cumplen las propiedades

a) $f + g = g + f$

b) $\forall\ f \in F\ \exists\ -f$ tal que $f + (-f) = \theta$

c) $\alpha(\beta f) = (\alpha\beta) f$

para la adición y multiplicación por un escalar definidas en IX.4.2.

3.- Para cada uno de los siguientes conjuntos de funciones, determinar si es linealmente dependiente o independiente en el intervalo indicado.

a) $A = \{x^2 - 3x + 1,\ \operatorname{sen}^2 x,\ 2x + \cos^2 x\}$ en $(-\infty, +\infty)$

b) $B = \{x - 2,\ x^2 + 5,\ x^2 - x + 7\}$ en $(-1, 3)$

c) $C = \{2 \operatorname{sen}^2 x,\ 3,\ -5 \cos^2 x\}$ en $(5, \infty)$

d) $D = \{f, g, h\}$ en $(0, 5)$

donde:

$$f(x) = \begin{cases} x^2 & \forall\ x < 1 \\ 1 & \forall\ x \geqslant 1 \end{cases}$$

$$g(x) = \begin{cases} x & \forall\ x < 2 \\ \operatorname{sen}^2 x & \forall\ x \geqslant 2 \end{cases}$$

$$h(x) = \begin{cases} x^2 + x + 2 & \forall\ x < 4 \\ \cos^2 x & \forall\ x \geqslant 4 \end{cases}$$

4.- Demostrar que si $\{f_1, f_2, \ldots, f_n\}$ es linealmente dependiente en (a,b), entonces $W(x) = 0 \quad \forall\, x \in (a,b)$

5.- Demostrar que si f y g son dos funciones derivables en I y $g(x) \neq 0 \quad \forall\, x \in I$, entonces

$$W(x) = 0 \quad \forall\, x \in I \Longrightarrow \{f,g\} \text{ linealmente dependiente en } I.$$

(Sugerencia: calcular $\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right)$)

6.- Demostrar que los conjuntos

$$\{1, x, x^2, \ldots, x^n\}$$

y

$$\{e^x, e^{2x}, \ldots, e^{nx}\}$$

son linealmente independientes para cualquier $n \in N$.

7.- Hallar el subespacio generado por

$$S = \{e^{4x} + \text{sen } x,\ e^{4x} - 3 \text{ sen } x,\ e^{4x}\}$$

y dar una base y la dimensión de dicho subespacio.

8.- Si S es el subconjunto de F tal que

$$S = \{f \mid f'' - 4f = \theta\}\ , \text{ donde } f''(x) = \frac{d^2 f(x)}{dx^2}$$

a) Demostrar que S es un espacio vectorial sobre R.

b) Si $f(x) = e^{2x}$ y $g(x) = e^{-2x}$, demostrar que $\{f,g\}$ es linealmente independiente.

c) Sabiendo que $\dim S = 2$, hallar el conjunto de soluciones de la ecuación $y'' - 4y = 0$.

IX.5 ESPACIOS CON PRODUCTO INTERNO

Como hemos visto, los elementos fundamentales que constituyen un espacio vectorial son: un conjunto de vectores, un conjunto de escalares y dos operaciones, llamadas adición y multiplicación por un escalar.

Tanto la definición de espacio vectorial como sus consecuencias inmediatas (teoremas) se refieren a propiedades que podríamos denominar "algebraicas", puesto que aluden al comportamiento de vectores y escalares respecto a dichas operaciones. Aun las nociones de independencia lineal, base y dimensión pueden considerarse de naturaleza algebraica.

De esta manera, no han sido aún considerados conceptos tales como magnitud, distancia y ángulo, llamados conceptos "métricos", los cuales se refieren a propiedades que pueden ser medidas.

Existen diversas formas de introducir en un espacio vectorial dichos conceptos. Una de ellas, la que aquí adoptaremos, consiste en definirlos a partir de una operación conocida como producto interno.

- Producto interno.

Puede decirse que el concepto de producto interno es la generalización, a un espacio vectorial cualquiera, del producto escalar en R^n†. Cuatro de sus propiedades fundamentales se emplean como los axiomas que debe satisfacer una función dada para ser considerada producto interno, como se enuncia a continuación.

†*También conocido como "producto punto".*

IX.5.1 DEFINICION

Sea V un espacio vectorial sobre C. Un producto interno en V es una función de V×V en C que asigna a cada pareja ordenada $(\bar{u},\bar{v})$ de vectores de V un escalar $(\bar{u}|\bar{v}) \in C$, llamado el producto de $\bar{u}$ y $\bar{v}$, que satisface las siguientes propiedades

i) $(\bar{u}|\bar{v}) = \overline{(\bar{v}|\bar{u})}$

ii) $(\bar{u}|\bar{v}+\bar{w}) = (\bar{u}|\bar{v}) + (\bar{u}|\bar{w})$

iii) $(\alpha\bar{u}|\bar{v}) = \alpha(\bar{u}|\bar{v})$

iv) $(\bar{u}|\bar{u}) > 0$ si $\bar{u} \neq \bar{0}$

En la definición anterior el símbolo $\overline{(\bar{v}|\bar{u})}$ representa al conjugado del número complejo $(\bar{v}|\bar{u})$.

La barra horizontal tendrá entonces dos significados: el de conjugado y el de vector. Obviamente tendrá el primer significado cuando se emplee sobre símbolos que representen escalares y el segundo sobre símbolos que representen vectores.

Cuando la función $(\cdot|\cdot)$ de la definición IX.5.1 es de V×V en R se dice que es un producto interno real.

En este caso la propiedad i) es simplemente

$$(\bar{u}|\bar{v}) = (\bar{v}|\bar{u})$$

llamada "conmutatividad" del producto interno.

En este caso también, el campo sobre el que se define el espacio V es el de los números reales.

Es claro, sin embargo, que cualquier enunciado que se demuestre para un espacio vectorial sobre C con producto interno complejo será válido también para un espacio vectorial sobre R con producto interno real.

El ejemplo más conocido de producto interno es, seguramente, el producto escalar en R^n, el cual está definido por

$$(\overline{x}|\overline{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

donde $\overline{x} = (x_1, x_2, \ldots, x_n)$, $\overline{y} = (y_1, y_2, \ldots, y_n) \in R^n$.

Es decir

$$(\overline{x}|\overline{y}) = \overline{x}\cdot\overline{y}$$

El producto escalar en R^n satisface las condiciones de la definición IX.5.1, puesto que, además de ser una función de $R^n \times R^n$ en R (producto interno real), los postulados i) a iv) son equivalentes a las siguientes expresiones

$$\forall\ \overline{x}, \overline{y}, \overline{z} \in R^n \quad y \quad \alpha \in R$$

i) $\overline{x}\cdot\overline{y} = \overline{y}\cdot\overline{x}$

ii) $\overline{x}\cdot(\overline{y}+\overline{z}) = (\overline{x}\cdot\overline{y}) \quad (\overline{x}\cdot\overline{z})$

iii) $(\alpha\overline{x})\cdot\overline{y} = \alpha(\overline{x}\cdot\overline{y})$

iv) $\overline{x}\cdot\overline{x} > 0$, si $\overline{x} \neq \overline{0}$

las cuales corresponden a conocidas propiedades del producto escalar, como seguramente recordará el lector.

En particular, en R^2 dicho producto interno queda definido por

$$(\overline{x}|\overline{y}) = x_1y_1 + x_2y_2 \qquad (I)$$

donde $\overline{x} = (x_1, x_2)$, $\overline{y} = (y_1, y_2)$.

En este mismo espacio pueden definirse otros productos internos, como por ejemplo

$$(\bar{x}|\bar{y}) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 \qquad \text{(II)}$$

donde $\bar{x} = (x_1, x_2)$, $\bar{y} = (y_1, y_2)$.

En efecto, para la función $(\cdot|\cdot)$ definida por la expresión (II) tenemos:

i) $$\begin{aligned}(\bar{y}|\bar{x}) &= 2y_1x_1 + y_1x_2 + y_2x_1 + y_2x_2\\ &= 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2\\ (\bar{y}|\bar{x}) &= (\bar{x}|\bar{y}).\end{aligned}$$

ii) $$\begin{aligned}(\bar{x}|\bar{y}+\bar{z}) &= 2x_1(y_1+z_1)+x_1(y_2+z_2)+x_2(y_1+z_1)+x_2(y_2+z_2)\\ &= 2x_1y_1+2x_1z_1+x_1y_2+x_1z_2+x_2y_1+x_2z_1+x_2y_2+x_2z_2\\ &= (2x_1y_1+x_1y_2+x_2y_1+x_2y_2)+(2x_1z_1+x_1z_2+x_2z_1+x_2z_2)\\ (\bar{x}|\bar{y}+z) &= (\bar{x}|\bar{y}) + (\bar{x}|\bar{z}).\end{aligned}$$

iii) $$\begin{aligned}(\alpha\bar{x}|\bar{y}) &= 2\alpha x_1y_1 + \alpha x_1y_2 + \alpha x_2y_1 + \alpha x_2y_2\\ &= \alpha(2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2)\\ (\alpha\bar{x}|\bar{y}) &= \alpha(\bar{x}|\bar{y}).\end{aligned}$$

iv) $$\begin{aligned}(\bar{x}|\bar{x}) &= 2x_1^2 + x_1x_2 + x_2x_1 + x_2^2\\ &= x_1^2 + (x_1^2 + 2x_1x_2 + x_2^2)\\ (\bar{x}|\bar{x}) &= x_1^2 + (x_1 + x_2)^2 > 0, \quad \text{si} \quad \bar{x} \neq (0,0).\end{aligned}$$

por lo que, de IX.5.1, dicha función es también un producto interno en R^2.

Con este ejemplo se muestra que puede haber más de un producto interno en un espacio vectorial.

Otros ejemplos de producto interno son los siguientes:

1) El producto escalar en C^n, o producto interno usual, definido por

$$(\bar{z}|\bar{w}) = z_1\bar{w}_1 + z_2\bar{w}_2 + \ldots + z_n\bar{w}_n$$

donde $\bar{z} = (z_1, z_2, \ldots, z_n)$, $\bar{w} = (w_1, w_2, \ldots, w_n)$ y $\bar{w}_i$ es el conjugado de w_i.

2) En el espacio V de las funciones reales de variable real, continuas en el intervalo (a,b), la función $(\cdot|\cdot)$ definida por

$$(f|g) = \int_a^b f(x)\ g(x)\ dx$$

es un producto interno real.

3) En el espacio V de las matrices de m×n con elementos en R, el siguiente es un producto interno real

$$(A|B) = tr(A^T B)$$

4) En el espacio V de matrices de m×n con elementos en C, la función de V×V en C definida por

$$(A|B) = tr(A^* B)$$

es un producto interno.

Como consecuencia de la definición IX.5.1, cualquier producto interno tiene las propiedades que se enuncian a continuación.

IX.5.2 TEOREMA

Sea V un espacio vectorial sobre C y sea $(\cdot|\cdot)$ un producto interno en V; entonces, $\forall\ \bar{u}, \bar{v} \in V$ y $\alpha \in C$:

i) $(\bar{u}|\alpha\bar{v}) = \bar{\alpha}(\bar{u}|\bar{v})$

ii) $(\bar{u}|\bar{u}) \in R$

iii) $(\bar{0}|\bar{u}) = 0 = (\bar{u}|\bar{0})$

iv) $(\bar{u}|\bar{u}) = 0 \iff \bar{u} = \bar{0}$

DEMOSTRACION

i) $(\bar{u}|\alpha\bar{v}) = \overline{(\alpha\bar{v}|\bar{u})}$ por i) de IX.5.1

$= \overline{\alpha(\bar{v}|\bar{u})}$ por iii) de IX.5.1

$= \bar{\alpha}\overline{(\bar{v}|\bar{u})}$ por vi) de II.1.6

$(\bar{u}|\alpha\bar{v}) = \bar{\alpha}(\bar{u}|\bar{v})$ por i) de IX.5.1

lo que demuestra la proposición i).

Si $(\cdot|\cdot)$ es un producto interno real, dicha proposición es simplemente

$$(\bar{u}|\alpha\bar{v}) = \alpha(\bar{u}|\bar{v})$$

ii) Por i) de IX.5.1 se tiene que

$$(\bar{u}|\bar{u}) = \overline{(\bar{u}|\bar{u})}$$

En consecuencia, por ii) de II.1.6

$$(\bar{u}|\bar{u}) \in R.$$

Esta propiedad nos indica que $(\bar{u}|\bar{u})$ es siempre un número real, aun cuando $(\cdot|\cdot)$ sea un producto interno complejo.

iii) $(\bar{0}|\bar{u}) = (0\bar{v}|\bar{u})$ para cualquier vector $\bar{v} \varepsilon V$, por ii) de IX.1.3 En consecuencia, por iii) de IX.5.1

$$(\bar{0}|\bar{u}) = 0(\bar{v}|\bar{u})$$

por lo que

$$(\bar{0}|\bar{u}) = 0.$$

Además, por i) de IX.5.1 y tomando en cuenta que el conjugado de 0 es 0, de la expresión anterior se tiene que

$$(\bar{u}|\bar{0}) = 0.$$

lo que completa la demostración de la propiedad iii).

iv) Del resultado anterior se sigue que

$$\bar{u} = \bar{0} \implies (\bar{u}|\bar{u}) = 0$$

y en consecuencia

$$(\bar{u}|\bar{u}) = 0 \quad \text{si y sólo si} \quad \bar{u} = \bar{0}$$

□

IX.5.3 TEOREMA (Desigualdad de Cauchy-Schwarz)

Sea V un espacio vectorial sobre C y sea $(\cdot|\cdot)$ un producto interno en V; entonces, $\forall\ \bar{u}, \bar{v}\ \varepsilon\ V$:

$$|(\bar{u}|\bar{v})|^2 \leqslant (\bar{u}|\bar{u})(\bar{v}|\bar{v})$$

donde $|(\bar{u}|\bar{v})|$ es el módulo de $(\bar{u}|\bar{v})$.

Además, la igualdad se cumple si y sólo si $\bar{u}$ y $\bar{v}$ son linealmente dependientes.

DEMOSTRACION

a) Supongamos primero que $\overline{u}$ y $\overline{v}$ son linealmente independientes; esto es, $\overline{u} \neq \overline{0}$, $\overline{v} \neq \overline{0}$ y además

$$\overline{u} \neq -\lambda\overline{v}, \ \forall \ \lambda \in C.$$

entonces

$$\overline{u} + \lambda\overline{v} \neq \overline{0}, \ \forall \ \lambda \in C.$$

y por iv) de IX.5.1

$$(\overline{u}+\lambda\overline{v}|\overline{u}+\lambda\overline{v}) > 0, \ \forall \ \lambda \in C \qquad (1)$$

Por otra parte

$$\begin{aligned}
(\overline{u}+\lambda\overline{v}|\overline{u}+\lambda\overline{v}) &= (\overline{u}+\lambda\overline{v}|\overline{u})+(\overline{u}+\lambda\overline{v}|\lambda\overline{v}) && \text{por ii) de IX.5.1}\\
&= \overline{(\overline{u}|\overline{u}+\lambda\overline{v})}+\overline{(\lambda\overline{v}|\overline{u}+\lambda\overline{v})} && \text{por i) de IX.5.1}\\
&= \overline{(\overline{u}|\overline{u})+(\overline{u}|\lambda\overline{v})}+\overline{(\lambda\overline{v}|\overline{u})+(\lambda\overline{v}|\lambda\overline{v})} && \text{por ii) de X.5.1}\\
&= \overline{(\overline{u}|\overline{u})}+\overline{(\overline{u}|\lambda\overline{v})}+\overline{(\lambda\overline{v}|\overline{u})}+\overline{(\lambda\overline{v}|\lambda\overline{v})} && \text{por v) de II.1.6}\\
&= (\overline{u}|\overline{u})+(\lambda\overline{v}|\overline{u})+(\overline{u}|\lambda\overline{v})+(\lambda\overline{v}|\lambda\overline{v}) && \text{por i) de IX.5.1}\\
&= (\overline{u}|\overline{u})+\lambda(\overline{v}|\overline{u})+(\overline{u}|\lambda\overline{v})+\lambda(\overline{v}|\lambda\overline{v}) && \text{por iii) de IX.5.1}\\
(\overline{u}+\lambda\overline{v}|\overline{u}+\lambda\overline{v}) &= (\overline{u}|\overline{u})+\lambda(\overline{v}|\overline{u})+\overline{\lambda}(\overline{u}|\overline{v})+\lambda\overline{\lambda}(\overline{v}|\overline{v}) && \text{por i) de IX.5.2}
\end{aligned}$$

llevando esta última expresión a (1) se tiene que

$$(\overline{u}|\overline{u}) + \lambda(\overline{v}|\overline{u}) + \overline{\lambda}(\overline{u}|\overline{v}) + \lambda\overline{\lambda}(\overline{v}|\overline{v}) > 0, \ \forall \ \lambda \in C$$

En particular, si $\lambda = -\dfrac{(\overline{u}|\overline{v})}{(\overline{v}|\overline{v})}$ se tendrá que

$$(\overline{u}|\overline{u}) - \frac{(\overline{u}|\overline{v})(\overline{v}|\overline{u})}{(\overline{v}|\overline{v})} > 0$$

esto es

$$(\overline{u}|\overline{u})(\overline{v}|\overline{v}) > (\overline{u}|\overline{v})(\overline{v}|\overline{u})$$

$$(\overline{u}|\overline{u})(\overline{v}|\overline{v}) > (\overline{u}|\overline{v})\overline{(\overline{u}|\overline{v})}$$

$$(\overline{u}|\overline{u})(\overline{v}|\overline{v}) > |(\overline{u}|\overline{v})|^2$$

b) Supongamos ahora que $\overline{u}$ y $\overline{v}$ son linealmente dependientes. Si $\overline{u} = \overline{0}$ o $\overline{v} = \overline{0}$ es inmediato que la igualdad se verifica. Si ambos son vectores no nulos, entonces

$$\overline{u} = \alpha\overline{v}$$

para algún $\alpha \neq 0$, por lo que

$$|(\overline{u}|\overline{v})| = |(\alpha\overline{v}|\overline{v})|$$

y en consecuencia

$$|(\overline{u}|\overline{v})| = |\alpha(\overline{v}|\overline{v})| \qquad \text{por iii) de IX.5.1}$$

$$= |\alpha|\,|(\overline{v}|\overline{v})| \qquad \text{por i) de II.2.3}$$

$$|(\overline{u}|\overline{v})| = |\alpha|(\overline{v}|\overline{v}) \qquad \text{por iv) de IX.5.1}$$

de donde

$$|(\overline{u}|\overline{v})|^2 = |\alpha|^2 (\overline{v}|\overline{v})(\overline{v}|\overline{v}) \qquad (2)$$

pero, como $\overline{u} = \alpha\overline{v}$, se tiene

$$(\bar{u}|\bar{u}) = (\alpha\bar{v}|\alpha\bar{v})$$

$$= \alpha\bar{\alpha}(\bar{v}|\bar{v})$$

$$(\bar{u}|\bar{u}) = |\alpha|^2 \; (\bar{v}|\bar{v})$$

por lo que, sustituyendo en (2) se llega a

$$|(\bar{u}|\bar{v})|^2 = (\bar{u}|\bar{u})\,(\bar{v}|\bar{v})$$

y la demostración queda completa.

□

A los espacios vectoriales dotados de producto interno se les conoce frecuentemente como "espacios euclidianos"; aunque algunos autores reservan este término para los espacios con producto interno real, distinguiéndolos de los espacios con producto interno complejo a los que llaman "espacios unitarios".

- Norma, distancia y ángulo.

La idea de magnitud (o tamaño) de un vector se introduce formalmente en un espacio vectorial con el concepto de norma; ésta se define a partir del producto interno como sigue.

IX.5.4 DEFINICION

Sea V un espacio vectorial sobre C y sea $(\cdot|\cdot)$ un producto interno en V. Se llama norma de $\bar{v} \in V$, y se representa con $\|\bar{v}\|$, al número real no negativo definido por

$$\|\bar{v}\| = (\bar{v}|\bar{v})^{1/2}$$

De esta manera, la norma es una función de V en el conjunto de los números reales no negativos.

Empleando el concepto de norma, la desigualdad de Cauchy-Schwarz puede expresarse como

$$|(\bar{u}|\bar{v})|^2 \leqslant \|\bar{u}\|^2 \ \|\bar{v}\|^2$$

o bien

$$|(\bar{u}|\bar{v})| \leqslant \|\bar{u}\| \ \|\bar{v}\|$$

donde se nota que, para cualquier par de vectores, el "tamaño" del producto es menor o igual que el producto de los tamaños.

Cabe hacer notar que, en la expresión anterior, el producto del lado izquierdo es un producto de vectores; mientras que el del lado derecho es un producto de números reales.

Como se sigue de la definición IX.5.4, la norma de un vector depende del producto interno que se haya elegido.

Por ejemplo, para el vector $\bar{v} = (4, -3) \in R^2$ se tiene, utilizando el producto interno usual en R^2

$$\|\bar{v}\| = (16 + 9)^{1/2} = \sqrt{25} = 5.$$

mientras que, utilizando el producto internc definido por la expresión $(\bar{x}|\bar{y}) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$

se tiene

$$\|\bar{v}\| = (32 - 12 - 12 + 9)^{1/2} = \sqrt{17}$$

Así, un mismo vector puede tener diferentes normas.

Esta situación, aparentemente contradictoria, es similar al hecho de asignar diferentes números a una misma magnitud, dependiendo de la escala o unidad de medida que se emplee.

En el siguiente teorema se enuncian las propiedades fundamentales que satisface toda norma, independientemente del producto interno del que provenga.

IX.5.5 TEOREMA

Si V es un espacio vectorial con producto interno, entonces $\forall\ \bar{u}, \bar{v} \in V$ y $\alpha \in C$:

i) $\|\bar{v}\| \geq 0$

ii) $\|\bar{v}\| = 0$ si y sólo si $\bar{v} = \bar{0}$

iii) $\|\alpha\bar{v}\| = |\alpha|\ \|\bar{v}\|$

iv) $\|\bar{u} + \bar{v}\| \leq \|\bar{u}\| + \|\bar{v}\|$

DEMOSTRACION

Las primeras tres propiedades se deducen inmediatamente de la definición de norma y de los axiomas que definen un producto interno. La propiedad iv), conocida como desigualdad del triángulo, se demuestra a continuación:

$$\|\bar{u}+\bar{v}\|^2 = (\bar{u}+\bar{v}\,|\,\bar{u}+\bar{v}) \qquad \text{por IX.5.4}$$

$$= (\bar{u}+\bar{v}\,|\,\bar{u}) + (\bar{u}+\bar{v}\,|\,\bar{v}) \qquad \text{por ii) de IX.5.1}$$

$$= \overline{(\bar{u}\,|\,\bar{u}+\bar{v})} + \overline{(\bar{v}\,|\,\bar{u}+\bar{v})} \qquad \text{por i) de IX.5.1}$$

$$= \overline{(\bar{u}\,|\,\bar{u}) + (\bar{u}\,|\,\bar{v})} + \overline{(\bar{v}\,|\,\bar{u}) + (\bar{v}\,|\,\bar{v})} \qquad \text{por ii) de IX.5.1}$$

$$= \overline{(\bar{u}\,|\,\bar{u})} + \overline{(\bar{u}\,|\,\bar{v})} + \overline{(\bar{v}\,|\,\bar{u})} + \overline{(\bar{v}\,|\,\bar{v})} \qquad \text{por v) de II.1.6}$$

$$= (\bar{u}\,|\,\bar{u}) + \overline{(\bar{u}\,|\,\bar{v})} + (\bar{u}\,|\,\bar{v}) + (\bar{v}\,|\,\bar{v}) \qquad \text{por i) de IX.5.1}$$

$$\|\bar{u}+\bar{v}\|^2 = \|\bar{u}\|^2 + \overline{(\bar{u}\,|\,\bar{v})} + (\bar{u}\,|\,\bar{v}) + \|\bar{v}\|^2 \qquad \text{por IX.5.4}$$

Por otra parte, como el lector puede demostrar

$$\overline{z} + z \leqslant 2\ |z|,\ \forall\ z\ \varepsilon\ C.$$

por lo que

$$\overline{(\bar{u}|\bar{v})} + (\bar{u}|\bar{v}) \leqslant 2\ |(\bar{u}|\bar{v})|$$

y por la desigualdad de Cauchy Schwarz

$$\overline{(\bar{u}|\bar{v})} + (\bar{u}|\bar{v}) \leqslant 2\ \|\bar{u}\|\ \ \|\bar{v}\|$$

En consecuencia

$$\|\bar{u} + \bar{v}\|^2 \leqslant \|\bar{u}\|^2 + 2\ \|\bar{u}\|\ \ \|\bar{v}\| + \|\bar{v}\|^2$$

$$\|\bar{u} + \bar{v}\|^2 \leqslant (\|\bar{u}\| + \|\bar{v}\|)^2$$

de donde

$$\|\bar{u} + \bar{v}\| \leqslant \|\bar{u}\| + \|\bar{v}\|$$

Como se quería.

$\square$

Se dice que un vector $\bar{v}$ es unitario cuando $\|\bar{v}\| = 1$.

El lector puede demostrar fácilmente que para cualquier vector $\bar{v}$ de un espacio con producto interno, el vector

$$\left(\frac{1}{\|\bar{v}\|}\right)\bar{v}$$

es un vector unitario.

Empleando el concepto de norma, podemos introducir en un espacio vectorial el concepto de distancia entre vectores de la siguiente manera.

IX.5.6 DEFINICION

Sea V un espacio vectorial con producto interno, y sean $\bar{u}$, $\bar{v} \in V$. Se llama distancia de $\bar{u}$ a $\bar{v}$, y se representa con $d(\bar{u},\bar{v})$, al número real definido por

$$d(\bar{u},\bar{v}) = ||\bar{v} - \bar{u}||$$

De esta manera, la distancia es una función de V×V en el conjunto de los números reales no negativos y, como puede demostrarse, tiene las propiedades que se enuncian a continuación.

IX.5.7 TEOREMA

Si V es un espacio vectorial con producto interno, entonces $\forall\ \bar{u}, \bar{v}, \bar{w} \in V$:

i) $d(\bar{u},\bar{v}) \geq 0$

ii) $d(\bar{u},\bar{v}) = 0$ si y sólo si $\bar{u} = \bar{v}$

iii) $d(\bar{u},\bar{v}) = d(\bar{v},\bar{u})$

iv) $d(\bar{u},\bar{w}) \leq d(\bar{u},\bar{v}) + d(\bar{v},\bar{w})$

En los espacios con producto interno real se acostumbra definir el ángulo entre dos vectores de la siguiente manera

IX.5.8 DEFINICION

Sea V un espacio vectorial con producto interno real, y sean $\bar{u}$, $\bar{v}$ dos vectores no nulos de V. Se llama ángulo entre $\bar{u}$ y $\bar{v}$ al número real θ, en el intervalo $0 \leqslant \theta \leqslant \pi$, tal que

$$\cos \theta = \frac{(\bar{u}|\bar{v})}{\|\bar{u}\| \, \|\bar{v}\|}$$

De la desigualdad de Cauchy - Schwarz se sigue que, si $\bar{u}, \bar{v} \neq \bar{0}$ entonces

$$-1 \leqslant \frac{(\bar{u}|\bar{v})}{\|\bar{u}\| \, \|\bar{v}\|} \leqslant 1$$

por lo que existe un número real θ (y sólo uno), en el intervalo $0 \leqslant \theta \leqslant \pi$, cuyo coseno es igual al mencionado cociente.

En ocasiones se generaliza el concepto de ángulo entre vectores al caso de espacios con producto interno complejo, reemplazando la expresión de la definición IX.5.8 por

$$\cos \theta = \frac{R(\bar{u}|\bar{v})}{\|\bar{u}\| \, \|\bar{v}\|}$$

donde $R(\bar{u}|\bar{v})$ representa la parte real de $(\bar{u}|\bar{v})$.

Es claro que la distancia y el ángulo entre vectores dependen también del producto interno elegido.

Así, por ejemplo, para los vectores

$$\bar{u} = (4, -3) \quad y \quad \bar{v} = (0, 2)$$

de R^2 se tiene, con el producto interno usual

$$d(\overline{u},\overline{v}) = \|(-4,\ 5)\| = (16 + 25)^{1/2} = \sqrt{41} \doteq 6.40$$

y

$$\cos\theta = \frac{0\ -6}{(5)(2)} = \frac{-3}{5} \therefore \theta = \text{ang cos } \left(\frac{-3}{5}\right) \doteq 2.21$$

mientras que, con el producto interno definido por

$$(\overline{x}|\overline{y}) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$$

se tiene, para estos mismos vectores

$$d(\overline{u},\overline{v}) = \|(-4,\ 5)\| = (32 - 40 + 25)^{1/2} = \sqrt{17} \doteq 4.12$$

y

$$\cos\theta = \frac{0 + 8 + 0 - 6}{(\sqrt{17})(\sqrt{4})} = \frac{1}{\sqrt{17}} \therefore \theta = \text{ang cos } \frac{1}{\sqrt{17}} \doteq 1.33$$

Los conceptos de norma, distancia y ángulo que se han definido en esta sección constituyen una generalización, a espacios vectoriales cualesquiera, de sus respectivas ideas geométricas.

Así, por ejemplo, podemos hablar de la distancia entre las matrices

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

o del ángulo entre las funciones

$$f(x) = x \quad \text{y} \quad g(x) = e^x$$

En efecto, de IX.5.6 y IX.5.4

$$d(A,B) = \|B-A\| = (B-A|B-A)^{1/2}$$

y para las matrices dadas

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

por lo que, empleando el producto interno del ejemplo 3)

$$(B-A|B-A) = tr\left[(B-A)^T(B-A)\right] = tr\begin{bmatrix} 20 & 8 & 10 \\ 8 & 4 & 2 \\ 10 & 2 & 10 \end{bmatrix} = 34$$

y finalmente

$$d(A,B) = \sqrt{34} \doteq 5.83$$

Para las funciones f y g definidas anteriormente tenemos

$$\cos\theta = \frac{(f|g)}{||f||\ ||g||}$$

por lo que, usando el producto interno del ejemplo 2) en el intervalo (0,1) tenemos

$$(f|g) = \int_0^1 xe^x dx = \left[xe^x - e^x\right]_0^1 = 1$$

$$||f||^2 = (f|f) = \int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$$

$$||g||^2 = (g|g) = \int_0^1 e^{2x} dx = \left[\frac{1}{2}e^{2x}\right]_0^1 = \frac{e^2-1}{2}$$

$$\cos\theta = \frac{1}{\left(\frac{1}{\sqrt{3}}\right)\sqrt{\frac{e^2-1}{2}}} = \sqrt{\frac{6}{e^2-1}} \quad \therefore \quad \theta \doteq 0.25$$

- Ortogonalidad

También el concepto geométrico de ortogonalidad, o perpendicularidad, tiene una generalización a espacios cualesquiera con producto interno.

Como el lector recordará, en el espacio cartesiano dos vectores son ortogonales cuando su producto escalar es igual a cero, puesto que tal condición equivale a que los segmentos dirigidos formen un ángulo de 90°. Así, la noción de ortogonalidad se generaliza a un espacio cualquiera con producto interno de la siguiente manera.

IX.5.9 DEFINICION

Sea V un espacio vectorial con producto interno. Dos vectores $\bar{u}$, $\bar{v} \in V$ son ortogonales si $(\bar{u}|\bar{v}) = 0$

En particular, el vector cero es ortogonal a cualquier vector del espacio. De hecho, es el único vector con tal propiedad, como el lector podrá demostrar.

Observe que, de acuerdo a la definición anterior, dos vectores ortogonales no nulos forman un ángulo de 90°; sin embargo, el recíproco de este enunciado no es válido en un espacio con producto interno complejo.

En efecto, si $\bar{u}$, $\bar{v}$ son ortogonales, por IX.5.9

$$(\bar{u}|\bar{v}) = \bar{0}$$

además, como $\bar{u}, \bar{v} \neq \bar{0}$, por ii) de IX.5.5, $||\bar{u}|| \neq 0$ y $||\bar{v}|| \neq 0$; por lo que, de IX.5.3

$$\cos\theta = \frac{0}{||\bar{u}||\,||\bar{v}||} = 0, \text{ y en consecuencia } \theta = \frac{\pi}{2}$$

esto es, los vectores forman un ángulo de 90°.

Supongamos ahora que $\bar{u}$ y $\bar{v}$ forman un ángulo de 90°, entonces

$$\theta = \frac{\pi}{2} \qquad \text{y} \qquad \cos\theta = 0$$

por lo que

$$\frac{R(\bar{u}|\bar{v})}{||\bar{u}||\,||\bar{v}||} = 0$$

de donde

$$R(\bar{u}|\bar{v}) = 0$$

es decir, la parte real de $(\bar{u}|\bar{v})$ es igual a cero. Esto no implica que $\bar{u}$ y $\bar{v}$ sean ortogonales.

Uno de los resultados más importantes relacionado con la ortogonalidad de dos vectores es la generalización del llamado "Teorema de Pitágoras", la cual se enuncia a continuación y cuya demostración se deja al lector.

IX.5.10 TEOREMA (DE PITAGORAS)

Sea V un espacio con producto interno y sean $\bar{u}, \bar{v} \in V$. Si $\bar{u}$ y $\bar{v}$ son ortogonales entonces

$$||\bar{u}+\bar{v}||^2 = ||\bar{u}||^2 + ||\bar{v}||^2$$

El concepto de ortogonalidad se emplea también para conjuntos de vectores.

Se considera que un conjunto es ortogonal cuando cada uno de sus vectores es ortogonal a los demás elementos del conjunto, como lo establece la siguiente definición

IX.5.11 DEFINICION

Sea V un espacio con producto interno y sea $S = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ un conjunto de vectores de V. Se dice que S es un conjunto ortogonal cuando

$$(\overline{v}_i | \overline{v}_j) = 0, \ \forall\, i \neq j$$

si además $\|\overline{v}_i\| = 1$, $\forall\, i$, el conjunto S es ortonormal.

Los conjuntos ortogonales, y en particular los ortonormales, tienen propiedades importantes, entre las cuales se encuentra la siguiente.

IX.5.12 TEOREMA

Un conjunto ortogonal de vectores no nulos es linealmente independiente.

DEMOSTRACION

Sea $S = \{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n\}$ un conjunto ortogonal de vectores no nulos, y consideremos la expresión

$$\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n = \overline{0}$$

Multiplicando ambos miembros por un vector cualquiera $\overline{v}_i \in S$ tenemos

$$(\overline{v}_i | \alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n) = (\overline{v}_i | \overline{0})$$

Entonces, por iii) de IX.5.2

$$(\overline{v}_i | \alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n) = 0$$

Aplicando reiteradamente ii) de IX.5.1 y i) de IX.5.2, llegamos a

$$\bar{\alpha}_1(\bar{v}_i|\bar{v}_1) + \bar{\alpha}_2(\bar{v}_i|\bar{v}_2) + \ldots + \bar{\alpha}_n(\bar{v}_i|\bar{v}_n) = 0$$

Como S es un conjunto ortogonal, todos los productos $(\bar{v}_i|\bar{v}_j)$ son iguales a cero para $i \neq j$; por lo que la expresión anterior se reduce a

$$\bar{\alpha}_i(\bar{v}_i|\bar{v}_i) = 0$$

y como $\bar{v}_i \neq \bar{0}$, por iv) de IX.5.1, $(\bar{v}_i|\bar{v}_i) > 0$, y en consecuencia

$$\bar{\alpha}_i = 0$$

de donde

$$\alpha_i = 0$$

Como lo anterior sucede para cualquier valor de i; por IX.2.3, S es linealmente independiente.

□

En muchos problemas relacionados con espacios vectoriales la elección de una base adecuada puede simplificar la solución de éstos. Generalmente la mejor elección es una base de tipo ortonormal.

Las coordenadas de un vector en una base ortonormal tienen una expresión muy simple en términos del producto de dicho vector por cada uno de los elementos de la base, como lo establece el siguiente enunciado.

IX.5.13 TEOREMA

Sea V un espacio con producto interno y sea $B = \{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$ una base ortogonal. Entonces, $\forall\ \bar{v} \in V$

$$(\bar{v})_B = (\alpha_i)\ , \text{ donde } \alpha_i = \frac{(\bar{v}|\bar{e}_i)}{(\bar{e}_i|\bar{e}_i)}$$

En particular, si B es una base ortonormal

$$\alpha_i = (\bar{v}|\bar{e}_i)$$

DEMOSTRACION

Sea

$$\bar{v} = \sum_{j=1}^{n} \alpha_j \bar{e}_j$$

entonces, multiplicando por un vector $\bar{e}_i$ arbitrario de la base B

$$(\bar{e}_i|\bar{v}) = (\bar{e}_i|\sum_{j=1}^{n} \alpha_j \bar{e}_j) = \sum_{j=1}^{n} (\bar{e}_i|\alpha_j \bar{e}_j) = \sum_{j=1}^{n} \bar{\alpha}_j (\bar{e}_i|\bar{e}_j)$$

pero como B es ortogonal, $(\bar{e}_i|\bar{e}_j) = 0$ para $i \neq j$; por lo que

$$(\bar{e}_i|\bar{v}) = \bar{\alpha}_i (\bar{e}_i|\bar{e}_i)$$

y en consecuencia

$$(\bar{v}|\bar{e}_i) = \overline{\bar{\alpha}_i (\bar{e}_i|\bar{e}_i)} = \alpha_i (\bar{e}_i|\bar{e}_i)$$

de donde

$$\alpha_i = \frac{(\bar{v}|\bar{e}_i)}{(\bar{e}_i|\bar{e}_i)}$$

como se quería.

Además, si B es ortonormal $(\bar{e}_i|\bar{e}_i) = \|\bar{e}_i\|^2 = 1$; por lo que

$$\alpha_i = (\overline{v}|\overline{e}_i)$$

lo cual completa la demostración.

$\square$

El concepto de ortogonalidad se generaliza también a conjuntos infinitos.

Un ejemplo muy conocido lo encontramos en la sucesión (infinita) de funciones

$$\{1, \cos x, \cos 2x, \ldots, \cos nx, \ldots\}$$

que juega un papel importante en la teoría y aplicaciones de las series de Fourier.

Esta sucesión constituye un conjunto ortogonal en el espacio de las funciones continuas en el intervalo $(0, 2\pi)$, para el producto interno definido por

$$(f|g) = \int_0^{2\pi} f(x)\ g(x)\ dx$$

puesto que

$$\int_0^{2\pi} \cos mx . \cos nx\ dx = 0, \text{ si } m \neq n$$

Además, ya que

$$\int_0^{2\pi} 1\ dx = 2\pi\ , \quad \int_0^{2\pi} \cos^2 nx\ dx = \pi$$

se tiene que la sucesión

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \ldots$$

es un conjunto ortonormal.

- Proceso de Gram-Schmidt.

Siempre es posible obtener una base ortogonal (y por consiguiente una ortonormal) a partir de una base cualquiera, mediante el llamado "proceso de Gram-Schmidt", el cual se describe a continuación.

A partir de un conjunto

$$G = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$$

que genera a un espacio V, el procedimiento construye un conjunto

$$G_0 = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n\}$$

que es ortogonal y que también genera al espacio V.

La idea básica consiste en formar uno a uno los vectores $\bar{w}_i$ de tal manera que cada nuevo vector sea ortogonal a todos los vectores obtenidos anteriormente.

Para facilitar la comprensión del método y deducir sus expresiones, analicemos primero el caso correspondiente a n = 2.

Sea V un espacio con producto interno y sea $G = \{\bar{v}_1, \bar{v}_2\}$ un generador de V.

El primer vector de G_0 puede ser cualquier vector de G, puesto que no requiere ser ortogonal a ningún otro. Hacemos entonces

$$\bar{w}_1 = \bar{v}_1 \qquad (1)$$

A continuación proponemos un vector $\bar{w}_2$ de la forma

$$\bar{w}_2 = \bar{v}_2 - \alpha_1 \bar{w}_1 \qquad (2)$$

buscando que el conjunto $\{\overline{w}_1, \overline{w}_2\}$ sea tambiên un generador de V y que $\overline{w}_2$ sea ortogonal a $\overline{w}_1$.

La primera condiciôn se cumple para cualquier valor del escalar α_1, puesto que todo vector $\overline{u} \varepsilon V$ es de la forma

$$\overline{u} = \lambda_1\overline{v}_1 + \lambda_2\overline{v}_2$$

pero de las expresiones (1) y (2)

$$\overline{v}_1 = \overline{w}_1$$

y

$$\overline{v}_2 = \overline{w}_2 + \alpha_1\overline{w}_1$$

por lo que

$$\overline{u} = \lambda_1\overline{w}_1 + \lambda_2(\overline{w}_2 + \alpha_1\overline{w}_1)$$

$$\overline{u} = (\lambda_1 + \lambda_2\alpha_1)\ \overline{w}_1 + \lambda_2\overline{w}_2$$

y el vector $\overline{u}$ es una combinaciôn lineal de $\overline{w}_1$ y $\overline{w}_2$.

Por otra parte, la segunda condiciôn se cumple si

$$(\overline{w}_2|\overline{w}_1) = 0$$

esto es, cuando

$$(\overline{v}_2 - \alpha_1\overline{w}_1|\overline{w}_1) = 0$$

Es decir,

$$(\overline{v}_2|\overline{w}_1) - \alpha_1(\overline{w}_1|\overline{w}_1) = 0$$

$$\alpha_1 = \frac{(\overline{v}_2|\overline{w}_1)}{(\overline{w}_1|\overline{w}_1)}$$

En consecuencia, el vector

$$\bar{w}_2 = \bar{v}_2 - \frac{(\bar{v}_2 | \bar{w}_1)}{(\bar{w}_1 | \bar{w}_1)} \bar{w}_1 \qquad (2')$$

satisface las dos condiciones requeridas.

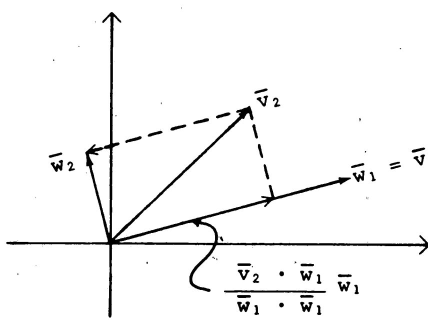
El vector

$$\frac{(\bar{v}_2 | \bar{w}_1)}{(\bar{w}_1 | \bar{w}_1)} \bar{w}_1$$

recibe el nombre de "proyección de $\bar{v}_2$ sobre $\bar{w}_1$", como una generalización del concepto geométrico de proyección asociado a dicha expresión en el caso del producto interno usual en R^2 y en R^3.

De esta manera, el vector $\bar{w}_2$ se obtiene restando a $\bar{v}_2$ su proyección sobre $\bar{w}_1$.

Una interpretación geométrica de lo anterior en el espacio R^2 es la siguiente



En resumen, si V es un espacio con producto interno y $G = \{\bar{v}_1, \bar{v}_2\}$ es un generador de V, entonces el conjunto $G_0 = \{\bar{w}_1, \bar{w}_2\}$, donde

$$\overline{w}_1 = \overline{v}_1 \qquad (1)$$

$$\overline{w}_2 = \overline{v}_2 - \frac{(\overline{v}_2|\overline{w}_1)}{(\overline{w}_1|\overline{w}_1)}\overline{w}_1 \qquad (2')$$

es un generador ortogonal de V.

Consideremos ahora el caso correspondiente a $n = 3$. Es decir, sea V un espacio con producto interno y sea

$$G = \{\overline{v}_1, \overline{v}_2, \overline{v}_3\}$$

un generador de V.

Es claro que para obtener los dos primeros vectores de

$$G_0 = \{\overline{w}_1, \overline{w}_2, \overline{w}_3\}$$

podemos emplear las expresiones (1) y (2').

Entonces, para obtener el tercer vector hacemos

$$\overline{w}_3 = \overline{v}_3 - \beta_1\overline{w}_1 - \beta_2\overline{w}_2 \qquad (3)$$

buscando que el conjunto $\{\overline{w}_1, \overline{w}_2, \overline{w}_3\}$ sea un generador de V y que $\overline{w}_3$ sea ortogonal tanto a $\overline{w}_1$ como a $\overline{w}_2$.

Nuevamente la primera condición se cumple para cualquier valor de β_1 y β_2, puesto que, $\forall\ \overline{u} \in V$

$$\overline{u} = \lambda_1\overline{v}_1 + \lambda_2\overline{v}_2 + \lambda_3\overline{v}_3$$

y de las expresiones (1), (2) y (3)

$$\overline{u} = \lambda_1\overline{w}_1 + \lambda_2(\overline{w}_2 + \alpha_1\overline{w}_1) + \lambda_3(\overline{w}_3 + \beta_1\overline{w}_1 + \beta_2\overline{w}_2)$$

$$\overline{u} = (\lambda_1 + \lambda_2\alpha_1 + \lambda_3\beta_1)\overline{w}_1 + (\lambda_2 + \lambda_3\beta_2)\overline{w}_2 + \lambda_3\overline{w}_3$$

por lo que $\bar{u} \in L(\{\bar{w}_1, \bar{w}_2, \bar{w}_3)\}$.

Además, $\bar{w}_3$ es ortogonal a $\bar{w}_1$ si

$$(\bar{w}_3|\bar{w}_1) = 0$$

Esto es,

$$(\bar{v}_3 - \beta_1\bar{w}_1 - \beta_2\bar{w}_2|\bar{w}_1) = 0$$

$$(\bar{v}_3|\bar{w}_1) - \beta_1(\bar{w}_1|\bar{w}_1) - \beta_2(\bar{w}_2|\bar{w}_1) = 0$$

pero como $(\bar{w}_2|\bar{w}_1) = 0$ se tiene

$$(\bar{v}_3|\bar{w}_1) - \beta_1(\bar{w}_1|\bar{w}_1) = 0$$

esto es

$$\beta_1 = \frac{(\bar{v}_3|\bar{w}_1)}{(\bar{w}_1|\bar{w}_1)}$$

En forma análoga se concluye que $\bar{w}_3$ es ortogonal a $\bar{w}_2$ si

$$\beta_2 = \frac{(\bar{v}_3|\bar{w}_2)}{(\bar{w}_2|\bar{w}_2)}$$

y en consecuencia

$$\bar{w}_3 = \bar{v}_3 - \frac{(\bar{v}_3|\bar{w}_1)}{(\bar{w}_1|\bar{w}_1)}\bar{w}_1 - \frac{(\bar{v}_3|\bar{w}_2)}{(\bar{w}_2|\bar{w}_2)}\bar{w}_2 \qquad (3')$$

Es decir, $\bar{w}_3$ se obtiene restando a $\bar{v}_3$ sus proyecciones sobre $\bar{w}_1$ y $\bar{w}_2$.

Generalizando los resultados anteriores se llega al siguiente enunciado.

IX.5.14 TEOREMA (Proceso de Gram-Schmidt)

Sea V un espacio con producto interno y sea $G = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ un generador de V. El conjunto $G_0 = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n\}$ donde

$$\bar{w}_1 = \bar{v}_1$$

$$\bar{w}_i = \bar{v}_i - \sum_{k=1}^{i-1} \frac{(\bar{v}_i|\bar{w}_k)}{(\bar{w}_k|\bar{w}_k)} \bar{w}_k \text{ , para } i = 2, 3, \ldots, n.$$

es un generador ortogonal de V.

Con relación a este proceso conviene hacer las siguientes observaciones

- Cada uno de los vectores $\bar{w}_i$ (del segundo en adelante) se obtiene restando al correspondiente $\bar{v}_i$ sus proyecciones sobre los vectores $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_{i-1}$, obtenidos previamente.
- En cada paso de este proceso, los vectores $\{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_i\}$ constituyen un generador ortogonal del espacio generado por el conjunto $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_i\}$.
- Si todos los vectores de G_0 son distintos de cero, dicho conjunto es una base de V.
- Si algunos vectores de G_0 son iguales a cero, entonces el subconjunto de G_0 formado por sus vectores no nulos es una base de V.
- Si G es una base de V entonces G_0 es una base ortogonal de V.
- Para obtener una base ortonormal a partir de una base ortogonal $B = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_r\}$, bastará con multiplicar cada uno de

los vectores $\bar{w}_i$ de B por el escalar $\dfrac{1}{\|\bar{w}_i\|}$.

A manera de ejemplo, obtengamos una base ortonormal del espacio V generado por los vectores

$$\bar{v}_1 = (1, 0, -1)$$

$$\bar{v}_2 = (-2, 1, 1)$$

$$\bar{v}_3 = (-1, 1, 0)$$

con las operaciones usuales de adición, multiplicación por un escalar y producto interno en R^3.

Primero obtendremos un generador ortogonal de dicho espacio, mediante el proceso de Gram-Schmidt.

Para ello hacemos

$$\bar{w}_1 = \bar{v}_1 = (1, 0, -1)$$

$$\bar{w}_2 = \bar{v}_2 - \frac{(\bar{v}_2|\bar{w}_1)}{(\bar{w}_1|\bar{w}_1)}\bar{w}_1 = (-2, 1, 1) - \frac{-3}{2}(1, 0, -1) = (-\tfrac{1}{2}, 1, -\tfrac{1}{2})$$

$$\bar{w}_3 = \bar{v}_3 - \frac{(\bar{v}_3|\bar{w}_1)}{(\bar{w}_1|\bar{w}_1)}\bar{w}_1 - \frac{(\bar{v}_3|\bar{w}_2)}{(\bar{w}_2|\bar{w}_2)}\bar{w}_2 =$$

$$= (-1, 1, 0) - \frac{-1}{2}(1, 0, -1) - \frac{3/2}{3/2}(-\tfrac{1}{2}, 1, -\tfrac{1}{2}) = (0, 0, 0)$$

Por lo que

$$G_0 = \{(1, 0, -1), (-\tfrac{1}{2}, 1, -\tfrac{1}{2}), (0, 0, 0)\}$$

es un generador ortogonal de V y el conjunto

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

es una base ortogonal de dicho espacio.

Para obtener una base ortonormal calculamos

$$||\bar{w}_1|| = (\bar{w}_1|\bar{w}_1)^{1/2} = \sqrt{2}$$

$$||\bar{w}_2|| = (\bar{w}_2|\bar{w}_2)^{1/2} = \sqrt{\frac{3}{2}}$$

y en consecuencia el conjunto

$$B' = \left\{ \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}} \right), \left(\frac{-1}{\sqrt{6}}, \sqrt{\frac{2}{3}}, \frac{-1}{\sqrt{6}} \right) \right\}$$

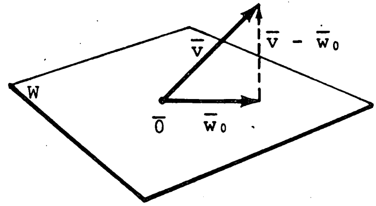
es una base ortonormal de V.

- El teorema de proyección

Uno de los resultados más importantes de la teoría de los espacios con producto interno es el llamado teorema de proyección (o teorema de la mejor aproximación), debido a que constituye el fundamento de diversos métodos de optimización.

Es fácil comprender los conceptos que intervienen en dicho teorema, así como el problema de que se ocupa, a través de su interpretación geométrica en el espacio cartesiano.

Consideremos para ello el subespacio de R^3 representado por el plano W de la siguiente figura, y un vector arbitrario $\bar{v} \in R^3$.



Se trata de encontrar un vector $\bar{w}_0$ del plano W que sea "el más cer cano" a $\overline{v}$ (o "el más aproximado"), en el sentido de que la distancia entre $\overline{v}$ y $\overline{w}_0$ sea la menor distancia posible entre $\overline{v}$ y cualquier vector de W.

Si $\overline{v} \in W$ la solución es trivial, puesto que el vector buscado es el propio vector $\overline{v}$; empero, si $\overline{v} \notin W$ cabe hacerse las siguientes preguntas:

1) ¿Existe algún vector $\overline{w}_0 \in W$ tal que $\|\overline{v} - \overline{w}_0\|$ sea mínima?
2) ¿Es dicho vector único?
3) ¿Cuáles son las características que lo definen?

Para el problema geométrico planteado en R^3, la intuición sugiere que tal vector existe, que es único, y que es precisamente la proyección (ortogonal) de $\overline{v}$ sobre el plano W, como se infiere de la figura anterior.

Sin embargo, para demostrar que esto es válido en un espacio cualquiera con producto interno, se requieren algunas definiciones generales y resultados previos que veremos a continuación

IX.5.15 DEFINICION

Sea V un espacio con producto interno y sea S un subconjunto de V. Se dice que un vector $\overline{v} \in V$ es ortogonal al conjunto S si

$$(\overline{v}|\overline{u}) = 0 \quad \forall\, \overline{u} \in S$$

El conjunto de todos los vectores de V ortogonales a S se denota con $S^{\perp}$; esto es:

$$S^{\perp} = \{\overline{v} \in V \mid (\overline{v}|\overline{u}) = 0,\ \forall\, \overline{u} \in S\}$$

Como el lector podrá demostrar, el conjunto $S^{\perp}$ es un subespacio de V, tanto si S es un subespacio de V como si no lo es.

Cuando S es un subespacio de V el conjunto $S^{\perp}$ se conoce como el "complemento ortogonal" de S, debido a que tiene la propiedad que se enuncia a continuación.

IX.5.16 TEOREMA

Sea V un espacio con producto interno y sea W un subespacio de V de dimensión finita. Entonces, cualquier vector $\bar{v} \in V$ puede expresarse en forma única como

$$\bar{v} = \bar{w} + \bar{w}'$$

donde $\bar{w} \in W$ y $\bar{w}' \in W^{\perp}$

DEMOSTRACION

Probaremos primero que existen dos vectores $\bar{w}$ y $\bar{w}'$ con las características propuestas y mostraremos a continuación que tales vectores son únicos.

Sea $B = \{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$ una base ortonormal de W. Definamos a $\bar{w}$ como

$$\bar{w} = \sum_{i=1}^{n} (\bar{v}|\bar{e}_i)\bar{e}_i \qquad (1)$$

y hagamos

$$\bar{w}' = \bar{v} - \bar{w} \qquad (2)$$

Es claro que $\bar{w}$ y $\bar{w}'$, así definidos, son tales que

$$\bar{v} = \bar{w} + \bar{w}'$$

y además $\bar{w} \in W$.

Para probar que $\bar{w}' \in W^{\perp}$ consideremos el producto de $\bar{w}'$ por un ele mento $\bar{e}_j$ cualquiera de la base B. Entonces, por (2)

$$(\bar{w}'|\bar{e}_j) = (\bar{v} - \bar{w}|\bar{e}_j)$$

$$(\bar{w}'|\bar{e}_j) = (\bar{v}|\bar{e}_j) - (\bar{w}|\bar{e}_j) \qquad (3)$$

pero, por definición de $\bar{w}$ se tiene que

$$(\bar{w}|\bar{e}_j) = (\sum_{i=1}^{n} (\bar{v}|\bar{e}_i)\ \bar{e}_i|\bar{e}_j)$$

Aplicando reiteradamente ii) y iii) de IX.5.1

$$(\bar{w}|\bar{e}_j) = \sum_{i=1}^{n} \left[(\bar{v}|\bar{e}_i)(\bar{e}_i|\bar{e}_j) \right]$$

Entonces, tomando en cuenta que B es un conjunto ortonormal se concluye que

$$(\bar{w}|\bar{e}_j) = (\bar{v}|\bar{e}_j)$$

Llevando este resultado a la expresión (3) se tiene

$$(\bar{w}'|\bar{e}_j) = 0, \quad \forall\ \bar{e}_j \in B$$

Es decir, $\bar{w}'$ es ortogonal al conjunto B y en consecuencia, es ortogonal al espacio W generado por B; esto es

$$\bar{w}' \in W^{\perp}$$

como se quería.

Para mostrar que $\bar{w}$ y $\bar{w}'$ son únicos, supongamos otros dos vectores $\bar{z} \in W$ y $\bar{z}' \in W^{\perp}$ tales que

$$\bar{v} = \bar{z} + \bar{z}'$$

como también

$$\bar{v} = \bar{w} + \bar{w}'$$

se tiene que

$$\bar{w} + \bar{w}' = \bar{z} + \bar{z}'$$

y en consecuencia

$$\bar{w} - \bar{z} = \bar{z}' - \bar{w}'$$

pero $\bar{w} - \bar{z} \in W$ y $\bar{z}' - \bar{w}' \in W^{\perp}$, por lo que $\bar{w} - \bar{z}$ es ortogonal y a la vez igual a $\bar{z}' - \bar{w}'$.

En consecuencia, como el vector cero es el único vector ortogonal a si mismo, se tiene

$$\bar{w} - \bar{z} = \bar{z}' - \bar{w}' = \bar{0}$$

de donde

$$\bar{z} = \bar{w} \quad y \quad \bar{z}' = \bar{w}'$$

por lo que la descomposición a que se refiere IX.5.16 es única. □

El vector $\bar{w}$ de la expresión (1) se conoce como la proyección de $\bar{v}$ sobre el espacio W, debido a que es la suma de las proyecciones de $\bar{v}$ sobre cada uno de los elementos de una base de W.

IX.5.17 DEFINICION

Sean V un espacio con producto interno, W un subespacio de V de dimensión finita y $\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$ una base ortonormal de W. Si $\bar{v} \in V$, el vector

$$\sum_{i=1}^{n} (\bar{v}|\bar{e}_i)\bar{e}_i$$

se llama la proyección de $\bar{v}$ sobre W

Con estos elementos podemos ahora enunciar y demostrar el teorema de proyección.

IX.5.18 TEOREMA (DE PROYECCION)
Sea V un espacio con producto interno y sea W un subespacio de V. Para cada vector $\bar{v} \in V$ existe uno y sólo un vector $\bar{w}_0 \in W$ tal que
$\|\bar{v} - \bar{w}_0\| < \|\bar{v} - \bar{w}\|, \forall \bar{w} \in W, \bar{w} \neq \bar{w}_0$
Dicho vector es la proyección de $\bar{v}$ sobre W.

DEMOSTRACION

Sea $\bar{v} \in V$, por IX.5.16 existen dos únicos vectores $\bar{w}_0 \in W$ y $\bar{w}_0' \in W^{\perp}$ tales que

$$\bar{v} = \bar{w}_0 + \bar{w}_0'$$

donde $\bar{w}_0$ es la proyección de $\bar{v}$ sobre W y

$$\bar{w}_0' = \bar{v} - \bar{w}_0 \in W^{\perp}$$

Además, para cualquier vector $\bar{w} \in W$ se tiene que

$$\bar{v} - \bar{w} = (\bar{v} - \bar{w}_0) + (\bar{w}_0 - \bar{w})$$

Como $(\bar{v} - \bar{w}_0) \in W^{\perp}$ y $(\bar{w}_0 - \bar{w}) \in W$, dichos vectores son ortogonales; por lo que, de IX.5.10

$$\|(\bar{v} - \bar{w}_0) + (\bar{w}_0 - \bar{w})\|^2 = \|\bar{v} - \bar{w}_0\|^2 + \|\bar{w}_0 - \bar{w}\|^2$$

es decir

$$\|\bar{v} - \bar{w}\|^2 = \|\bar{v} - \bar{w}_0\|^2 + \|\bar{w}_0 - \bar{w}\|^2$$

Además, si $\overline{w} \neq \overline{w}_0$ se tendrá $||\overline{w}_0 - \overline{w}|| > 0$, por lo que

$$||\overline{v} - \overline{w}||^2 > ||\overline{v} - \overline{w}_0||^2$$

y en consecuencia

$$||\overline{v} - \overline{w}|| > ||\overline{v} - \overline{w}_0||$$

como se quería. $\square$

IX.5.19 EJERCICIOS

1.- Verificar que

$$(\overline{x}|\overline{y}) = 3x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2$$

donde $\overline{x} = (x_1, x_2)$, $\overline{y} = (y_1, y_2) \in R^2$

es un producto interno en R^2. Con dicho producto calcular $(\overline{a}|\overline{b})$ y $(\overline{b}|\overline{c})$

para $\overline{a} = (-3, 1)$, $\overline{b} = (-2, -2)$, $\overline{c} = (-1, 4)$

2.- Demostrar que

si $\overline{z} = (z_1, z_2, \ldots, z_n)$, $\overline{w} = (w_1, w_2, \ldots, w_n) \in C^n$,

entonces, la función definida por

$$(\overline{z}|\overline{w}) = z_1\overline{w}_1 + z_2\overline{w}_2 + \ldots + z_n\overline{w}_n$$

donde $\overline{w}_i$ es el conjugado de w_i, es un producto interno en C^n; mientras que la definida por

$$(\overline{z}|\overline{w}) = z_1w_1 + z_2w_2 + \ldots + z_nw_n$$

no lo es ¿Cuáles axiomas no se verifican?

3.- a) Demostrar que

$$(A|B) = tr(A^{*}B)$$

es un producto interno en el espacio de las matrices de $m \times n$ con elementos complejos.

b) Calcular $(A|B)$ con dicho producto y verificar que se cumple la desigualdad de Cauchy-Schwarz para las matrices

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

4.- Calcular la distancia y el ángulo entre los vectores $\bar{x} = (3, -2)$ y $\bar{y} = (2, 1)$

a) Con el producto interno usual en R^2

b) Con el producto interno definido por

$$(\bar{x}|\bar{y}) = 3x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2$$

5.- En el espacio vectorial de las funciones continuas en el intervalo $[-1, 1]$

a) Demostrar que la función definida como

$$(f|g) = \int_{-1}^{1} f(t)g(t)dt$$

es un producto interno

b) Si $f(t) = a$, $g(t) = bt^3$ y $h(t) = at+b$ determinar la pareja de enteros a, b para la cual dos de las funciones son ortogonales y dos forman un ángulo de $\frac{\pi}{3}$.

c) Para las funciones del inciso anterior, hallar las distancias de f a g y de g a h, y verificar que

$$d(f,h) \leq d(f,g) + d(g,h)$$

6.- Demostrar que en un espacio V con producto interno, se cumple que $\forall\ \bar{u}, \bar{v}, \bar{w} \in V$

a) $d(\bar{u}, \bar{w}) \leq d(\bar{u}, \bar{v}) + d(\bar{v}, \bar{w})$

b) $d(\bar{u}, \bar{v}) = d(\bar{v}, \bar{u})$

7.- Demostrar el teorema de Pitágoras (IX.5.10) en un espacio con producto interno cualquiera, y demostrar que en un espacio con producto interno real $(\bar{u}|\bar{v}) = 0$ es además una condición necesaria para la igualdad

$$||\bar{u} + \bar{v}||^2 = ||\bar{u}||^2 + ||\bar{v}||^2$$

8.- Con el producto interno definido en R^2 por

$$(\bar{x}|\bar{y}) = 3x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2 \quad \forall\ \bar{x} = (x_1, x_2),\ \bar{y} = (y_1, y_2)$$

a) Calcular el valor de p tal que $B_1 = \{(1, 1), (2, p)\}$ sea una base ortogonal de R^2

b) A partir de B_1, obtener una base ortonormal B_2 de R^2.

c) Determinar las coordenadas en B_2 del vector $(1, -1) \in R^2$, utilizando el teorema IX.5.13.

9.- Sean V un espacio con producto interno, $G = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ un generador de V y $G_0 = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n\}$ tal que

$$\bar{w}_1 = \bar{v}_1$$

$$\bar{w}_i = \bar{v}_i - \sum_{k=1}^{i-1} \frac{(\bar{v}_i|\bar{w}_k)}{(\bar{w}_k|\bar{w}_k)}\,\bar{w}_k, \text{ para } i = 2, 3, \ldots, n$$

Demostrar que el subconjunto de G_0 formado por sus vectores no nulos es una base ortogonal de V.

10.- Considerando el producto interno usual en R^3, obtener una base ortonormal del espacio generado por cada uno de los siguientes conjuntos

a) $A = \{(-1, 2, 1), (0, 1, 2), (-2, 3, 0)\}$

b) $B = \{(1, 0, -2), (2, 1, -1), (-1, 3, 1)\}$

11.- Para el producto interno usual en R^3, obtener el complemento ortogonal $S_i^{\perp}$ de cada uno de los siguientes subespacios de R^3, y dar una interpretación geométrica de dichos complementos.

a) $S_1 = \{(0, 0, z) \mid z \in R\}$

b) $S_2 = \{(x, x, 0) \mid x \in R\}$

c) $S_3 = \{(x, x, z) \mid x, z \in R\}$

d) $S_4 = R^3$

e) $S_5 = \{(0, 0, 0)\}$

12.- Si V es un espacio con producto interno demostrar que

a) Para $S \subset V$, $S^{\perp}$ es un subespacio de V.

b) $V^{\perp} = \{\bar{0}_V\}$

c) $\{\bar{0}_V\}^{\perp} = V$

d) Para un subespacio W de V, $(W^{\perp})^{\perp} = W$

13.- Sean V el espacio de los polinomios de coeficientes reales y grado menor que 3 sobre R, $W = \{ax^2 + bx - 2a - b \mid a, b \in R\}$ un subespacio de V, y el producto interno en V definido por

$$(p|q) = \sum_{i=0}^{2} p(i)\, q(i)$$

a) Obtener una base ortonormal de W

b) Determinar la proyección sobre W, de cada uno de los polinomios

$$p(x) = x^2 - x + 1$$

$$q(x) = 3x - 2$$

14.- Sea $W = \left\{ \begin{bmatrix} a & b \\ -b & 2a \end{bmatrix} \middle| \; a, b \in R \right\}$ un subespacio del espacio de las matrices cuadradas de orden 2 sobre R con producto interno definido por $(A|B) = \operatorname{tr}(AB^T)$.

Obtener las matrices pertenecientes a W, más próximas a

$M = \begin{bmatrix} 2 & -1 \\ 1 & -2 \end{bmatrix}$ y $N = \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}$, respectivamente.

15.- Obtener el vector $\hat{x} \in R^2$ que mejor se aproxime a un vector $\bar{x}$ tal que

$$\bar{y} = M\bar{x}^T \quad \text{para} \quad \bar{y} = (1, 2, -1) \quad \text{y} \quad M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}$$

considerando el producto interno usual en R^3. Esto es,

$$||\bar{y} - M\hat{x}^T|| \leq ||\bar{y} - M\bar{x}^T|| \;\forall\; \bar{x} \in R^2$$

La técnica se conoce como estimación por mínimos cuadrados.

CAPITULO X TRANSFORMACIONES LINEALES

INTRODUCCION

En diversas ramas de la matemática, la física y las ciencias sociales, es frecuente utilizar modelos que emplean funciones vectoriales de variable vectorial; es decir, funciones del tipo $\bar{w} = f(\bar{v})$ donde $\bar{w}$ y $\bar{v}$ son vectores. A tales funciones se les conoce usualmente como "Transformaciones".†

En este capítulo nos ocuparemos de una clase especial de transformaciones, denominadas lineales, que son las más simples y también las de mayor aplicación.

En la práctica, muchos problemas que involucran transformaciones más generales suelen resolverse aproximando éstas a transformaciones lineales.

† *Algunos autores utilizan términos como "aplicación" o "mapeo" en lugar de transformación.*

X.1 CONCEPTOS FUNDAMENTALES

- Transformación. Dominio, codominio, recorrido y núcleo.

X.1.1 DEFINICION

Si V y W son espacios vectoriales una función $T: V \to W$ recibe el nombre de transformación. Los espacios V y W se llaman, respectivamente, dominio y codominio de la transformación

Así, por ejemplo, la función $T: R^3 \to R^2$ definida por la regla

$$T(x, y, z) = (x, y)$$

es una transformación. Los espacios vectoriales R^3 y R^2 son, respectivamente, el dominio y el codominio de T.

Respecto a la expresión anterior cabe aclarar que se acostumbra escribir $T(x, y, z)$, en lugar de $T[(x, y, z)]$, con el propósito de simplificar la notación.

Para la transformación T definida anteriormente se tiene, por ejemplo, que

$$T(1, 2, 3) = (1, 2)$$

Esto es, la imagen del vector (1, 2, 3) bajo la transformación T es el vector (1, 2); también se dice que T transforma al vector (1, 2, 3) en el vector (1, 2). Otra manera de indicar esto es la siguiente

$$(1, 2, 3) \xrightarrow{T} (1, 2)$$

Algunos otros ejemplos de imágenes de vectores de R^3 bajo esta mis ma transformación son

$$T(-2, 7, 1) = (-2, 7)$$

$$T(0, 0, 4) = (0, 0), \text{ etc.}$$

Esta transformación tiene la siguiente interpretación geométrica: si (x, y, z) representa un segmento dirigido cualquiera del espacio cartesiano tridimensional, T transforma dicho segmento en su proyección sobre el plano X - Y.

Consideremos ahora otra transformación de R^3 en R^2, definida por

$$S(x, y, z) = (y, 3y)$$

Como se ve, la imagen de cualquier vector de R^3 bajo esta transformación es una pareja ordenada cuya segunda componente es el triple de la primera. Por ello, no todos los vectores de R^2 son imagen de algún vector del dominio, sino únicamente aquellos de la forma (a, 3a).

Al conjunto formado por todos estos vectores se le conoce como "recorrido" de la transformación S.

En general, se llama recorrido de una transformación al conjunto de todos los vectores que son imagen de algún vector del dominio.

Otro conjunto importante definido por una transformación es el núcleo.

Se llama núcleo de una transformación al conjunto de vectores cuya imagen es el vector cero.

Por ejemplo, para la transformación S definida anteriormente cualquier vector cuya segunda componente sea nula tiene como imagen al vector cero de R^2, por lo que pertenece al núcleo de S.

De esta manera, los conceptos de recorrido y núcleo pueden definirse como sigue

X.1.2 DEFINICION

Sea $T: V \to W$ una transformación:

i) Se llama recorrido de T al conjunto

$$T(V) = \{T(\bar{v}) \mid \bar{v} \in V\}$$

ii) Se llama núcleo de T al conjunto

$$N(T) = \{\bar{v} \mid T(\bar{v}) = \bar{0}_W\}$$

Como se sigue de la definición anterior, el recorrido es un subconjunto del codominio mientras que el núcleo lo es del dominio.

Tales subconjuntos pueden ser propios o impropios, dependiendo de la transformación de que se trate.

Por ejemplo, para la transformación S definida anteriormente, un vector (a, b) de R^2 pertenece al recorrido de S si existe al menos un vector (x, y, z) de R^3 tal que

$$S(x, y, z) = (a, b)$$

pero como

$$S(x, y, z) = (y, 3y)$$

dicha condición se cumple si y sólo si

$$(a, b) = (y, 3y), \text{ para algún } y \in R.$$

esto es, si

$$a = y$$

y

$$b = 3y, \text{ para algún } y \in R$$

por lo que

$$S(R^3) = \{(y, 3y) \mid y \in R\}$$

y el recorrido de S es un subconjunto propio de R^2.

Para esta misma transformación, un vector (x, y, z) pertenece al núcleo de S si es tal que

$$S(x, y, z) = (0, 0)$$

esto es

$$(y, 3y) = \bar{0}$$

condición que se sumple si y sólo si

$$y = 0$$

por lo que

$$N(S) = \{(x, 0, z) \mid x, z \in R\}$$

y el núcleo de S es un subconjunto propio de R^3.

Para el caso de la transformación $T: R^3 \rightarrow R^2$ definida por

$$T(x, y, z) = (x, y)$$

el recorrido es igual al codominio; esto es

$$T(R^3) = R^2$$

mientras que el núcleo

$$N(T) = \{(0, 0, z) \mid z \in R\}$$

es un subconjunto propio del dominio.

Un ejemplo donde el núcleo es igual al dominio lo constituye la llamada "transformación cero" O: $R^3 \to R^2$ definida por

$$O(x, y, z) = (0, 0)$$

En este caso el recorrido de la transformación está constituido únicamente por el vector cero de R^2.

- Linealidad. Recorrido y núcleo de una transformación lineal.

Para introducir la noción de linealidad consideremos nuevamente la transformación S definida por

$$S(x, y, z) = (y, 3y)$$

y elijamos un par de vectores del dominio, por ejemplo

$$\overline{v}_1 = (1, 2, 3) \quad y \quad \overline{v}_2 = (0, -3, 4)$$

Si sumamos dichos vectores y calculamos la imagen de la suma bajo la transformación S obtendremos

$$\overline{v}_1 + \overline{v}_2 = (1, 2, 3) + (0, -3, 4) = (1, -1, 7)$$

y

$$S(\overline{v}_1 + \overline{v}_2) = S(1, -1, 7) = (-1, -3)$$

A este mismo resultado se llega sumando las imágenes de los vectores $\overline{v}_1$ y $\overline{v}_2$; ya que

$$S(\overline{v}_1) = S(1, 2, 3) = (2, 6)$$

y

$$S(\overline{v}_2) = S(0, -3, 4) = (-3, -9)$$

por lo que

$$S(\overline{v}_1) + S(\overline{v}_2) = (2, 6) + (-3, -9) = (-1, -3)$$

Así, encontramos que "la imagen de la suma es igual a la suma de las imágenes", para los vectores $\overline{v}_1 = (1, 2, 3)$ y $\overline{v}_2 = (0, -3, 4)$.

Cabe ahora preguntarse si esto se cumple para cualquier par de vectores $\overline{v}_1$, $\overline{v}_2$ de R^3; es decir, ¿es cierto que

$$S(\overline{v}_1 + \overline{v}_2) = S(\overline{v}_1) + S(v_2);\ \forall\ \overline{v}_1, \overline{v}_2 \in R^3?$$

La respuesta es afirmativa, como se demuestra a continuación.

Sean $\overline{v}_1 = (x_1, x_2, x_3)$ y $\overline{v}_2 = (y_1, y_2, y_3)$ dos vectores cualesquiera de R^3.

Entonces

$$S(\overline{v}_1 + \overline{v}_2) = S[(x_1, x_2, x_3) + (y_1, y_2, y_3)]$$

Para la adición usual en R^3 se tiene que

$$S(\overline{v}_1 + \overline{v}_2) = S(x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

y de la regla que define a S

$$S(\overline{v}_1 + \overline{v}_2) = (x_2 + y_2, 3[x_2 + y_2])$$

Además, por la distributividad de la multiplicación sobre la adición de números reales se tiene que

$$S(\overline{v}_1 + \overline{v}_2) = (x_2 + y_2, 3x_2 + 3y_2) \qquad --- (1)$$

Por otra parte, aplicando S a los vectores $\overline{v}_1$ y $\overline{v}_2$.

$$S(\overline{v}_1) = S(x_1, x_2, x_3) = (x_2, 3x_2)$$

y

$$S(\overline{v}_2) = S(y_1, y_2, y_3) = (y_2, 3y_2)$$

sumando estas imágenes mediante la adición usual en R^2 se obtiene

$$S(\overline{v}_1) + S(\overline{v}_2) = (x_2 + y_2, 3x_2 + 3y_2) \qquad --- (2)$$

En consecuencia, de (1) y (2) se sigue que

$$S(\overline{v}_1 + \overline{v}_2) = S(\overline{v}_1) + S(\overline{v}_2); \ \forall\ \overline{v}_1, \overline{v}_2 \in R^3$$

como lo habíamos anticipado.

Esta es una de las dos condiciones que debe cumplir la transformación S para ser considerada lineal. La otra condición es la siguiente

$$S(\alpha\overline{v}_1) = \alpha S(\overline{v}_1); \ \forall\ \overline{v}_1 \in R^3, \alpha \in R$$

Es decir, "la imagen del producto de un escalar por un vector es igual al producto del escalar por la imagen del vector".

Veamos si esto se cumple para cualquier escalar y cualquier vector, en el caso de la transformación S del ejemplo.

Sean $\overline{v}_1 = (x_1, x_2, x_3) \in R^3$ y $\alpha \in R$

Entonces

$$S(\alpha\overline{v}_1) = S[\alpha(x_1, x_2, x_3)]$$

Para el producto por un escalar definido en forma usual en R^3 se tiene

$$S(\alpha\overline{v}_1) = S(\alpha x_1, \alpha x_2, \alpha x_3)$$

por lo que, aplicando la transformación S

$$S(\alpha\overline{v}_1) = (\alpha x_2, 3\alpha x_2) \qquad --- (3)$$

por otra parte, al aplicar S al vector $\overline{v}_1$ se obtiene

$$S(\overline{v}_1) = S(x_1, x_2, x_3) = (x_2, 3x_2)$$

Entonces, multiplicando α por la imagen de $\overline{v}_1$ mediante el producto por un escalar definido en forma usual en R^2

$$\alpha S(\overline{v}_1) = \alpha(x_2, 3x_2) = (\alpha x_2, \alpha[3x_2])$$

y de las propiedades de la multiplicación en R se tiene

$$\alpha S(\overline{v}_1) = (\alpha x_2, 3\alpha x_2) \qquad --- (4)$$

En consecuencia, de (3) y (4)

$$S(\alpha\overline{v}_1) = \alpha S(\overline{v}_1); \quad \forall\ \overline{v}_1 \in R^3,\ \alpha \in R$$

Así, la transformación S del ejemplo cumple también con la segunda condición mencionada, por lo que se trata de una transformación lineal.

En general, la definición de transformación lineal suele enunciarse de la siguiente manera

X.1.3 DEFINICION

Sean V y W dos espacios vectoriales sobre un campo K. Una transformación T: V → W es lineal si $\forall\ \bar{v}_1,\ \bar{v}_2 \in V$ y $\forall\ \alpha \in K$:

i) $T(\bar{v}_1 + \bar{v}_2) = T(\bar{v}_1) + T(\bar{v}_2)$

y

ii) $T(\alpha\bar{v}_1) = \alpha T(\bar{v}_1)$

Cabe hacer notar que el símbolo + que aparece en el miembro izquierdo de la expresión i) representa a la adición definida en V; mientras que en el miembro derecho de dicha expresión el mismo símbolo representa a la adición en W.

De manera similar, el producto en el miembro izquierdo de la expresión ii) se refiere al producto de un escalar de K por un vector de V y en el miembro derecho al producto por un vector de W.

Las expresiones i) y ii) de la definición anterior, que en el estudio de sistemas físicos se conocen como propiedades de superposición y homogeneidad, pueden reunirse en una sola expresión para establecer la siguiente condición de linealidad, equivalente a la definición X.1.3:

$$T(\alpha\bar{v}_1 + \bar{v}_2) = \alpha T(\bar{v}_1) + T(\bar{v}_2);\ \forall\ \bar{v}_1,\ \bar{v}_2 \in V \text{ y } \forall\ \alpha \in K$$

A continuación se presentan algunos ejemplos de transformaciones que, como el lector podrá demostrar, son lineales.

1. Si A es una matriz de m×n con elementos en R, la expresión

   $T(\bar{v}) = A\bar{v},\ \forall\ \bar{v} \in R^n$, donde $\bar{v}$ y $T(\bar{v})$ son vectores columna, define una transformación de R^n en R^m.

2. Sea F el espacio vectorial de las funciones reales de variable real:

a) Si S es el subespacio de F constituido por todas las funciones derivables en un intervalo dado, la transformación

$$D: S \to F$$

definida por

$$D(f) = f', \ \forall \ f \in S$$

se conoce como "operador derivación" y transforma a una función cualquiera f en su derivada f'.

b) Si S es el subespacio de F constituido por todas las funciones continuas en un intervalo $[a, b]$, la transformación

$$J: S \to F$$

definida por

$$[J(f)](x) = \int_a^x f(t)\ dt, \ \forall \ x \in [a, b]$$

se conoce como "operador integración" y transforma a una función cualquiera f en la función de x definida por

$\int_a^x f(t)\ dt.$

3. Si V y W son dos espacios cualesquiera:

a) La transformación

$$O: V \to W$$

definida por

$$O(\bar{v}) = \bar{0}_w; \ \forall \ \bar{v} \in V$$

se conoce como "transformación cero" de V a W.

b) La transformación

$$I_v: V \rightarrow V$$

definida por

$$I_v(\bar{v}) = \bar{v}; \ \forall \ \bar{v} \in V$$

se llama "transformación identidad" en V.

4. Si V es un espacio con producto interno, S un subespacio de V y $B = \{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_r\}$ una base ortonormal de S, la transformación

$$T: V \rightarrow S$$

definida por

$$T(\bar{v}) = \sum_{i=1}^{r} (\bar{v} \mid \bar{e}_i)\bar{e}_i \ ; \ \forall \ \bar{v} \in V$$

transforma un vector cualquiera $\bar{v} \in V$ en su proyección sobre el subespacio S.

Las transformaciones lineales tienen la siguiente propiedad

X.1.4 TEOREMA

Si $T: V \rightarrow W$ es una transformación lineal entonces

$$T(\bar{0}_v) = \bar{0}_w$$

DEMOSTRACION

Si T es lineal, por ii) de X.1.3

$$T(\alpha\cdot\overline{0}_V) = \alpha T(\overline{0}_V),\ \forall\ \alpha \in K$$

Entonces, por i) de IX.1.3

$$T(\overline{0}_V) = \alpha T(\overline{0}_V),\ \forall\ \alpha \in K$$

En particular, si $\alpha = 0$ por ii) de IX.1.3 se tiene

$$T(\overline{0}_V) = \vec{0}_W$$

lo cual demuestra el teorema.

□

Como ya hemos visto, el recorrido de una transformación es un subconjunto del codominio y el núcleo es un subconjunto del dominio. Si la transformación es lineal dichos subconjuntos son además subespacios, como se indica a continuación

X.1.5 TEOREMA

Si $T: V \to W$ es una transformación lineal, entonces:

i) $T(V)$ es un subespacio de W

ii) $N(T)$ es un subespacio de V

DEMOSTRACION

Como $T(V)$ es un subconjunto de W, de acuerdo con el teorema IX.1.6 bastará con probar que $T(V)$ es cerrado para la adición y la multiplicación por un escalar.

Sean $\overline{w}_1$ y $\overline{w}_2$ dos vectores cualesquiera de $T(V)$

Por i) de X.1.2 existen dos vectores $\overline{v}_1$, $\overline{v}_2 \in V$ tales que

$$\overline{w}_1 = T(\overline{v}_1) \quad y \quad \overline{w}_2 = T(\overline{v}_2)$$

se tiene entonces que

$$\overline{w}_1 + \overline{w}_2 = T(\overline{v}_1) + T(\overline{v}_2)$$

y como T es lineal, por i) de X.1.3 se tiene

$$\overline{w}_1 + \overline{w}_2 = T(\overline{v}_1 + \overline{v}_2)$$

pero $\overline{v}_1$, $\overline{v}_2 \in V$ y V es un espacio vectorial, por lo que $\overline{v}_1 + \overline{v}_2 \in V$ y en consecuencia

$$\overline{w}_1 + \overline{w}_2 \in T(V) \qquad --- (1)$$

Además, si α es un escalar de K

$$\alpha\overline{w}_1 = \alpha T(\overline{v}_1)$$

y como T es lineal

$$\alpha\overline{w}_1 = T(\alpha\overline{v}_1)$$

pero $\alpha\overline{v}_1 \in V$ por lo que

$$\alpha\overline{w}_1 \in T(V) \qquad --- (2)$$

En consecuencia, de (1) y (2), T(V) es un subespacio de W.

Se deja al lector como ejercicio la demostración de que el núcleo es un subespacio del dominio. □

Así, por ejemplo, para la transformación $S: R^3 \rightarrow R^2$ definida por

$$S(x, y, z) = (y, 3y)$$

encontramos que su recorrido es el conjunto

$$S(R^3) = \{(y,\ 3y) \mid y \in R\}$$

Es fácil verificar que dicho conjunto es un subespacio de R^2.

Para la misma transformación se obtuvo que

$$N(S) = \{(x,\ 0,\ z) \mid x,\ z \in R\}$$

el cual, como el lector puede comprobar, es un subespacio de R^3.

Para la transformación S del ejemplo anterior fue muy sencillo obtener su recorrido (inclusive por simple inspección), aunque por lo general esto no resulta tan fácil.

No obstante, para determinar el recorrido de una transformación lineal específica podemos aprovechar la siguiente propiedad.

X.1.6 TEOREMA

Sea $T: V \to W$ una transformación lineal. Si $B = \{\bar{v}_1,\ \bar{v}_2, \dots,\ \bar{v}_n\}$ es una base de V, entonces el conjunto $G = \{T(\bar{v}_1),\ T(\bar{v}_2), \dots,\ T(\bar{v}_n)\}$ es un generador de T(V).

DEMOSTRACION

Sea $B = \{\bar{v}_1,\ \bar{v}_2, \dots,\ \bar{v}_n\}$ una base de V. Si $\bar{w}$ es un vector cualquiera de T(V), entonces existe algún vector $\bar{v} \in V$ tal que

$$\bar{w} = T(\bar{v})$$

Como B es una base de V

$$\bar{v} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \dots + \alpha_n\bar{v}_n$$

por lo que

$$\bar{w} = T(\alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \ldots + \alpha_n\bar{v}_n)$$

y como T es lineal

$$\bar{w} = \alpha_1\ T(\bar{v}_1) + \alpha_2\ T(\bar{v}_2) + \ldots + \alpha_n\ T(\bar{v}_n)$$

En consecuencia, el conjunto $G = \{T(\bar{v}_1),\ T(\bar{v}_2), \ldots,\ T(\bar{v}_n)\}$ es un generador de $T(V)$.

□

Si el conjunto G del teorema anterior es linealmente independiente entonces es una base de T(V), y si es linealmente dependiente puede obtenerse una base de T(V) a partir de él.

Por ejemplo, para la transformación $T\colon R^3 \to R^3$ definida por

$$T(x, y, z) = (3x + y,\ 6x - z,\ 2y + z)$$

Consideremos, por sencillez, la base canónica de R^3

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

para lo cual se tiene

$$T(1, 0, 0) = (3, 6, 0)$$
$$T(0, 1, 0) = (1, 0, 2)$$
$$T(0, 0, 1) = (0, -1, 1)$$

Entonces, del teorema X.1.6, el conjunto

$$G = \{(3, 6, 0),\ (1, 0, 2),\ (0, -1, 1)\}$$

es un generador del recorrido de T.

Como G es un subconjunto de R^3, podemos obtener L(G) como el espacio renglón de la matriz

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

reduciéndola a su forma canónica escalonada

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

Así, una base del recorrido es

$$\{(1, 0, 2), (0, 1, -1)\}$$

de donde

$$T(R^3) = \{(a, b, 2a - b) \mid a, b \in R\}$$

Para este mismo ejemplo obtengamos ahora el núcleo de la transformación. Se requiere entonces determinar bajo qué condiciones un vector $(x, y, z) \in R^3$ es tal que

$$T(x, y, z) = (0, 0, 0)$$

Aplicando la regla de correspondencia se tiene

$$(3x + y, 6x - z, 2y + z) = (0, 0, 0)$$

lo cual es equivalente al sistema homogéneo

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

cuya solución general es

$$x = x$$
$$y = -3x$$
$$z = 6x$$

por lo que

$$N(T) = \{(x, -3x, 6x) \mid x \in R\}$$

El teorema X.1.5 establece que el recorrido y el núcleo de una transformación lineal son espacios vectoriales y, por lo tanto, tienen dimensión.

Como el lector puede darse cuenta, para el ejemplo anterior las dimensiones de estos espacios son

$$\dim T(R^3) = 2 \qquad y \qquad \dim N(T) = 1$$

si sumamos dichas dimensiones obtendremos un número que coincide con la dimensión del dominio.

Esto sucede siempre que trabajamos con espacios de dimensión finita, como se establece a continuación

X.1.7 TEOREMA

Si V es un espacio de dimensión finita y $T: V \to W$ es una transformación lineal, entonces

$$\dim V = \dim T(V) + \dim N(T)$$

DEMOSTRACION

Como V es de dimensión finita y N(T) es un subespacio de V, por

IX.2.13, se tiene que

$$\dim N(T) \leq \dim V$$

A continuación se demuestra el teorema para

$$0 < \dim N(T) < \dim V$$

La demostración del mismo para los casos extremos, $\dim N(T) = 0$ y $\dim N(T) = \dim V$, se deja al lector como ejercicio.

Sean $k = \dim N(T)$, $A = \{\overline{n}_1, \overline{n}_2, \ldots, \overline{n}_k\}$ una base de $N(T)$ y $\dim V = k + r$

Como A es un conjunto linealmente independiente constituido por vectores de V, podemos ampliarlo hasta obtener una base de V de la forma

$$B = \{\overline{n}_1, \overline{n}_2, \ldots, \overline{n}_k, \overline{v}_1, \overline{v}_2, \ldots, \overline{v}_r\}$$

Probaremos entonces que el conjunto

$$C = \{T(\overline{v}_1), T(\overline{v}_2), \ldots, T(\overline{v}_r)\}$$

es una base de T(V).

1) Probemos que el conjunto C es un generador de T(V):

Sea $\overline{w}$ un vector cualquiera de T(V), entonces existe algún $\overline{v} \in V$ tal que

$$\overline{w} = T(\overline{v})$$

En consecuencia, existen escalares $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_r$ tales que

$$\overline{w} = T(\alpha_1\overline{n}_1 + \ldots + \alpha_k\overline{n}_k + \beta_1\overline{v}_1 + \ldots + \beta_r\overline{v}_r)$$

y como T es lineal

$$\overline{w} = \alpha_1 T(\overline{n}_1) + \ldots + \alpha_k T(\overline{n}_k) + \beta_1 T(\overline{v}_1) + \ldots + \beta_r T(\overline{v}_r)$$

Además $\overline{n}_1, \ldots, \overline{n}_k \in N(T)$, por lo que

$$\overline{w} = \alpha_1 \overline{0}_W + \ldots + \alpha_k \overline{0}_W + \beta_1 T(\overline{v}_1) + \ldots + \beta_r T(\overline{v}_r)$$

de donde, finalmente

$$\overline{w} = \beta_1 T(\overline{v}_1) + \ldots + \beta_r T(\overline{v}_r)$$

y el conjunto C es un generador de T(V).

2) Probemos que el conjunto C es linealmente independiente:

Sean $\lambda_1, \ldots, \lambda_r$ tales que

$$\lambda_1 T(\overline{v}_1) + \ldots + \lambda_r T(\overline{v}_r) = \overline{0}_W$$

Como T es lineal se tiene

$$T(\lambda_1 \overline{v}_1 + \ldots + \lambda_r \overline{v}_r) = \overline{0}_W$$

por lo que $\lambda_1 \overline{v}_1 + \ldots + \lambda_r \overline{v}_r \in N(T)$, y como A es una base de N(T)

$$\lambda_1 \overline{v}_1 + \ldots + \lambda_r \overline{v}_r = \gamma_1 \overline{n}_1 + \ldots + \gamma_k \overline{n}_k$$

esto es

$$\gamma_1 \overline{n}_1 + \ldots + \gamma_k \overline{n}_k - \lambda_1 \overline{v}_1 - \ldots - \lambda_r \overline{v}_r = \overline{0}_V$$

Pero como B es linealmente independiente, la expresión anterior implica

$$\gamma_1 = \ldots = \gamma_k = -\lambda_1 = \ldots = -\lambda_r = 0$$

por lo que

$$\lambda_1 = \ldots = \lambda_r = 0$$

y el conjunto C es linealmente independiente.

En consecuencia, de 1) y 2), el conjunto C es una base de T(V), por lo que dim T(V) = r y la demostración queda completa. $\square$

X.1.8 EJERCICIOS

1.- Para cada una de las siguientes transformaciones

$T: R^3 \rightarrow R^3$ tal que $T(x, y, z) = (3x + y, x - z, 2x + y + z)$

$S: R^3 \rightarrow R^4$ tal que $S(x, y, z) = (2x - z, \text{sen } y, z - 2, - x + z$

$Q: R^3 \rightarrow R^3$ tal que $Q(x, y, z) = (x + y, y - 1, xy)$

a) Obtener la imagen bajo la transformación de los vectores (1, 0, 2) y (1, -3, 1).

b) Determinar si la transformación es lineal.

c) Definir el núcleo y el recorrido de la transformación.

2.- Sea el espacio vectorial de todas las funciones reales de variable real, continuas en el intervalo $(0, \infty)$ y sea $T: V \rightarrow V$ tal que

$$|T(f)|(x) = f(x) + \int_0^x f(t)\, dt$$

a) Obtener la imagen bajo T de las funciones f, g y θ definidas por

$$f(x) = \text{sen } 2x, \quad g(x) = e^{-x}, \quad \theta(x) = 0 \quad \forall\, x \in R$$

b) Demostrar que T es lineal.

3.- Demostrar que el núcleo de una transformación lineal es un subespacio del dominio de la transformación.

4.- Demostrar que si T: $V \rightarrow W$ es una transformación lineal, entonces

$$\dim V = \dim T(V) + \dim N(T)$$

para los casos en que

a) $\dim N(T) = 0$

b) $\dim N(T) = \dim V$

5.- Para la transformación T: $V \rightarrow W$ donde V es el espacio de los polinomios de grado menor que 3 con coeficientes reales, W el de las matrices cuadradas de orden 2 con elementos reales, y T está definida por

$$T(ax^2 + bx + c) = \begin{bmatrix} a + c & b + 2c \\ 2a - b & -a - c \end{bmatrix}$$

a) Verificar que T es lineal

b) Determinar el núcleo de T y una base de él.

c) Determinar el recorrido de T y una base de él.

d) Calcular las dimensiones del núcleo y del recorrido de la transformación.

X.2 REPRESENTACION MATRICIAL DE UNA TRANSFORMACION LINEAL

En el caso de espacios de dimensión finita las transformaciones lineales pueden ser representadas por medio de matrices, las cuales se emplean para resolver diversos problemas con grandes ventajas desde el punto de vista operacional.

Con objeto de introducir el concepto de matriz asociada a una transformación lineal, consideremos la transformación $T: R^3 \to R^2$ definida por

$$T(x, y, z) = (x + 2y, 3x - z)$$

y tratemos de encontrar una matriz A tal que el producto de ésta por cualquier vector del dominio nos proporcione la imagen de dicho vector bajo la transformación T; esto es, una matriz A tal que

$$A\bar{v} = T(\bar{v})$$

Como en este caso $\bar{v}$ es un vector de R^3 y $T(\bar{v})$ es un vector de R^2, la igualdad anterior sólo puede lograrse mediante una matriz A de 2×3, considerando a $\bar{v}$ y $T(\bar{v})$ como vectores columna.

En consecuencia, la matriz A deberá ser de la forma

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

y satisfacer la igualdad

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + 2y \\ 3x - z \end{bmatrix} \quad \forall\, x, y, z \in R$$

Es decir

$$a_{11}x + a_{12}y + a_{13}z = x + 2y$$

y

$$a_{21}x + a_{22}y + a_{23}z = 3x - z \ , \ \forall \ x, y, z \in R$$

lo cual se cumple para los valores

$$a_{11} = 1, \ a_{12} = 2, \ a_{13} = 0$$

y

$$a_{21} = 3, \ a_{22} = 0, \ a_{23} = -1$$

por lo que la matriz

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

satisface las condiciones pedidas.

A esta matriz se le conoce como "matriz asociada a la transforma - ción T" (definida anteriormente) y se le representa con M(T).

Veamos ahora que relación tienen los elementos de esta matriz con los vectores de una base del dominio.

Si calculamos las imágenes de los vectores de la base canónica

$$E = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

bajo la transformación T, obtenemos

$$T(1, 0, 0) = (1, 3)$$
$$T(0, 1, 0) = (2, 0)$$
$$T(0, 0, 1) = (0, -1)$$

Como el lector puede darse cuenta, estas imágenes (consideradas co mo vectores columna) son precisamente las columnas que integran la matriz asociada a T.

Se puede demostrar fácilmente que éste es un resultado de carácter general. Por ello, para obtener la matriz asociada a una transformación lineal bastará calcular las imágenes de los vectores que integran la base canónica del dominio.

Por ejemplo, para la transrormación S; $R^3 \to R^3$ definida por

$$S(x, y, z) = (3x + y, 6x - z, 2y + z)$$

se tiene

$$S(1, 0, 0) = (3, 6, 0)$$
$$S(0, 1, 0) = (1, 0, 2)$$
$$S(0, 0, 1) = (0, -1, 1)$$

por lo que su matriz asociada es

$$M(S) = \begin{bmatrix} 3 & 1 & 0 \\ 6 & 0 & -1 \\ 0 & 2 & 1 \end{bmatrix}$$

En efecto, si multiplicamos un vector cualquiera de R^3 por esta matriz se obtiene

$$\begin{bmatrix} 3 & 1 & 0 \\ 6 & 0 & -1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x + y \\ 6x - z \\ 2y + z \end{bmatrix}$$

que es la imagen de dicho vector bajo la transformación S.

- Matriz asociada, referida a dos bases cualesquiera.

En los ejemplos anteriores las columnas de la matriz asociada a la transformación son las imágenes de los vectores de la base canónica del dominio, y al multiplicar dicha matriz por un vector se obtiene la imagen de éste bajo la transformación. Esto es posible debido a que tanto el dominio como el codominio son espacios del tipo R^n.

Sin embargo, las ideas anteriores pueden generalizarse al caso de espacios vectoriales cualesquiera, simplemente reemplazando los vectores por sus respectivos vectores de coordenadas.

De esta manera, si $T: V \rightarrow W$ es una transformación lineal, A es una base de V y B es una base de W, en lugar del vector $\bar{v} \in V$ tendremos a su vector de coordenadas en la base A: $(\bar{v})_A$, en lugar de $T(\bar{v}) \in W$ tendremos a su vector de coordenadas en la base B: $[T(\bar{v})]_B$. La matriz que al multiplicar al vector $(\bar{v})_A$ nos da como resultado $[T(\bar{v})]_B$ se llama "matriz asociada a T referida a las bases A y B", y se representa con $M_B^A(T)$; esto es

$$M_B^A(T) \; (\bar{v})_A = [T(\bar{v})]_B$$

Las columnas de dicha matriz son los vectores de coordenadas, en la base B, de las imágenes de los elementos que integran la base A, como lo establece el siguiente teorema

X.2.1 TEOREMA

Sean V y W dos espacios vectoriales con dim V = n y dim W = m; y sean $A = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ y $B = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_m\}$ bases de V y W, respectivamente.

Si $T: V \to W$ es una transformación lineal, existe una y sólo una matriz $M_B^A(T)$, de m×n, tal que

$$M_B^A(T)\ (\bar{v})_A = [T(\bar{v})]_B\ ,\ \forall\ \bar{v} \in V$$

Las n columnas de dicha matriz son los vectores

$$[T(\bar{v}_1)]_B\ ,\ [T(\bar{v}_2)]_B\ , \ldots,\ [T(\bar{v}_n)]_B$$

DEMOSTRACION

Sean V y W dos espacios vectoriales con dim V = n y dim W = m, y sean $A = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ y $B = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_m\}$ dos bases cualesquiera de V y W respectivamente.

Consideremos una transformación lineal $T: V \to W$.

Si $\bar{v}$ es un vector cualquiera de V, éste puede expresarse unívocamente en términos de la base A como

$$\bar{v} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \ldots + \alpha_n\bar{v}_n \qquad ---\ (1)$$

por lo que

$$T(\bar{v}) = T(\alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \ldots + \alpha_n\bar{v}_n)$$

y como T es lineal

$$T(\bar{v}) = \alpha_1 T(\bar{v}_1) + \alpha_2 T(\bar{v}_2) + \ldots + \alpha_n T(\bar{v}_n) \qquad --- (2)$$

Por otra parte, los vectores $T(\bar{v}_1), T(\bar{v}_2), \ldots, T(\bar{v}_n)$ son elementos de W, por lo que pueden expresarse unívocamente en términos de la base B; es decir

$$\left.\begin{aligned} T(\bar{v}_1) &= a_{11}\bar{w}_1 + a_{21}\bar{w}_2 + \ldots + a_{m1}\bar{w}_m \\ T(\bar{v}_2) &= a_{12}\bar{w}_1 + a_{22}\bar{w}_2 + \ldots + a_{m2}\bar{w}_m \\ &\vdots \\ T(\bar{v}_n) &= a_{1n}\bar{w}_1 + a_{2n}\bar{w}_2 + \ldots + a_{mn}\bar{w}_m \end{aligned}\right\} \qquad --- (3)$$

Llevando estas expresiones a (2) se tiene que

$$T(\bar{v}) = \alpha_1(a_{11}\bar{w}_1 + a_{21}\bar{w}_2 + \ldots + a_{m1}\bar{w}_m) + \alpha_2(a_{12}\bar{w}_1 + a_{22}\bar{w}_2 + \ldots + a_{m2}\bar{w}_m) +$$

$$+ \ldots + \alpha_n(a_{1n}\bar{w}_1 + a_{2n}\bar{w}_2 + \ldots + a_{mn}\bar{w}_m)$$

de donde

$$T(\bar{v}) = (\alpha_1 a_{11} + \alpha_2 a_{12} + \ldots + \alpha_n a_{1n})\bar{w}_1 + (\alpha_1 a_{21} + \alpha_2 a_{22} + \ldots + \alpha_n a_{2n})\bar{w}_2 +$$

$$+ \ldots + (\alpha_1 a_{m1} + \alpha_2 a_{m2} + \ldots + \alpha_n a_{mn})\bar{w}_m$$

Por la definición IX.2.14, las sumas encerradas entre paréntesis son las coordenadas de $T(\bar{v})$ en la base B, por lo que podemos escribir

$$\left[T(\bar{v})\right]_B = \begin{bmatrix} \alpha_1 a_{11} + \alpha_2 a_{12} + \ldots + \alpha_n a_{1n} \\ \alpha_1 a_{21} + \alpha_2 a_{22} + \ldots + \alpha_n a_{2n} \\ \vdots \\ \alpha_1 a_{m1} + \alpha_2 a_{m2} + \ldots + \alpha_n a_{mn} \end{bmatrix}$$

o bien, como producto de matrices

$$[T(\bar{v})]_B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

De (1), los escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ son las coordenadas de $\bar{v}$ en la base A, por lo que

$$[T(\bar{v})]_B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} (\bar{v})_A$$

y finalmente

$$[T(\bar{v})]_B = M_B^A(T)\,(\bar{v})_A$$

donde

$$M_B^A(T) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Además, de (3)

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} = [T(\bar{v}_1)]_B \,, \quad \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} = [T(\bar{v}_2)]_B \,, \ldots, \quad \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = [T(\bar{v}_n)]_B$$

por lo que las columnas de $M_B^A(T)$ son los vectores de coordenadas, en la base B, de las imágenes de los elementos que integran la base A.

Para demostrar que esta matriz es única, supongamos otra matriz

$$N_B^A(T) = \left[b_{ij}\right]$$

tal que

$$N_B^A(T)(\bar{v})_A = \left[T(\bar{v})\right]_B$$

Como el vector de coordenadas de $T(\bar{v})$ en la base B es único, se tendrá que

$$N_B^A(T)(\bar{v})_A = M_B^A(T)(\bar{v})_A \quad , \forall\ \bar{v} \in V$$

es decir

$$\begin{aligned}
\alpha_1 a_{11} + \alpha_2 a_{12} + \ldots + \alpha_n a_{1n} &= \alpha_1 b_{11} + \alpha_2 b_{12} + \ldots + \alpha_n b_{1n} \\
\alpha_1 a_{21} + \alpha_2 a_{22} + \ldots + \alpha_n a_{2n} &= \alpha_1 b_{21} + \alpha_2 b_{22} + \ldots + \alpha_n b_{2n} \\
&\vdots \\
\alpha_1 a_{m1} + \alpha_2 a_{m2} + \ldots + \alpha_n a_{mn} &= \alpha_1 b_{m1} + \alpha_2 b_{m2} + \ldots + \alpha_n b_{mn}
\end{aligned}$$

para cualquier valor de $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Haciendo $\alpha_k = 1$ y $\alpha_j = 0 \ \forall\ j \neq k$, se obtiene

$$\begin{aligned}
a_{1k} &= b_{1k} \\
a_{2k} &= b_{2k} \\
&\vdots \\
a_{mk} &= b_{mk}
\end{aligned}$$

y al recorrer k todos los valores 1, 2,.., n, se concluye que

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

En consecuencia

$$M_B^A(T) = N_B^A(T)$$

Esto completa la demostración.

$\square$

De acuerdo con el teorema anterior, la matriz $M_B^A(T)$ nos permite calcular la imagen de un vector cualquiera $\bar{v}$ del dominio mediante el siguiente procedimiento, que podríamos considerar indirecto:

1) Determinar las coordenadas de $\bar{v}$ en la base A.
2) Multiplicar la matriz $M_B^A(T)$ por el vector $(\bar{v})_A$.
3) Obtener el vector $T(\bar{v})$ a partir de sus coordenadas en la base B.

$$\begin{tikzcd}[column sep=huge, row sep=huge]
\bar{v} \arrow[r, dashed, "\textit{aplicación de la regla}"] \arrow[d, "\textit{cálculo de coordenadas}"', "1"] & T(\bar{v}) \\
(\bar{v})_A \arrow[r, "2", "\textit{multiplicación por } M_B^A(T)"'] & {[T(\bar{v})]_B} \arrow[u, "3"', "\textit{multiplicación por escalares y adiciones}"]
\end{tikzcd}$$

El teorema X.2.1 es de importancia por diversas razones. Primeramente, porque reduce el problema de aplicar una transformación al de multiplicar una matriz por un vector, lo cual nos permite aprovechar las propiedades y procedimientos ya conocidos del álgebra de matrices, incluyendo las posibilidades de su manejo por computadora.

Otra consecuencia importante de dicho enunciado es que la matriz

$M_B^A(T)$ depende de las bases A y B. Generalmente dichas bases se eligen de manera que el cálculo de coordenadas resulte lo más sencillo posible; sin embargo, podemos también seleccionar dichas bases buscando que la matriz sea muy simple (con la mayor cantidad de ceros, por ejemplo). De este último aspecto nos ocuparemos más adelante cuando tratemos el tema de la diagonalización.

Para ilustrar los conceptos inherentes al teorema X.2.1 mediante un ejemplo, consideremos el espacio

$$V = \{ax^2 + bx + c \mid a, b, c \in R\}$$

de todos los polinomios con coeficientes reales de grado menor que tres, y el espacio

$$W = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \middle| \; a, b, c \in R \right\}$$

de matrices simétricas de orden dos con elementos en R; y sea $T: V \rightarrow W$ la transformación definida por

$$T(ax^2 + bx + c) = \begin{bmatrix} a+c & 3b \\ 3b & 2a+2c \end{bmatrix}$$

Podemos seleccionar las siguientes bases para V y W, buscando simplificar al máximo el cálculo de coordenadas

$$A = \{x^2, x, 1\}$$

y

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

Para hallar la matriz asociada a T referida a estas bases obtene -

mos

$$T(x^2) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

$$T(x) = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}$$

$$T(1) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

cuyas coordenadas en la base B son

$$[T(x^2)]_B = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

$$[T(x)]_B = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$$

$$[T(1)]_B = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

En consecuencia, la matriz asociada a T referida a las bases que elegimos es

$$M_B^A(T) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

Si queremos emplear esta matriz para obtener la imagen del vector $\overline{v} = 3x^2 - 2x + 4$, por ejemplo, hacemos lo siguiente:

1) Determinamos las coordenadas de $\overline{v}$ en la base A. En este caso, por simple inspección se obtiene

$$(\overline{v})_A = \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix}$$

2) premultiplicamos este vector por $M_B^A(T)$

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 2 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ -6 \\ 14 \end{bmatrix}$$

3) y como el resultado es $[T(\overline{v})]_B$, entonces

$$T(\overline{v}) = 7 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (-6) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 14 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & -6 \\ -6 & 14 \end{bmatrix}$$

Observe que el mismo resultado se obtiene aplicando directamente la regla al vector $\overline{v}$

$$3x^2 - 2x + 4 \xrightarrow{T} \begin{bmatrix} 7 & -6 \\ -6 & 14 \end{bmatrix}$$

Como vimos al inicio de esta sección, cuando los espacios V y W del teorema X.2.1 son del tipo R^n y A, B son las bases canónicas de V y W, la matriz $M_B^A(T)$ se denota simplemente con M(T) y se llama matriz asociada a T.

- Matriz de transición y matriz identidad.

El cambio de coordenadas de una base a otra puede efectuarse también multiplicando una matriz por un vector. Esta matriz se conoce como "matriz de transición" o "matriz de cambio de base".

Para obtener dicha matriz podemos apoyarnos en el teorema X.2.1, considerando el cambio de coordenadas como una transformación identidad y determinando su matriz asociada referida a las bases en cuestión.

De esta manera, si $A = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ y $B = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n\}$ son dos bases de un espacio vectorial V, conocemos las coordenadas de un vector $\bar{v}$ en la base A y deseamos obtener sus coordenadas en la base B; consideramos una transformación $T: V \rightarrow V$ definida como

$$T(\bar{v}) = \bar{v}, \ \forall \ \bar{v} \in V$$

y obtenemos su matriz asociada, referida a las bases A y B.

Puesto que $T(\bar{v}_i) = \bar{v}_i, \ \forall \ \bar{v}_i \in A$, las columnas de la matriz $M_B^A(T)$ son en este caso los vectores de coordenadas en la base B de los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ que forman la base A. Entonces, si multiplicamos dicha matriz por el vector $(\bar{v})_A$ se obtiene $(\bar{v})_B$, como queríamos.

Así, por ejemplo, si $A = \{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ y $B = \{\bar{w}_1, \bar{w}_2, \bar{w}_3\}$ son dos bases de un espacio V (de dimensión tres) y los vectores de A pueden expresarse en términos de los vectores de B como

$$\begin{aligned} \bar{v}_1 &= 2\bar{w}_1 + \bar{w}_2 - \bar{w}_3 \\ \bar{v}_2 &= 4\bar{w}_2 - 3\bar{w}_3 \\ \bar{v}_3 &= \bar{w}_1 + \bar{w}_3 \end{aligned}$$

entonces la matriz

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

es la matriz de transición de la base A a la base B.

De esta manera, por ejemplo, el vector

$$\bar{v} = 5\bar{v}_1 - \bar{v}_2 - 3\bar{v}_3$$

cuyas coordenadas en la base A son

$$(\bar{v})_A = (5, -1, -3)^T$$

tiene por coordenadas en la base B

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

Es decir

$$(\bar{v})_B = (7, 1, -5)^T$$

como puede verificarse fácilmente.

Respecto a la matriz de transición de una base A a una base B, conviene hacer notar que es una matriz no singular (debido a la independencia lineal del conjunto $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$), y su inversa es la matriz de transición de la base B a la base A, como el lector podrá demostrar.

Salvo el caso anterior, en el que interesa considerar dos bases distintas de un mismo espacio por tratarse de una transformación de coordenadas, cuando se tiene una transformación del tipo $T: V \rightarrow V$ lo usual es considerar una sola base para el espacio V, y la matriz $M_A^A(T)$ se llama matriz asociada a T referida a la base A.

En el caso particular de la transformación identidad se tiene la siguiente propiedad

> X.2.2 TEOREMA
>
> Si V es un espacio vectorial de dimensión n, entonces la matriz asociada a la transformación identidad $I: V \to V$, referida a cualquier base de V, es la matriz identidad I_n

DEMOSTRACION

Sea $A = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ una base de V; por X.2.1 se tiene que

$$M_A^A(T) = \left[[T(\bar{v}_1)]_A \ [T(\bar{v}_2)]_A \ \cdots \ [T(\bar{v}_n)]_A\right]$$

si $T: V \to V$ es la transformación identidad

$$\begin{aligned} T(\bar{v}_1) &= \bar{v}_1 \\ T(\bar{v}_2) &= \bar{v}_2 \\ &\vdots \\ T(\bar{v}_n) &= \bar{v}_n \end{aligned}$$

por lo que

$$\begin{aligned} [T(\bar{v}_1)]_A &= (1, 0, \ldots, 0)^T \\ [T(\bar{v}_2)]_A &= (0, 1, \ldots, 0)^T \\ &\vdots \\ [T(\bar{v}_n)]_A &= (0, 0, \ldots, 1)^T \end{aligned}$$

y en consecuencia

$$M_A^A(T) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & \dots & 1 \end{bmatrix} = I_n$$

$\square$

- Rango de la matriz asociada y dimensión del recorrido.

Regresemos al ejemplo de la transformación T: V → W definida por

$$T(ax^2 + bx + c) = \begin{bmatrix} a + c & 3b \\ 3b & 2a + 2c \end{bmatrix}$$

para la cual se obtuvo

$$M_B^A(T) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

si calculamos el rango de esta matriz se obtiene

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

$$R\left[M_B^A(T)\right] = 2.$$

Por otra parte, por simple inspección puede obtenerse en este caso el recorrido de T, que es

$$T(V) = \left\{ \begin{bmatrix} x & y \\ y & 2x \end{bmatrix} \middle| \; x, y \in R \right\}$$

por lo que

$$\dim T(V) = 2.$$

Es decir, el rango de la matriz asociada a la transformación coincidió con la dimensión del recorrido de ésta. Tal relación se verifica para cualquier transformación lineal, independientemente de las bases que se elijan para la representación matricial, como lo establece el siguiente teorema.

X.2.3 TEOREMA

Sea $T: V \to W$ una transformación lineal y sean A, B dos bases de V y W, respectivamente; entonces:

$$R\left[M_B^A(T)\right] = \dim T(V)$$

DEMOSTRACION

Sea $T: V \to W$ una transformación lineal y sean $A = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ y $B = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_m\}$ bases de V y W respectivamente.

Por IX.3.6, el rango de $M_B^A(T)$ es la dimensión del espacio columna de dicha matriz y, por X.2.1, las columnas de $M_B^A(T)$ son las coordenadas en B de los vectores $T(\bar{v}_i)$, $i = 1, 2, \ldots, n$.

Además, por X.1.6, $G = \{T(\bar{v}_1), T(\bar{v}_2), \ldots, T(\bar{v}_n)\}$ es un generador de $T(V)$ cuya dimensión es igual al número máximo de vectores linealmente independientes de G.

Para demostrar que estos dos números son iguales bastará con probar que un conjunto $\{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_r\}$ es linealmente independiente si y sólo si el conjunto $\{(\bar{u}_1)_B, (\bar{u}_2)_B, \ldots, (\bar{u}_r)_B\}$ es linealmente independiente.

En efecto

Sean $(\bar{u}_i)_B = (\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{im})^T$, con $i = 1, 2, \ldots, r$

Si el conjunto $\{(\bar{u}_1)_B, (\bar{u}_2)_B, \ldots, (\bar{u}_r)_B\}$ es linealmente independiente entonces, por IX.2.3,

$$\alpha_1 \begin{bmatrix} \lambda_{11} \\ \lambda_{12} \\ \vdots \\ \lambda_{1m} \end{bmatrix} + \alpha_2 \begin{bmatrix} \lambda_{21} \\ \lambda_{22} \\ \vdots \\ \lambda_{2m} \end{bmatrix} + \ldots + \alpha_r \begin{bmatrix} \lambda_{r1} \\ \lambda_{r2} \\ \vdots \\ \lambda_{rm} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

se satisface sólo con

$$\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$$

Igualdad entre vectores de R^m que puede expresarse como:

$$\left.\begin{array}{l} \alpha_1\lambda_{11} + \alpha_2\lambda_{21} + \ldots + \alpha_r\lambda_{r1} = 0 \\ \alpha_1\lambda_{12} + \alpha_2\lambda_{22} + \ldots + \alpha_r\lambda_{r2} = 0 \\ \vdots \\ \alpha_1\lambda_{1m} + \alpha_2\lambda_{2m} + \ldots + \alpha_r\lambda_{rm} = 0 \end{array}\right\} \quad (1)$$

Por otra parte, ya que B es base

$\beta_1\bar{w}_1 + \beta_2\bar{w}_2 + \ldots + \beta_m\bar{w}_m = \bar{0}$ implica que $\beta_1 = \beta_2 = \ldots = \beta_m = 0$

Si hacemos estas β_j, $j = 1, 2, \ldots, m$ iguales a los primeros miembros de (1) tendremos que

$$(\alpha_1\lambda_{11} + \alpha_2\lambda_{21} + \ldots + \alpha_r\lambda_{r1})\overline{w}_1 + (\alpha_1\lambda_{12} + \alpha_2\lambda_{22} + \ldots + \alpha_r\lambda_{r2})\overline{w}_2 + \ldots$$
$$\ldots + (\alpha_1\lambda_{1m} + \alpha_2\lambda_{2m} + \ldots + \alpha_r\lambda_{rm})\overline{w}_m = \overline{0}$$

Efectuando operaciones y reordenando tendremos

$$\alpha_1(\lambda_{11}\overline{w}_1 + \lambda_{12}\overline{w}_2 + \ldots + \lambda_{1m}\overline{w}_m) + \alpha_2(\lambda_{21}\overline{w}_1 + \lambda_{22}\overline{w}_2 + \ldots + \lambda_{2m}\overline{w}_m) + \ldots$$
$$\ldots + \alpha_r(\lambda_{r1}\overline{w}_1 + \lambda_{r2}\overline{w}_2 + \ldots + \lambda_{rm}\overline{w}_m) = \overline{0}$$

pero lo que está en los paréntesis son los vectores $\overline{u}_i$ expresados como combinaciones lineales de los elementos de la base B, por lo que

$$\alpha_1\overline{u}_1 + \alpha_2\overline{u}_2 + \ldots + \alpha_r\overline{u}_r = \overline{0}$$

y ya que los únicos valores posibles para las αs son

$$\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$$

de IX.2.3, el conjunto $\{\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_r\}$ es linealmente independiente.

De manera análoga puede probarse que si el conjunto $\{\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_r\}$ es linealmente independiente, entonces el conjunto $\{(\overline{u}_1)_B, (\overline{u}_2)_B, \ldots, (\overline{u}_r)_B\}$ también es linealmente independiente.

□

X.2.4 EJERCICIOS

1.- Sea la transformación lineal $T: R^2 \to R^3$ definida por

$T(x, y) = (x - y, 2x - 2y, - x + y)$

a) Obtener la matriz $M(T)$

b) Utilizando la matriz obtenida determinar la imagen bajo T de los vectores $(1, 2)$, $(0, 3)$ y $(-1, 5)$.

c) Determinar la matriz $M_B^A(T)$ para $A = \{(2, 1), (1, 0)\}$ y $B = \{(1, 0, 0), (0, -1, 0), (0, 0, -3)\}$

d) Verificar que los rangos de $M(T)$ y $M_B^A(T)$ son iguales.

2.- Sean $S: R^2 \to R^3$, $A = \{(1, 1), (0, -1)\}$,

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

$$M_B^A(S) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Obtener la regla de correspondencia de la transformación S.

3.- Sean $A = \{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ y $B = \{\bar{w}_1, \bar{w}_2, \bar{w}_3\}$ dos bases de un espacio V sobre R y tales que sus elementos están relacionados por

$$\bar{v}_1 = \bar{w}_1 + 2\bar{w}_2 + \bar{w}_3$$
$$\bar{v}_2 = 2\bar{w}_2 + \bar{w}_3$$
$$\bar{v}_3 = \bar{w}_1 - \bar{w}_2$$

Sea $T: V \to V$ una transformación tal que

$$T(\bar{v}_1) = \bar{w}_1$$
$$T(\bar{v}_2) = \bar{w}_2$$
$$T(\bar{v}_3) = \bar{w}_3$$

Obtener $M_B^A(T)$, $M_A^B(T)$, $M_A^A(T)$ y $M_B^B(T)$

4.- Sean V el espacio de las matrices cuadradas de orden 2 sobre R y $S: V \to V$ la transformación lineal definida por

$$S(A) = A + A^T, \forall A \in V$$

a) Hallar $M(S)$

b) Determinar rango de $M(S)$

c) Calcular las dimensiones del núcleo y del recorrido de S.

5.- Sean $A = \{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ y $B = \{\bar{w}_1, \bar{w}_2, \bar{w}_3\}$ dos bases de un espacio V de dimensión tres, si

$$\bar{v}_1 = 2\bar{w}_1 + \bar{w}_2 - \bar{w}_3$$
$$\bar{v}_2 = 2\bar{w}_1 + 3\bar{w}_3$$
$$\bar{v}_3 = \bar{w}_1 + \bar{w}_2 - 3\bar{w}_3$$

a) Obtener las coordenadas en B de $\bar{u}_1$ y $\bar{u}_2$ tales que

$(\bar{u}_1)_A = (1, 0, -1)^T$ y $(\bar{u}_2)_A = (0, -2, 1)^T$

b) Obtener las coordenadas en A de $\bar{u}_3$ y $\bar{u}_4$ tales que

$(\bar{u}_3)_B = (2, -1, 0)^T$ y $(\bar{u}_4)_B = (1, 1, -1)^T$

6.- Sea $T: R^3 \to R^3$ una transformación definida por

$$T(i) = 2i + 3k$$
$$T(j + k) = 2j$$
$$T(i + j) = i - j$$

donde $\{i, j, k\}$ es la base canónica de R^3

a) Determinar la regla de correspondencia de T

b) Obtener $M(T)$

c) Obtener $M_A^A(T)$ para $A = \{(-1, 0, 0), (0, 3, -1), (1, 1, 0)\}$

X.3 ALGEBRA DE LAS TRANSFORMACIONES LINEALES

Como vimos en la sección X.1, una transformación es una función cuyo dominio y codominio son espacios vectoriales. En consecuencia, las operaciones que usualmente se realizan con funciones pueden también efectuarse con transformaciones.

En esta sección nos ocuparemos de tales operaciones; en particular de la adición, la multiplicación por un escalar y la composición, haciendo énfasis en algunas de sus características para el caso de transformaciones lineales.

Empezaremos por definir la relación de igualdad entre transformaciones, la cual se desprende de la igualdad de funciones, conforme lo establece el siguiente enunciado.

> X.3.1 DEFINICION
>
> Sean S y T dos transformaciones de V en W. Se dice que S y T son iguales, lo cual se denota mediante S = T, cuando
>
> $$S(\bar{v}) = T(\bar{v}), \; \forall\, \bar{v} \in V$$

- Adición y multiplicación por un escalar.

La definición de adición y multiplicación por un escalar para el caso de transformaciones también es una consecuencia de su definición para el caso más general de funciones, como se aprecia en el siguiente enunciado

X.3.2 DEFINICION

Sean S y T dos transformaciones de V en W, y sea K el campo sobre el cual está definido el espacio W:

i) La suma de S y T es una transformación de V en W, denotada con S+T y definida por

$$(S+T)(\bar{v}) = S(\bar{v}) + T(\bar{v})\,;\ \forall\ \bar{v} \in V$$

ii) El producto de un escalar $\alpha \in K$ por la transformación S es una transformación de V en W, denotada con αS y definida por

$$(\alpha S)(\bar{v}) = \alpha \cdot S(\bar{v})\,;\ \forall\ \bar{v} \in V$$

Así, por ejemplo, para las transformaciones S y T, ambas de R^3 en R^2, definidas por

$$S(x, y, z) = (x + 3z,\ 2x - y)$$

y

$$T(x, y, z) = (x + y,\ x + y)$$

Si $\bar{v}$ es el vector de R^3

$$\bar{v} = (1, 0, -3)$$

las imágenes de dicho vector bajo las transformaciones S y T son, respectivamente

$$S(\bar{v}) = (-8, 2)$$

y

$$T(\bar{v}) = (1, 1)$$

La suma de tales vectores es la pareja ordenada (-7, 3). Este vector es, por X.3.2, la imagen de $\bar{v}$ bajo la transformación S+T; es decir

$$(S+T)(\bar{v}) = (-7,\ 3)$$

Para obtener la regla de correspondencia de S+T bastará con sumar las imágenes de un vector arbitrario del dominio bajo las transformaciones S y T; esto es:

$$(S+T)(x,y,z) = S(x,y,z) + T(x,y,z) = (x+3z,2x-y) + (x+y,x+y)$$

y la regla de correspondencia de S+T es

$$(S+T)(x,\ y,\ z) = (2x + y + 3z,\ 3x)$$

Para la misma transformación S y el mismo vector $\bar{v}$ del ejemplo anterior, la imagen de $\bar{v}$ bajo la transformación 3S es

$$(3S)(\bar{v}) = 3S(\bar{v}) = 3(-8,\ 2) = (-24,\ 6)$$

y la regla de correspondencia de dicha transformación es

$$(3S)(x,y,z) = 3S(x,y,z) = 3(x+3z,2x-y) = (3x+9z,6x-3y)$$

Respecto a la adición y a la multiplicación por un escalar, las transformaciones lineales tienen las propiedades que veremos a continuación

X.3.3 TEOREMA

Si S y T son transformaciones lineales, entonces S+T y αS también son lineales.

DEMOSTRACION

Sean $S: V \to W$ y $T: V \to W$ dos transformaciones lineales, K el campo sobre el que están definidos V y W, y sean $\bar{v}_1,\ \bar{v}_2 \in V$ y $\lambda \in K$, entonces

$$(S+T)(\lambda\bar{v}_1+\bar{v}_2) = S(\lambda\bar{v}_1+\bar{v}_2) + T(\lambda\bar{v}_1+\bar{v}_2) \qquad \text{por i) de X.3.2}$$

$$= [\lambda S(\bar{v}_1)+S(\bar{v}_2)]+[\lambda T(\bar{v}_1)+T(\bar{v}_2)] \qquad \text{por linealidad de S y de T.}$$

$$= [\lambda S(\bar{v}_1)+\lambda T(\bar{v}_1)]+[S(\bar{v}_2)+T(\bar{v}_2)] \qquad \text{por ii) y v) de IX.1.1}$$

$$= \lambda[S(\bar{v}_1)+T(\bar{v}_1)]+[S(\bar{v}_2)+T(\bar{v}_2)] \qquad \text{por vii) de IX.1.1}$$

$$(S+T)(\lambda\bar{v}_1+\bar{v}_2) = \lambda(S+T)(\bar{v}_1)+(S+T)(\bar{v}_2) \qquad \text{por i) de X.3.2}$$

y en consecuencia S+T es lineal.

La prueba de que αS es lineal para cualquier escalar α se deja al lector como ejercicio.

□

Si consideramos el conjunto de todas las transformaciones lineales que pueden definirse de un espacio V en un espacio W, conjuntamente con las operaciones de adición y multiplicación por un escalar definidas por X.3.2, tendremos un ejemplo más de espacio vectorial, como lo establece el siguiente teorema.

X.3.4 TEOREMA

Sean V y W dos espacios vectoriales sobre un campo K, y sea L(V, W) el conjunto de todas las transformaciones lineales de V en W. El conjunto L(V, W) es un espacio vectorial sobre K.

DEMOSTRACION

La cerradura de L(V, W) para la adición y la multiplicación por un escalar, postulados i) y vi) de la definición de espacio vectorial, se sigue directamente de X.3.3.

A continuación demostramos que se satisfacen los postulados iii), iv) y viii) de la misma definición y se sugiere al lector, como ejercicio, verificar que se satisfacen algunos de los postulados restantes.

iii) $\exists\ O \in L(V, W)$ tal que $O+T = T,\ \forall\ T \in L(V, W)$.

Sea $O: V \to W$ la transformación lineal definida por

$$O(\bar{v}) = \bar{0}_W,\ \forall\ \bar{v} \in V$$

(verifique que dicha transformación es lineal)

Entonces, si T es una transformación cualquiera de $L(V, W)$, $\forall\ \bar{v} \in V$ se tiene que

$$\begin{aligned} (O+T)(\bar{v}) &= O(\bar{v}) + T(\bar{v}) && \text{por i) de X.3.2} \\ &= \bar{0}_W + T(\bar{v}) && \text{por definición de O} \\ (O+T)(\bar{v}) &= T(\bar{v}) && \text{por iii) de IX.1.1} \end{aligned}$$

Finalmente, por X.3.1.

$$O+T = T$$

lo que demuestra el enunciado iii).

iv) $\forall\ T \in L(V, W)\ \exists\ -T \in L(V, W)$ tal que $-T+T = O$.

Sea $-T: V \to W$ la transformación lineal definida por

$$(-T)(\bar{v}) = -[T(\bar{v})];\ \forall\ \bar{v} \in V$$

Entonces, $\forall\ \bar{v} \in V$ se tiene que

$[-T+T](\bar{v}) = (-T)(\bar{v}) + T(\bar{v})$ por i) de X.3.2

$= -[T(\bar{v})] + T(\bar{v})$ por definición de -T

$[-T+T](\bar{v}) = \bar{0}_W$ por iv) de IX.1.1

Finalmente, por definición de O

$-T+T = O$

lo que demuestra el enunciado iv).

viii) $(\alpha + \beta)T = \alpha T + \beta T;\ \forall\ T \in L(V, W)$ y $\forall\ \alpha, \beta \in K$.

Sea $T: V \to W$ una transformación lineal cualquiera. Entonces, $\forall\ \bar{v} \in V$ y $\forall\ \alpha, \beta \in K$ se tiene que

$[(\alpha + \beta)T](\bar{v}) = (\alpha + \beta)\ T(\bar{v})$ por ii) de X.3.2

$= \alpha T(\bar{v}) + \beta T(\bar{v})$ por viii) de IX.1.1

$[(\alpha + \beta)T](\bar{v}) = (\alpha T + \beta T)(\bar{v})$ por i) de X.3.2

Finalmente, por X.3.1

$(\alpha + \beta)T = \alpha T + \beta T$

lo que demuestra el enunciado viii).

□

Veamos ahora la relación que existe entre las operaciones con transformaciones lineales y las operaciones con matrices. Para ello consideremos nuevamente las transformaciones S y T definidas por

$S(x, y, z) = (x + 3z, 2x - y)$

y

$T(x, y, z) = (x + y, x + y)$

cuya suma está definida por la regla

$$(S+T)(x, y, z) = (2x + y + 3z, 3x)$$

Si determinamos las matrices asociadas a S y a T obtendremos

$$M(S) = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 0 \end{bmatrix} \quad y \quad M(T) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

sumando estas matrices se llega a

$$M(S) + M(T) = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 0 & 0 \end{bmatrix}$$

Por otra parte, la matriz asociada a S+T es

$$M(S+T) = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 0 & 0 \end{bmatrix}$$

Como el lector puede darse cuenta, esta matriz es igual a la que se obtuvo sumando las matrices asociadas a S y a T; es decir

$$M(S+T) = M(S) + M(T) \qquad - - - (1)$$

Una situación similar ocurre con el producto por un escalar.

Por ejemplo, para la misma transformación S se obtuvo

$$(3S)(x, y, z) = (3x + 9z, 6x - 3y)$$

Cuya matriz asociada es

$$M(3S) = \begin{bmatrix} 3 & 0 & 9 \\ 6 & -3 & 0 \end{bmatrix}$$

la cual coincide con la que se obtiene al multiplicar por 3 la matriz asociada a S, puesto que

$$3\ M(S) = 3\begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 9 \\ 6 & -3 & 0 \end{bmatrix}$$

y en consecuencia

$$M(3S) = 3M(S)$$

En general, para cualquier escalar α se tiene que

$$M(\alpha S) = \alpha M(S) \qquad \text{- - - (2)}$$

Las expresiones (1) y (2) se verifican para cualquier S, T y α, como demostraremos más adelante. Esto nos permite realizar operaciones de adición y multiplicación por un escalar, para transformaciones lineales, efectuando dichas operaciones con las matrices asociadas.

Así, por ejemplo, para sumar las transformaciones S y T, de R^3 en R^3, definidas por

$$S(x, y, z) = (3x - 4y, 2y + z, -2x)$$
$$T(x, y, z) = (2y + z, 0, 3x + y + 2z)$$

podemos obtener primero sus matrices asociadas

$$M(S) = \begin{bmatrix} 3 & -4 & 0 \\ 0 & 2 & 1 \\ -2 & 0 & 0 \end{bmatrix} \quad y \quad M(T) = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 0 \\ 3 & 1 & 2 \end{bmatrix}$$

y después sumarlas

$$M(S) + M(T) = \begin{bmatrix} 3 & -2 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Por la expresión (1) esta matriz es igual a la matriz asociada a S+T; por lo que

$$M(S+T) = \begin{bmatrix} 3 & -2 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Una vez que tenemos la matriz asociada a S+T, su regla de corres - pondencia se obtiene, como ya vimos, multiplicando esta matriz por un vector arbitrario del dominio. Así se llega finalmente a

$$(S+T)(x, y, z) = (3x - 2y + z, 2y + z, x + y + 2z)$$

Para multiplicar por un escalar, por ejemplo $\frac{1}{2}$, la transformación S, bastará con multiplicar su matriz asociada

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

por dicho escalar

$$\frac{1}{2} M(S) = \begin{bmatrix} \frac{3}{2} & -2 & 0 \\ 0 & 1 & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix}$$

Entonces, la matriz asociada a $\frac{1}{2}S$ es

$$M(\tfrac{1}{2}S) = \begin{bmatrix} \frac{3}{2} & -2 & 0 \\ 0 & 1 & \frac{1}{2} \\ -1 & 0 & 0 \end{bmatrix}$$

de donde

$$(\tfrac{1}{2}S)(x, y, z) = (\tfrac{3}{2}x - 2y, y + \tfrac{z}{2}, -x)$$

Como el lector se habrá dado cuenta, las expresiones

$$M(S+T) = M(S) + M(T) \qquad \text{- - - (1)}$$

y

$$M(\alpha S) = \alpha M(S) \qquad \text{- - - (2)}$$

nos indican que el espacio de las transformaciones lineales de R^m en R^n y el espacio de las matrices de $m \times n$ son isomorfos, bajo el concepto de matriz asociada.

Desde un punto de vista más general, puede concluirse que el isomorfismo se presenta entre el espacio de las transformaciones lineales de V en W (donde V y W son espacios cualesquiera de dimensión m y n respectivamente) y el espacio de las matrices de $m \times n$; en este caso bajo el concepto de matriz asociada referida a dos bases. Al respecto se tiene el siguiente teorema

X.3.5 TEOREMA

Sean V y W dos espacios vectoriales sobre un campo K, con dim V = m y dim W = n, y sea L(V, W) el conjunto de todas las transformaciones lineales de V en W. Si A y B son bases de V y W respectivamente; entonces $\forall$ S, T ε L(V, W) y $\forall$ α ε K:

i) $M_B^A(S+T) = M_B^A(S) + M_B^A(T)$

ii) $M_B^A(\alpha S) = \alpha M_B^A(S)$

DEMOSTRACION

i) Sean S, T ε L(V, W), y $\bar{v}$ un vector cualquiera de V. Por definición de M_B^A (X.2.1) se tiene que

$$M_B^A(S+T)(\bar{v})_A = \left[(S+T)(\bar{v})\right]_B \quad \text{---} \ (1)$$

Por otra parte, por ii) de VI.3.3 sabemos que

$$\left[M_B^A(S) + M_B^A(T)\right](\bar{v})_A = M_B^A(S)(\bar{v})_A + M_B^A(T)(\bar{v})_A$$

y por definición de M_B^A (X.2.1)

$$\left[M_B^A(S) + M_B^A(T)\right](\bar{v})_A = \left[S(\bar{v})\right]_B + \left[T(\bar{v})\right]_B$$

Además, puesto que

$$(\bar{v}_1)_A + (\bar{v}_2)_A = (\bar{v}_1 + \bar{v}_2)_A; \quad \forall\ \bar{v}_1, \bar{v}_2\ \varepsilon\ V \quad \text{(demuéstrelo)}$$

se tendrá que

$$\left[M_B^A(S) + M_B^A(T)\right](\bar{v})_A = \left[S(\bar{v}) + T(\bar{v})\right]_B$$

y en consecuencia, por i) de X.3.2.

$$\left[M_B^A(S) + M_B^A(T)\right](\bar{v})_A = \left[(S+T)(\bar{v})\right]_B \qquad --- (2)$$

Entonces, por la unicidad de M_B^A establecida en X.2.1, de (1) y (2) se sigue que

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

como se quería.

La demostración de la parte ii) se deja al lector como ejercicio.

□

- Composición

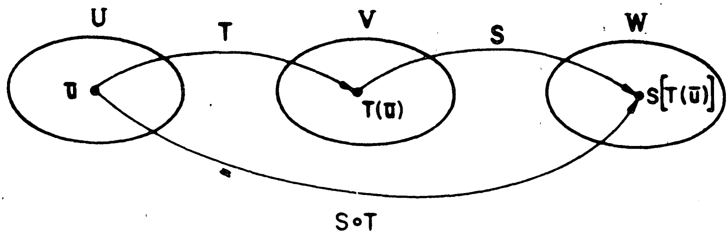
Una operación que resulta de especial interés por sus aplicaciones es la composición de transformaciones, la cual es un caso particular de la composición de funciones ya conocida por el lector.

Efectuar la composición de dos transformaciones equivale a reunir sus efectos en una sola transformación, como se describe a continuación.

Si $T: U \rightarrow V$ es una transformación y $\bar{u}$ es un vector de U; al aplicar a $\bar{u}$ la transformación T obtenemos el vector $T(\bar{u})$, que es un vector de V. Si además está definida una transformación $S: V \rightarrow W$, y aplicamos ésta al vector $T(\bar{u})$, obtendremos el vector $S\left[T(\bar{u})\right]$, que pertenece al espacio W.

A partir de ambas transformaciones puede definirse una transformación de U en W que asigne directamente al vector $\bar{u} \in U$ el vector $S\left[T(\bar{u})\right] \in W$. Dicha transformación se representa con SoT (léase "S composición T") y se conoce como la composición de S y T.

Estas ideas se representan gráficamente en la siguiente figura



De acuerdo con lo anterior, podemos establecer formalmente la definición de composición como sigue.

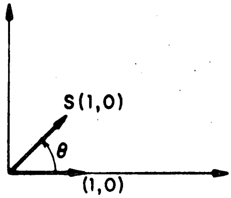
> X.3.6 DEFINICION
>
> Si $T: U \to V$ y $S: V \to W$ son dos transformaciones, $S \circ T$ es una transformación de U en W definida por
>
> $$(S \circ T)(\bar{u}) = S[T(\bar{u})];\ \forall\ \bar{u} \in U$$

Así, por ejemplo, si $T: R^3 \to R^2$ es aquella que transforma un vector cualquiera del espacio cartesiano en su proyección sobre el plano X - Y; es decir

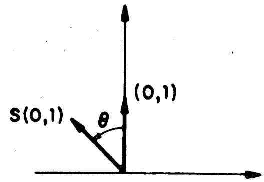
$$T(x, y, z) = (x, y);\ \forall\ (x, y, z) \in R^3$$

y $S: R^2 \to R^2$ es una transformación que consiste en girar un vector cualquiera del plano un ángulo θ en sentido positivo; entonces la transformación $S \circ T$ asocia a cada vector del espacio su proyección sobre el plano girada un ángulo θ en sentido positivo.

Puesto que la transformación S de este ejemplo es una transformación lineal, podemos obtener fácilmente su regla de correspondencia a partir de su matriz asociada. Se tiene entonces que



$S(1,\ 0) = (\cos\theta,\ \text{sen}\theta)$



$S(0,\ 1) = (-\text{sen}\theta,\ \cos\theta)$

por lo que

$$M(S) = \begin{bmatrix} \cos\theta & -\text{sen}\theta \\ \text{sen}\theta & \cos\theta \end{bmatrix}$$

de donde

$$S(x,\ y) = (x\cos\theta - y\,\text{sen}\theta,\ x\,\text{sen}\theta + y\cos\theta);\ \forall\ (x,\ y)\ \varepsilon\ R^2$$

Si ahora queremos obtener la regla de correspondencia de SoT, de la definición X.3.6 se sigue que

$$\begin{aligned}(SoT)(x,y,z) &= S[T(x,y,z)] \\ &= S(x,y) \\ (SoT)(x,y,z) &= (x\cos\theta - y\,\text{sen}\theta,\ x\,\text{sen}\theta + y\cos\theta);\ \forall\ (x,y,z)\ \varepsilon\ R^3\end{aligned}$$

Esta última expresión es la regla de correspondencia de SoT, para las transformaciones S y T del ejemplo.

Es importante hacer notar que la composición no es una operación conmutativa. En algunos casos resulta que SoT $\neq$ ToS y en algunos otros existe SoT pero no existe ToS, como en el caso del ejemplo anterior.

En efecto, para las transformaciones del ejemplo anterior se tiene, por X.3.6, que

$$(T \circ S)(x, y) = T[S(x, y)]$$

pero $S(x, y) \in R^2$ y T actúa sobre vectores de R^3, por lo que no puede obtenerse $T[S(x, y)]$ y la transformación $T \circ S$ no está definida.

En el ejemplo siguiente existen $S \circ T$ y $T \circ S$, pero son diferentes.

Sean S y T las transformaciones de R^2 en R^2 definidas por

$$S(x, y) = (-2y, 3x + y)$$

y

$$T(x, y) = (x + 2y, -2x - y)$$

Entonces

$$\begin{aligned}(S \circ T)(x, y) &= S[T(x, y)] \\ &= S(x + 2y, -2x - y) \\ &= (-2[-2x - y], + 3[x + 2y] + [-2x - y]) \\ (S \circ T)(x, y) &= (4x + 2y, x + 5y)\end{aligned}$$

por otra parte

$$\begin{aligned}(T \circ S)(x, y) &= T[S(x, y)] \\ &= T(-2y, 3x + y) \\ &= (-2y + 2[3x + y], -2[-2y] - [3x + y]) \\ (T \circ S)(x, y) &= (6x, -3x + 3y)\end{aligned}$$

y en consecuencia

$$S \circ T \neq T \circ S.$$

como se había dicho.

La composición de dos transformaciones lineales siempre es una transformación lineal, como lo establece el siguiente enunciado

> **X.3.7 TEOREMA**
>
> Si $T: U \rightarrow V$ y $S: V \rightarrow W$ son transformaciones lineales, entonces $S \circ T$ es una transformación lineal.

DEMOSTRACION

Sean $T: U \rightarrow V$ y $S: V \rightarrow W$ dos transformaciones lineales, $\bar{u}_1$, $\bar{u}_2$ dos vectores cualesquiera de U y α un escalar arbitrario del campo sobre el que están definidos los espacios U, V y W. Entonces:

$$\begin{aligned}
(S \circ T)(\alpha\bar{u}_1 + \bar{u}_2) &= S\left[T(\alpha\bar{u}_1 + \bar{u}_2)\right] && \text{por X.3.6} \\
&= S\left[\alpha T(\bar{u}_1) + T(\bar{u}_2)\right] && \text{por linealidad de T} \\
&= \alpha S\left[T(\bar{u}_1)\right] + S\left[T(\bar{u}_2)\right] && \text{por linealidad de S} \\
(S \circ T)(\alpha\bar{u}_1 + \bar{u}_2) &= \alpha\left[(S \circ T)(\bar{u}_1)\right] + (S \circ T)(\bar{u}_2) && \text{por X.3.6}
\end{aligned}$$

y en consecuencia, de X.1.3, $S \circ T$ es lineal. □

Para introducir la relación entre la composición de transformaciones y las operaciones con matrices, consideremos nuevamente las transformaciones S y T del ejemplo anterior, definidas por

$$S(x, y) = (-2y,\ 3x + y)$$

y

$$T(x, y) = (x + 2y,\ -2x - y)$$

para las cuales se obtuvo

$$(S \circ T)(x, y) = (4x + 2y,\ x + 5y)$$

Si determinamos las matrices asociadas a S y a T obtendremos

$$M(S) = \begin{bmatrix} 0 & -2 \\ 3 & 1 \end{bmatrix} \quad y \quad M(T) = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$$

Multiplicando estas matrices se llega a

$$M(S)\ M(T) = \begin{bmatrix} 4 & 2 \\ 1 & 5 \end{bmatrix}$$

Por otra parte, al obtener la matriz asociada a $S \circ T$ encontramos

$$M(S \circ T) = \begin{bmatrix} 4 & 2 \\ 1 & 5 \end{bmatrix}$$

por lo que

$$M(S \circ T) = M(S)\ M(T)$$

Para estas mismas transformaciones, el lector puede verificar que

$$M(T \circ S) = M(T)\ M(S)$$

Sin embargo, también se tiene que

$$M(S \circ T) \neq M(T \circ S)$$

para las transformaciones de este ejemplo.

En general, la composición de transformaciones lineales es equivalente a la multiplicación de matrices asociadas, como lo establece el siguiente enunciado cuya demostración se deja al lector como ejercicio.

X.3.8 **TEOREMA**

Si $T: U \to V$ y $S: V \to W$ son transformaciones lineales y A, B, C son bases de U, V y W respectivamente, entonces

$$M_C^A(S \circ T) = M_C^B(S)\ M_B^A(T)$$

De esta manera, las propiedades de las operaciones con transformaciones lineales son similares a las de las operaciones con matrices. En particular, las propiedades relativas a la composición se enuncian en el siguiente teorema

X.3.9 **TEOREMA**

Sean U, V, W y X espacios vectoriales sobre un campo K; y F, G, H, S, T transformaciones lineales cualesquiera entre los espacios que se indica

$F: U \to V$

$G: U \to V$

$H: W \to X$

$S: V \to W$

$T: V \to W$

entonces:

i) $S \circ (F+G) = S \circ F + S \circ G$

ii) $(S+T) \circ F = S \circ F + T \circ F$

iii) $\alpha(S \circ F) = (\alpha S) \circ F = S \circ (\alpha F),\ \forall\ \alpha \in K$

iv) $H \circ (S \circ F) = (H \circ S) \circ F$

v) $T \circ I_V = T,\ I_W \circ T = T$

donde I_V e I_W son las transformaciones identidad en los espacios V y W respectivamente.

DEMOSTRACION

Demostraremos los incisos i) y iv), dejando al lector como ejercicio la demostración de los restantes.

i) Sean $S: V \to W$, $F: U \to V$ y $G: U \to V$, tres transformaciones lineales, y $\bar{u} \in U$, entonces

$$[S\circ(F+G)](\bar{u}) = S[(F+G)(\bar{u})] \qquad \text{por X.3.6}$$

$$= S[F(\bar{u}) + G(\bar{u})] \qquad \text{por i) de X.3.2}$$

$$= S[F(\bar{u})] + S[G(\bar{u})] \qquad \text{por linealidad de S}$$

$$= (S\circ F)(\bar{u}) + (S\circ G)(\bar{u}) \qquad \text{por X.3.6}$$

$$[S\circ(F+G)](\bar{u}) = [S\circ F + S\circ G](\bar{u}) \qquad \text{por i) de X.3.2}$$

y finalmente por X.3.1

$$S\circ(F+G) = S\circ F + S\circ G$$

iv) Sean $F: U \to V$, $S: V \to W$, $H: W \to X$ tres transformaciones lineales, y $\bar{u} \in U$, entonces

$$[H\circ(S\circ F)](\bar{u}) = H[(S\circ F)\bar{u}] \qquad \text{por X.3.6}$$

$$= H(S[F(\bar{u})]) \qquad \text{por X.3.6}$$

$$= (H\circ S)[F(\bar{u})] \qquad \text{por X.3.6}$$

$$[H\circ(S\circ F)](\bar{u}) = [(H\circ S)\circ F](\bar{u}) \qquad \text{por X.3.6}$$

y por X.3.1

$$H\circ(S\circ F) = (H\circ S)\circ F$$

como queríamos.

□

- Inversa de una transformación.

Un problema que se presenta con frecuencia en las aplicaciones consiste en encontrar un vector cuando de él se conoce su imagen bajo determinada transformación.

Como ejemplo de este tipo de problema, busquemos un vector $\bar{v}$ cuya imagen bajo la transformación $T: R^2 \rightarrow R^2$ definida por

$$T(x, y) = (3y, y)$$

sea el vector

$$T(\bar{v}) = (6, 2)$$

Para dicha transformación, existen varios vectores cuya imagen bajo T es el vector (6, 2). Por citar algunos tenemos los siguientes

$$(1, 2) \xrightarrow{T} (6, 2)$$
$$(-3, 2) \xrightarrow{T} (6, 2)$$
$$(0, 2) \xrightarrow{T} (6, 2), \text{ etc.}$$

por lo que en este caso nuestro problema no tiene solución única.

Sin embargo, si la transformación en cuestión consiste en girar un vector del plano un ángulo θ, digamos 60°, en sentido positivo; para determinar el vector $\bar{v}$ cuya imagen sea un vector conocido $\bar{w}$, bastará con girar el vector $\bar{w}$ un ángulo de 60° en el sentido negativo, lo cual nos conduce a un único vector $\bar{v}$.

Si llamamos T a la transformación que gira al vector 60° en sentido positivo y H a la que lo gira 60° en sentido negativo, tendremos que H nos permite "regresar" cualquier vector del recorrido de

T al vector del dominio del cual procede.

En otras palabras, la transformación H "cancela" los efectos producidos por la transformación T. También T cancela los efectos producidos por H.

A la transformación H se le conoce como inversa de T y se le representa con T^{-1}.

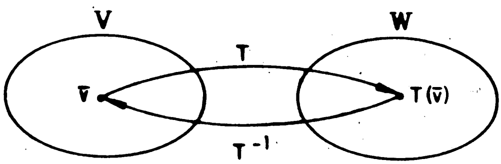
En general, si $T: V \rightarrow W$ es una transformación, se llama inversa de T y se representa con T^{-1} a una transformación de W en V tal que

$$T^{-1}[T(\bar{v})] = \bar{v}, \ \forall \ \bar{v} \in V \qquad - - - (1)$$

y

$$T[T^{-1}(\bar{w})] = \bar{w}, \ \forall \ \bar{w} \in W \qquad - - - (2)$$

Una representación gráfica para este concepto es la siguiente



Empleando la definición de composición, las expresiones (1) y (2) pueden escribirse de la siguiente manera

$$(T^{-1} \circ T)(\bar{v}) = \bar{v}; \ \forall \ \bar{v} \in V$$

$$(T \circ T^{-1})(\bar{w}) = \bar{w}; \ \forall \ \bar{w} \in W$$

Considerando la definición de transformación identidad, dichas expresiones son equivalentes a

$$T^{-1} \circ T = I_v$$

y

$$T \circ T^{-1} = I_w$$

por lo que el concepto de transformación inversa puede también definirse, en términos más "algebraicos", de la siguiente manera.

X.3.10 DEFINICION

Si $T: V \to W$ es una transformación, se llama inversa de T a una transformación $T^{-1}: W \to V$ tal que

i) $T^{-1} \circ T = I_v$

y

ii) $T \circ T^{-1} = I_w$

donde I_v e I_w son las transformaciones identidad en V y en W respectivamente.

Cabe comentar que algunos autores definen como "inversa por la izquierda" de T a una transformación T^{-1} que satisface la expresión i), y como "inversa por la derecha" a una que satisface la expresión ii).

Sin embargo, se puede demostrar (y se sugiere al lector hacerlo) que una inversa por la izquierda es también inversa por la derecha y, en consecuencia, es una inversa en los términos de la definición X.3.10.

Con relación al concepto de inversa surgen de inmediato dos preguntas

- ¿Todas las transformaciones tienen inversa?
- ¿Es única la inversa de una transformación?

La respuesta a la primera pregunta es negativa, como se sigue de los ejemplos anteriores.

En efecto, para la transformación $T: R^2 \to R^2$ definida por

$$T(x, y) = (3y, y)$$

es imposible definir una transformación T^{-1} tal que

$$T^{-1}[T(x, y)] = (x, y)$$

puesto que existen varios vectores del dominio con una misma imagen bajo T.

Por ejemplo, ¿cuál sería la imagen bajo T^{-1} del vector (6, 2)?

$$\begin{array}{cccc} & \overset{T^{-1}}{\nearrow} & (1, 2) & \\ (6, 2) & \longrightarrow & (-3, 2) & ? \\ & \searrow & (0, 2) & \\ & \vdots & & \end{array}$$

Podría ser cualquiera de los vectores indicados en el esquema anterior, lo cual no es permitido para una transformación.

Otra razón por la que no puede definirse en este caso una transformación $T^{-1}: R^2 \to R^2$, es que existen vectores del codominio que no son imagen bajo T de vector alguno del dominio.

Por ejemplo, el vector (1, 1) pertenece a R^2 pero no pertenece al recorrido de T, por lo que no se le puede asignar imagen bajo T^{-1}.

$$(1, 1) \xrightarrow{T^{-1}} ?$$

Las dos situaciones por las cuales la transformación anterior care ce de inversa no se presentan en el caso de la transformación que gira un vector 60° en sentido positivo, puesto que, en este caso:

1) Todos los vectores del dominio tienen diferentes imágenes bajo la transformación, por lo que se dice que ésta es "uno a uno".

2) Todos los vectores del codominio son imagen de algún vector del dominio, bajo la transformación en cuestión, por lo que se dice que ésta es "sobre".

Como veremos a continuación, tales características constituyen una condición necesaria y suficiente para la existencia de la inversa de una transformación.

X.3.11 DEFINICION

Sea $T; V \rightarrow W$ una transformación, se dice que

i) T es uno a uno si $T(\bar{v}_1) = T(\bar{v}_2) \Rightarrow \bar{v}_1 = \bar{v}_2$; $\forall\ \bar{v}_1, \bar{v}_2 \in V$

ii) T es sobre si $T(V) = W$

iii) T es biyectiva si es uno a uno y es sobre

X.3.12 TEOREMA

Sea $T: V \rightarrow W$ una transformación. T^{-1} existe si y sólo si T es biyectiva

DEMOSTRACION

Sea T una transformación biyectiva

Si $\bar{w} \in T(V)$, por i) de X.3.11, existe un único vector $\bar{v} \in V$ tal que

$T(\bar{v}) = \bar{w}$

Sea S una transformación definida por

$S(\bar{w}) = \bar{v}$

entonces, $S[T(\bar{v})] = \bar{v}$ — de sustituir $\bar{w}$ por $T(\bar{v})$

$(S \circ T)(\bar{v}) = \bar{v}$ — por X.3.6

es decir, $S \circ T = I_V$

Además, por ii) de X.3.11,

$\forall\ \bar{w} \in W$ existe $\bar{v} \in V$ tal que

$T(\bar{v}) = \bar{w}$

$T[S(\bar{w})] = \bar{w}$ — por definición de S

$(T \circ S)(\bar{w}) = \bar{w}$ — por X.3.6

esto es, $T \circ S = I_W$

Por lo tanto, de X.3.10, $S = T^{-1}$ y la inversa existe.

Ahora demostraremos que si T^{-1} existe, T es biyectiva

Sean $\bar{v}_1$, $\bar{v}_2 \in V$ tales que, $T(\bar{v}_1) = T(\bar{v}_2)$

$(T^{-1} \circ T)(\bar{v}_1) = T^{-1}[T(\bar{v}_1)]$ — por X.3.6

$(T^{-1} \circ T)(\bar{v}_1) = \bar{v}_1$ — por i) de X.3.10

por lo tanto $T^{-1}[T(\bar{v}_1)] = \bar{v}_1$ — - - - (1)

$(T^{-1} \circ T)(\bar{v}_2) = T^{-1}[T(\bar{v}_2)]$ — por X.3.6.

$(T^{-1} \circ T)(\bar{v}_2) = \bar{v}_2$ — por i) de X.3.10

luego $T^{-1}[T(\bar{v}_2)] = \bar{v}_2$ --- (2)

De (1) y (2) $T(\bar{v}_1) = T(\bar{v}_2)$ implica $\bar{v}_1 = \bar{v}_2$

y T es uno a uno

Sea ahora $\bar{w} \varepsilon W$

$(T \circ T^{-1})(\bar{w}) = \bar{w}$ por ii) de X.3.10

$T[T^{-1}(\bar{w})] = \bar{w}$ por X.3.6

es decir, $\forall \bar{w} \varepsilon W$ existe un $T^{-1}(\bar{w}) = \bar{v} \varepsilon V$ tal que

$T(\bar{v}) = \bar{w}$

luego $T(V) = W$ y T es sobre

Por lo tanto, de iii) de X.3.11, T es biyectiva y la demostración termina. □

Con relación a la unicidad de la inversa se tiene que, cuando T^{-1} existe, dicha transformación es única. Esta propiedad se establece, junto con algunas otras, en el siguiente teorema.

X.3.13 TEOREMA

Si $F: U \rightarrow V$ y $T: V \rightarrow W$ son dos transformaciones biyectivas, y λ es un escalar del campo sobre el que están definidos V y W, entonces:

i) T^{-1} es única

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

iii) $(T \circ F)^{-1} = F^{-1} \circ T^{-1}$

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

DEMOSTRACION

Se demuestran a continuación las propiedades i) y iii)

i) Sean H y M dos inversas de T. Entonces, por i) de X.3.10

$$H \circ T = I_V \quad y \quad M \circ T = I_V$$

En consecuencia

$$H \circ T = M \circ T$$

por lo que

$$(H \circ T) \circ H = (M \circ T) \circ H$$

Entonces, por iv) de X.3.9

$$H \circ (T \circ H) = M \circ (T \circ M)$$

y por ii) de X.3.10

$$H \circ I_W = M \circ I_W$$

Finalmente, por v) de X.3.9

$$H = M$$

y la inversa es única.

iii) $(F^{-1}\circ T^{-1})\circ(T\circ F) = [(F^{-1}\circ T^{-1})\circ T]\circ F$ por iv) de X.3.9

$= [F^{-1}\circ(T^{-1}\circ T)]\circ F$ por iv) de X.3.9

$= (F^{-1}\circ I_V)\circ F$ por i) de X.3.10

$= F^{-1}\circ F$ por v) de X.3.9

$(F^{-1}\circ T^{-1})\circ(T\circ F) = I_V$ por i) de X.3.10

En consecuencia, $F^{-1}\circ T^{-1}$ es la inversa de $T\circ F$, como se pretendía demostrar.

□

Como el lector se habrá dado cuenta, lo que hasta ahora hemos visto con relación a la inversa se refiere a transformaciones cualesquiera, y los enunciados son aplicables incluso al caso más general de funciones. En lo que sigue aplicaremos estas ideas al caso particular de las transformaciones lineales.

X.3.14 TEOREMA

Sea $T: V \rightarrow W$ una transformación lineal. Si T^{-1} existe entonces es una transformación lineal

DEMOSTRACION

Sean $\bar{v}_1, \bar{v}_2 \in V$ y $\bar{w}_1, \bar{w}_2 \in W$ tales que

$$T(\bar{v}_1) = \bar{w}_1 \qquad --- (1)$$

$$T(\bar{v}_2) = \bar{w}_2 \qquad --- (2)$$

$$(T^{-1} \circ T)(\bar{v}_1) = \bar{v}_1 \qquad \text{por i) de X.3.10}$$

$$T^{-1}[T(\bar{v}_1)] = \bar{v}_1 \qquad \text{por X.3.6}$$

sustituyendo (1) es la última igualdad, tenemos

$$T^{-1}(\bar{w}_1) = \bar{v}_1 \qquad --- (3)$$

de manera semejante

$$T^{-1}(\bar{w}_2) = \bar{v}_2 \qquad --- (4)$$

Sean K el campo sobre el que están definidos V y W, y $\alpha \in K$, entonces

$$\alpha\bar{v}_1 + \bar{v}_2 = (T^{-1} \circ T)(\alpha\bar{v}_1 + \bar{v}_2) \qquad \text{por i) de X.3.10}$$

$$= T^{-1}[T(\alpha\bar{v}_1 + \bar{v}_2)] \qquad \text{por X.3.6}$$

$$= T^{-1}[\alpha T(\bar{v}_1) + T(\bar{v}_2)] \qquad \text{por linealidad de T}$$

$$\alpha\bar{v}_1 + \bar{v}_2 = T^{-1}(\alpha\bar{w}_1 + \bar{w}_2) \qquad \text{de sustituir (1) y (2)}$$

Sustituyendo (3) y (4) en el primer miembro tenemos

$$\alpha T^{-1}(\bar{w}_1) + T^{-1}(\bar{w}_2) = T^{-1}(\alpha\bar{w}_1 + \bar{w}_2)$$

y de X.1.3, T^{-1} es lineal, como queríamos.

□

En el caso particular de una transformación lineal cuyo dominio es un espacio de dimensión finita, la condición necesaria y suficiente para la existencia de la inversa puede enunciarse de la siguiente manera.

X.3.15 TEOREMA

Sean V un espacio de dimensión finita y $T: V \to W$ una transformación lineal. T^{-1} existe si y sólo si dim V = dim W y $N(T) = \{\overline{0}_v\}$

DEMOSTRACION

Sean dim V = dim W, $N(T) = \{\overline{0}_v\}$, y $\overline{v}_1, \overline{v}_2 \in V$ tales que

$$T(\overline{v}_1) = T(\overline{v}_2)$$

es decir, $T(\overline{v}_1) - T(\overline{v}_2) = \overline{0}_w$

por linealidad de T, $T(\overline{v}_1 - \overline{v}_2) = \overline{0}_w$

luego, por ii) de X.1.2, $\overline{v}_1 - \overline{v}_2 \in N(T)$

ya que $N(T) = \{\overline{0}_v\}$, $\overline{v}_1 - \overline{v}_2 = \overline{0}_v$

esto es, $\overline{v}_1 = \overline{v}_2$

por lo tanto, $T(\overline{v}_1) = T(\overline{v}_2)$ implica $\overline{v}_1 = \overline{v}_2$

y T es uno a uno.

Además, de IX.2.11, dim N(T) = 0, luego, de X.1.7,

$$\dim V = \dim T(V)$$

pero, dim V = dim W, por hipótesis

entonces, dim W = dim T(V)

y, por i) de X.1.5, $W = T(V)$

es decir, T es sobre y, de X.3.12, T^{-1} existe.

Ahora supongamos que T^{-1} existe entonces,

T es biyectiva	por X.3.12
$N(T) = \{0_v\}$	por X.1.4 y i) de X.3.11
$\dim W = \dim T(V)$	por ii) de X.3.11
$\dim V = \dim W$	por X.1.7 y IX.2.11

y la demostración termina.

□

Observe que las condiciones que establece el teorema X.3.15 para la existencia de la inversa son equivalentes a que la matriz asociada a la transformación sea cuadrada y no singular.

En otras palabras, la transformación T tiene inversa si y sólo si la matriz $M_B^A(T)$ tiene inversa.

Supongamos ahora que $T: V \to W$ es una transformación lineal para la cual existe su inversa $T^{-1}: W \to V$, que $\dim V = n$ y que A, B son bases de V y W respectivamente.

Entonces, por i) de X.3.10

$$T^{-1} \circ T = I_V$$

de donde

$$M_A^A(T^{-1} \circ T) = M_A^A(I_v)$$

pero, por X.2.2 la matriz $M_A^A(I_v)$ es la matriz identidad I_n, por lo que

$$M_A^A(T^{-1} \circ T) = I_n$$

Por otra parte, de X.3.8 se tiene que

$$M_A^A(T^{-1} \circ T) = M_A^B(T^{-1})\ M_B^A(T)$$

y en consecuencia

$$M_A^B(T^{-1})\ M_B^A(T) = I_n$$

por lo que $M_A^B(T^{-1})$ es la matriz inversa de $M_B^A(T)$, es decir

$$M_A^B(T^{-1}) = [M_B^A(T)]^{-1}$$

De esta manera, la matriz asociada a T^{-1} (referida a las bases B y A) puede obtenerse invirtiendo la matriz asociada a T (referida a las bases A y B).

Con lo anterior hemos demostrado el siguiente teorema

X.3.16 TEOREMA

Sean $T: V \to W$ una transformación lineal, V un espacio de dimensión finita y A, B bases de V y W respectivamente:

i) T^{-1} existe si y sólo si $M_B^A(T)$ es no singular

ii) Si T^{-1} existe, entonces $M_A^B(T^{-1}) = [M_B^A(T)]^{-1}$

Apliquemos ahora este teorema a las transformaciones de los dos ejemplos que hemos visto en este apartado.

Como en ambos casos se trata de transformaciones de R^2 en R^2, utilizaremos, por sencillez, la base canónica de R^2 para el dominio y el codominio.

Para la transformación definida por

$$T(x, y) = (3y, y)$$

su matriz asociada es

$$M(T) = \begin{bmatrix} 0 & 3 \\ 0 & 1 \end{bmatrix}$$

la cual es, claramente, una matriz singular.

En consecuencia, por i) de X.3.16, no existe T^{-1}.

Para la transformación S que consiste en girar un vector del plano un ángulo θ en sentido positivo, encontramos que

$$M(S) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

En particular, si $\theta = 60°$ se tiene

$$M(S) = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

Para determinar si esta matriz es no singular y obtener su inversa, procedemos como en la sección VI.4

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

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

Concluimos entonces que M(S) es no singular y su inversa es

$$[M(S)]^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ & \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

En consecuencia, por i) de X.3.16 existe S^{-1}, y por ii) de X.3.16

$$M(S^{-1}) = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ & \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

de donde se obtiene

$$S^{-1}(x,\ y) = \left(\frac{x + \sqrt{3}y}{2}\ ,\ \frac{-\sqrt{3}x\ +\ y}{2}\right)$$

que es la regla de correspondencia de S^{-1}.

Finalmente aplicaremos el teorema anterior a un ejemplo en el que no se tengan bases canónicas.

Para ello consideremos una transformación T del espacio vectorial real

$$V = \{ax^2 + bx + c \mid a,\ b,\ c \in R\}$$

en el espacio vectorial real

$$W = \left\{ \begin{bmatrix} a & b \\ & \\ b & c \end{bmatrix} \middle| \ a,\ b,\ c \in R \right\}$$

definida por

$$T(ax^2 + bx + c) = \begin{bmatrix} a + c & 3b \\ & \\ 3b & a - c \end{bmatrix}$$

y consideremos las siguientes bases para V y W respectivamente

$A = \{x^2, x, 1\}$

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

La matriz asociada a T, referida a estas bases, resulta ser

$$M_B^A(T) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

Para determinar si esta matriz es no singular, y obtener su inversa, procedemos como sigue

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

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

se tiene entonces que $M_B^A(T)$ es no singular y que su inversa es

$$[M_B^A(T)]^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}$$

En consecuencia, por i) de X.3.16 existe T^{-1} y por ii) de X.3.16 sabemos que

$$M_A^B(T^{-1}) = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}$$

Para obtener la regla de correspondencia de T^{-1} a partir de esta matriz, debemos calcular las coordenadas de una matriz arbitraria de W

$$\bar{w} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

en la base B, las cuales son

$$(\bar{w})_B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Entonces, por X.2.1 se tiene que

$$[T^{-1}(\bar{w})]_A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \frac{a+c}{2} \\ \frac{b}{3} \\ \frac{a-c}{2} \end{bmatrix}$$

por lo que la regla de correspondencia de T^{-1} es

$$T^{-1}\left(\begin{bmatrix} a & b \\ b & c \end{bmatrix}\right) = \frac{a+c}{2}x^2 + \frac{b}{3}x + \frac{a-c}{2}$$

Para verificar que ésta es la regla de correspondencia de T^{-1}, apliquemos sucesivamente las transformaciones T y T^{-1} a un vector arbitrario del espacio V. Se obtiene entonces que

$$ax^2 + bx + c \xrightarrow{T} \begin{bmatrix} a+c & 3b \\ 3b & a-c \end{bmatrix} \xrightarrow{T^{-1}} ax^2 + bx + c$$

como cabía esperar.

X.3.17 EJERCICIOS

1.- Sean V y W dos espacios vectoriales sobre un campo K y L(V, W) el conjunto de todas las transformaciones lineales de V en W. Demostrar que $\forall$ T, S ε L(V, W) y $\forall$ α, β ε K

i) $\alpha(T+S) = \alpha T + \alpha S$

ii) $\alpha(\beta T) = (\alpha\beta)T$

2.- Sean S: V $\to$ W, T: V $\to$ W dos transformaciones lineales, A y B bases de V y W respectivamente y K el campo sobre el que están definidos V y W.

Demostrar que

a) $M_B^A(\alpha S) = \alpha M_B^A(S)$, $\forall\ \alpha \ \varepsilon\ K$; b) $M_B^A(S) = M_B^A(T) \Rightarrow S = T$

3.- Para las siguientes transformaciones lineales

S: $R^3 \to R^2$ tal que $S(x, y, z) = (x + y, y + z)$

T: $R^2 \to R^2$ tal que $T(x, y) = (x + y, x - y)$

U: $R^3 \to R^3$ tal que $U(x, y, z) = (x, y, 0)$

V: $R^2 \to R^2$ tal que $V(x, y) = (-x, x + y)$

Obtener, si existen

a) S+T; b) 2T-V; c) S∘U; d) U∘S; e) (S∘U) + 3S

4.- Sean $T: U \to V$, $S: V \to W$ dos transformaciones lineales y A, B, C bases de U, V y W, respectivamente.

Demostrar que

$$M_C^A(S \circ T) = M_C^B(S)\ M_B^A(T)$$

5.- Sea T la transformación que aplicada a un vector del plano lo transforma en su simétrico respecto al eje y; y sea S tal que

$$M_B^A(S) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$ para $A = \{(1,1),(0,-1)\}$ y $B=\{(1,0,1),(0,1,1),(1,1,0)\}$

Hallar la regla de correspondencia de $S \circ T$.

6.- Sean U, V, W espacios vectoriales sobre un campo K, I_V e I_W las transformaciones identidad en V y W, respectivamente, y $S: U \to V$, $T: V \to W$ dos transformaciones lineales. Demostrar que

i) $\alpha(T \circ S) = (\alpha T) \circ S = T \circ (\alpha S), \forall\ \alpha \in K$

ii) $T \circ I_V = T$, $I_W \circ T = T$

7.- Para cada una de las siguientes transformaciones

$T(x, y, z) = (x + z, x - z, y)$

$S(x, y, z) = (x + y, z - x, y + z)$

$H = U \circ U$, donde $U(x, y, z) = (0, x, y)$

$D(ax^3 + bx^2 + cx) = 3ax^2 + 2bx + c$

Determinar si existe o no su inversa y en caso afirmativo obtenerla.

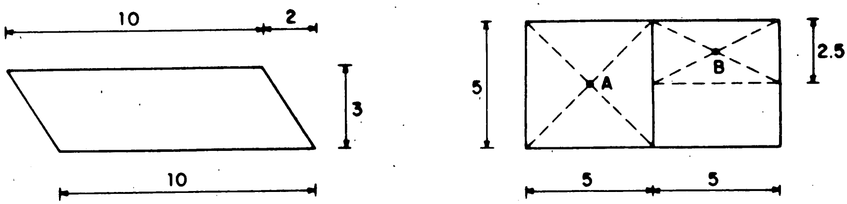
8.- Sea $T: V \to W$ una transformación biyectiva y λ un escalar del campo sobre el que están definidos V y W.

Demostrar que

i) $(T^{-1})^{-1} = T$

ii) $(\lambda T)^{-1} = \lambda^{-1}\, T^{-1}$, si $\lambda \neq 0$

9.- La figura de la derecha puede obtenerse a partir de la figura de la izquierda mediante una transformación lineal:



a) Obtener la regla de correspondencia de dicha transformación.

b) Investigar si es posible localizar en la figura de la izquierda los puntos cuyas imágenes bajo la transformación son los A y B señalados en la figura de la derecha. De ser posible, localizarlos

X.4 VALORES Y VECTORES CARACTERISTICOS

En esta sección trataremos con transformaciones de un espacio vectorial en si mismo; esto es, transformaciones del tipo

$$T: V \to V$$

a las que se conoce con el nombre de "operadores".

Para este tipo de transformaciones pueden haber vectores que no se modifiquen al aplicar la transformación, o cuya modificación consista únicamente en quedar multiplicados por un escalar. Es decir, pueden existir vectores tales que

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

donde λ es un escalar.

Geométricamente, si $V = R^n$ esto significa que dichos vectores no cambian de dirección al aplicar la transformación, sino que sólo cambian de tamaño (si $\lambda \neq 1$) o de sentido (si $\lambda < 0$).

A tales vectores se les llama "vectores característicos" del operador T, y a los escalares se les conoce como "valores característicos" de dicho operador.

Se tienen varios sinónimos para estos conceptos y entre los más comunes podemos citar los siguientes: eigenvector y eigenvalor, vector propio y valor propio, autovector y autovalor, etc.

Los valores y vectores característicos nos permiten encontrar propiedades "intrínsecas" de los operadores; esto es, propiedades que son independientes del sistema de referencia que se elija. Además constituyen una valiosa herramienta para la solución de diversos problemas de aplicación.

En la presente sección nos ocuparemos primero de su definición y características fundamentales, así como de un procedimiento para obtenerlos en el caso de espacios de dimensión finita. Posteriormente los aplicaremos al análisis de la existencia y obtención de una representación matricial diagonal.

4 - Valores y vectores característicos. Definición y propiedades.

Para introducir los conceptos de valor y vector característico mediante un ejemplo, consideremos el operador $T: R^2 \to R^2$ definido por

$$T(x, y) = (2x + y, 6x + y)$$

Si aplicamos dicho operador al vector

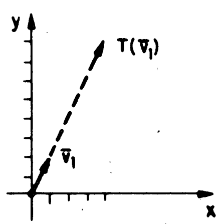
$$\bar{v}_1 = (1, 2)$$

encontramos que

$$T(\bar{v}_1) = T(1, 2) = (4, 8) = 4(1, 2) = 4\bar{v}_1$$

Esto es; la transformación solamente lo multiplicó por 4.

Geométricamente podemos decir que no le cambió la dirección sino únicamente su tamaño, el cual aumentó cuatro veces, como se aprecia en la siguiente figura



Se dice entonces que el número 4 es un valor característico de este operador y que el vector $\bar{v}_1 = (1, 2)$ es un vector característico de T correspondiente al valor 4.

Si aplicamos ahora el operador T a los vectores

$$\bar{v}_2 = (3, 6)$$

y

$$\bar{v}_3 = (-2, -4)$$

encontramos que

$$T(\bar{v}_2) = T(3, 6) = (12, 24) = 4(3, 6)$$

y

$$T(\bar{v}_3) = T(-2, -4) = (-8, -16) = 4(-2, -4)$$

por lo que $\bar{v}_2$ y $\bar{v}_3$ también son vectores característicos de T correspondientes al valor 4.

Sin embargo, no todos los vectores característicos corresponden al mismo valor. Por ejemplo, para los vectores

$$\bar{v}_4 = (1, -3)$$

y

$$\bar{v}_5 = (-2, 6)$$

se tiene

$$T(\bar{v}_4) = T(1, -3) = (-1, 3) = (-1)(1, -3) = (-1)\bar{v}_4$$

y

$$T(\bar{v}_5) = T(-2, 6) = (2, -6) = (-1)(-2, 6) = (-1)\bar{v}_5$$

por lo que $\bar{v}_4$ y $\bar{v}_5$ son vectores característicos de T correspondientes al valor característico -1.

Es fácil darse cuenta que no todos los vectores del dominio son vectores característicos del operador, ya que en general éste modifica tanto la magnitud como la dirección de los vectores.

Esto último sucede, por ejemplo, con los vectores

$$\bar{v}_6 = (1,\ 3)$$

$$\bar{v}_7 = (-1,\ 5)$$

$$\bar{v}_8 = (-1,\ 0)$$

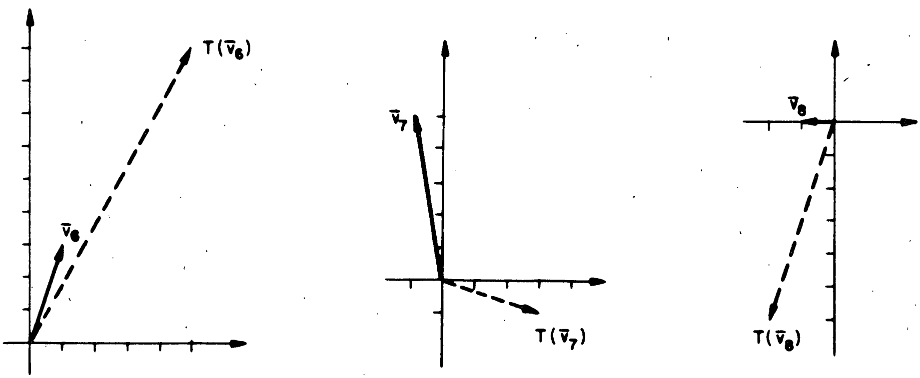
para los cuales se tiene que

$$T(\bar{v}_6) = T(1,\ 3) = (5,\ 9)$$

$$T(\bar{v}_7) = T(-1,\ 5) = (3,\ -1)$$

$$T(\bar{v}_8) = T(-1,\ 0) = (-2,\ -6)$$

por lo que T modifica la magnitud y dirección de éstos, como se aprecia en la siguiente figura



Enunciaremos a continuación una definición general para los conceptos discutidos en el ejemplo.

X.4.1 DEFINICION

Sea V un espacio vectorial sobre K y sea T: V $\rightarrow$ V un operador li neal. Si existe un vector $\bar{v} \in V$, $\bar{v} \neq 0$, tal que

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

para algún escalar $\lambda \in K$; entonces se dice que λ es un valor característico de T y que $\bar{v}$ es un vector característico de T correspondiente a λ.

Observe que la definición anterior excluye al vector cero como vector característico. Esto se debe a la conveniencia de que todo vector característico corresponda a un solo valor característico.

Empero, esta definición permite al escalar cero ser un valor característico.

Veamos a continuación algunos ejemplos más de valores y vectores característicos.

1. Para la transformación identidad

   I: V $\rightarrow$ V

   todos los vectores no nulos de V son vectores característicos correspondientes al valor 1, puesto que

   $$I(\bar{v}) = \bar{v} = 1\cdot\bar{v};\ \forall\ \bar{v} \in V$$

2. Para la transformación cero

   O: V $\rightarrow$ V

   todos los vectores no nulos de V son vectores característicos correspondientes al valor 0, puesto que

$0(\bar{v}) = \bar{0} = 0 \cdot \bar{v}; \; \forall \; \bar{v} \in V$

3. Para un operador lineal cualquiera

$T: V \to V$

todos los vectores no nulos del núcleo son vectores característicos correspondientes al valor 0.

En especial, si $N(T) = \{\bar{0}_V\}$ entonces el escalar cero no es un valor característico de T.

4. Para el operador derivación definido por

$D(f) = f'$

en el espacio de las funciones reales de variable real, sus vectores característicos son aquellas funciones f no nulas tales que

$f' = \lambda f$

para algún $\lambda \in R$. Esta es una ecuación diferencial cuyas soluciones están dadas por la expresión

$f(x) = ce^{\lambda x}$

donde c es una constante arbitraria.

En efecto, se tiene que

$D(ce^{\lambda x}) = \lambda ce^{\lambda x}$

por lo que todas las funciones exponenciales de la forma $ce^{\lambda x}$ son vectores característicos del operador D correspondientes al valor λ, $\forall \; \lambda \in R$. Además, todos los números reales son valores característicos del operador D.

Una de las propiedades más útiles de los vectores característicos es la que establece el siguiente teorema.

> X.4.2 TEOREMA
>
> Sea $T: V \to V$ un operador lineal y sean $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ vectores característicos de T correspondientes a los valores $\lambda_1, \lambda_2, \ldots, \lambda_k$ respectivamente. Si $\lambda_i \neq \lambda_j \; \forall \; i \neq j$, entonces el conjunto $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k\}$ es linealmente independiente.

DEMOSTRACION

Se hará por inducción matemática

I) Si $k = 1$

por iv) de IX.1.3, considerando que, de X.4.1, $\bar{v}_1 \neq \bar{0}$

$$\alpha_1 \bar{v}_1 = \bar{0} \implies \alpha_1 = 0$$

por lo que $\bar{v}_1$ es linealmente independiente.

II) Suponemos que el conjunto $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k\}$ es linealmente independiente para $k = n$.

Por ii) de IX.2.3

$$\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 + \ldots + \alpha_n \bar{v}_n = \bar{0} \implies \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0 \qquad ---(1)$$

Para $k = n + 1$ la ecuación de dependencia es

$$\beta_1 \bar{v}_1 + \beta_2 \bar{v}_2 + \ldots + \beta_n \bar{v}_n + \beta_{n+1} \bar{v}_{n+1} = \bar{0} \qquad ---(2)$$

Aplicando el operador T a ambos miembros tenemos, por linealidad de T,

$$\beta_1 T(\overline{v}_1) + \beta_2 T(\overline{v}_2) + \dots + \beta_n T(\overline{v}_n) + \beta_{n+1} T(\overline{v}_{n+1}) = \overline{0}$$

Por X.4.1

$$\beta_1 \lambda_1 \overline{v}_1 + \beta_2 \lambda_2 \overline{v}_2 + \dots + \beta_n \lambda_n \overline{v}_n + \beta_{n+1} \lambda_{n+1} \overline{v}_{n+1} = \overline{0} \qquad \text{---(3)}$$

Por otra parte, multiplicando (2) por λ_{n+1} se obtiene

$$\beta_1 \lambda_{n+1} \overline{v}_1 + \beta_2 \lambda_{n+1} \overline{v}_2 + \dots + \beta_n \lambda_{n+1} \overline{v}_n + \beta_{n+1} \lambda_{n+1} \overline{v}_{n+1} = \overline{0} \qquad \text{---(4)}$$

Restando (4) de (3) y factorizando

$$\beta_1 (\lambda_1 - \lambda_{n+1}) \overline{v}_1 + \beta_2 (\lambda_2 - \lambda_{n+1}) \overline{v}_2 + \dots + \beta_n (\lambda_n - \lambda_{n+1}) \overline{v}_n = \overline{0}$$

y de (1),

$$\begin{aligned} \beta_1 (\lambda_1 - \lambda_{n+1}) &= \alpha_1 = 0 \\ \beta_2 (\lambda_2 - \lambda_{n+1}) &= \alpha_2 = 0 \\ &\vdots \\ \beta_n (\lambda_n - \lambda_{n+1}) &= \alpha_n = 0 \end{aligned}$$

Además, $\lambda_1, \lambda_2, \dots, \lambda_{n+i}$ son diferentes por hipótesis

es decir, $\lambda_i - \lambda_{n+1} \neq 0$, $i = 1, 2, \dots, n$

luego $\beta_1 = \beta_2 = \dots = \beta_n = 0$

de (2), $\beta_{n+1} \overline{v}_{n+1} = 0$

y en virtud de que, por X.4.1, $\overline{v}_{n+1} \neq 0$ se tiene $\beta_{n+1} = 0$

luego, $\forall\, k \in N$

$$\beta_1 \overline{v}_1 + \beta_2 \overline{v}_2 + \dots + \beta_k \overline{v}_k = \overline{0} \implies \beta_1 = \beta_2 = \dots = \beta_k = 0$$

y el conjunto $\{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_k\}$ es linealmente independiente.

$\square$

Es importante hacer notar que el recíproco de este teorema no es válido. Es decir, si un conjunto de vectores característicos es linealmente independiente, los vectores no necesariamente corresponden a valores característicos diferentes.

Por ejemplo, para la transformación identidad $I: R^3 \rightarrow R^3$, todos los vectores no nulos de R^3 son vectores característicos de I asociados al valor 1, por lo que la base canónica de R^3

$$\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

es un conjunto linealmente independiente formado por tres vectores característicos correspondientes al mismo valor.

Los valores y vectores característicos tienen las siguientes propiedades elementales, cuya demostración se deja al lector como ejercicio.

X.4.3 TEOREMA

Sean V un espacio vectorial sobre K, $T: V \rightarrow V$ un operador lineal y $\bar{v}$ un vector característico de T correspondiente al valor λ:

i) El escalar λ es único.

ii) $\forall\, k \in K,\ k \neq 0$, el vector $k\bar{v}$ es un vector característico de T correspondiente a λ.

iii) Si $\bar{u}$ es un vector característico de T correspondiente a λ y $\bar{u} \neq -\bar{v}$, entonces $\bar{u} + \bar{v}$ es un vector característico de T correspondiente a λ.

- Espacios característicos y espacios invariantes.

Analizando las propiedades ii) y iii) del teorema anterior podemos darnos cuenta que al agregar el vector cero al conjunto de vectores característicos correspondientes a un determinado valor se obtiene un subespacio de V. Es decir

> X.4.4 TEOREMA
>
> Si $T: V \to V$ es un operador lineal y λ es un valor característico de T, entonces el conjunto
>
> $$E(\lambda) = \{\overline{v} \mid \overline{v} \in V \quad y \quad T(\overline{v}) = \lambda\overline{v}\}$$
>
> es un subespacio de V.

DEMOSTRACION

Sean $\overline{v}_1$, $\overline{v}_2 \in E(\lambda)$; esto es

$$T(\overline{v}_1) = \lambda\overline{v}_1$$

y

$$T(\overline{v}_2) = \lambda\overline{v}_2$$

Entonces

$$T(\overline{v}_1 + \overline{v}_2) = T(\overline{v}_1) + T(\overline{v}_2) \qquad \text{por i) de X.1.3}$$

$$= \lambda\overline{v}_1 + \lambda\overline{v}_2 \qquad \text{por hipótesis}$$

$$T(\overline{v}_1 + \overline{v}_2) = \lambda(\overline{v}_1 + \overline{v}_2) \qquad \text{por vii) de IX.1.1}$$

y en consecuencia

$$\overline{v}_1 + \overline{v}_2 \in E(\lambda) \qquad - - - (1)$$

Además, si α es un escalar del campo sobre el que está definido V

$$T(\alpha\overline{v}_1) = \alpha T(\overline{v}_1) \qquad \text{por ii) de X.1.3}$$

$$= \alpha(\lambda\overline{v}_1) \qquad \text{por hipótesis}$$

$$T(\alpha\overline{v}_1) = (\alpha\lambda)\overline{v}_1 \qquad \text{por ix) de IX.1.1}$$

$T(\alpha\bar{v}_1) = (\lambda\alpha)\bar{v}_1$ por vi) de VIII.3.4

$T(\alpha\bar{v}_1) = \lambda(\alpha\bar{v}_1)$ por ix) de IX.1.1

de donde

$\alpha\bar{v}_1 \in E(\lambda)$ - - - (2)

Finalmente, por IX.1.6, de (1) y (2) se sigue que $E(\lambda)$ es un subespacio de V.

□

A dicho subespacio se le denomina espacio característico, como lo indica la siguiente definición.

X.4.5 DEFINICION

Si $T: V \to V$ es un operador lineal y λ es un valor característico de T, el conjunto

$$E(\lambda) = \{\bar{v} \mid \bar{v} \in V \quad y \quad T(\bar{v}) = \lambda\bar{v}\}$$

se llama espacio característico de T correspondiente al valor λ.

Es claro que todos los vectores de un espacio característico se transforman en vectores del mismo espacio al aplicarles la transformación; esto es

$$\text{si } \bar{v} \in E(\lambda) \quad \text{entonces} \quad T(\bar{v}) \in E(\lambda)$$

En general, cualquier subconjunto del dominio con esta propiedad se denomina "invariante" bajo la transformación, como se establece a continuación

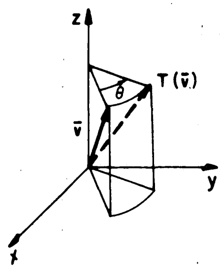
X.4.6 DEFINICION

Sean $T: V \to V$ un operador lineal y U un subespacio de V, se dice que U es invariante bajo T si

$$\forall\ \bar{u} \in U \quad \text{se tiene que} \quad T(\bar{u}) \in U$$

De esta manera, los espacios característicos de una transformación son invariantes bajo ésta, aunque no a la inversa.

Por ejemplo, consideremos la transformación $T: R^3 \to R^3$ que consiste en girar un vector cualquiera del espacio un ángulo θ positivo alrededor del eje z, Esto es



$$T(x,\ y,\ z) = (x \cos\theta - y \operatorname{sen}\theta,\ x \operatorname{sen}\theta + y \cos\theta,\ z)$$

Al aplicar esta transformación a cualquier vector del plano X,Y se obtiene siempre un vector del mismo plano, por lo que el subespacio de R^3

$$\{(x,\ y,\ 0)\ |\ x,\ y \in R\}$$

es un invariante bajo T; pero no es un espacio característico de la transformación cuando $\theta \neq n\,\pi$.

Por otra parte, el subespacio de R^3 constituido por todos los vectores sobre el eje z; esto es

$\{(0, 0, z) | z \in R\}$

es otro invariante bajo T y es también un espacio característico del operador.

- Caso de dimensión finita. Polinomio característico.

En un espacio de dimensión finita, el problema de obtener los valores y vectores característicos de un operador lineal puede resolverse con ayuda de los determinantes y los sistemas de ecuaciones lineales, mediante el procedimiento que se presenta a continuación.

Consideremos un espacio vectorial V sobre un campo K, con dim V = n, y sea $T: V \rightarrow V$ un operador lineal.

Deseamos obtener los vectores no nulos $\bar{v}$ y los escalares λ tales que

$$T(\bar{v}) = \lambda\bar{v} \qquad --- (1)$$

Como V es de dimensión finita, podemos plantear la igualdad anterior en términos de vectores de coordenadas y utilizar el concepto de matriz asociada a una transformación.

De esta manera, si B es una base de V y $M_B^B(T)$ es la matriz asociada a T referida a dicha base, se tiene que

$$[T(\bar{v})]_B = M_B^B(T)(\bar{v})_B$$

por lo que la expresión (1) es equivalente a

$$M_B^B(T)(\bar{v})_B = [\lambda\bar{v}]_B \qquad --- (2)$$

Esta expresión es una igualdad en K^n, donde el primer miembro es el producto de una matriz de $n \times n$ por una matriz de $n \times 1$.

Entonces, por las propiedades del álgebra de matrices la expresión (2) es equivalente a

$$M_B^B(T)(\bar{v})_B - [\lambda\bar{v}]_B = \bar{0}$$

$$M_B^B(T)(\bar{v})_B - \lambda(\bar{v})_B = \bar{0}$$

$$[M_B^B(T) - \lambda I_n](\bar{v})_B = \bar{0} \qquad - - - (3)$$

Esta última expresión es un sistema homogéneo de ecuaciones lineales cuya matriz de coeficientes es $M_B^B(T) - \lambda I_n$.

Con el propósito de simplificar la notación hagamos

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

$$I_n = I$$

$$(\bar{v})_B = \bar{x}$$

se tiene entonces el sistema homogéneo

$$(A - \lambda I)\bar{x} = \bar{0} \qquad - - - (4)$$

el cual tiene soluciones no triviales si y sólo si

$$\det (A - \lambda I) = 0 \qquad - - - (5)$$

Como $A - \lambda I$ es una matriz de n×n, el primer miembro de la expresión (5) es un polinomio en λ de grado n cuyas raíces son los únicos valores que hacen posible la existencia de soluciones no triviales para el sistema (4). En consecuencia, dichas raíces serán los valores característicos del operador siempre que pertenezcan al campo K.

Al polinomio det $(A - \lambda I)$ se le conoce como "polinomio característico" y a la expresión (5) como "ecuación característica" del operador T.

Por otra parte, para cada valor de λ la expresión (4) es un sistema homogéneo indeterminado cuyas soluciones no nulas son los vectores de coordenadas, en la base B, de los vectores característicos buscados.

Es claro que si $V = R^n$ y B es la base canónica de R^n, entonces las soluciones no nulas $\bar{x}$ del sistema (4) son directamente los vectores característicos del operador.

Para ilutrar lo anterior mediante un ejemplo, consideremos nuevamente el operador $T: R^2 \to R^2$ del inicio de esta sección, definido por

$$T(x, y) = (2x + y, 6x + y)$$

Si elegimos como B a la base canónica de R^2, la matriz asociada referida a dicha base es

$$M(T) = \begin{bmatrix} 2 & 1 \\ 6 & 1 \end{bmatrix}$$

por lo que la matriz $A - \lambda I$ del sistema (4) es

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

y en consecuencia

$$\det (A - \lambda I) = (2 - \lambda)(1 - \lambda) -6 = \lambda^2 - 3\lambda - 4$$

es el polinomio característico del operador, que en este caso es de grado 2.

Las raíces de dicho polinomio son los valores

$$\lambda_1 = 4$$

y

$$\lambda_2 = -1$$

Estos dos números reales son los valores característicos del operador T del ejemplo.

Para obtener sus vectores característicos correspondientes bastará con sustituir dichos valores en la expresión (4) y hallar las soluciones no triviales del sistema correspondiente a cada valor.

Así, para $\lambda_1 = 4$ el sistema $(A - \lambda I)\overline{x} = \overline{0}$ es

$$\begin{bmatrix} -2 & 1 \\ 6 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

y su solución general es

$$x_2 = 2x_1$$

De esta manera, los vectores característicos de T correspondientes al valor $\lambda_1 = 4$ son todos los vectores de la forma

$(a, 2a)$ con $a \neq 0$.

Por otra parte, para $\lambda_2 = -1$ tenemos el sistema

$$\begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

cuya solución general es

$$x_2 = -3x_1$$

por lo que los vectores característicos correspondientes a λ_2 son todos aquellos de la forma

$$(a, -3a) \quad \text{con } a \neq 0$$

Si aplicamos el procedimiento descrito anteriormente al operador que consiste en girar un vector cualquiera del plano un ángulo de 60° en sentido positivo, el cual puede considerarse como un operador $T: R^2 \to R^2$ definido por

$$T(x, y) = (\frac{1}{2} x - \frac{\sqrt{3}}{2} y, \frac{\sqrt{3}}{2} x + \frac{1}{2} y)$$

y cuya matriz asociada es

$$M(T) = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

se obtiene el polinomio característico

$$\det (A - \lambda I) = \begin{vmatrix} \frac{1}{2} - \lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} - \lambda \end{vmatrix} = \lambda^2 - \lambda + 1$$

cuyas raíces son los números complejos

$$\lambda_1 = \frac{1}{2} + \frac{\sqrt{3}}{2} i$$

y

$$\lambda_2 = \frac{1}{2} - \frac{\sqrt{3}}{2} i$$

En este caso las raíces del polinomio no pertenecen al campo sobre el que está definido el espacio V, puesto que R^2 es un espacio vectorial sobre R (y no sobre C).

Por lo tanto, el operador así definido no tiene valores caracterís ticos y, en consecuencia, tampoco tiene vectores característicos.

Geométricamente esto significa que ningún vector no nulo del plano conserva su dirección después de aplicarle dicho operador, ya que esto equivale a girarlo un ángulo de 60°.

Finalmente, apliquemos el procedimiento descrito para obtener los espacios característicos del operador lineal T: $R^3 \to R^3$ cuya matriz asociada es

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

Para ello, encontramos primero sus valores característicos a partir del polinomio

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 3 & 3 \\ 0 & 4-\lambda & 4 \\ 0 & -5 & -5-\lambda \end{vmatrix} = -\lambda^3 + \lambda^2 + 2\lambda$$

cuyas raíces son

$$\lambda_1 = 0$$

$$\lambda_2 = 2$$

$$\lambda_3 = -1$$

Para cada uno de estos valores obtengamos ahora sus vectores característicos correspondientes.

Para $\lambda_1 = 0$ se tiene el sistema

$$\begin{bmatrix} 2 & 3 & 3 \\ 0 & 4 & 4 \\ 0 & -5 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

cuya solución general es

$$x = 0$$
$$y = k$$
$$z = -k$$

Entonces, los vectores característicos de T correspondientes al valor $\lambda_1 = 0$ son todos aquellos de la forma

$$(0,\ k,\ -k),\ \text{con } k \neq 0$$

y el espacio característico correspondiente a dicho valor es

$$E(\lambda_1) = \{(0,\ k,\ -k) \mid k \in R\}$$

Es claro que, en este caso, $\dim E(\lambda_1) = 1$.

Para $\lambda_2 = 2$ el sistema es

$$\begin{bmatrix} 0 & 3 & 3 \\ 0 & 2 & 4 \\ 0 & -5 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

y su solución general es

$$x = k$$
$$y = 0$$
$$z = 0$$

por lo que

$$E(\lambda_2) = \{(k,\ 0,\ 0) \mid k \in R\}$$

y $\dim E(\lambda_2) = 1$

Para $\lambda_3 = -1$ el sistema es

$$\begin{bmatrix} 3 & 3 & 3 \\ 0 & 5 & 4 \\ 0 & -5 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

y su solución general es

$$x = k$$
$$y = 4k$$
$$z = -5k$$

por lo que

$$E(\lambda_3) = \{(k, 4k, -5k) \mid k \in R\}$$

y también dim $E(\lambda_3) = 1$

En este ejemplo el escalar cero resultó ser un valor característico del operador, y su espacio característico correspondiente coincide con el núcleo de T. Como el lector puede verificar.

Regresando al caso general, diremos que a las soluciones no nulas del sistema (4) se les conoce también como vectores característicos *de la matriz* A, y a las soluciones de la ecuación (5) como valores característicos de dicha matriz.

En forma análoga, al polinomio det $(A - \lambda I)$ se le llama también polinomio característico de A.

Cabe hacer notar que la matriz A depende de la elección de la base B, por lo que las soluciones del sistema (4) dependerán también de dicha elección. Esto era de esperarse puesto que las soluciones no nulas del sistema (4) son los vectores de coordenadas, en la base B, de los vectores característicos del operador.

Aparentemente las soluciones de la ecuación (5) también dependen de la matriz A; sin embargo, como se demostrará más adelante, todas las representaciones matriciales de un operador T tienen el mismo polinomio característico y, por lo tanto, los mismos valores característicos.

- Diagonalización. Matrices similares.

Como vimos en la sección X.2, la representación matricial de una transformación lineal depende de las bases que se elijan para el dominio y el codominio.

En el caso particular de los operadores suele utilizarse la misma base para el dominio y el codominio. Dicha base ha sido seleccionada, hasta ahora, buscando simplificar al máximo *la obtención* de la matriz asociada; sin embargo, esto no necesariamente coincide con simplificar al máximo *la forma* de la matriz que se obtiene (su aspecto).

Que la matriz asociada sea de forma sencilla ofrece ciertas ventajas pues, además de que permite identificar más fácilmente la información contenida en ella, su manejo algebraico se simplifica.

Entre los tipos más sencillos de matrices están las diagonales. Por ello, a continuación se tratará el problema de encontrar una representación diagonal para un operador, referida a una misma base del dominio y del codominio.

Con esta última restricción no siempre es posible encontrar una representación diagonal para cualquier operador. Las condiciones bajo las cuales existe tal representación, así como las características de ésta, se describen en el siguiente enunciado.

X.4.7

Sean V un espacio vectorial de dimensión n y T: V → V un operador lineal.

Existe una matriz diagonal asociada a T, referida a una base, si y sólo si existe una base de V formada por vectores característicos de T. En tal caso, la matriz asociada a T, referida a esta base, es una matriz diagonal cuyos elementos d_{ii} son los valores característicos correspondientes.

DEMOSTRACION

Supongamos primero que existe una base de V formada por vectores característicos de T. Esto es,

sea

$$B = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$$

una base de V tal que

$$T(\bar{v}_i) = \lambda_i \bar{v}_i, \text{ para } i = 1, 2, \ldots, n$$

Entonces, por IX.2.14

$$\begin{aligned} [T(\bar{v}_1)]_B &= (\lambda_1, 0, \ldots, 0)^T \\ [T(\bar{v}_2)]_B &= (0, \lambda_2, \ldots, 0)^T \\ &\vdots \\ [T(\bar{v}_n)]_B &= (0, 0, \ldots, \lambda_n)^T \end{aligned}$$

por lo que, de X.2.1

$$M_B^B(T) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

y T tiene una representación matricial diagonal, formada con los valores característicos de T correspondientes a los vectores característicos que constituyen la base B.

Recíprocamente, supongamos que T tiene una representación diagonal

$$D = \text{diag}(d_{11}, d_{22}, \dots, d_{nn})$$

referida a la base

$$B = \{\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n\}$$

Entonces, por X.2.1 sabemos que

$$[T(\bar{v}_1)]_B = (d_{11}, 0, \dots, 0)^T$$

$$[T(\bar{v}_2)]_B = (0, d_{22}, \dots, 0)^T$$

$$\vdots$$

$$[T(\bar{v}_n)]_B = (0, 0, \dots, d_{nn})^T$$

y en consecuencia, por IX.2.14

$$T(\bar{v}_i) = d_{ii}\bar{v}_i, \text{ para } i = 1, 2, \dots, n$$

por lo que los vectores $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$ son valores característicos de T correspondientes a los valores $d_{11}, d_{22}, \dots, d_{nn}$; con lo que la demostración queda completa.

$\square$

De esta manera, el problema de hallar una representación matricial diagonal para un operador T en un espacio de dimensión n, se reduce al de encontrar un conjunto linealmente independiente formado por n vectores característicos de T.

Así, por ejemplo, para el operador $T: R^3 \to R^3$ cuya matriz asociada, referida a la base canónica de R^3, es

$$M(T) = \begin{bmatrix} 2 & 3 & 3 \\ 0 & 4 & 4 \\ 0 & -5 & -5 \end{bmatrix}$$

sus valores característicos, obtenidos en un ejemplo anterior, son

$$\lambda_1 = 0$$
$$\lambda_2 = 2$$
$$\lambda_3 = -1$$

y sus correspondientes vectores característicos, son de la forma

$(0, k, -k)$, con $k \neq 0$

$(k, 0, 0)$, con $k \neq 0$

$(k, 4k, -5k)$, con $k \neq 0$

En consecuencia, como sus valores característicos son diferentes, por X.4.2 el conjunto formado por los vectores característicos

$$B = \{(0, 1, -1), (1, 0, 0), (1, 4, -5)\}$$

es linealmente independiente, y por IX.2.12 es una base de R^3.

Entonces, por X.4.7, existe una representación matricial diagonal del operador T. Para la base B elegida dicha representación es la siguiente

$$M_B^B(T) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Como se deduce de este ejemplo, para que un operador tenga representación diagonal es condición suficiente que sus valores característicos sean diferentes; sin embargo, tal condición no es necesaria, como se muestra mediante el ejemplo siguiente.

Para el operador lineal $T: R^3 \rightarrow R^3$ cuya matriz asociada, referida a la base canónica de R^3, es

$$M(T) = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

sus valores característicos son

$$\lambda_1 = 7$$
$$\lambda_2 = 1$$
$$\lambda_3 = 1$$

y sus espacios característicos son

$$E(\lambda_1) = \{(a, 2a, 3a) \mid a \in R\}$$
$$E(\lambda_2) = E(\lambda_3) = \{(a, b, -a-b) \mid a, b \in R\}$$

En este caso, el espacio $E(\lambda_2)$, que corresponde a un valor característico repetido, es un espacio de dimensión 2. En consecuencia, puede obtenerse una base de dicho espacio formada por dos vectores característicos asociados al mismo valor, como la siguiente

$\{(1, 0, -1), (0, 1, -1)\}$

De esta manera, el conjunto

$$B = \{(1, 2, 3), (1, 0, -1), (0, 1, -1)\}$$

es una base de R^3 formada por vectores característicos de T. Entonces, a pesar de tener un valor característico repetido, el operador tiene una representación diagonal, que para la base elegida es la siguiente

$$M_B^B(T) = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Un caso diferente de operador con valores característicos repetidos lo constituye la transformación lineal $T: R^3 \to R^3$ cuya matriz asociada, referida a la base canónica de R^3, es

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

Para este operador los valores característicos son

$$\lambda_1 = 1$$
$$\lambda_2 = 2$$
$$\lambda_3 = 2$$

y sus espacios característicos

$$E(\lambda_1) = \{(a, 0, 2a) \mid a \in R\}$$
$$E(\lambda_2) = E(\lambda_3) = \{(a, a, 2a) \mid a \in R\}$$

En este caso, tanto $E(\lambda_1)$ como $E(\lambda_2)$ son espacios de dimensión uno, por lo que a lo más podremos obtener un conjunto formado por dos vectores característicos de T que sea linealmente independiente. En consecuencia, por X.4.7, el operador T no tiene representación matricial diagonal referida a una sola base de R^3.

Veamos ahora la relación que existe entre las diversas matrices que representan a un mismo operador.†

X.4.8	TEOREMA Si dos matrices M y N representan al mismo operador lineal, entonces existe una matriz no singular P tal que $N = P^{-1}M\,P$

DEMOSTRACION

Sea V un espacio vectorial de dimensión n y sea $T: V \to V$ un operador lineal.

Si M y N son matrices asociadas al operador T referidas a las bases A y B, respectivamente, del espacio V; entonces, para cualquier vector $\bar{v} \in V$:

$$[T(\bar{v})]_A = M(\bar{v})_A \qquad --- \ (1)$$

y

$$[T(\bar{v})]_B = N(\bar{v})_B \qquad --- \ (2)$$

Por otra parte, si P es la matriz de transición de la base B a la base A, se tiene que

$$P(\bar{v})_B = (\bar{v})_A$$

†*En lo que sigue diremos que una matriz M "representa" al operador lineal $T: V \to V$ si M es la matriz asociada a T referida a una base de V.*

sustituyendo esta expresión en (1) se llega a

$$[T(\bar{v})]_A = MP(\bar{v})_B$$

Como la matriz de transición es no singular $\exists\ P^{-1}$; por lo que, de la expresión anterior

$$P^{-1}[T(\bar{v})]_A = P^{-1}M\,P(\bar{v})_B \qquad --- (3)$$

Además, como P^{-1} es la matriz de transición de la base A a la base B, se tiene que

$$P^{-1}[T(\bar{v})]_A = [T(\bar{v})]_B$$

por lo que, de (2)

$$P^{-1}[T(\bar{v})]_A = N(\bar{v})_B \qquad --- (4)$$

entonces, de (3) y (4)

$$N(\bar{v})_B = P^{-1}M\,P(\bar{v})_B \ ; \ \forall\ \bar{v} \in V$$

y en consecuencia

$$N = P^{-1}M\,P$$

como se quería.

$\square$

Es claro que si existe una matriz P tal que

$$N = P^{-1}M\,P$$

también existe una matriz Q tal que

$$M = Q^{-1}N\,Q$$

lo cual se sigue de la expresión anterior haciendo $Q = P^{-1}$ y despejando a M.

Por otra parte, el recíproco del teorema X.4.8 también es válido†; es decir, si M y N son dos matrices de $n\times n$ y existe una matriz P tal que $N = P^{-1}M\,P$, entonces M y N representan al mismo operador lineal.

Se tiene en consecuencia que dos matrices M y N representan al mismo operador lineal si y sólo si existe una matriz no singular P tal que $N = P^{-1}M\,P$.

La relación algebraica que establece el enunciado anterior entre las matrices M y N se conoce con el nombre de "similitud", esto es, se dice que M y N son similares.

X.4.9 DEFINICION

Dos matrices A y B de $n\times n$, son similares si existe una matriz no singular C tal que

$$B = C^{-1}A\,C$$

Se tiene entonces el siguiente teorema

X.4.10 TEOREMA

Dos matrices representan al mismo operador lineal si y sólo si son similares.

Las matrices similares tienen algunas propiedades interesantes que

†*Se sugiere al lector demostrarlo.*

conviene mencionar. Por ejemplo, tienen el mismo determinante.

X.4.11 TEOREMA

Si A y B son matrices similares entonces det A = det B

DEMOSTRACION

Si A y B son similares, por X.4.9

$$B = C^{-1} A\ C$$

y por lo tanto

$$\det B = \det(C^{-1} A\ C)$$

de donde

$$\det B = (\det C^{-1})(\det AC) \qquad \text{por iii) de VII.2.3}$$

$$= \frac{1}{\det C} (\det AC) \qquad \text{por VII.4.5}$$

$$\det B = \frac{1}{\det C} (\det A)(\det C) \qquad \text{por iii) de VII.2.3}$$

y finalmente

$$\det B = \det A$$

□

Con ayuda de este resultado podemos probar el siguiente enunciado importante

X.4.12 TEOREMA

Dos matrices similares tienen el mismo polinomio característico y, por tanto, los mismos valores característicos

DEMOSTRACION

Sean A y B dos matrices similares, entonces

$$B - \lambda I = P^{-1}A\,P - \lambda I \qquad \text{por X.4.9}$$

$$= P^{-1}A\,P - \lambda P^{-1}I\,P \qquad \text{por VI.3.5 y VI.4.1}$$

$$= P^{-1}(AP - \lambda IP) \qquad \text{por i) de VI.3.3}$$

$$B - \lambda I = P^{-1}(A - \lambda I)P \qquad \text{por ii) de VI.3.3}$$

luego, por X.4.9, $(B - \lambda I)$ y $(A - \lambda I)$ son similares y, de X.4.11,

$$\det(B - \lambda I) = \det(A - \lambda I)$$

como queríamos □

Como consecuencia de este teorema y de X.4.10 se concluye lo siguiente

X.4.13	TEOREMA Todas las representaciones matriciales de un operador lineal tienen el mismo polinomio característico y, por lo tanto, los mismos valores característicos.

Este resultado garantiza la unicidad (salvo en el orden) del conjunto de valores característicos de un operador, obtenido mediante el procedimiento descrito en el apartado anterior.

Finalmente, abordaremos el problema de "diagonalizar" una matriz; es decir, encontrar una matriz diagonal que sea similar a la matriz en cuestión.

Para resolver este problema podemos apoyarnos en el teorema X.4.7 del inicio de este apartado, cuyo equivalente para el caso de matrices puede enunciarse de la siguiente manera.

X.4.14 TEOREMA

Una matriz A de n×n es similar a una matriz diagonal D si y sólo si existe un conjunto linealmente independiente formado por n vectores característicos de A. En tal caso, existe una matriz no singular P tal que $D = P^{-1}A\ P$, donde D es una matriz diagonal cuyos elementos d_{ii} son los valores característicos de A, y P tiene como columnas a n vectores característicos de A correspondientes a dichos valores.

Tomando en cuenta que cualquier matriz A de n×n puede considerarse como la representación matricial de un operador lineal T en cierta base, el teorema X.4.14 es una consecuencia inmediata de los teoremas X.4.7 y X.4.10, y de la definición X.4.9, salvo la afirmación de que P tiene como columnas a n vectores característicos de A. Esto último se demuestra a continuación.

Sea $S = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ un conjunto linealmente independiente formado por n vectores característicos de A, correspondientes a los valores característicos $\lambda_1, \lambda_2, \ldots, \lambda_n$, y sea P una matriz cuyas columnas son dichos vectores; esto es

$$P = [\bar{v}_1 \quad \bar{v}_2 \quad \ldots \quad \bar{v}_n]$$

entonces

$$AP = A[\bar{v}_1 \quad \bar{v}_2 \quad \ldots \quad \bar{v}_n]$$
$$AP = [A\bar{v}_1 \quad A\bar{v}_2 \quad \ldots \quad A\bar{v}_n]$$

$$AP = [\lambda_1\bar{v}_1 \quad \lambda_2\bar{v}_2 \quad \ldots \quad \lambda_n\bar{v}_n]$$

$$= [\bar{v}_1 \quad \bar{v}_2 \quad \ldots \quad \bar{v}_n] \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

$$AP = P \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

Como el conjunto S es linealmente independiente, P es no singular y existe P^{-1}. En consecuencia

$$P^{-1}A\ P = P^{-1}P \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

$$P^{-1}A\ P = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

como se quería demostrar.

□

Una matriz A que satisface las condiciones del teorema X.4.14 se dice que es "diagonalizable".

Por ejemplo, para diagonalizar la matriz

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

obtenemos sus valores característicos

$$\lambda_1 = 1$$

$$\lambda_2 = 4$$

$$\lambda_3 = -2$$

y sus correspondientes vectores característicos

$(k, -2k, 4k)$, con $k \neq 0$

$(k, \quad k, \quad k)$, con $k \neq 0$

$(-5k, 7k, k)$, con $k \neq 0$

Como los valores característicos son diferentes, el conjunto

$$\{(1, -2, 4), (1, 1, 1), (-5, 7, 1)\}$$

es linealmente independiente.

En consecuencia, por X.4.14 la matriz A es diagonalizable y la matriz

$$P = \begin{bmatrix} 1 & 1 & -5 \\ -2 & 1 & 7 \\ 4 & 1 & 1 \end{bmatrix}$$

es tal que

$$P^{-1} A\, P = \text{diag}(1, 4, -2)$$

como el lector puede verificar.

Es claro que la matriz P de X.4.14, a la que algunos autores llaman "diagonalizante" o "diagonalizadora", no es única.

Por ejemplo, podemos emplear otros vectores característicos asociados a los mismos valores y obtener otra matriz

$$P_2 = \begin{bmatrix} \frac{1}{2} & 3 & 5 \\ -1 & 3 & -7 \\ 2 & 3 & -1 \end{bmatrix}$$

tal que

$$P_2^{-1} A\, P_2 = \text{diag}(1, 4, -2)$$

o cambiar el orden en las columnas de P y obtener

$$P_3 = \begin{bmatrix} 1 & 1 & -5 \\ 1 & -2 & 7 \\ 1 & 4 & 1 \end{bmatrix}$$

Con lo que se tiene otra matriz "diagonalizadora" puesto que

$$P_3^{-1} A\, P_3 = \mathrm{diag}(4,\ 1,\ -2)$$

sólo que la matriz diagonal obtenida en este caso es diferente a la de los dos anteriores, debido a que los valores característicos se encuentran en distinto orden.

- Teorema de Cayley - Hamilton

Uno de los resultados más interesantes que relacionan a una matriz con su polinomio característico es el que se conoce como teorema de Cayley-Hamilton, el cual establece que "toda matriz cuadrada es una raíz de su polinomio característico".

Antes de demostrar el teorema debemos precisar algunos términos mediante la siguiente definición

X.4.15 DEFINICION

Si f es el polinomio en x con coeficientes en K definido por

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

y A es una matriz de m×m con elementos en K, entonces

$$f(A) = a_n A^n + a_{n-1} A^{n-1} + \ldots + a_1 A + a_0 I_m$$

En particular, se dice que A es una raíz del polinomio f si $f(A) = 0$, donde 0 es la matriz nula de m×m.

Así, por ejemplo, para el polinomio

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

y la matriz

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

tenemos que

$$p(A) = A^2 - 4A + 3I$$

$$= \begin{bmatrix} 3 & 0 & -6 \\ -9 & 9 & -3 \\ -5 & 4 & 6 \end{bmatrix} - 4 \begin{bmatrix} -1 & 2 & 0 \\ 1 & 1 & -3 \\ 3 & -2 & 0 \end{bmatrix} + 3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$p(A) = \begin{bmatrix} 10 & -8 & -6 \\ -13 & 8 & 9 \\ -17 & 12 & 9 \end{bmatrix}$$

Para el mismo polinomio y la matriz

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

se tiene que

$$p(B) = B^2 - 4B + 3I$$

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

$$p(B) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

por lo que B es una raíz del polinomio p.

X.4.16 TEOREMA (DE CAYLEY-HAMILTON)

Si A es una matriz de $n\times n$ y

$$f(\lambda) = \det(A - \lambda I)$$

entonces

$$f(A) = 0$$

DEMOSTRACION

Sea A una matriz de $n\times n$, entonces los menores de $A - \lambda I$ son de orden $n - 1$ y sus cofactores son polinomios en λ de grado $n - 1$, luego a la matriz cuyos elementos son los cofactores de la transpuesta de $A - \lambda I$ la podemos expresar como

$$\operatorname{Adj}(A - \lambda I) = \lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} +\ldots+ \lambda B_1 + B_0$$

donde las B_i, $i = 0, 1,\ldots, n - 1$ son matrices de $n\times n$

Ya que, $\quad (A - \lambda I)\operatorname{Adj}(A - \lambda I) = [\det(A - \lambda I)]I$

y $\quad f(\lambda) = \det(A - \lambda I) = a_n\lambda^n + a_{n-1}\lambda^{n-1} +\ldots+ a_1\lambda + a_0$

se tiene

$$(A - \lambda I)(\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} +\ldots+ \lambda B_1 + B_0) =$$

$$= (a_n\lambda^n + a_{n-1}\lambda^{n-1} +\ldots+ a_1\lambda + a_0)I$$

y efectuando operaciones

$$\lambda^{n-1}AB_{n-1}+\lambda^{n-2}AB_{n-2}+\ldots+\lambda AB_1+AB_0-\lambda^nB_{n-1}-\lambda^{n-1}B_{n-2}-\ldots-\lambda^2B_1-\lambda B_0 =$$

$$= a_n\lambda^nI+a_{n-1}\lambda^{n-1}I+\ldots+a_1\lambda I+a_0I$$

agrupando y reordenando

$$- \lambda^n B_{n-1} + \lambda^{n-1}(AB_{n-1} - B_{n-2}) +\ldots+ \lambda(AB_1 - B_0) + AB_0 =$$

$$= \lambda^n a_n I + \lambda^{n-1} a_{n-1} I +\ldots+ \lambda a_1 I + a_0 I$$

lo que se cumple ∀ λ, en consecuencia

$$- B_{n-1} = a_n I$$

$$AB_{n-1} - B_{n-2} = a_{n-1} I$$

$$\vdots$$

$$AB_1 - B_0 = a_1 I$$

$$AB_0 = a_0 I$$

multiplicando ambos miembros de estas igualdades por A^n, A^{n-1},..., A e I, respectivamente,

$$- A^n B_{n-1} = a_n A^n$$

$$A^n B_{n-1} - A^{n-1} B_{n-2} = a_{n-1} A^{n-1}$$

$$\vdots$$

$$A^2 B_1 - AB_0 = a_1 A$$

$$AB_0 = a_0 I$$

sumándolas ahora, miembro a miembro

$$0 = a_n A^n + a_{n-1} A^{n-1} +\ldots+ a_1 A + a_0 I$$

esto es, $0 = f(A)$

como queríamos

□

Así, por ejemplo, para la matriz

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

cuyo polinomio característico es

$$f(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 3 & 3 \\ 0 & 4-\lambda & 4 \\ 0 & -5 & -5-\lambda \end{vmatrix} = - \lambda^3 + \lambda^2 + 2\lambda$$

se tiene que

$$f(A) = - A^3 + A^2 + 2A$$

por lo que

$$f(A) = - \begin{bmatrix} 8 & 9 & 9 \\ 0 & 4 & 4 \\ 0 & -5 & -5 \end{bmatrix} + \begin{bmatrix} 4 & 3 & 3 \\ 0 & -4 & -4 \\ 0 & 5 & 5 \end{bmatrix} + 2 \begin{bmatrix} 2 & 3 & 3 \\ 0 & 4 & 4 \\ 0 & -5 & -5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

como lo indica el teorema.

Si A es una matriz cualquiera de $n \times n$, el teorema de Cayley-Hamilton puede emplearse para expresar su potencia enésima y demás potencias superiores como combinaciones lineales de las potencias - - - - - A^{n-1}, A^{n-2}, ..., A, I.

Por ejemplo, para la matriz

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

cuyo polinomio característico es

$$-\lambda^3 + 5\lambda^2 - 8\lambda + 4$$

Por el teorema X.4.16 sabemos que

$$-A^3 + 5A^2 - 8A + 4I = 0$$

de donde

$$A^3 = 5A^2 - 8A + 4I$$

Es decir, A^3 está expresada como combinación lineal de A^2, A, I.

Para la cuarta potencia tenemos, de la expresión anterior, que

$$\begin{aligned} A^4 &= (5A^2 - 8A + 4I)A \\ &= 5A^3 - 8A^2 + 4A \\ &= 5(5A^2 - 8A + 4I) - 8A^2 + 4A \\ A^4 &= 17A^2 - 36A + 20I \end{aligned}$$

y para la quinta potencia

$$\begin{aligned} A^5 &= (17A^2 - 36A + 20I)A \\ &= 17A^3 - 36A^2 + 20A \\ &= 17(5A^2 - 8A + 4I) - 36A^2 + 20A \\ A^5 &= 49A^2 - 116A + 68I \end{aligned}$$

etc.

El teorema también puede utilizarse para expresar A^{-1} como una combinación lineal de las potencias A^{n-1}, A^{n-2}, ..., A, I, cuando A es no singular.

Así, por ejemplo, para la misma matriz A del ejemplo anterior, de la expresión

$$- A^3 + 5A^2 - 8A + 4I = 0$$

se tiene que

$$- A^3 + 5A^2 - 8A = - 4I$$

$$- \frac{1}{4}(-A^3 + 5A^2 - 8A) = I$$

$$\frac{1}{4}A^3 - \frac{5}{4}A^2 + 2A = I$$

$$A(\frac{1}{4}A^2 - \frac{5}{4}A + 2I) = I$$

y en consecuencia

$$A^{-1} = \frac{1}{4}A^2 - \frac{5}{4}A + 2I$$

X.4.17 EJERCICIOS

1.- Sean $T: R^2 \rightarrow R^2$ un operador lineal y $\overline{v} = (1,1), \overline{w} = (1, -1)$ dos vectores característicos de T correspondientes a los valores $\lambda_1 = 2$ y $\lambda_2 = 0$, respectivamente. Determinar la regla que define a T.

2.- Demostrar que un vector característico de un operador lineal, corresponde a uno y sólo un valor característico de dicho operador.

3.- Sea

$$M(T) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ r & 0 & s \end{bmatrix}$$

la matriz asociada al operador lineal $T: R^3 \rightarrow R^3$

a) Determinar los valores de r y s tales que el espacio $U = \{(a, b, a - b) \mid a, b \in R\}$ sea un invariante de T.

b) Obtener los valores y los espacios característicos de T.

4.- Obtener los espacios característicos de la transformación del ejercicio 9 de X.3.17, y dar una interpretación geométrica de ellos.

5.- De cada uno de los siguientes operadores lineales

$T: R^3 \rightarrow R^3$ tal que $T(x,y,z) = (-x+z,\ 2x+3y+4z,\ -x-3z)$

$S: R^3 \rightarrow R^3$ tal que $S(x,y,z) = (4x+2y-z,\ -2x-y+2z,\ x+2y+2z)$

a) Hallar los valores característicos.

b) Obtener la dimensión de cada uno de los espacios característicos.

c) Determinar si existe una representación matricial diagonal, y en caso afirmativo obtenerla.

6.- Demostrar que si A es similar a B y B es similar a C, entonces

a) A y C son similares

b) tr A = tr C

7.- Demostrar que si M y N son dos matrices de $n \times n$ y existe una matriz P tal que $N = P^{-1} M P$, entonces M y N representan al mismo operador lineal.

8.- Sean $T: V \rightarrow V$ un operador lineal, $A = \{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ una base de

V y $M(T) = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 2 & -1 & 1 \end{bmatrix}$ la representación matricial de T refe-

rida a la base A.

Obtener la matriz N(T) referida a la base $B = \{\bar{w}_1, \bar{w}_2, \bar{w}_3\}$ de V, tal que

$$\bar{w}_1 = 2\bar{v}_1 + \bar{v}_2 + 3\bar{v}_3$$
$$\bar{w}_2 = \bar{v}_1 - \bar{v}_2 + \bar{v}_3$$
$$\bar{w}_3 = -\bar{v}_1 + 3\bar{v}_2$$

y verificar que N y M son similares.

9.- Verificar que las matrices

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad y \quad \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

tienen los mismos valores característicos pero no son similares. ¿Contradice este ejemplo el teorema X.4.12?

10.- Determinar si cada una de las siguientes matrices es diagonalizable y en caso afirmativo obtener una matriz P tal que $P^{-1}AP$ sea una matriz diagonal

a) $A = \begin{bmatrix} 2 & -1 \\ 4 & -3 \end{bmatrix}$ b) $A = \begin{bmatrix} 1 & 4 & 0 \\ -1 & 5 & 0 \\ 1 & 3 & 2 \end{bmatrix}$ c) $A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 3 & 1 \\ 1 & -1 & -1 \end{bmatrix}$

11.- Utilizar el teorema de Cayley - Hamilton para obtener

a) A^4 donde $A = \begin{bmatrix} 1 & 2 \\ 5 & -2 \end{bmatrix}$

b) B^{-1} donde $B = \begin{bmatrix} 2 & -5 & 1 \\ -1 & 1 & -1 \\ -2 & 6 & -3 \end{bmatrix}$

X.5 OPERADORES EN ESPACIOS CON PRODUCTO INTERNO

En esta sección estudiaremos algunos tipos especiales de operadores que actúan en espacios vectoriales donde se tiene definido un producto interno.

En particular nos ocuparemos de aquellos para los cuales existe una base ortonormal del espacio formada por vectores característicos del operador. Para estos operadores se tiene una representación matricial diagonal y una descomposición en términos de proyecciones ortogonales sobre sus espacios característicos, como veremos al final de la sección.

- El adjunto de un operador lineal.

Los tipos especiales de operadores que nos interesan pueden quedar definidos y clasificados convenientemente a partir del concepto de operador adjunto, cuya definición y propiedades veremos a continuación.

X.5.1 DEFINICION

Sea V un espacio con producto interno y sea T: V → V un operador lineal. Un operador T*: V → V se dice que es adjunto de T si

$$(T(\bar{u})|\bar{v}) = (\bar{u}|T^*(\bar{v})), \ \forall\, \bar{u}, \bar{v} \in V$$

Para demostrar la existencia y unicidad del adjunto, en el caso de espacios de dimensión finita, nos apoyaremos en el teorema X.5.2 que se presenta a continuación.

En dicho teorema se emplea una función ϕ: V → K, la cual puede ser considerada como una transformación debido a que todo campo K es un espacio vectorial sobre sí mismo. Sin embargo, en el enunciado se emplea el término "funcional" debido a que es frecuente distinguir con tal nombre a este tipo de funciones de un espacio vectorial en su campo de escalares.

X.5.2 TEOREMA

Sea V un espacio vectorial sobre un campo K, de dimensión finita y con producto interno. Si $\phi: V \to K$ es una funcional lineal, entonces existe un vector único $\bar{x} \in V$ tal que

$$\phi(\bar{u}) = (\bar{u}|\bar{x}), \; \forall \; \bar{u} \in V$$

DEMOSTRACION

Sea $B = \{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$ una base ortonormal de V.

Para demostrar que existe $\bar{x}$ hagamos

$$\bar{x} = \sum_{i=1}^{n} \overline{\phi(\bar{e}_i)}\, \bar{e}_i \qquad ---(1)$$

De esta manera, para un vector arbitrario $\bar{e}_j$ de la base B se tiene, por las propiedades del producto interno

$$(\bar{e}_j|\bar{x}) = (\bar{e}_j| \sum_{i=1}^{n} \overline{\phi(\bar{e}_i)}\, \bar{e}_i)$$

$$= \sum_{i=1}^{n} (\bar{e}_j|\overline{\phi(\bar{e}_i)}\, \bar{e}_i)$$

$$(\bar{e}_j|\bar{x}) = \sum_{i=1}^{n} \phi(\bar{e}_i)(\bar{e}_j|\bar{e}_i)$$

y como B es un conjunto ortonormal, por IX.5.11

$$(\bar{e}_j|\bar{x}) = \phi(\bar{e}_j) \qquad ---(2)$$

Consideremos ahora un vector cualquiera $\bar{u} \in V$. Puesto que B es una base de V existen escalares únicos $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que

$$\bar{u} = \sum_{j=1}^{n} \alpha_j \bar{e}_j$$

por lo que

$$(\bar{u}|\bar{x}) = (\sum_{j=1}^{n} \alpha_j \bar{e}_j | \bar{x})$$

y de las propiedades del producto interno

$$(\bar{u}|\bar{x}) = \sum_{j=1}^{n} \alpha_j (\bar{e}_j | \bar{x})$$

entonces, de (2)

$$(\bar{u}|\bar{x}) = \sum_{j=1}^{n} \alpha_j \phi(\bar{e}_j)$$

y como ϕ es lineal

$$(\bar{u}|\bar{x}) = \phi(\sum_{j=1}^{n} \alpha_j \bar{e}_j)$$

por lo que

$$(\bar{u}|\bar{x}) = \phi(\bar{u})$$

y la expresión (1) define un vector $\bar{x}$ con las características deseadas.

Para demostrar que $\bar{x}$ es único supongamos otro vector $\bar{y} \in V$ tal que

$$\phi(\bar{u}) = (\bar{u}|\bar{y}), \ \forall \ \bar{u} \in V$$

Entonces se tendrá que

$$(\bar{u}|\bar{x}) = (\bar{u}|\bar{y}), \ \forall \ \bar{u} \in V$$

y por las propiedades de K y del producto interno en V

$$(\bar{u}|\bar{x}) - (\bar{u}|\bar{y}) = 0$$

$$(\bar{u}|\bar{x} - \bar{y}) = 0, \forall\, \bar{u} \in V$$

En particular, si $\bar{u} = \bar{x} - \bar{y}$ se tendrá que

$$(\bar{x} - \bar{y} \mid \bar{x} - \bar{y}) = 0$$

y por iv) de IX.5.2

$$\bar{x} - \bar{y} = \bar{0}$$

En consecuencia

$$\bar{x} = \bar{y}$$

y la demostración de X.5.2 queda completa. □

Demostraremos ahora la existencia y unicidad del adjunto de un operador lineal T, en un espacio V de dimensión finita y con producto interno.

Sea $\bar{v}$ un vector arbitrario del espacio V y definamos la funcional $\phi_v : V \to K$ como

$$\phi_v(\bar{u}) = (T(\bar{u})|\bar{v}), \forall\, \bar{u} \in V$$

Es claro que ϕ_v es una funcional lineal ya que, por linealidad de T y por las propiedades del producto interno, se tiene que $\forall\, \bar{u}_1, \bar{u}_2 \in V$ y $\alpha \in K$:

$$\phi_v(\alpha\bar{u}_1 + \bar{u}_2) = (T(\alpha\bar{u}_1 + \bar{u}_2)|\bar{v})$$

$$= (\alpha T(\bar{u}_1) + T(\bar{u}_2)|\bar{v})$$

$$\phi_v(\alpha\bar{u}_1 + \bar{u}_2) = (\alpha T(\bar{u}_1)|\bar{v}) + (T(\bar{u}_2)|\bar{v})$$

$$\phi_v(\alpha\bar{u}_1 + \bar{u}_2) = \alpha(T(\bar{u}_1)|\bar{v}) + (T(\bar{u}_2)|\bar{v})$$

$$\phi_v(\alpha\bar{u}_1 + \bar{u}_2) = \alpha\phi_v(\bar{u}_1) + \phi_v(\bar{u}_2)$$

Entonces, por X.5.2 existe un único $\bar{x}_v \in V$ tal que

$$\phi_v(\bar{u}) = (\bar{u}|\bar{x}_v);\ \forall\ \bar{u} \in V$$

Por lo tanto, para cada vector $\bar{v} \in V$ existe un único $\bar{x}_v \in V$ tal que

$$(T(\bar{u})|\bar{v}) = (\bar{u}|\bar{x}_v)$$

En consecuencia, el operador $T^*: V \to V$ definido por

$$T^*(\bar{v}) = \bar{x}_v,\ \forall\ \bar{v} \in V$$

es tal que

$$(T(\bar{u})|\bar{v}) = (\bar{u}|T^*(\bar{v}));\ \forall\ \bar{u},\ \bar{v} \in V$$

Para demostrar que T^* es único supongamos otro operador $S: V \to V$ tal que

$$(T(\bar{u})|\bar{v}) = (\bar{u}|S(\bar{v})),\ \forall\ \bar{u},\ \bar{v} \in V$$

entonces se tendrá que

$$(\bar{u}|T^*(\bar{v})) = (\bar{u}|S(\bar{v}));\ \forall\ \bar{u},\ \bar{v} \in V$$

por lo que

$$T^*(\bar{v}) = S(\bar{v})\quad \forall\ \bar{v} \in V$$

y en consecuencia

$$T^* = S.$$

El operador T^*, cuya existencia y unicidad acabamos de probar, tiene además la propiedad de ser un operador lineal, como se demuestra a continuación.

En efecto, de

$$(T(\bar{u})|\bar{v}) = (\bar{u}|T^*(\bar{v})), \forall\ \bar{u}, \bar{v} \in V$$

se sigue que, $\forall\ \bar{v}_1, \bar{v}_2 \in V$ y $\alpha \in K$

$$(\bar{u}|T^*(\alpha\bar{v}_1 + \bar{v}_2)) = (T(\bar{u})|\alpha\bar{v}_1 + \bar{v}_2)$$

y en consecuencia

$$\begin{aligned}(\bar{u}|T^*(\alpha\bar{v}_1 + \bar{v}_2)) &= (T(\bar{u})|\alpha\bar{v}_1) + (T(\bar{u})|\bar{v}_2) && \text{por ii) de IX.5.1}\\ &= \bar{\alpha}(T(\bar{u})|\bar{v}_1) + (T(\bar{u})|\bar{v}_2) && \text{por i) de IX.5.2}\\ &= \bar{\alpha}(\bar{u}|T^*(\bar{v}_1)) + (\bar{u}|T^*(\bar{v}_2)) && \text{por X.5.1}\\ &= (\bar{u}|\alpha T^*(\bar{v}_1)) + (\bar{u}|T^*(\bar{v}_2)) && \text{por i) de IX.5.2}\\ (\bar{u}|T^*(\alpha\bar{v}_1 + \bar{v}_2)) &= (\bar{u}|\alpha T^*(\bar{v}_1) + T^*(\bar{v}_2)) && \text{por ii) de IX.5.1}\end{aligned}$$

Como esta expresión se cumple para todo vector $\bar{u} \in V$ se tiene que

$$T^*(\alpha\bar{v}_1 + \bar{v}_2) = \alpha T^*(\bar{v}_1) + T^*(\bar{v}_2)$$

por lo que T^* es lineal.

Hemos demostrado así el siguiente teorema

X.5.3 TEOREMA

Si V es un espacio vectorial de dimensión finita y con producto interno, entonces para cada operador lineal $T: V \to V$ existe un único adjunto T^*, que también es lineal.

Cabe comentar que este teorema no se cumple para espacios de dimen sión infinita, debido a que en este caso no es posible garantizar la existencia del adjunto. Aunque la unicidad y linealidad de T^* si se verifican cuando éste existe.

Muchas propiedades que involucran el concepto de adjunto son independientes de la dimensión del espacio, por lo que en lo sucesivo sólo indicaremos que la dimensión es finita cuando tal restricción se requiera por otra razón diferente a la existencia de T^*, la cual se presupondrá siempre que se hable de éste.

Regresando al caso de dimensión finita, hay una relación entre las representaciones matriciales de los operadores T y T^* cuando éstas se encuentran referidas a una base ortonormal.

Dicha relación, que puede emplearse para obtener el adjunto de un operador dado, es la siguiente: si A es la representación matri - cial de T en una base ortonormal, entonces A^* (la conjugada-transpuesta de A) es la representación matricial de T^* en dicha base. Se tiene así el siguiente teorema.

X.5.4 TEOREMA

Sea V un espacio vectorial de dimensión finita con producto in - terno y sea B una base ortonormal de V. Si $T: V \rightarrow V$ es un opera dor lineal, entonces

$$M_B^B(T^*) = \left[M_B^B(T)\right]^*$$

DEMOSTRACION

Sea $B = \{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$ una base ortonormal de V, y sean

$$A = [a_{ij}] = M_B^B(T)$$

y

$$N = [n_{ij}] = M_B^B(T^*)$$

las representaciones matriciales de T y T^* referidas a la base B.

Entonces, por X.2.1, a_{ij} es la i-ésima coordenada de $T(\bar{e}_j)$ en la base B, y como B es una base ortonormal, por IX.5.13 podemos escribir

$$a_{ij} = (T(\bar{e}_j)|\bar{e}_i) \qquad ---(1)$$

Mediante un razonamiento similar, n_{ij} es la i-ésima coordenada de $T^*(\bar{e}_j)$ en la base B; por lo que

$$n_{ij} = (T^*(\bar{e}_j)|\bar{e}_i) \qquad ---(2)$$

Luego, por X.5.1, de la expresión (1) se sigue que

$$a_{ij} = (\bar{e}_j|T^*(\bar{e}_i))$$

y por i) de IX.5.1

$$\bar{a}_{ij} = (T^*(\bar{e}_i)|\bar{e}_j)$$

En consecuencia

$$\bar{a}_{ji} = (T^*(\bar{e}_j)|\bar{e}_i) = n_{ij}$$

Finalmente, por VI.7.11, VI.7.6 y VI.7.1 se concluye que

$$A^* = N$$

Es decir

$$[M_B^B(T)]^* = M_B^B(T^*)$$

como lo indica el teorema.

□

Consideremos, por ejemplo, el operador $T: C^2 \rightarrow C^2$ definido por

$$T(x, y) = (-x - 5iy, [2 + i]y)$$

Para obtener el adjunto de T podemos seleccionar una base ortonormal de C^2 para el producto interno usual; por ejemplo

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

Para dicha base tenemos

$$M_B^B(T) = \begin{bmatrix} -1 & -5i \\ 0 & 2+i \end{bmatrix}$$

La conjugada transpuesta de esta matriz es

$$[M_B^B(T)]^* = \begin{bmatrix} -1 & 0 \\ 5i & 2-i \end{bmatrix}$$

por lo que, de X.5.4

$$M_B^B(T^*) = \begin{bmatrix} -1 & 0 \\ 5i & 2-i \end{bmatrix}.$$

y la regla de correspondencia del adjunto de T es

$$T^*(x, y) = (-x, 5ix + [2 - i]y)$$

Es importante remarcar que el teorema X.5.4 es para matrices referidas a una base ortonormal. Cuando la base no es ortonormal no existe una relación sencilla entre las representaciones matriciales de T y T*.

A continuación se presentan algunas otras propiedades que relacionan a un operador T con su adjunto T^*.

X.5.5 TEOREMA

Sean V un espacio con producto interno, W un subespacio de V y $T: V \to V$ un operador lineal:

i) $T^*(V) = N(T)^{\perp}$

ii) $N(T^* \circ T) = N(T)$

iii) Si W es invariante bajo T entonces $W^{\perp}$ es invariante bajo T^*

DEMOSTRACION

i) Sea $\bar{u} \in N(T)$.

Como $T(\bar{u}) = \bar{0}$, por iii) de IX.5.2 se tiene que

$$(T(\bar{u})|\bar{v}) = 0, \forall \bar{v} \in V.$$

Entonces, por X.5.1 podemos escribir

$$(\bar{u}|T^*(\bar{v})) = 0, \forall \bar{v} \in V$$

por lo que, de IX.5.15

$$\bar{u} \in T^*(V)^{\perp}$$

y en consecuencia

$$N(T) \subseteq T^*(V)^{\perp} \qquad ---(1)$$

Sea ahora $\bar{w} \in T^*(V)^{\perp}$

Entonces, por IX.5.15

$(\bar{w}|T^*(\bar{v})) = 0, \ \forall \ \bar{v} \in V$

por lo que, de X.5.1

$(T(\bar{w})|\bar{v}) = 0, \ \forall \ \bar{v} \in V$

en consecuencia, de IX.5.15

$T(\bar{w}) \in V^{\perp}$

y por b) del ejercicio 12 de IX.5.19

$T(\bar{w}) = \bar{0}_V$

esto es

$\bar{w} \in N(T)$

y por lo tanto

$T^*(V)^{\perp} \subseteq N(T)$ - - -(2)

De (1) y (2) se sigue que

$T^*(V)^{\perp} = N(T)$

por lo que, del ejercicio 12, inciso d), de IX.5.19

$T^*(V) = N(T)^{\perp}$

lo que demuestra el enunciado i).

ii) Sea $\bar{u} \in N(T^* \circ T)$.

$(T(\bar{u})|T(\bar{u})) = (\bar{u}|T^*[T(\bar{u})])$ por X.5.1

$= (\bar{u}|(T^* \circ T)(\bar{u}))$ por X.3.6

$(T(\bar{u})|T(\bar{u})) = (\bar{u}|\bar{0})$ por X.1.2

$(T(\bar{u}) | T(\bar{u})) = 0$ por iii) de IX.5.2

En consecuencia, por iv) de IX.5.2

$$T(\bar{u}) = \bar{0}$$

por lo que

$$\bar{u} \in N(T)$$

se tiene así que

$$N(T^* \circ T) \subseteq N(T) \qquad ---(3)$$

Por otra parte, si $\bar{w} \in N(T)$ se tendrá que

$(T^* \circ T)(\bar{w}) = T^*[T(\bar{w})]$ por X.3.6

$= T^*(\bar{0})$ por X.1.2

$(T^* \circ T)(\bar{w}) = \bar{0}$ por X.5.3 y X.1.4

de donde $\bar{w} \in N(T^* \circ T)$ y

$$N(T) \subseteq N(T^* \circ T) \qquad ---(4)$$

Finalmente, de (3) y (4)

$$N(T^* \circ T) = N(T)$$

Como lo indica el teorema.

iii) Sea $\bar{w} \in W$.

Si W es invariante bajo T, por X.4.6

$$T(\bar{w}) \in W$$

Entonces, $\forall\, \bar{u} \in W^{\perp}$ se tendrá que

$$(T(\bar{w})|\bar{u}) = 0$$

por lo que, de X.5.1

$$(\bar{w}|T^*(\bar{u})) = 0$$

y en consecuencia

$$T^*(\bar{u}) \in W^\perp$$

Luego, por X.4.6, $W^\perp$ es invariante bajo T^*.

□

Finalmente, en el siguiente teorema enunciamos algunas otras propiedades del adjunto, cuya demostración se sugiere al lector como ejercicio, las cuales resultan ser análogas a las de la conjugación-transposición de matrices que se vio en el capítulo VI. Dicha analogía resulta natural en virtud de la relación establecida por el teorema X.5.4.

X.5.6 TEOREMA

Sea V un espacio vectorial sobre K, con producto interno. Si S y T son operadores lineales en V y α es un escalar de K, entonces:

i) $(T^*)^* = T$

ii) $(\alpha T)^* = \bar{\alpha} T^*$

iii) $(S + T)^* = S^* + T^*$

iv) $(S \circ T)^* = T^* \circ S^*$

- Operadores normales, hermitianos, antihermitianos y unitarios.

Estudiaremos ahora algunos tipos especiales de operadores, entre los cuales la clase más amplia que trataremos es la de los operadores normales.

X.5.7 DEFINICION

Sea V un espacio con producto interno y sea $T: V \to V$ un operador lineal. Se dice que T es normal si $T \circ T^* = T^* \circ T$.

De esta definición se sigue de inmediato que si T es normal T^* también es normal y viceversa.

Los operadores normales tienen las siguientes propiedades

X.5.8 TEOREMA

Sea V un espacio con producto interno y sea $T: V \to V$ un operador normal:

i) $\|T(\bar{v})\| = \|T^*(\bar{v})\|$, $\forall\, \bar{v} \in V$

ii) Si $T(\bar{v}) = \lambda\bar{v}$ entonces $T^*(\bar{v}) = \bar{\lambda}\bar{v}$.

iii) Si $\bar{v}_1$, $\bar{v}_2$ son vectores característicos de T correspondientes a los valores λ_1, λ_2 y $\lambda_1 \neq \lambda_2$, entonces $(\bar{v}_1|\bar{v}_2) = 0$.

DEMOSTRACION

i) $\|T(\bar{v})\|^2 = (T(\bar{v})|T(\bar{v}))$ por IX.5.4

$= (\bar{v}|T^*[T(\bar{v})])$ por X.5.1

$\|T(\bar{v})\|^2 = (\bar{v}|(T^* \circ T)(\bar{v}))$ por X.3.6

$$\|T(\bar{v})\|^2 = (\bar{v}|(T \circ T^*)(\bar{v}))$$ por X.5.7

$$= (\bar{v}|T[T^*(\bar{v})])$$ por X.3.6

$$= (\bar{v}|(T^*)^*[T^*(\bar{v})])$$ por i) de X.5.6.

$$= (T^*(\bar{v})|T^*(\bar{v}))$$ por X.5.1

$$\|T(\bar{v})\|^2 = \|T^*(\bar{v})\|^2$$ por IX.5.4

de donde

$$\|T(\bar{v})\| = \|T^*(\bar{v})\|$$

Es decir, las imágenes asignadas a un vector $\bar{v}$ cualquiera por un operador normal y por su adjunto "tienen el mismo tamaño".

ii) Primero demostraremos que $T - \alpha I$ es un operador normal para cualquier escalar α.

Por X.5.6 y tomando en cuenta que $I^* = I$ se tiene

$$(T - \alpha I)^* = T^* - \bar{\alpha} I$$

Entonces, por las propiedades del álgebra de transformaciones se tiene que

$$(T - \alpha I) \circ (T - \alpha I)^* = (T - \alpha I) \circ (T^* - \bar{\alpha} I)$$

$$(T - \alpha I) \circ (T - \alpha I)^* = T \circ T^* - \alpha T^* - \bar{\alpha} T + \alpha \bar{\alpha} I$$

y también que

$$(T - \alpha I)^* \circ (T - \alpha I) = (T^* - \bar{\alpha} I) \circ (T - \alpha I)$$

$$(T - \alpha I)^* \circ (T - \alpha I) = T^* \circ T - \bar{\alpha} T - \alpha T^* + \alpha \bar{\alpha} I$$

Como T es normal $T \circ T^* = T^* \circ T$, por lo que

$$(T - \alpha I) \circ (T - \alpha I)^* = (T - \alpha I)^* \circ (T - \alpha I)$$

y $T - \alpha I$ es normal.

Sea ahora $\bar{v}$ tal que $T(\bar{v}) = \lambda\bar{v}$

y formemos un operador S como

$$S = T - \lambda I$$

Entonces, por el resultado anterior S es normal y por i) de X.5.8

$$\|S(\bar{v})\| = \|S^*(\bar{v})\|$$

y en consecuencia

$$\|S(\bar{v})\| = \|(T - \lambda I)^*(\bar{v})\| \qquad \text{por definición de } S$$

$$= \|(T^* - \bar{\lambda}I)(\bar{v})\| \qquad \text{por X.5.6}$$

$$\|S(\bar{v})\| = \|T^*(\bar{v}) - \bar{\lambda}\bar{v}\| \qquad \text{por X.3.2}$$

Por otra parte, como $S = T - \lambda I$ y $T(\bar{v}) = \lambda\bar{v}$ se tiene

$$\|S(\bar{v})\| = \|(T - \lambda I)(\bar{v})\| = \|\lambda\bar{v} - \lambda\bar{v}\| = \|\bar{0}\| = 0$$

Llevando este resultado a la expresión anterior

$$0 = \|T^*(\bar{v}) - \bar{\lambda}\bar{v}\|$$

por lo que, de ii) de IX.5.5

$$T^*(\bar{v}) - \bar{\lambda}\bar{v} = \bar{0}$$

y en consecuencia

$$T^*(\bar{v}) = \bar{\lambda}\bar{v}$$

como se quería.

Este resultado puede interpretarse de dos maneras distintas, aunque equivalentes:

1) Si λ es un valor característico de T entonces $\overline{\lambda}$ es un valor característico de T*.

2) Todo vector característico de T es también un vector característico de T* (aunque no necesariamente correspondiente al mismo valor que para T).

iii) Sean $\overline{v}_1$, $\overline{v}_2$ vectores característicos de T correspondientes a los valores λ_1, λ_2 respectivamente.

Por X.5.1 podemos escribir

$$(T(\overline{v}_1)|\overline{v}_2) = (\overline{v}_1|T^*(\overline{v}_2))$$

De donde, por hipótesis y por el resultado ii) de este teorema

$$(\lambda_1\overline{v}_1|\overline{v}_2) = (\overline{v}_1|\overline{\lambda}_2\overline{v}_2)$$

Ahora, por las propiedades del producto interno

$$\lambda_1(\overline{v}_1|\overline{v}_2) = \lambda_2(\overline{v}_1|\overline{v}_2)$$

y en consecuencia

$$(\lambda_1 - \lambda_2)(\overline{v}_1|\overline{v}_2) = 0$$

Entonces, si $\lambda_1 \neq \lambda_2$ se tiene que $\lambda_1 - \lambda_2 \neq 0$ y de la expresión anterior se sigue que

$$(\overline{v}_1|\overline{v}_2) = 0$$

Es decir que, para los operadores normales, los vectores ca -

racterísticos asociados a valores distintos son ortogonales.

□

Algunos casos particulares de operadores normales reciben nombres especiales debido a que sus características propias son de interés para la teoría de operadores lineales.

Estos tipos de operadores suelen distinguirse con diferentes nombres cuando el espacio vectorial está definido sobre el campo de los números complejos y cuando está definido sobre el campo de los números reales. En la siguiente definición los nombres correspondientes al segundo caso, que en adelante llamaremos "caso real" por brevedad, aparecen entre paréntesis.

X.5.9 DEFINICION

Sea V un espacio vectorial sobre C (sobre R) con producto interno y sea $T: V \to V$ un operador lineal, se dice que

i) T es hermitiano (simétrico) si $T^* = T$

ii) T es antihermitiano (antisimétrico) si $T^* = -T$

iii) T es unitario (ortogonal) si $T^* = T^{-1}$

Es frecuente encontrar en los textos el término "Autoadjunto" en lugar de "Hermitiano". La razón resulta obvia al observar la definición. Como el lector puede verificar fácilmente, todos los operadores que acabamos de definir son operadores normales. Estos operadores tienen las siguientes propiedades con relación a los valores característicos.

X.5.10 TEOREMA

Sea V un espacio vectorial con producto interno y sea λ un valor característico del operador lineal $T: V \to V$

i) si $T^* = T$ entonces λ es real

ii) si $T^* = -T$ entonces λ es imaginario

iii) si $T^* = T^{-1}$ entonces $|\lambda| = 1$

DEMOSTRACION

i) Sea λ un valor característico de T; entonces

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

Luego, por ii) de X.5.8

$$T^*(\bar{v}) = \bar{\lambda}\bar{v}$$

Como $T^* = T$ se tiene que

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

en consecuencia

$$\lambda\bar{v} = \bar{\lambda}\bar{v}$$

y como $\bar{v} \neq \bar{0}$

$$\lambda = \bar{\lambda}$$

por lo que $\lambda \in R$.

ii) Mediante un razonamiento similar, pero haciendo $T^* = -T$, se tiene

$$\lambda = -\bar{\lambda}$$

por lo que $R(\lambda) = 0$.

iii) Nuevamente, de

$$T^*(\bar{v}) = \bar{\lambda}\bar{v}$$

como $T^* = T^{-1}$

$$T^{-1}(\bar{v}) = \bar{\lambda}\bar{v}$$

por lo que

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

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

$$\bar{v} = \bar{\lambda}\lambda\bar{v}$$

y en consecuencia

$$\bar{\lambda}\lambda = 1$$

de donde $|\lambda| = 1$

□

El teorema X.5.10 nos permite obtener algunas conclusiones interesantes respecto a los tipos de operadores definidos en X.5.9, entre las que se encuentran las siguientes.

El enunciado i) nos dice que los valores característicos de un operador hermitiano, cuando existen, son números reales.

Este mismo enunciado conduce a un resultado trivial en el caso de operadores simétricos, puesto que el campo de donde se toman los escalares, y por tanto los valores característicos, es el de los números reales.

El enunciado ii) nos indica que los valores característicos de un operador antihermitiano, cuando existen, son números imaginarios puros; es decir, números complejos cuya parte real es nula.

En el caso de operadores antisimétricos dicho enunciado nos permite concluir que, para este tipo de operadores, el único valor característico posible es el cero.

El enunciado iii) asegura que los valores característicos de un operador unitario, de existir, son números complejos cuyo módulo es de tamaño uno.

Para los operadores ortogonales, este enunciado indica que sus valores característicos, de existir, son necesariamente igual a 1 o a -1.

Finalmente, cabe hacer énfasis en que el teorema X.5.10 no se ocupa de la existencia de valores característicos, sino de las propiedades de éstos cuando existen.

El problema de la existencia de valores característicos es diferente en el caso de un espacio vectorial complejo y en el de un espacio vectorial real, por lo que ambos deben abordarse por separado, como lo haremos en el siguiente apartado.

Antes presentaremos una caracterización para las representaciones matriciales de los tipos de operadores definidos en X.5.9, cuando están referidas a bases ortonormales.

X.5.11 TEOREMA

Sean V un espacio con producto interno, B una base ortonormal de V, $T: V \to V$ un operador lineal y A la representación matricial de T referida a la base B:

i) $T^* = T$ si y sólo si $A^* = A$

ii) $T^* = - T$ si y sólo si $A^* = - A$

iii) $T^* = T^{-1}$ si y sólo si $A^* = A^{-1}$

DEMOSTRACION

Se demuestra a continuación solamente el enunciado i).

Ia prueba de que $T^* = T \Longrightarrow A^* = A$ es inmediata, puesto que, si A es la representación matricial de T en la base B, como $T^* = T$ se tiene que

$$A = M_B^B(T) = M_B^B(T^*)$$

y entonces, de X.5.4 se sigue que

$$A = A^*$$

Supongamos ahora que $A^* = A$, y sea $B = \{\bar{e}_1, \bar{e}_2, ..., \bar{e}_n\}$.

Si $\bar{u}$, $\bar{v}$ son dos vectores cualesquiera de V, entonces

$$\bar{u} = \sum_{j=1}^{n} \alpha_j \bar{e}_j \quad y \quad \bar{v} = \sum_{i=1}^{n} \beta_i \bar{e}_i$$

por lo que

$$(T(\bar{u}) | \bar{v}) = (T(\sum_{j=1}^{n} \alpha_j \bar{e}_j) | \sum_{i=1}^{n} \beta_i \bar{e}_i)$$

$$(T(\bar{u})|\bar{v}) = \left(\sum_{j=1}^{n} \alpha_j T(\bar{e}_j) \,\Big|\, \sum_{i=1}^{n} \beta_i \bar{e}_i\right)$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \alpha_j T(\bar{e}_j) \,\Big|\, \beta_i \bar{e}_i\right)$$

$$(T(\bar{u})|\bar{v}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_j \bar{\beta}_i (T(\bar{e}_j)|\bar{e}_i) \qquad ---(1)$$

donde $\bar{\beta}_i$ representa al conjugado de β_i.

Por otra parte

$$(\bar{u}|T(\bar{v})) = \left(\sum_{j=1}^{n} \alpha_j \bar{e}_j \,\Big|\, T\left(\sum_{i=1}^{n} \beta_i \bar{e}_i\right)\right)$$

$$= \left(\sum_{j=1}^{n} \alpha_j \bar{e}_j \,\Big|\, \sum_{i=1}^{n} \beta_i T(\bar{e}_i)\right)$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \alpha_j \bar{e}_j \,\Big|\, \beta_i T(\bar{e}_i)\right)$$

$$(\bar{u}|T(\bar{v})) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \bar{\beta}_i (\bar{e}_j|T(\bar{e}_i)) \qquad ---(2)$$

Además, como vimos en la demostración de X.5.4

$$a_{ij} = (T(\bar{e}_j)|\bar{e}_i) \qquad ---(3)$$

de donde

$$a_{ji} = (T(\bar{e}_i)|\bar{e}_j)$$

$$\bar{a}_{ji} = (\bar{e}_j|T(\bar{e}_i)) \qquad ---(4)$$

Llevando las expresiones (3) y (4) a (1) y (2) obtenemos

$$(T(\bar{u})|\bar{v}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_j \bar{\beta}_i \, a_{ij}$$

y

$$(\bar{u}|T(\bar{v})) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_j \bar{\beta}_i \, \bar{a}_{ji}$$

Finalmente, como $A^* = A$ se tiene que $\bar{a}_{ji} = a_{ij}$ por lo que

$$(T(\bar{u})|\bar{v}) = (\bar{u}|T(\bar{v}))$$

y en consecuencia $T^* = T$, lo que completa la demostración de i).

$\square$

Considerando las definiciones de la sección VI.7 y tomando en cuenta que $A^* = A^T$ cuando A es una matriz con elementos en R, el teorema X.5.11 nos permite concluir lo siguiente.

Un operador es hermitiano (simétrico) si y sólo si su representación matricial en una base ortonormal es una matriz hermitiana (simétrica). A una conclusión semejante se llega para operadores antihermitianos (antisimétricos) y unitarios (ortogonales).

Así por ejemplo, el operador $T: R^2 \rightarrow R^2$ definido por

$$T(x, y) = (x + 2y, 2x - 3y)$$

es un operador simétrico puesto que su representación matricial en la base canónica de R^2, que es una base ortonormal, es la matriz

$$M(T) = \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}$$

que es una matriz simétrica.

Si en lugar de la base canónica elegimos la siguiente base ortonormal de R^2

$$B = \{(\frac{1}{\sqrt{10}}, \frac{-3}{\sqrt{10}}), (\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}})\}$$

obtendremos la representación matricial

$$M_B^B(T) = \begin{bmatrix} -\frac{19}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{9}{5} \end{bmatrix}$$

que es también una matriz simétrica.

- Diagonalización en espacios con producto interno.

En este apartado estableceremos una condición necesaria y suficiente para la existencia de una base ortonormal formada por vectores característicos de un operador.

Para ello requerimos introducir algunos conceptos y resultados adicionales que veremos a continuación. El primero de tales conceptos es el de suma de dos subconjuntos de un espacio vectorial, que se define de la siguiente manera.

X.5.12 DEFINICION

Sea V un espacio vectorial sobre un campo K, y sean S_1, S_2 dos subconjuntos de V. La suma de S_1 y S_2 es el conjunto

$$S_1 + S_2 = \{\bar{v}_1 + \bar{v}_2 \mid \bar{v}_1 \in S_1 \text{ y } \bar{v}_2 \in S_2\}$$

Cuando S_1 y S_2 son, además, subespacios de V, el conjunto $S_1 + S_2$ es también un subespacio de V. El lector podrá demostrar esto fácilmente.

Un caso especial de suma que resulta de gran utilidad es el de suma directa, cuya definición es la siguiente

X.5.13 DEFINICION

Sea V un espacio vectorial sobre un campo K y sean W_1, W_2 dos subespacios de V. Se dice que V es la suma directa de W_1 y W_2 y se escribe

$$V = W_1 \oplus W_2$$

cuando

$$V = W_1 + W_2 \quad \text{y} \quad W_1 \cap W_2 = \{\overline{0}\}$$

Como consecuencia de esta última condición, las sumas directas tienen la siguiente propiedad importante

X.5.14 TEOREMA

Sea V un espacio vectorial sobre un campo K y sean W_1, W_2 dos subespacios de V:

$V = W_1 \oplus W_2$ si y sólo si todo vector $\overline{v} \in V$

puede ser expresado en forma única como

$\overline{v} = \overline{w}_1 + \overline{w}_2$, donde $\overline{w}_1 \in W_1$ y $\overline{w}_2 \in W_2$.

DEMOSTRACION

Supongamos que $V = W_1 \oplus W_2$.

Entonces, por X.5.13, $V = W_1 + W_2$ y por X.5.12 cualquier vector $\overline{v} \in V$ puede expresarse como

$\overline{v} = \overline{w}_1 + \overline{w}_2$, donde $\overline{w}_1 \in W_1$ y $\overline{w}_2 \in W_2$

Para probar que tal expresión es única consideremos otros dos vectores $\bar{z}_1 \in W_1$ y $\bar{z}_2 \in W_2$ tales que

$$\bar{v} = \bar{z}_1 + \bar{z}_2$$

Entonces se tendrá que

$$\bar{w}_1 + \bar{w}_2 = \bar{z}_1 + \bar{z}_2$$

y por lo tanto

$$\bar{w}_1 - \bar{z}_1 = \bar{z}_2 - \bar{w}_2$$

pero $(\bar{w}_1 - \bar{z}_1) \in W_1$, $(\bar{z}_2 - \bar{w}_2) \in W_2$ y, por X.5.13, $W_1 \cap W_2 = \{\bar{0}\}$; por lo que la igualdad anterior implica que

$$\bar{w}_1 - \bar{z}_1 = \bar{0}$$

y

$$\bar{z}_2 - \bar{w}_2 = \bar{0}$$

de donde $\bar{w}_1 = \bar{z}_1$ y $\bar{w}_2 = \bar{z}_2$

En consecuencia, la expresión $\bar{v} = \bar{w}_1 + \bar{w}_2$, con $\bar{w}_1 \in W_1$ y $\bar{w}_2 \in W_2$, es única.

Supongamos ahora que $\forall\ \bar{v} \in V$ es única la expresión

$$\bar{v} = \bar{w}_1 + \bar{w}_2, \text{ donde } \bar{w}_1 \in W_1 \text{ y } \bar{w}_2 \in W_2$$

Claramente $V = W_1 + W_2$ puesto que $\forall\ \bar{v} \in V$ existen vectores $\bar{w}_1 \in W_1$ y $\bar{w}_2 \in W_2$ tales que $\bar{v} = \bar{w}_1 + \bar{w}_2$.

Para mostrar que $W_1 \cap W_2 = \{\bar{0}\}$ consideremos un vector - - - - - - $\bar{z} \in (W_1 \cap W_2)$.

Entonces $\bar{z} \in W_1$, por lo que

$$\bar{z} = \bar{z} + \bar{0}, \text{ donde } \bar{z} \in W_1, \text{ y } \bar{0} \in W_2 \qquad ---(1)$$

pero también $\bar{z} \in W_2$, por lo que

$$\bar{z} = \bar{0} + \bar{z}, \text{ donde } \bar{0} \in W_1 \quad y \quad \bar{z} \in W_2 \qquad ---(2)$$

Además $\bar{z} \in V$, por lo que la expresión

$$\bar{z} = \bar{w}_1 + \bar{w}_2, \text{ donde } \bar{w}_1 \in W_1 \quad y \quad \bar{w}_2 \in W_2$$

es única. En consecuencia, de (1) y (2) se sigue que

$$z = \bar{0}$$

por lo que

$$V = W_1 \oplus W_2$$

□

Con relación a las bases y dimensiones de los subespacios que intervienen en una suma directa se tiene el siguiente resultado para el caso de dimensión finita.

X.5.15 TEOREMA

Sean V un espacio vectorial de dimensión finita, W_1, W_2 dos subespacios de V tales que

$$V = W_1 \oplus W_2$$

y A, B bases de W_1 y W_2 respectivamente, entonces:

i) $\dim V = \dim W_1 + \dim W_2$

ii) $A \cup B$ es una base de V.

DEMOSTRACION

Sean $A = \{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\}$ y $B = \{\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n\}$ bases de W_1 y W_2

respectivamente.

Por X.5.14, $\forall\ \bar{v} \in V$ existen $\bar{w}_1 \in W_1$ y $\bar{w}_2 \in W_2$ tales que

$$\bar{v} = \bar{w}_1 + \bar{w}_2$$

Como A y B son bases de W_1 y W_2

$$\bar{w}_1 = \sum_{i=1}^{n} \alpha_i \bar{x}_i$$

y

$$\bar{w}_2 = \sum_{i=1}^{n} \beta_i \bar{y}_i$$

por lo que

$$\bar{v} = \sum_{i=1}^{m} \alpha_i \bar{x}_i + \sum_{i=1}^{n} \beta_i \bar{y}_i$$

y el conjunto

$$A \cup B = \{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m, \bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n\}$$

es un generador de V.

Para mostrar que es linealmente independiente consideremos la expresión

$$\lambda_1\bar{x}_1 + \ldots + \lambda_m\bar{x}_m + \mu_1\bar{y}_1 + \ldots + \mu_n\bar{y}_n = \bar{0}$$

esto es

$$\sum_{i=1}^{m} \lambda_i \bar{x}_i + \sum_{i=1}^{n} \mu_i \bar{y}_i = \bar{0}$$

puesto que el vector cero puede expresarse como

$$\bar{0} + \bar{0} = \bar{0}$$

y $\bar{0} \in V$, por X.5.14 la descomposición anterior es única. En consecuencia

$$\sum_{i=1}^{m} \lambda_i \overline{x}_i = \overline{0} \quad \text{y} \quad \sum_{i=1}^{n} \mu_i \overline{y}_i = \overline{0}$$

pero A y B son bases de W_1 y W_2, por lo que las expresiones anteriores implican que

$$\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0 \quad \text{y} \quad \mu_1 = \mu_2 = \ldots = \mu_n = 0$$

En consecuencia $A \cup B$ es una base de V y

$$\dim V = m + n = \dim W_1 + \dim W_2$$

$\square$

Como consecuencia de los teoremas IX.5.16, X.5.14 y X.5.15 se concluye el siguiente resultado para espacios con producto interno.

X.5.16 TEOREMA

Si V es un espacio de dimensión finita con producto interno y W es un subespacio de V, entonces:

i) $V = W \oplus W^{\perp}$

ii) $\dim V = \dim W + \dim W^{\perp}$

Cuando $V = W \oplus W^{\perp}$ la expresión

$$\overline{v} = \overline{w}_1 + \overline{w}_2, \text{ donde } \overline{w}_1 \in W \quad \text{y} \quad \overline{w}_2 \in W^{\perp}$$

se conoce como "descomposición ortogonal" (o descomposición en dos direcciones ortogonales) del vector $\overline{v}$. La razón es evidente puesto que, en este caso, $\overline{w}_1$ y $\overline{w}_2$ son vectores ortogonales.

Consideremos ahora un espacio vectorial V, de dimensión finita y con producto interno, y pasemos a ocuparnos de las condiciones que debe satisfacer un operador lineal $T: V \to V$ para que exista una ba se ortonormal de V formada por vectores característicos de T.

En este punto el problema básico es la existencia de valores característicos del operador y, por lo tanto, de vectores característicos. Dicho problema es de naturaleza diferente cuando el espacio vectorial está definido sobre el campo de los números complejos y cuando está definido sobre el campo de los números reales.

En el primer caso, por el teorema fundamental del álgebra (III.3.2), el polinomio característico tiene al menos un raíz en C, la cual es un valor característico del operador.

Sin embargo, si el espacio vectorial está definido sobre R, la raíz cuya existencia garantiza el mencionado teorema puede no ser un número real y, en consecuencia, no ser un valor característico del operador.

De hecho, en el caso real existen operadores que no tienen valores característicos. Vimos un ejemplo de ello en la sección X.4 con el operador $T: R^2 \to R^2$ que gira un vector cualquiera del plano un ángulo de 60°.

Requerimos entonces una condición suficiente para garantizar la existencia de valores característicos de un operador en el caso real. Tal condición es la simetría del operador, como lo indica el siguiente teorema.

X.5.17 TEOREMA

Sea V un espacio vectorial sobre R, de dimensión finita y con producto interno, y sea $T: V \to V$ un operador lineal. Si T es simétrico entonces tiene al menos un valor característico.

DEMOSTRACION

Sea T un operador simétrico y sea A la representación matricial de T en una base ortonormal de R. Por X.5.11, A es una matriz real simétrica y, por lo tanto, hermitiana.

Definamos ahora un operador $H: C^n \rightarrow C^n$ mediante

$$H(\bar{v}) = A\bar{v}, \ \forall \ \bar{v} \ \varepsilon \ C^n$$

El operador H, así definido, es hermitiano para el producto interno usual en C^n; ya que $\forall \ \bar{u}; \ \bar{v} \ \varepsilon \ C^n$:

$$(H(\bar{u})|\bar{v}) = (A\bar{u}|\bar{v})$$

$$= \sum_{i=1}^{n} \left[\left(\sum_{j=1}^{n} a_{ij}u_j\right)(\bar{v}_i)\right]^{\dagger}$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} u_j a_{ij} \bar{v}_i\right)$$

$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} u_i a_{ji} \bar{v}_j\right)$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} u_i a_{ji} \bar{v}_j\right)$$

$$(H(\bar{u})|\bar{v}) = \sum_{i=1}^{n} u_i \left(\sum_{j=1}^{n} a_{ji} \bar{v}_j\right)$$

pero A es una matriz hermitiana, por lo que $a_{ji} = \bar{a}_{ij}$ y entonces

$$(H(\bar{u})|\bar{v}) = \sum_{i=1}^{n} u_i \overline{\left(\sum_{j=1}^{n} a_{ij} v_j\right)}$$

$$= (\bar{u}|A\bar{v})$$

$$(H(\bar{u})|\bar{v}) = (\bar{u}|H(\bar{v}))$$

† *Aquí* $\bar{v}_i$ *representa el conjugado de la componente i-ésima de* $\bar{v}$.

y en consecuencia H es un operador hermitiano.

Puesto que C^n es un espacio vectorial sobre C, por el teorema fundamental del álgebra H: $C^n \rightarrow C^n$ tiene al menos un valor característico λ, y como H es hermitiano, por i) de X.5.10, λ es real.

Además, como la representación matricial de H en la base

$$\{(1, 0,\ldots, 0), (0, 1,\ldots, 0),\ldots, (0, 0,\ldots, 1)\}$$

es precisamente la matriz A, λ es un valor característico de T.

□

Ahora que tenemos una condición suficiente para la existencia de valores característicos de un operador en el caso real, podemos demostrar el teorema X.5.18 que se enuncia a continuación.

En realidad se trata de dos teoremas, uno para el caso complejo y otro para el caso real; pero debido a que sus argumentos son semejantes los presentamos en un mismo enunciado, válido para el caso complejo, indicando entre paréntesis los términos correspondientes al caso real.

X.5.18 TEOREMA

Sea V un espacio vectorial sobre C (sobre R), de dimensión finita y con producto interno, y sea T: V → V un operador lineal.

Existe una base ortonormal de V formada por vectores característicos de T si y sólo si T es normal (simétrico).

DEMOSTRACION

La demostración de ambos teoremas se hará también simultáneamente, haciendo las aclaraciones necesarias para la validez de los argu - mentos en el caso real.

Se demostrará primero que si T es normal (simétrico) entonces exis te una base ortonormal de V formada por vectores característicos de T. La demostración se hará por inducción matemática sobre la dimensión de V.

Para dim V = 1 la validez del enunciado es inmediata puesto que T tiene un valor característico (aun en el caso real ya que el polinomio característico es de primer grado con coeficientes en R y su raíz es real). Para cualquier vector característico $\overline{v}$ correspon - diente a dicho valor el conjunto

$$\{\frac{1}{||\overline{v}||}\overline{v}\}$$

es una base de V formada por un vector característico de T de norma uno.

Sea ahora dim V = n > 1.

El polinomio característico de T es de grado n y por el teorema fundamental del álgebra (por el teorema X.5.17) T tiene al menos un valor característico λ_1 y un vector característico correspon - diente $\overline{v}_1$. Hagamos

$$\overline{e}_1 = \frac{1}{||\overline{v}_1||}\overline{v}_1$$

y sea W el subespacio de V generado por $\overline{e}_1$; es decir

$$W = \{\alpha\overline{e}_1 \mid \alpha \in K\}$$

W es invariante bajo T^* ya que $\forall\ \overline{w} \in W$:

$$\begin{aligned} T^*(\overline{w}) &= T^*(\alpha\overline{e}_1) \\ &= \alpha T^*(\overline{e}_1) \\ &= \alpha(\overline{\lambda}_1\overline{e}_1) \\ &= \overline{\lambda}_1(\alpha\overline{e}_1) \\ T^*(\overline{w}) &= \overline{\lambda}_1\overline{w} \end{aligned}$$

por lo que $T^*(\overline{w}) \in W$.

En consecuencia, por X.5.5, $W^\perp$ es invariante bajo $(T^*)^*$ y, por X.5.6, es invariante bajo T.

Entonces podemos definir un operador

$$S: W^\perp \to W^\perp$$

como

$$S(\overline{v}) = T(\overline{v});\ \forall\ \overline{v} \in W^\perp.$$

(Nota: Algunos autores llaman al operador S "la restricción de T a $W^\perp$" y lo representan con $T_{W^\perp}$, debido a que tiene el mismo criterio de asignación que T sólo que actúa en un subespacio de V. Observe que para poder definir la restricción de un operador $T: V \to V$ a un subespacio de V, en este caso $W^\perp$, se requiere que dicho subespacio sea invariante bajo T).

El operador S es claramente un operador normal (simétrico), puesto que T lo es.

Además, por X.5.16

$$\dim W^\perp = \dim V - \dim W = n - 1$$

y podemos aplicar a $S: W^\perp \to W^\perp$ la hipótesis de inducción. Esto es, considerar que existe una base ortonormal de $W^{\perp}$ formada por $n - 1$ vectores característicos de S, digamos

$$\{\overline{e}_2, \overline{e}_3, \ldots, \overline{e}_n\}$$

los cuales son también vectores característicos de T.

Luego, por X.5.16 nuevamente

$$V = W \oplus W^{\perp}$$

y por X.5.15 se tiene que el conjunto

$$\{\overline{e}_1, \overline{e}_2, \overline{e}_3, \ldots, \overline{e}_n\}$$

es una base ortonormal de V formada por vectores característicos de T.

La demostración del recíproco es más simple y se presenta a continuación.

Sea $B = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_n\}$ una base ortonormal de V formada por vectores característicos de T, con $T(\overline{e}_i) = \lambda_i \overline{e}_i$.

Si $\overline{v}$ es un vector cualquiera de V se tiene que

$$\overline{v} = \sum_{i=1}^{n} \alpha_i \overline{e}_i$$

y en consecuencia

$$(T\circ T^*)(\overline{v}) = T[T^*(\sum_{i=1}^{n} \alpha_i \overline{e}_i)] \qquad \text{por X.3.6}$$

$$= T[\sum_{i=1}^{n} \alpha_i T^*(\overline{e}_i)] \qquad \text{por X.5.3}$$

$$(T\circ T^*)(\overline{v}) = T(\sum_{i=1}^{n} \alpha_i \overline{\lambda}_i \overline{e}_i) \qquad \text{por ii) de X.5.8}$$

$$(T \circ T^*)(\bar{v}) = \sum_{i=1}^{n} \alpha_i \bar{\lambda}_i T(\bar{e}_i)$$ puesto que T es lineal

$$(T \circ T^*)(\bar{v}) = \sum_{i=1}^{n} \alpha_i \bar{\lambda}_i \lambda_i \bar{e}_i$$ por hipótesis

por otra parte

$$(T^* \circ T)(\bar{v}) = T^*[T(\sum_{i=1}^{n} \alpha_i \bar{e}_i)]$$ por X.3.6

$$= T^*(\sum_{i=1}^{n} \alpha_i T(\bar{e}_i))$$ por linealidad de T

$$= T^*(\sum_{i=1}^{n} \alpha_i \lambda_i \bar{e}_i)$$ por hipótesis

$$= \sum_{i=1}^{n} \alpha_i \lambda_i T^*(\bar{e}_i)$$ por X.5.3

$$= \sum_{i=1}^{n} \alpha_i \lambda_i \bar{\lambda}_i \bar{e}_i$$ por ii) de X.5.8

$$(T^* \circ T)(\bar{v}) = \sum_{i=1}^{n} \alpha_i \bar{\lambda}_i \lambda_i \bar{e}_i$$ por vi) de VIII.3.4

se tiene entonces que

$$T \circ T^* = T^* \circ T$$

por lo que T es normal.

Además, para el caso real se tiene que

$$T^*(\bar{v}) = T^*(\sum_{i=1}^{n} \alpha_i \bar{e}_i)$$

$$= \sum_{i=1}^{n} \alpha_i T^*(\bar{e}_i)$$

$$T^*(\bar{v}) = \sum_{i=1}^{n} \alpha_i \bar{\lambda}_i \bar{e}_i$$

$$T^*(\bar{v}) = \sum_{i=1}^{n} \alpha_i \lambda_i \bar{e}_i$$

$$T^*(\bar{v}) = T(\bar{v})$$

por lo que $T^* = T$ y T es simétrico.

Esto completa la demostración del teorema.

$\square$

El teorema X.5.18 establece una condición necesaria y suficiente para la existencia de una base ortonormal formada por vectores característicos del operador.

Por X.4.7 resulta que esta condición es suficiente para la existencia de una representación matricial diagonal.

Tal condición, sin embargo, no es necesaria; puesto que la base de vectores característicos a que se refiere X.4.7 puede no ser una base ortonormal.

Un ejemplo de esto último lo tenemos en la transformación $T: R^2 \to R^2$ definida por

$$T(x, y) = (2x + y, 6x + y)$$

que utilizamos en la sección X.4.

Para dicho operador se obtuvo que sus valores característicos son

$$\lambda_1 = 4$$

y

$$\lambda_2 = -1$$

y sus correspondientes vectores característicos son de la forma

$(a,\ 2a)$ con $a \neq 0$

y

$(b,\ -3b)$ con $b \neq 0$

De esta manera, cualquier conjunto

$$B = \{(a,\ 2a),\ (b,\ -3b)\} \text{ con } a,\ b \neq 0$$

es una base de R^2 formada por vectores característicos de T y, por X.4.7, su representación matricial en dicha base es

$$M_B^B(T) = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$$

la cual es una matriz diagonal.

Se deja al lector como ejercicio demostrar que en este ejemplo no existe una base ortonormal de R^2 formada por vectores característicos de T, para el producto interno usual en R^2.

Para concluir este apartado enunciaremos la generalización de los conceptos de suma y suma directa al caso de n subespacios y estableceremos algunos resultados relacionados con estos conceptos que se emplearán en el apartado siguiente.

X.5.19 DEFINICION

Sea V un espacio vectorial sobre un campo K y sean $S_1,\ S_2, \ldots,\ S_n$ subconjuntos de V. La suma de estos subconjuntos es el conjunto

$$S_1 + S_2 + \ldots + S_n = \{\overline{v}_1 + \overline{v}_2 + \ldots + \overline{v}_n \mid \overline{v}_i \in S_i;\ i = 1,\ 2, \ldots,\ n\}$$

al que representamos mediante

$$\sum_{i=1}^{n} S_i$$

Es claro que cuando $S_1, S_2, \ldots, S_n$ son subespacios de V, el conjunto $\sum_{i=1}^{n} S_i$ es también un subespacio de V.

Para el concepto de suma directa definido en X.5.13 se tiene la siguiente generalización

X.5.20 DEFINICION

Sea V un espacio vectorial sobre un campo K y sean $W_1, W_2, \ldots, W_n$ subespacios de V. Se dice que V es la suma directa de - - - - - $W_1, W_2, \ldots, W_n$ y se escribe

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_n$$

cuando

$$V = \sum_{i=1}^{n} W_i \quad y \quad W_i \cap \sum_{j \neq i} W_j = \{\overline{0}\}, \text{ para } i = 1, 2, \ldots, n$$

El símbolo $\sum_{j \neq i} W_j$ se emplea para representar la suma de los subespacios W_j para todos los valores de j, entre 1 y n, que son diferentes del valor i.

El teorema X.5.14 también se geraliza de la siguiente manera

X.5.21 TEOREMA

Sea V un espacio vectorial sobre un campo K y sean $W_1, W_2, \ldots, W_n$ subespacios de V:

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_n$$

si y sólo si todo vector $\overline{v} \in V$ puede ser expresado de manera única como

$$\overline{v} = \overline{w}_1 + \overline{w}_2 + \ldots + \overline{w}_n, \text{ donde } \overline{w}_i \in W_i$$

Los resultados de X.5.15 también son válidos en el caso general; esto es, si

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_n$$

y B_i es una base de W_i, entonces

$$\dim V = \dim W_1 + \dim W_2 + \dots + \dim W_n$$

y $\bigcup_{i=1}^{n} B_i$ es una base de V

donde $\bigcup_{i=1}^{n} B_i$ representa la unión de las bases $B_1, B_2, \dots, B_n$

Un resultado importante que relaciona el concepto de suma directa con el problema de diagonalización es el siguiente

X.5.22 TEOREMA

Sea V un espacio vectorial de dimensión n y sea $T: V \to V$ un operador lineal. Existe una matriz diagonal asociada a T, referida a una base, si y sólo si V es la suma directa de los espacios característicos de T.

DEMOSTRACION

Sean $\lambda_1, \lambda_2, \dots, \lambda_k$ los diferentes valores característicos de T, y sea $E(\lambda_i)$ el espacio característico de T correspondiente al valor λ_i.

Supongamos que existe una matriz diagonal asociada a T, referida a una base.

Entonces, por X.4.7 existe una base de V formada por vectores característicos de T y cualquier vector de V es una combinación li -

neal de dichos vectores característicos; en consecuencia

$$V = \sum_{i=1}^{k} E(\lambda_i)$$

Para mostrar que se trata de una suma directa, sean

$$\bar{w}_1 \in E(\lambda_1),\ \bar{w}_2 \in E(\lambda_2), \ldots,\ \bar{w}_k \in E(\lambda_k)$$

Entonces, para un vector cualquiera $\bar{v} \in V$ se tiene que

$$\bar{v} \in \left[E(\lambda_i) \cap \sum_{j \neq i} E(\lambda_j)\right]$$

si y sólo si

$$\bar{v} \in E(\lambda_i) \quad \text{y} \quad \bar{v} \in \sum_{j \neq i} E(\lambda_j)$$

Esto es

$$\bar{v} = \alpha_i \bar{w}_i$$

y

$$\bar{v} = \sum_{j \neq i} \alpha_j \bar{w}_j$$

de donde

$$\alpha_i \bar{w}_i = \sum_{j \neq i} \alpha_j \bar{w}_j$$

por lo que

$$\alpha_i \bar{w}_i - \sum_{j \neq i} \alpha_j \bar{w}_j = \bar{0}$$

Pero como $\bar{w}_1$, $\bar{w}_2, \ldots, \bar{w}_k$ corresponden a valores característicos diferentes, por X.4.2 el conjunto formado por dichos vectores es linealmente independiente y de la expresión anterior se sigue que

$$\alpha_i = 0 \quad \text{y} \quad \alpha_j = 0, \text{ para } j = 1, 2, \ldots, n \text{ con } j \neq i$$

En consecuencia

$$\overline{v} = \overline{0}$$

y

$$E(\lambda_i) \cap \sum_{j \neq i} E(\lambda_j) = \{ \overline{0} \}$$

por lo que V es la suma directa de sus espacios característicos; esto es

$$V = E(\lambda_1) \oplus E(\lambda_2) \oplus \ldots \oplus E(\lambda_k)$$

Sea ahora B_i una base de $E(\lambda_i)$; entonces, sus elementos son vectores característicos de T.

Además si

$$V = E(\lambda_1) \oplus E(\lambda_2) \oplus \ldots \oplus E(\lambda_k)$$

por la generalización de X.5.15, $\bigcup_{i=1}^{k} B_i$ es una base de V que está formada por vectores característicos de T y en consecuencia, de X.4.7, existe una matriz diagonal asociada a T referida a dicha base, con lo que se termina la demostración. $\square$

- Proyecciones ortogonales y el teorema espectral.

Veremos ahora como se puede "descomponer" un operador normal T, expresándolo como combinación lineal de ciertos operadores conocidos como proyecciones ortogonales.

Introduciremos primero algunos conceptos y resultados que son necesarios, empezando con el de proyección.

X.5.23 DEFINICION

Sea V un espacio vectorial y sea W_1 un subespacio de V. Si existe un subespacio W_2 de V tal que

$$V = W_1 \oplus W_2$$

y

$$\bar{v} = \bar{w}_1 + \bar{w}_2, \text{ donde } \bar{w}_1 \in W_1 \text{ y } \bar{w}_2 \in W_2$$

entonces el operador $P: V \rightarrow V$ definido por

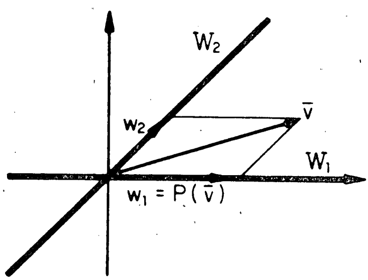
$$P(\bar{v}) = \bar{w}_1, \ \forall \ \bar{v} \in V$$

se llama la proyección sobre W_1 a lo largo de W_2

Con relación a esta definición cabe hacer notar que para un W_1 fijo pueden existir diversos subespacios W_2. De esta manera, el término "a lo largo de" proporciona al operador P la unicidad requerida por el enunciado.

La siguiente figura muestra una interpretación geométrica de la definición X.5.23. En este caso

$$V = R^2, \ W_1 = \{(x, 0) \mid x \in R\} \text{ y } W_2 = \{(x, x) \mid x \in R\}$$



X.5.24 TEOREMA

Sea V un espacio vectorial y sean W_1, W_2 dos subespacios de V. Si P: V $\rightarrow$ V es la proyección sobre W_1 a lo largo de W_2 entonces

$$P(V) = W_1 \quad y \quad N(P) = W_2$$

Es decir que el recorrido de P es el subespacio W_1 y su núcleo es W_2

DEMOSTRACION

Sean $\overline{w}_1 \in W_1$ y $\overline{w}_2 \in W_2$, entonces

$$\overline{w}_1 + \overline{w}_2 \in V, \text{ por X.5.13}$$

luego, existe un vector $\overline{v} \in V$ tal que

$$\overline{v} = \overline{w}_1 + \overline{w}_2$$

es decir, $\forall\ \overline{w}_1 \in W_1$ existe un $\overline{v} \in V$ tal que

$$P(\overline{v}) = \overline{w}_1$$

y, por i) de X.1.2

$$P(V) = W_1$$

Por otra parte, ya que W_2 es subespacio de V, todo vector $\overline{w}_2 \in W_2$ pertenece también a V y, por X.5.14, puede expresarse en forma única como

$$\overline{w}_2 = \overline{u}_1 + \overline{u}_2 \quad \text{donde} \quad \overline{u}_1 \in W_1 \quad y \quad \overline{u}_2 \in W_2$$

Pero, por X.5.13, el único vector que pertenece a $W_1 \cap W_2$ es el ce ro por lo que

$$\bar{w}_2 = \bar{0} + \bar{u}_2$$

y, por X.5.23, $P(\bar{w}_2) = \bar{0}$

luego, por ii) de X.1.2

$$N(P) = W_2$$

□

Un tipo especialmente interesante de proyección se tiene cuando el subespacio W_2 de la definición X.5.23 es el complemento ortogonal de W_1. En este caso la proyección recibe el nombre de "proyección ortogonal" o simplemente "proyección".

X.5.25 DEFINICION

Sean V un espacio vectorial con producto interno, W un subespacio de V y $W^{\perp}$ el complemento ortogonal de W. Si

$$\bar{v} = \bar{w} + \bar{w}', \text{ donde } \bar{w} \in W \text{ y } \bar{w}' \in W^{\perp}$$

entonces el operador $P: V \rightarrow V$ definido por

$$P(\bar{v}) = \bar{w}$$

se llama la proyección ortogonal sobre W.

Este tipo de proyección tiene ciertas características importantes.

En particular, si W es un subespacio de V y P es la proyección ortogonal sobre W; entonces, para cualquier vector $\bar{v} \in V$, $P(\bar{v})$ es el vector de W que más se aproxima a $\bar{v}$.

El lector puede verificar fácilmente que esto coincide con el re -

sultado IX.5.18 (el teorema de proyección). Es conveniente remarcar, sin embargo, que la proyección (ortogonal) que se define en esta sección es un operador, mientras que la proyección a que refiere el teorema IX.5.18 es un vector.

Otras propiedades importantes de las proyecciones ortogonales, cuya demostración se sugiere al lector, son las siguientes.

X.5.26 TEOREMA

Sea V un espacio con producto interno y sea $P: V \to V$ un operador lineal. P es una proyección ortogonal (sobre algún subespacio de V) si y sólo si:

i) $P(V)^{\perp} = N(P)$ y $N(P)^{\perp} = P(V)$

ii) $P \circ P = P = P^*$.

Podemos ahora enunciar, a manera de teorema, la forma como puede descomponerse un operador normal en términos de proyecciones ortogonales.

X.5.27 TEOREMA (ESPECTRAL)

Sea V un espacio vectorial sobre C (sobre R), de dimensión finita y con producto interno, y sea $T: V \to V$ un operador normal (simétrico).

Si $\lambda_1, \lambda_2, \ldots, \lambda_k$ son los diferentes valores característicos de T, $E(\lambda_i)$ es el espacio característico correspondiente a λ_i y P_i es la proyección ortogonal sobre $E(\lambda_i)$, entonces:

i) $T = \lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_k P_k$

ii) $P_1 + P_2 + \ldots + P_k = I$

iii) $P_i \circ P_j = 0$, para $i \neq j$

DEMOSTRACION

i) Como T es normal (simétrico), por X.5.18 existe una base ortonormal de V formada por vectores característicos de T y, por X.5.22, V es la suma directa de sus espacios característicos, esto es

$$V = E(\lambda_1) \oplus E(\lambda_2) \oplus \dots \oplus E(\lambda_k)$$

entonces, por X.5.21, para cualquier vector $\bar{v} \in V$ se tiene que

$$\bar{v} = \bar{w}_1 + \bar{w}_2 + \dots + \bar{w}_k, \text{ donde } \bar{w}_i \in E(\lambda_i) \qquad ---(1)$$

en consecuencia, como T es un operador lineal

$$T(\bar{v}) = T(\bar{w}_1) + T(\bar{w}_2) + \dots + T(\bar{w}_k)$$

y como $\bar{w}_i \in E(\lambda_i)$

$$T(\bar{v}) = \lambda_1\bar{w}_1 + \lambda_2\bar{w}_2 + \dots + \lambda_k\bar{w}_k, \text{ donde } \bar{w}_i \in E(\lambda_i) \qquad ---(2)$$

Por otra parte, sea S_i la suma directa de todos los subespacios $E(\lambda_j)$, con $j \neq i$. Si $\bar{x}$ es un vector cualquiera de S_i se tendrá que

$$\bar{x} = \sum_{j\neq i} \bar{y}_j, \text{ donde } \bar{y}_j \in E(\lambda_j)$$

En consecuencia, si $\bar{w}$ es un vector cualquiera de $E(\lambda_i)$ se tendrá que

$$(\bar{w}|\bar{x}) = (\bar{w}|\sum_{j\neq i} \bar{y}_j)$$

y por las propiedades del producto interno

$$(\bar{w}|\bar{x}) = \sum_{j\neq i} (\bar{w}|\bar{y}_j)$$

Pero como T es normal, por iii) de X.5.8, $(\bar{w}|\bar{y}_j) = 0$ para

$j \neq i$, y en consecuencia

$$(\overline{w}|\overline{x}) = 0$$

De esta manera se tiene que S_i es el complemento ortogonal de $E(\lambda_i)$

Entonces, de la expresión (2) se sigue, por X.5.25, que el operador $P_i: V \to V$ definido por

$$P_i(\overline{v}) = \overline{w}_i, \ \forall \ \overline{v} \ \varepsilon \ V$$

es la proyección ortogonal sobre $E(\lambda_i)$.

Llevando este resultado a (2) escribimos

$$T(\overline{v}) = \lambda_1 P_1(\overline{v}) + \lambda_2 P_2(\overline{v}) + \ldots + \lambda_k P_k(\overline{v})$$

y por X.3.2

$$T(\overline{v}) = (\lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_k P_k)(\overline{v})$$

en consecuencia, por X.3.1

$$T = \lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_k P_k$$

ii) Consideremos nuevamente la expresión

$$\overline{v} = \overline{w}_1 + \overline{w}_2 + \ldots + \overline{w}_k, \text{ donde } \overline{w}_i \ \varepsilon \ E(\lambda_i) \qquad ---(1)$$

Aplicando a ésta el resultado al que llegamos en el desarrollo anterior se tiene que

$$\overline{v} = P_1(\overline{v}) + P_2(\overline{v}) + \ldots + P_k(\overline{v})$$

y en consecuencia

$$\overline{v} = (P_1 + P_2 + \ldots + P_k)(\overline{v})$$

Luego, por definición de transformación identidad y de igualdad de transformaciones

$$I = P_1 + P_2 + \ldots + P_k$$

como se quería.

iii) Sea $\bar{v}$ un vector arbitrario de V, por X.3.6

$$(P_i \circ P_j)(\bar{v}) = P_i[P_j(\bar{v})]$$

entonces

$$(P_i \circ P_j)(\bar{v}) = P_i(\bar{w}_j) \qquad ---(3)$$

Además, como vimos en la demostración de i), cuando $i \neq j$ se tiene que $\bar{w}_j$ pertenece al complemento ortogonal de $E(\lambda_i)$; y como P_i es la proyección ortogonal sobre $E(\lambda_i)$, por X.5.25

$$P_i(\bar{w}_j) = \bar{0}$$

Llevando este resultado a (3)

$$(P_i \circ P_j)(\bar{v}) = \bar{0}$$

De donde

$$P_i \circ P_j = 0$$

Esto completa la demostración del teorema.

□

A la expresión i) del teorema anterior se le conoce como "descomposición espectral" del operador T y, como se sigue de lo que hemos visto, tal descomposición es única (salvo en el orden de los sumandos).

Así, por ejemplo, para el operador $T; R^3 \to R^3$ definido por

$$T(x, y, z) = (3x + 2y, 2x + 3y, 5z)$$

se tiene que

$$M(T) = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

es una matriz simétrica, así que, por i) de X.5.11, T es un operador simétrico.

Entonces, por X.5.27, T tiene una descomposición en términos de proyecciones ortogonales sobre sus espacios característicos.

Para obtener dichas proyecciones, así como la descomposición espectral de T, se determinan primero los valores característicos del operador a partir del polinomio

$$\begin{vmatrix} 3-\lambda & 2 & 0 \\ 2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} = (3 - \lambda)^2(5 - \lambda) - 4(5 - \lambda) = (5-\lambda)(5-\lambda)(1-\lambda)$$

cuyas tres raíces son los valores característicos

5, 5 y 1

de los cuales tenemos sólo dos diferentes. Llamemos a éstos λ_1 y λ_2; es decir

$$\lambda_1 = 5$$
y
$$\lambda_2 = 1$$

Sus correspondientes espacios característicos son

$$E(\lambda_1) = \{(a, a, b) \mid a, b \in R\}$$

y

$$E(\lambda_2) = \{(c, -c, 0) \mid c \in R\}$$

Estos espacios tienen la siguiente interpretación geométrica.

$E(\lambda_1)$, que es un espacio de dimensión dos, corresponde a un plano perpendicular al plano X - Y y que forma un ángulo de 45° con el plano X - Z; mientras que $E(\lambda_2)$, que es un espacio de dimensión uno, corresponde a una recta alojada en el plano X - Y y de ecuación $y = -x$. Claramente, $E(\lambda_1)$ y $E(\lambda_2)$ son complementos ortogonales.

Como vimos en la demostración de X.5.27, R^3 es la suma directa de los espacios característicos de T; esto es

$$R^3 = E(\lambda_1) \oplus E(\lambda_2)$$

por lo que la expresión

$$(x, y, z) = (a, a, b) + (c, -c, 0)$$

es única.

En efecto, se tiene que $\forall\ (x, y, z) \in R^3$

$$(x,y,z) = \left(\frac{x+y}{2}, \frac{x+y}{2}, z\right) + \left(\frac{x-y}{2}, \frac{-x+y}{2}, 0\right)$$

lo que determina las proyecciones

$$P_1(x,y,z) = \left(\frac{x+y}{2}, \frac{x+y}{2}, z\right) \qquad ---(1)$$

y

$$P_2(x,y,z) = \left(\frac{x-y}{2}, \frac{-x+y}{2}, 0\right) \qquad ---(2)$$

Como era de esperarse, se verifica que

$$5P_1(x,y,z) + P_2(x,y,z) = 5\left(\frac{x+y}{2}, \frac{x+y}{2}, z\right) + \left(\frac{x-y}{2}, \frac{-x+y}{2}, 0\right)$$

$$= (3x + 2y,\ 2x + 3y,\ 5z)$$

$$5P_1(x,y,z) + P_2(x,y,z) = T(x,y,z)$$

por lo que

$$5P_1 + P_2 = T$$

con P_1 y P_2 definidos por (1) y (2), es la descomposición espectral del operador T del ejemplo.

X.5.28 EJERCICIOS

1.- Obtener el vector $\bar{v} \in R^3$ tal que

$$\phi(\bar{u}) = (\bar{u}|\bar{v})\ \forall\ \bar{u} \in R^3$$

considerando el producto usual en R^3 y si

$$\phi(x, y, z) = 2x - 3y + z$$

2.- Si $T: C^2 \to C^2$ es un operador lineal definido por

$$T(x, y) = (x + iy, -3ix + (1 - i)y)$$

a) Determinar la regla que define a T^*

b) Verificar que se satisface $(T(\bar{u})|\bar{v}) = (\bar{u}|T^*(\bar{v}))$ con el producto interno usual en C^2

3.- Para el operador $T: R^3 \to R^3$ definido por

$$T(x, y, z) = (x - y, 3y + z, 2x - 2y)$$

verificar que

$$T^*(V) = N(T)^{\perp} \quad \text{y} \quad N(T^* \circ T) = N(T)$$

4.- Demostrar que

$$(\alpha T)^* = \bar{\alpha} T^*$$

y

$$(S \circ T)^* = T^* \circ S^*$$

de dos maneras diferentes, una de ellas directamente a partir de la definición de T^*.

5.- Para cada uno de los operadores $T: C^2 \to C^2$ definidos por las siguientes reglas:

a) $T(x, y) = ((1 - i)x + iy, ix + (1 + i)y)$

b) $T(x, y) = (ix + (2 - i)y, x + iy)$

Determinar si es normal y en caso afirmativo verificar que se cumplen las propiedades ii) y iii) de X.5.8.

6.- Hallar los valores de $p, q \in R$ tales que el operador $T: C^2 \to C^2$ definido por $T(x, y) = (px + (2 + i)y, (p - i)x + iqy)$ sea hermitiano. Considérese el producto interno usual en C^2.

7.- Demostrar que si T es un operador unitario en un espacio vectorial V con producto interno, entonces

a) $(T(\bar{v}) | T(\bar{w})) = (\bar{v} | \bar{w}) \ \forall \ \bar{v}, \bar{w} \in V$

b) $\|T(\bar{v})\| = \|\bar{v}\|, \ \forall \ \bar{v} \in V$

Es decir, T preserva tanto el producto interno como la norma.

8.- a) Demostrar que R^3 es la suma directa de los subespacios

$W_1 = \{(x,x,z) | x, z \in R\}$ y $W_2 = \{(x,2x,x) | x \in R\}$

b) Hallar otras dos descomposiciones de R^3 en sumas directas una de las cuales sea como en X.5.16 y verificar que se cumple X.5.15

9.- Para cada uno de los operadores $T: V \to V$ siguientes

a) $T(x,y) = ((2-2i)x - iy, -ix + (2+2i)y)$ donde $V = C^2$

b) $T(x,y) = (2x - 3y, -3x + y)$ donde $V = R^2$

verificar que satisfacen las condiciones de X.5.18 y hallar una base de V formada por vectores característicos de T.

10.- Verificar que el operador $T: R^3 \to R^3$ definido por

$T(x, y, z) = (3x, x + y - 2z, x - 2y + z)$

es diagonalizable y que su dominio es la suma directa de sus espacios característicos.

11.- a) Obtener la proyección sobre el subespacio de R^3

$W_1 = \{(x, y, x) \mid x, y \in R\}$, a lo largo de cada uno de los siguientes subespacios

$W_2 = \{(x, 2x, 3x) \mid x \in R\}$

$W_3 = \{(0, 0, z) \mid z \in R\}$

b) Obtener la proyección ortogonal sobre W_1

c) Hallar la proyección ortogonal sobre cada uno de los si - guientes subespacios de R^3

$W_4 = \{(x, 0, 0) \mid x \in R\}$

$W_5 = \{(0, y, 0) \mid y \in R\}$

$W_6 = \{(x, y, 0) \mid x, y \in R\}$

$W_7 = \{(x, y, z) \mid x, y, z \in R\}$

12.- Sean V un espacio vectorial con producto interno y $P: V \to V$ un operador lineal. Demostrar que P es una proyección ortogo nal (sobre algún subespacio de V) si y sólo si

$P \circ P = P = P^*$

13.- Obtener la descomposición espectral de los siguientes operadores

a) $T: R^3 \rightarrow R^3$; $T(x,y,z) = (-3x+2y-z,\ 2x-2z,\ -x-2y-3z)$

b) $S: C^3 \rightarrow C^3$; $S(x,y,z) = (x+iy,\ x+(2+i)y,\ z)$