ALGEBRA

Fecha	Tema	Horario	Profesor
Del 22 de Febrero al 19 de Marzo	1. CONJUNTOS	de 20 a 22 h	Ing. Héctor Godínez Cabrer

Notación y ejemplos, relaciones en conjuntos y operaciones, producto cartesiano.

2. NUMEROS REALES E IMAGINARIOS

Números naturales, enteros, racionales, irracionales y reales, números complejos.

3. ESTRUCTURAS ALGEBRAICAS

Operaciones binarias y sus propiedades, grupos, anillos y campos.

4. CONJUNTOS METRICOS

Definición y ejemplos, conceptos topológicos y funciones.

5. ESPACIOS VECTORIALES

Definición y propiedades, dependencia lineal, base y dimensión, subespacios, transformaciones lineales.

6. MATRICES

Definición y ejemplos, operaciones con matrices, matrices adjuntas e inversas, rango, operaciones elementales y equivalencia de matrices, determinantes, sistemas de ecuaciones lineales.

DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

ACTUALIZACION EN MATEMATICAS PARA INGENIEROS

A L G E B R A

FEBRERO, 1982

cio de Minería Calle de Tacuba 5 primer piso México 1, D. F. Tel: 521-40-20 Apdo. Postal M-2285

CONJUNTOS

POR

ARTURO DELGADO R.

Profesor de la División de Estudios de Posgrado de la Facultad de Ingeniería de la U. N. A. M.

CONJUNTOS. (o clases). Introducción: La teoría de los conjuntos es hoy en día básica en el estudio de casi todas las ramas de las matemáticas: teoría de probabilidades, análisis matemático, circuitos eléctricos, lógica matemática, etc.

El establecimiento de la teoría de los conjuntos se atribuye a Georg Cantor (1845-1918).

El concepto de "conjunto" no se define en forma precisa. Para explicar lo que se entiende por conjunto, se da la idea intuitiva:

"Un conjunto es una colección o agregado de objetos bien definidos, los cuales deben poseer una propiedad o atributo característico que no deje lugar a duda si dicho objeto pertenece o no a la colección*"

Ejemplo: Los ex-presidentes de la República mexicana; otro ejemplo de conjunto podría ser el de los números primos mayores que 10 y menores que 211; etc.

Los elementos de un conjunto pueden quedar definidos al enumerar éstos; por ejemplo: el conjunto constituido por los tres números: 2, 5, 9,; etc.

Pueden, en otros casos, distinguirse los elementos de un conjunto, citando una propiedad común a todos ellos; por ejemplo: el conjunto de los números pares.

* Obsérvese que no se exige homogeneidad, cada objeto definido en un conjunto dado se denomina "elemento" del conjunto.

De lo anterior se infiere que los conjuntos pueden tener un número finito o infinito de elementos.

Notación. Se acostumbra designar los conjuntos con letras mayúsculas.

Los elementos de un conjunto se distinguen con letras minúsculas.

Cuando un conjunto queda definido al enumerar sus elementos, se escribe:

$$A = \{a,e,i,o,u\} = \{e,u,a,o,i\}, \text{ etc.}$$

$$B = \{2,5,9\} = \{2,9,5\} = \{9,5,2\}, \text{ etc.}$$

Para indicar que el elemento 5 está en el conjunto B se anota:

$$5 \in B$$

Si un elemento no está en un conjunto dado, se denota:

$$8 \notin B$$

Cuando el conjunto se define expresando una propiedad común de todos sus elementos, se emplea una línea vertical | (o a veces dos puntos :) cuyo significado es "tal que"

Ejemplos:

$$C = \{x \mid x > 25\}$$

$D = \{l \mid l$ es un libro de la Biblioteca Nacional$\}$

Subconjunto. Se dice que el conjunto P es un "subconjunto" del conjunto R, si cada elemento de P es también elemento de R. La misma idea se expresa diciendo que P está contenido en R; o bien, que R contiene a P. (P está incluído en R). La noción de subconjunto se expresa en forma simbólica como sigue:

$$P \subseteq R$$

$$R \supseteq P$$

Si el conjunto P tiene menos elementos que el conjunto R, se dice que P es un subconjunto "propio" de R, y ello se indica:

$$P \subset R$$

Ejemplo: sea: A el conjunto de letras del alfabeto castellano; siendo B el conjunto de vocales del mismo alfabeto, podemos entonces escribir

$$A \supset B \quad o \quad B \subset A$$

Después de dar estas definiciones preliminares, enfocaremos nuestra atención a la elaboración de definiciones y reglas que nos permiten construir una álgebra de conjuntos.

Igualdad.- Se afirma que dos conjuntos A y B son iguales si, y sólo si, ambos conjuntos consisten de los mismos elementos; esto es, cada elemento de A pertenece a B y viceversa. Ejemplo

$$A = \{a,b,c\} \;;\; B = \{x,y,z\} = \{x,z,y,z,z,y,x\}$$

Si A = B, debe tenerse que el elemento a de A es exactamente el mismo que alguno de los elementos de B; así por ejemplo puede tenerse: $a = z$; $b = x$; $c = y$

Una técnica usual para demostrar la igualdad de dos conjuntos consiste en hacer ver que se cumplan las dos condiciones siguientes:

Si $A \subseteq B$; además $B \subseteq A$, entonces $A = B$

Conjunto vacío (ó nulo).- Conjunto vacío es aquel conjunto que no contiene elementos. El conjunto vacío se representa con el símbolo ϕ. Ejemplo:

$$\phi = \{ \text{seres humanos vivos mayores de 1000 años} \}$$

Obsérvese que $\phi \neq \{\phi\}$; puesto que el conjunto ϕ del primer miembro, por definición, no contiene ningún elemento; en tanto que el conjunto $\{\phi\}$ del segundo miembro contiene un elemento.

El conjunto vacío ϕ se considera como un subconjunto de cualquier conjunto sea cual fuere este último. $\phi \subseteq A$, para todo A.

Conjunto universal (ó universo).- El conjunto universal es el conjunto del cual se derivan todos los subconjuntos que pueden intervenir en un problema particular. Designaremos el conjunto universal con el símbolo $\bar{U}$.

El propósito del conjunto $\bar{U}$ es el de evitar paradojas.

Ejemplo:

Si consideramos como universo el conjunto de números reales, se tiene por ejemplo: $\ln 1 = 0$; en tanto que el logaritmo natural de un número negativo no existe. En cambio, si el universo lo constituye el conjunto de números complejos, se encuentra lo siguiente:

$$\ldots -1) = i\pi$$

Cuando se presen[illegible] ignar valores a una variable x en una ecuació[illegible] de valores de x son admisibles; este conjunto de valores [illegible]ituye el universo.

Obsérvese que el conjunto $\bar{U}$ puede ser considerado como un subconjunto de sí mismo. Otra observación: $\phi \subseteq A \subseteq \bar{U}$; para todo A.

Conjunto potencia.- Sea A un conjunto constituido por un número n de elementos. El conjunto potencia de A es el conjunto cuyos elementos son precisamente cada uno de los 2^n subconjuntos que es factible formar al combinar los n elementos del conjunto A.

El conjunto vacío ϕ, y el conjunto universal $\bar{U}$ deben formar parte de los 2^n subconjuntos elementos del conjunto potencia; En efecto:

Sea C_n^r $\binom{n}{r}$ las combinaciones de n objetos tomados en subconjuntos de r objetos ($r = n$). El número total posible de subconjuntos que resultan es:

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots \binom{n}{n-1} + \binom{n}{n} = (1 + 1)^n = 2^n$$

donde: $\binom{r}{0}$ nos da el conjunto vacío ϕ en tanto que $\binom{n}{n}$ da el conjunto universal $\bar{U}$.

Notación: El conjunto potencia de A se designa 2^A. Ejemplo: Sea $A = \{a, b\}$

$$2^A = \{\phi\ \{a\}, \{b\}, \{a, b\}\}$$

Definición simbólica de conjunto potencia:

$$2^A = \{B \mid B \subseteq A\}$$

Ejemplos: 1) Si $A \subseteq B$ y $B \subseteq A$, demostrar que $A = B$

Solución:

Si $x \in A \rightarrow x \in B$, puesto que $A \subseteq B$

Por otro lado:

Si $x \in B \rightarrow x \in A$, ya que $B \subseteq A$

$\rightarrow$ A y B tienen los mismos elementos $\rightarrow A = B$

2) Si $A \subseteq B$ y $B \subseteq C$, demostrar que $A \subseteq C$

Debemos hacer ver que cada elemento en A está también en C.

Solución:

Si $x \in A \rightarrow x \in B$, ya que $A \subseteq B$; pero por ser $B \subseteq C \rightarrow x \in C$; luego $x \in A \rightarrow x \in C$, por lo cual se deduce que $A \subseteq C$

3) Si $A \subset B$ y $B \subseteq C$, demostrar que $A \subset C$.

Puesto que $A \subset B$ y $B \subseteq C$, resulta que, cuando más, $A \subseteq C$

Pero si $A = C$; siendo por hipótesis $B \subseteq C$, $\rightarrow B \subseteq A$, lo cual contradice la hipótesis de que $A \subset B$; por lo cual $A \neq C$ resultando, por lo tanto, $A \subset C$

4) Determinar el conjunto potencia 2^A; siendo, $A = \{\{a,b\},c\}$.

Solución

Puesto que el conjunto A tiene 2 elementos: $\{a,b\}$ y $c \rightarrow 2^A$ contendrá $2^2 = 4$ elementos:

$$2^A = \{\emptyset \quad \{\{a, b\}\} \quad \{c\} \quad \{\{a, b\}, c\}\}$$

5) Diga si las siguientes afirmaciones son correctas o incorrectas:

a) $\emptyset = \{0\}$

b) $0 = \{\emptyset\}$

c) $\{x\} \subset \{\{x\}\}$

d) $\{x\} \in \{\{x\}\}$

e) $\subset \{\{x\}\}$

f) $\{\{x\},\{y\}\} \subset \{\{x\}, \{x, y\}\}$

Solución:

a) Incorrecto: $\emptyset$ no tiene elementos; $\{0\}$ tiene un elemento.

b) Incorrecto: 0 no es un conjunto

c) Incorrecto: elemento $\neq$ conjunto (ver inciso d)

d) Correcto: $\{x\}$ es elemento del conjunto $\{\{x\}\}$

e) Correcto: $\emptyset \subseteq A$, para todo conjunto A

f) Incorrecto: $\{y\} \notin \{\{x\},\{x, y\}\}$

6) Si a,b,c son elementos cualesquiera, determinar qué relación debe existir entre ellos, de modo que sea cierta la igualdad:

$$\{\{a,b\}\} = \{\{a,b,c\}, \{c,b\}\}$$

Dado que el conjunto del 1er. miembro está constituido por un solo elemento que es el conjunto a,b , para que pueda verificarse la igualdad, debe tenerse que en el segundo miembro:

$$\{a,b,c\} = \{c,b\}$$

debiendo estos dos conjuntos ser iguales al conjunto $\{a,b\}$ del primer miembro:

$$\{a,b\} = \{a,b,c\} = \{c,b\}$$

por lo cual debe tenerse que $a = c$

$$\{a,b\} = \{a,b,a\} = \{a,b\}$$

En conclusión: si $a = c$:

$$\{\{a,b\}\} = \{\{a,b,a\}, \{a,b\}\}$$

Operaciones con conjuntos.-

Unión.- Se entiende por "unión" de dos conjuntos A y B, el conjunto C que resulta al considerar C constituido por los elementos que pertenecen a A ó a B ó a ambos.

Notación: $C = A \cup B$

Ejemplo: sean $A = \{1,2,3,4\}$; $B = \{3,4,5\}$

$$C = A \cup B = \{1,2,3,4,5\}$$

$$A \cup B = \{x | x \in A \text{ ó } x \in B\}$$

Intersección.- Dados los conjuntos A y B, el conjunto "intersección" D tiene por elementos aquellos que son comunes a A y a B:

Notación: $D = A \cap B$

$D = A \cap B = \{x | x \in A \text{ y } x \in B\}$,

Ejemplo: $A = \{a,b,c,d\}$; $B = \{c,d,e\}$

$$D = A \cap B = \{c,d\}$$

Conjuntos ajenos (ó desunidos).- Dos conjuntos A y B son ajenos si no tienen ningún elemento común; esto es, si su intersección es el conjunto vacío: $A \cap B = \emptyset$. Ejemplo:

$$A = \{1,2,3\}; \quad B = \{5,6\}.$$

$$A \cap B = \emptyset$$

Diferencia.- La "diferencia" E de dos conjuntos A y B es el conjunto cuyos elementos están en A y no pertenecen a B

Notación: $E = A-B$

$$E = A-B = \{x | x \in A, x \notin B\}$$

Ejemplo: $A = \{1,2,3,4\}$; $B = \{3,4,5\}$

$E = A-B = \{1,2\}$

Complemento.- El "complemento" del conjunto A es el conjunto A', cuyos elementos son los elementos del universo, que no pertenecen a A. (otra notación: $A' = \bar{A} = A^c$)

$$A' = \bar{U}-A = \{x | x \in \bar{U}, x \notin A\}$$

Ejemplo: Sean $\bar{U} = \{a,e,i,o,u\}$; $A = \{a,e\}$

$A' = \{i,o,u\}$

Los conceptos de unión e intersección pueden generalizarse para más de dos conjuntos:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n$$

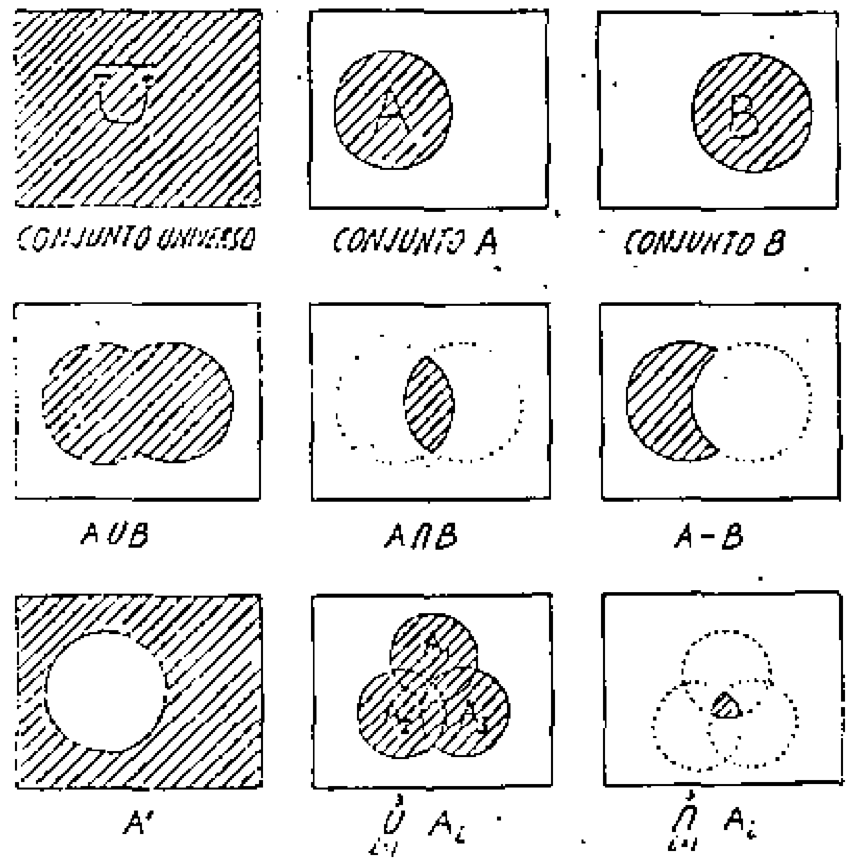
$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n$$

$$\bigcup_{S} A_i = \{x | x \in A_i \text{ para al menos una } i \in S\}$$

$$\bigcap_{S} A_i = \{x | x \in A_i \text{ para toda } i \in S\}$$

siendo el conjunto de subconjuntos $\{A_i | i \in S\}$

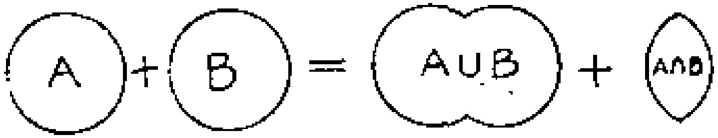
Diagramas de Venn (ó de Euler).- Son esquemas en los cuales los conjuntos se representan como áreas planas.



Fórmulas generales para relacionar número de elementos en conjuntos.- Relacionaremos el número de elementos que hay en dos conjuntos A y B, con el número de elementos en los conjuntos $A \cup B$ y $A \cap B$.

Designemos el número de elementos que hay en un conjunto, digamos el conjunto A como $n\,A$. $n\,A$ es un número, no un conjunto!

Establezcamos la siguiente igualdad entre áreas;



ó sea:

$$n\,A + n\,B = n(A \cup B) + n(A \cap B)$$

Problema: Todos los alumnos de una clase de gimnasia se inscriben para practicar natación (N) ó atletismo (A), ó ambos. Si hay 200 alumnos en la clase, y en la lista de inscritos en natación hay 104 nombres; en tanto que para atletismo hay inscritos 130 ¿cuántos alumnos se inscribieron para practicar ambos deportes?

Solución: a) Empleando la fórmula

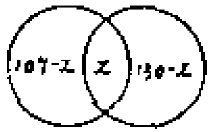
$$n(N \cup A) = 200$$

$$n\,N = 104; \quad n\,A = 130$$

$$n\,N + n\,A = n(N \cup A) + n(N \cap A)$$

$$\therefore \quad n(N \cap A) = 104 + 130 - 200 = 34$$

b) con auxilio de diagrama de Venn.



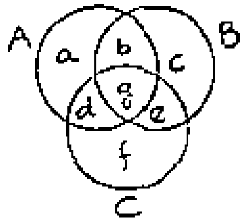
$$(104-x) + x + (130-x) = 200$$

$$234 - x = 200$$

$$x = 34 = n(N \cap A)$$

Comprobar, para 3 conjuntos A, B, C, la siguiente fórmula:

$nA + nB + nC = n(A \cup B \cup C) + n(A \cap B) + n(A \cap C) + n(B \cap C) - n(A \cap B \cap C)$ comprobación:



$$nA = a + b + d + g$$

$$nB = b + c + e + g$$

$$nC = d + e + f + g$$

$$nA + nB + nC =$$

$$= a + 2b + c + 2d + 2e + f + 3g =$$

$$= (a+b+d+g) + (b+c+e+g) + (d+e+f+g) =$$

$$= (a+b+c+d+e+f+g) + (b+g) + (d+g) + (e+g) - g$$

$$= n(A \cup B \cup C) + n(A \cap B) + n(A \cap C) + n(B \cap C) - n(A \cap B \cap C)$$

Fórmula general para relacionar el número de elementos que existen en m conjuntos*

$$n\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{C_m^1} nA_i - \sum_{i<j}^{C_m^2} n(A_i \cap A_j) +$$

$$+ \sum_{i<j<k}^{C_m^3} n(A_i \cap A_j \cap A_k) + \ldots + (-1)^{p+1} \sum_{i<j<\ldots<p}^{C_m^p} n(A_i \cap \ldots \cap A_p) +$$

$$\ldots + (-1)^{m+1} \sum_{i<j<\ldots<m}^{C_m^m} n(A_i \cap \ldots \cap A_m)$$

donde

$$1 \leq p \leq m;\ i,j,k,\ldots = 1,2,3,\ldots,m$$

C_m^p número de combinaciones de m objetos tomados p a p,

$$C_m^p = C_m^{m-p}$$

$$C_m^p = \frac{m(m-1)(m-2)\ldots(m-p+1)}{p!}$$

$p! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-1)p$

$C_m^m = C_m^0 = 1$, para todo $m = 0,1,2,\ldots$

* La fórmula se demuestra por inducción matemática en la segunda parte de estas notas.

Ejemplos:

i) $m = 3$

Solución:

$$n(A_1 \cup A_2 \cup A_3) = nA_1 + nA_2 + nA_3 - n(A_1 \cap A_2) - n(A_1 \cap A_3) - n(A_2 \cap A_3) +$$

$$+ n(A_1 \cap A_2 \cap A_3)$$

Resultado que concuerda con el último problema, si $A_1 = A$; $A_2 = B$; $A_3 = C$.

ii) $m = 4$

Solución:

$$n(A_1 \cup A_2 \cup A_3 \cup A_4) = nA_1 + nA_2 + nA_3 + nA_4 - n(A_1 \cap A_2) - n(A_1 \cap A_3) - n(A_1 \cap A_4) -$$

$$- n(A_2 \cap A_3) - n(A_2 \cap A_4) - n(A_3 \cap A_4) + n(A_1 \cap A_2 \cap A_3) + n(A_1 \cap A_2 \cap A_4) +$$

$$+ n(A_1 \cap A_3 \cap A_4) + n(A_2 \cap A_3 \cap A_4) - n(A_1 \cap A_2 \cap A_3 \cap A_4)$$

iii) $m = 5$

Solución

$$n(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5) = nA_1 + nA_2 + nA_3 + nA_4 + nA_5 -$$

$$- n(A_1 \cap A_2) - n(A_1 \cap A_3) - n(A_1 \cap A_4) - n(A_1 \cap A_5) - n(A_2 \cap A_3) -$$

$$- n(A_2 \cap A_4) - n(A_2 \cap A_5) - n(A_3 \cap A_4) - n(A_3 \cap A_5) - n(A_4 \cap A_5) +$$

$+n(A_1 \cap A_2 \cap A_3)+n(A_1 \cap A_2 \cap A_4)+n(A_1 \cap A_2 \cap A_5)+n(A_1 \cap A_3 \cap A_4)+$

$+n(A_1 \cap A_3 \cap A_5)+n(A_1 \cap A_4 \cap A_5)+n(A_2 \cap A_3 \cap A_4)+n(A_2 \cap A_3 \cap A_5)+$

$+n(A_2 \cap A_4 \cap A_5)+n(A_3 \cap A_4 \cap A_5)-n(A_1 \cap A_2 \cap A_3 \cap A_4)-$

$-n(A_1 \cap A_2 \cap A_3 \cap A_5)-n(A_1 \cap A_2 \cap A_4 \cap A_5)-n(A_1 \cap A_3 \cap A_4 \cap A_5)-$

$-n(A_2 \cap A_3 \cap A_4 \cap A_5)+n(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5)$

Observaciones:

1) El número de términos que debe tener el desarrollo de cada suma $\sum^{C_m^p}$ es precisamente C_m^p. Por ejemplo la suma

$\sum_{i<j}^{C_m^2}$ tiene $C_m^2 = \frac{m(m-1)}{1.2} = \frac{m^2-m}{2}$ términos.

2) Todos los términos que provienen del desarrollo de cualquier suma $\sum$, deben tener el mismo signo.

3) Los signos de las diversas sumas $\sum$, son alternados, y existen m sumas $\sum$ en el segundo miembro.

4) El número total de términos que se obtienen al desarrollar todas las sumas $\sum$ del segundo miembro es 2^m-1. Por ejemplo, en el desarrollo para $m = 5$ se obtuvieron: $2^5-1 = 32-1 = 31$ términos.

5) El último término del desarrollo puede escribirse:

$$\sum_{1<2<\ldots<m}^{C_m^m} n(A_1 \cap \ldots \cap A_m) = n\left(\bigcap_{i=1}^{m} A_i\right)$$

6) La asociatividad de $\bigcup_{i=1}^{m} A_i$ y de $\bigcap_{i=1}^{m} A_i$ se demuestra más adelante.

Problema: En una bolsa hay 50 objetos, de los cuales:

14 son de color gris	$:nG = 14$
13 son radiactivos	$:nR = 13$
9 son en forma de cubo	$:nC = 9$
5 son grises y reactivos	$:n(G \cap R) = 5$
7 son grises y cúbicos	$:n(G \cap C) = 7$
4 son radiactivos y cúbicos	$:n(R \cap C) = 4$
3 son grises, radiactivos y cúbicos	$:n(G \cap R \cap C) = 3$

¿Cuántos objetos no son ni grises, ni radiactivos, ni cúbicos?

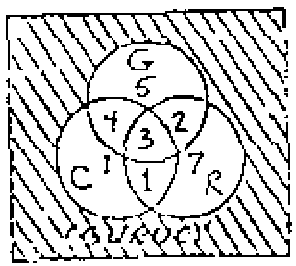
Solución: a) Se busca : $x = n(G \cup R \cup C)' \to x = n(U - G \cup R \cup C)$

$\therefore$ aplicando la fórmula: $14+13+9 = n(G \cup R \cup C)+5+7+4-3$

$\to n(G \cup R \cup C) = 23$

$\to x = 50-23 = 27$

b) Empleando diagramas de Venn:



$n(G \cup R \cup C) =$

$= 4+5+3+2+1+7+1 =$

$= 23$

$U = 50$

$\rightarrow n(G \cup R \cup C)' = 50-23=27$

$\therefore$ <u>27</u> objetos (de los 50) no son, ni grises, ni radiactivos, ni mé - bicos.

<u>Algunos métodos empleados para demostrar relaciones entre conjuntos.</u>- Aun cuando para establecer la validez de una ecuación entre conjuntos es suficiente y basta con emplear una, cualquiera de las técnicas que se detallarán a continuación, sucede, en algunos casos, que un método en particular resulta, para determinados estudiantes, más claro o expedito que los otros métodos.

Para ilustrar los diversos métodos, nos referimos a un ejemplo concreto: supongamos que se desea probar la veracidad de la ecuación $A - B = A \cap B'$

1^{er} método: A partir de las definiciones:

$$A-B = \{x|x \in A \text{ y } x \notin B\}$$

$$= \{x|x \in A \text{ y } x \in B'\}$$

$$= A \cap B'$$

$2°$ método: Haciendo ver que $A-B \subset A \cap B'$; además $A \cap B' \subset A-B \rightarrow$ $\rightarrow A-B = A \cap B'$.

a) Sea $x \in A-B$

$\rightarrow x \in A, x \notin B \rightarrow x \in B'$

$x \in A$ y $x \in B' \rightarrow x \in A \cap B'$

$\rightarrow A-B \subset A \cap B'$ ... (1)

b) Sea $x \in A \cap B'$

$\rightarrow x \in A$ y $x \in B'$

$\rightarrow x \in A$ y $x \notin B \rightarrow x \in A-B$

$\rightarrow A \cap B' \subset A-B$ ... (2)

de (1) y (2) resulta: $A-B = A \cap B'$

3^{er} método: Construyendo una "tabla de verdad".- Dado que un elemento x sólo puede "estar" o "no estar" en un determinado conjunto; indicamos con V (Verdad) en la tabla si x "está" en el conjunto, y con una F(Falso) si x "no está" en el conjunto.

Las combinaciones posibles para conjuntos A y B son las siguientes:

V ↗ V → VV
V ↘ F → VF

F ↗ V → FV
F ↘ F → FF

Se demuestra que la igualdad entre los dos miembros de una ecuación entre conjuntos queda establecida si las columnas correspondientes a cada miembro de la ecuación que en la tabla de verdad tie son las mismas entradas (V ó F) en los renglones correspondientes.

$x \in A$	$x \in B$	$x \in B'$	$x \in A-B$	$x \in A \cap B'$
V	V	F	F	F
V	F	V	V	V
F	V	F	F	F
F	F	V	F	F

1ª Observación: Una tabla de verdad debe tener 2^n renglones; siendo n el número de conjuntos diferentes involucrados en la proposición que se desea demostrar (en el ejemplo anterior, los conjuntos son A y B; por lo tanto $n = 2$; $2^n = 4$ renglones).

2ª Observación.- Las ecuaciones que relacionan operaciones entre conjuntos, en general, no admiten demostración mediante el diagrama de Venn, ya que un diagrama de Venn no puede incluir todas las situaciones posibles, como lo hace, por ejemplo, una tabla de verdad.

Más adelante se presentan ejemplos que ilustran lo expuesto en la 2ª observación.

<u>Principio de "dualidad"</u>.- Si se intercambian las operaciones $\cap$ por $\cup$; cambiando asimismo los conjuntos U por ϕ en cualquier expresión entre conjuntos, la expresión que resulta se denomina "dual" de la expresión original.

Ejemplo: sea la ecuación:

$$(A \cap U) \cap (\phi \cup A') = \phi$$

la dual es:

$$(A \cup \phi) \cup (U \cap A') = U$$

"Si ciertos axiomas implican sus propios duales, entonces el dual de cualquier teorema que es consecuencia de los axiomas es también consecuencia de los axiomas".

Por lo anterior, dado cualquier teorema y su demostración, el dual del teorema puede ser demostrado de manera análoga, usando el dual de cada paso sucesivo de la demostración original.

De modo que, si un teorema es verdadero, también lo es el dual de ese mismo teorema

LEYES DEL ALGEBRA DE CONJUNTOS

	LEY	DUAL
1.- IDEMPOTENCIA	$A \cup A = A$	$A \cap A = A$
2.- CONMUTATIVIDAD	$A \cup B = B \cup A$	$A \cap B = B \cap A$
3.- ASOCIATIVIDAD	$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$
4.- DISTRIBUTIVIDAD	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
5.- IDENTIDAD	$A \cup \phi = A$ $A \cup U = U$	$A \cap U = A$ $A \cap \phi = \phi$
6.- COMPLEMENTO	$A \cup A' = U$ $U' = \phi$	$A \cap A' = \phi$ $\phi' = U$
7.- DE MORGAN	$A' \cup B' = (A \cap B)'$	$A' \cap B' = (A \cup B)'$
8.- INVOLUCION	$(A')' = A$	

Enseguida daremos un ejemplo en el que se ilustra cómo, mediante la aplicación de los mismos razonamientos, es factible deducir de una ley y, en forma paralela, la dual.

Supongamos que se desea, a partir de la ley de idempotencia, deducir la ley de identidad, así como su dual:

$A = A \cup A$	hipótesis	$A = A \cap A$
$= (A \cup A) \cap U$	de 5	$= (A \cap A) \cup \phi$
$= (A \cup A) \cap (A \cup A')$	de 6	$= (A \cap A) \cup (A \cap A')$
$= A \cup (A \cap A')$	de 4	$= A \cap (A \cup A')$
$= A \cup \phi$	de 6	$= A \cap U$
$A = A \cup \phi$	Identidad	$A = A \cap U$

Se comprueba que los pasos seguidos en la transformación de la columna de la izquierda son precisamente los duales de cada uno de los pasos seguidos en la deducción realizada en la columna derecha.

Obsérvese asimismo que no todas las leyes del álgebra listadas en la tabla de la página anterior son independientes.

Como ejercicio, demostraremos la ley

3.- Asociatividad:

1er. método:

$A\cup(B\cup C) = \{x \mid x \in A \text{ ó } x \in B\cup C\}$

$= \{x \mid x \in A \text{ ó } x \in B \text{ ó } x \in C\}$

$= \{x \mid (x \in A \text{ ó } x \in B) \text{ ó } x \in C\}$

$= (A\cup B)\cup C$

2o. método:

a) Sea $x \in A\cup(B\cup C)$

$\Rightarrow x \in A \text{ ó } x \in B\cup C \Rightarrow x \in B \text{ ó } x \in C$

o sea $x \in A$ ó $x \in B$ ó $x \in C$

Si $x \in A \Rightarrow x \in A\cup B \Rightarrow x \in (A\cup B)\cup C$

Si $x \in B \Rightarrow x \in A\cup B \Rightarrow x \in (A\cup B)\cup C$

Si $x \in C \Rightarrow x \in C\cup(A\cup B) \Rightarrow x \in (A\cup B)\cup C$ (de 2)

$\Rightarrow A\cup(B\cup C) \subseteq (A\cup B)\cup C$ ...(1)

b) Si $x \in (A\cup B)\cup C$

$x \in A\cup B$ ó $x \in C$

$\Rightarrow x \in A$ ó $x \in B$ ó $x \in C$

Si $x \in A \Rightarrow x \in A\cup(B\cup C)$

Si $x \in B \Rightarrow x \in (B\cup C) \Rightarrow x \in A\cup(B\cup C)$

Si $x \in C \Rightarrow x \in (B\cup C) \Rightarrow x \in A\cup(B\cup C)$

$\Rightarrow (A\cup B)\cup C \subseteq A\cup(B\cup C)$ ...(2)

de (1) y (2): $(A\cup B)\cup C = A\cup(B\cup C) = A\cup B\cup C$

3er método

$x \in A$	$x \in B$	$x \in C$	$x \in A\cup B$	$x \in (A\cup B)\cup C$	$x \in B\cup C$	$x \in A\cup(B\cup C)$
V	V	V	V	V	V	V
V	V	F	V	V	V	V
V	F	V	V	V	V	V
V	F	F	V	V	F	V
F	V	V	V	V	V	V
F	V	F	V	V	V	V
F	F	V	F	V	V	V
F	F	F	F	F	F	F

F V V F F V F

Ejercicios.- Con base en las leyes del álgebra de conjuntos, demostrar:

1.- $A'\cap B'\cap C' = (A\cup B\cup C)'$

$A'\cap B'\cap C' = (A'\cap B')\cap C'$ (de 3)

$= [(A'\cap B')'\cup(C')']'$ (de 7)

$= [(A\cup B)\cup C]'$ (de 7 y 8)

$= (A\cup B\cup C)'$ (de 3)

2.- $A\cup(A'\cap B) = A\cup B$

$A\cup(A'\cap B) = (A\cup A')\cap(A\cup B)$ (de 4)

$= U\cap(A\cup B)$ (de 6)

$= A\cup B$ (de 5)

3.- $[(A'\cup B)\cap(C\cup A')]' = A\cap(B'\cup C')$

$[(A'\cup B)\cap(C\cup A')]' = [(A'\cup B)\cap(A'\cup C)]'$ (de 2)

$= [A'\cup(B\cap C)]'$ (de 4)

$= (A')'\cap(B\cap C)'$ (de 7)

$= A\cap(B\cap C)'$ (de 8)

$= A\cap(B'\cup C')$ (de 7)

4.- $(A\cup B)\cup(A'\cup B') = U$

$(A\cup B)\cup(A'\cup B') = (A\cup A')\cup(B\cup B')$ (de 2 y 3)

$= U\cup U$ (de 6)

$= U$ (de 1)

5.- Con base en la fórmula dada anteriormente: $nA+nB = n(A\cup B)+n(A\cap B)$, deducir la fórmula para relacionar el número de elementos en tres conjuntos, $nA+nB+nC$.

Solución:

Por asociatividad: $n(A\cup B\cup C) = n(A\cup(B\cup C))$

por la fórmula dada para 2 conjuntos:

$n(A\cup(B\cup C)) = nA+n(B\cup C)-n[A\cap(B\cup C)]$

aplicando nuevamente la misma fórmula:

$n(A\cup(B\cup C)) = nA+nB+nC-n(B\cap C)-n(A\cap(B\cup C))$

de la distributividad:

$n(A\cap(B\cup C)) = n[(A\cap B)\cup(A\cap C)]$

aplicando la fórmula para 2 conjuntos a este último:

$$n(A \cup B \cup C) = nA+nB+nC-n(B\cap C)-n(A\cap B)-$$

$$+n(A \cup C)+n\left[(A\cap B)\cap(A\cap C)\right]$$

$$= nA+nB+nC-n(A\cap B)-n(A\cap C)-n(B\cap C)+n(A\cap B\cap C)$$

o bien:

$$nA+nB+nC = n(A\cup B\cup C)+n(A\cap B)+n(A\cap C)+n(B\cap C)-n(A\cap B\cap C)$$

6.- En una encuesta efectuada a un grupo de 100 estudiantes universitarios acerca de qué idioma extranjero estaban estudiando en ese semestre, se obtuvieron las siguientes respuestas:

Idioma:	Número de estudiantes:
Francés (F)	26
Inglés (I)	48
Inglés (I) y Ruso (R)	8
Francés, pero no Ruso	23
Solamente Francés	18
Francés e Inglés	8
Ningún idioma	24

Se pregunta:

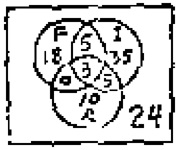
i) ¿Cuántos cursan Ruso? $nR = ?$

ii) ¿Cuántos toman Francés y Ruso, pero no llevan Inglés?

iii) ¿Cuántos cursan Francés y Ruso o Inglés, o ambos?

Solución:

Del diagrama de Venn:



i) $nR = 18$

ii) $n(F\cap R\cap I') = 0$

iii) $n(F\cap(R\cup I)) = 8$

7.- Se tienen 3 objetos: A, B, C, y se sabe que ninguno de los 3 pesa más de 1 kg ni menos de 0.5 kg. Para determinar con exactitud el peso que tiene cada uno de los objetos, se emplea una balanza; pero ésta únicamente registra, con precisión, pesos entre 1 y 2 kg.

El problema consiste en determinar, con la balanza dada, con absoluta precisión, el peso de cada uno de los 3 objetos, efectuando únicamente 3 pesadas en la balanza.

Solución: Para poder emplear la balanza, deberán pesarse 2 objetos conjuntamente.

Nomenclatura: $\bar{U} = \{A, B, C\}$

nA = peso de A

nB = peso de B

nC = peso de C

$n\bar{U}$ = peso conjunto de los 3 objetos.

$S_1 = \{B,C\}$; $S_2 = \{A,C\}$; $S_3 = \{A,B\}$.

nS_1 = peso de B y C juntos = $nB+nC$

nS_2 = peso de A y C juntos = $nA+nC$

nS_3 = peso de A y B juntos = $nA+nB$

Resulta entonces:

$S_1 \cup S_2 = S_2 \cup S_3 = S_3 \cup S_1 = U$

$S_1 \cap S_2 = C$

$S_2 \cap S_3 = A$

$S_3 \cap S_1 = B$

Hemos visto que:

$$nS_1+nS_2 = n(S_1 \cup S_2)+n(S_1 \cap S_2)$$

$$= nU+nC$$

$$nS_1+nS_2 = nU+nC \quad \ldots (1)$$

$$nS_1+nS_3 = nU+nB \quad \ldots (2)$$

$$\underline{nS_2+nS_3 = nU+nA} \quad \ldots (3)$$

$$2(nS_1+nS_2+nS_3) = 3nU+nA+nB+nC = 4nU$$

$$\therefore nU = \tfrac{1}{2}(nS_1+nS_2+nS_3) \quad (4)$$

(4) en (3):

$$nS_2+nS_3 = \tfrac{1}{2}(nS_1+nS_2+nS_3)+nA$$

Simplificando, y despejando nA se obtiene la primera de las siguientes 3 ecuaciones; las 2 siguientes resultan de manera análoga:

$$nA = \tfrac{1}{2}(nS_2+nS_3-nS_1) \qquad \text{(de (4) y (3))}$$

$$nB = \tfrac{1}{2}(nS_1+nS_3-nS_2) \qquad \text{(de (4) y (2))}$$

$$nC = \tfrac{1}{2}(nS_1+nS_2-nS_3) \qquad \text{(de (4) y (1))}$$

8.- Al director de una escuela secundaria mixta (coeducacional) se le presentan los siguientes datos estadísticos tendientes a reflejar el efecto de la práctica de los deportes en el aprove - chamiento académico de los alumnos:

La muestra consistió de 100 alumnos, 50 hombres y 50 mujeres. Pasó año el 60 %, del cual 28 son mujeres y 32, hombres.
De los 100 que constituyen la muestra, 56 practican deporte y, de estos 56, hay 36 hombres y 20 mujeres. Se observa que de los 60 aprobados, 34 son deportistas, y de estos 34, hay 30 hombres.

El director pregunta: ¿Cuántas mujeres no-deportistas no pasaron año?

1ª Solución:

$nU = 100$, $nH = nM = 50$; $nA = 60$;

$n(A \cap M) = 28$; $n(A \cap H) = 32$; $nD = 56$

$n(D \cap H) = 36$; $n(D \cap M) = 20$; $n(A \cap D) = 34$;

$n(A \cap D \cap H) = 30$; INCOGNITA: $N(M \cap D' \cap A')$

pero: $n(M \cap D' \cap A') = n(H' \cap D' \cap A')$

del problema 1.- anterior:

$$n(M' \cap D' \cap A') = n(M \cup D \cup A)'$$

pero $n(M \cup D \cup A)' = n\bar{U} - n(M \cup D \cup A)$

del problema 5.- anterior:

$$n(M \cup A \cup D) = nM + nA + nD - n(M \cap A) - n(M \cap D) - n(A \cap D) + n(M \cap A \cap D)$$

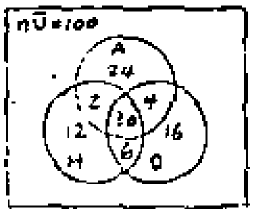
$$= 50+60+56-32-36-34+30$$

$$= 94$$

$$n(M \cap D' \cap A') = n\bar{U} - n(M \cup D \cup A) = 100-94 = 6$$

o sea: El número de mujeres, que no practican deporte y que reprobaron el año es 6.

El diagrama de Venn correspondiente a este problema es como sigue:

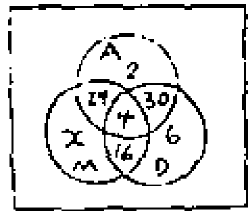


$n(M \cup A \cup D) = 94$

2ª Solución:

$$n(M \cup A \cup D) = nM + nA + nD - n(M \cap A) - n(M \cup D) - n(A \cap D) + n(M \cap A \cap D)$$

Donde se conocen todos los datos del segundo miembro. Para completar los datos, puede recurrirse al siguiente diagrama de Venn:



$\therefore n(M \cup A \cup D) = x+82$

$$\therefore \quad x+82 = 50+60+56-28-20-34+4$$

$$x+82 = 170-82 = 88$$

$$x = 6$$

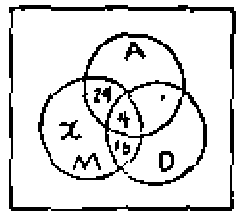
3ª Solución:

Del diagrama de Venn se observa que:

$$nM = x + n(A \cap M) + n(M \cap D) - n(M \cap A \cap D)$$

$$50 = x+28+20-4 = x+44$$

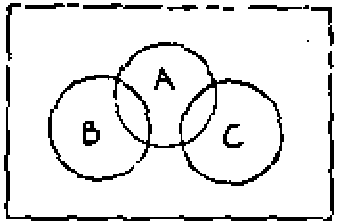
$$\therefore x = 6$$



Dibujar diagramas de Venn, tales que verifiquen las siguientes ecuaciones:

1.- $A \cup B \cap C = (B-C) \cap (A \cap C)$

Solución:



Del diagrama:

$A \cap B \cap C = \phi$

$B - C = B$

$\Rightarrow (B-C) \cap (A \cap C) = B \cap A \cap C$

$\Rightarrow (B-C) \cap (A \cap C) = A \cap B \cap C$

2.- $A \cap B \cap C = A \cap A'$

Solución:

Dado que $A \cap A' = \phi$; (para cualquier A) $\Rightarrow$ que para que se veri- fique la ecuación en 2, basta que $A \cap B \cap C = \phi$; como sucede, por ejemplo, en el diagrama de Venn del problema 1.- anterior (existen muchas más posibilidades)

3.- $A-(B-C) = (A-B) \cup C$ ...(1)

Transformando el 1^{er} miembro de (1):

$A-(B-C) = A \cap (B-C)'$; (Demostrado por 3 métodos)

$= A \cap (B \cap C')'$; (por la misma razón de antes)

$= A \cap (B' \cup C)$; (Por De Morgan)

$= (A \cap B') \cup (A \cap C)$; (Por distributividad)

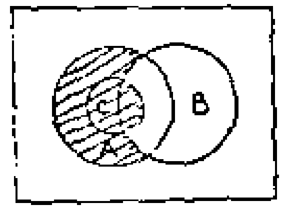
$= (A-B) \cup (A \cap C)$ ...(2)

Comparando los 2^{os} miembros de (1) y (2), siendo iguales los 1^{os} miembros:

Si $A \cap C = C \Rightarrow (1) = (2)$

pero $A \cap C = C \Rightarrow C \subseteq A$.

De ahí que el diagrama de Venn apropiado puede ser el siguiente:



Del diagrama:

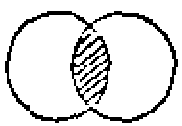
$A-(B-C) = (A-B) \cup C$

Es muy importante observar que las 3 fórmulas dadas no tienen validez general, como puede demostrarse, por ejemplo, construyendo las tablas de verdad correspondientes a cada caso.

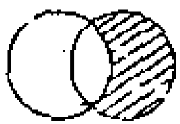
Problema: colocar los paréntesis necesarios en el segundo miembro de modo que resulte cierta la siguiente igualdad para todo conjunto A, B:

$$A-[B-(A\cap B)] = A-B-B-A$$

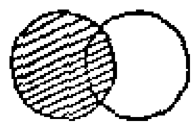
Solución: Dado que el conjunto del primer miembro sí está bien definido, se podrá tener idea de qué conjunto representa si nos auxiliamos de diagramas de Venn:



$A\cap B$



$B-(A\cap B)$



$A-[B-(A\cap B)]$

Por lo que parece ser que $A-[B-(A\cap B)] = A$.

Lo que puede comprobarse (demostrarse) con una tabla de verdad (para todo A, B)

$x\in$ A	B	$A\cap B$	$B-(A\cap B)$	$A-[B-(A\cap B)]$
V	V	V	F	V
V	F	F	F	V
F	V	F	V	F
F	F	F	F	F

$\rightarrow A=A-[B-(A\cap B)]$

Volviendo ahora al segundo miembro: y observando que $B-B=\emptyset$

$A-[(B-B)-A] = A-[\emptyset-A] = A-\emptyset = A.$

Por lo cual, la igualdad propuesta se verifica para todo conjunto A, B, si:

$$A-[B-(A\cap B)] = A-[(B-B)-A]$$

Obsérvese que no funciona ninguna de las siguientes posibilidades de colocación de paréntesis en el segundo miembro:

$$(A-B)-(B-A) = A-B \neq A$$

$$[A-(B-B)]-A = \emptyset \neq A$$

$$[(A-B)-B]-A = (A-B)-A = \emptyset \neq A$$

$$A-[B-(B-A)] = A-B \neq A$$

Teorema: La condición necesaria y suficiente para que un conjunto B sea el complemento de otro conjunto A, es que se verifiquen simultáneamente las dos condiciones siguientes:

$$a) \quad A \cap B = \phi$$

$$y \quad b) \quad A \cup B = \bar{U}; \quad \text{(b) es la dual de a))}$$

Demostración: condición necesaria:

Hipótesis: $B = A'$

$$A \cap B = A \cap A' = \phi$$

asimismo: $B = A' \rightarrow A \cup B = A \cup A' = \bar{U}$

condición suficiente

Hipótesis: $A \cap B = \phi$, y $A \cup B = \bar{U}$

$$B = B \cap \bar{U}$$

$$= B \cap (A \cup A')$$

$$= (B \cap A) \cup (B \cap A')$$

$$= (A \cap B) \cup (B \cap A')$$

de la hipótesis: $A \cap B = \phi$

$$\rightarrow B = \phi \cup (B \cap A')$$

$$= (A' \cap A) \cup (A' \cap B)$$

$$= A' \cap (A \cup B)$$

de la hipótesis: $A \cup B = \bar{U}$

$$B = A' \cap \bar{U}$$

$$\rightarrow B = A'$$

Problema.- Con base en el teorema anterior, demostrar la ley de De Morgan:

$$(A \cup B)' = A' \cap B'$$

Bastará hacer ver que:

a) $(A \cup B) \cup (A' \cap B') = \bar{U}$

ó b) $(A \cup B) \cap (A' \cap B') = \phi$

puesto que si hacemos $A \cup B = C$; $A' \cap B' = D$

$$C \cup D = \bar{U} \rightarrow C' = D$$

$$ó \quad C \cap D = \phi \rightarrow C' = D$$

Demostración de a)

$$(A \cup B) \cup (A' \cap B') = \left[(A \cup B) \cup A' \right] \cap \left[(A \cup B) \cup B' \right]$$

$$= \left[A' \cup (A \cup B) \right] \cap \left[A \cup (B \cup B') \right]$$

$$= \left[(A' \cup A) \cup B \right] \cap \left[A \cup (B \cup B') \right]$$

$$= \left[\bar{U} \cup B \right] \cap \left[A \cup \bar{U} \right]$$

$$= \bar{U} \cap \bar{U} = \bar{U}$$

$$\rightarrow (A \cup B)' = A' \cap B'$$

Aplicación de la teoría de conjuntos a razonamientos lógicos.-

Supongamos que se establece la siguiente relación entre varios conjuntos

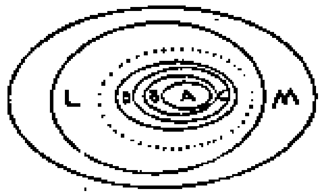
$A \subset B$

$B \subset C$

$C \subset D$

. . .

$L \subset M$



La conclusión es: $A \subset M$

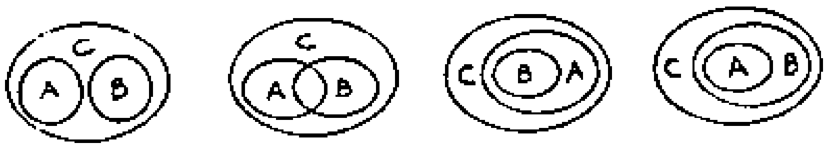
Debe observarse que la conclusión anterior es válida únicamente si cada conjunto aparece no más de una vez del lado derecho o del lado izquierdo.

No puede concluirse nada concreto acerca de los conjuntos A y B en el siguiente ejemplo:

$A \subset C$

$B \subset C$

En todos los casos siguientes se cumplen las condiciones dadas:

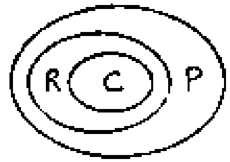


Las leyes de los conjuntos y el uso de diagramas de Venn pueden conducir a la determinación de la validez de ciertos tipos de razonamientos, como se ilustra en los ejemplos dados a continuación:

1.- Determinar si es válido el siguiente razonamiento:

Premisas: { Todos los cuadrados son rectángulos.
Todos los rectángulos son paralelogramos.

Conclusión: { Por lo tanto, todos los cuadrados son paralelogramos.



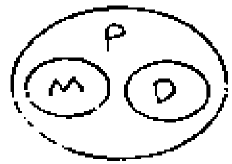
$C \subset R \subset P$

$\therefore$ el razonamiento es válido

2.- ¿Es válido el siguiente razonamiento?

Premisas: { Algunos problemas se resuelven matemáticamente
Algunos problemas son difíciles

Conclusión: { Por lo tanto, algunos problemas matemáticos son difíciles.



Aun cuando la conclusión es correcta, ésta se obtiene de un razonamiento falso (ó inválido)

Empleando las leyes de conjuntos es posible demostrar que los conjuntos M y D del problema anterior no necesariamente tienen intersección: Del diagrama:

$$P \cup M = P, \quad P \cup D = P$$

$$(P \cup M) \cap (P \cup D) = P \cap P = P$$

$$\therefore \quad P \cup (M \cap D) = P \ldots (1)$$

Sabemos además que $P \cup \phi = P \ldots (2)$

Comparando (1) y (2) vemos que

$M \cap D$ puede ser ϕ

$\therefore$ M y D no necesariamente tienen intersección.

3.- Analizar el siguiente razonamiento

Premisas: Todos los niños son felices / Gentes felices no asesinan

Conclusión: Por lo tanto, ningún niño es asesino

Del diagrama podemos concluir que el razonamiento es válido:

Analíticamente: $N \cap F = N$; $F \cap A = \phi$

$N = N \cap F$

$N \cap A = (N \cap F) \cap A$

$= N \cap (F \cap A)$

$N \cap A = N \phi$

$N \cap A = \phi$

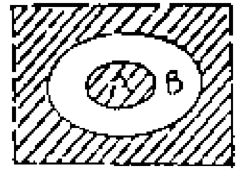
∴ No existen gentes que sean simultáneamente niños y asesinos

∴ ningún niño es asesino.

Expresión de A B en función de las operaciones unión e intersección

Dado que es válida la conclusión A C si se sabe que A B y B C, conviene expresar el símbolo en función de la unión e intersección de dos conjuntos.

Supongamos que A B; lo cual se indica mediante el siguiente diagrama de Venn:



Puede afirmarse que, si $A \subset B$, las áreas rayadas en el diagrama no se superponen. Esta condición es equivalente a:

$$A \cap B' = \phi$$

Puesto que la recíproca también es cierta, se concluye que:

$$A \subset B \leftrightarrow A \cap B' = \phi \quad ...(1)$$

así mismo, de (1)

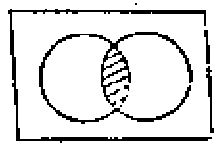
$$B' \subset A' \leftrightarrow B' \cap (A')' = \phi$$

$$B' \subset A' \leftrightarrow A \cap B' = \phi \quad ...(2)$$

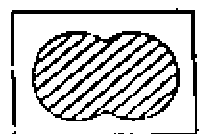
de (1) y (2)

$$A \subset B \leftrightarrow B' \subset A' \quad ...(3)$$

Por otro lado:



$A \cap B \neq A \rightarrow A \not\subset B$

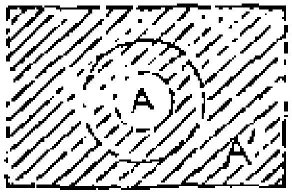


$A \cup B \neq B \rightarrow A \not\subset B$

$$A \subseteq B \leftrightarrow A \cap B = A \qquad \ldots (4)$$

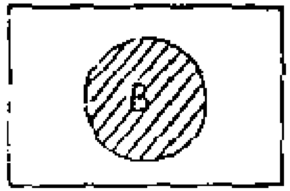
$$A \subseteq B \leftrightarrow A \cup B = B \qquad \ldots (5)$$

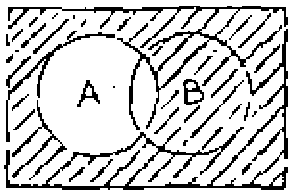
También: $A \subseteq B \leftrightarrow A' \cup B = U$ ...(6) (dual de (1))



$A' \cup B = U$

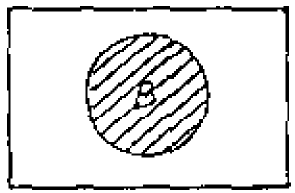
$\rightarrow A \subset B$





$A' \cup B \neq U$

$\rightarrow A \not\subset B$



Demostraciones analíticas:

1.- Se afirma que $A \subseteq B \leftrightarrow A \cap B = A$

Demostración:

1ª parte: hipótesis: $A \subseteq B$

$\rightarrow$ si $x \in A \rightarrow x \in B$; por lo cual, si $x \in A$ y $x \in B$ $\therefore$
$x \in A \cap B \rightarrow A \subset A \cap B$ ...(i)

por otro lado, si $x \in A \cap B \rightarrow x \in A$ y $x \in B$

$\rightarrow x \in A \rightarrow A \cap B \subseteq A$ ...(ii)

de (i) y (ii) resulta: $A \cap B = A$

2ª parte: hipótesis: $A \cap B = A$

Demostración: si $A \cap B = A \rightarrow A \subset A \cap B$ ...(α)

pero, para todo $A, B \in U$: $A \cap B \subset B$ ...(β)

De (α) y (β) se concluye que $A \subseteq B$

2.- Demostrar que $A \cap B = A \rightarrow A \cap B' = \phi$

En efecto: $A = A \cap B \rightarrow A \cap B' = (A \cap B) \cap B' =$

$= A \cap (B' \cap B) = A \cap \phi = \phi$

$\therefore$ $A \cap B = A$ equivale a $A \cap B' = \phi$

3.- Demostrar que $A \cap B' = \phi \rightarrow A' \cup B = U$.

Aplicando la ley de De Morgan a $A \cap B' = \phi$ se obtiene: $A' \cup B = U$.

4.- Demostrar que $\phi = A \cap B' \rightarrow A \cup B = B$.

$A \cup B = (A \cap B') \cup B = (A \cup B) \cap (B' \cup B) = (A \cup B) \cap U = A \cup B$

$\therefore$ $\phi = A \cap B'$ equivale a $A \cup B = B$

En resumen:

$$A \subset B \begin{cases} B' \subset A' \\ A \cap B' = \phi \\ A \cap B = A \\ A \cup B = B \\ A' \cup B = U \end{cases}$$

Problema:

¿Qué conclusiones pueden derivarse de las siguientes proposiciones?

a) Todos los estudiantes inscritos en deportes juegan o beisbol o futbol (o ambos).

b) No es permitido jugar beisbol y ser además miembro del equipo de natación.

c) Los que no forman parte del equipo de natación, deben practicar atletismo.

Solución: Definamos los siguientes conjuntos:

$\mathcal{U}$ = Todos los estudiantes inscritos en deportes

B = Conjunto que juega beisbol.

F = Conjunto que juega futbol.

N = Conjunto en el equipo de natación.

A = Conjunto en el equipo de atletismo.

Simbólicamente, las proposiciones dadas se escriben:

a) $B \cup F = \mathcal{U}$

b) $B \cap N = \phi$

c) $N' \subset A$

Hemos visto que a) y b) son equivalentes a:

a) $B \cup F = \mathcal{U} \equiv B' \subset F$ ó $F' \subset B$

b) $B \cap N = \phi \equiv B \subset N'$ ó $N \subset B'$

de lo cual tomamos:

$$\left.\begin{array}{l} a)\ F' \subset B \\ b)\ B \subset N' \end{array}\right\} \rightarrow F' \subset N' \text{ ó } N \subset F$$

$$\text{de}\quad \left.\begin{array}{l} b)\ B \subset N' \\ c)\ N' \subset A \end{array}\right\} \rightarrow B \subset A$$

puesto que

$$\left.\begin{array}{l} F' \subset B \\ B \subset A \end{array}\right\} \rightarrow F' \subset A$$

En resumen, las conclusiones son:

1.- $F' \subset N' \equiv N \subset F$

Todos los miembros del equipo de natación juegan futbol.

2.- $B \subset A$

Todos los que juegan beisbol están en el equipo de atletismo.

3.- $F' \subset A$

Todos los que no juegan futbol están en el equipo de atletismo.

Teorema.- Dos conjuntos A y B son iguales si, y sólo si, se verifica una, cualquiera, de las dos condiciones que se expresan a continuación:

a) $(A-B) \cup (B-A) = \phi$

ó b) $(A \cap B') \cup (A' \cap B) = \phi$

Demostración de b)

Hipótesis: $(A \cap B') \cup (A' \cap B) = \phi$

Para que la unión de dos conjuntos pueda ser igual al conjunto ϕ es necesario que cada conjunto sea ϕ, puesto que $\phi \cup \phi = \phi$

$$\left.\begin{array}{l} A \cap B' = \phi \to A \subset B \\ A' \cap B = \phi \to B \subset A \end{array}\right\} A = B$$

Recíprocamente:

Hipótesis $A = B$

$$(A \cap B') \cup (A' \cap B) = (A \cap A') \cup (A' \cap A) =$$

$$= \phi \cup \phi = \phi$$

Problema: Con base en el último teorema, demostrar que en el problema anterior, es factible

Solución: Debemos buscar bajo [illegible] se verifica la ecuación:

$$(N \cap F') \cup (N' \cap F) = \phi$$

De la 1ª conclusión obtenida:

$$N \subset F \to N \cap F' = \phi$$

$$\therefore \quad \phi \cup (N' \cap F) = \phi$$

$$\to N' \cap F = \phi$$

lo anterior se verifica si $F \subset N \to F$ puede ser igual a N

B = A	F = N

Del diagrama se comprueba que se cumplen todas las condiciones dadas y las conclusiones obtenidas, esto es:

$$\left.\begin{array}{l} B \cup F = U \\ B \cap N = \phi \\ N' \subset A \end{array}\right\} \text{DATOS}$$

$$\left.\begin{array}{l} N \subset F \\ B \subset A \\ F' \subset A \end{array}\right\} \text{CONCLUSIONES}$$

<u>Pares ordenados</u>.- Sean dos elementos: a y b; si designamos como primer elemento a uno de ellos, digamos a, y al elemento b como segundo, tendremos definido un par ordenado, el que se representa como: (a, b).

<u>Igualdad</u>.- Dos pares ordenados son iguales si, y sólo si, sus primeros elementos son iguales; siendo asimismo iguales sus segundos elementos; esto es: $(a, b) = (c, d) \leftrightarrow a = c,\ b = d$

Observaciones:

1.- $(2, 3) \neq (1, 2)$

2.- Los pares ordenados pueden estar constituidos con el primer elemento igual al segundo: $(a, a) = (5, 5)$

$(b, b) \neq (1, 2)$

Los puntos del plano coordenado son ejemplos conocidos de pares ordenados.

Se da como definición de par ordenado la siguiente:

$(a,b) = \{\{a\},\{a,b\}\}$; en donde $\{a,b\}$ es un conjunto (par no ordenado), y el conjunto $\{a\}$ determina cuál de los elementos del conjunto $\{a,b\}$ debe considerarse como primer elemento del par.

Con base en la definición anterior, se demostrarán algunas de las afirmaciones hechas anteriormente:

I.- $(a,b) = (c,d) \longleftrightarrow a = c;\ b = d$

II.- $(a,b) \neq (b,a)$, si $a \neq b$

III.- Si un par ordenado tiene sus dos coordenadas iguales, resulta $(a,a) = \{\{a\}\}$

Demostraciones:

I.- a) Hipótesis: $(a,b) = (c,d)$

Se demostrará que $a = c;\ b = d$.

En efecto: $(a,b) = \{\{a\},\{a,b\}\}$; asimismo: $(c,d) = \{\{c\},\{c,d\}\}$

por hipótesis: $\{\{a\},\{a,b\}\} = \{\{c\},\{c,d\}\}$ ...(1)

i) si $a = b \Rightarrow \{\{a\},\{a,b\}\} = \{\{a\},\{a,a\}\} = \{\{a\},\{a\}\} = \{\{a\}\}$

(con lo cual queda demostrada la afirmación III).de (1)

$\therefore \{\{c\},\{c,d\}\} = \{\{a\}\}, \Rightarrow \{\{c\}\} = \{\{a\}\}; \Rightarrow \{c\} = \{a\}, \Rightarrow a = c.$

puesto que si $a = b \Rightarrow b = a$, se obtiene asimismo $b = d$

ii) si $a \neq b$, resulta que, por tener ambos miembros de la igualdad (1) el mismo número de elementos, $\Rightarrow c \neq d$; obteniéndose de (1) que:

$\{a\} \in \{\{c\},\{c,d\}\}$

puesto que: $\{a\} \neq \{c,d\} \Rightarrow \{a\} = \{c\}$

$\therefore a = c$

Por otro lado, de (1):

$\{a,b\} \in \{\{c\},\{c,d\}\}$

pero $\{a,b\} \neq \{c\}$

$\Rightarrow \{a,b\} = \{c,d\}$

puesto que resultó $a = c, \Rightarrow b = d$; ya que $a \neq b$; con lo cual se completa la demostración de la parte I a)

I.-b) Hipótesis: $a = c;\ b = d$

Se demostrará que $(a,b) = (c,d)$

En efecto, de la hipótesis resultan:

$\{a\} = \{c\}; \{a,b\} = \{c,d\}$

$\therefore \{\{a\},\{a,b\}\} = \{\{c\},\{c,d\}\}$

$\rightarrow (a,b) = (c,d)$

II.- Hipótesis $a \neq b$; conclusión: $(a,b) \neq (b,a)$.

Por definición:

$(a,b) = \{\{a\},\{a,b\}\}$; $(b,a) = \{\{b\},\{b,a\}\}$

si: $a \neq b$, $\rightarrow \{a\} \neq \{b\}$, $\rightarrow \{\{a\},\{a,b\}\} \neq \{\{b\},\{b,a\}\}$

$\rightarrow (a,b) \neq (b,a)$

Obsérvese que lo anterior se verifica aun cuando: $\{a,b\} = \{b,a\}$, para toda a, b.

<u>Producto cartesiano de conjuntos</u>.- Dados dos conjuntos A y B, el "producto cartesiano" de A y B es el conjunto cuyos elementos son todos los pares ordenados (a, b) tales que $a \in A$ y $b \in B$:

$$A \times B = \{(a,b) | a \in A,\ b \in B\}$$

Ejemplo: Sea $A = \{a,b\}$; $B = \{x, y, z\}$

$A \times B = \{(a, x), (a, y), (a, z), (b, x), (b, y), (b, z)\}$

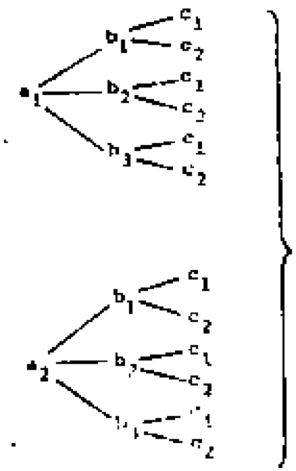
Obsérvese que: $n(A \times B) = nA \times nB$; en el ejemplo: $nA = 2$; $nB = 3 \rightarrow nA \times nB = 6 \rightarrow n(A \times B) = 6$

La definición anterior puede generalizarse para formar productos

cartesianos de cualquier número de conjuntos:

$A \times B \times \dots \times N = \{(a,b,\dots,n) | a \in A, b \in B, \dots, n \in N\}$

Ejemplo: Sean: $A = \{a_1,a_2\}$; $B = \{b_1,b_2,b_3\}$; $C = \{c_1,c_2\}$



resultan 12 ternas ordenadas

$nA \times nB \times nC = n(A \times B \times C) = 12$

$A \times B \times C = \{(a_1b_1c_1)(a_1b_1c_2)(a_1b_2c_1)(a_1b_2c_2)\dots\dots(a_2b_3c_1)(a_2b_3c_2)\}$

Ejercicios:

1.- Demostrar que: $A \times (B \cap C) = (A \times B) \cap (A \times C)$

a) $A \times (B \cap C) = \{(x,y) | x \in A,\ y \in B \cap C\}$

$= \{(x,y) | x \in A,\ y \in B \text{ y } y \in C\}$

$= \{(x,y) | (x,y) \in A \times B \text{ y } (x,y) \in A \times C\}$

$= (A \times B) \cap (A \times C)$

$\{a\} = \{c\}$; $\{a,b\} = \{c,d\}$

$\therefore \{\{a\},\{a,b\}\} = \{\{c\},\{c,d\}\}$

$\Rightarrow (a,b) = (c,d)$

II.- Hipótesis $a \neq b$; conclusión: $(a,b) \neq (b,a)$.

Por definición:

$(a,b) = \{\{a\},\{a,b\}\}$; $(b,a) = \{\{b\},\{b,a\}\}$

si: $a \neq b$, $\Rightarrow \{a\} \neq \{b\}$, $\Rightarrow \{\{a\},\{a,b\}\} \neq \{\{b\},\{b,a\}\}$

$\Rightarrow (a,b) \neq (b,a)$

Obsérvese que lo anterior se verifica aun cuando: $\{a,b\} = \{b,a\}$, para toda a, b.

Producto cartesiano de conjuntos.- Dados dos conjuntos A y B, el "producto cartesiano" de A y B es el conjunto cuyos elementos son todos los pares ordenados (a, b) tales que $a \in A$ y $b \in B$:

$$A \times B = \{(a,b) \mid a \in A,\ b \in B\}$$

Ejemplo: Sea $A = \{a,b\}$; $B = \{x, y, z\}$

$A \times B = \{(a, x), (a, y), (a, z), (b, x), (b, y), (b, z)\}$

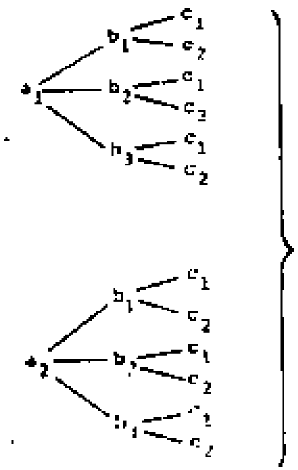
Obsérvese que: $n(A \times B) = nA \times nB$; en el ejemplo: $nA = 2$; $nB = 3 \Rightarrow nA \times nB = 6 \Rightarrow n(A \times B) = 6$

La definición anterior puede generalizarse para formar productos

cartesianos de cualquier número de conjuntos:

$A \times B \times \ldots \times N = \{(a,b,\ldots,n) \mid a \in A, b \in B, \ldots, n \in N\}$

Ejemplo: Sean: $A = \{a_1,a_2\}$; $B = \{b_1,b_2,b_3\}$; $C = \{c_1,c_2\}$



resultan 12 ternas ordenadas

$nA \times nB \times nC = n(A \times B \times C) = 12$

$A \times B \times C = \{(a_1b_1c_1)(a_1b_1c_2)(a_1b_2c_1)(a_1b_2c_2)\ldots\ldots(a_2b_3c_1)(a_2b_3c_2)\}$

Ejercicios:

1.- Demostrar que: $A \times (B \cap C) = (A \times B) \cap (A \times C)$

a) $A \times (B \cap C) = \{(x,y) \mid x \in A,\ y \in B \cap C\}$

$= \{(x,y) \mid x \in A,\ y \in B \text{ y } y \in C\}$

$= \{(x,y) \mid (x,y) \in A \times B \text{ y } (x,y) \in A \times C\}$

$= (A \times B) \cap (A \times C)$

otro método:

b) Sea $(x,y) \in A \times (B \cap C)$

$\rightarrow x \in A,\ y \in B \cap C$

$\rightarrow x \in A,\ y \in B$ y $y \in C$

$\rightarrow (x,y) \in (A \times B)$ y $(x,y) \in (A \times C)$

$\rightarrow (x,y) \in (A \times B) \cap (A \times C)$

$\rightarrow A \times (B \cap C) \subset (A \times B) \cap (A \times C) \ldots (1)$

Sea $(x,y) \in (A \times B) \cap (A \times C)$

$\rightarrow (x,y) \in A \times B$ y $(x,y) \in A \times C$

$\rightarrow x \in A$ y $y \in B$ y C

$\rightarrow x \in A$ y $y \in B \cap C$

$\rightarrow (x,y) \in A \times (B \cap C)$

$\rightarrow (A \times B) \cap (A \times C) \subset A \times (B \cap C) \ldots (2)$

de (1) y (2):

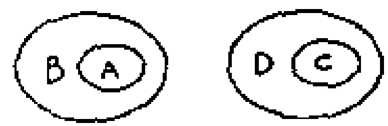
$A \times (B \cap C) = (A \times B) \cap (A \times C)$

c) Tabla de verdad:

$x \in A$	$y \in B$	$y \in C$	$y \in B \cap C$	$(x,y) \in A \times (B \cap C)$	$(x,y) \in A \times B$	$(x,y) \in A \times C$	$(x,y) \in (A \times B) \cap (A \times C)$
V	V	V	V	V	V	V	V
V	V	F	F	F	V	F	F
V	F	V	F	F	F	V	F
V	F	F	F	F	F	F	F
F	V	V	V	F	F	F	F
F	V	F	F	F	F	F	F
F	F	V	F	F	F	F	F
F	F	F	F	F	F	F	F

7.- Demostrar que si $A \subset B$ y $C \subset D \rightarrow (A \times C) \subset (B \times D)$

a) Sea $(x,y) \in A \times C$



$\rightarrow x \in A$ y $y \in C$

pero, por hipótesis: $A \subset B$ y $C \subset D$

$\rightarrow$ si $x \in B$ y $y \in D \rightarrow (x,y) \in B \times D$

$\rightarrow$ si $(x,y) \in (A \times C) \rightarrow (x,y) \in (B \times D)$

$\rightarrow (A \times C) \subset (B \times D)$

otro método:

b) $(A \times C) \subset (B \times D)$

$\therefore \quad (A \times C) \cup (B \times D) = B \times D$

$\therefore \quad (A \times C) \cap (B \times D) = A \times C$

pag 48

1	2	3	4	5	6	7	8	9	10	11
				de 1 y 3						de 5 y 10
x∈A	x∈B	y∈C	y∈D	(x,y)∈	(x,y)∈B×D				$\bigvee_{i=6}^{9} C_i$	(x,y)∈
x∉A	x∉B	y∉D	y∉C	A × C	1 y 3	1 y 4	2 y 3	2 y 4		(A×C)∪(B×D)
V	V	V	V	V	V				V	V
V	V	V	F	V	V				V	V
V	V	F	V	F	F	V			V	V
V	V	F	F	F	F	F	F	F	F	F
V	F	V	V	V	V				V	V
V	F	V	F	V	V				V	V
V	F	F	V	F	F	V			V	V
V	F	F	F	F	F	F	F	F	F	F
F	V	V	V	F	F	F	V		V	V
F	V	V	F	F	F	F	V		V	V
F	V	F	V	F	F	F	F	V	V	V
F	V	F	F	F	F	F	F	F	F	F
F	F	V	V	F	F	F	F	F	F	F
F	F	V	F	F	F	F	F	F	F	F
F	F	F	V	F	F	F	F	F	F	F
F	F	F	F	F	F	F	F	F	F	F

La tabla nos demuestra que

$(A \times C) \cup (B \times D) = B \times D$

$\therefore (A \times C) \subset (B \times D)$

Algunos libros de referencia:

1.- "Sets, logic and axiomatic theories". R. R. Stoll, (Edit. Freeman) 1961.

2.- "Introduction to logic and sets". R. R. Christian. (Edit. Blaisdell) 1965.

3.- "Introduction to the foundation of mathematics". R.L. Wilder, (Edit. Wiley) 1952.

4.- "Set theory and related topics". S. Lipschutz, (Edit. Schaum) 1964.

5.- "Notas sobre conjuntos y números" J. Salazar R. (Edit. C.F.E.) 1967.

6.- "Foundations of modern mathematics". L. Mehlenbacher (Edit. Prindle, Weber & Schmidt) 1967.

7.- "The algebraic foundations of mathematics" R. A. Beaumont R. S. Pierce, (Edit. Addison-We y) 1963.

8.- "Introducción a la teoría de conjuntos" L. Oubiña (Edit. Univ. de Buenos Aires). 1969

9.- "Conjuntos" Aplicaciones matemáticas a la administración. Ariel Kleiman y Elena K. de Kleiman. (Edit. Limusa-Wiley, S.A., México) 1972.

EL SISTEMA DE LOS NUMEROS REALES

por

EDUARDO SOLAR GONZALEZ

LEDA SPEZIALE DE GUZMAN

profesores de la División de Ciencias Básicas
de la Facultad de Ingeniería de la U. N. A. M.

CAPITULO I EL SISTEMA DE LOS NUMEROS REALES

CONTENIDO

CAPITULO 1   EL SISTEMA DE LOS NUMEROS REALES

INTRODUCCION

Es difícil precisar cuándo aparecen en la historia los primeros números, pero tenemos la certeza de que éstos fueron los llamados números naturales (1, 2, 3,...), los cuales surgieron de la necesidad de contar.

Tan pronto como el hombre dejó de ser nómada para dedicarse a la agricultura y el comercio, tuvo necesidad de desarrollar formas cada vez mejores para contar y para representar los números. El estudio de cómo tales formas han evolucionado hasta convertirse en las actuales merecería por sí solo un libro completo.

El hombre primitivo requería únicamente de los primeros números naturales para contar. Sin embargo, con el advenimiento de la civilización hubo de utilizar números más y más grandes, hasta que los antiguos griegos dieron el salto al audaz concepto de infinito, noción que requirió de un alto grado de abstracción por ir más allá de toda experiencia física.

Curiosamente, el paso de los enteros positivos a los negativos resultó más difícil. Los números negativos aparecieron como soluciones de ecuaciones tan pronto como los matemáticos se ocuparon del álgebra. Los griegos, más preocupados por la geometría que por el álgebra, descartaron los números negativos al no poder representarlos gráficamente y adaptarlos a su geometría; mientras que los chinos y los hindúes, por el contrario, reconocieron la existencia de números negativos aún antes de la era cristiana. Los números negativos, sin embargo, no fueron incorporados completamente al cuerpo de las matemáticas sino hasta 1545 con la publicación de la obra de Girolamo Cardano, "Ars Magna".

Los números racionales (o fracciones) son más antiguos que los números negativos. Estos números aparecen en los más primitivos escritos matemáticos como el papiro Rhind, obra legada por la cultura egipcia.

Los números irracionales tienen también una larga historia. La necesidad de estos números se había presentado ya a los antiguos griegos en sus estudios geométricos; sin embargo, no se encontraron métodos satisfactorios para la construcción de estos números a partir de los racionales sino hasta el siglo XIX con los trabajos de Richard Dedekind. Fue en este siglo cuando los matemáticos lograron dar unidad y fundamento lógico al estudio de los números.

En el año de 1889, Giuseppe Peano propuso cinco axiomas o postulados que caracterizan completamente al conjunto de los números naturales; estos cinco postulados se utilizaron como punto de partida para una construcción total del sistema de los números reales.

Actualmente, el estudio de los números reales se puede emprender desde dos puntos de vista fundamentalmente distintos.

Uno de ellos es el conocido como "método constructivo", en el cual se estudian primero los números naturales y a partir de ellos se in - troducen los enteros, que a su vez sirven como base para definir los números racionales; por último, introduciendo los números irraciona - les, se completa la construcción.

El otro punto de vista es el llamado "enfoque axiomático", en el cual los números reales se definen como entes que satisfacen un conjunto de propiedades ya conocidas.

El enfoque axiomático aprovecha, por así decirlo, los resultados de esfuerzos de generaciones de matemáticos, para tomarlos como punto de partida. El método constructivo, por el contrario, se identifica más con la evolución histórica de las matemáticas. Las propiedades que en el segundo enfoque se aceptan como axiomas, en el método constructivo constituyen teoremas y por tanto deben ser demostradas.

La presentación que aquí haremos se asemeja más al método constructivo, pues creemos que esto contribuye a la comprensión de los concep - tos fundamentales. Sin embargo, no nos preocuparemos demasiado por cubrir ciertos aspectos formales de la construcción, pues ello podría desviar la atención hacia aspectos que no son de importancia fundamen tal en una primera aproximación al estudio de los números. El lector interesado en los aspectos formales puede recurrir a las referencias (1) y (2) que, esperamos, le dejarán satisfecho.

I.1 CONTAR

Seguramente todos los que estudien este libro sabrán contar, pero qui zá muy pocos se habrán preguntado qué es contar.

Contar es, en esencia, comparar.

El primer paso hacia contar lo constituye la noción de pluralidad o

cantidad, noción primitiva que poseen ya inteligencias poco desarro- lladas, como las de algunos animales. Los números vienen después y constituyen un patrón o referencia para establecer la comparación en tre cantidades.

Un pequeño de dos años que juega con tres esferas en su cuna percibe cuando se le esconde alguna de ellas; esto es, posee la noción de can tidad, aunque no sabe contar.

Si las esferas con que el niño juega fueran diez, en lugar de tres, difícilmente percibiría cuando una de ellas le fuera escondida. Sin embargo, si comparara las esferas con cada uno de sus dedos, segura - mente se daría cuenta cuando faltara alguna. ¡Estaría contando!

El proceso de contar, así pues, se fundamenta en el de "aparear"; proceso al que los matemáticos llaman "establecer una correspondencia uno a uno".

Para abundar sobre este concepto, consideremos un salón donde hay si- llas y personas. Para saber si la cantidad de sillas es igual a la de personas bastará con pedir a éstas que se sienten. Si sobran si - llas o personas sabremos que hay más de las primeras o más de las se- gundas, pero si no sobran sillas y todas las personas están sentadas diremos que hay tantas sillas como personas, esto es, existe una "co- rrespondencia uno a uno" entre el conjunto de sillas y el conjunto de personas que se encuentran en el salón.

Para contar, sin embargo, se requiere además seleccionar un patrón que nos sirva como referencia para comparar. Este patrón es la llama da "escala de los números naturales" que todos conocemos:

1, 2, 3, 4, 5, ...;

Cuando contamos un conjunto de objetos, lo que hacemos es poner en co rrespondencia uno a uno los objetos que contamos con los primeros números naturales, a partir del uno; el último número con el que establecemos la correspondencia nos indica la cantidad de objetos que tiene el conjunto. Así, si deseamos contar las letras de la palabra "ALGEBRA", establecemos la correspondencia:

A	L	G	E	B	R	A			
↕	↕	↕	↕	↕	↕	↕			
1	2	3	4	5	6	7	8	9	...

y encontramos que tiene siete letras.

Cabe resaltar que la palabra "siete" es un nombre que se ha asignado al número, de la misma manera que el carácter "7" es un símbolo que se utiliza para representarlo. Los símbolos que se emplean para representar números se llaman numerales.

Es importante no confundir al número, que es una idea abstracta, con su nombre o con su numeral, que no son más que una palabra o un símbolo que se emplean para distinguirlo. El número siete, por ejemplo, constituye una idea única; aunque en diferentes lenguas y épocas se hayan empleado otras palabras y numerales para referirse a ella; como la palabra "seven" en inglés y el numeral "VII" en la numeración romana.

1.2 LOS NUMEROS NATURALES

A continuación definiremos los números naturales, las operaciones de adición y multiplicación y algunos otros conceptos relacionados con ellos. El lector podrá preguntarse, no sin razón, cuál es el objeto de definir conceptos "tan conocidos". Una respuesta es que estamos tratando de construir el sistema de los números reales, y para ello necesitamos partir de bases sólidas definidas formalmente. Con esta idea buscamos definiciones que nos permitan demostrar propiedades de los números; propiedades que les dan unidad y gracias a las cuales podemos hablar de ellos como un sistema.

La tarea que ahora emprendemos puede parecer algo ardua, especialmente porque estamos en la etapa de tratar de establecer los conceptos más básicos. Sirva de consuelo notar que no fué sino hasta 1889, siglos después de que todo mundo usaba números y hacía operaciones con ellos, cuando un matemático — Peano — fué capaz de plantear las ideas que ahora nos ocupan.

- Postulados de Peano

Los cinco axiomas a que nos referimos en la introducción y que definen al conjunto de los números naturales, son los siguientes:

I.2.1 DEFINICION (Postulados de Peano)

El conjunto N de los números naturales es tal que:

i) $1 \in N$.

ii) Para cada $n \in N$ existe un único $n^* \in N$, llamado el siguiente de n.

iii) Para cada $n \in N$ se tiene que $n^* \neq 1$.

iv) Si $m, n \in N$ y $m^* = n^*$, entonces $m = n$.

v) Todo subconjunto de S de N que tenga las propiedades:

a) $1 \in S$

b) $k \in S$ implica que $k^* \in S$

es el mismo conjunto N

Los postulados i) a v) coinciden con algunas de las propiedades que intuitivamente aceptamos para los números naturales; empero, la importancia de estos postulados estriba en que son suficientes para deducir, a partir de ellos, todas las propiedades de los números naturales.

El primero de estos postulados establece que el número uno (que se considera conocido) es un número natural.

El segundo nos indica que si elegimos un número natural, sea cual fuere éste, a dicho número corresponde uno y sólo un número natural llamado su "siguiente".

El postulado iii) arroja como consecuencia que el número uno es el primer número natural, puesto que no es el siguiente de ninguno.

El postulado iv) nos dice que dos números naturales cuyos siguientes sean iguales son, en realidad, el mismo número. De este postulado y de la unicidad del siguiente, establecida en el postulado ii), se sigue también que dos números naturales diferentes tienen diferentes siguientes.

El postulado v) nos dice que podemos alcanzar cualquier número natural partiendo del uno y recorriendo los siguientes uno a uno hasta llegar al número natural deseado. Este postulado, conocido a menudo como "principio de inducción", es el fundamento del método de demostración por Inducción Matemática que trataremos posteriormente.

Como el lector se habrá dado cuenta, los postulados de Peano establecen las propiedades "intrínsecas", por llamarles de alguna manera, de los números naturales; otras propiedades son las algebraicas, que seguramente el lector ya conoce, las cuales son consecuencia de la introducción de las operaciones con números naturales, de las que no nos hemos ocupado aún.

- La adición en N

I.2.2 DEFINICIÓN

i) $n + 1 = n^*$, para todo $n \in N$

ii) $n + m^* = (n + m)^*$, siempre que $n + m$ esté definido

La definición anterior puede parecernos en principio extraña; sin embargo, recordemos que sólo podemos apoyarnos en los postulados de Peano para establecerla, ya que dichos postulados constituyen el fundamento que hemos elegido para desarrollar el sistema de los Números Naturales.

Para ilustrar el empleo de la definición I.2.2 en el cálculo de una suma, obtengamos el valor de $8 + 3$ siguiendo paso a paso la definición.

Para poder aplicar ii), escribimos

$$8 + 3 = 8 + 2^*$$

puesto que 3 es el siguiente de 2. Ahora, de ii) se sigue que

$$8 + 3 = (8 + 2)^*$$

Puesto que la definición no nos dice cual es el valor de $8 + 2$ aplicamos nuevamente ii), para lo cual hacemos

$$8 + 3 = (8 + 1^*)^*$$

puesto que 2 es el siguiente de 1. Entonces, de ii) tenemos que

$$8 + 3 = ((8 + 1)^*)^*$$

Ahora podemos ya aplicar el inciso i) de la definición, de donde

$$8 + 3 = ((8^*)^*)^*$$

Como $8^* = 9$, $9^* = 10$ y $10^* = 11$, queda

$$8 + 3 = (9^*)^*$$
$$8 + 3 = 10^*$$
$$8 + 3 = 11$$

con lo que obtenemos el resultado que el lector ya conocía. □

Como puede verse en el ejemplo anterior la definición I.2.2 es recursiva, y además nos dice que para sumar $n + m$ debemos recorrer m números naturales consecutivos a partir de n.

La adición, así definida, satisface las siguientes propiedades, que seguramente ya son del conocimiento del lector.

I.2.3 TEOREMA

Para todo $m, n, p \in N$:

i)	$m + n \in N$	cerradura
ii)	$m + (n + p) = (m + n) + p$	asociatividad
iii)	$m + n = n + m$	conmutatividad
iv)	si $m + p = n + p$, entonces $m = n$	cancelación

DEMOSTRACION

i) Sea m un número natural y formemos el conjunto

$$S = \{n \mid n \in N \text{ y } (m + n) \in N\}$$

esto es: un elemento del conjunto S es un número natural n que al sumarse con m da como resultado otro número natural.

Para demostrar la propiedad de cerradura bastará con probar que S es el conjunto de todos los números naturales, lo que haremos apoyándonos en el quinto postulado de Peano como sigue.

I) $m + 1 = m^*$, por i) de la definición I.2.2

$m + 1 \in N$, por ii) de la definición I.2.1.

Luego $1 \in S$.

II) Sea $k \in S$; es decir: $k \in N$ y $m + k \in N$ - - - - (1)

Como $m + k \in N$, del inciso ii) de I.2.1

$(m + k)^* \in N$

y del inciso ii) de I.2.2

$m + k^* \in N$ - - - - (2)

Además, como $k \in N$, del inciso ii) de I.2.1

$k^* \in N$ - - - - (3)

Finalmente, de (1), (2) y (3) se tiene que

$k \in S$ implica que $k^* \in S$.

Con lo que hemos probado que S es el conjunto de todos los números naturales.

Las propiedades ii) y iii) pueden demostrarse en forma análoga; para una demostración de iv) puede consultarse la referencia 2 pág. 96. □

Inducción Matemática

La idea fundamental bajo la cual se desarrolló la demostración anterior puede generalizarse para cualquier enunciado relativo a los números naturales. Las demostraciones así realizadas reciben el nombre de "demostraciones por inducción matemática" y su desarrollo general, con fundamento en el quinto postulado de Peano, es el siguiente:

Sea P(n) una proposición enunciada para todos los números naturales, y sea

$S = \{n \mid n \in N$ y $P(n)$ es verdadera$\}$

Para demostrar que P(n) es verdadera para todos los números naturales, bastará con probar que $S = N$; para lo cual hacemos lo siguiente:

I) Verificar que P(1) es verdadera (esto equivale a verificar que $1 \in S$).

II) Demostrar que si P(k) es verdadera entonces P(k + 1) es verdadera (esto equivale a demostrar que $k \in S \Rightarrow (k + 1) \in S$).

A partir de I) y II), y del quinto postulado de Peano, podemos concluir que P(n) es verdadera para todo $n \in N$.

I.2.4 EJEMPLOS

a) Consideremos el conjunto de los números impares

$\{x \mid x = 2n - 1,\ n \in N\} = \{1,3,5,7,9,...\}$

Si sumamos sus dos primeros elementos obtenemos

$1 + 3 = 4$

Este resultado puede también expresarse como

$1 + 3 = 2^2$

De manera semejante

$1 + 3 + 5 = 3^2$

$1 + 3 + 5 + 7 = 4^2$

Esto nos hace suponer que la suma de los n primeros números impares es igual a n^2. Es decir:

$1 + 3 + 5 + \ldots + (2n - 1) = n^2$ - - - - $P(n)$

La generalización anterior, representada por $P(n)$, se obtuvo del análisis de tres casos particulares, los que fueron extrapolados con ayuda del supuesto "sentido común"; sin embargo, hasta ahora sólo podemos asegurar que la proposición $P(n)$ es verdadera para $n = 2$, $n = 3$ y $n = 4$. Para tener la certeza de que $P(n)$ se cumple sin importar cual sea el número natural n, haremos la demostración, por inducción matemática, del siguiente enunciado:

$1 + 3 + 5 + \ldots + (2n - 1) = n^2$, para todo $n \in N$.

Demostración

I) Para $n = 1$, el primer miembro de $P(n)$ tiene sólo un término y $P(1)$ es

$$1 = (1)^2$$

Por tanto $P(1)$ es verdadera.

II) A partir de que $P(k)$ es verdadera debemos concluir que $P(k + 1)$ es también verdadera.

Suponemos entonces que $P(k)$ es verdadera; es decir:

$1 + 3 + 5 + \ldots + (2k - 1) = k^2$ - - - - (1)

(donde el primer miembro representa la suma de los k primeros números impares).

Como $P(k + 1)$ es

$1 + 3 + 5 + \ldots + (2k + 1) = (k + 1)^2$ - - - - (2)

(donde el primer miembro representa la suma de los $k + 1$ primeros números impares), debemos demostrar que esta igualdad es verdadera partiendo de la expresión (1); lo cual puede hacerse de la siguiente manera:

Sumando en ambos miembros de (1) el número $2(k + 1) - 1$, que es el número impar siguiente de $2k - 1$, tenemos

$1 + 3 + 5 + \ldots + (2k - 1) + (2(k + 1)-1) = k^2 + (2(k + 1)-1)$

(donde el primer miembro también representa la suma de los $k + 1$ primeros números impares). Esta expresión puede escribirse como

$1 + 3 + 5 + \ldots + (2k + 1) = k^2 + 2k + 1$

o también como

$1 + 3 + 5 + \ldots + (2k + 1) = (k + 1)^2$

que coincide con la expresión (2), por lo que $P(k + 1)$ es verdadera.

Con lo anterior hemos demostrado que si $P(k)$ es verdadera entonces $P(k + 1)$ es verdadera y la prueba termina $\square$

b) Como otro ejemplo demostraremos ahora el siguiente enunciado:

$n^2 + n$ es un número par, $\forall\, n \in N$

Demostración

I) $P(1)$ dice que

$1^2 + 1$ es un número par

$P(1)$ es verdadera ya que $1^2 + 1 = 2$ y 2 es un número par.

II) Hipótesis: P(k) es verdadera; por lo que

$k^2 + k$ es un número par

Por otra parte

$$(k + 1)^2 + (k + 1) = k^2 + 2k + 1 + k + 1$$
$$= k^2 + k + 2k + 2$$
$$(k + 1)^2 + (k + 1) = (k^2 + k) + 2(k + 1)$$

Como $(k + 1) \in N$, $2(k + 1)$ es un número par y, de la expresión anterior, $(k + 1)^2 + (k + 1)$ será un número par si $k^2 + k$ es par.

En consecuencia, de la hipótesis se sigue que

$(k + 1)^2 + (k + 1)$ es un número par

Hemos demostrado que si P(k) es verdadera entonces P(k + 1) es verdadera y la prueba termina.

□

- La multiplicación en N

En forma similar a como se hizo para la adición, la multiplicación puede definirse como sigue.

I.2.5 DEFINICION

i) $n \cdot 1 = n$

ii) $n \cdot m^* = (n \cdot m) + n$

La multiplicación, así definida, satisface las propiedades que se resumen en el siguiente Teorema.

I.2.6 TEOREMA

Para todo $m, n, p \in N$:

i)	$m \cdot n \in N$	cerradura
ii)	$m \cdot (n \cdot p) = (m \cdot n) \cdot p$	asociatividad
iii)	$m \cdot n = n \cdot m$	conmutatividad
iv)	si $m \cdot p = n \cdot p$, entonces $m = n$	cancelación

Las propiedades i) a iii) pueden demostrarse directamente por inducción matemática y su demostración se deja al lector como ejercicio. Para una demostración de la propiedad iv) puede consultarse la referencia 2 pág. 96.

Tomadas simultáneamente, las operaciones de adición y multiplicación satisfacen la siguiente ley distributiva.

I.2.7 TEOREMA

Para todo $m, n, p \in N$:

$m \cdot (n + p) = (m \cdot n) + (m \cdot p)$

DEMOSTRACION

Sean n y p dos números naturales, y consideremos el enunciado

$m \cdot (n + p) = (m \cdot n) + (m \cdot p)$, $\forall\, m \in N$

el cual, por el inciso iii) del teorema I.2.6, es equivalente a

$(n + p) \cdot m = (n \cdot m) + (p \cdot m)$, $\forall\, m \in N$

que demostraremos por inducción matemática.

I) P(1) dice que $(n + p) \cdot 1 = (n \cdot 1) + (p \cdot 1)$, y es verdadera ya que

$(n + p) \cdot 1 = n + p$	por i) de I.2.5
$(n + p) \cdot 1 = (n \cdot 1) + (p \cdot 1)$	por i) de I.2.5

II) Hipótesis: $P(k)$ es verdadera, esto es

$(n + p)k = (n \cdot k) + (p \cdot k)$

Entonces

$$
\begin{aligned}
(n + p) \cdot (k + 1) &= (n + p) \cdot k^* && \text{por i) de I.2.2} \\
&= (n + p) \cdot k + (n + p) && \text{por ii) de I.2.5} \\
&= (n \cdot k) + (p \cdot k) + (n + p) && \text{por hipótesis} \\
&= n \cdot k + (p \cdot k + n) + p && \text{por ii) de I.2.3} \\
&= n \cdot k + (n + p \cdot k) + p && \text{por iii) de I.2.3} \\
&= (n \cdot k + n) + (p \cdot k + p) && \text{por ii) de I.2.3} \\
&= (n \cdot k^*) + (p \cdot k^*) && \text{por ii) de I.2.5} \\
(n + p) \cdot (k + 1) &= n \cdot (k + 1) + p \cdot (k + 1) && \text{por i) de I.2.2}
\end{aligned}
$$

Con lo que hemos probado que si $P(k)$ es verdadera entonces $P(k + 1)$ es verdadera, lo cual completa la demostración. □

Orden en N

Cuando hablamos de cantidades es frecuente efectuar comparaciones y decimos, por ejemplo, que cierta cantidad es menor que otra. Puesto que los números naturales nos sirven para expresar cantidades, es necesario definir la relación "menor que" en N.

Decimos, por ejemplo, que 3 es menor que 5. Si observamos que existe un número natural, en este caso el 2, tal que

$3 + 2 = 5$

podemos establecer, a partir de la suma, una definición para la relación "menor que", que coincida con la idea intuitiva que de ella tenemos.

I.2.8 DEFINICION

Dados dos números naturales n y m, decimos que n es menor que m, lo que representamos mediante $n < m$, si

$$\exists\, x \in N \text{ tal que } n + x = m.$$

Los números naturales satisfacen la siguiente propiedad, llamada Ley de Tricotomía.

I.2.9 TEOREMA

Si m y n son números naturales cualesquiera, entonces se verifica una y sólo una de las siguientes proposiciones:

i) $n < m$

ii) $n = m$

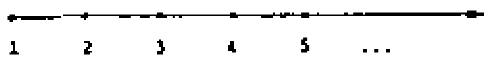
iii) $m < n$

Para una demostración, el lector puede consultar la referencia 3 pág. 36

La definición I.2.8 y el teorema I.2.9, establecen un orden en el conjunto N ya que:

$1 + 1 = 2$	por lo que	$1 < 2$
$2 + 1 = 3$	por lo que	$2 < 3$
$3 + 1 = 4$	por lo que	$3 < 4$
$4 + 1 = 5$	por lo que	$4 < 5$
⋮		⋮

y podemos representar gráficamente a los números naturales como puntos igualmente espaciados sobre una recta, a la que llamamos recta numérica:



En ella, cuando $m < n$ se tendrá que el punto que representa a m se encuentra a la izquierda del que representa a n.
Con base en I.2.8 puede demostrarse que la relación "menor que" en N tiene las siguientes propiedades.

I.2.10 TEOREMA

Para todo $m, n, p \in N$:

i) $m < n \Rightarrow m + p < n + p$

ii) $m < n \Rightarrow mp < np$

iii) $m < n$ y $n < p \Rightarrow m < p$

En ocasiones resulta más cómodo emplear la relación "menor que" en el sentido inverso; esto es, decir que un número es mayor que otro, por lo que estableceremos la siguiente definición.

I.2.11 DEFINICION

Dados dos números naturales m y n, decimos que m es mayor que n, lo que representamos mediante $m > n$, si $n < m$.

Es claro que la relación "mayor que" tiene propiedades análogas a las establecidas en el teorema I.2.10.

I.2.12 EJERCICIOS

1) Demostrar que para todo $m, n, p \in N$:

a) $m + (n + p) = (m + n) + p$
b) $1 + n = n + 1$
c) $m + n = n + m$

2) Demostrar que para todo $m, n, p \in N$:

a) $m \cdot n \in N$
b) $m \cdot (n \cdot p) = (m \cdot n) \cdot p$
c) $m \cdot n = n \cdot m$

3) Demostrar que si $m, n \in N$ entonces $m + n \neq m$

4) Demostrar que para todo $m, n \in N$:

a) $m^* + n = m + n^*$
b) $m^* \cdot n \neq m \cdot n^*$; si $m \neq n$
c) $m^* + n^* = [(m + n)^*]^*$

5) Demostrar que si $n \in N$ entonces:

a) $m + m = 2m$
b) $\underbrace{m + m + \dots + m}_{n \text{ sumandos}} = n \cdot m$

6) Si $n, p \in N$ se define

$$p^n = \underbrace{p \cdot p \cdot \dots \cdot p}_{n \text{ factores}}$$

Demostrar que para todo $m, n, p, q \in N$:

a) $p^m \cdot p^n = p^{m+n}$
b) $(p^m)^n = p^{m \cdot n}$
c) $(p \cdot q)^n = p^n \cdot q^n$

7) Demostrar que para todo $m, n, p \in N$:

a) $m < n \Rightarrow m + p < n + p$
b) $m > n \Rightarrow m \cdot p > n \cdot p$
c) $m < n$ y $n < p \Rightarrow m < p$

I.3 LOS NUMEROS ENTEROS

El sistema de los números naturales, con la operación de adición, manifiesta su primera deficiencia tan pronto como empezamos a plantear ecuaciones en dicho sistema, ya que una ecuación del tipo

$n + x = m$; con $m, n \in N$ - - - - (A)

puede tener solución en N o no tenerla.

Por ejemplo, la solución de la ecuación

$3 + x = 7$

es el número que sumado a 3 nos arroje como resultado 7. En este caso, el número natural cuatro satisface la condición pedida; sin embargo, si consideramos la ecuación

$3 + x = 1$

encontraremos que no existe número natural alguno que la satisfaga.

Esta deficiencia crea la necesidad de ampliar el conjunto de los números naturales. Surgen así los números negativos y el cero que, con los naturales, constituyen el conjunto de los números enteros.

- La diferencia de números naturales

I.3.1 DEFINICION

Sea la ecuación:

$n + x = m$; con $m, n \in N$

A su solución; es decir, al número x que sumado a n nos da como resultado m, lo llamaremos la diferencia $m - n$.

Según el teorema I.2.5 para $m - n$ se presentan tres casos:

i) $n < m$

Este es el único caso en que la ecuación (A) tiene solución en N, lo cual es consecuencia inmediata de la definición I.2.8, por lo que $m - n$ es un número natural.

ii) $n = m$

En este caso la ecuación (A) no tiene solución en N y toma la forma

$n + x = n$

En consecuencia $m - n$ no es un número natural, lo definimos como el cero y lo representamos con 0.

iii) $m < n$

En este caso, como en el anterior, la ecuación (A) no tiene solución en N, por lo que $m - n$ tampoco es un número natural; lo definimos como un número entero negativo y lo representamos con $-(n-m)$, donde $(n - m) \in N$.

Así, por ejemplo, la solución de la ecuación

$3 + x = 1$

es el número que sumado a 3 nos da como resultado 1. A dicho número, que es un entero negativo, lo representamos con -2.

Como vimos, el "número" $m - n$ no es siempre un número natural, por lo que el concepto de número tal y como lo hemos manejado hasta ahora resulta ya demasiado restrictivo. A partir de aquí consideraremos como números a todos los entes matemáticos que resulten de la ampliación progresiva del conjunto N, por lo que la palabra número tendrá una acepción cada vez más amplia.

De acuerdo con lo anterior, a los números que se obtienen mediante la diferencia de dos números naturales les llamaremos números enteros y al conjunto que forman lo representaremos con Z. Esto es:

I.3.2 DEFINICION

$$Z = \{x \mid x = m - n;\ m, n \in N\}$$

Es claro que el subconjunto Z^+ de Z, definido por

$$Z^+ = \{x \mid x = m - n;\ m, n \in N;\ m > n\}$$

al cual se le conoce como "conjunto de los enteros positivos", es precisamente el conjunto de los números naturales, por lo que

$$N \subseteq Z$$

- La igualdad en Z.

Como consecuencia de la definición I.3.1, un número entero puede ser expresado en la forma $m - n$ de una infinidad de maneras distintas.

Por ejemplo, el número $x = -2$ puede expresarse como $1 - 3$ ya que es solución de la ecuación

$$3 + x = 1$$

Es claro, sin embargo, que dicho número puede además expresarse como $2 - 4$, ya que también es solución de

$$4 + x = 2$$

así como de muchas otras ecuaciones del tipo $m + x = n$, con $m, n \in N$.

Debido a esto es necesario establecer un criterio que nos permita decidir si dos números enteros expresados como la diferencia de dos naturales son iguales o no lo son. Dicho criterio es el siguiente.

I.3.3 DEFINICION

Sean $a = m - n$, $b = p - q$ dos números enteros, con $m, n, p, q \in N$. Entonces:

$$a = b \quad \text{si} \quad m + q = n + p$$

Así, con base en esta definición podemos concluir que $a = 1 - 3$ y $b = 2 - 4$ representan al mismo número entero, ya que

$$1 + 4 = 3 + 2$$

y se satisface la igualdad en N que requiere I.3.3 para establecer la igualdad en Z.

• La adición en Z.

Como $N \subset Z$, la adición en Z debe producir los mismos resultados que la adición en N cuando los enteros que se suman son positivos, lo que conduce a la siguiente definición:

I.3.4 DEFINICION

Sean $a = m - n$, $b = p - q$ dos números enteros, con $m, n, p, q \in N$. El número $a + b$ se define como:

$$a + b = (m + p) - (n + q).$$

Cabe hacer notar que, en la definición anterior, el símbolo + que está a la izquierda del signo igual representa la adición de números enteros (la cual se está definiendo), mientras que el símbolo + que está a la derecha representa la adición de números naturales.

Para ilustrar el empleo de la definición I.3.4 calcularemos a continuación la suma $8 + 3$. Con objeto de hacer notar la diferencia entre las operaciones de adición en Z y en N, en la discusión que sigue representaremos estas operaciones con dos símbolos diferentes:

$\oplus$ para la adición en Z

$+$ para la adición en N

Los números enteros 8 y 3 pueden expresarse como

$$8 = 9 - 1$$

$$3 = 7 - 4$$

y aplicando I.3.4 tenemos

$$8 \oplus 3 = (9 + 7) - (1 + 4)$$

De la definición I.2.2 se sigue que

$$8 \oplus 3 = 16 - 5$$

es decir: es el número que sumado a 5 da como resultado 16, por lo que

$$8 \oplus 3 = 11.$$

Cabe hacer notar que, como 8 y 3 son también números naturales, el mismo resultado debería obtenerse empleando directamente la definición I.2.2. En efecto, en el ejemplo ilustrativo de I.2.2, se obtuvo que

$$8 + 3 = 11$$

□

Como el lector habrá notado, en I.3.4 hemos definido la adición en $\mathbb{Z}$ a partir de la adición en $\mathbb{N}$ y la definición de número entero.

La definición I.3.4 nos permite demostrar las propiedades que comúnmente empleamos cuando tratamos con la adición de números enteros, algunas de las cuales se resumen en el siguiente teorema. Cabe hacer notar que la adición en $\mathbb{Z}$ satisface todas las propiedades que establecimos para la adición en $\mathbb{N}$, sin embargo, la adición en $\mathbb{Z}$ cuenta con otras propiedades adicionales, como la existencia de un elemento idéntico y la existencia de elementos inversos.

I.3.5 TEOREMA

Para todo $a, b, c \in \mathbb{Z}$:

i)	$a + b \in \mathbb{Z}$	cerradura
ii)	$a + (b + c) = (a + b) + c$	asociatividad
iii)	$a + b = b + a$	conmutatividad
iv)	si $a + c = b + c$, entonces $a = b$	cancelación
v)	$a + 0 = a$	elemento idéntico
vi)	$\exists\, -a \in \mathbb{Z}$ tal que $a + (-a) = 0$	elementos inversos

DEMOSTRACION

Se demostrarán únicamente i), v) y vi)

i) Sean $a = m - n$ y $b = p - q$, dos números enteros cualesquiera.

De la definición I.3.4

$$a + b = (m + p) - (n + q)$$

como $m, n, p, q \in \mathbb{N}$, del inciso i) de I.2.3

$$(m + p) \in \mathbb{N} \text{ y } (n + q) \in \mathbb{N}$$

por lo que, de I.3.2

$$(m + p) - (n + q) \in \mathbb{Z}$$

En consecuencia

$a + b \in \mathbb{Z}$, como se quería.

v) Se definió el cero como el número x tal que

$$n + x = n, \text{ con } n \in \mathbb{N}$$

por lo que puede expresarse como

$$0 = n - n$$

Entonces, si $a = m - n$ es un número entero

$$a + 0 = (m - n) + (n - n)$$

De la definición I.3.4

$$a + 0 = (m + n) - (n + n)$$

Por lo que $a + 0$ es tal que

$$(n + n) + (a + 0) = m + n$$

Ahora, por las propiedades ii), iii) y iv) de I.3.5 se tiene que

$$n + [n + (a + 0)] = m + n$$

$$[n + (a + 0)] + n = m + n$$

$$n + (a + 0) = m$$

En consecuencia

$$a + 0 = m - n$$

Esto es

$a + 0 = a$, como se quería.

vi) Sea $a = m - n$ un número entero. Dicho número es tal que

$$n + a = m, \text{ con } m, n \in N.$$

En consecuencia, la solución de la ecuación

$$m + x = n, \text{ con } m, n \in N$$

será el número $x = n - m$ que, por I.3.2, es un número entero.

Consideremos ahora la suma

$$a + x = (m - n) + (n - m)$$

De la definición I.3.4

$$a + x = (m + n) - (n + m)$$

y por el inciso iii) de I.2.3

$$a + x = (m + n) - (m + n)$$

En consecuencia

$$a + x = 0$$

Luego, existe $x \in Z$ tal que $a + x = 0$

y la prueba termina. □

A dicho número x (que como se ve en la prueba depende de a) se le denomina "el inverso de a para la suma" y se le representa con $-a$.

La sustracción en Z puede definirse ahora a partir de la adición y de vi) de I.3.5, de la siguiente manera.

I.3.6 DEFINICION

Sean $a, b \in Z$, el número $a - b$ se define como

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

y puede demostrarse que esta operación tiene la propiedad de cerradura; esto es

$$a - b \in Z; \ \forall\, a, b \in Z.$$

- La multiplicación en Z

En forma similar a como se hizo para la adición, la multiplicación en Z puede definirse como sigue

I.3.7 DEFINICION

Sean $a = m - n$, $b = p - q$ dos números enteros, con $m, n, p, q \in N$. El número $a \cdot b$ se define como

$$a \cdot b = (m \cdot p + n \cdot q) - (n \cdot p + m \cdot q)$$

El símbolo $\cdot$ que está a la izquierda del signo igual representa la multiplicación en Z (la cual se está definiendo), mientras que los símbolos $\cdot$ y $+$ que están a la derecha representan la multiplicación y la adición de números naturales. Como puede verse, la multiplicación en Z quedó definida a partir de la multiplicación y la adición en N y la definición de número entero.

La multiplicación, así definida, satisface las propiedades enunciadas a continuación.

I.3.8 TEOREMA

Para todo $a, b, c \in Z$:

i)	$a\cdot b \in Z$	cerradura
ii)	$a\cdot(b\cdot c) = (a\cdot b)\cdot c$	asociatividad
iii)	$a\cdot b = b\cdot a$	conmutatividad
iv)	si $a\cdot c = b\cdot c$ y $c \neq 0$, entonces $a = b$	cancelación
v)	$a\cdot 1 = a$	elemento idéntico

DEMOSTRACION

Se demostrarán únicamente las propiedades iii) y v).

iii) Sean $a = m - n$ y $b = p - q$ dos números enteros cualesquiera

$a\cdot b = (m - n)\cdot(p - q)$

$= (mp + nq) - (np + mq)$ — por I.3.7

$= (mp + nq) - (mq + np)$ — por iii) de I.2.3

$= (pm + qn) - (qm + pn)$ — por iii) de I.2.6

$= (p - q)\cdot(m - n)$ — por I.3.7

$a\cdot b = b\cdot a$ — como se quería

v) En i) de I.2.2 se definió

$n + 1 = n^*$

por lo que el número 1 puede expresarse como

$1 = n^* - n, \ \forall n \in N$

Entonces, si $a = m - n$ es un número entero

$a\cdot 1 = (m - n)\cdot(n^* - n)$

$= (mn^* + nn) - (nn^* + mn)$ — por I.3.7

$= [(mn + m) + nn] - [(nn + n) + mn]$ — por ii) de I.2.5

$a\cdot 1 = [(mn + nn) + m] - [(mn + nn) + n]$ — por ii) y iii) de I.2.3

por lo que, $a\cdot 1$ es tal que

$[(mn + nn) + n] + (a\cdot 1) = (mn + nn) + m$

y, de ii) de I.3.5

$(mn + nn) + [n + (a\cdot 1)] = (mn + nn) + m$

En consecuencia, de iii) y iv) de I.3.5

$n + (a\cdot 1) = m$

Es decir

$a\cdot 1 = m - n$

$a\cdot 1 = a$

y la prueba termina

□

Como el lector habrá notado, en la demostración anterior hemos omitido el símbolo · en una parte del proceso. Debido a que esta omisión es usual en los libros de matemáticas y facilita la escritura, en adelante nosotros también omitiremos el símbolo $\cdot$ en la multiplicación, así como algunos otros símbolos que no son indispensables pero que hemos empleado anteriormente para enfatizar los primeros conceptos.

Tomadas simultáneamente, la adición y la multiplicación satisfacen la siguiente propiedad distributiva.

> **I.3.9 TEOREMA**
>
> Para todo $a, b, c \in Z$:
>
> $$a\,(b + c) = ab + ac$$

DEMOSTRACION

Sean $a = m - n$, $b = p - q$, $c = r - s$, tres números enteros cualesquiera. Entonces

$$a\,(b + c) = (m - n)\,[(p - q) + (r - s)]$$

$$= (m - n)\,[(p + r) - (q + s)] \qquad \text{por I.3.4}$$

$$= [m(p + r) + n(q + s)] - [n(p + r) + m(q + s)] \qquad \text{por I.3.3}$$

$$= [(mp + mr) + (nq + ns)] - [(np + nr) + (mq + ms)] \qquad \text{por I.2.7}$$

$$= [(mp + nq) + (mr + ns)] - [(np + mq) + (nr + ms)] \qquad \text{por ii) y iii) de I.2.3}$$

$$= [(mp + nq) - (np + mq)] + [(mr + ns) - (nr + ms)] \qquad \text{por I.3.4}$$

$$= (m - n)(p - q) + (m - n)(r - s) \qquad \text{por I.3.7}$$

$$a(b + c) = ab + ac$$

con lo que concluye la demostración. □

La introducción del cero y los negativos trae como consecuencia la aparición de algunas propiedades adicionales para la multiplicación en Z, propiedades de las que no disponíamos para la multiplicación en N, algunas de las cuales se enuncian en el siguiente teorema:

> **I.3.10 TEOREMA**
>
> Para todo $a, b \in Z$:
>
> i) $a \cdot 0 = 0$
>
> ii) $(-a)(b) = -(ab)$ primera regla de los signos
>
> iii) $(-a)(-b) = ab$ segunda regla de los signos

DEMOSTRACION

Se demostrarán únicamente i) y ii)

i) $0 + 0 = 0$ por v) de I.3.5

$a(0 + 0) = a \cdot 0$ multiplicando por $a \in Z$

$a \cdot 0 + a \cdot 0 = a \cdot 0$ por I.3.9

$a \cdot 0 + a \cdot 0 = a \cdot 0 + 0$ por v) de I.3.5

$$a\cdot 0 + a\cdot 0 = 0 + a\cdot 0 \qquad \text{por iii) de I.3.5}$$

$$a\cdot 0 = 0 \qquad \text{por iv) de I.3.5}$$

ii) Sean $a,\ b \in \mathbb{Z}$

$$\begin{aligned}
(-a)(b) + ab &= (b)(-a) + ba && \text{por iii) de I.3.8}\\
&= b(-a + a) && \text{por I.3.9}\\
&= b\,[a + (-a)] && \text{por iii) de I.3.5}\\
&= b\cdot 0 && \text{por vi) de I.3.5}\\
(-a)(b) + ab &= 0 && \text{por i) de I.3.10}\\
ab + (-a)(b) &= 0 && \text{por iii) de I.3.5}
\end{aligned}$$

De esta última expresión se sigue que $(-a)(b)$ es el inverso de ab para la suma; esto es

$$(-a)(b) = -(ab)$$

como se quería.

□

- Orden en Z

Para los números enteros podemos también definir la relación "menor que", como una generalización de la que hemos definido para los naturales.

I.3.11 DEFINICION

Sean $a,\ b \in \mathbb{Z}$:

i) $a < b$ si $\exists\, n \in \mathbb{N}$ tal que $a + n = b$

ii) $a > b$ si $b < a$

Los números enteros también satisfacen la Ley de Tricotomía, enunciada para los números naturales en el teorema I.2.9; y la relación "menor que" tiene en Z las siguientes propiedades

I.3.12 TEOREMA

Para todo $a,b,c \in \mathbb{Z}$:

i) $a < b \Rightarrow a + c < b + c$

ii) $a < b$ y $c > 0 \Rightarrow ac < bc$
$a < b$ y $c < 0 \Rightarrow ac > bc$

iii) $a < b$ y $b < c \Rightarrow a < c$

cuya demostración se deja al lector como ejercicio.

Los números enteros quedan representados también en la recta numérica como puntos igualmente espaciados, sólo que ahora la recta se extiende indefinidamente en ambos sentidos

... -3 -2 -1 0 1 2 3 ...

Si $a < b$ se tendrá que el punto que representa a a estará a la izquierda del que representa a b.

Como ahora la recta se extiende en ambos sentidos no tiene un punto inicial, como sucede para los números naturales, por lo que se considera como punto de referencia el punto que representa al cero. Los números que se encuentran representados a la derecha de dicho punto se dice que son "positivos" y los que están a la izquierda "negativos".

I.3.13 DEFINICION

Sea $a \in Z$:

a es positivo si $a > 0$

a es negativo si $a < 0$

En particular, el cero no es positivo ni negativo.

Al conjunto Z^+ que se definió a continuación de I.3.2 como:

$Z^+ = \{x \mid x = m - n;\ m,n \in N;\ m > n\}$

se le conoce como el conjunto de los "enteros positivos". Puede demostrarse que todo elemento de dicho conjunto es positivo en el sentido que establece la definición I.3.13; esto es:

$x \in Z^+$ si y sólo si $x \in Z$ y $x > 0$.

I.3.12 EJERCICIOS

1) Demostrar a partir de I.3.3 que para todo $m, n \in N$:

a) $m - m^* = n - n^*$

b) $m^* - n \neq m - n^*$

2) Demostrar que para todo $a, b, c \in Z$:

a) $a + b = b + a$

b) $a + c = b + c \Rightarrow a = b$

3) Demostrar que si $a, b, c \in Z$ entonces:

a) $a - b \in Z$

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

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

4) Demostrar que para todo $a, b \in Z$:

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

b) $(-a)(-b) = ab$

5) Demostrar que si $a, b, c \in Z$ entonces:

a) $a(b - c) = ab - ac$

6) Demostrar que si $a, b, c \in Z$:

a) $a > 0 \Leftrightarrow -a < 0$

b) $a - b < a + b$, si $b > 0$

c) $a < b$ y $c > 0 \Rightarrow ac < bc$

$a < b$ y $c < 0 \Rightarrow ac > bc$

1.4 LOS NUMEROS RACIONALES

Al inicio de la sección 1.3 planteamos una ecuación del tipo

$n + x = m$; con $m, n \in N$ - - - - (A)

y vimos que los números naturales son insuficientes para proporcionar soluciones en todos los casos. Esta deficiencia se superó construyendo el conjunto Z, en el cual siempre hallamos solución a una ecuación de tal tipo.

El conjunto de los números enteros, sin embargo, también presenta deficiencias. En efecto, si planteamos ahora una ecuación en términos de la multiplicación como

$bx = a$; con $a, b \in Z$ - - - - (B)

encontramos que no siempre tiene solución en Z.

Por ejemplo, la solución de la ecuación

$3x = 12$

es el número entero 4, que multiplicado por 3 nos da como resultado 12. Sin embargo, si consideramos la ecuación

$3x = 7$

encontramos que no existe un número entero que multiplicado por 3 dé como resultado 7.

Esta deficiencia crea la necesidad de ampliar ahora el conjunto de los números enteros, con lo que surgen los números fraccionarios. Estos números junto con los enteros forman el conjunto de los números racionales.

- El cociente de números enteros.

1.4.1 DEFINICION

Sea la ecuación:

$bx = a$; con $a, b \in Z$

A su solución; es decir, al número x que multiplicado por b nos da como resultado a, le llamaremos el cociente de a entre b y lo representaremos con $\frac{a}{b}$.

Como el lector recordará de sus cursos elementales de aritmética, un entero $b \neq 0$ se dice factor de un entero a si existe un $c \in Z$ tal que

$bc = a$

Así, para el cociente $\frac{a}{b}$ podemos distinguir tres casos:

i) b es factor de a.

Este es el único caso en que la ecuación (B) tiene solución en Z, por lo que $\frac{a}{b}$ es un número entero.

ii) b no es factor de a y $b \neq 0$

En este caso la ecuación (B) no tiene solución en Z, por lo que $\frac{a}{b}$ no es un número entero y decimos que es un número fraccionario.

Así, por ejemplo, la solución de la ecuación

$3x = 7$

es el número que multiplicado por 3 nos da como resultado 7. A dicho número lo representamos con $\frac{7}{3}$ y es un número fraccionario.

iii) $b = 0$.

Este es un caso que merece especial atención, en virtud de que $\frac{a}{b}$ no está definido.

En efecto; si $a \neq 0$, $\frac{a}{b}$ no existe; ya que no hay un número que multiplicado por cero dé como resultado un número diferente de cero.

Por otra parte, si $a = 0$, $\frac{a}{b}$ no está determinado; ya que cualquier número multiplicado por cero de como resultado el cero.

En cualquier caso, el número $\frac{a}{b}$ con $b = 0$ no está definido.

De acuerdo con lo anterior, a los números que se obtienen como el cociente de dos números enteros a, b con $b \neq 0$ les llamaremos números racionales y al conjunto que forman lo representaremos con Q. Esto es:

1.4.2 DEFINICION

$$Q = \{x \mid x = \tfrac{a}{b},\ a,\ b \in Z,\ b \neq 0\}$$

Es claro que el subconjunto de Q definido por

$$\{x \mid x = \tfrac{a}{b},\ a,\ b \in Z,\ b = 1\}$$

es precisamente el conjunto de los números enteros, por lo que

$$Z \subset Q$$

- La igualdad en Q

Seguramente el lector ha manejado ya ampliamente los números racio-

nales en forma de "quebrados", y se habrá dado cuenta que la expresión de un número racional en la forma $\frac{a}{b}$ no es única.

Por ejemplo, el número por el que debemos multiplicar 12 para obtener 28 puede representarse como

$$\frac{28}{12},\ \frac{-14}{-6},\ \frac{7}{3},\ \text{etc.}$$

donde la última se conoce como su "mínima expresión".

En general, si $x = \frac{a}{b}$ es un número racional con $a,\ b \in Z$ y $b > 0$, si el único factor común de a y b es el número uno decimos que $\frac{a}{b}$ es la mínima expresión del número racional x.

De manera recíproca, si $x = \frac{a}{b}$ es un número racional y $k \neq 0$ es un número entero, $\frac{ka}{kb}$ representa también al número x.

De acuerdo con lo anterior, resulta natural considerar que dos números racionales $\frac{a}{b}$ y $\frac{c}{d}$ son iguales cuando

$$a = kc \text{ y } b = kd$$

o cuando

$$c = ka \text{ y } d = kb$$

para algún $k \in Z$, $k \neq 0$

Si se cumple alguna de estas dos condiciones se tendrá que

$$ad = bc$$

y viceversa, lo que conduce a la siguiente definición

I.4.3 DEFINICION

Sean $\frac{a}{b}$, $\frac{c}{d}$ dos números racionales, con $a,b,c,d \in \mathbb{Z}$ y $b,\ d \neq 0$, entonces

$$\frac{a}{b} = \frac{c}{d} \qquad \text{si} \qquad ad = bc$$

la cual establece la igualdad de números racionales en términos de la igualdad de números enteros.

- La adición en Q

I.4.4 DEFINICION

Sean $\frac{a}{b}$, $\frac{c}{d}$ dos números racionales, donde $a,b,c,d \in \mathbb{Z}$ y $b,d \neq 0$.

El número $\frac{a}{b} + \frac{c}{d}$ se define como

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

La adición en Q, así definida, tiene las propiedades que se enuncian a continuación

I.4.5 TEOREMA

Para todo $x,\ y,\ z \in \mathbb{Q}$:

i)	$x + y \in \mathbb{Q}$	cerradura
ii)	$x + (y + z) = (x + y) + z$	asociatividad
iii)	$x + y = y + x$	conmutatividad
iv)	si $x + z = y + z$, entonces $x = y$	cancelación
v)	$x + 0 = x$	elemento idéntico
vi)	$\exists\ -x \in \mathbb{Q}$ tal que $x + (-x) = 0$	elementos inversos

DEMOSTRACION

Se demostrará únicamente vi)

Sea $x = \frac{a}{b}$, con $a,\ b \in \mathbb{Z}$ y $b \neq 0$

Por vi) de I.3.3, $\exists\ -a \in \mathbb{Z}$ y de I.4.2, $\frac{-a}{b} \in \mathbb{Q}$.

Ahora

$$\begin{aligned}
\frac{a}{b} + \frac{-a}{b} &= \frac{ab + (-a)b}{b \cdot b} && \text{por I.4.4} \\
&= \frac{ba + b(-a)}{b \cdot b} && \text{por iii) de I.3.8} \\
&= \frac{b[a + (-a)]}{b \cdot b} && \text{por I.3.9} \\
&= \frac{[a + (-a)]}{b} && \text{por I.4.3} \\
&= \frac{0}{b} && \text{por vi) de I.3.5} \\
&= \frac{0}{1} && \text{por I.4.3}
\end{aligned}$$

$$\frac{a}{b} + \frac{-a}{b} = 0$$

Con lo que $-x = \frac{-a}{b}$ y la prueba termina. $\square$

La sustracción en Q puede definirse ahora a partir de la adición y de vi) de I.4.5, de la siguiente manera:

I.4.6 DEFINICION

Sean $\frac{a}{b}, \frac{c}{d} \in Q$, al número $\frac{a}{b} - \frac{c}{d}$ se define como

$$\frac{a}{b} - \frac{c}{d} = \frac{a}{b} + \frac{-c}{d}$$

Como consecuencia de i) de I.4.5, la sustracción es cerrada en Q; esto es

$\forall x, y \in Q: \; x - y \in Q.$

- La Multiplicación en Q

I.4.7 DEFINICION

Sean $\frac{a}{b}, \frac{c}{d}$ dos números racionales, donde $a, b, c, d, \in Z$ y $b, d \neq 0$.

El número $\frac{a}{b} \cdot \frac{c}{d}$ se define como

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$$

La multiplicación en Q, así definida, satisface las propiedades que establece el siguiente teorema.

I.4.8. TEOREMA

Para todo $x, y, z \in Q$:

i)	$xy \in Q$	cerradura
ii)	$x(yz) = (xy)z$	asociatividad
iii)	$xy = yx$	conmutatividad
iv)	si $xz = yz$ y $z \neq 0$, entonces $x = y$	cancelación
v)	$x \cdot 1 = x$	elemento idéntico
vi)	si $x \neq 0 \; \exists \; x^{-1} \in Q$ tal que $xx^{-1} = 1$	elementos inversos

DEMOSTRACION

Demostraremos a continuación i), ii) y vi)

i) Sean $x = \frac{a}{b}$, $y = \frac{c}{d}$ dos números racionales, con lo que $a, b, c, d, \in Z$ y $b, d \neq 0$

De I.4.7 se sigue que

$$xy = \frac{ac}{bd}$$

donde $ac \in Z$ y $bd \in Z$ por ii de I.3.8.

Para demostar que $xy \in Q$ se requiere además probar que $bd \neq 0$. Esto lo haremos a continuación empleando un método de demostración indirecta conocido como prueba por contradicción o por reducción al absurdo.

Supongamos que $bd = 0$. Entonces

$bd = b \cdot 0$ — por i) de I.3.10

$db = 0 \cdot b$ — por iii) de I.3.8

Como $b \neq 0$, de iv) de I.3.8 se sigue que

$d = 0$

lo cual contradice la condición $d \neq 0$ establecida por $y \in Q$. En consecuencia la hipótesis $bd = 0$ es falsa por lo que

$bd \neq 0$

como se quería.

ii) Sean $x = \frac{a}{b}$, $y = \frac{c}{d}$, $t = \frac{e}{f}$, tres números racionales

$$x(yt) = \frac{a}{b}\left(\frac{c}{d}\cdot\frac{e}{f}\right)$$

$$= \frac{a}{b}\left(\frac{ce}{df}\right) \qquad \text{por I.4.7}$$

$$= \frac{a(ce)}{b(df)} \qquad \text{por I.4.7}$$

$$= \frac{(ac)e}{(bd)f} \qquad \text{por ii) de I.3.8}$$

$$= \left(\frac{ac}{bd}\right)\frac{e}{f} \qquad \text{por I.4.7}$$

$$= \left(\frac{a}{b}\cdot\frac{c}{d}\right)\frac{e}{f} \qquad \text{por I.4.7}$$

$$x(yt) = (xy)t \qquad \text{como se quería.}$$

vi) Sea x un número racional diferente de cero; esto es

$x = \frac{a}{b}$ con $a, b \in Z$, $a, b \neq 0$

Como $a, b \in Z$ y $a \neq 0$, por I.4.2 $\exists\ \frac{b}{a} \in Q$.

Ahora

$$\frac{a}{b}\cdot\frac{b}{a} = \frac{ab}{ba} \qquad \text{por I.4.7}$$

$$\frac{a}{b}\cdot\frac{b}{a} = \frac{ab}{ab} \qquad \text{por iii) de I.3.8}$$

De v) de I.3.8, $\frac{a}{a} = 1$; por lo que

$$\frac{a}{b}\cdot\frac{b}{a} = 1$$

Entonces $x^{-1} = \frac{b}{a}$ y la prueba termina. □

Al número x^{-1} (que como se ve en la prueba depende de x) se le denomina el "inverso de x para la multiplicación". Cabe enfatizar aquí que todo número racional, con excepción del cero, tiene un inverso multiplicativo en Q.

La división en Q puede definirse ahora a partir de la multiplicación y de vi) de I.4.8, de la siguiente manera.

I.4.9 DEFINICION

Sean $\frac{a}{b}$, $\frac{c}{d}$ dos números racionales y $\frac{c}{d} \neq 0$

El número $\frac{a}{b} \div \frac{c}{d}$ se define como

$$\frac{a}{b} \div \frac{c}{d} = \frac{a}{b}\cdot\frac{d}{c}$$

Como consecuencia de i) de I.4.8, la división en Q satisface la si -

guiente propiedad

$\forall\, x, y \in Q,\ y \neq 0:\ x \div y \in Q.$

Por otra parte, puede demostrarse que los teoremas I.3.9 y I.3.10 también son válidos para los números racionales; esto es

I.4.10 TEOREMA

Para todo $x, y, z \in Q$:

$x(y + z) = xy + xz$

I.4.11 TEOREMA

Para todo $x, y \in Q$:

i) $x \cdot 0 = 0$

ii) $(-x)(y) = -(xy)$

iii) $(-x)(-y) = xy$

- Orden en Q

Todo número racional puede expresarse como una fracción con denominador positivo, ya que

$$\frac{a}{b} = \frac{(-1)a}{(-1)b} = \frac{-a}{-b}$$

Así, si $b > 0$ el denominador de $\frac{a}{b}$ es positivo; mientras que si $b < 0$, $\frac{a}{b}$ puede expresarse como $\frac{-a}{-b}$ donde $-b$ es positivo.

Con ayuda de este resultado podemos establecer la relación "menor que" en Q a partir de la relación "menor que" en Z, de la siguiente manera.

I.4.12 DEFINICION

Sean $\frac{a}{b}, \frac{c}{d} \in Q$, donde $b, d \in Z^+$;

i) $\frac{a}{b} < \frac{c}{d}$ si $ad < bc$

ii) $\frac{a}{b} > \frac{c}{d}$ si $\frac{c}{d} < \frac{a}{b}$

Como consecuencia de esta definición y de la Ley de Tricotomía en Z, la relación "menor que" en Q satisface también dicha ley, así como las siguientes propiedades

I.4.13 TEOREMA

Para todo $x, y, z \in Q$:

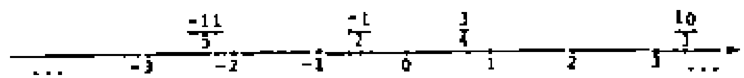
i) $x < y \Rightarrow x + z < y + z$

ii) $x < y$ y $z > 0 \Rightarrow xz < yz$
$x < y$ y $z < 0 \Rightarrow xz > yz$

iii) $x < y$ y $y < z \Rightarrow x < z$

En forma similar a como se definió en Z, diremos que un número $x \in Q$ es positivo si $x > 0$ y es negativo si $x < 0$.

Los elementos de Q pueden también ser representados en la recta numérica, donde si $x < y$, el punto que representa a x se encuentra a la izquierda del que representa a y



Los números racionales poseen una propiedad conocida como "densidad", según la cual entre dos números racionales diferentes siempre hay otro número racional; como lo establece el siguiente teorema.

I.4.14 TEOREMA

Para todo $x, y \in Q$, con $x < y$, $\exists\, z \in Q$ tal que:

$$x < z < y$$

DEMOSTRACION

Como $x < y$

$x + y < y + y$ por i) de I.4.13

$\frac{1}{2}(x + y) < \frac{1}{2}(y + y)$ por ii) de I.4.13

$\frac{1}{2}(x + y) < y$ - - - - (1)

Además, de $x < y$

$x + x < y + x$ por i) de I.4.13

$\frac{1}{2}(x + x) < \frac{1}{2}(y + x)$ por ii) de I.4.13

$\frac{1}{2}(x + x) < \frac{1}{2}(x + y)$ por iii) de I.4.5

$x < \frac{1}{2}(x + y)$ - - - - (2)

Por lo tanto, de (1) y (2) se sigue que $z = \frac{1}{2}(x + y)$ es tal que

$x < z < y$

Además, por la cerradura de Q para la adición y la multiplicación, se tiene que $\frac{1}{2}(x + y) \in Q$, con lo que concluye la demostración. □

Cabe hacer notar que los números naturales y los enteros no poseen la propiedad de densidad.

- **Expresión decimal de un número racional**

Hemos definido al número racional $\frac{a}{b}$ como el cociente de los números enteros a y b, siempre que $b \neq 0$. Si consideramos el número racional $\frac{7}{4}$, por ejemplo, y efectuamos la división de 7 entre 4 como se acostumbra en la aritmética, obtendremos

$$\frac{7}{4} = 1,75$$

A 1,75 se le conoce como la expresión decimal del número racional $\frac{7}{4}$.

Todo número racional tiene una expresión decimal, por ejemplo

$$\frac{3}{8} = 0,375$$

$$\frac{7}{33} = 0,212121\ldots$$

$$\frac{149}{132} = 1,12878787\ldots$$

Se dice que una expresión decimal es periódica cuando un dígito o grupo de dígitos se repite indefinidamente a partir de un cierto lugar a la derecha del punto decimal; ya sea desde el principio, como en 0,212121..., o empezando más adelante, como en 1,12878787...

De acuerdo con esto, una expresión decimal que aparentemente termina como 0,375, también es periódica ya que puede escribirse como 0,375000...

En general, con respecto a la expresión decimal de un número racional podemos establecer el siguiente enunciado.

I.4.15 TEOREMA

Todo número racional tiene una expresión decimal periódica.

Para demostrar I.4.15 requeriremos de otro teorema conocido como el "Algoritmo de la división para números enteros", que trataremos a continuación.

Como hemos visto, la ecuación $3x = 7$ no tiene solución en $\mathbb{Z}$, ya que no existe un número entero x que multiplicado por 3 dé como resultado 7.

A cambio, podemos encontrar los números enteros 2 y 1 tales que

$$3(2) + 1 = 7$$

En general, siempre es posible hallar esos dos números de acuerdo con lo que establece el siguiente teorema.

ALGORITMO DE LA DIVISION PARA NUMEROS ENTEROS†

Dados dos números enteros a y b, con $b > 0$, existen dos enteros únicos q y r, con $0 \leq r < b$, tales que

$$a = bq + r$$

Los números a, b, q y $r \in \mathbb{Z}$ reciben el nombre de dividendo, divisor, cociente y residuo respectivamente. Cabe hacer notar que aquí el término "cociente" tiene una interpretación más restringi-

† El estudiante puede consultar la demostración en la referencia 4, pág. 25.

da que la que empleamos al inicio de la sección I.4; ya que aquí el cociente es siempre un número entero.

La relación $a = bq + r$; que está planteada en términos de números enteros exclusivamente, puede ser enunciada en $\mathbb{Q}$ como

$$\frac{a}{b} = q + \frac{r}{b}$$

expresión que nos recuerda la forma como llevamos a cabo el proceso de dividir en la aritmética; esto es, obteniendo un cociente y un residuo.

Antes de pasar a la demostración de I.4.15 recordaremos el significado de la notación decimal y su relación con el algoritmo de la división para enteros a través de un ejemplo:

La expresión

$$1.1287878787...$$

representa la suma

$$1 + \frac{1}{10} + \frac{2}{10^2} + \frac{8}{10^3} + \frac{7}{10^4} + \frac{8}{10^5} + \frac{7}{10^6} + \frac{8}{10^7} + \frac{7}{10^8} + \cdots$$

Así, si buscamos la expresión decimal del número $\frac{149}{132}$ podemos proceder de la siguiente manera

o) Para obtener la parte entera vemos que

$$149 = 132(1) + 17, \text{ donde } 0 \leq 17 < 132$$

entonces

$$\frac{149}{132} = 1 + \frac{17}{132}$$

i) Para la primera cifra decimal vemos que

$$10 \times 17 = 170 = 132(1) + 38, \text{ donde } 0 \leq 38 < 132$$

por lo que $\frac{17}{132} = \frac{1}{10} + \frac{38}{10 \times 132}$

entonces

$$\frac{149}{132} = 1 + \frac{1}{10} + \frac{38}{10 \times 132}$$

ii) Para la segunda cifra decimal

$10 \times 38 = 380 = 132(2) + 116$, donde $0 \leq 116 < 132$

por lo que $\frac{38}{10 \times 132} = \frac{2}{10^2} + \frac{116}{10^2 \times 132}$

entonces

$$\frac{149}{132} = 1 + \frac{1}{10} + \frac{2}{10^2} + \frac{116}{10^2 \times 132}$$

iii) Para la tercera

$10 \times 116 = 1160 = 132(8) + 104$, donde $0 \leq 104 < 132$

por lo que $\frac{116}{10^2 \times 132} = \frac{8}{10^3} + \frac{104}{10^3 \times 132}$

entonces

$$\frac{149}{132} = 1 + \frac{1}{10} + \frac{2}{10^2} + \frac{8}{10^3} + \frac{104}{10^3 \times 132}$$

iv) Para la cuarta

$10 \times 104 = 1040 = 132(7) + 116$, donde $0 \leq 116 < 132$

por lo que $\frac{104}{10^3 \times 132} = \frac{7}{10^4} + \frac{116}{10^4 \times 132}$

entonces

$$\frac{149}{132} = 1 + \frac{1}{10} + \frac{2}{10^2} + \frac{8}{10^3} + \frac{7}{10^4} + \frac{116}{10^4 \times 132}$$

v) Para la quinta

$10 \times 116 = 1160 = 132(8) + 104$, donde $0 \leq 104 < 132$

por lo que $\frac{116}{10^4 \times 132} = \frac{8}{10^5} + \frac{104}{10^5 \times 132}$

entonces

$$\frac{149}{132} = 1 + \frac{1}{10} + \frac{2}{10^2} + \frac{8}{10^3} + \frac{7}{10^4} + \frac{8}{10^5} + \frac{104}{10^5 \times 132}$$

Podemos ver que la relación $a = bq + r$ establecida en el paso iii) se ha presentado nuevamente en el paso v); ésto se debe a que el residuo obtenido en ii) se ha repetido en el paso iv). Como consecuencia de ello, los cocientes obtenidos en los pasos iii) y v) se repetirán indefinidamente. Esto es

$$\frac{149}{132} = 1 + \frac{1}{10} + \frac{2}{10^2} + \frac{8}{10^3} + \frac{7}{10^4} + \frac{8}{10^5} + \frac{7}{10^6} + \frac{8}{10^7} + \frac{7}{10^8} + \cdots$$

o bien

$$\frac{149}{132} = 1.12878787\ldots$$

por lo que la expresión decimal de $\frac{149}{132}$ es periódica.

Consideremos ahora un número racional positivo $\frac{a}{b}$, con $a,b > 0$;

o) Por el algoritmo de la división para enteros, existen $q_0, r_0 \in Z$ tales que

$a = b\,q_0 + r_0$, donde $0 \leq r_0 < b$

entonces $\frac{a}{b} = q_0 + \frac{r_0}{b}$

i) Ahora, como $r_0 \in Z$ tenemos que $10r_0 \in Z$ y existen $q_1, r_1 \in Z$ tales que

$10r_0 = b\,q_1 + r_1$, donde $0 \leq r_1 < b$

por lo que $\frac{r_0}{b} = \frac{q_1}{10} + \frac{r_1}{10 \times b}$

entonces $\frac{a}{b} = q_0 + \frac{q_1}{10} + \frac{r_1}{10 \times b}$

ii) Ahora, como $r_1 \in Z$ tenemos que $10r_1 \in Z$ y existen $q_2, r_2 \in Z$ tales que

$10r_1 = b\,q_2 + r_2$, donde $0 \leq r_2 < b$

por lo que $\frac{r_1}{10 \times b} = \frac{q_2}{10^2} + \frac{r_2}{10^2 \times b}$

entonces $\frac{a}{b} = q_0 + \frac{q_1}{10} + \frac{q_2}{10^2} + \frac{r_2}{10^2 \times b}$

(iii) Ahora, como $r_2 \in \mathbb{Z}$ tenemos que $10r_2 \in \mathbb{Z}$ y existen $q_3, r_3 \in \mathbb{Z}$ tales que

$10r_2 = b\, q_3 + r_3$, donde $0 \leq r_3 < b$

por lo que $\frac{r_2}{10^2 \times b} = \frac{q_3}{10^3} + \frac{r_3}{10^3 \times b}$

entonces $\frac{a}{b} = q_0 + \frac{q_1}{10} + \frac{q_2}{10^2} + \frac{q_3}{10^3} + \frac{r_3}{10^3 \times b}$

En consecuencia, la expresión decimal de $\frac{a}{b}$ será

$\frac{a}{b} = q_0 . q_1\, q_2\, q_3 \ldots$

El proceso puede continuarse indefinidamente; sin embargo, como los residuos $r_0, r_1, r_2, r_3, \ldots$, son números enteros tales que $0 \leq r_i < b$, a lo más podrían existir b residuos diferentes. Cuando alguno de los residuos obtenidos se presenta por segunda vez se inicia el segundo ciclo del período, el cual se repite indefinidamente.

Si $\frac{a}{b}$ es negativo, entonces escribimos $\frac{a}{b} = -\frac{-a}{b}$ donde $\frac{-a}{b}$ es positivo y por tanto tiene una expresión decimal periódica.

Esto completa la demostración de I.4.15.

□

El recíproco del teorema I.4.15 también es válido; es decir:

I.4.16 TEOREMA

Toda expresión decimal periódica representa a un número racional.

La demostración de este teorema esta fundamentada en el concepto de límite, por lo que no la haremos aquí.† Sin embargo, presentamos a continuación un ejemplo que muestra la idea bajo la cual puede obtenerse un número racional como cociente de enteros a partir de su expresión decimal.

Consideremos la expresión decimal

1.40666...

Buscamos dos números enteros a,b tales que

$\frac{a}{b} = 1.40666\ldots$

Como tenemos dos dígitos antes de presentarse el período por primera vez, multiplicamos por 10^2 para obtener

$$10^2\, \frac{a}{b} = 140.666\ldots \qquad \text{(i)}$$

En vista de que el período consta de un dígito, multiplicamos la expresión (i) por 10^1 para obtener:

$$10^3\, \frac{a}{b} = 1406.666\ldots \qquad \text{(ii)}$$

Puesto que la parte decimal de estas dos últimas expresiones es la misma, restando (i) de (ii) obtenemos un número entero; esto es

$$10^3\, \frac{a}{b} - 10^2\, \frac{a}{b} = 1406 - 140$$

$$(10^3 - 10^2)\, \frac{a}{b} = 1266$$

$$900\, \frac{a}{b} = 1266$$

por lo que

$$\frac{a}{b} = \frac{1266}{900} = \frac{211}{150}$$

† El lector interesado puede consultarla en la referencia 3 pág. 64.

I.4.16 EJERCICIOS

1) Demostrar a partir de I.4.3 que si a, b, k son números enteros diferentes de cero, entonces:

   a) $\frac{a}{ka} = \frac{b}{kb}$

   b) $\frac{ka}{b} \neq \frac{a}{kb}$

2) Demostrar que para todo $x, y \in Q$:

   a) $x + y \in Q$
   b) $x + 0 = x$
   c) $x - y \in Q$

3) Demostrar que para todo $x, y \in Q$:

   a) $xy = yx$
   b) $x \cdot 1 = x$
   c) $x \cdot 0 = 0$

4) Demostrar que si $x, y \in Q$:

   $xy = 0 \Rightarrow x = 0$ ó $y = 0$

5) Sean $x, y \in Q$, demostrar que

   a) Si $y \neq 0$: $x \div y \in Q$.
   b) En general: $x \div (y \div z) \neq (x \div y) \div z$.
   c) Si $z \neq 0$: $(y \div z)x = (xy) \div z$

6) Sean $x, y, z \in Q$. Demostrar que:

   a) $x + z < y + z \Leftrightarrow x < y$

   b) si $z > 0$: $xz < yz \Leftrightarrow x < y$
      si $z < 0$: $xz < yz \Leftrightarrow x > y$

7) a) Empleando el algoritmo de la división para números enteros, hallar la expresión decimal de: $\frac{12}{11}, \frac{5}{6}, \frac{3}{13}$

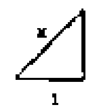
   b) Expresar cada uno de los siguientes decimales periódicos como el cociente de dos números enteros: 0.73333333..., 1.77272727..., 1.25925925...

I.5 LOS NUMEROS REALES

Consideremos ahora la ecuación

$$x^2 = 2 \qquad \text{- - - - [C]}$$

que se presenta al tratar de obtener la magnitud de la hipotenusa de un triángulo rectángulo cuyos catetos son de longitud unitaria:



$$x^2 = 1^2 + 1^2$$

Esta ecuación no tiene solución en Q, ya que no existe un número racional x tal que su cuadrado sea igual a 2, como demostraremos a continuación. A dicho número; esto es, al número cuyo cuadrado es igual a 2, se le representa mediante el símbolo $\sqrt{2}$.

Antes de demostrar que este número no es racional probaremos el siguiente resultado importante

I.5.1 LEMA

Si a^2 es un número par y $a \in Z$, entonces a es un número par

En efecto:

Si a es un número impar es de la forma

$a = 2n + 1$, donde $n \in Z$

entonces, su cuadrado

$$a^2 = (2n+1)(2n+1) = 4n^2 + 4n + 1 = 2(2n^2 + 2n) + 1$$

es de la forma

$$a^2 = 2m + 1, \text{ donde } m \in \mathbb{Z}$$

y es por tanto un número impar.

En consecuencia, si a^2 es par a no puede ser impar; por lo que, si $a \in \mathbb{Z}$ entonces a es un número par, lo que demuestra el lema.

Demostraremos ahora el siguiente teorema

I.5.2 TEOREMA

El número x tal que $x^2 = 2$ no es un número racional

DEMOSTRACION. (por reducción al absurdo)

Supongamos que x es un número racional; esto es, existen dos números enteros p y q tales que la mínima expresión de x es

$$x = \frac{p}{q}.$$

Como x es tal que $x^2 = 2$, entonces $(\frac{p}{q})^2 = 2$ por lo que

$$p^2 = 2q^2 \qquad ----(1)$$

Como $q \in \mathbb{Z}$, $q^2 \in \mathbb{N}$ y p^2 es un número par. En consecuencia, de I.5.1, p es un número par y es de la forma

$$p = 2n, \text{ donde } n \in \mathbb{Z} \qquad ----(2)$$

de (1) y (2) se sigue que

$$(2n)^2 = 2q^2$$

o sea

$$2n^2 = q^2$$

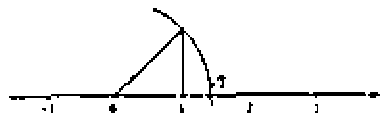
por lo que q^2 es un número par. En consecuencia, de I.5.1, q es un número par y es de la forma

$$q = 2m, \text{ donde } m \in \mathbb{Z} \qquad ----(3)$$

De (2) y (3), p y q tienen como factor común al número 2, por lo que $\frac{p}{q}$ no es la mínima expresión de x, lo cual contradice la hipótesis. Concluimos entonces que x no es un número racional, lo que demuestra el teorema. □

Como vimos en la sección anterior los números racionales tienen representación en la recta numérica, y en virtud de la densidad de $\mathbb{Q}$ (Teorema I.4.14) podría pensarse que estos son suficientes para "llenar" la recta; es decir, que todos los puntos de la recta corresponden a algún número racional, lo cual es falso.

Desde la antigüedad los geómetras griegos se dieron cuenta que no todos los puntos de la recta corresponden a números racionales. Por ejemplo, la siguiente construcción geométrica permite localizar a $\sqrt{2}$ en la recta numérica.



Existen muchos otros números que, como $\sqrt{2}$, tienen representación en la recta numérica y no son racionales. A este tipo de números se les conoce como números irracionales.

Como consecuencia de los teoremas I.4.15 y I.4.16, los números irracionales tienen expresión decimal no periódica, característica que los distingue de los números racionales.

Los números irracionales pueden ser de dos tipos: los que son solución de alguna ecuación algebraica con coeficientes enteros (como $\sqrt{2}$) a los que se llama irracionales algebraicos, y los que no son solución de una ecuación de tal tipo, a los que se llama irracionales trascendentes. Como ejemplo de estos últimos tenemos los números π y e de relevante importancia en las matemáticas.

En la sección I.3 se construyeron los números enteros a partir de los naturales, y en la I.4 los racionales a partir de los enteros, de tal forma que

$$N \subset Z \subset Q$$

La construcción de los números irracionales a partir de los racionales cae fuera del alcance de este libro, por lo que no se hará aquí.[†]

Al conjunto que contiene tanto a los números racionales como a los irracionales se le conoce como el conjunto de los números reales y se le representa con R.

Este conjunto viene a satisfacer la necesidad de un conjunto de nú-

[†] El estudiante puede consultarla en cualquiera de las referencias 1 a 3.

meros que represente a todos los puntos de la recta; es decir, a cada número real corresponde un punto de la recta y viceversa.

Al conjunto de los números irracionales se le representa comúnmente con Q', por lo que podemos escribir

$$R = Q \cup Q'$$

y

$$\emptyset = Q \cap Q'$$

y se cumple además que

$$N \subset Z \subset Q \subset R.$$

- Propiedades de las operaciones en R

A partir de la construcción formal de los números irracionales es posible definir las operaciones de adición y multiplicación para los números reales, de tal forma que se conserven las propiedades que estas operaciones tienen en Q. En virtud de que no hemos desarrollado aquí dicha construcción omitiremos también las correspondientes definiciones de las operaciones, aceptando que tienen las propiedades que se enuncian a continuación.

I.5.3 TEOREMA

Para todo $x, y, z \in R$:

i)	$x + y \in R$ $xy \in R$	cerradura
ii)	$x + (y + z) = (x + y) + z$ $x(yz) = (xy)z$	asociatividad
iii)	$x + y = y + x$ $xy = yx$	conmutatividad
iv)	$x + 0 = x$ $x \cdot 1 = x$	elemento idéntico
v)	$\exists\ -x \in R$ tal que $x + (-x) = 0$ $\exists\ x^{-1} \in R$ tal que $xx^{-1} = 1$, si $x \neq 0$	elementos inversos
vi)	$x(y + z) = xy + xz$	distributividad

Como consecuencia de las propiedades establecidas en el teorema I.5.3 se pueden deducir todas las propiedades algebraicas del sistema de los números reales; en particular las que enunciaremos a continuación, cuya demostración se deja al lector como ejercicio.

I.5.4 TEOREMA

Para todo $x, y \in R$:

i) $x \cdot 0 = 0$

ii) $(-x)(y) = -(xy)$

iii) $(-x)(-y) = xy$

A partir de I.5.3 pueden también definirse las operaciones de sustracción y división en R como sigue

I.5.5 DEFINICION

Sean x, y dos números reales

i) El número $x - y$ se define como: $x - y = x + (-y)$

ii) Si $y \neq 0$ el número $x \div y$ se define como: $x \div y = xy^{-1}$

Cabe hacer notar que el resultado de efectuar la división de x entre y, que denotamos mediante $x \div y$, coincide con el concepto de "cociente" establecido en la definición I.4.1; ya que, de I.5.5

$$x \div y = xy^{-1}$$

Entonces

$$(x \div y)(y) = (xy^{-1})y = x(y^{-1}y) = x \cdot 1 = x.$$

por lo que $x \div y$ es el número que multiplicado por y nos da como resultado x; es decir $\frac{x}{y}$.

Debido a que esta última forma presenta mayores ventajas en el manejo de las expresiones algebraicas, en adelante usaremos el símbolo $\frac{x}{y}$ para representar también al número $x \div y$.

- Orden en R

A partir de la construcción formal de los irracionales se puede definir también la relación "menor que" en R, de tal forma que si x, y son números reales tales que $x < y$ entonces el punto que representa a x en la recta numérica se encuentra a la izquierda del que representa a y. Aceptaremos también que dicha relación conserva las propiedades que tiene en Q; es decir

I.5.6 TEOREMA

Si $x, y \in R$ entonces se verifica una y sólo una de las siguientes proposiciones:

i) $x < y$

ii) $x = y$

iii) $y < x$

I.5.7 TEOREMA

Para todo $x, y, z \in R$:

i) $x < y \Rightarrow x + z < y + z$

ii) $x < y$ y $z > 0 \Rightarrow xz < yz$
$x < y$ y $z < 0 \Rightarrow xz > yz$

iii) $x < y$ y $y < z \Rightarrow x < z$

También se definen en R la relación "mayor que" y los términos "positivo" y "negativo", de la misma manera que se definen en Q; es decir:

$x > y$ si $y < x$
x es positivo si $x > 0$
x es negativo si $x < 0$

La relación "menor que", sus propiedades y conceptos relacionados son de gran utilidad para describir y manejar intervalos de valores para variables reales. En particular, las propiedades enunciadas por el teorema I.5.7 nos permiten "despejar" una variable en una relación de desigualdad, cuando esto es posible, como veremos en los ejemplos que se presentan a continuación.

I.5.8 EJEMPLOS

a) Obtener el intervalo de valores de x para los que se satisface la siguiente desigualdad.

$$\frac{2x - 3}{x + 2} < \frac{1}{3} \quad \text{con } x \neq -2$$

Solución

Lo primero que se ocurre es multiplicar por 3 y por $x + 2$ ambos miembros de la desigualdad, pero como no conocemos el valor de x no sabemos si $x + 2$ es positivo o negativo, por lo que debemos considerar dos casos:

Caso 1) $x + 2 > 0$

multiplicando por $x + 2$ y por 3:

$$3(2x - 3) < x + 2$$

o sea:

$$6x - 9 < x + 2$$

sumando 9 y restando x en ambos miembros

$$5x < 11$$

multiplicando por $\frac{1}{5}$

$$x < \frac{11}{5}$$

Caso 2) $x + 2 < 0$

multiplicando por $x + 2$ y por 3:

$$3(2x - 3) > x + 2$$

o sea:

$$6x - 9 > x + 2$$

sumando 9 y restando x en ambos miembros

$$5x > 11$$

multiplicando por $\frac{1}{5}$

$$x > \frac{11}{5}$$

Como $x + 2 > 0$, $x > -2$

y se obtiene finalmente:

$$-2 < x < \frac{11}{5}$$

Como $x + 2 < 0$, $x < -2$

Por tanto, no existen valores de x que satisfagan simultáneamente:

$$x < -2$$

y

$$x > \frac{11}{5}$$

La solución de la desigualdad propuesta es el conjunto de valores de x tales que:

$$-2 < x < \frac{11}{5}$$

b) Obtener el conjunto de valores de x para los cuales se satisface la siguiente desigualdad:

$\frac{3}{x} < 5$ con $x \neq 0$

Caso 1) $x > 0$

$3 < 5x \therefore \frac{3}{5} < x$ con $x > 0$

luego $\frac{3}{5} < x < \infty$

Caso 2) $x < 0$

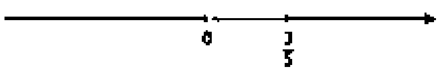
$\frac{3}{x} < 5 \therefore 3 > 5x$ y $\frac{3}{5} > x$ con $x < 0$

luego $-\infty < x < 0$

El conjunto buscado es:

$$\{x \mid x \in R \text{ y } \tfrac{3}{5} < x < \infty\} \cup \{x \mid x \in R \text{ y } -\infty < x < 0\}$$

O bien, gráficamente:



- Propiedad de completitud.

Los teoremas I.5.3 a I.5.7 nos muestran que el sistema de los números reales tiene las mismas propiedades algebraicas y de orden que el sistema de los números racionales; sin embargo, sabemos que el sistema de los números reales es más amplio y versátil puesto que en R podemos resolver ecuaciones como $x^2 = 2$, para las cuales no existe solución en Q. Cabe preguntarse entonces si los números reales tienen alguna propiedad adicional que no satisfagan los números racionales. Tal propiedad existe y se le conoce como "propiedad de completitud".

Antes de enunciar la propiedad de completitud es conveniente introducir algunos conceptos e ilustrarlos a través de ejemplos.

I.5.9 DEFINICION

Sea S un subconjunto de R. Un elemento $t \in R$ es una cota superior de S si:

$$x \leq t, \ \forall x \in S$$

De la definición anterior se deduce que si t es una cota superior de S, cualquier otro número real mayor que t será también una cota superior de S.

Si un conjunto tiene cotas superiores se dice que está acotado superiormente.

I.5.10 DEFINICION

Sea S un subconjunto de R. Un elemento $m \in R$ se llama elemento máximo de S si:

i) $x \leq m,\ \forall x \in S$

y

ii) $m \in S$.

lo que denotamos mediante $\max S = m$.

Es decir; un elemento es máximo de un conjunto S si es una cota superior de S y además pertenece a S.

A diferencia de las cotas superiores, el elemento máximo de un conjunto, si existe, es único.

En efecto; sean m y m' dos elementos máximos de S. Como $m \in S$ y m' es máximo

$$m \leq m'$$

por otra parte, como $m' \in S$ y m es máximo

$$m' \leq m$$

en consecuencia

$$m' = m$$

como se quería.

Como puede demostrarse fácilmente, el elemento máximo de un conjunto es la menor de sus cotas superiores.

Es posible, sin embargo, hallar conjuntos acotados superiormente para los cuales no existe elemento máximo; para tales conjuntos se tiene un concepto que sustituye al de máximo en el sentido de "la menor de las cotas superiores". Tal concepto es el de supremo que se define a continuación.

I.5.11 DEFINICION

Sea S un subconjunto de R. Un elemento $p \in R$ se llama supremo de S si:

i) $x \leq p,\ \forall x \in S$

y

ii) $q \in R$ y $x \leq q,\ \forall x \in S \implies p \leq q$

lo que denotamos mediante $\sup S = p$.

Es decir; un elemento p es el supremo de un conjunto S si p es una cota superior de S y ningún número menor que p es cota superior de S.

En forma similar a como se hizo para el máximo se puede demostrar que el supremo de un conjunto, si existe, es único.

Cabe resaltar aquí que la única diferencia entre los conceptos de elemento máximo y supremo de un conjunto S, estriba en que el máximo debe ser un elemento del conjunto S mientras que el supremo puede ser un elemento de S o no serlo.

I.5.12 EJEMPLOS

a) Sea $A = \{x \mid x \in R,\ x \leq 1\} \subset R$

Este conjunto está acotado superiormente.

Su elemento máximo es el 1.

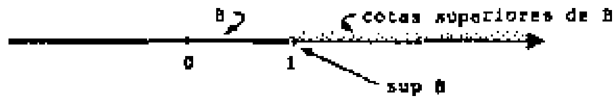
Su supremo es también el 1.

b) Sea $B = \{x \mid x \in R,\ x < 1\} \subset R$

Este conjunto está acotado superiormente.

No tiene elemento máximo

Su supremo es el 1.



I.5.13 TEOREMA (completitud de R)†

Todo subconjunto no vacío de R que está acotado superiormente tiene un supremo que pertenece a R.

† El lector puede consultar la demostración en la referencia 2 pág. 251.

Este teorema, cuya demostración omitiremos, nos garantiza la existencia de una mínima cota superior para cualquier conjunto de números reales acotado superiormente, y establece además que la mínima cota es un número real; es decir, que pertenece a R.

La propiedad de completitud permite establecer el concepto de continuidad en R, el cual es de importancia fundamental en el estudio del cálculo.

Para mostrar que los números racionales no poseen la propiedad de completitud establecida para los reales en el teorema I.5.13, consideremos el siguiente ejemplo que se inicia buscando una aproximación al valor de $\sqrt{2}$ mediante expresiones decimales con un número finito de cifras:

Recordemos que el símbolo $\sqrt{2}$ representa a un número x tal que $x^2 = 2$. Así, para una primera aproximación vemos que

$$(1)^2 = 1 \qquad y \qquad (2)^2 = 4$$

por lo que

$$1 < \sqrt{2} < 2$$

En consecuencia, si deseamos aproximarnos "por la izquierda" al valor de $\sqrt{2}$, la primera aproximación será 1.

Para obtener ahora una aproximación con una cifra decimal, observamos que

si x =	1.1	1.2	1.3	1.4	1.5
entonces x^2 =	1.21	1.44	1.69	1.96	2.25

por lo que

$$1.4 < \sqrt{2} < 1.5$$

y la siguiente aproximación será 1.4.

Para obtener una aproximación con dos cifras decimales observamos que

si $x =$	1.41	1.42
entonces $x^2 =$	1.9881	2.0164

por lo que

$$1.41 < \sqrt{2} < 1.42$$

y la aproximación será ahora 1.41.

Para tres y cuatro cifras decimales obtendríamos, respectivamente

si $x =$	1.411	1.412	1.413	1.414	1.415
entonces $x^2 =$	1.990921	1.993744	1.996569	1.999396	2.002225

y

si $x =$	1.4141	1.4142	1.4143
entonces $x^2 =$	1.99967881	1.99996164	2.00024449

y el proceso podría continuarse indefinidamente en virtud de que, por I.4.15 y I.4.16, la expresión decimal de $\sqrt{2}$ es no periódica.

Conviene hacer hincapié en que las aproximaciones obtenidas anteriormente, es decir

$$1,\ 1.4,\ 1.41,\ 1.414,\ 1.4142$$

son expresiones decimales con un número finito de cifras y, por tanto, son números racionales.

Consideremos ahora el conjunto S que consta de las aproximaciones de $\sqrt{2}$ que pueden obtenerse mediante el proceso que hemos descrito anteriormente; es decir:

$$S = \{1,\ 1.4,\ 1.41,\ 1.414,\ 1.4142,\ 1.41421,\ 1.414213,\ \ldots\}$$

Por construcción, el conjunto S tiene las siguientes características

1) Cada aproximación es mayor que la anterior
2) S contiene un número infinito de elementos
3) $S \neq \emptyset$, $S \subset Q$ y S está acotado superiormente

Las propiedades enunciadas en el punto 3) satisfacen las condiciones del teorema I.5.13 al reemplazar en él R por Q, sin embargo la conclusión no sería válida ya que S no tiene un supremo en Q. Es claro que $\sqrt{2}$ es la mínima cota superior de S y $\sqrt{2} \notin Q$.

Como consecuencia de lo anterior vemos que el conjunto de los números racionales carece de la propiedad de completitud.

- Valor absoluto de un número real.

Dado que existe una correspondencia uno a uno entre los puntos de la recta y los números reales, cabe esperar que el concepto geométrico de distancia entre dos puntos tenga su equivalente en el conjunto de los números reales. Dicho concepto, especialmente útil para el cálculo, puede ser introducido con ayuda del valor absoluto de un número real, que definiremos a continuación

I.5.14 DEFINICION

Sea x un número real. El valor absoluto de x, que representaremos con $|x|$, se define como

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

Así, por ejemplo

$\left|\frac{1}{2}\right| = \frac{1}{2}$, ya que $\frac{1}{2} > 0$

$|0| = 0$, ya que $0 = 0$

$|-7| = 7$, ya que $-7 < 0$

En general el valor absoluto de $x \in \mathbb{R}$ será un número real no negativo; es decir, positivo o cero. Las principales propiedades del valor absoluto se enuncian en el siguiente teorema

I.5.15 TEOREMA

Para todo $x, y \in \mathbb{R}$:

i) $|x| \geq 0$. Además $|x| = 0 \Leftrightarrow x = 0$

ii) $|xy| = |x| \cdot |y|$

iii) $|x + y| \leq |x| + |y|$

DEMOSTRACION

La propiedad i) es consecuencia inmediata de la definición I.5.14.

La propiedad ii) puede demostrarse fácilmente a partir del teorema I.5.4. La propiedad (iii), conocida como "desigualdad del triángulo", se demuestra a continuación.

1) Si $x = 0$ o $y = 0$ la igualdad es obvia.

2) Si $x > 0$ y $y > 0$ tenemos que $x + y > 0$, por lo que

$$|x + y| = x + y = |x| + |y|$$

y se satisface la igualdad.

3) Si $x < 0$ y $y < 0$ tenemos que $x + y < 0$, por lo que

$$|x + y| = -(x + y) = (-x) + (-y) = |x| + |y|$$

y se satisface la igualdad.

4) Nos queda por considerar el caso en que x, y tienen signos contrarios. Supongamos $x > 0$ y $y < 0$:

Si $y = -x$, $x + y = 0$ y la desigualdad es obvia.

Si $y < -x$, $x + y < 0$, por lo que

$$|x + y| = -(x + y) = (-x) + (-y) = -|x| + |y| < |x| + |y|$$

Si $y > -x$, $x + y > 0$, por lo que

$$|x + y| = x + y = x - (-y) = |x| - |y| < |x| + |y|$$

con lo que hemos demostrado iii) de I.5.15

A partir de la definición I.5.14 y de la representación de los números reales como puntos de la recta, se observa que mientras más grande es $|x|$ más lejos del origen se encuentra el punto que representa a x. Debido a esto, al número $|x|$ se le conoce como la distancia de x al cero.

De acuerdo con lo anterior, si a es un número real positivo se tendrá que un punto x está situado entre $-a$ y a cuando (y solamente cuando) $|x| < a$. Esta idea puede generalizarse a través del siguiente teorema.

I.5.16 TEOREMA

Sea $a \in \mathbb{R}$ con $a \geq 0$; $\forall\, x \in \mathbb{R}$ se tiene que:

$$|x| \leq a \iff -a \leq x \leq a$$

DEMOSTRACION

i) Sea $|x| \leq a$ --- (1)

De la definición I.5.14 se sigue que

$|x| = x$ ó $|x| = -x$, $\forall\, x \in \mathbb{R}$

Por lo que:

a) Si $|x| = x$, de (1)

$|x| \leq a \Rightarrow x \leq a$ --- (2)

b) Si $|x| = -x$, de (1)

$|x| \leq a \Rightarrow -x \leq a \Rightarrow x \geq -a$ --- (3)

por tanto, de (2) y (3) se sigue que

$|x| \leq a \Rightarrow -a \leq x \leq a$, $\forall\, x \in \mathbb{R}$.

ii) Supongamos ahora que $-a \leq x \leq a$ --- (4)

a) Si $x \geq 0$, $|x| = x$ y de (4)

$|x| \leq a$ --- (5)

b) Si $x < 0$, $|x| = -x$ y de (4)

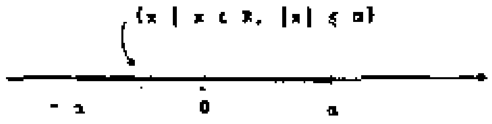
$-a \leq -|x| \Rightarrow a \geq |x|$ --- (6)

por lo que, de (5) y (6) se tiene que

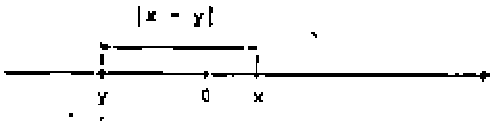
$-a \leq x \leq a \Rightarrow |x| \leq a$, $\forall\, x \in \mathbb{R}$.

y la demostración queda completa.

El teorema I.5.16 tiene la siguiente interpretación geométrica:



Con ayuda del valor absoluto, la distancia entre dos números reales cualesquiera x, y puede definirse como el número real no negativo $|x - y|$. A continuación se ilustra geométricamente el caso en que $x > 0$, $y < 0$:



I.5.17 EJERCICIOS

1) Demostrar a partir de I.5.3 que, si $x, y, z \in \mathbb{R}$:

a) $x + z = y + z \Rightarrow x = y$

b) $xz = yz \Rightarrow x = y$, si $z \neq 0$.

c) $x \cdot 0 = 0$

d) $(-x)(y) = -(xy)$

e) $xy = 0 \Rightarrow x = 0$ ó $y = 0$

2) Sean $x, y, z \in R$, demostrar que:

a) $x + y \in R$

b) $\frac{x}{y} \in R$, si $y \neq 0$

c) $x(y + z) = xy + xz$

d) $x\left(\frac{y}{z}\right) = \frac{xy}{z}$, si $z \neq 0$

3) Para cada una de las siguientes desigualdades, determinar el conjunto de valores de x que la satisfacen

a) $\frac{2 - x}{3} < 2(x - 2)$

b) $2 > \frac{x + 4}{x}$

c) $\frac{3x + 7}{x - 3} + 2 < 3$

4) De manera similar a como se hizo en I.5.9, el concepto de cota inferior se define como sigue:

Si $S \subset R$, un elemento $f \in R$ es una cota inferior de S si $x \geq f$, $\forall x \in S$. Además un conjunto que tiene cotas inferiores se dice que está acotado inferiormente.

Para cada uno de los conjuntos del ejercicio anterior decir si están acotados superiormente, inferiormente o de ambas maneras.

5) De manera similar a como se hizo en I.5.10 y I.5.11 se definen los conceptos de máximo e ínfimo de un conjunto.

Para cada uno de los conjuntos del ejercicio 3 hallar su máximo, mínimo, supremo e ínfimo, si existen.

6) Sean $x, y \in R$, demostrar que:

a) $|x| = 0 \iff x = 0$

b) $|xy| = |x| \cdot |y|$

c) $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$, si $y \neq 0$

d) $|x - y| = |y - x|$

7) Para cada una de las siguientes afirmaciones, demostrar su validez o dar un contraejemplo en caso de ser falsa:

a) $x \leq 3 \implies |x| \leq 3$

b) $|x| \leq 5 \implies -3 \leq x \leq 4$

c) $|x - 2| < 1 \iff 1 < x < 3$

d) $|x + 1| < 3 \iff -2 < x < 2$

I.6 MAS SOBRE INDUCCION MATEMATICA

Tan pronto como propusimos el primer teorema de la sección I.2 fue necesario recurrir a la inducción matemática para poder demostrarlo, ya que nos encontrábamos en los inicios de la construcción del sistema de los números reales y no disponíamos de más elementos para realizar la prueba que las propiedades que definen a los números naturales, es decir, los postulados de Peano. Es por ello que todas las propiedades de las operaciones en N debieron demostrarse por inducción matemática.

Al pasar a la sección I.3 definimos las operaciones en Z a partir de las operaciones en N, por lo que sus propiedades pudieron demostrarse a partir de las propiedades en N sin recurrir a la inducción matemática.

A partir de dicha sección, y debido a que nuestro interés fundamental era el de establecer las principales propiedades algebraicas de los diferentes sistemas de números, no fue necesario volver a utilizar la demostración por inducción. Sin embargo, existen otras proposiciones importantes en dichos sistemas, así como en otros que se tratarán más adelante, las cuales sólo pueden ser demostradas por inducción matemática.

Debido a su importancia como método de demostración, en esta sección presentaremos algunas observaciones y ejemplos adicionales que esperamos contribuyan a una mejor comprensión del método y que permitan su utilización en un contexto más amplio que el de la sección I.2.

Empezaremos por señalar que el nombre no es muy afortunado, ya que la palabra "inducción" suele asociarse al proceso que consiste en obtener una conclusión a partir del análisis de varios casos particulares, proceso que nada tiene que ver con el método de demostración que nos ocupa.

Mediante un proceso como el anteriormente descrito, al que se conoce como razonamiento inductivo, podemos llegar a proponer un enunciado de carácter general a partir del análisis de varios casos particulares (como se hizo con $P(n)$ en el ejemplo I.2.4 - a), pero nunca a demostrarlo.

A diferencia del razonamiento inductivo, el cual sólo puede asegurar la validez de la conclusión para los casos particulares de los cuales fue obtenido, el método de demostración por inducción matemática nos garantiza la validez de la conclusión para todos los casos.

Como hemos visto, una prueba por inducción matemática consta de dos partes cualitativamente diferentes:

I) Una verificación: $P(1)$ es cierta.

II) La demostración de una implicación: $P(k)$ es cierta $\Rightarrow$ $P(k+1)$ es cierta.

El papel que juegan estas dos partes en la prueba es similar al que se describe en la siguiente situación.

Supongamos que estamos frente a una escalera que tiene una infinidad

de peldaños y que deseamos tener la seguridad de que podemos llegar a cualquiera de sus peldaños. Esta seguridad nos la pueden proporcionar los dos hechos siguientes:

1) Podemos subir al primer peldaño.

2) Estando en un peldaño cualquiera podemos subir al siguiente.

En efecto: por 1) podemos situarnos en el primer peldaño. Estando en el primer peldaño, por 2) podemos situarnos en el segundo. Estando en el segundo peldaño, por 2) (nuevamente) podemos situarnos en el tercero, y así sucesivamente podemos llegar a cualquiera de ellos.

En una prueba por inducción la parte I nos garantiza que existe un valor para el cual la proposición es cierta, mientras que la parte II nos dice que si la proposición es cierta para un valor entonces será cierta para el siguiente valor (aunque esta parte no garantiza que exista un valor para el cual la proposición sea cierta).

Puesto que la primera parte de una prueba por inducción es una simple verificación, usualmente la dificultad de la prueba estriba en demostrar la implicación de la segunda parte, lo cual requiere en algunos casos de una buena dosis de ingenio; empero, ambas partes son indispensables para la prueba. A continuación presentamos un ejemplo en el que puede demostrarse la implicación y, sin embargo, la proposición es falsa.

Sea

$$2 + 4 + 6 + \ldots + 2n = n^2 + n + 1 \qquad --- P(n)$$

y supongamos que P(k) es cierta; esto es:

$$2 + 4 + 6 + \ldots + 2k = k^2 + k + 1$$

Sumando en ambos miembros $2(k + 1)$:

$$\begin{aligned} 2 + 4 + 6 + \ldots + 2k + 2(k + 1) &= (k^2 + k + 1) + 2(k + 1) \\ &= k^2 + k + 1 + 2k + 2 \\ &= (k^2 + 2k + 1) + (k + 1) + 1 \end{aligned}$$

$$2 + 4 + 6 + \ldots + 2k + 2(k + 1) = (k + 1)^2 + (k + 1) + 1$$

con lo que hemos demostrado que si P(k) es cierta entonces P(k + 1) es cierta; sin embargo, la prueba por inducción no puede completarse ya que no es posible hallar un valor para el cual la proposición sea cierta (demuéstrelo).

Hay proposiciones que aunque no se cumplen para todos los números naturales son válidas a partir de un cierto valor. Tal es el caso de la siguiente proposición

si $x \in \mathbb{R}$ y $x > 0$, entonces $(x + 1)^n > x^n + 1$ ---- P(n)

puesto que para $n = 1$

$$(x + 1)^1 = x^1 + 1$$

y la proposición no se cumple.

Para $n = 2$ tenemos que

$$(x + 1)^2 > x^2 + 1$$

ya que

$$(x + 1)^2 = x^2 + 2x + 1$$

y, como $x > 0$, $2x > 0$; por lo que

$$x^2 + 2x + 1 > x^2 + 1$$

de donde

$$(x + 1)^2 > x^2 + 1$$

y la proposición es válida para $n = 2$.

Del análisis de P(n) puede verse que a medida que n aumenta $(x + 1)^n$ crece más rápidamente que $x^n + 1$ por lo que, como la proposición fue válida para $n = 2$, cabe esperar que seguirá siendo válida para valores de n mayores que dos.

En efecto, demostraremos a continuación que la validez de P(k) implica la validez de P(k + 1). Hipótesis: P(k) es cierta, luego

$$(x + 1)^k > x^k + 1$$

Como $x > 0$, multiplicando por $x + 1$ tenemos

$$(x + 1)^k (x + 1) > (x^k + 1)(x + 1)$$

Esto es

$$(x + 1)^{k+1} > x^{k+1} + x^k + x + 1$$

Ahora, como $x > 0$ se tendrá que $x^k + x > 0$, por lo que

$$x^{k+1} + x^k + x + 1 > x^{k+1} + 1$$

En consecuencia

$$(x + 1)^{k+1} > x^{k+1} + 1$$

por lo que P(k + 1) es cierta.

Finalmente, como la proposición es cierta para $n = 2$, de lo anterior se sigue que también es cierta para cualquier valor mayor que dos.

Como vimos en el ejemplo anterior, es posible utilizar la inducción matemática para demostrar la validez de una proposición a partir de un valor fijo n_0. En estos casos la parte I de la prueba consistirá en verificar que la proposición es cierta para el valor n_0, la parte II consistirá nuevamente en demostrar que la validez de $P(k)$ implica la de $P(k + 1)$, teniendo en cuenta que ahora k es mayor o igual que n_0, y la conclusión será que la proposición es cierta para todo valor de n mayor o igual que n_0.

I.6.1 EJERCICIOS

1) Demostrar por inducción matemática que las siguientes fórmulas son válidas para todo número natural n:

   a) $1 + 2 + 3 + \ldots + n = \dfrac{n\,(n + 1)}{2}$

   b) $1 + 2 + 4 + \ldots + 2^n = 2^{n+1} - 1$, ¿qué representa aquí n?

   c) $a + aq + aq^2 + \ldots + aq^{n-1} = \dfrac{a(1 - q^n)}{1 - q}$, $\forall\ a, q \in R$.

2) Obsérvese que:

$$1 + \frac{1}{1} = 2$$

$$\left(1 + \frac{1}{1}\right)\left(1 + \frac{1}{2}\right) = 3$$

$$\left(1 + \frac{1}{1}\right)\left(1 + \frac{1}{2}\right)\left(1 + \frac{1}{3}\right) = 4$$

   proponer una ley general y demostrarla por inducción matemática.

3) Hallar la fórmula que simplifica el producto

$$\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{4}\right) \ldots \left(1 - \frac{1}{n+1}\right)$$

   y demostrarla por inducción.

4) Para cada una de las siguientes proposiciones, hallar el menor valor de n para el cual se cumple y demostrarla por inducción matemática:

   a) $\dfrac{n^3 - n}{3} \in N$.

   b) $n! > n^3$, donde $n! = 1 \times 2 \times 3 \times \ldots \times n$, $\forall\ n \in N$ y $0! = 1$

   c) $(x + 1)^n > nx^2 + nx + 1$, $\forall\ x \in R$, $x > 0$.

5) Demostrar por inducción matemática las siguientes propiedades del valor absoluto:

   a) $|x^n| = |x|^n$, $\forall\ n \in N$

   b) $|a_1\, a_2\, a_3 \ldots a_n| = |a_1|\,|a_2|\,|a_3| \ldots |a_n|$; $\forall\ n \in N$, $n \geq 2$

   c) $|a_1 + a_2 + a_3 + \ldots + a_n| \leq |a_1| + |a_2| + |a_3| + \ldots + |a_n|$; $\forall\ n \in N$, $n \geq 2$

6) Demostrar por inducción matemática las siguientes relaciones trigonométricas:

   a) $\operatorname{sen}(\theta + n\pi) = (-1)^n \operatorname{sen}\theta$; $n = 0, 1, 2, \ldots$

   b) $\operatorname{sen}\left[\theta + (2n - 1)\,\dfrac{\pi}{2}\right] = (-1)^{n-1} \cos\theta$; $\forall\ n \in N$

I.7 REFERENCIAS

1.- E. G. H. Landau
The Foundations of Analysis
Chelsea Publishing Co.
Nueva York, 1951.

2.- Rosa Beaumont y Richard Pierce
The Algebraic Foundations of Mathematics
Addison Wesley Publishing Co.
Massachusetts, 1963.

3.- Frank Ayres
Algebra Moderna
Serie Schaum, Mc Graw Hill
México, 1969

4.- Marie Weiss y Roy Dubisch
Algebra Superior
Limusa Wiley
México, 1967

NUMEROS COMPLEJOS

por

EDUARDO SOLAR GONZALEZ

LEDA SPEZIALE DE GUZMAN

profesores de la División de Ciencias Básicas
de la Facultad de Ingeniería de la U.N.A.M.

CAPITULO II NUMEROS COMPLEJOS

CONTENIDO

CAPITULO II NUMEROS COMPLEJOS

INTRODUCCION

Los números complejos aparecen en el horizonte de las matemáticas con la introducción de los números imaginarios. Estos surgieron en el álgebra como consecuencia de una necesidad, de la misma manera que los números negativos.

Girolamo Cardano, eminente matemático italiano del siglo XVI, fué el primero en reconocer la verdadera importancia de las raíces negativas al establecer la teoría general de las ecuaciones de tercero y cuarto grado. Se dió cuenta de la necesidad de números negativos y llegó a hablar incluso de raíces cuadradas de números negativos aunque, al parecer, no llegó a precisar el concepto de número imaginario.

Rafael Bombelli, también italiano, continuó la obra de Cardano y, en una obra publicada en 1572, señaló que los números imaginarios eran indispensables para la solución de ecuaciones de la forma $x^2 + c = 0$, donde c es un número positivo.

Los números imaginarios fueron atacados a partir de entonces, fueron declarados por muchos como "imposibles" o "inexistentes", por el mero hecho de no poder relacionarlos con experiencias de su vida cotidiana.

Pero una ecuación como $x^2 + 1 = 0$ no iba a quedarse sin solución. Las matemáticas requerían de los imaginarios para desarrollarse y, finalmente, éstos se impusieron.

Un número imaginario representa una idea abstracta pero muy precisa. La respuesta a la pregunta: ¿Qué número al ser multiplicado por sí mismo es igual a -1? sólo puede concebirse con la ayuda del imaginario más conocido, al que Euler representó con el símbolo i que todavía se emplea.

II.1 FORMA BINOMICA

Una vez aceptada la existencia de i como un número tal que $i \cdot i = -1$, un número imaginario queda definido como todo aquel de la forma bi, donde b es cualquier número real. Por ejemplo:

$$3i, \; -7i, \; \frac{1}{5}i, \; \sqrt{2}i$$

son números imaginarios.

Los imaginarios, con las mismas reglas de operación que los números reales, y considerando $i^2 = -1$, proporcionan soluciones a toda ecuación de la forma $x^2 + c = 0$, donde c es un número positivo.

Por ejemplo, las soluciones de la ecuación

$$x^2 + 25 = 0$$

son los números $5i$ y $-5i$ puesto que:

$$(5i)^2 = (5i)(5i) = 25i^2 = 25(-1) = -25$$

y

$$(-5i)^2 = (-5i)(-5i) = 25i^2 = 25(-1) = -25$$

En general, las soluciones de la ecuación

$$x^2 + c = 0, \quad \text{con} \quad c > 0,$$

son los números imaginarios $\sqrt{c}\, i$ y $-\sqrt{c}\, i$, donde $\sqrt{c} \in R$.

Por otra parte si tenemos una ecuación como

$$x^2 + 4x + 13 = 0$$

y aplicamos la fórmula general de la ecuación de segundo grado, obtenemos

$$x = -2 \pm \sqrt{-9} = -2 \pm 3i$$

y nos encontramos con que cada una de las soluciones es la "suma" de un número real con un número imaginario, lo cual debe interpretarse como un nuevo tipo de número. Surgen así los números complejos cuyo conjunto, que representamos con C, se define como sigue

II.1.1 DEFINICION

$$C = \{z \mid z = a + bi,\ a, b \in R,\ i^2 = -1\}$$

Así, por ejemplo, $-2 + 3i$, $-2 - 3i$, $1 + \sqrt{5}i$, $\frac{1}{2} - \pi i$ son números complejos.

Para manejar los números complejos necesitamos saber cuándo dos de

ellos son iguales, por lo que establecemos la siguiente definición

II.1.2 DEFINICION

Sean $z_1 = a + bi$, $z_2 = c + di$ dos números complejos con $a, b, c, d \in R$, entonces

$$z_1 = z_2 \quad \text{si} \quad a = c \quad \text{y} \quad b = d$$

Como puede verse la igualdad en C requiere de dos igualdades entre números reales. Así, si $a \neq c$ o $b \neq d$ se tendrá que $z_1 \neq z_2$.

- La adición y la multiplicación en C.

Las operaciones de adición y multiplicación de números complejos se definen, en términos de la adición y multiplicación de números reales, de la manera siguiente

II.1.3 DEFINICION

Sean $z_1 = a + bi$, $z_2 = c + di$ dos números complejos, donde $a, b, c, d \in R$.

i) el número $z_1 + z_2$ se define como

$$z_1 + z_2 = (a + c) + (b + d)i$$

ii) el número $z_1 z_2$ se define como

$$z_1 z_2 = (ac - bd) + (ad + bc)i$$

Estas operaciones tienen las siguientes propiedades

II.1.4 TEOREMA

Para todo $z_1, z_2, z_3 \in C$:

i)	$z_1 + z_2 \in C$	
	$z_1 \, z_2 \in C$	cerradura
ii)	$z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$	
	$z_1 \, (z_2 z_3) = (z_1 z_2) z_3$	asociatividad
iii)	$z_1 + z_2 = z_2 + z_1$	
	$z_1 \, z_2 = z_2 \, z_1$	conmutatividad
iv)	$z_1 + (0 + 0i) = z_1$	
	$z_1 \, (1 + 0i) = z_1$	elemento idéntico
v)	$\exists \, -z_1 \in C$ tal que $z_1 + (-z_1) = 0 + 0i$	
	si $z_1 \neq 0 + 0i$ $\exists \, z_1^{-1} \in C$ tal que $z_1 \, z_1^{-1} = 1 + 0i$	elementos inversos
vi)	$z_1 \, (z_2 + z_3) = z_1 z_2 + z_1 z_3$	distributividad

DEMOSTRACION

Se demostrarán únicamente i), iii) y v)

i) Sean $z_1 = a + bi$, $z_2 = c + di$ dos números complejos

$z_1 + z_2 = (a + c) + (b + d)i$ por i) de II.1.3

como $a, b, c, d \in R$ de i) de I.5.3 $(a + c), (b + d) \in R$

por lo que, de II.1.1

$(a + c) + (b + d)i \in C$

en consecuencia

$z_1 + z_2 \in C$ como se quería

Por otra parte

$z_1 z_2 = (ac - bd) + (ad + bc)i$ Por ii) de II.1.3

Como $a, b, c, d \in R$, de i) de I.5.3 se tiene que

$(ac - bd), (ad + bc) \in R$

por lo tanto, de II.1.1

$(ac - bd) + (ad + bc)i \in C$

y en consecuencia

$z_1 z_2 \in C$ como se quería

iii) Sean $z_1 = a + bi$, $z_2 = c + di$ dos números complejos

$z_1 + z_2 = (a + bi) + (c + di)$

$= (a + c) + (b + d)i$ por i) de II.1.3

$= (c + a) + (d + b)i$ por iii) de I.5.3

$= (c + di) + (a + bi)$ por i) de II.1.3

$z_1 + z_2 = z_2 + z_1$

Además

$z_1 z_2 = (a + bi)(c + di)$

$= (ac - bd) + (ad + bc)i$ por ii) de II.1.3

$= (ca - db) + (cb + da)i$ por iii) de I.5.3

$= (c + di)(a + bi)$ por ii) de II.1.3

$z_1 z_2 = z_2 z_1$

v) Sea $z_1 = a + bi$ un número complejo, con $a, b \in R$

De v) de I.5.3 $\exists \, -a, -b \in R$, y en consecuencia

$(-a) + (-b)i \in C$ por II.1.1

Sumando este número a z_1 se obtiene

$(a+bi)+[(-a)+(-b)i] = [a+(-a)] + [b+(-b)]i$ por i) de II.1.3

$(a+bi)+[(-a)+(-b)i] = 0 + 0i$ por v) de I.5.3

con lo que $-z_1 = (-a) + (-b)i$

Sea ahora $z_1 = a + bi$ con $a, b \in \mathbb{R}$ y $a, b \neq 0$

$\therefore \exists \; \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \in \mathbb{R}$, y de II.1.1

$$\frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2} i \in C$$

Aplicando ii) de II.1.3

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

$$= \frac{a^2 + b^2}{a^2 + b^2} + \frac{ab - ab}{a^2 + b^2} i$$

$$(a + bi)\left(\frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2} i\right) = 1 + 0i$$

con lo que $z_1^{-1} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2} i$

y termina la demostración. □

De acuerdo con II.1.1 un número complejo es de la forma

$$z = a + bi, \text{ con } a, b \in \mathbb{R}$$

En particular, si $b = 0$ el número complejo queda como

$$z = a + 0i$$

que puede ser considerado como el número real a, ya que la correspondencia

$$a + 0i \longleftrightarrow a$$

se conserva a través de las operaciones de adición y multiplicación.

En efecto, sean

$$a + 0i \longleftrightarrow a$$

y $c + 0i \longleftrightarrow c$

$$(a + 0i) + (c + 0i) = (a + c) + 0i \longleftrightarrow a + c$$

y $(a + 0i)(c + 0i) = (ac) + 0i \longleftrightarrow ac$

De manera semejante, a todo número complejo de la forma $0 + bi$ lo podemos considerar como el número imaginario bi.

De acuerdo con lo anterior, tanto el conjunto de los números reales como el conjunto de los números imaginarios son subconjuntos de C.

Con la definición formal del conjunto de los números complejos y de las operaciones de adición y multiplicación en C, podemos ahora verificar que el número complejo $x_1 = -2 + 3i$ es una solución de la ecuación

$$x^2 + 4x + 13 = 0$$

Para ello, consideramos que $4 = 4 + 0i$ y $13 = 13 + 0i$, y sustituimos $x_1 = -2 + 3i$ en el miembro izquierdo de la ecuación, obteniendo

$$(-2 + 3i)^2 + (4 + 0i)(-2 + 3i) + (13 + 0i)$$

de ii) de II.1.3 tenemos

$$[(4 - 9) + (-6 - 6)i] + [(-8 + 0) + (12 + 0)i] + (13 + 0i) =$$

$$= (-5 - 12i) + (-8 + 12i) + (13 + 0i)$$

y de i) de II.1.3

$$(-5 - 8 + 13) + (-12 + 12 + 0)i = 0 + 0i$$

$$= 0 \quad \text{como queríamos}$$

□

Como veremos más adelante, un número complejo puede ser representado en varias formas. A la forma $a + bi$, que hemos estado manejando, se le conoce como forma binómica debido a su apariencia de binomio.

Cuando un número complejo z está expresado en forma binómica, es decir como

$$z = a + bi$$

a los números reales a y b se les conoce, respectivamente, como parte real y parte imaginaria de z. Algunas veces se emplea para esta idea la siguiente notación:

$R(z) = a$

$I(z) = b$

- El conjugado de un número complejo

II.1.5 DEFINICION

Sea $z = a + bi$ un número complejo

El conjugado de z, que representaremos con $\bar{z}$, se define como

$$\bar{z} = a - bi$$

El conjugado tiene las propiedades que se enuncian en el siguiente teorema

II.1.6 TEOREMA

Para todo $z_1, z_2 \in C$:

i) $\bar{\bar{z}}_1 = z_1$

ii) $z_1 = \bar{z}_1 \Leftrightarrow z_1 \in R$

iii) $z_1 + \bar{z}_1 \in R$

iv) $z_1 \bar{z}_1 \in R$

v) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$

vi) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$

DEMOSTRACION

Se demostrarán únicamente ii), iii) y vi)

ii) Sea $z_1 = a + bi$ un número complejo

Probaremos primero que $z_1 = \bar{z}_1 \implies z_1 \in R$

$z_1 = \bar{z}_1 \implies a + bi = a - bi$ por II.1.5

$b = -b$ por II.1.2

por lo que $b = 0$

$\therefore z_1 = a + 0i = a \in R$

Ahora probaremos que $z_1 \in R \implies z_1 = \bar{z}_1$

$z_1 = a + 0i$

$\bar{z}_1 = a - 0i$ por II.1.5

$\bar{z}_1 = a + 0i$, ya que el cero es igual a su inverso aditivo

$z_1 = \bar{z}_1$ por II.1.2

Por lo tanto queda demostrado que

$z_1 = \bar{z}_1 \Leftrightarrow z_1 \in R$

iii) Sea $z_1 = a + bi$ un número complejo

$\bar{z}_1 = a - bi$ por II.1.5

$z_1 + \bar{z}_1 = (a + a) + (b - b)i$ por i) de II.1.3

$= 2a + 0i$

$z_1 + \bar{z}_1 = 2a \in R$

vi) Sean $z_1 = a + bi$, $z_2 = c + di$ dos números complejos

$\overline{z_1 z_2} = \overline{(a + bi)(c + di)}$

$= \overline{(ac - bd) + (ad + bc)i}$ por ii) de II.1.3

$= (ac - bd) - (ad + bc)i$ por II.1.5

$\overline{z_1 z_2} = (ac - bd) + (-ad - bc)i$

$$\overline{z_1 z_2} = [ac - (-b)(-d)] + [a(-d) + (-b)c]i$$

$$= (a - bi)(c - di) \qquad \text{por ii) de II.1.1}$$

$$\overline{z_1 z_2} = \bar{z}_1 \, \bar{z}_2 \qquad \text{por II.1.5}$$

y la prueba termina.

□

- La sustracción y la división en C.

La sustracción y la división en C se definen a partir de la adición y la multiplicación en C, respectivamente, y de v) de II.1.4, de la siguiente manera

II.1.7 DEFINICION

Sean $z_1 = a + bi$, $z_2 = c + di$ dos números complejos, donde $a, b, c, d \in \mathbb{R}$

i) el número $z_1 - z_2$ se define como

$$z_1 - z_2 = z_1 + (-z_2)$$

ii) si $z_2 \neq 0 + 0i$ el número $\frac{z_1}{z_2}$ se define como

$$\frac{z_1}{z_2} = z_1 \, z_2^{-1}$$

De la definición anterior se pueden obtener las siguientes fórmulas de uso práctico

de v) de II.1.4 $\quad -(c + di) = (-c) + (-d)i$

de i) de II.1.7 $\quad (a + bi) - (c + di) = (a + bi) + [(-c) + (-d)i]$

y de i) de II.1.3 $\quad \boxed{(a + bi) - (c + di) = (a - c) + (b - d)i}$

que es la fórmula que se emplea para la sustracción.

De v) de II.1.4 $\quad (c + di)^{-1} = \frac{c}{c^2 + d^2} + \frac{-d}{c^2 + d^2} i$

de ii) de II.1.7 $\quad \frac{a + bi}{c + di} = (a + bi)\left(\frac{c}{c^2 + d^2} + \frac{-d}{c^2 + d^2} i\right)$

y de ii) de II.1.3 $\quad \boxed{\frac{a + bi}{c + di} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2} i}$

que es la fórmula que se emplea para la división.

Así, si $z_1 = 1 + 8i$ y $z_2 = 2 + i$

$$z_1 - z_2 = (1 - 2) + (8 - 1)i$$

$$z_1 - z_2 = -1 + 7i$$

$$\frac{z_1}{z_2} = \frac{2 + 8}{4 + 1} + \frac{16 - 1}{4 + 1} i$$

$$= \frac{10}{5} + \frac{15}{5} i$$

$$\frac{z_1}{z_2} = 2 + 3i$$

El resultado de la división puede también obtenerse multiplicando dividendo y divisor por el conjugado del divisor y considerando a los números complejos como binomios, como puede verse a continuación

$$\frac{z_1}{z_2} = \frac{1 + 8i}{2 + i}$$

$$= \frac{(1 + 8i)(2 - i)}{(2 + i)(2 - i)}$$

$$= \frac{2 - i + 16i - 8i^2}{2^2 - i^2}$$

$$= \frac{(2 + 8) + 15i}{4 + 1}$$

$$= \frac{10}{5} + \frac{15}{5} i$$

$\frac{z_1}{z_2} = 2 + 3i$, que coincide con el resultado obtenido antes.

En general, todas las operaciones con números complejos expresados en forma binómica pueden efectuarse considerándolos como si fueran binomios, tomando en cuenta que las potencias de i superiores a uno deben reducirse de acuerdo a lo siguiente:

$i^2 = -1$

$i^3 = i^2 i = (-1)i = -i$

$i^4 = i^3 i = (-i)i = 1$

$i^5 = i^4 i = (1)i = i$

$i^6 = i^5 i = ii = -1$

.
.
.

y en adelante los valores se repiten periódicamente.

II.1.8 EJEMPLOS

Sean $z_1 = -5 - 2i$, $z_2 = -1 + i$

a) $z_1 + z_2 = (-5 - 2i) + (-1 + i) = -5 - 1 - 2i + i = -6 - i$

b) $z_1 - z_2 = (-5 - 2i) - (-1 + i) = -5 - 2i + 1 - i = -4 - 3i$

c) $z_1 z_2 = (-5 - 2i)(-1 + i) = 5 + 5i + 2i - 2i^2 = 5 - 5i + 2i + 2$

$z_1 z_2 = 7 - 3i$

d) $\dfrac{z_1}{z_2} = \dfrac{-5 - 2i}{-1 + i} = \dfrac{(-5 - 2i)(-1 - i)}{(-1 + i)(-1 - i)} = \dfrac{5 + 2i + 5i + 2i^2}{(-1)^2 - (i)^2}$

$\dfrac{z_1}{z_2} = \dfrac{(5 - 2) + (2 + 5)i}{1 + 1} = \dfrac{3 + 7i}{2} = \dfrac{3}{2} + \dfrac{7}{2} i$

II.1.9 EJERCICIOS

1) Demostrar que para todo $z_1, z_2, z_3 \in C$

a) $z_1 + (0 + 0i) = z_1$

b) $z_1(1 + 0i) = z_1$

c) $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$

2) Demostrar que para todo $z_1, z_2 \in C$

a) $\bar{z}_1 = -z_1 \Leftrightarrow R(z) = 0$

b) $z_1 \bar{z}_1 \in R$

3) Demostrar que las operaciones con números complejos pueden efectuarse considerándolos como binomios; es decir, que considerándolos como binomios se obtienen los mismos resultados que con las expresiones II.1.3 y las fórmulas de uso práctico que se emplean para la sustracción y la división.

4) Obtener todos los valores $x, y \in R$ que satisfacen las siguientes igualdades

a) $\dfrac{x + yi}{x - yi} = x + yi$

b) $(x + yi)^2 = (x - yi)^2$ donde $z^2 = zz$

5) Si $z_1 = -i$, $z_2 = 3$, $z_3 = \sqrt{2} + i$, $z_4 = -2 + 3i$ obtener el resultado de las siguientes operaciones

a) $\dfrac{z_1}{z_3} - z_4$

b) $\dfrac{\bar{z}_1 z_2}{z_3 \bar{z}_4}$

c) $\dfrac{z_1 - z_4}{z_3} + \bar{z}_2 z_4$

6) Expresar el resultado de las siguientes operaciones en la forma $a + bi$

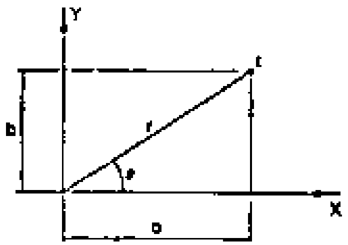
a) $\dfrac{i^7 + i^5 + i^4}{i^3 + i^3 + i^2}$

b) $\dfrac{\overline{(2 - i)} + (2 + i)(1 + i)}{(2 - i)(2 + i)}$

c) $\dfrac{(3 - 2i)(4 - i)^2 + (4 - i)(-1 + 5i)}{(4 - i)i}$

II.2 FORMA POLAR O TRIGONOMETRICA

A cada pareja ordenada de números reales (a,b) corresponde uno y sólo un número complejo a + bi y viceversa, por lo que podemos representar a dicho número complejo como un punto de coordenadas (a,b) en el plano cartesiano, donde su parte real a queda representada en el eje x, y su parte imaginaria b en el eje y.



A esta representación de los números complejos en el plano se le conoce como "Diagrama de Argand", y a los ejes x, y se les llama, respectivamente, "eje real" y "eje imaginario".

El punto de coordenadas (a,b) también está determinado por los parámetros (r,θ) de la figura, conocidos como coordenadas polares del punto.

Las coordenadas cartesianas (a,b) se obtienen a partir de las polares (r,θ) mediante las siguientes fórmulas de transformación

$$\left.\begin{array}{l} a = r\cos\theta \\ b = r\operatorname{sen}\theta \end{array}\right| \quad - - - - (A)$$

En consecuencia, el número complejo z = a + bi puede también expresarse como

$$z = (r\cos\theta) + (r\operatorname{sen}\theta)i$$

$$z = r(\cos\theta + i\operatorname{sen}\theta)$$

que es la llamada forma polar o trigonométrica del número complejo z. Podemos emplear una abreviatura para simplificar esta última expresión, ya que en ambas funciones trigonométricas se trata del mismo ángulo. Usaremos entonces la expresión "cis θ" para representar al factor cos θ + i sen θ, con lo que podemos escribir z = r cis θ. Formalizaremos lo anterior mediante la siguiente definición

II.2.1 DEFINICION

$$r\operatorname{cis}\theta = (r\cos\theta) + (r\operatorname{sen}\theta)i$$

En consecuencia, para expresar el número complejo z = a + bi en forma polar escribiremos

$$z = r\operatorname{cis}\theta$$

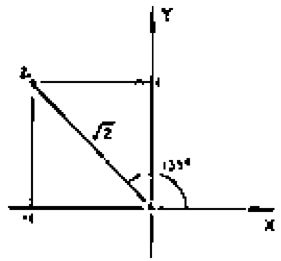
donde las coordenadas polares (r,θ) se obtienen a partir de las cartesianas (a,b) mediante las expresiones:

$$\left.\begin{array}{l} r = \sqrt{a^2 + b^2} \\ \theta = \text{ang tan}\,\dfrac{b}{a} \end{array}\right| \quad - - - - (B)$$

Así, por ejemplo, si tenemos el número complejo

$$z_1 = -1 + i$$

y queremos expresarlo en forma polar, aplicando las expresiones (B) obtenemos



$$r = \sqrt{(-1)^2 + (1)^2}$$

$$r = \sqrt{2}$$

$$\theta = \text{ang tan}\,\frac{1}{-1}$$

$$\theta = 135°$$

de donde $z_1 = \sqrt{2}\operatorname{cis} 135°$

Si tenemos ahora un número complejo en forma polar

$z_2 = 2 \text{ cis } 240°$

y queremos expresarlo en forma binómica, aplicando las expresiones (A) obtenemos

$a = 2 \cos \theta$

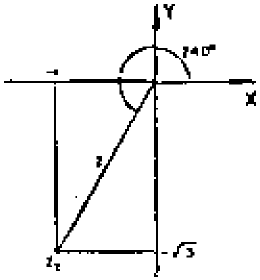
$= 2(-\frac{1}{2})$

$a = -1$

$b = 2 \text{ sen } \theta$

$= 2(-\frac{\sqrt{3}}{2})$

$b = -\sqrt{3}$



y finalmente $z_2 = -1 - \sqrt{3}\, i$

En la forma polar al número real r, que es una distancia y por tanto un número no negativo, se le conoce como el "módulo" del número complejo, y al ángulo θ como su "argumento". Para el caso particular del número 0 + 0i, su módulo es cero y su argumento se considera arbitrario.

Consideremos el caso de los números complejos $z_1 = 2 \text{ cis } 120°$ y $z_2 = 2 \text{ cis } 480°$.

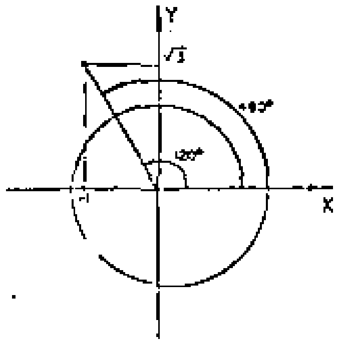
Al transformarlos a la forma binómica encontramos que

$z_1 = -1 + \sqrt{3}\, i$ y

$z_2 = -1 + \sqrt{3}\, i$

por lo que, de II.1.2

$z_1 = z_2$



En general, si tenemos dos números complejos expresados en forma polar con módulos iguales y argumentos que difieran un múltiplo entero de 360°, dichos números quedarán representados en el plano por el mismo punto; en consecuencia, en su forma binómica tendrán la misma parte real y la misma parte imaginaria, por lo que serán iguales, como lo establece el siguiente teorema.

II.2.2 TEOREMA

Sean $z_1 = r_1 \text{ cis } \theta_1$ y $z_2 = r_2 \text{ cis } \theta_2$;

$z_1 = z_2 \Leftrightarrow r_1 = r_2 \quad y \quad \theta_1 = \theta_2 + k\,(360°)$

con $k = 0, 1, 2, 3, \ldots$

DEMOSTRACION

i) Sean $z_1 = r_1 \text{ cis } \theta_1$

$z_2 = r_2 \text{ cis } \theta_2$

por II.2.1

$z_1 = (r_1 \cos \theta_1) + (r_1 \text{ sen } \theta_1) i$

$z_2 = (r_2 \cos \theta_2) + (r_2 \text{ sen } \theta_2) i$

Entonces, si $z_1 = z_2$ de II.1.2 se tiene que

$$\left.\begin{aligned} r_1 \cos \theta_1 &= r_2 \cos \theta_2 \\ r_1 \text{ sen } \theta_1 &= r_2 \text{ sen } \theta_2 \end{aligned}\right| \quad ---\to (1)$$

elevando al cuadrado

$r_1^2 \cos^2 \theta_1 = r_2^2 \cos^2 \theta_2$

$r_1^2 \text{ sen}^2 \theta_1 = r_2^2 \text{ sen}^2 \theta_2$

sumando miembro a miembro las igualdades y factorizando

$r_1^2 (\cos^2 \theta_1 + \text{sen}^2 \theta_1) = r_2^2 (\cos^2 \theta_2 + \text{sen}^2 \theta_2)$

por lo que

$r_1^2 = r_2^2$

como $r_1, r_2 \geq 0$ se sigue que

$r_1 = r_2$ ---- (2)

sustituyendo este resultado en (1) obtenemos

$\cos \theta_1 = \cos \theta_2$

y

$\text{sen}\, \theta_1 = \text{sen}\, \theta_2$

por lo que, de la trigonometría

$\theta_1 = \theta_2 + k(360°)$, con $k = 0, 1, 2, \ldots$ ---- (3)

En consecuencia, de (2) y (3) se tiene que

$z_1 = z_2 \Rightarrow r_1 = r_2$ y $\theta_1 = \theta_2 + k(360°)$, $k = 0, 1, 2, \ldots$

ii) Sean ahora $r_1 = r_2$ y $\theta_1 = \theta_2 + k(360°)$, con $k = 0, 1, 2, \ldots$

Entonces por la periodicidad de las funciones seno y coseno

$r_1 \cos \theta_1 = r_2 \cos \theta_2$

$r_1 \,\text{sen}\, \theta_1 = r_2 \,\text{sen}\, \theta_2$

y de II.1.2

$(r_1 \cos \theta_1) + (r_1 \,\text{sen}\, \theta_1)i = (r_2 \cos \theta_2) + (r_2 \,\text{sen}\, \theta_2)i$

esto es

$r_1 \text{ cis } \theta_1 = r_2 \text{ cis } \theta_2$

y finalmente

$z_1 = z_2$ con lo que termina la demostración. □

Como consecuencia de II.2.2 un número complejo puede tener más de un

argumento. Llamaremos "argumento principal" del número complejo, al ángulo θ tal que

$0° \leq \theta < 360°$

Así, como ejemplo tenemos que el argumento principal de

$z_1 = \sqrt{2} \text{ cis } 750°$ es $\theta_1 = 30°$

ya que $750° = 30° + 2(360°)$ y $0° \leq 30° < 360°$

y el argumento principal de $z_2 = 2 \text{ cis } 360°$ es $\theta_2 = 0°$

- La multiplicación y la división de números complejos en forma polar.

Una de las ventajas del manejo de números complejos en su forma polar, es la sencillez con que pueden efectuarse algunas operaciones, entre ellas la multiplicación y la división que se reducen a multiplicar módulos y sumar argumentos en el primer caso, y a dividir módulos y restar argumentos en el segundo. Así, por ejemplo, si

$z_1 = 6 \text{ cis } 120°$ y $z_2 = 2 \text{ cis } 40°$

tenemos que

$z_1 z_2 = (6)(2) \text{ cis } (120° + 40°) = 12 \text{ cis } 160°$

$\frac{z_1}{z_2} = \frac{6}{2} \text{ cis } (120° - 40°) = 3 \text{ cis } 80°.$

Formalizaremos lo anterior mediante el siguiente teorema.

II.2.3 TEOREMA

Sean $z_1 = r_1 \text{ cis } \theta_1$ y $z_2 = r_2 \text{ cis } \theta_2$, entonces:

i) $z_1 z_2 = r_1 r_2 \text{ cis } (\theta_1 + \theta_2)$

ii) $\frac{z_1}{z_2} = \frac{r_1}{r_2} \text{ cis } (\theta_1 - \theta_2)$

DEMOSTRACION

i) Sean $z_1 = r_1 \text{ cis } \theta_1$ y $z_2 = r_2 \text{ cis } \theta_2$

esto es

$$z_1 = (r_1 \cos \theta_1) + (r_1 \operatorname{sen} \theta_1)i$$

$$z_2 = (r_2 \cos \theta_2) + (r_2 \operatorname{sen} \theta_2)i$$

de i) de II.1.3

$$z_1 z_2 = \left[(r_1\cos\theta_1)(r_2\cos\theta_2) - (r_1\operatorname{sen}\theta_1)(r_2\operatorname{sen}\theta_2)\right] +$$

$$+ \left[(r_1\cos\theta_1)(r_2\operatorname{sen}\theta_2) + (r_1\operatorname{sen}\theta_1)(r_2\cos\theta_2)\right]i$$

$$z_1 z_2 = r_1 r_2 (\cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1\operatorname{sen}\theta_2) +$$

$$+ r_1 r_2 (\cos\theta_1\operatorname{sen}\theta_2 + \operatorname{sen}\theta_1\cos\theta_2)i$$

de las identidades trigonométricas

$$\cos(\theta_1 + \theta_2) = \cos\theta_1 \cos\theta_2 - \operatorname{sen}\theta_1 \operatorname{sen}\theta_2$$

$$\operatorname{sen}(\theta_1 + \theta_2) = \cos\theta_1 \operatorname{sen}\theta_2 + \operatorname{sen}\theta_1 \cos\theta_2$$

tenemos que

$$z_1 z_2 = r_1 r_2 \cos(\theta_1 + \theta_2) + r_1 r_2 \operatorname{sen}(\theta_1 + \theta_2)i$$

esto es

$z_1 z_2 = r_1 r_2 \text{ cis } (\theta_1 + \theta_2)$ como queríamos.

De manera semejante puede demostrarse la parte ii).

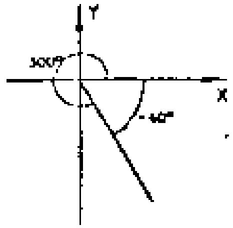
□

Al efectuar la división de números complejos en forma polar, puede suceder que el argumento del divisor sea mayor que el del dividendo y en ese caso se tendrá que el resultado es un número complejo con argumento negativo. Los argumentos negativos deben interpretarse como ángulos medidos en el otro sentido; esto es, en el que giran las manecillas del reloj.

Un argumento negativo está fuera del intervalo $0° \le \theta < 360°$, así,

para obtener el argumento principal correspondiente deberá sumarse 360° tantas veces como se requiera para quedar dentro de dicho intervalo.

Ejemplo



$$\frac{3 \text{ cis } 40°}{5 \text{ cis } 100°} = \frac{3}{5} \text{ cis } (-60°)$$

$$= \frac{3}{5} \text{ cis } (-60° + 360°)$$

$$\frac{3 \text{ cis } 40°}{5 \text{ cis } 100°} = \frac{3}{5} \text{ cis } 300°$$

- Potencias y raíces de números complejos

II.2.4 DEFINICION

Sean $z \in C$ y $n \in N$

La potencia enésima de z, que representaremos con z^n, se define como

$$z^n = \underbrace{zzz \ldots z}_{n \text{ factores}}$$

Con la definición anterior

$$z^2 = zz$$

si $z = r \text{ cis } \theta$

$$z^2 = rr \text{ cis } (\theta + \theta) = r^2 \text{ cis } (2\theta)$$

De manera semejante

$$z^3 = zzz$$

$$= z^2 z$$

$$z^3 = \left[r^2 \text{ cis } (2\theta)\right]\left[r \text{ cis } \theta\right]$$

$z^3 = r^2 r \text{ cis } (2\theta + \theta)$

$z^3 = r^3 \text{ cis } (3\theta)$

En general, cuando el número complejo está en forma polar podemos obtener sus potencias naturales con la llamada fórmula de De Moivre, que se expresa en el siguiente teorema y cuya demostración se deja al lector como ejercicio.

II.2.5 TEOREMA

Para todo número natural n:

$(r \text{ cis } \theta)^n = r^n \text{ cis } (n\theta)$

II.2.6 EJEMPLOS

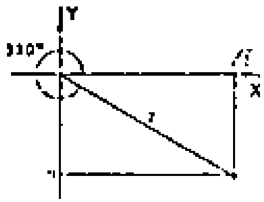
Sean $z_1 = \sqrt{2} \text{ cis } 70°$ y $z_2 = 3 \text{ cis } 225°$

a) $z_1^2 = (\sqrt{2})^2 \text{cis } (2 \cdot 70) = 2 \text{ cis } 140°$

b) $z_2^3 = (3)^3 \text{ cis } (3 \cdot 225) = 27 \text{ cis } 675 = 27 \text{ cis } 315°$

c) $z_2^4 = (3)^4 \text{ cis } (4 \cdot 225) = 81 \text{ cis } 900 = 81 \text{ cis } 180°$

Si un número complejo está en forma binómica, generalmente es más fácil obtener sus potencias transformándolo a la forma polar y aplicando el teorema II.2.5, como se ilustra a continuación



d) Obtener $(\sqrt{3} - i)^3$

$r = \sqrt{(\sqrt{3})^2 + (-1)^2} = \sqrt{3+1} = 2$

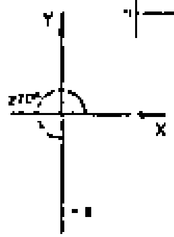
$\theta = \text{ang tan } \dfrac{-1}{\sqrt{3}} = 330°$

$\sqrt{3} - i = 2 \text{ cis } 330°$

$(\sqrt{3} - i)^3 = 8 \text{ cis } 990°$

$= 8 \text{ cis } 270°$

$(\sqrt{3} - i)^3 = -8i$



e) Obtener la cuarta potencia de $\left[\dfrac{z_2 + z_3}{z_4}\right] \bar{z}_1$, donde:

$z_1 = \sqrt{2} \text{ cis } 90°$, $z_2 = 3 + 3\sqrt{3}i$, $z_3 = 2 \text{ cis } 60°$ y $z_4 = -2$.

Para sumar los números z_2 y z_3 pasamos z_3 a forma binómica

$z_3 = 2 \text{ cis } 60° = 2 \cos 60° + (2 \text{ sen } 60°)i = 1 + \sqrt{3}i$

entonces

$z_3 + z_2 = (1 + \sqrt{3}i) + (3 + 3\sqrt{3}i) = 4 + 4\sqrt{3}i$

ahora

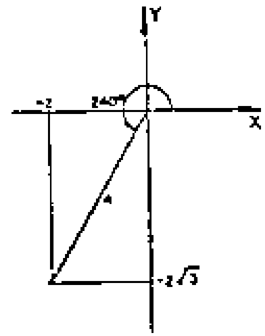
$\dfrac{z_2 + z_3}{z_4} = \dfrac{4 + 4\sqrt{3}i}{-2} = -2 - 2\sqrt{3}i$

para este número

$r = \sqrt{(-2)^2 + (-2\sqrt{3})^2} = \sqrt{4 + 12} = 4$

$\theta = \text{ang tan } \dfrac{-2\sqrt{3}}{-2} = 240°$

$\therefore \dfrac{z_2 + z_3}{z_4} = 4 \text{ cis } 240°$



por otra parte

$z_1 = \sqrt{2} \text{ cis } 90° = 0 + \sqrt{2}i$

por lo que $\bar{z}_1 = 0 - \sqrt{2}i = \sqrt{2} \text{cis } 270°$

así

$\left[\dfrac{z_2 + z_3}{z_4}\right] \bar{z}_1 = (4\text{cis}240°)(\sqrt{2}\text{cis}270°) = 4\sqrt{2}\text{cis}510° = 4\sqrt{2}\text{cis}150°$

y finalmente la cuarta potencia buscada es

$(4\sqrt{2}\text{cis}150°)^4 = (4\sqrt{2})^4 \text{ cis } (4 \cdot 150°) = 1024\text{cis}600° = 1024\text{cis}240°$

□

II.2.7 DEFINICION

Sean $z \in C$ y $n \in N$

Si $\omega^n = z$ decimos que ω es raíz enésima de z, y lo representamos mediante $\omega = \sqrt[n]{z}$.

Para obtener una expresión que nos permita calcular las raíces de un número complejo, consideremos los números

$z = r \text{ cis } \theta$ y $\omega = \rho \text{ cis } \alpha$

si ω es raíz enésima de z, entonces

$(\rho \text{ cis } \alpha)^n = r \text{ cis } \theta$

por lo que, de II.2.5

$\rho^n \text{ cis } n\alpha = r \text{ cis } \theta$

En consecuencia, de II.2.2

$\rho^n = r$

y

$n\alpha = \theta + k\,(360°)$, con $k = 0, 1, 2$

es decir

$$\rho = \sqrt[n]{r}$$

y

$$\alpha = \frac{\theta + k\,(360°)}{n}$$

en donde hemos representado con $\sqrt[n]{r}$ al número real no negativo cuya enésima potencia es igual a r.

Por lo que

$\omega = \sqrt[n]{r} \text{ cis } \frac{\theta + k\,(360°)}{n}$, con $k = 0, 1, 2, \ldots$

Veamos ahora qué valores toma el argumento de ω para los valores de $k = 0, 1, 2, \ldots$

para $k = 0$

$$\omega_0 = \sqrt[n]{r} \text{ cis } \frac{\theta}{n}$$

para $k = 1$

$$\omega_1 = \sqrt[n]{r} \text{ cis } \frac{\theta + 360°}{n} = \sqrt[n]{r} \text{ cis } \left[\frac{\theta}{n} + \frac{360°}{n}\right]$$

$\vdots$

para $k = n - 1$

$$\omega_{n-1} = \sqrt[n]{r} \text{ cis } \frac{\theta + (n-1)360°}{n} = \sqrt[n]{r} \text{ cis } \left[\frac{\theta}{n} + \frac{(n-1)360°}{n}\right]$$

para $k = n$

$$\omega_n = \sqrt[n]{r} \text{ cis } \frac{\theta + n(360°)}{n} = \sqrt[n]{r} \text{ cis } \left[\frac{\theta}{n} + 360°\right]$$

y de II.2.2 $\omega_n = \omega_0$

para $k = n + 1$

$$\omega_{n+1} = \sqrt[n]{r} \text{ cis } \frac{\theta + (n+1)360°}{n} = \sqrt[n]{r} \text{ cis } \left[\frac{\theta}{n} + \frac{360°}{n} + 360°\right]$$

y de II.2.2 $\omega_{n+1} = \omega_1$

De lo anterior podemos observar que la raíz enésima de un número complejo no es única y que existen exactamente n raíces diferentes correspondientes a los valores de $k = 0, 1, \ldots, (n - 1)$, ya que

$\omega_n = \omega_0$

$\omega_{n+1} = \omega_1$

$\omega_{n+2} = \omega_2$

$\vdots$

De esta manera hemos demostrado el siguiente teorema

II.2.8 TEOREMA

Para todo número natural n:

$$\sqrt[n]{r \text{ cis } \theta} = \sqrt[n]{r} \text{ cis } \frac{\theta + k\,(360°)}{n}$$

con $k = 0, 1, 2, \ldots, (n - 1)$

Estas n raíces quedan representadas en el Diagrama de Argand por n puntos sobre una circunferencia con centro en el origen y radio igual a $\sqrt[n]{r}$.

II.2.9 EJEMPLOS

a) Obtener las raíces cúbicas de $z = -4\sqrt{3} - 4i$ y representarlas en el Diagrama de Argand.

Primero pasamos $z = -4\sqrt{3} - 4i$ a su forma polar

$r = \sqrt{16 \times 3 + 16} = 8$

$\theta = \text{ang tan} \dfrac{-4}{-4\sqrt{3}} = 210°$

$z = 8 \text{ cis } 210°$

Ahora, del teorema II.2.8

$\sqrt[3]{z} = \sqrt[3]{8} \text{ cis } \dfrac{210° + k(360°)}{3}$

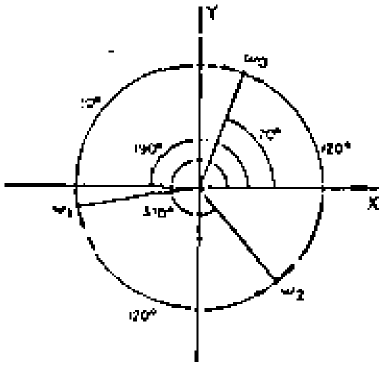
con $k = 0, 1, 2$

por lo que

para $k = 0$, $\omega_0 = 2 \text{ cis } 70°$

para $k = 1$, $\omega_1 = 2 \text{ cis } 190°$

para $k = 2$, $\omega_2 = 2 \text{ cis } 310°$



b) Obtener los valores de x tales que

$x^3 + 8 = 0$

Despejando a x tenemos

$x^3 = -8$, $x = \sqrt[3]{-8}$

la forma polar de -8 es $8 \text{ cis } 180°$, con lo que

$\sqrt[3]{-8} = \sqrt[3]{8 \text{ cis } 180°} = 2 \text{ cis } \dfrac{180 + k(360°)}{3}$, con $k = 0, 1, 2$.

para $k = 0$, $x_0 = 2 \text{ cis } 60° = 1 + \sqrt{3}i$

para $k = 1$, $x_1 = 2 \text{ cis } 180° = -2$

para $k = 2$, $x_2 = 2 \text{ cis } 300° = 1 - \sqrt{3}i$

Como podemos observar hay tres valores de x que satisfacen la ecuación dada, un número real (-2) y dos números complejos $(1 + \sqrt{3}i$ y $1 - \sqrt{3}i)$.

□

Hasta ahora hemos definido z^n para $n \in N$; sin embargo, a la raíz enésima de un número real x, suele representársele con $x^{1/n}$, lo que sugiere una extensión de la definición II.2.4 a casos en donde el exponente no es un número natural, lo que haremos de acuerdo a la siguiente definición

II.2.10 DEFINICION

Sean $z \in C$ y $m, n \in N$:

$z^0 = 1$

$z^{-n} = \dfrac{1}{z^n}$

$z^{\frac{m}{n}} = \sqrt[n]{z^m}$

Para ilustrar esta definición veamos los siguientes ejemplos

a) $z = (2 \text{ cis } 35°)^{-4} = \dfrac{1}{(2 \text{ cis } 35°)^4} = \dfrac{1 \text{ cis } 0°}{16 \text{ cis } 140°} = \dfrac{1}{16} \text{ cis } (-140°)$

$\therefore z = \dfrac{1}{16} \text{ cis } 220°$

b) Las soluciones de la ecuación $w^5 - (4 \text{ cis } 20°)^3 = 0$, son

$w = (4 \text{ cis } 20°)^{\frac{3}{5}} = \sqrt[5]{(4 \text{ cis } 20°)^3} = \sqrt[5]{64 \text{ cis } 60°}$

$\therefore w = \sqrt[5]{64} \text{ cis } \dfrac{60 + k(360°)}{5}$,

con $k = 0, 1, 2, 3, 4$

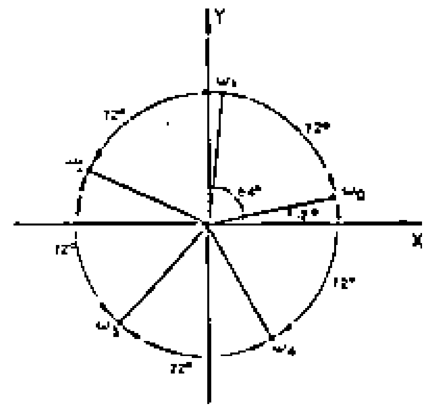
$\omega_0 = 2 \sqrt[5]{2} \text{ cis } 12°$

$\omega_1 = 2 \sqrt[5]{2} \text{ cis } 84°$

$\omega_2 = 2 \sqrt[5]{2} \text{ cis } 156°$

$\omega_3 = 2 \sqrt[5]{2} \text{ cis } 228°$

$\omega_4 = 2 \sqrt[5]{2} \text{ cis } 300°$



II.2.11 EJERCICIOS

1) Obtener la forma polar de los siguientes números complejos

$z_1 = 2 - 2i,\ z_2 = -3,\ z_3 = 5i,\ z_4 = -2i,\ z_5 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$

2) Obtener la forma binómica de los siguientes números complejos

$z_1 = \text{cis } 150°,\ z_2 = 4 \text{ cis } 210°,\ z_3 = 2\sqrt{2} \text{ cis } 315°$

3) Demostrar por inducción matemática que para todo $n \in N$

$(r \text{ cis } \theta)^n = r^n \text{ cis } (n\theta)$ (fórmula de De Moivre)

4) Efectuar las siguientes operaciones

a) $\dfrac{(1 - i) - (3 + i)}{2 \text{ cis } 120°}$

b) $(1 - i)^4 \cdot \dfrac{2 \text{ cis } 60°}{-\sqrt{3} + i}$

5) Obtener $z \in C$ tal que

$(z = 2\bar{z} + (\sqrt{3} - i)^4 \left(\frac{1}{8} \text{ cis } 30°\right)$

6) a) Determinar las soluciones de la ecuación

$z^4 + a^4 = 0$ para cualquier $a \in R$

b) Determinar los valores de $a \in R$ para los que $z_1 = 1 + i$ y $z_2 = 1 - i$ son soluciones de la ecuación y obtener las otras soluciones.

II.3 FORMA DE EULER O EXPONENCIAL

En el siglo XVIII, el matemático suizo Leonard Euler estableció la relación

$e^{\theta i} = \cos \theta + i \text{ sen } \theta \quad \forall\ \theta \in R$

que nos permite escribir el número complejo $z = r \text{ cis } \theta$ en la forma

$z = re^{\theta i}$

conocida como forma de Euler o forma exponencial; en la cual r es el módulo y θ el argumento expresado en radianes.

Por ejemplo, la forma de Euler de los números complejos $z_1 = 2 \text{ cis } 225°$, $z_2 = 3 \text{ cis } 180°$ y $z_3 = \sqrt{2} \text{ cis } 60°$ es:

$z_1 = 2e^{\frac{5\pi}{4}i}$, $z_2 = 3e^{\pi i}$ y $z_3 = \sqrt{2}\, e^{\frac{\pi}{3}i}$

Con base en el teorema II.2.2 podemos establecer la relación de igualdad entre números complejos expresados en forma de Euler, de acuerdo con el siguiente teorema

II.3.1 TEOREMA

Sean $z_1 = r_1 e^{\theta_1 i}$, $z_2 = r_2 e^{\theta_2 i}$;

$z_1 = z_2 \iff r_1 = r_2$ y $\theta_1 = \theta_2 + k\,(2\pi)$

con $k = 0, 1, 2, \ldots$

- Operaciones con números complejos en forma de Euler.

Los teoremas II.2.3, II.2.5 y II.2.8, establecidos para los números complejos en forma polar, tienen expresiones análogas para los números complejos expresados en forma de Euler, las cuales se presentan a continuación

$$(r_1e^{\theta_1 i})(r_2e^{\theta_2 i}) = (r_1r_2)e^{(\theta_1 + \theta_2)i}$$

$$\frac{r_1e^{\theta_1 i}}{r_2e^{\theta_2 i}} = \frac{r_1}{r_2}\, e^{(\theta_1 - \theta_2)i}$$

$$(re^{\theta i})^n = r^n\, e^{n\theta i}$$

$$\sqrt[n]{re^{\theta i}} = \sqrt[n]{r}\, e^{\frac{\theta + k(2\pi)}{n} i} \qquad \text{con } k = 0,\ 1,\ 2,\ \ldots,\ (n-1)$$

Estas fórmulas nos permiten efectuar operaciones de multiplicación y división directamente en la forma exponencial, así como obtener potencias y raíces de números complejos expresados en dicha forma.

II.3.2 EJEMPLOS

Dados $z_1 = \sqrt{3}\, e^{\frac{\pi}{2}i}$, $z_2 = e^{\pi i}$, $z_3 = 8e^{3i}$, $z_4 = 5e^{\frac{3}{2}\pi i}$

efectuar las siguientes operaciones:

a) z_1z_2 b) $\frac{z_1}{z_2}$ c) $(z_3)^{\frac{2}{3}}$ d) $\frac{z_1 + z_2}{z_4}$

Solución

a) $z_1z_2 = (\sqrt{3}e^{\frac{\pi}{2}i})(e^{\pi i}) = \sqrt{3}\, e^{\frac{3}{2}\pi i}$

b) $\frac{z_1}{z_2} = \frac{\sqrt{3}\, e^{\frac{\pi}{2}i}}{e^{\pi i}} = \sqrt{3}\, e^{-\frac{\pi}{2}i} = \sqrt{3}\, e^{\frac{3}{2}\pi i}$

ya que el argumento principal θ, en radianes, es tal que

$0 \leq \theta < 2\pi$

c) $(z_3)^{\frac{2}{3}} = \sqrt[3]{(8e^{3i})^2} = \sqrt[3]{64\, e^{6i}} = \sqrt[3]{64}\, e^{\frac{6 + k(2\pi)}{3}i}$

con $k = 0,\ 1,\ 2$

$$(z_3)^{\frac{2}{3}} = \sqrt[3]{64}\, e^{\frac{6 + k(2\pi)}{3}i} = \begin{cases} 4\, e^{2i} & \text{con } k = 0 \\ 4\, e^{(2 + \frac{2}{3}\pi)i} & \text{con } k = 1 \\ 4\, e^{(2 + \frac{4}{3}\pi)i} & \text{con } k = 2 \end{cases}$$

d) Para efectuar la adición indicada en el numerador, transformaremos los números z_1 y z_2 a su forma binómica y, posteriormente, la suma $z_1 + z_2$ a la forma de Euler.

$$z_1 = \sqrt{3}\, e^{\frac{\pi}{2}i} = (\sqrt{3} \cos \tfrac{\pi}{2}) + (\sqrt{3} \operatorname{sen} \tfrac{\pi}{2})i = 0 + \sqrt{3}\, i$$

$$z_2 = e^{\pi i} = (\cos \pi) + (\operatorname{sen} \pi)i = -1 + 0i$$

$$z_1 + z_2 = -1 + \sqrt{3}\, i = 2\, e^{\frac{2}{3}\pi i}$$

Ahora, efectuando la división obtenemos

$$\frac{z_1 + z_2}{z_4} = \frac{2\, e^{\frac{2}{3}\pi i}}{5\, e^{\frac{3}{2}\pi i}} = \frac{2}{5}\, e^{-\frac{5}{6}\pi i} = \frac{2}{5}\, e^{\frac{7}{6}\pi i} \qquad \square$$

- Logaritmo natural de un número complejo.

Una vez que hemos manejado la expresión $e^{\theta i}$ podemos aceptar la existencia de exponentes complejos. Esto nos permite generalizar el concepto de logaritmo para el caso de los números complejos, como sigue

> II.3.3 DEFINICION
>
> Sea $z \in C$
>
> El logaritmo natural de z, que representaremos con $L(z)$, se define como
>
> $$L(z) = w \quad \text{si} \quad e^{w} = z$$

A partir de esta definición deduciremos una fórmula para la obten - ción de logaritmos de números complejos

Sean $z = r\,e^{\theta i}$ y $w = a + bi$

si $w = L(z)$, entonces por II.3.3

$e^{w} = z$; es decir

$e^{a + bi} = r\,e^{\theta i}$

$e^{a}\,e^{bi} = r\,e^{\theta i}$

En consecuencia, de II.3.1

$e^{a} = r$

y

$b = \theta + k(2\pi)$, con $k = 0, 1, 2, \ldots$

Es decir

$a = Lr$

y

$b = \theta + 2k\pi$, con $k = 0, 1, 2, \ldots$

En donde Lr representa el logaritmo natural del número real no nega- tivo r.

Con lo anterior hemos demostrado el siguiente teorema

> II.3.4 TEOREMA
>
> Si $z = r\,e^{\theta i}$, entonces:
>
> $L(z) = Lr + (\theta + 2k\pi)i$, con $k = 0, 1, 2, \ldots$

Así, por ejemplo, el logaritmo natural del número complejo $z = 2\,e^{\frac{\pi}{3}i}$ es:

$L(z) = L(2) + (\frac{\pi}{3} + 2k\pi)i$, con $k = 0, 1, 2, \ldots$

Es decir

$L(z) = L(2) + \frac{6k + 1}{3}\pi i$, con $k = 0, 1, 2, \ldots$

Esta expresión representa a una infinidad de números complejos, uno para cada valor de k, esto es:

$L(2) + \frac{\pi}{3}i$, para $k = 0$

$L(2) + \frac{7}{3}\pi i$, para $k = 1$

$L(2) + \frac{13}{3}\pi i$, para $k = 2$

$\vdots$ $\vdots$

Cuando nos interesa un solo logaritmo, en general se considera el que corresponde al argumento principal del número; es decir, al que se obtiene con $k = 0$ cuando $0 \leq \theta < 2\pi$. A este logaritmo se le cono ce como "logaritmo principal".

II.3.5 EJEMPLOS

Obtener el logaritmo principal de cada uno de los siguientes números:

a) cis 45° b) $-\sqrt{3} - i$ c) -4

Solución

a) $L(\text{cis } 45°) = L(e^{\frac{\pi}{4}i}) = L(1) + \frac{\pi}{4}i = 0 + \frac{\pi}{4}i \doteq 0.7854\,i$

b) $L(-\sqrt{3} - i) = L(2\,e^{\frac{7}{6}\pi i}) = L(2) + \frac{7}{6}\pi i \doteq 0.6932 + 3.6652\,i$

c) $L(-4) = L(4\,e^{\pi i}) = L(4) + \pi i \doteq 1.3863 + 3.1416\,i$

El argumento correspondiente a un número real negativo es π, por lo que su logaritmo natural no es un número real; es por esto que cuan- do se considera el logaritmo natural como una función real de varia - ble real no está definida para números negativos.

II.3.6 EJERCICIOS

1) Efectuar las siguientes operaciones

a) $1 - e^{\pi i}$ b) $\dfrac{1 - e^{\frac{\pi}{2}i}}{1 + e^{\frac{\pi}{2}i}}$ c) $1 + e^{4\pi i}$

2) Representar en el Diagrama de Argand las soluciones de la ecuación

$$\frac{4 - 4i}{z^{\frac{1}{4}}} = 2\, e^{\pi i}$$

3) Dados $z_1 = 1 + i$, $z_2 = \sqrt{2}\, e^{\frac{\pi}{4}i}$, $z_3 = e^{3\pi i}$, $z_4 = 8 \text{ cis } 30°$

obtener los números $z \in C$, que satisfacen a la ecuación

$$z_1 + z_2 = \frac{z_3 z_4}{z^{\frac{1}{2}}}$$

4) Obtener todos los valores de $x, y \in R$ tales que:

a) $x + yi = x e^{yi}$

b) $e^{x + yi} = -1$

5) Obtener todos los números $z \in C$ tales que:

a) $z = \sqrt[4]{(-9)(3\, e^{\frac{\pi}{2}i})}$

b) $z = L \left[\dfrac{(1 + i) + \sqrt{2} \text{ cis } 45°}{-\sqrt{2}} \right]$

ESPACIOS VECTORIALES

TRANSFORMACIONES LINEALES

SISTEMAS DE ECUACIONES
LINEALES Y MATRICES

Capítulos V, VI y IV, respectivamente, de los APUNTES DE ALGEBRA *Editados por la Sección de Matemáticas de la Facultad de Ingeniería de la* UNAM, *elaborados por:*

JOSE B. CALVO A.
MANUELA GARIN DE A.
FRANCISCO GONZALEZ Z.
RONALD LOPEZ G.
RAFAEL LOPEZ L.
JESUS MERCADO M.
GUILLERMO MONSIVAIS G.
JESUS E. NOLASCO M.
GUSTAVO PEREZ D.
ISAAC SCHERSON S.
GUILLERMO SILIS G.
EDUARDO SOLAR G.
LEDA SPEZIALE DE G.

ESPACIOS VECTORIALES

INTRODUCCION

En varias ramas de la matemática, se presentan conjuntos donde resulta importante considerar "combinaciones lineales" de sus elementos. El álgebra lineal, trata las propiedades comunes de todos aquellos sistemas algebraicos que constan de un conjunto más una noción de "combinación lineal" de sus elementos.

El estudiante ha tenido ya experiencia con vectores del espacio euclidiano tridimensional, y en particular con combinaciones lineales de tales vectores. En este capítulo estableceremos la definición de una estructura algebraica, llamada espacio vectorial, que en la experiencia ha demostrado ser la más útil abstracción de un tipo de sistemas como el espacio euclidiano tridimensional.

El espacio vectorial de las ternas.

Consideremos los vectores

$\bar{v}_1 = (3, -1, 2)$ y $\bar{v}_2 = (1, 2, 0)$

Como sabemos, la suma de $\bar{v}_1$ y $\bar{v}_2$ es el vector:

$\bar{v}_1 + \bar{v}_2 = (4, 1, 2)$

y el producto del escalar $3 \in R$ por $\bar{v}_1$ es el vector:

$3\,\bar{v}_1 = (9, -3, 6)$

En general, si

$\bar{a} = (a_1, a_2, a_3)$ y $\bar{b} = (b_1, b_2, b_3)$

son dos vectores del tipo de $\bar{v}_1$ y $\bar{v}_2$, es decir, ternas ordenadas de números reales, la suma se define como

$\bar{a} + \bar{b} = (a_1+b_1, a_2+b_2, a_3+b_3)$ . . . (1)

y el producto de un escalar α por un vector $\bar{a}$ se define como:

$\alpha\,\bar{a} = (\alpha a_1, \alpha a_2, \alpha a_3)$ . . . . (2)

Al conjunto formado por todas las ternas ordenadas de números reales, lo representaremos con R^3. Simbólicamente:

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

En R^3, las operaciones definidas por las expresiones (1) y (2), tienen las siguientes propiedades.

Para todas las ternas ordenadas de números reales $\bar{a}, \bar{b}, \bar{c}$ y todos los escalares $\alpha, \beta \in R$ se cumple que:

1) R^3 es cerrado respecto a la suma y al producto por un escalar.
2) $(\bar{a}+\bar{b})+\bar{c} = \bar{a}+(\bar{b}+\bar{c})$.
3) $\bar{a}+\bar{0}=\bar{0}+\bar{a}=\bar{a}$, donde $\bar{0}=(0,0,0)$.
4) $\bar{a}+(-\bar{a})=(-\bar{a})+\bar{a}=\bar{0}$.
5) $\bar{a}+\bar{b}=\bar{b}+\bar{a}$.
6) $\alpha(\bar{a}+\bar{b})=\alpha\bar{a}+\alpha\bar{b}$.
7) $(\alpha+\beta)\bar{a}=\alpha\bar{a}+\beta\bar{a}$.
8) $\alpha(\beta\bar{a})=(\alpha\beta)\bar{a}$.
9) $1\bar{a}=\bar{a}$.

Como puede verse, R^3 forma un grupo abeliano con respecto a la suma; sin embargo, posee algunas propiedades adicionales (6 a 9) con respecto al producto por un escalar.

Existen varios sistemas que tienen las mismas características que R^3, a este tipo de sistemas les llamaremos ESPACIOS VECTORIALES.

Ejemplo V. 1.

Un sistema que posee la misma estructura (de espacio vectorial) que R^3, es el conjunto de todos los polinomios de grado ≤ 3 con las operaciones de suma y producto por un escalar. Demostraremos a continuación que se cumplen todas las propiedades que enunciamos para R^3.

1) Sean $\rho_1 = a_3x^3 + a_2x^2 + a_1x + a_0$
y $\rho_2 = b_3x^3 + b_2x^2 + b_1x + b_0$

dos polinomios cualesquiera de grado ≤ 3 con coeficientes reales.

Entonces:

$$\rho_1 + \rho_2 = (a_3 + b_3)x^3 + (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0)$$

es otro polinomio de grado ≤ 3, y si $\alpha \in R$:

$$\alpha\rho_1 = (\alpha a_3)x^3 + (\alpha a_2)x^2 + (\alpha a_1)x + (\alpha a_0)$$

también es un polinomio de grado ≤ 3.

Por lo que el conjunto de todos los polinomios de grado ≤ 3, es cerrado respecto a las operaciones de suma y producto por un escalar.

2) Sean

$$\rho_1 = a_3x^3 + a_2x^2 + a_1x + a_0$$
$$\rho_2 = b_3x^3 + b_2x^2 + b_1x + b_0$$
$$\rho_3 = c_3x^3 + c_2x^2 + c_1x + c_0$$

entonces:

$$(\rho_1 + \rho_2) + \rho_3 = (a_3 + b_3)x^3 + (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0) + (c_3x^3 + c_2x^2 + c_1x + c_0)$$
$$(\rho_1 + \rho_2) + \rho_3 = (a_3 + b_3 + c_3)x^3 + (a_2 + b_2 + c_2)x^2 + (a_1 + b_1 + c_1)x + (a_0 + b_0 + c_0)$$
$$(\rho_1 + \rho_2) + \rho_3 = a_3x^3 + a_2x^2 + a_1x + a_0 + (b_3 + c_3)x^3 + (b_2 + c_2)x^2 + (b_1 + c_1)x + (b_0 + c_0)$$
$$(\rho_1 + \rho_2) + \rho_3 = \rho_1 + (\rho_2 + \rho_3)$$

por lo que la suma es asociativa.

3) Si $\bar{0} = 0x^3 + 0x^2 + 0x + 0$

entonces: $\rho_1 + \bar{0} = \bar{0} + \rho_1 = \rho_1$

por lo que el polinomio $\bar{0}$ es el idéntico para la suma.

4) Sea $\rho_1 = a_3x^3 + a_2x^2 + a_1x + a_0$

existe el polinomio

$$-\rho_1 = -a_3x^3 - a_2x^2 - a_1x - a_0$$

tal que

$$\rho_1 + (-\rho_1) = (a_3 - a_3)x^3 + (a_2 - a_2)x^2 + (a_1 - a_1)x + (a_0 - a_0) = \bar{0}$$

y además

$$-\rho_1 + \rho_1 = \bar{0}$$

por lo que $-\rho_1$ es el inverso aditivo de ρ_1.

5) $\rho_1 + \rho_2 = (a_3 + b_3)x^3 + (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0)$

$\rho_1 + \rho_2 = (b_3 + a_3)x^3 + (b_2 + a_2)x^2 + (b_1 + a_1)x + (b_0 + a_0)$

$\rho_1 + \rho_2 = \rho_2 + \rho_1$

por tanto, la suma de polinomios es conmutativa.

6) Si $\alpha \in R$:

$\alpha(\rho_1 + \rho_2) = \alpha(a_3 + b_3)x^3 + \alpha(a_2 + b_2)x^2 + \alpha(a_1 + b_1)x + \alpha(a_0 + b_0)$

$\alpha(\rho_1 + \rho_2) = (\alpha a_3 + \alpha b_3)x^3 + (\alpha a_2 + \alpha b_2)x^2 + (\alpha a_1 + \alpha b_1)x + (\alpha a_0 + \alpha b_0)$

$\alpha(\rho_1 + \rho_2) = \alpha\rho_1 + \alpha\rho_2$.

El producto por un escalar es distributivo respecto a la suma de polinomios.

7) Si $\alpha, \beta \in R$:

$(\alpha + \beta)\rho_1 = (\alpha + \beta)a_3x^3 + (\alpha + \beta)a_2x^2 + (\alpha + \beta)a_1x + (\alpha + \beta)a_0$

$(\alpha + \beta)\rho_1 = \alpha a_3x^3 + \alpha a_2x^2 + \alpha a_1x + \alpha a_0 + \beta a_3x^3 + \beta a_2x^2 + \beta a_1x + \beta a_0$

$(\alpha + \beta)\rho_1 = \alpha\rho_1 + \beta\rho_1$.

El producto por un escalar es distributivo respecto a la suma de escalares.

8) $\alpha(\beta\rho_1) = \alpha(\beta a_3x^3 + \beta a_2x^2 + \beta a_1x + \beta a_0)$

$\alpha(\beta\rho_1) = (\alpha\beta)a_3x^3 + (\alpha\beta)a_2x^2 + (\alpha\beta)a_1x + (\alpha\beta)a_0$

$\alpha(\beta\rho_1) = (\alpha\beta)\rho_1$.

El producto por un escalar es asociativo.

9) $1p_1=1(a_3x^3+a_2x^2+a_1x+a_0)=a_3x^3+a_2x^2+a_1x+a_0=p_1$.

$1 \in R$ es el idéntico para el producto por un escalar.

En consecuencia, el conjunto de todos los polinomios de grado ≤ 3 también es un espacio vectorial para las operaciones de suma y producto por un escalar.

- -

Ejemplo V. 2.

Veamos si las matrices también forman un espacio vectorial.

Considérese el conjunto M de todas las matrices de orden $m \times n$ con elementos en C.

$M=\{(a_{ij}) | a_{ij} \in C \quad (i=1, 2, \ldots m, j=1, 2, \ldots n)\}$

De la definición de suma de matrices y de producto de un escalar por una matriz, se desprenden las siguientes propiedades para todas las matrices A, B, C de M y todos los escalares α, β de C:

1) M es cerrado con respecto a la suma y el producto por un escalar.

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

3) $A+0=0+A=A$, (Aquí 0 representa la matriz nula).

4) $A+(-A)=(-A)+A=0$.

5) $A+B=B+A$.

6) $\alpha(A+B)=\alpha A+\alpha B$.

7) $(\alpha+\beta)A=\alpha A+\beta A$.

8) $\alpha(\beta A)=(\alpha\beta)A$.

9) $1A=A$.

por lo que el conjunto de todas las matrices de orden $m \times n$ tiene la misma estructura de espacio vectorial que R^3.

- -

Hemos visto tres sistemas distintos que tienen la misma estructura, a los que hemos llamado espacios vectoriales. Los vectores de R^3 tienen una representación geométrica (segmentos dirigidos), mas no así los polinomios y las matrices, sin embargo, por el hecho de ser elementos de un espacio vectorial, también les llamaremos vectores.

V. 1. DEFINICION DE ESPACIO VECTORIAL.

Definición.

Un espacio vectorial V sobre un campo K, es un conjunto no vacío en el cual se definen las siguientes dos operaciones:

Una operación que asigna a cualesquiera elementos $\overline{u}$, $\overline{v}$ $\in$ V un elemento $\overline{u}+\overline{v}\in V$ al que se conoce como la suma de $\overline{u}$ y $\overline{v}$.

Una operación que asigna a cada elemento $\overline{v}\in V$ y a cada escalar $\alpha\in K$, un elemento $\alpha\overline{v}\in V$ al que se conoce como el producto de α por $\overline{v}$.

Las cuales cumplen las siguientes propiedades.

Para cualesquiera elementos $\overline{u}$, $\overline{v}$, $\overline{w}$ $\in$ V:

1) $(\overline{u}+\overline{v})+\overline{w}=\overline{u}+(\overline{v}+\overline{w})$.

2) $\exists\overline{0}\in V$ tal que $\overline{u}+\overline{0}=\overline{0}+\overline{u}=\overline{u}$,
(a $\overline{0}$ lo llamamos vector cero).

3) $\forall\ \overline{u}\in V$, $\exists\ -\overline{u}\in V$ tal que $\overline{u}+(-\overline{u})=-\overline{u}+\overline{u}=\overline{0}$.

4) $\overline{u}+\overline{v}=\overline{v}+\overline{u}$.

Para cualesquiera α, $\beta\in K$ y $\overline{u}$, $\overline{v}\in V$:

5) $\alpha(\overline{u}+\overline{v})=\alpha\overline{u}+\alpha\overline{v}$.

6) $(\alpha+\beta)\overline{u}=\alpha\overline{u}+\beta\overline{u}$.

7) $\alpha(\beta\overline{u})=(\alpha\beta)\overline{u}$.

8) $1\overline{u}=\overline{u}$,
(aquí 1 representa la unidad de K).

Consecuencias elementales de la definición.

Hay hechos simples que se desprenden casi inmediatamente de la definición de Espacio Vectorial, algunos de los cuales enunciamos a continuación en forma de teoremas.

Teorema V. 1

Si $\overline{0}$ es el vector cero de V, entonces:
$\forall\ \alpha\in K$ se tiene que $\alpha\overline{0}=\overline{0}$

Demostración.

Si α es un escalar cualquiera de K y $\overline{0}$ es el vector cero de V, entonces, por las propiedades 2) y 5) de la definición:

$$\alpha\overline{0}=\alpha(\overline{0}+\overline{0})=\alpha\overline{0}+\alpha\overline{0}$$

sumando $-(\alpha\overline{0})\in V$, por la propiedad 3) tenemos:

$$\alpha\overline{0}+(-(\alpha\overline{0}))=\alpha\overline{0}+\alpha\overline{0}+(-(\alpha\overline{0}))$$

$$\overline{0}=(\alpha\overline{0}+\alpha\overline{0})+(-(\alpha\overline{0}))$$

por 1) tenemos:

$$\overline{0}=\alpha\overline{0}+(\alpha\overline{0}+(-(\alpha\overline{0}))$$

de 3) se sigue que:

$$\overline{0}=\alpha\overline{0}+\overline{0}$$

finalmente, por 2):

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

Teorema V. 2.

Si 0 es el elemento cero de K, entonces:
$\forall\ \overline{v}\in V$ se tiene que $0\overline{v}=\overline{0}$

Demostración.

Si 0 es el escalar cero y $\overline{v}$ un vector cualquiera de V, entonces, como 0 es el idéntico aditivo en K:

$$0\overline{v}=(0+0)\overline{v}$$

por la propiedad 6) de la definición

$$0\overline{v}=0\overline{v}+0\overline{v}$$

sumando $-(0\overline{v})\in V$, por la propiedad 3) tenemos:

$$0\overline{v}+(-0\overline{v})=(0\overline{v}+0\overline{v})+(-(0\overline{v}))$$

$$\bar{0}=(0\bar{v}+0\bar{v})+(-(0\bar{v}))$$

por 1) tenemos

$$\bar{0}=0\bar{v}+(0\bar{v}+(-(0\bar{v}))$$

de 3) se sigue que

$$\bar{0}=0\bar{v}+\bar{0}$$

finalmente, por 2):

$$\bar{0}=0\bar{v}\ ,\quad \forall\ \bar{v}\in V\ .$$

Teorema V. 3.

Si 1 es la unidad de K, entonces: $\forall\ \bar{v}\in V$ se tiene que $(-1)\bar{v}=-\bar{v}$

Demostración.

Si $\bar{v}$ es un vector cualquiera de V, entonces:

$$\bar{0}=0\bar{v}=(1-1)\bar{v}=1\bar{v}+(-1)\bar{v}=\bar{v}+(-1)\bar{v}$$

en consecuencia, $(-1)\ \bar{v}$ es el inverso aditivo de $\bar{v}$ al que hemos representado con $-\bar{v}$. Esto demuestra el teorema.

En general, el vector $(-\alpha)\ \bar{v}$ es el inverso aditivo de $\alpha\bar{v}$, ya que:

$$(-\alpha)\bar{v}=(-1\alpha)\bar{v}=(-1)\alpha\bar{v}=-(\alpha\bar{v})$$

Por otra parte:

$$\alpha(-\bar{v})=\alpha(-1\bar{v})=(-\alpha)\bar{v}=-(\alpha\bar{v})$$

por lo que

$$\boxed{(-\alpha)\bar{v}=\alpha(-\bar{v})=-\alpha\bar{v}}$$

Así, la resta o sustracción de vectores se obtiene a partir de la suma de la siguiente manera:

Definición.

Si $\bar{u}$ y $\bar{v}$ son dos vectores de V:

$$\bar{u}-\bar{v}=\bar{u}+(-\bar{v})$$

Ejemplo V. 3.

El conjunto de todos los puntos de la recta que pasa por el origen y tiene como números directores 1, 2, 3, forma un espacio vectorial sobre el campo de los números reales.

La ecuación de dicha recta en forma simétrica es:

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

Por lo tanto, los puntos de dicha recta tendrán por coordenadas:

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

y el conjunto de puntos de la recta será:

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

cuyos elementos son ternas ordenadas de números reales. Entonces, la suma y el producto por un escalar están definidos de igual forma que en R^3.

Verifiquemos que S satisface todas las propiedades que definen un espacio vectorial.

Las operaciones deben ser cerradas:

Sean $\bar{v}_1=(x_1,\ 2x_1,3x_1)$ y $\bar{v}_2=(x_2,\ 2x_2,\ 3x_2)$ dos elementos cualesquiera de S y α cualquier escalar de R:

$$\bar{v}_1+\bar{v}_2=(x_1,\ 2x_1,\ 3x_1)+(x_2,\ 2x_2,\ 3x_2)=(x_1+x_2,2(x_1+x_2),3(x_1+x_2))\in S$$

y:

$$\alpha\bar{v}_1=\alpha(x_1,2x_1,3x_1)=(\alpha x_1,2\alpha x_1,3\alpha x_1)\in S.$$

Propiedades para la suma.

Sean $\bar{v}_1=(x_1,2x_1,3x_1)$, $\bar{v}_2=(x_2,2x_2,3x_2)$ y $\bar{v}_3=(x_3,2x_3,3x_3)$ tres elementos de S:

1) $(\bar{v}_1+\bar{v}_2)+\bar{v}_3=((x_1+x_2),2(x_1+x_2),3(x_1+x_2))+(x_3,2x_3,3x_3)$

$(\bar{v}_1+\bar{v}_2)+\bar{v}_3=((x_1+x_2+x_3),2(x_1+x_2+x_3),3(x_1+x_2+x_3))$

$(\bar{v}_1+\bar{v}_2)+\bar{v}_3=(x_1,2x_1,3x_1)+((x_2+x_3),2(x_2+x_3),3(x_2+x_3))$

$(\bar{v}_1+\bar{v}_2)+\bar{v}_3=\bar{v}_1+(\bar{v}_2+\bar{v}_3)$

por lo que la suma es asociativa en S.

2) Existe $\bar{0}=(0,0,0)\in S$ tal que:

$\bar{v}_1+\bar{0}=(x_1+0,2x_1+0,3x_1+0)=(0+x_1,0+2x_1,0+3x_1)$

$\bar{v}_1+\bar{0}=\bar{0}+\bar{v}_1=\bar{v}_1.$

3) Para todo $\bar{v}_1$, existe $-\bar{v}_1=(-x_1,-2x_1,-3x_1)\in S$ tal que:

$\bar{v}_1+(-\bar{v}_1)=(x_1,2x_1,3x_1)+(-x_1,-2x_1,-3x_1)=\bar{0}$

$-\bar{v}_1+\bar{v}_1=(-x_1,-2x_1,-3x_1)+(x_1,2x_1,3x_1)=\bar{0}$

4) $\bar{v}_1+\bar{v}_2=(x_1,2x_1,3x_1)+(x_2,2x_2,3x_2)$

$\bar{v}_1+\bar{v}_2=(x_1+x_2,\ 2x_1+2x_2,\ 3x_1+3x_2)$

$\bar{v}_1+\bar{v}_2=(x_2+x_1,\ 2x_2+2x_1,\ 3x_2+3x_1)$

$\bar{v}_1+\bar{v}_2=\bar{v}_2+\bar{v}_1$

por tanto la suma es conmutativa en S.

Propiedades para el producto por un escalar.

Sean $\bar{v}_1$, $\bar{v}_2$ elementos de S y α, β escalares de R:

5) $\alpha(\bar{v}_1+\bar{v}_2)=\alpha(x_1+x_2,2x_1+2x_2,3x_1+3x_2)$

$\alpha(\bar{v}_1+\bar{v}_2)=(\alpha x_1+\alpha x_2,2\alpha x_1+2\alpha x_2,3\alpha x_1+3\alpha x_2)$

$\alpha(\bar{v}_1+\bar{v}_2)=(\alpha x_1,2\alpha x_1,3\alpha x_1)+(\alpha x_2,2\alpha x_2,3\alpha x_2)$

$\alpha(\bar{v}_1+\bar{v}_2)=\alpha\bar{v}_1+\alpha\bar{v}_2.$

El producto por un escalar es distributivo sobre la suma de vectores.

6) $(\alpha+\beta)\bar{v}_1=((\alpha+\beta)x_1,2(\alpha+\beta)x_1,3(\alpha+\beta)x_1)$

$(\alpha+\beta)\bar{v}_1=(\alpha x_1+\beta x_1,2\alpha x_1+2\beta x_1,3\alpha x_1+3\beta x_1)$

$(\alpha+\beta)\bar{v}_1=(\alpha x_1,2\alpha x_1,3\alpha x_1)+(\beta x_1,2\beta x_1,3\beta x_1)$

$(\alpha+\beta)\bar{v}_1=\alpha\bar{v}_1+\beta\bar{v}_1.$

El producto por un escalar es distributivo sobre la suma de escalares.

7) $\alpha(\beta\bar{v}_1)=\alpha(\beta x_1,2\beta x_1,3\beta x_1)$

$\alpha(\beta\bar{v}_1)=((\alpha\beta)x_1,2(\alpha\beta)x_1,3(\alpha\beta)x_1)$

$\alpha(\beta\bar{v}_1)=(\alpha\beta)\bar{v}_1.$

El producto por un escalar es asociativo.

8) $1\bar{v}_1=1(x_1,2x_1,3x_1)=(x_1,2x_1,3x_1)=\bar{v}_1.$

La unidad de R es el idéntico para el producto por un escalar.

De todo lo anterior se sigue que, el conjunto

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

forma un espacio vectorial sobre el campo de los números reales.

- -

Antes de pasar a la sección siguiente, hay que hacer notar que los teoremas V.1, V.2, V.3 y la definición de resta o sustracción de vectores, fueron enunciados a partir de la definición de espacio vectorial, y se cumplen para los vectores de cualquier conjunto que tenga dicha estructura. Con esta misma idea, se enunciarán los teoremas y definiciones a lo largo de todo el capítulo.

V. 2. SUBESPACIOS DE UN ESPACIO VECTORIAL.

En el ejemplo V.3, analizamos el conjunto

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

que está formado por ternas ordenadas de números reales, por lo que S es un subconjunto de R^3. Además, se demostró que S también forma un espacio vectorial sobre R. Diremos entonces que S es un subespacio de R^3.

> Definición.
>
> Dado un espacio vectorial V y un subconjunto S de V, decimos que S es un subespacio de V, si S es también un espacio vectorial respecto a las operaciones definidas en V.

Ejemplo V.4

En el ejemplo V.1, demostramos que el conjunto de todos los polinomios de grado ≤ 3 es un espacio vectorial sobre R.

Si consideramos al conjunto de todos los polinomios de grado ≤ 2, vemos que éste es un subconjunto del primero y en forma análoga al ejemplo en cuestión, puede probarse que forma un espacio vectorial por sí mismo. En consecuencia, se trata de un subespacio.

El siguiente teorema establece un criterio sencillo para determinar si un subconjunto de un espacio vectorial es además, un subespacio.

> Teorema V.4
>
> Sea S un subconjunto de un espacio vectorial V. S es un subespacio de V si y sólo si es cerrado respecto a la suma y el producto por un escalar definidos en V.

Demostración.

Como S contiene elementos de V, se satisfacen las propiedades 1), 4), 5), 6), 7) y 8) de la definición. Por tanto, bastará con probar que se cumplen las propiedades 2) y 3).

En efecto:

Sea $\bar{v}$ un vector cualquiera de S y sea 0 el escalar cero de K. Como S es cerrado respecto al producto por un escalar y por el teorema V.2:

$0\bar{v}=\bar{0} \in S.$

Sea $\bar{v}$ un vector cualquiera de S y -1 el inverso aditivo de la unidad de K. Como S es cerrado respecto al producto por un escalar y por el teorema V. 3,

$(-1)\bar{v}=-\bar{v} \in S.$

Con lo que se completa la demostración.

Ejemplo V.5.

El conjunto de todas las matrices de la forma:

$$\begin{bmatrix} a & 2a \\ b & 0 \end{bmatrix} \quad (\text{donde } a, b \in C)$$

es un subespacio del espacio vectorial de matrices cuadradas de orden 2 sobre el campo de los números complejos.

En efecto:

Sea $A=\left\{\begin{bmatrix} a & 2a \\ b & 0 \end{bmatrix} \middle| \ a, b \in C\right\}$, que es un subconjunto de las matrices de orden 2.

Verifiquemos que A es cerrado respecto a la suma y el producto por un número complejo.

Para ello, considérense dos matrices cualesquiera de A:

$$\begin{bmatrix} a_1 & 2a_1 \\ b_1 & 0 \end{bmatrix} \quad y \quad \begin{bmatrix} a_2 & 2a_2 \\ b_2 & 0 \end{bmatrix}$$

y un número complejo α.

a) Para la suma:

$$\begin{bmatrix} a_1 & 2a_1 \\ b_1 & 0 \end{bmatrix} + \begin{bmatrix} a_2 & 2a_2 \\ b_2 & 0 \end{bmatrix} = \begin{bmatrix} a_1+a_2 & 2(a_1+a_2) \\ b_1+b_2 & 0 \end{bmatrix}$$

que es otra matriz de A.

b) Para el producto por un escalar.

$$\alpha \begin{bmatrix} a_1 & 2a_1 \\ b_1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha a_1 & 2\alpha a_1 \\ \alpha b_1 & 0 \end{bmatrix}$$

que también pertenece a A.

Entonces, según el teorema V. 4, A es un subespacio del espacio vectorial de las matrices cuadradas de orden 2.

- -

Ejemplo V. 6.

La intersección de dos subespacios es un nuevo subespacio.

En efecto:

Sean V un espacio vectorial sobre un campo K y A, B dos subespacios de V tales que:

$A \cap B \neq \emptyset$

Entonces, $\forall\ \bar{v}_1, \bar{v}_2 \in (A \cap B)$ se tiene que

$\bar{v}_1, \bar{v}_2 \in A$.	$\bar{v}_1, \bar{v}_2 \in B$.
Como A es un subespacio de V, es cerrado para la suma, por tanto:	Como B es un subespacio de V, es cerrado para la suma, por tanto:
$\bar{v}_1+\bar{v}_2 \in A$	$\bar{v}_1+\bar{v}_2 \in B$

por lo que:

$$\bar{v}_1+\bar{v}_2 \in (A \cap B)$$

En forma análoga, si $\alpha \in K$ y $\bar{v} \in (A \cap B)$, $\alpha\bar{v} \in A$ y $\alpha\bar{v} \in B$,

por lo que

$$\alpha\bar{v} \in (A \cap B)$$

Del teorema V. 4 se sigue que $A \cap B$ es un nuevo subespacio de V.

- -

Ejemplo V. 7.

Considérese el conjunto de todos los polinomios de grado igual a 2; aunque es un subconjunto del conjunto de polinomios de grado ≤ 3, no forma un subespacio de éste. Para demostrarlo bastará con probar que no satisface la cerradura respecto a la suma.

Sean:

$f(x)=x^2+bx+c \quad y \quad g(x)=-x^2+b'x+c'$

dos polinomios de segundo grado. Su suma será el polinomio:

$f(x)+g(x)=(b+b')x+c+c'$

que no es un polinomio de segundo grado.

Por lo tanto, el conjunto de los polinomios de segundo grado no forma un espacio vectorial.

- -

V. 3. DEPENDENCIA LINEAL.

Consideremos los siguientes vectores de R^3:

$\bar{a}=(3, -1, 2)$ $\bar{b}=(1, -1, -1)$

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

$\bar{c}=\bar{a}-2\bar{b}$ . . . . (1)

dicho vector será:

$\bar{c}=(1, 1, 4)$

Diremos entonces que el vector $\bar{c}$ se obtuvo a partir de una combinación lineal de los vectores $\bar{a}$ y $\bar{b}$.

En general, una combinación lineal de m vectores es la suma de dichos vectores multiplicados por ciertos escalares.

Definición.

Una combinación lineal de los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m$ de un espacio vectorial V sobre un campo K, es una expresión de la forma:

$$\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_m\bar{v}_m$$

donde $\alpha_1, \alpha_2, \ldots, \alpha_m \in K$.

Al conjunto de todas las combinaciones lineales de un conjunto S de vectores, lo representaremos con $L(S)$.

Teorema V. 5.

El conjunto de todas las combinaciones lineales de un conjunto

$S=\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m\}$

de vectores de un espacio vectorial V sobre un campo K, es un subespacio de V.

Demostración.

Sea $L(S)$ el conjunto de todas las combinaciones lineales del conjunto

$S=\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m\}$

y sean $\bar{a}$ y $\bar{b}$ dos vectores cualesquiera de $L(S)$:

$\bar{a}=\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_m\bar{v}_m$

$\bar{b}=\beta_1\bar{v}_1+\beta_2\bar{v}_2+\ldots+\beta_m\bar{v}_m$

donde $\alpha i, \beta i \in K$.

La suma de $\bar{a}$ y $\bar{b}$:

$\bar{a}+\bar{b}=(\alpha_1+\beta_1)\bar{v}_1+(\alpha_2+\beta_2)\bar{v}_2+\ldots+(\alpha_m+\beta_m)\bar{v}_m$

es otro elemento de $L(S)$.

Y el producto de $\lambda \in K$ por $\bar{a}$:

$\lambda\bar{a}=(\lambda\alpha_1)\bar{v}_1+(\lambda\alpha_2)\bar{v}_2+\ldots+(\lambda\alpha_m)\bar{v}_m$

es otro elemento de $L(S)$.

Entonces, por el teorema V.4 , $L(S)$ es un subespacio de V.

Volvamos ahora con los vectores $\bar{a}$, $\bar{b}$ y $\bar{c}$ mencionados. La expresión (1) puede escribirse como:

$\bar{a}-2\bar{b}-\bar{c}=\bar{0}$

Vemos entonces que el vector cero puede obtenerse a partir de una combinación lineal de los vectores $\bar{a}$, $\bar{b}$ y $\bar{c}$, con escalares NO TODOS NULOS. En este caso, diremos que los vectores $\bar{a}$, $\bar{b}$ y $\bar{c}$ son linealmente DEPENDIENTES.

Es claro que, dado un conjunto de vectores cualesquiera, el vector cero siempre puede obtenerse a partir de una combinación lineal de ellos con todos los escalares nulos;

Definición.

Un conjunto $S=\{\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n\}$ de vectores de un espacio vectorial V sobre un campo K, es linealmente dependiente si existen escalares de K, $\alpha_1, \alpha_2, \dots, \alpha_n$, no todos nulos, que satisfacen la ecuación:

$$\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\dots+\alpha_n\bar{v}_n=\bar{0}.$$

En caso contrario, si la ecuación sólo admite la solución trivial $(\alpha_1=\alpha_2=\dots=\alpha_n=0)$, diremos que S es linealmente independiente.

Ejemplo V. 8.

Demostrar que el conjunto

$$A=\{(1,-2,3),(-2,-3,1),(-2,4,-6)\}$$

de vectores de R^3, es linealmente dependiente, y que el conjunto

$$B=\{(1,0,1),(1,1,0),(0,1,1)\}$$

es linealmente independiente.

Solución:

a) De acuerdo con la definición, el conjunto A será linealmente dependiente si la ecuación

$$\alpha_1(1,-2,3)+\alpha_2(-2,-3,1)+\alpha_3(-2,4,-6)=(0,0,0)$$

que llamaremos de dependencia lineal, se satisface para escalares $\alpha_1, \alpha_2, \alpha_3 \in R$ no todos nulos. Demostraremos a continuación que dichos escalares existen.

Efectuando operaciones:

$$(\alpha_1-2\alpha_2-2\alpha_3, -2\alpha_1-3\alpha_2+4\alpha_3, 3\alpha_1+\alpha_2-6\alpha_3)=(0,0,0).$$

Por igualdad de vectores en R^3, obtenemos el siguiente sistema homogéneo:

$$\begin{aligned} \alpha_1-2\alpha_2-2\alpha_3&=0\\ -2\alpha_1-3\alpha_2+4\alpha_3&=0\\ 3\alpha_1+\alpha_2-6\alpha_3&=0 \end{aligned}$$

Veamos cual es el rango de la matriz de coeficientes:

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

Como el rango de la matriz es $R(A)=2$, menor que el número de incógnitas, el sistema tiene soluciones diferentes de la trivial. Por lo tanto, existen escalares no todos nulos $\alpha_1, \alpha_2, \alpha_3 \in R$ que satisfacen la ecuación de dependencia lineal. Hemos demostrado que el conjunto A es linealmente dependiente.

Si queremos obtener los escalares $\alpha_1, \alpha_2, \alpha_3$, resolvamos el sistema:

$$\begin{aligned} \alpha_1-2\alpha_2-2\alpha_3&=0\\ \alpha_2&=0 \end{aligned}$$

una de cuyas soluciones es:

$$\alpha_1=2,\ \alpha_2=0,\ \alpha_3=1$$

Entonces, los vectores de A son tales que

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

y podemos expresar al menos uno de ellos como una combinación lineal de los otros dos.

b) En forma análoga, la ecuación de dependencia lineal para los vectores de B es:

$$\alpha_1(1,0,1)+\alpha_2(1,1,0)+\alpha_3(0,1,1)=(0,0,0)$$

de donde obtenemos el siguiente sistema homogéneo:

$$\alpha_1+\alpha_2 \quad = 0$$
$$\alpha_2+\alpha_3 = 0$$
$$\alpha_1+ \quad \alpha_3 = 0$$

cuya matriz de coeficientes es de rango 3, por lo que sólo admite la solución trivial ($\alpha_1=\alpha_2=\alpha_3=0$).

Hemos demostrado que los vectores de B son linealmente independientes.(1) En consecuencia, ninguno de los vectores de B puede expresarse como una combinación lineal de los otros dos.

- -

Ejemplo V. 9.

Demostrar que los polinomios

$f(x)=x^2+2x+1$, $g(x)=5x^2+3x-2$ y $h(x)=2x^2-3x-5$

son linealmente dependientes, obteniendo al polinomio cero como una combinación lineal de ellos, con escalares no todos nulos.

Solución.

Formemos la ecuación de dependencia lineal

$\alpha f(x)+\beta g(x)+\gamma h(x)=0$, o sea:

$$\alpha(x^2+2x+1)+\beta(5x^2+3x-2)+\gamma(2x^2-3x-5)=\bar{0}$$

Efectuando operaciones y agrupando:

$$(\alpha+5\beta+2\gamma)x^2+(2\alpha+3\beta-3\gamma)x+(\alpha-2\beta-5\gamma)=\bar{0}$$

de donde obtenemos el sistema:

$$\alpha+5\beta+2\gamma = 0$$
$$2\alpha+3\beta-3\gamma = 0$$
$$\alpha-2\beta-5\gamma = 0$$

---

(1) Cuando un conjunto de vectores es linealmente independiente (dependiente), es común decir que los vectores son linealmente independientes (dependientes).

cuya matriz $A=\begin{bmatrix}1 & 5 & 2\\ 2 & 3 & -3\\ 1 & -2 & -5\end{bmatrix}$ es de rango 2.

Por lo tanto, el sistema es indeterminado. Una solución particular es:

$\alpha=-3 \qquad \beta=1 \qquad \gamma=-1$

por lo que, una combinación lineal de los polinomios es:

$$-3f(x)+g(x)-h(x)=0$$

Con esto hemos demostrado que $f(x)$, $g(x)$, $h(x)$ son polinomios linealmente dependientes.

- -

Ejemplo V. 10.

Veamos si las matrices

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

son linealmente dependientes o independientes.

La ecuación de dependencia lineal tomará en este caso la forma:

$$\alpha\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\beta\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}+\gamma\begin{bmatrix}1 & 1\\ 2 & 0\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

efectuando operaciones:

$$\begin{bmatrix}\alpha+\gamma & \beta+\gamma\\ \beta+2\gamma & \alpha\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

de donde, por igualdad de matrices:

$$\alpha+\gamma = 0$$
$$\beta+\gamma = 0$$
$$\beta+2\gamma = 0$$
$$\alpha = 0$$

sistema que sólo admite la solución trivial, por lo que las matrices propuestas son linealmente independientes.

\- -

Ejemplo V. 11.

De acuerdo con la definición, la dependencia o independencia lineal de un mismo conjunto de vectores, puede variar según la naturaleza de los escalares de K, como veremos en el siguiente ejemplo:

Consideremos el conjunto de los números complejos:

$C = \{a+bi \mid a, b \in R,\ i = \sqrt{-1}\}$

Como se vió en el Capítulo II, dicho conjunto forma un grupo abeliano para la suma y tiene las siguientes propiedades para el producto por un número real:

C es cerrado para el producto por un número real.

$\forall\ \alpha, \beta \in R$ y $z_1, z_2 \in C$:

1) $\alpha(z_1+z_2) = \alpha z_1+\alpha z_2$

2) $(\alpha+\beta)\ z_1 = \alpha z_1+\beta z_1$

3) $\alpha(\beta z_1) = (\alpha\beta)\ z_1$

4) $1 \cdot z_1 = z_1$

por lo que el conjunto C de los números complejos forma un espacio vectorial sobre el campo de los números reales.

Si en esta estructura analizamos la dependencia lineal del conjunto

$S = \{3+i,\ -2i\}$

planteando la ecuación

$\alpha(3+i) + \beta(-2i) = 0+0i$

donde los escalares $\alpha, \beta \in R$, obtendremos que:

$$\left.\begin{array}{l} 3\alpha = 0 \\ \alpha - 2\beta = 0 \end{array}\right\} \implies \alpha = \beta = 0$$

por lo que el conjunto S es linealmente INDEPENDIENTE.

Por otra parte, es claro que el conjunto C también forma un espacio vectorial sobre el campo de los números complejos.

En este caso, el producto de un escalar es la multiplicación ordinaria de números complejos y se verifican todas las propiedades que definen un espacio vectorial.

Analicemos en esta estructura la dependencia lineal del mismo conjunto

$S = \{3+i,\ -2i\}$

planteando la ecuación

$\alpha(3+i) + \beta(-2i) = 0+0i$

donde ahora los escalares $\alpha, \beta \in C$.

Sean $\alpha = a+bi$ y $\beta = c+di$, entonces:

$(a+bi)\ (3+i) + (c+di)\ (-2i) = 0+0i$

$(3a-b+2d) + (a+3b-2c)i = 0+0i$

de donde se obtiene el sistema:

$3a-b+2d = 0$

$a+3b-2c = 0$

una de cuyas soluciones es

$a=2, \quad b=0, \quad c=1; \quad d=-3.$

Entonces existen escalares complejos no nulos $\alpha = 2$ y $\beta = 1-3i$ tales que

$\alpha(3+i) + \beta(-2i) = 0+0i$

por lo que el conjunto S es linealmente DEPENDIENTE.

\- -

Enunciaremos a continuación dos importantes teoremas que son consecuencias inmediatas de la definición de dependencia lineal:

Teorema V. 6.

Todo conjunto que contiene al vector cero es linealmente dependiente.

Demostración.

Sea $\{\bar{0}, \bar{v}_1, \ldots, \bar{v}_m\}$ un conjunto que contiene al vector cero; entonces, se cumple siempre que

$$\alpha_1\bar{0}+\alpha_2\bar{v}_1+\ldots+\alpha_{m+1}\bar{v}_m=\bar{0}$$

para $\alpha_2=\ldots=\alpha_{m+1}=0$

y cualquier escalar $\alpha_1\neq 0$.

Teorema V. 7.

Todo subconjunto de un conjunto de vectores linealmente independiente, es también independiente.

Demostración.

Sea $S=\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ un conjunto de vectores linealmente independientes; entonces

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

si y sólo si $\alpha_1=\alpha_2=\ldots=\alpha_n=0$

por lo que $\alpha_i v_i=\bar{0}$ $\forall i$, y podemos cancelar en la expresión anterior tantos términos como se deseen.

De aquí que, para cualquier subconjunto de S

$S'=\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m\}$ (donde $m\leq n$)

se tendrá que: $\beta_1\bar{v}_1+\beta_2\bar{v}_2+\ldots+\beta_m\bar{v}_m=\bar{0}$

si y sólo si $\beta_1=\beta_2=\ldots=\beta_m=0$, por lo que S es linealmente independiente.

V. 4. BASE Y DIMENSION DE UN ESPACIO VECTORIAL.

Consideremos el siguiente conjunto de vectores de R^3:

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

Como veremos a continuación, cualquier vector de R^3 puede expresarse como una combinación lineal de los vectores de G.

En efecto, sea (x, y, z) un vector cualquiera de R^3; demostraremos que siempre existen escalares $\alpha, \beta, \gamma, \delta \in R$ tales que:

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

De la expresión anterior se obtiene el siguiente sistema de ecuaciones:

$$\alpha+\beta+2\delta = x$$
$$-\beta-\gamma-2\delta = y$$
$$\alpha+\gamma+2\delta = z$$

cuya matriz ampliada es

$$(A,\bar{b})=\left[\begin{array}{cccc|c} 1 & 1 & 0 & 2 & x \\ 0 & -1 & -1 & -2 & y \\ 1 & 0 & 1 & 2 & z \end{array}\right]$$

Reduciendo esta matriz a su forma escalonada tenemos:

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

donde vemos que $R(A)=R(A,\bar{b})=3$, por lo que el sistema es compatible y existen los escalares $\alpha, \beta, \gamma, \delta \in R$ para cualquier valor de x, y, z.

Hemos demostrado que cualquier vector de R^3 puede expresarse como una combinación lineal de los vectores de G. Por este hecho, diremos que G es un conjunto generador de R^3, o que los vectores de G generan al espacio vectorial R^3.

Hay que notar, sin embargo, que G es un conjunto de vectores linealmente dependiente, ya que:

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

Ahora bien, si excluímos de G al vector $(2, -2, 2)$, los vectores restantes formarán el conjunto

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

el cual, como el estudiante puede comprobar, es también un generador de R^3. A diferencia de G, B es un conjunto linealmente independiente y entonces se dice que es una base de R^3.

El estudiante puede comprobar que si excluímos de B alguno de sus vectores, el nuevo conjunto que se obtiene ya no es un generador de R^3.

Definición.

Se llama base de un espacio vectorial V, a cualquier conjunto B de vectores de V tal que:

1o. B es linealmente independiente.

2o. Cualquier vector de V puede expresarse como combinación lineal de los vectores de B.

Ejemplo V. 12.

Consideremos nuevamente el conjunto de matrices cuadradas

$$A = \left\{ \begin{bmatrix} a & 2a \\ b & 0 \end{bmatrix} \;\middle|\; a, b \in C \right\}$$

el cual, como se demostró en el ejemplo V. 5, forma un espacio vectorial para las operaciones de suma y producto por un escalar.

Si observamos la forma de las matrices que están en el conjunto A, puede verse que una base de dicho espacio es el conjunto:

$$B = \left\{ \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

como demostraremos a continuación:

1) Las matrices son linealmente independientes.

En efecto, la ecuación de dependencia es

$$\lambda_1 \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

que, por igualdad de matrices, conduce al sistema

$$\lambda_1 = 0$$
$$2\lambda_1 = 0$$
$$\lambda_2 = 0$$

cuya única solución es la trivial, por lo que las matrices son independientes.

2) B genera a A.

En efecto, cualquier matriz de A es de la forma

$$\begin{bmatrix} a & 2a \\ b & 0 \end{bmatrix}$$

y puede expresarse como

$$\begin{bmatrix} a & 2a \\ b & 0 \end{bmatrix} = a \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

es decir, cualquier matriz de A es una combinación lineal de las matrices del conjunto B.

Con 1) y 2), hemos demostrado que el conjunto B es una base de A.

- -

En los dos casos anteriores, ambos espacios vectoriales pueden ser generados por un conjunto finito de vectores, por lo que diremos que son espacios de DIMENSION FINITA. Ahora bien, no todos los espacios vectoriales pueden ser generados por un conjunto finito de vectores. Por ejemplo, el espacio de todos los polinomios es generado por el conjunto infinito

$$\{1, x, x^2, x^3, \dots\}$$

y no puede ser generado por ningún conjunto finito. A este tipo de espacios les llamaremos de DIMENSION INFINITA.

Pasemos ahora a la tarea de establecer una definición de "dimensión". Aunque la palabra dimensión suele asociarse a un concepto geométrico, aquí trataremos de encontrar una definición apropiada en términos algebraicos.

Consideremos primero el caso de dimensión finita, para lo cual requerimos de los dos teoremas siguientes:

Teorema V. 8.

Sea V un espacio vectorial generado por un conjunto finito

$$B=\{\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n\}$$

de vectores linealmente independientes. Entonces, cualquier conjunto que contenga más de n elementos es linealmente dependiente.

Demostración.

Sea $S=\{\bar{w}_1, \bar{w}_2, \dots, \bar{w}_m\}$ un subconjunto de V con más de n vectores $(m>n)$. Como el conjunto $B=\{\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n\}$ es un generador de V, cualquier vector de S es una combinación lineal de elementos de B; es decir, existen escalares α_{ij} tales que:

$$\bar{w}_1=\alpha_{11}\bar{v}_1+\alpha_{12}\bar{v}_2+\dots+\alpha_{1n}\bar{v}_n$$
$$\bar{w}_2=\alpha_{21}\bar{v}_1+\alpha_{22}\bar{v}_2+\dots+\alpha_{2n}\bar{v}_n$$
$$\dots\dots\dots\dots\dots$$
$$\bar{w}_m=\alpha_{m1}\bar{v}_1+\alpha_{m2}\bar{v}_2+\dots+\alpha_{mn}\bar{v}_n$$

Formemos ahora la ecuación de dependencia lineal para los elementos de S:

$$\lambda_1\bar{w}_1+\lambda_2\bar{w}_2+\dots+\lambda_m\bar{w}_m=\bar{0}$$

Reemplazando los vectores $\bar{w}_i$ por sus respectivas expresiones en términos de los vectores de B, tenemos:

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

efectuando los productos y agrupando:

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

como $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$ son linealmente independientes, tenemos que:

$$\lambda_1\alpha_{11}+\lambda_2\alpha_{21}+\dots+\lambda_m\alpha_{m1}=0$$
$$\lambda_1\alpha_{12}+\lambda_2\alpha_{22}+\dots+\lambda_m\alpha_{m2}=0$$
$$\dots\dots\dots\dots\dots$$
$$\lambda_1\alpha_{1n}+\lambda_2\alpha_{2n}+\dots+\lambda_m\alpha_{mn}=0$$

el cual es un sistema homogéneo de n ecuaciones con m incógnitas $(\lambda_1, \lambda_2, \dots, \lambda_m)$ y como $n<m$, el sistema admite soluciones no triviales. Por lo tanto, los vectores $\bar{w}_1, \bar{w}_2, \dots, \bar{w}_m$ son linealmente dependientes.

Con ayuda del teorema anterior, puede demostrarse el siguiente resultado importante.

Teorema V. 9.

Todas las bases de un espacio vectorial de dimensión finita, tienen el mismo número de elementos.

Demostración.

Sean

$B_1=\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ y $B_2=\{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_m\}$

dos bases cualesquiera de un espacio vectorial V.

Como B_1 es una base de V, es un conjunto independiente y genera a V.

Entonces, si B_2 es una base, debe ser un conjunto independiente y, por el teorema V. 8, $m \leq n$.

El mismo razonamiento con B_1 y B_2 intercambiados prueba que $n \leq m$.

Por lo tanto $n=m$.

Este último teorema, nos permite establecer la siguiente

Definición.

La dimensión de un espacio vectorial es el número de elementos de una cualquiera de sus bases.

Representaremos la dimensión de V mediante el símbolo dim V.

Así, para el caso del espacio vectorial R^3, hemos visto que dim $R^3=3$, mientras que para el espacio A de matrices del ejemplo V. 12 se tiene que dim A=2.

Un caso particular se presenta con el conjunto cuyo único elemento es el vector cero. Este conjunto $\{\bar{0}\}$, el cual es cerrado para las operaciones de suma y producto por un escalar, y satisface todas las propiedades que definen a un espacio vectorial, recibe el nombre de "espacio cero" o "espacio nulo".

Al espacio nulo le asignaremos arbitrariamente la dimensión cero y diremos que es de dimensión finita.

La definición de dimensión puede generalizarse aún para el caso de espacios vectoriales de dimensión infinita. Sin embargo, un estudio formal de tales espacios vectoriales requiere un tratamiento bastante más complejo que el que aquí hemos empleado. En este curso, nos interesarán especialmente las características de los espacios vectoriales de dimensión finita.

Ejemplo V. 13.

En el ejemplo V. 3 demostramos que el conjunto

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

es un espacio vectorial sobre R (posteriormente se vió que es un subespacio de R^3). Obtengamos ahora la dimensión de dicho espacio. Para ello, obtengamos primero una base de S.

Como vemos, los elementos de S son ternas ordenadas de números reales, donde la segunda componente es el doble de la primera y la tercera componente es el triple de la primera, siendo la primera componente arbitraria. Entonces, es de esperarse que el conjunto

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

sea una base de S.

En efecto, como B tiene un sólo elemento y no es el vector cero, el conjunto B es independiente. Además, cualquier vector $(x, 2x, 3x)$ de S puede obtenerse como

$(x, 2x, 3x)=x(1, 2, 3)$

por lo que B es una base de S.

Ahora, como B contiene un solo vector, de la definición de dimensión se sigue que:

dim S= 1.

Como ejercicio adicional, podemos obtener otra base de S.

Dicha base, por el teorema V. 9 deberá contener un solo elemento. que puede ser cualquier vector no nulo de S. Por ejemplo, el conjunto

$\{(\sqrt{7}\ ,\ 2\sqrt{7}\ ,\ 3\sqrt{7}\ )\}$

es otra base de S, ya que es independiente y un vector cualquiera de S puede obtenerse como

$(x,\ 2x,\ 3x)=\frac{x}{\sqrt{7}}\ (\sqrt{7}\ ,\ 2\sqrt{7}\ ,\ 3\sqrt{7}\ ).$

Vale la pena hacer una observación adicional:

Cualquier vector $(x,\ 2x,\ 3x)\ \varepsilon\ S$ puede obtenerse también como:

$x(1,\ 0,\ 0)+2x(0,\ 1,\ 0)+3x(0,\ 0,\ 1)$

y como el estudiante puede verificar, el conjunto

$E=\{(1,\ 0,\ 0),\ (0,\ 1,\ 0),\ (0,\ 0,\ 1)\}$

es linealmente independiente. Sin embargo, no es una base de S ya que E no contiene vectores de S.

- -

Ejemplo V. 14.

Sea $G=\{(1,\ 0,\ -1),\ (2,\ 3,\ 2),\ (-1,\ -3,\ -3),\ (3,\ 3,\ 1)\}$ un conjunto de cuatro vectores de R^3. Veamos cual es la dimensión del espacio L(G) generado por los vectores de G.

Como G contiene cuatro vectores de R^3 y dim $R^3=3$, entonces, por el teorema V. 8, G es linealmente dependiente.

Analicemos entonces tres vectores de G, por ejemplo:

$(1,\ 0\ -1),\ (2,\ 3,\ 2)$ y $(-1,\ -3,\ -3)$

Estos vectores son también linealmente dependientes ya que:

$(-1,\ -3,\ -3)=(1,\ 0,\ -1)-(2,\ 3,\ 2)$

Analizando ahora los vectores:

$(1,\ 0,\ -1),\ (2,\ 3,\ 2)$ y $(3,\ 3,\ 1)$

se obtiene:

$(3,\ 3,\ 1)=(1,\ 0,\ -1)+(2,\ 3,\ 2).$

Puede comprobarse fácilmente que las combinaciones restantes de tres elementos de G, arrojan conjuntos linealmente dependientes. Sin embargo, podemos encontrar subconjuntos de G con dos vectores linealmente independientes. Por ejemplo:

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

Entonces, la dimensión del espacio generado por G es 2.

Ya que dim $R^3=3$, es claro que L(G) es un subconjunto propio de R^3, por lo que se dice que L(G) es un subespacio propio de R^3.

- -

Enunciaremos a continuación un teorema que puede ser de utilidad para obtener bases de espacios de dimensión conocida. El estudiante puede consultar la referencia 2 pag. 686 para una demostración.

<u>Teorema</u> V. 10

Sea V un espacio vectorial de dimensión finita, con dim V=n.

a) Cualquier conjunto de vectores de V linealmente independiente es un subconjunto de alguna base de V.

b) Cualquier conjunto de n vectores de V linealmente independiente es una base de V.

Coordenadas de un vector con respecto a una base.

En el ejemplo V. 12, demostramos que el conjunto

$$B=\left\{\begin{bmatrix}1 & 2\\ 0 & 0\end{bmatrix},\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\right\}$$

es una base del espacio vectorial de matrices.

$$A=\left\{\begin{bmatrix}a & 2a\\ b & 0\end{bmatrix}\ \middle|\ a,\ b\ \epsilon\ C\right\}$$

sobre el campo de los números complejos.

Si consideramos una matriz de A, por ejemplo:

$$M=\begin{bmatrix}i & 2i\\ -3 & 0\end{bmatrix}$$

dicha matriz puede expresarse como una combinación lineal de las matrices de la base B como:

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

A los escalares i y -3 de C, les llamamos coordenadas de M en la base B.

Estas coordenadas pueden representarse mediante el siguiente arreglo

$$(M)_B=(i,\ -3)$$

al que se conoce como "vector de coordenadas" de la matriz M en la base B. En ocasiones, se acostumbra representar las coordenadas mediante un vector columna. Para nuestro caso tendríamos

$$(M)_B=\begin{bmatrix}i\\ -3\end{bmatrix}$$

Hay que hacer notar que el orden en que aparecen los escalares en el vector de coordenadas depende del orden en que aparecen los vectores en la base, por lo que consideraremos a las bases como conjuntos ordenados. Así, si tenemos la base de A:

$$B'=\left\{\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix},\begin{bmatrix}1 & 2\\ 0 & 0\end{bmatrix}\right\}$$

los vectores de coordenadas de M en la base B' serán:

$$(M)_{B'}=(-3,\ i) \qquad (M)_{B'}=\begin{bmatrix}-3\\ i\end{bmatrix}$$

Definición.

Sea V un espacio vectorial sobre un campo K y sea

$$B=\{\bar{v}_1,\ \bar{v}_2,\ \ldots,\ \bar{v}_n\}$$

una base ordenada de V.

Si $\bar{v}$ es un vector de V tal que

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

entonces, a los escalares $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n\ \epsilon\ K$ les llamaremos coordenadas de $\bar{v}$ en la base B, y al arreglo

$$(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)$$

le llamaremos vector de coordenadas de $\bar{v}$ en la base B.

Como hemos visto, si B es una base del espacio vectorial V, cualquier vector de V puede expresarse como una combinación lineal de los vectores de la base. Demostraremos a continuación que dicha expresión es única:

En efecto, sea $B=\{\bar{v}_1,\ \bar{v}_2,\ \ldots,\ \bar{v}_n\}$ una base de V, y sea $\bar{v}\ \epsilon\ V$ un vector cualquiera del espacio. Entonces, $\bar{v}$ puede expresarse como:

$$\bar{v}=\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_n\bar{v}_n \quad \ldots\ldots (1)$$

Supongamos que $\bar{v}$ tiene otra expresión en términos de los vectores de la base:

$$\bar{v}=\beta_1\bar{v}_1+\beta_2\bar{v}_2+\ldots+\beta_n\bar{v}_n \quad \ldots\ldots (2)$$

restando (2) de (1) se obtiene

$$\bar{0}=(\alpha_1-\beta_1)\bar{v}_1+(\alpha_2-\beta_2)\bar{v}_2+\ldots+(\alpha_n-\beta_n)\bar{v}_n$$

Como $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ son linealmente independientes:

$\alpha_1 - \beta_1 = 0$

$\alpha_2 - \beta_2 = 0$

. . . .

$\alpha_n - \beta_n = 0$

entonces, podemos concluir que

$\alpha_1 = \beta_1$

$\alpha_2 = \beta_2$

. . . . .

$\alpha_n = \beta_n$

Hemos demostrado así que la expresión de un vector cualquiera $\bar{v} \in V$ en términos de los vectores de la base B es única. Es claro entonces que $\bar{v}$ tiene un solo vector de coordenadas en la base B.

Ejemplo V. 15.

Consideremos la base de R^3:

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

obtenida al principio de esta sección.

Elijamos ahora un vector de R^3; por ejemplo:

$\bar{v} = (3, -1, 2)$

y obtengamos las coordenadas de $\bar{v}$ en la base B.

Dichas coordenadas serán los escalares α, β y γ de R tales que

$(3, -1, 2) = \alpha\,(1, 0, 1) + \beta\,(1, -1, 0) + \gamma\,(0, -1, 1)$

La igualdad planteada entre los elementos de R^3 conduce al sistema

$3 = \alpha + \beta$

$-1 = -\beta - \gamma$

$2 = \alpha + \gamma$

cuya solución es

$\alpha = 2, \quad \beta = 1, \quad \gamma = 0$

que son las coordenadas buscadas.

Entonces, el vector de coordenadas de $\bar{v}$ en la base B será

$(\bar{v})_B = (2, 1, 0)$

o bien, en forma de columna

$$(\bar{v})_B = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Ahora bien, el conjunto

$E = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

es otra base de R^3.

Puede verse claramente que

$\bar{v} = (3, -1, 2) = 3(1, 0, 0) + (-1)(0, 1, 0) + 2(0, 0, 1)$

por lo que las coordenadas del vector $\bar{v}$ en dicha base son precisamente sus componentes, es decir

$(\bar{v})_E = (3, -1, 2) = \bar{v}$

A la base E de R algunos autores le llaman "base natural" o "base canónica". En general, si R^n es el conjunto de "eneadas" ordenadas de números reales, es decir:

$R^n = \{(x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in R\}$

su base canónica es el conjunto:

$E = \{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)\}$

- -

Ejemplo V. 16.

Sean $B_1 = \{(2, -1, 0), (3, 1, 1), (0, -1, 1)\}$

y $B_2 = \{(2, 0, 1), (1, 1, 1), (3, -2, -1)\}$

dos bases de R^3 y sea $\bar{v} \in R^3$ un vector cuyas coordenadas respecto a la base B_1 son:

$(\bar{v})_{B_1} = (1, 2, -1)$

¿Cuáles son las coordenadas de $\bar{v}$ en la base B_2?

Primero obtengamos el vector $\bar{v}$ en su forma de terna ordenada:

$\bar{v} = 1(2,-1,0)+2(3,1,1)-(0,-1,1)$

$\therefore \quad \bar{v} = (8, 2, 1)$

y como se hizo anteriormente:

$(8, 2, 1) = \alpha(2, 0, 1)+\beta(1, 1, 1)+\gamma\ (3, -2, -1)$

resultando el sistema:

$8 = 2\alpha+\beta+3\gamma$

$2 = \quad \beta-2\gamma$

$1 = \alpha+\beta-\gamma$

cuya solución es:

$\alpha = -\frac{11}{3} \qquad \beta = \frac{22}{3} \qquad \gamma = \frac{8}{3}$

por tanto

$(\bar{v})_{B_2} = (-\frac{11}{3}, \frac{22}{3}, \frac{8}{3})$

- -

V. 5. ESPACIOS VECTORIALES DE FUNCIONES.

El estudiante ha tenido ya amplia experiencia en el manejo de funciones, está familiarizado con conceptos tales como dominio de una función e imagen de un elemento según una regla de correspondencia y muy probablemente ha efectuado combinaciones lineales con algunas funciones. En esta sección nos ocuparemos del tratamiento de las funciones como elementos de un espacio vectorial.

A manera de resumen, presentamos algunas definiciones y conceptos básicos:

Función.

> Definición.
>
> Sean A y B dos conjuntos no vacíos cualesquiera. Una función f de A en B es una regla o criterio que asocia a cada elemento de A uno y sólo un elemento de B.

Para expresar que f es una función de A en B, escribiremos:

$$f: A \to B$$

A los conjuntos A y B se les llama dominio y codominio de la función respectivamente.

Si $x \in A$, el elemento de B asociado a x mediante la función f lo representamos con $f(x)$ y le llamamos "imagen de x según f". Esto puede ilustrarse mediante la siguiente figura:

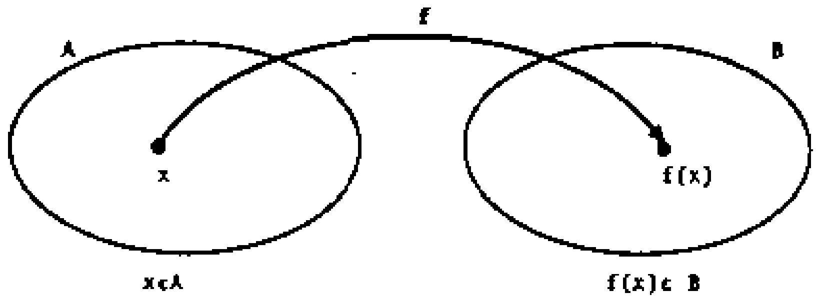


Figura V. 1. Ilustración del concepto de función.

Igualdad de funciones.

Definición.

Dos funciones f y g de un conjunto A en otro conjunto B son iguales si y sólo si la imagen de cualquier elemento del dominio según las funciones f y g es la misma. Simbólicamente:

$$f=g \iff f(x)=g(x) \quad \forall\, x \in A.$$

El espacio vectorial de las funciones reales de variable real.

Las clases más conocidas de funciones son aquellas que relacionan un número real con otro por medio de una regla de correspondencia f, es decir; funciones del tipo:

$$f:R \to R$$

como ejemplos de este tipo de funciones pueden citarse

$$f(x)= 2x^3 .$$

$$g(x)=\sqrt{3}\ \text{sen}\ (2x)$$

$$h(x)=x^2+e^x \qquad \text{etc.}$$

El conjunto de todas estas funciones forma un espacio vectorial para las operaciones de suma y producto por un escalar. Como sabemos, si f y g son dos funciones cualesquiera, la función suma f+g es aquella que relaciona al número real x con el número real f(x)+g(x) lo cual representamos mediante:

$$(f+g)(x) = f(x)+g(x) \qquad \ldots\ldots \quad (4)$$

y si $\alpha \in R$ y f es una función, el producto αf es una función tal que

$$(\alpha f)(x) = \alpha\ f(x) \qquad \ldots\ldots \quad (5)$$

El vector cero de este espacio es la llamada "función cero", la cual es una función f tal que $f(x)=0\ \forall\ x$.

Podemos encontrar algunos subespacios de este espacio vectorial, los cuales son llamados frecuentemente "espacios funcionales". Como ejemplos pueden citarse los siguientes:

1) El conjunto de todas las funciones definidas en un intervalo dado.
2) El conjunto de todos los polinomios.
3) El conjunto de todos los polinomios de grado $\leq n$, siendo n fijo.
4) El conjunto de todas las funciones continuas en un intervalo dado.
5) El conjunto de todas las funciones derivables en un punto dado.
6) El conjunto de todas las funciones integrables en un intervalo dado.
7) El conjunto de todas las funciones y tales que:

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

donde a y b son constantes dadas.

Etc.

Dependencia lineal.

Si representamos con F, al conjunto de todas las funciones reales de variable real, como F es un espacio vectorial sobre R, es claro que los conceptos de combinación, dependencia e independencia lineal son aplicables tanto a los elementos de F como a los elementos de cualquiera de sus subespacios.

Por ejemplo, las funciones f_1, f_2, f_3 $\in F$ tales que:

$f_1(x)=\cos^2 x$, $f_2(x)=\text{sen}^2 x$, $f_3(x)=5$ $\forall x \in R$ (2)

son linealmente dependientes.

En efecto, como

$\cos^2 x+\text{sen}^2 x=1$ $\forall x \in R$

se tiene que

$f_3(x)= 5\ (f_1(x)+f_2(x))$ $\forall x \in R$

entonces:

$f_3=5f_1+5f_2$

por lo que f_3 es una combinación lineal de las funciones f_1 y f_2.

Ejemplo V. 17.

Las funciones $3x^2+4$ y x^2+1 son linealmente independientes, ya que

$\alpha(3x^2+4)+\beta(x^2+1)=0$ $\forall x$ . . . .(1)

implica que

$\alpha=\beta=0$

En efecto, de (1) se sigue que

$(3\alpha+\beta)x^2+(4\alpha+\beta)=0$ $\forall x$

y por igualdad de polinomios

$3\alpha+\beta=0$

$4\alpha+\beta=0$

sistema que sólo admite la solución trivial

$\alpha=\beta=0$.

(2) En forma menos estricta, se acostumbra hablar de las funciones: $\cos^2 x$, $\text{sen}^2 x$ y 5, como eventualmente lo haremos.

El Wronskiano.

Estableceremos ahora la independencia lineal de las funciones del ejemplo anterior utilizando otro método. Si derivamos la expresión (1) se obtiene:

$\alpha(6x)+\beta(2x) = 0$ $\forall x$ . . . .(2)

Las expresiones (1) y (2) conducen al siguiente sistema:

$(3x^2+4)\alpha+(x^2+1)\beta=0$

$6x\ \alpha+\ 2x\ \ \beta=0$

Para cualquier x fija, este es un sistema homogeneo en las incógnitas α y β . El determinante de la matriz de coeficientes, llamado Wronskiano de las funciones $3x^2+4$ y x^2+1, es:

$$\begin{vmatrix} 3x^2+4 & x^2+1 \\ 6x & 2x \end{vmatrix} = 2x$$

Puesto que existe al menos un valor de x para el cual el determinante no es nulo (de hecho cualquier $x\neq 0$), la solución del sistema es $\alpha=\beta=0$ y las funciones son linealmente independientes.

Generalizaremos ahora el procedimiento anterior:

Definición.

Sean f_1, f_2, ..., f_n funciones derivables al menos n-1 veces en el intervalo (a, b). El determinante:

$$W(x)=\begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

Se llama Wronskiano de las funciones f_1, f_2, ..., f_n en el intervalo (a,b).

Teorema V. 11

Sea $W(x)$ el Wronskiano de las funciones $f_1, f_2, \ldots, f_n$ en el intervalo (a,b).

Si $W(x_0)\neq 0$ para algún x_0 en el intervalo entonces el conjunto $f_1, f_2, \ldots, f_n$ es linealmente independiente.

Demostración.

La ecuación de dependencia lineal es

$$\alpha_1 f_1(x)+\alpha_2 f_2(x)+\ldots+\alpha_n f_n(x)=0 \quad \forall\, x \in (a,b)$$

derivando n-1 veces:

$$\alpha_1 f_1'(x)+\alpha_2 f_2'(x)+\ldots+\alpha_n f_n'(x)=0$$
$$\alpha_1 f_1''(x)+\alpha_2 f_2''(x)+\ldots+\alpha_n f_n''(x)=0$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\alpha_1 f_1^{(n-1)}(x)+\alpha_2 f_2^{(n-1)}(x)+\ldots+\alpha_n f_n^{(n-1)}(x)=0$$

Si hacemos $x=x_0$ tenemos un sistema homogéneo donde $W(x_0)$ es el determinante de la matriz de coeficientes; como este determinante es diferente de cero el sistema admite sólo la solución trivial.

Entonces, el único conjunto de valores de $\alpha_1, \alpha_2, \ldots, \alpha_n$ para los cuales la ecuación

$$\alpha_1 f_1(x)+\alpha_2 f_2(x)+\ldots+\alpha_n f_n(x)=0$$

se satisface en todo el intervalo (a,b) es

$$\alpha_1=\alpha_2=\ldots=\alpha_n=0.$$

Esto prueba que las funciones son linealmente independientes.

Ejemplo V. 18.

Demostrar que las funciones f_1, f_2, f_3 tales que

$f_1(x)=\operatorname{sen}x \quad f_2(x)=\cos x \quad f_3(x)=x \quad \forall\, x \in R$

son linealmente independientes.

Solución.

El Wronskiano de f_1, f_2 y f_3 es:

$$W(x)=\begin{vmatrix} \operatorname{sen}x & \cos x & x \\ \cos x & -\operatorname{sen}x & 1 \\ -\operatorname{sen}x & -\cos x & 0 \end{vmatrix} = -x$$

ya que existen valores de x para los cuales $W(x)\neq 0$, por el teorema V. 11 las funciones f_1, f_2 y f_3 son linealmente independientes en el intervalo $(-\infty, \infty)$.

Ejemplo V. 19.

Sea $S=\{-e^{2t}+\operatorname{sen}t,\ 3e^{2t},\ 2\operatorname{sen}t-\frac{1}{2}e^{2t}\}$ un conjunto de funciones de P, determinar el subespacio generado por S así como la dimensión de dicho subespacio.

Solución.

Primero determinemos si S es un conjunto linealmente independiente, para lo cual obtenemos el Wronskiano de las funciones

$$W(t)=\begin{vmatrix} -e^{2t}+\operatorname{sen}t & 3e^{2t} & 2\operatorname{sen}t-\frac{1}{2}e^{2t} \\ -2e^{2t}+\cos t & 6e^{2t} & 2\cos t-e^{2t} \\ -4e^{2t}-\operatorname{sen}t & 12e^{2t} & -2\operatorname{sen}t-2e^{2t} \end{vmatrix}$$

multiplicando por -2 y -4 el primer renglón y sumando al segundo y tercer renglón respectivamente:

$$W(t)=\begin{vmatrix} -e^{2t}+\operatorname{sen}t & 3e^{2t} & 2\operatorname{sen}t-\frac{1}{2}e^{2t} \\ -2\operatorname{sen}t+\cos t & 0 & -4\operatorname{sen}t+2\cos t \\ -5\operatorname{sen}t & 0 & -10\operatorname{sen}t \end{vmatrix}$$

desarrollando por cofactores según la segunda columna

$$W(t) = -3e^{2t}\begin{vmatrix} -2 \text{ sen } t + \cos t & -4 \text{ sen } t + 2\cos t \\ -5 \text{ sen } t & -10 \text{ sen } t \end{vmatrix}$$

$$W(t) = -3e^{2t}(20 \text{ sen}^2 t - 10 \text{ sen}t \cos t - 20 \text{ sen}^2 t + 10 \text{ sen}t \cos t)$$

por lo que

$$W(t) = 0 \quad \forall\, t \in R.$$

En este caso, el teorema V. 11 no nos dice nada respecto a la independencia de las funciones. Para poder afirmar que son linealmente dependientes, debemos obtener tres escalares α, β y γ tales que

$$\alpha(-e^{2t}+\text{sen } t)+\beta(3e^{2t})+\gamma(2\text{sen } t - \tfrac{1}{2} e^{2t}) = 0 \qquad \ldots \quad (1)$$

Dichos escalares pueden calcularse mediante un sistema de tres ecuaciones obtenidas a partir de (1) para valores fijos de t, demostrando posteriormente que la expresión (1) se satisface para toda $t \in R$ con los mismos escalares obtenidos.

Para nuestro caso, haciendo $t=0$, $t=\frac{\pi}{2}$, $t=\pi$ en (1), se obtienen las ecuaciones:

$$-\alpha + 3\beta - \tfrac{1}{2}\gamma = 0$$

$$(1-e^{\pi})\alpha + 3e^{\pi}\beta + (2-\tfrac{1}{2}e^{\pi})\gamma = 0$$

$$-e^{2\pi}\alpha + 3e^{2\pi}\beta - \tfrac{1}{2}e^{2\pi}\gamma = 0$$

La solución general del sistema es

$$\alpha = 4\beta, \qquad \beta=\beta, \qquad \gamma=-2\beta$$

y una solución particular

$$\alpha = 4, \qquad \beta = 1, \qquad \gamma = -2$$

sustituyendo estos valores en (1) tenemos:

$$4(-e^{2t}+\text{sen } t)+3e^{2t}-2(2\text{sen}t-\tfrac{1}{2}e^{t})=0$$

Relación que se cumple $\forall\, t \in R$, por lo que las funciones son linealmente dependientes en el intervalo $(-\infty, \infty)$.

Ahora bien, como el estudiante puede demostrar fácilmente, cualquier subconjunto de S que contenga dos elementos es linealmente independiente. Por ejemplo el conjunto:

$$B = \{-e^{2t}+\text{sen } t,\ 3e^{2t}\}$$

el cual es una base de L(S), por lo que

$$\dim L(S) = 2$$

y el subespacio generado por S está constituido por todas las funciones de la forma

$$Ae^{-2t} + B \text{ sen } t.$$

- -

La no anulación de Wronskiano en al menos un punto del intervalo, se ha establecido como una condición suficiente para la independencia lineal de un conjunto de funciones (teorema V. 11). Esta condición no es necesaria, ya que el Wronskiano puede anularse en todo el intervalo siendo las funciones independientes como se ilustra en el siguiente ejemplo.

Ejemplo V. 20.

Las funciones f y g tales que:

$$f(x) = x^2 \quad \forall\, x \in (-1, 1)$$

y

$$g(x) = x|x| \quad \forall\, x \in (-1, 1)$$

son linealmente independientes en el intervalo (-1, 1) y el Wronskiano es nulo en todos los puntos de dicho intervalo.

En efecto, en el intervalo (-1, 0):

$$f(x)=x^2 \qquad y \qquad g(x)=-x^2$$

por lo que

$$W(x) = \begin{vmatrix} x^2 & -x^2 \\ 2x & -2x \end{vmatrix} = 0 \quad \forall\, x \in (-1, 0) \quad \ldots\ldots \quad (1)$$

y en el intervalo $(0, 1)$:

$$f(x) = x^2 \quad y \quad g(x) = x^2$$

por lo que

$$W(x) = \begin{vmatrix} x^2 & x^2 \\ 2x & 2x \end{vmatrix} = 0 \quad \forall\, x \in (0, 1) \quad \ldots\ldots \quad (2)$$

En consecuencia, de (1) y (2) se sigue que:

$$W(x) = 0 \quad \forall\, x \in (-1, 1)$$

Sin embargo, puede demostrarse fácilmente que f y g son linealmente independientes en dicho intervalo, ya que no existen dos escalares no nulos tales que:

$$\alpha f(x) + \beta\, g(x) = 0, \quad \forall\, x \in (-1, 1)$$

. . . .

BIBLIOGRAFIA

1.- Kenneth Hoffman / Ray Kunze
Algebra Lineal
Prentice Hall Internacional, 1973.

2.- Tom M. Apostol
Calculus Vol. I, 2a. edición
Editorial Reverté, S. A., 1972.

3.- Serge Lang
Linear Algebra
Addison Wesley Co.

ESTRUCTURAS ALGEBRAICAS

POR

EDUARDO SOLAR G.

Profesor de la División de Ciencias Básicas de la Facultad de Ingeniería de la U. N. A. M.

I.1) OPERACIONES BINARIAS

Sea S un conjunto no vacío, en el cual se define la relación de igualdad entre sus elementos: Una operación binaria $\triangle$ en dicho conjunto es una relación que asocia a cada par ordenado (a,b) de elementos de S, un elemento único $a \triangle b$, al cual podemos llamar resultado de la operación.

Si el resultado de la operación está en el conjunto S, se dice que S es cerrado con respecto a dicha operación; con lo cual la operación está definida en el conjunto.

Ejemplo I.1.1

Sea S el conjunto de los números naturales (N) y sea $\triangle$ la operación suma (+).

A cada par ordenado de números naturales (a,b) asociamos, mediante la operación suma, el número $a + b$, el cual es único y y está también en los naturales.

Por ejemplo:

al par ordenado (1,2) asociamos el número 3; es decir: $1 + 2 = 3$

al par ordenado (5,3) asociamos el número 8; es decir: $5 + 3 = 8$

al par ordenado (2,4) asociamos el número 6; es decir: $2 + 4 = 6$

etc.

Ejemplo I.1.2

Sea $S = \{\alpha, \beta, \gamma\}$ y sea Δ la operación definida por la siguiente tabla:

Δ	α	β	γ
α	α	β	γ
β	β	γ	β
γ	γ	α	α

$(\gamma \Delta \beta = \alpha)$

(NOTA: El primer elemento de la pareja debe buscarse en la columna de la izquierda y el segundo en el renglón superior. El resultado de la operación será el elemento que se encuentre en la intersección del renglón correspondiente al primer elemento y la columna correspondiente al segundo, como se ilustra).

$$\gamma \Delta \beta = \alpha$$

o sea:

al par ordenado (γ, β) asociamos el elemento α; es decir: $\gamma \Delta \beta = \alpha$

al par ordenado (α, β) asociamos el elemento β; es decir: $\alpha \Delta \beta = \beta$

al par ordenado (β, γ) asociamos el elemento β; es decir: $\beta \Delta \gamma = \beta$

etc.

Ejemplo I.1.3

La división es una operación binaria no definida en el conjunto de los números enteros (Z), puesto que:

$\forall\ a,b \in Z$ se cumple que $\frac{a}{b} \in Z$

pues no obstante que

para $-4,2 \in Z$, $\frac{-4}{2} \in Z$

se tiene también que

para $2,3 \in Z$, $\frac{2}{3} \notin Z$;

con lo cual ya no se cumple para todos los números enteros.

Propiedades de las operaciones binarias

1) Una operación binaria Δ en un conjunto S es "conmutativa", si para dos elementos cualesquiera a y b del conjunto se cumple que "a operado con b" es el mismo elemento que "b operado con a" o sea:

$\forall\ a,b \in S$ se cumple que $a \Delta b = b \Delta a$

En lo sucesivo utilizaremos solamente esta última forma para enunciar las demás propiedades.

Ejemplo I.1.4

a) Las operaciones de adición y multiplicación en el conjunto de los números reales (R) son conmutativas ya que:

$\forall\ a,b \in R$ se cumple que $a + b = b + a$

y

$\forall\ a,b \in R$ se cumple que $a \cdot b = b \cdot a$

b) La operación Δ en el conjunto S del ejemplo I.1.2 no es conmutativa, ya que

$\beta \Delta \gamma = \beta$ y $\gamma \Delta \beta = \alpha$

por lo tanto

$$\beta \,\Delta\, \gamma \neq \gamma \,\Delta\, \beta$$

2) Una operación binaria Δ en un conjunto S es "asociativa" si $\forall$ a, b, c $\in S$ se cumple que $a\Delta(b\Delta c) = (a\Delta b)\Delta c$.

Ejemplo I.1.5

a) La adición en los números naturales es asociativa, ya que:

$$\forall\, a,b,c \in \mathbb{N} \text{ se cumple que } a + (b + c) = (a + b) + c$$

b) La operación   en el conjunto S del ejemplo I.1.2 es asociativa, ya que

$$\forall\, a,b,c \in S \text{ se cumple que } a\Delta(b\Delta c) = (a\Delta b)\Delta c$$

Analicemos uno de los casos posibles:

$$\alpha \,\Delta\, (\beta \,\Delta\, \gamma) = \alpha \,\Delta\, \beta = \beta$$

$$(\alpha \,\Delta\, \beta) \,\Delta\, \gamma = \beta \,\Delta\, \gamma = \beta$$

por lo tanto

$$\alpha \,\Delta\, (\beta \,\Delta\, \gamma) = (\alpha \,\Delta\, \beta) \,\Delta\, \gamma = \beta$$

El análisis exhaustivo de todos los casos es una tarea considerable; sin embargo, se recomienda al estudiante, en este ejemplo, verificar cada una de las siguientes afirmaciones:

$$\alpha \,\Delta\, (\gamma \,\Delta\, \beta) = (\alpha \,\Delta\, \gamma) \,\Delta\, \beta$$

$$\beta \,\Delta\, (\alpha \,\Delta\, \gamma) = (\beta \,\Delta\, \alpha) \,\Delta\, \gamma$$

$$\beta \,\Delta\, (\gamma \,\Delta\, \alpha) = (\beta \,\Delta\, \gamma) \,\Delta\, \alpha$$

$$\gamma \,\Delta\, (\alpha \,\Delta\, \beta) = (\gamma \,\Delta\, \alpha) \,\Delta\, \beta$$

$$\gamma \,\Delta\, (\beta \,\Delta\, \alpha) = (\gamma \,\Delta\, \beta) \,\Delta\, \alpha$$

Suponemos pues que

$$\forall\, a,b,c \in S$$

se cumple que

$$a\Delta(b\Delta c) = (a\Delta b)\Delta c$$

en base a lo cual podemos afirmar que dicha operación es asociativa.

3) Se dice que un conjunto S está dotado de "Elemento Idéntico" con respecto a la operación binaria Δ si:

$$\exists\; u \in S \text{ tal que } \forall\, a \in S$$

se tiene que

$$a \,\Delta\, u = u \,\Delta\, a = a$$

independientemente de que la operación Δ sea conmutativa o no.

Se hace notar que puede existir un elemento que sea sólo idéntico por la izquierda; es decir: $u \,\Delta\, a = a$; o bien que sea sólo idéntico por la derecha; es decir $a \,\Delta\, u = a$

Ejemplo I.1.6

a) El elemento idéntico para la suma en el conjunto de los números reales es el número cero, ya que:

$$\exists\; 0 \in \mathbb{R} \text{ tal que } \forall\, a \in \mathbb{R}$$

se tiene que

$$a + 0 = 0 + a = a$$

b) El elemento idéntico para la multiplicación en el conjunto de los números reales es el número uno, ya que:

$\exists\ 1 \in R$ tal que $\forall\ a \in R$

se tiene que

$a \cdot 1 = 1 \cdot a = a$

c) Sea el conjunto de los números enteros y sea * la operación definida como

$x * y = 2x + y$

(donde el signo + representa la suma ordinaria)

El elemento idéntico izquierdo es $u = 0$ ya que:

$\forall\ a \in I$ se tiene que $0 * a = 0 + a = a$

mientras que para esta operación no existe elemento idéntico derecho.

d) Se deja al estudiante encontrar el elemento idéntico para la operación Δ del ejemplo I.1.2

4) Sea un conjunto S dotado de elemento idéntico u con respecto a una operación binaria Δ. Un elemento $\hat{a}$ se dice inverso de a si:

$\hat{a}\ \Delta\ a = a\ \Delta\ \hat{a} = u$

independientemente de que la operación Δ sea conmutativa o no lo sea.

Se hace notar que un elemento puede tener inverso sólo por la izquierda, es decir:

$\hat{a}\ \Delta\ a = u;$

o bien, tener inverso sólo por la derecha, es decir:

$a\ \Delta\ \hat{a} = u$

Ejemplo I.1.3

a) El elemento inverso de a para la suma en el conjunto de los reales es $\hat{a} = -a$ (llamado simétrico de a), ya que:

$\forall\ a \in R,\ \exists\ -a$ tal que $-a + a = a + (-a) = 0$

(recuérdese que el cero es el idéntico para la suma).

b) El elemento inverso para la multiplicación en el conjunto de los números reales es $\hat{a} = \frac{1}{a}$ (llamado recíproco de a), ya que:

$\forall\ a \neq 0 \in R,\ \exists\ \frac{1}{a}$ tal que $\frac{1}{a} \cdot a = a \cdot \frac{1}{a} = 1$

(recuérdese que el uno es el idéntico para la multiplicación).

c) Para la operación Δ en el conjunto S del ejemplo I.1.2 el inverso izquierdo de β es γ, ya que:

$\gamma\ \Delta\ \beta = \alpha$

(α es el idéntico para la operación Δ)

mientras que β no tiene inverso derecho, ya que:

$\nexists\ x \in S$ tal que $\beta\ \Delta\ x = \alpha$

de donde podemos concluir que β no tiene inverso.

Es conveniente hacer notar que mientras el idéntico es el mismo para todo el conjunto, los inversos son propios de cada elemento.

Sea ahora S un conjunto en el cual se definen las operaciones binarias * y Δ.

5) La operación * es "distributiva por la izquierda" sobre la operación Δ, si:

$\forall$ a, b, c $\in$ S

se cumple que:

$a*(b\Delta c) = (a*b)\Delta(a*c)$

6) La operación * es "distributiva por la derecha" sobre la operación Δ, si:

$\forall$ a, b, c $\in$ S

se cumple que

$(b\Delta c)*a = (b*a)\Delta(c*a)$

Una operación * es "totalmente distributiva" (o simplemente "distributiva") sobre otra operación Δ, si:

Es distributiva por la izquierda y por la derecha y ambos resultados son iguales; es decir, que además de cumplir con las propiedades 5 y 6, se debe cumplir que:

$a*(b\Delta c) = (b\Delta c)*a$

lo cual implica que:

$(a*b)\Delta(a*c) = (b*a)\Delta(c*a)$

Nótese que esto último requiere que la operación * sea conmutativa.

Ejemplo I.1.8

a) En el conjunto de los números reales, la multiplicación es distributiva sobre la suma, ya que:

$\forall$ a, b, c $\in$ R

se tiene que:

$a(b + c) = ab + ac$ (es distributiva por la izquierda)

$(b + c)a = ba + ca$ (es distributiva por la derecha).

y además:

$ab + ac = ba + ca$

debido a que la multiplicación es conmutativa

b) En el conjunto de los números reales, la suma no es distributiva sobre la multiplicación, ya que:

$a + (bc) \neq (a + b)(a + c)$

I.2) GRUPOS, ANILLOS Y CAMPOS

Un sistema algebraico es un conjunto no vacío S, en el cual se definen una o más operaciones binarias.

De acuerdo con el número de operaciones y con sus propiedades respecto a los elementos de S, los sistemas algebraicos poseen cierta estructura interna.

Estudiaremos a continuación la naturaleza de las principales estructuras algebraicas.

GRUPOS

La estructura algebraica más simple que aquí estudiaremos es la de Grupo.

Un conjunto "G" no vacío forma un grupo con respecto a la operación binaria *, si cumple con las siguientes propiedades:

G1) G es cerrado con respecto a la operación *:

$\forall$ a, b $\in$ G se cumple que $a*b \in G$

G2) La operación * es asociativa;

$\forall\, a,b,c \in G,\ a*(b*c) = (a*b)*c$

G3) Existe en el grupo un elemento idéntico con respecto a la operación *:

$\exists\, u \in G$ tal que $\forall\, a \in G,\ u*a = a = a*u$

G4) Para todo elemento $a \in G$ existe su inverso $\hat{a}$, que también está en el grupo:

$\forall\, a \in G,\ \exists\, \hat{a} \in G$ tal que $\hat{a} * a = u = a * \hat{a}$

Se hace notar que para que la estructura sea un "grupo", no es indispensable que la operación sea conmutativa. Si el grupo G, además de cumplir con las cuatro propiedades anteriores, cumple también con la siguiente:

G5)

$\forall\, a, b \in G,\ a*b = b*a$

se dice entonces que G forma un "Grupo Abeliano" o "Grupo Conmutativo" con respecto a la operación *.

Ejemplo I.2.1

El conjunto de los números enteros forma un grupo abeliano con respecto a la operación de suma o adición. (Se deja al estudiante comprobar que se cumple con las cinco propiedades)

Ejemplo I.2.2

¿Forman los enteros un grupo abeliano con respecto a la operación * definida a continuación:

$a*b = a - b$?

(donde el signo - representa la resta o sustracción ordinaria).

Solución:

1.- $\forall\, a, b \in Z,\ a - b \in Z$
(cumple con la cerradura)

2.- $\forall\, a, b \in Z$, se cumple que

$a - (b - c) = (a - b) - c$

puesto que:

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

$(a - b) - c = a - b - c$

de donde:

$a - (b - c) \neq (a - b) - c$

y la operación no es asociativa

Conclusión:

El conjunto de los enteros no forma un grupo con respecto a la operación *; y en consecuencia, tampoco podrá formar grupo abeliano.

Propiedades Elementales de los Grupos

A continuación enunciaremos algunas de las propiedades que cumplen los grupos debido a su estructura, las enunciaremos en forma de Teoremas y serán demostradas a partir de la definición de grupo, o bien haciendo uso de algún Teorema anterior previamente demostrado.

Teorema I. Ley de Cancelación

$\forall\, a, b, c \in G;\quad 1)\ a*b = a*c \Rightarrow b = c$

$\qquad\qquad\qquad 2)\ b*a = c*a \Rightarrow b = c$

Demostración:

Partamos de la expresión que se encuentra a la izquierda de la con dicional:

$a*b = a*c$

En base a G4 (existencia de inversos) podemos escribir:

G4: $\hat{a} * (a*b) = \hat{a} * (a*c)$

En base a G2 (ley asociativa) tenemos:

G2: $(\hat{a} * a)*b = (\hat{a} * a)*c$

En base a G4 (propiedades de los inversos) tenemos:

G4: $u*b = u*c$

En base a G3 (propiedades del idéntico) tenemos:

G3: $b = c$

Vemos que partiendo de la expresión de la izquierda y haciendo uso solamente de la definición de grupo, hemos llegado a la expresión de la derecha (en el sentido que indica la condicional), con lo que hemos demostrado que, en un grupo:

$a*b = a*c \Rightarrow b = c$

(Se deja al estudiante la demostración de la segunda parte de este Teorema).

En adelante nos concretaremos a anotar la propiedad o el Teorema que justifica cada paso.

Teorema II

$\forall$ a, b $\in$ G, las ecuaciones $a*x = b$, $y*a = b$, tienen soluciones únicas x y y en el grupo.

Demostración:

Sea la ecuación: $a*x = b$ ---- (1)

G4: $\exists\ \hat{a} \in G$

TI: $\hat{a} * (a*x) = \hat{a} * b$

G2: $(\hat{a} * a)*x = \hat{a} * b$

G4: $u*x = \hat{a} * b$

G3: $x = \hat{a} * b$

por lo que $x = \hat{a} * b$ es una solución de (1). Supongamos ahora que existen dos soluciones, x y x', de la ecuación (1). Entonces:

$ax = b$
$ax' = b$ de donde $ax = ax'$, TI: $x = x'$

con lo cual hemos demostrado que la solución $x = \hat{a} * b$ es única.

Análogamente, podemos demostrar que $y = b * \hat{a}$ es la única solución de la ecuación $y * a = b$.

Se hace notar que, en un grupo, las soluciones $x = \hat{a} * b$ y $y = b * \hat{a}$ son en general diferentes. En caso de que el grupo fuese abeliano tendríamos que $x = y$; es decir, que ambas ecuaciones tendrían la misma solución.

Teorema III

"El elemento idéntico en un grupo es único".

En efecto, el elemento idéntico es la única solución de la ecuación: $ax = a$.

Teorema IV

"El inverso de un elemento en un grupo es único".

En efecto, el inverso $\hat{a}$ de a es la única solución de la ecuación: $ax = u$

Teorema V

"El inverso de $\hat{a}$ es a".

Es decir: $\forall\ a \in G$: $\hat{(\hat{a})} = a$

Teorema VI:

"El inverso del resultado de una operación es el resultado de la operación de los inversos en orden contrario".

Es decir: $\forall\, a, b \in G,\ \widehat{(a * b)} = \hat{b} * \hat{a}$

Corolario: $\forall\, a, b, \dots, p, q \in G$

$\widehat{(a * b * \dots * p * q)} = \hat{q} * \hat{p} * \dots * \hat{b} * \hat{a}$

Se deja al estudiante la demostración de los teoremas V y VI.

En base a la propiedad asociativa de una operación * en un grupo, daremos la siguiente definición.

D1) Para cualquier $a \in G$ y $n \in N$:

$\overset{n}{a} = a * a * \dots * a$ (n términos)

$\overset{0}{a} = u$ (u = elemento idéntico para la operación *)

$\overset{-n}{a} = \widehat{(\overset{n}{a})} = \hat{a} * \hat{a} * \dots * \hat{a}$ (n términos)

En base a D1 se puede demostrar que:

D2) Para cualquier $a \in G$ y $m, n \in N$:

1.- $\overset{m}{a} * \overset{n}{a} = \overset{m+n}{a}$

y

2.- $\overset{n}{(\overset{m}{a})} = \overset{mn}{a}$

Ejemplo 1.2.3

a) Si $G = R$ y * es la multiplicación,

D1 nos dice que:

$\overset{n}{a} = a^n = a \cdot a \cdot \dots \cdot a$ (n factores)

$\overset{0}{a} = a^0 = 1$ (elemento idéntico para la multiplicación)

$\overset{-n}{a} = a^{-n} = \frac{1}{a} \cdot \frac{1}{a} \cdot \dots \cdot \frac{1}{a}$ (n factores)

y D2 establece que:

$a^m \cdot a^n = a^{m+n}$ y $(a^m)^n = a^{mn}$

b) Si $G = R$ y * es la suma,

D1 nos dice que:

$\overset{n}{a} = n \cdot a = a + a + \dots + a$ (n sumandos)

$\overset{0}{a} = 0 \cdot a = 0$ (elemento idéntico para la suma)

$\overset{-n}{a} = -n \cdot a = (-a) + (-a) + \dots + (-a)$ (n sumandos)

y D2 establece que:

$ma + na = (m + n)a$ y $m(n \cdot a) = (mn) \cdot a$

ANILLOS

Hemos estudiado hasta ahora las propiedades de una estructura algebraica en la que se define una sola operación. Las estructuras que manejaremos en adelante se definen con respecto a dos operaciones binarias diferentes.

Un conjunto "A" no vacío, forma un anillo con respecto a dos operaciones binarias $\oplus$ y $\odot$ si:

A1) El conjunto A forma un grupo abeliano con respecto a la operación $\oplus$.

A2) El conjunto A es cerrado con respecto a la operación $\odot$:

$\forall\, a, b \in A:\ a \odot b \in A$

A3) La operación $\odot$ es asociativa:

$\forall\, a, b, c \in A:\ a \odot (b \odot c) = (a \odot b) \odot c$

A4) La segunda operación ($\odot$) es distributiva sobre la primera ($\oplus$), tanto por la izquierda como por la derecha:

$\forall\, a, b, c \in A:\ a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$

$(b \oplus c) \odot a = (b \odot a) \oplus (c \odot a)$

Se hace notar que para que la estructura sea anillo no es indispensable que la segunda operación ($\otimes$) sea conmutativa. En caso de serlo, la estructura toma el nombre de "anillo conmutativo"; es decir si:

A5) $\forall\ a, b \in A;\ a \otimes b = b \otimes a$

Es evidente entonces, que para un anillo conmutativo la operación $\otimes$ es totalmente distributiva sobre $\oplus$.

Puesto que un anillo forma un grupo abeliano con respecto a la primera operación ($\oplus$), cumplirá para dicha operación con todas las propiedades de grupo que ya hemos enunciado.

Al elemento idéntico para la operación $\oplus$ le asignaremos el nombre de "elemento cero del anillo", y lo denotaremos con "z". Al elemento idéntico para la operación $\otimes$, <u>en caso de existir</u>, le llamaremos "<u>elemento unidad</u>", y lo denotaremos con "v".

Resulta evidente que un anillo puede o no tener "elemento unidad", pues para que una estructura forme anillo no es indispensable que exista el idéntico para la operación $\otimes$; en caso de existir éste, o sea:

$\exists\ v \in A$ tal que $\forall\ a \in A;\ v \otimes a = a \otimes v = a$

se dice que la estructura es un "<u>Anillo con unidad</u>".

Ejemplo 1.2.4

El conjunto de los números enteros (Z) forma un anillo con respecto a las operaciones de suma y multiplicación.

En efecto:

1) $\forall\ a,b \in Z,\ a + b \in Z$ (es cerrado con respecto a la suma).

2) $\forall\ a,b,c \in Z,\ a + (b + c) = (a + b) + c$ (la suma es asociativa)

3) $\exists\ 0 \in Z$ tal que $\forall\ a \in Z,\ 0 + a = a + 0 = a$ (existencia del idéntico aditivo).

4) $\forall\ a \in Z, \exists\ (-a) \in Z$ tal que $(-a) + a = a + (-a) = 0$ (existencia de los inversos aditivos)

Hasta ahora tenemos que Z forma un grupo con respecto a la suma.

5) $\forall\ a,b \in Z:\ a + b = b + a$
forma un grupo abeliano con respecto a la suma

6) $\forall\ a,b \in Z:\ ab \in Z$
(es cerrado con respecto a la multiplicación)

7) $\forall\ a,b \in Z:\ a(bc) = (ab)c$
(la multiplicación es asociativa)

8) $\forall\ a,b,c \in Z$ se tiene que: $a(b + c) = ab + ac$
$(b + c)a = ba + ca$
(la multiplicación es distributiva sobre la suma por ambos lados).

$\therefore$ los enteros forman un anillo con respecto a las operaciones de suma y multiplicación; y el cero del anillo es el número cero ($z = 0$)

Continuemos con el análisis de esta estructura:

9) $\forall\ a,b \in Z:\ ab = ba$
(la multiplicación es conmutativa)

$\therefore$ El anillo es, además, conmutativo

10) $\exists\ 1 \in Z$ tal que $\forall\ a \in Z:\ 1 \cdot a = a \cdot 1 = a$
(existe en el anillo el idéntico multiplicativo)

Por lo tanto, se trata de un anillo conmutativo con unidad ($v = 1$)

Ejemplo I.2.5

Sea el conjunto $S = \{a, b\}$. Investigar si forma anillo con respecto a las operaciones $\oplus$ y $\odot$, definidas por las siguientes tablas:

$\oplus$	a	b
a	a	b
b	b	a

$\odot$	a	b
a	a	a
b	a	b

Solución:

Veamos con respecto a la operación +:

1) $\forall\, x, y \in S:\ x + y \in S$
(S es cerrado con respecto a $\oplus$)

2) Probemos ahora algunas combinaciones diferentes para la asociatividad:

i) $a \oplus (b \oplus a) = (a \oplus b) \oplus a$
$a \oplus b = b \oplus a$
$b = b$ (se cumple)

ii) $b \oplus (a \oplus b) = (b \oplus a) \oplus b$
$b \oplus b = b \oplus b$ (se cumple)

iii) $b \oplus (b \oplus b) = (b \oplus b) \oplus b$
$b \oplus a = a \oplus b$
$b = b$ (se cumple)

Suponemos pues que la operación $\oplus$ es asociativa (no haremos el análisis exhaustivo de todos los casos posibles).

3) $a \in S$ tal que $\forall\, x \in S:\ a \oplus x = x \oplus a = x$
(existe el idéntico para la operación $\oplus$ y es "a".

4) El inverso de a es a, ya que: $a \oplus a = a$
El inverso de b es b, ya que: $b \oplus b = a$
(existen los inversos para $\oplus$)

5) $a \oplus b = b \oplus a = b$ (la operación $\oplus$ es conmutativa)

6) $\forall\, x, y \in S:\ x \odot y \in S$ (S es cerrado con respecto a $\odot$)

7) La operación $\odot$ es asociativa

i) $a \odot (b \odot a) = (a \odot b) \odot a$
$a \odot a = a \odot a$ (se cumple)

ii) $b \odot (a \odot b) = (b \odot a) \odot b$
$b \odot a = a \odot b$
$a = a$ (se cumple)

iii) $b \odot (b \odot b) = (b \odot b) \odot b$
$b \odot b = b \odot b$ (se cumple)

8) Probemos algunas combinaciones diferentes para la distributividad:

i) $a \odot (b \oplus a) = (a \odot b) \oplus (a \odot a)$
$a \odot b = a \oplus a$
$a = a$ (se cumple)

ii) $b \odot (a \oplus b) = (b \odot a) \oplus (b \odot b)$
$b \odot b = a \oplus b$
$b = b$ (se cumple por la izquierda)

iii) $(b \oplus a) \odot a = (b \odot a) \oplus (a \odot a)$
$b \odot a = a \oplus a$
$a = a$ (se cumple)

iv) $(a \oplus b) \odot b = (a \odot b) \oplus (b \odot b)$
$b \odot b = a \oplus b$
$b = b$ (se cumple por la derecha)

Por lo tanto, el conjunto S forma un anillo con respecto a las operaciones $\oplus$ y $\odot$

Se deja al estudiante investigar si el anillo es conmutativo y si tiene unidad.

CAMPOS

Hemos llegado finalmente al estudio de la estructura algebraica más completa, en la cual también se definen dos operaciones binarias: Un campo es un anillo conmutativo con unidad, en el cual existen los inversos para la segunda operación, con excepción del cero del anillo (z), el cual carece de dicho inverso.
Este concepto de campo nos conduce a la siguiente definición:

Un conjunto "C" de por lo menos dos elementos forma un campo con respecto a dos operaciones binarias $\oplus$ y $\otimes$ si:

C1) Forma un grupo abeliano con respecto a la operación $\oplus$, a cuyo elemento idéntico llamamos "elemento cero del campo" y representamos con "z"

C2) Sus elementos diferentes de z forman un grupo abeliano con respecto a la operación $\otimes$, a cuyo elemento idéntico llamamos "elemento unidad del campo", y lo representamos con "v".

C3) La operación $\otimes$ es distributiva sobre la operación $\oplus$.

Ejemplo I.2.6

Sea el conjunto A = {p, q, r, s}

Investigar si forma un campo con respecto a las operaciones $\oplus$ y $\otimes$ definidas a continuación:

$\oplus$	p	q	r	s
p	r	s	p	q
q	s	r	q	p
r	p	q	r	s
s	q	p	s	r

$\otimes$	p	q	r	s
p	s	p	r	q
q	p	q	r	s
r	r	r	r	r
s	q	s	r	p

Solución

1) A es cerrado con respecto a $\oplus$

2) i) $p \oplus (q \oplus r) = (p \oplus q) \oplus r$
$$p \oplus q = s \oplus r$$
$$s = s$$
ii) $s \oplus (q \oplus p) = (s \oplus q) \oplus p$
$$s \oplus s = p \oplus p$$
$$r = r$$
iii) $r \oplus (s \oplus q) = (r \oplus s) \oplus q$
$$r \oplus p = s \oplus q$$
$$p = p$$

3) Existe el idéntico para la $\oplus$ (z = r)

4) Existen los inversos para la $\oplus$ y son únicos en cada caso
$\hat{p} = p$, $\hat{q} = q$, $\hat{r} = r$, $\hat{s} = s$

5) La operación $\oplus$ es conmutativa (nótese la simetría de la tabla).

∴ El conjunto A forma un grupo abeliano con respecto a la operación $\oplus$.

Además:

6) A es cerrado con respecto a la $\odot$

7) i) $p \odot (q \odot r) = (p \odot q) \odot r$

$$p \odot r = p \odot r$$

ii) $s \odot (r \odot q) = (s \odot r) \odot q$

$$s \odot r = r \odot q$$

$$r = r$$

iii) $s \odot (q \odot p) = (s \odot q) \odot p$

$$s \odot p = s \odot p$$

La operación $\odot$ es asociativa

8) i) $p \odot (q \oplus r) = (p \odot q) \oplus (p \odot r)$

$$p \odot q = p \oplus r$$

$$p = p$$

ii) $s \odot (r \oplus q) = (s \odot r) \oplus (s \odot q)$

$$s \odot q = r \oplus s$$

$$s = s$$

$\odot$ es distributiva por la izquierda sobre $\oplus$

iii) $(p \oplus q) \odot r = (p \odot r) \oplus (q \odot r)$

$$s \odot r = r \oplus r$$

$$r = r$$

iv) $(s \oplus p) \odot q = (s \odot q) \oplus (p \odot q)$

$$q \odot q = s \oplus p$$

$$q = q$$

$\odot$ es distributiva por la derecha sobre $\oplus$

9) La operación $\odot$ es conmutativa

(Nótese la simetría de la tabla)

10) Existe el idéntico para la $\odot$ $(u = q)$

11) $\hat{p} = s,\ \hat{q} = q,\ \hat{s} = p$

Todos los elementos tienen inverso para la $\odot$ excepto r, que es el cero.

Por lo tanto, el conjunto A forma un campo con respecto a las operaciones $\oplus$ y $\odot$

CONJUNTOS METRICOS

CONCEPTOS TOPOLOGICOS

ESTUDIO GENERAL DE LAS FUNCIONES

POR:

ARTURO DELGADO R.

Profesor de la División de Estudios de Posgrado de la Facultad de Ingeniería de la U. N. A. M.

Conjuntos métricos. Métrica. Sea M un conjunto cualquiera. Se dice que S es una "métrica" con respecto a M si, y sólo si para dos elementos $p,q \in M$ existe un número real asociado $S(p,q)$ denominado "distancia" de p a q; que tiene las siguientes propiedades:

1.- $S(p,q) \geq 0$

2.- $S(p,q) = 0 \iff p = q$

3.- $S(p,q) = S(q,p)$

4.- $S(p,q) + S(q,r) \geq S(p,r)$ para todo $p,q,r \in M$

Estas propiedades coinciden con el concepto intuitivo de distancia.

"Conjunto métrico" se define como el par (M,S) consistente en un conjunto M y una métrica S para M.

Los elementos de M suelen llamarse "puntos"

Ejemplos de conjuntos métricos

1) $M_1 = \{x \mid x \in R^1\}$; $S_1\ (p,q) = |x_1 - x_2|$; $x_1 - x_2 \in M_1$

$M_1 = E^1$ = <u>Espacio Euclidiano</u> unidimensional

2) $M_2 = (x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in R^1$; para toda

$p = (x_1, x_2, \ldots, x_n)$, $q = (y_1, y_2, \ldots, y_n) \in M_2$

$$S_2(p,q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$$

$M_2 = E^n$ = <u>Espacio Euclidiano</u> n - dimensional

3) $M_3 = \{x \mid x \in R^1\}$; $p = x_1$, $q = x_2$

$$S_3(p,q) = \frac{|x_1 - x_2|}{1 + |x_1 - x_2|}$$

Demostración de que $S_3(p,q)$ constituye una métrica. Debemos hacer ver que se cumplen las 4 propiedades que definen una métrica: (Obsérvese que $M_3 = M_1$)

1.- $S_3(p,q) \geq 0$, en efecto:

si $x_1 \neq x_2$, sea $|x_1 - x_2| = n > 0$

$$S_3(p,q) = \frac{n}{1+n} > 0$$

2.- Si $x_1 = x_2 \Rightarrow |x_1 - x_2| = 0$

$$S_3(x_1, x_2) = \frac{0}{1} = 0$$

recíprocamente: $S_3(p,q) = 0 = \frac{|x_1 - x_2|}{1+|x_1 - x_2|}$

$$\Rightarrow |x_1 - x_2| = 0 \quad \Rightarrow x_1 = x_2$$

3.- $S_3(p,q) = S_3(q,p)$

$$\frac{|x_1 - x_2|}{1+|x_1 - x_2|} = \frac{|x_2 - x_1|}{1+|x_2 - x_1|}$$

4.- $S(p,q) + S(q,r) \geq S(p,r)$

Sea $p = x_1$, $q = x_2$, $r = x_3$

$$\frac{|x_1 - x_2|}{1+|x_1 - x_2|} + \frac{|x_2 - x_3|}{1+|x_2 - x_3|} \geq$$

$$\geq \frac{|x_1 - x_2|}{1+|x_1 - x_2|+|x_2 - x_3|} + \frac{|x_2 - x_3|}{1+|x_1 - x_2|+|x_2 - x_3|} =$$

$$= \frac{|x_1 - x_2|+|x_2 - x_3|}{1+|x_1 - x_2|+|x_2 - x_3|} \geq \frac{|x_1 - x_2 + x_2 - x_3|}{1+|x_1 - x_2 + x_2 - x_3|} =$$

$$= \frac{|x_1 - x_3|}{1+|x_1 - x_3|}$$

$$\Rightarrow \frac{|x_1 - x_2|}{1+|x_1 - x_2|} + \frac{|x_2 - x_3|}{1+|x_2 - x_3|} \geq \frac{|x_1 - x_3|}{1+|x_1 - x_3|}$$

4) Sea M_4 cualquier conjunto, siendo $S_4(p,q)$ tal que, para toda $p,q \in M_4$:

$$S_4(p,q) = \begin{cases} 0, & \text{si } p = q \\ 1, & \text{si } p \neq q \end{cases}$$

Demostración de que $S_4(p,q)$ puede aceptarse como métrica válida:

1.- $S_4(p,q) \geq 0$ (por definición)

2.- $S_4(p,q) = 0$, si $p = q$ (por definición)

3.- $S_4(p,q) = S_4(q,p)$

si $p \neq q$:

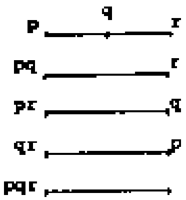
$$S_4(p,q) = S_4(q,p) = 1 \quad \text{(por definición)}$$

si $p = q$:

$$S_4(p,p) = S_4(q,q) = 0$$

4.- $S_4(p,q) + S_4(q,r) \geq S_4(p,r)$

De acuerdo con la posición relativa que ocupen los puntos p,q,r, pueden presentarse los siguientes casos:

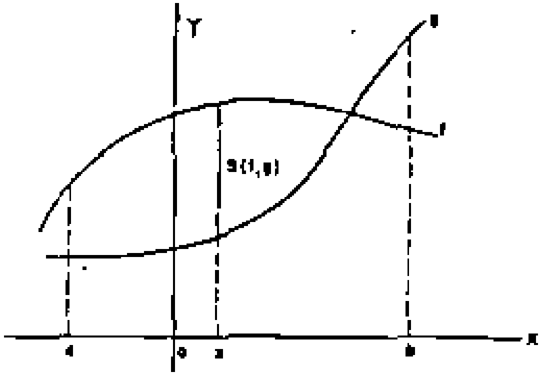


$S(p,q)$	+	$S(q,r)$		$S(p,r)$
1	+	1	>	1
0	+	1	=	1
1	+	1	>	0
1	+	0	=	1
0	+	0	=	0

$$\Rightarrow S(p,q) + S(q,r) \geq S(p,r)$$

5) Sea M_5 el conjunto de funciones continuas en el intervalo cerrado $\boxed{a,b}$; definiéndose para dos funciones $f,g \in M_5$:

$S_5(f,g) = \max |f(x) - g(x)|;\ x \in \boxed{a,b}$



Demostración de las propiedades.

Puesto que, por definición f y g son funciones continuas $\in \boxed{a,b} \Rightarrow |f(x) - g(x)|$ es asimismo función continua,

$\Rightarrow \max |f(x) - g(x)| \in \boxed{a,b}$, de ahí que tenga sentido $S_5(f,g)$ en la forma en que se define.

1.- $\max |f(x) - g(x)| \geq 0$

2.- $\max |f(x) - g(x)| = 0 \Longleftrightarrow f(x) = g(x)$ para toda $x \in \boxed{a,b}$

3.- $\max |f(x) - g(x)| = \max |g(x) - f(x)|$

4.- Si $f,g,h \in M_5$:

$|f(x) - h(x)| = |f(x) - g(x) + g(x) - h(x)| \leq$
$\leq |f(x) - g(x)| + |g(x) - h(x)|$

$\Rightarrow \max |f(x) - h(x)| \leq \max \left[|f(x) - g(x)| + |g(x) - h(x)|\right] \leq$
$\leq \max |f(x) - g(x)| + \max |g(x) - h(x)|$

$\Rightarrow S_5(f,g) + S_5(g,h) \geq S_5(f,h)$

6) Sea M_5 el mismo conjunto M_5 del problema anterior: defínese ahora para el espacio:

$S_6(f,g) = \int_a^b |f(x) - g(x)|\, dx$

Demostración de las propiedades:

1.- $\int_a^b |f(x) - g(x)|\, dx \geq 0$

2.- $\int_a^b |f(x) - g(x)|\, dx = 0 \Longleftrightarrow f(x) - g(x) \equiv 0$, puesto que $a \neq b$

3.- $\int_a^b |f(x) - g(x)|\, dx = \int_a^b |g(x) - f(x)|\, dx$

4.- $\int_a^b |f(x) - g(x)|\, dx + \int_a^b |g(x) - h(x)|\, dx =$

$= \int_a^b (|f(x) - g(x)| + |g(x) - h(x)|)\, dx \geq$

$\geq \int_a^b |f(x) - g(x) + g(x) - h(x)|\, dx = \int_a^b |f(x) - h(x)|\, dx$

$\Rightarrow \int_a^b |f(x) - g(x)|\, dx + \int_a^b |g(x) - h(x)|\, dx \geq$

$\geq \int_a^b |f(x) - h(x)|\, dx$

o sea

$S_6(f,g) + S_6(g,h) \geq S_6(f,h)$

$\forall\ f,g,h \in M_5$

Los ejemplos anteriores hacen ver que un espacio puede ser convertido en un espacio métrico en más de una manera; esto es, para un mismo espacio, pueden existir varias métricas admisibles.

Entornos (o vecindades). Sea (M,S) cualquier conjunto métrico; siendo p un punto fijo en M. Sea además $\varepsilon > 0$, $\varepsilon \in R^1$.
Definimos como entorno (o vecindad) $V(p,\varepsilon)$ de p, con radio ε, al conjunto:

$V(p,\varepsilon) = \{q \mid q \in M \text{ y } S(p,q) < \varepsilon\}$

Ejemplos:

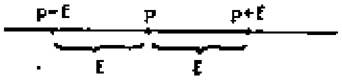
1.- En E^1; $V(p,\varepsilon)$ es el intervalo abierto: $(p-\varepsilon,\ p+\varepsilon)$; puesto que:

$V(p,\varepsilon) = \{x \mid p-\varepsilon < x < p+\varepsilon\}$

o

$V(p,\varepsilon) = \{x \mid |x-p| < \varepsilon\}$

Los extremos del intervalo (puntos $p-\varepsilon$ y $p+\varepsilon$) no están en el entorno.



2.- En E^3; $V(p,\varepsilon)$ es el conjunto de puntos dentro de la esfera con centro $p=(a,b,c)$ y radio ε:

$V(p,\varepsilon) = \{(x,y,z) \mid \sqrt{(x-a)^2+(y-b)^2+(z-c)^2} < \varepsilon\}$

La superficie exterior (o frontera) de la esfera <u>no</u> está en el entorno $V(p,\varepsilon)$

3.- Sea A cualquier conjunto cuya métrica se define del siguiente modo:

$$S(p,q) = \begin{cases} 0, & \text{si } p=q \\ 1, & \text{si } p \neq q \end{cases}$$

$$\Rightarrow \begin{cases} V(p,\varepsilon) = p, & \text{si } \varepsilon \leq 1 \\ V(p,\varepsilon) = A, & \text{si } \varepsilon > 1 \end{cases}$$

<u>Entorno incompleto</u> (o vecindad reducida o agujerada). Es un entorno que contiene un punto menos que $V(p,\varepsilon)$.

Notación: $\dot{V}(p,\varepsilon)$ o $V'(p,\varepsilon)$

$\dot{V}(p,\varepsilon) = \{q \mid 0 < S(p,q) < \varepsilon\}$

o

$\dot{V}(p,\varepsilon) = V(p,\varepsilon) - \{p\} = V'(p,\varepsilon)$

<u>Punto interior</u>. Sea el conjunto $A \subset (M,S)$. Se dice que p es un punto "interior" de A, si existe $V(p,\varepsilon)$ tal que $V(p,\varepsilon) \subset A$.

<u>Punto exterior</u>. El punto q es punto "exterior" de A, si existe $V(q,\varepsilon)$ tal que:

$V(q,\varepsilon) \subset (M,S) - A$

Haciendo, para abreviar, $(M,S) \equiv M$; resulta que un punto exterior de A es un punto interior de $M - A$

<u>Punto frontera</u>. Si todo entorno $V(p,\varepsilon)$ contiene un punto de A y un punto de $M - A$, se dice que p es punto "frontera" de A. Es obvio que p es un punto frontera de A si, y sólo si es punto frontera de $M - A$.

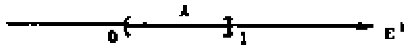
Ejemplos:

1) Sea $M = E^1$; siendo $A = \{x \mid 0 < x \leq 1\}$

El interior de A es $(0,1)$

La frontera de A es $\{0,1\}$; son 2 puntos

El exterior de A es $E^1 - [0,1] = \{x \mid x > 1\} \cup \{x \mid x < 0\}$

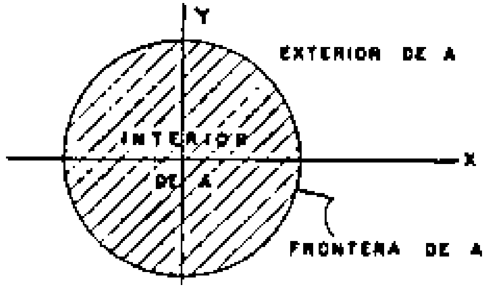


2) Sea $M = E^2$; $A = \{(x,y) \mid x^2+y^2 \leq 4\}$

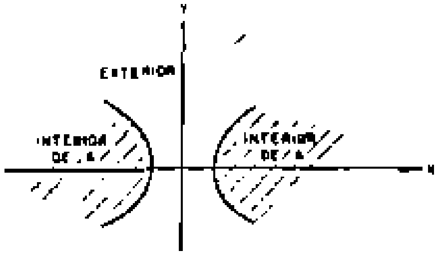
El interior de A es $\{(x,y) \mid x^2+y^2 < 4\}$

El exterior de A es $\{(x,y) \mid x^2+y^2 > 4\}$

La frontera de A es la circunferencia $\{(x,y) \mid x^2+y^2 = 4\}$



3) Sea $M = E^2$; $A = \{(x,y) \mid x^2 - y^2 > 1\}$



EL INTERIOR DE $A = A = \{(x,y) \mid x^2 - y^2 > 1\}$

EL EXTERIOR DE $A = \{(x,y) \mid x^2 - y^2 < 1\}$

LA FRONTERA DE $A = \{(x,y) \mid x^2 - y^2 = 1\}$

Conjunto abierto. Sea $A \subset M$. Se dice que el conjunto A es un conjunto "abierto", si, y sólo si no contiene ningún punto de su frontera.

Conjunto cerrado. El conjunto A es un conjunto cerrado si, y sólo si contiene todos sus puntos frontera.

De las definiciones anteriores se deduce que A es un conjunto abierto, si todo punto de A es al mismo tiempo punto interior de A.

El conjunto del ejemplo 1) anterior no es ní conjunto cerrado ni conjunto abierto.

El conjunto del ejemplo 2) es conjunto cerrado.

En el ejemplo 3) el conjunto es abierto.

El complemento de un conjunto abierto es un conjunto cerrado; siendo el complemento de un conjunto cerrado, un conjunto abierto.

Punto de acumulación (o punto límite). Sea $A \subset M$; $p \in M$. Se dice que p es un punto de "acumulación" de A, si todo entorno incompleto $V'(p,\varepsilon)$ contiene un punto de A. Dicho de otro modo, p es un punto de acumulación de A, si cada entorno $V(p,\varepsilon)$ contiene un punto de A, que no sea el punto p.

El conjunto de puntos de acumulación en el ejemplo 1) anterior es:

$$[0,1] = \{x \mid 0 \leq x \leq 1\}$$

En el ejemplo 2), el conjunto de puntos de acumulación de A, es el mismo conjunto A.

El conjunto de puntos de acumulación del conjunto A, en el ejemplo 3) es:

$$\{(x,y) \mid x^2 - y^2 \geq 1\}$$

Si p es un punto de acumulación de un conjunto $A \subset M \Rightarrow$ que todo entorno $V(p,\varepsilon)$ contiene un número infinito de puntos del conjunto A.

Un conjunto cerrado contiene todos sus puntos de acumulación; recíprocamente, un conjunto que contiene todos sus puntos de acumulación, es un conjunto cerrado.

Funciones. La noción de función es uno de los conceptos básicos de mayor trascendencia en el estudio de las matemáticas.

En el capítulo anterior se precisó la idea de función, describiéndola ésta como una relación en la cual ningún par ordenado tiene el mismo primer elemento. En este capítulo se hará un estudio más completo de las funciones, clasificándolas y estableciendo sus propiedades.

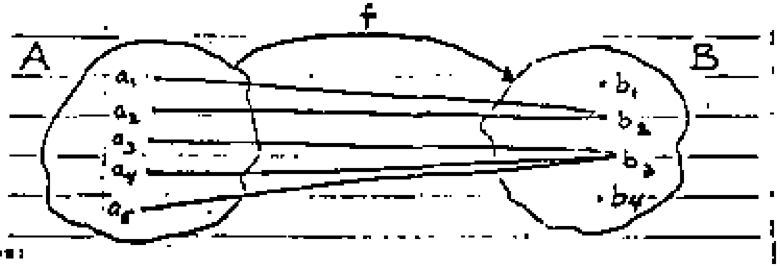
Función unívoca (o función en). Este es el tipo más general de función. Una función unívoca de un conjunto A en un conjunto B es una regla (o mapeo) que asocia a cada elemento $a \in A$, un único elemento $b \in B$

Notaciones:

$f: A \to B$

ó $A \xrightarrow{f} B$

De manera gráfica, podemos representar la función unívoca como sigue:



Observaciones:

1.- Todos los elementos del conjunto A deben tener una imagen en el conjunto B; pudiendo, en la función unívoca, ser un elemento $b \in B$ imagen de más de un elemento del conjunto A.

2.- La noción de función involucra:

i) Un conjunto A denominado "dominio" de f

ii) Un conjunto B denominado "codominio" (o contradominio de f)

iii) Una regla que establece qué elementos del conjunto B son imágenes de los elementos del conjunto A.

Se acostumbra designar las imágenes $a_i \in A$ como $f(a_i) \in B$; ésto es, de la figura anterior,

$$f(a_1) = f(a_2) = b_2$$

$$f(a_3) = f(a_4) = f(a_5) = b_3$$

$$f(A) = \{b_2, b_3\}$$

$$f = \{(a_1,b_2),\ (a_2,b_2),\ (a_3,b_3),(a_4,b_3),(a_5,b_3)\}$$

3.- Si $f: R^1 \to R^1$, la función f se denomina "función real de variable real".

Si el dominio y el codominio están constituidos por el mismo conjunto A, se dice que f es un "operador" (o una transformación) en A.

$A \xrightarrow{f} A$

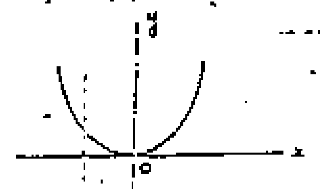
Se llama "rango" de f al conjunto constituido por todas las imágenes $f(A) \subset B$ de elementos $a \in A$

$f(A) \subseteq B$

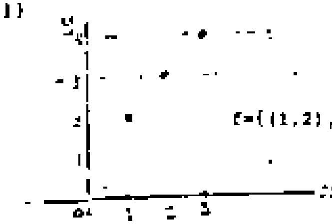
Geométricamente: Una línea vertical corta la gráfica de una función unívoca en un solo punto $(f: R^1 \to R^1)$

Ejemplo: 1) $A = B = R^1$

$f: A \to B$ tal que $y = f(x) = x^2$



2) $f = \{(x, x+1) \mid x \in \{1,2,3\}\}$

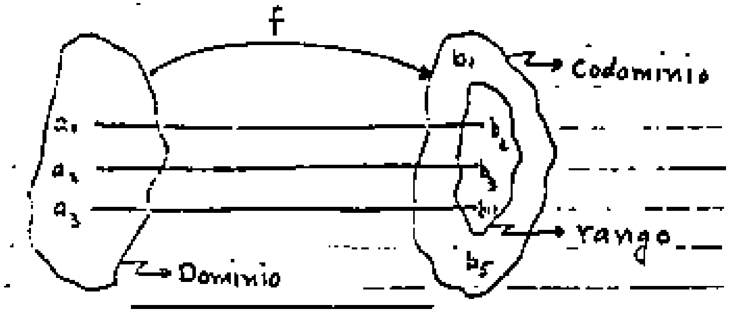


Función biunívoca (inyectiva o uno a uno). Una función $f: A \to B$ se denomina biunívoca, si para todo par de elementos $a_1, a_2 \in A$ distintos, las imágenes bajo f son asimismo distintas; ésto es:

$f(a_1) = f(a_2) \Rightarrow a_1 = a_2$

Ejemplos:

1)



2) $A = B = R^1$

$f: R^1 \to R^1$; siendo $y = x$

f no es biunívoca, pues: $f(1) = f(-1) = 1$

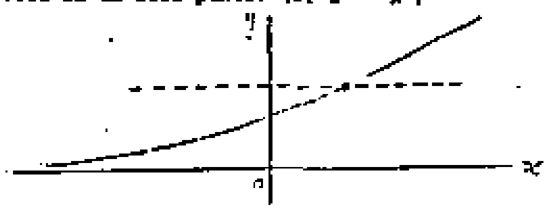
3) $f: R^1 \to R^1$, tal que $f(x) = x$

f sí es biunívoca

4) $f: R^1 \to R^1$, tal que $f(x) = e^x$

f sí es biunívoca

Geométricamente: Una línea horizontal corta la gráfica de una función biunívoca en un solo punto. ($f: R^1 \to R^1$)



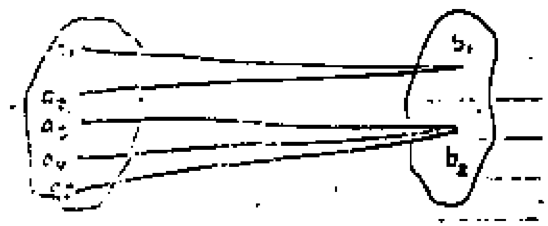
Función sobre (o suprayectiva). Una función $f: A \to B$ es "sobre" si para todo elemento $y \in B$ existe al menos un elemento $x \in A$ tal que $f(x) = y$.

Dicho de otro modo, una función es "sobre" cuando su rango y su codominio son el mismo conjunto; esto es:

$f(A) = B$

Ejemplos: $A = \{a_1, a_2, a_3, a_4, a_5\}$; $B = \{b_1, b_2\}$

1)



2) $f: R^1 \to R^1$ tal que $f(x) = x^2$

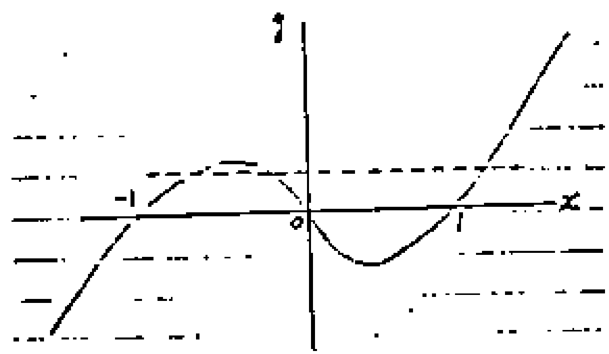
no es sobre: $-1 \in R^1$; pero $f(x) = -1$ no existe

3) $f: R^1 \to R^1$ tal que $g(x) = x^3 - x$

sí es sobre; puesto que el rango es R^1

Geométricamente: No existe ninguna línea horizontal que no corte la gráfica de una función sobre por lo menos en un punto. ($f: R^1 \to R^1$)

Como ejemplo, consideremos la gráfica que corresponde al ejercicio 3) último:



Función biyectiva. Una función $f: A \to B$ se denomina "biyectiva" si es, al mismo tiempo, biunívoca y sobre.

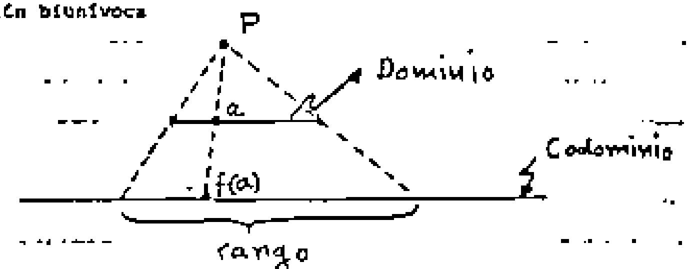
Ejemplo:

$f: R^1 \to R^1$; siendo $f(x) = x^3$

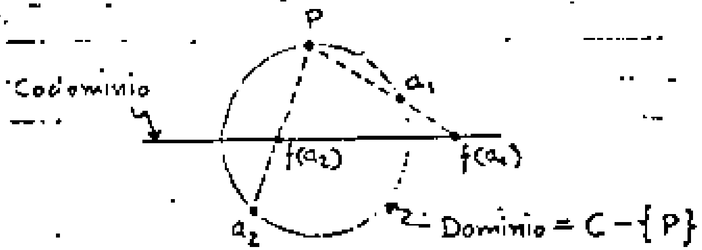
f si es biyectiva

1a. Observación. El concepto de función que se ha expuesto es tan amplio, que permite designar (y operar) funciones de muy variada índole, que en nada se parecen a las funciones numéricas.

Ejemplo 1) Función biunívoca



2) Función biyectiva, (biunívoca y sobre)



Por el punto P se traza una semirrecta, la cual, al cortar el dominio y el codominio define los puntos a_i y $f(a_i)$ respectivamente.

Nótese que en funciones biunívocas, dos pares ordenados distintos no pueden tener el mismo segundo elemento:

si $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$

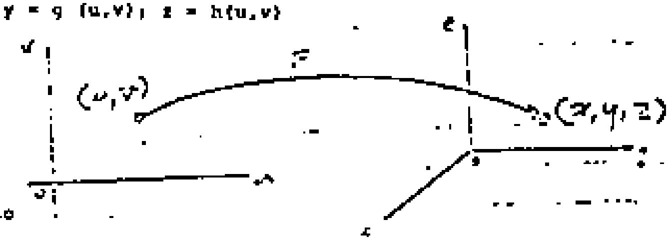
3) Curvas $f: E^1 \to E^m$

a) Si $m = 2$; siendo: $x = f(t)$; $y = g(t)$; $a \leq t \leq b$

b) Si $m = 3$; $x = f_1(t)$; $y = f_2(t)$, $z = f_3(t)$; $a \leq t \leq b$

4) Superficies. $f: E^2 \to E^3$

$x = f(u,v)$; $y = g(u,v)$; $z = h(u,v)$

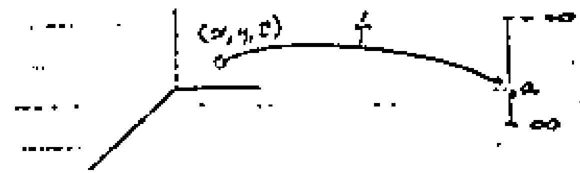


5) Generalizando

$f: E^n \to E^m$

$m \lesseqgtr n$

a) $f: E^3 \to E^1$; $f(x,y,z) = a$



b) En general, a las funciones $f: E^n \to E^m$ se les suele denominar "Transformaciones"

$$f: \begin{cases} y_1 = f_1\,(x_1, x_2, \ldots, x_n) \\ y_2 = f_2\,(x_1, x_2, \ldots, x_n) \\ - - - - - - - - - - - - \\ y_m = f_m\,(x_1, x_2, \ldots, x_n) \end{cases}$$

Donde: $n > m$, $n = m$, o $n < m$

2a. Observación. La gráfica de una función $f: A \to B$ se define del siguiente modo:

$f^* = \{(a,b) \mid a \in A,\ b = f(a)\}$

Resultando de la definición que:

a) $f^* \subseteq A \times B$

b) Para cada $a \in A$ existe una pareja ordenada $(a,b) \in f^*$

c) Si $(a,b) \in f^*$ y $(a,c) \in f^* \Longrightarrow b = c$

d) La gráfica f^* puede ser un conjunto en un espacio de más de 3 dimensiones, en cuyo caso, sólo se concibe en forma abstracta.

Ejemplo. Sea $f: D \to E^1$; siendo $D \subset E^3$:

$f^* = \{(a,b) \mid a \in D,\ b = f(a)\}$

Si $a = (x,y,z)$; $b = f(a)$

$\Longrightarrow b = f(x,y,z) \Longrightarrow f^* \subseteq E^3 \times E^1$

$\Longrightarrow (a,b) = (x,y,z,b)$

$\Longrightarrow$ La gráfica f^*, de la función f, resulta ser un conjunto de un espacio tetradimensional.

$f^* = \{(x,y,z,b) \mid (x,y,z) \in D,\ b = f(x,y,z)\}$

Aun cuando la construcción visual de la gráfica no es posible en casos como el que ilustra el ejemplo anterior, la noción de "gráfica" sigue siendo una idea útil y muy conveniente.

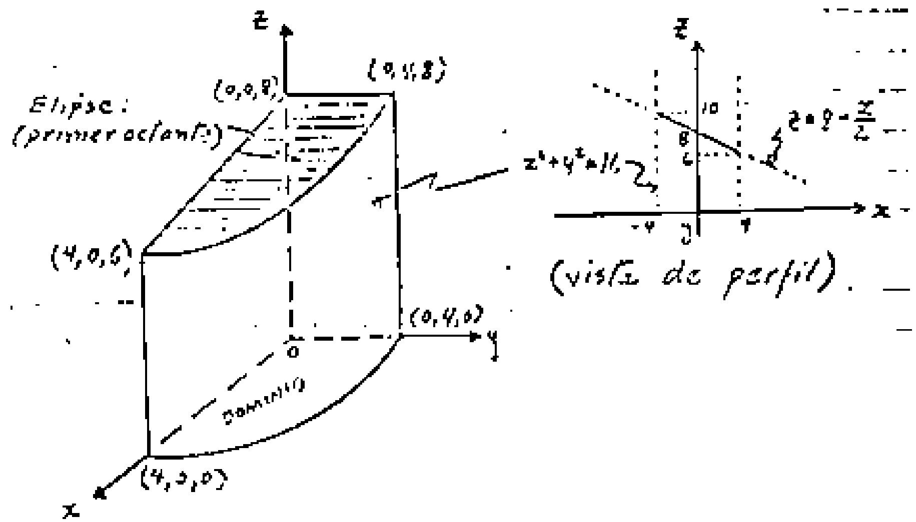
3a. Observación. Además de la forma 4] anterior para definir en forma paramétrica una superficie, la siguiente función representa una superficie.

$f: D \to E^1$; donde el dominio $D \subset E^2$

Ejemplo:

$f(x,y) = 8 - \frac{x}{2}$; $D = \{(x,y) \mid x^2 + y^2 \le 16\}$

Resulta la superficie una elipse cuyo plano está en el plano proyectante: $z = 8 - \frac{x}{2}$; estando el perímetro de la elipse definido por la intersección del plano anterior con el cilindro $x^2 + y^2 = 16$.



La función no es biunívoca; puesto que, para x = constante, z tiene el mismo valor, independientemente del valor de y; por ejemplo:

$(x_1, y_1) = (2,1) \Rightarrow z = 7$; $(x_1, y_2) = (2,2) \Rightarrow z = 7$

La función no es sobre, pues siendo el codominio E^1, el rango es:

$[6,10] \neq E^1 = (-\infty,\infty)$

Desde el punto de vista geométrico, una función $f: E^2 \to E^1$; $z = f(x,y)$ será biunívoca (uno a uno) si todo plano paralelo al plano x,y de ecuación $z = z_0$ ($z_0 \in$ rango de f), sólo contiene un punto de la gráfica de f.

La función f será suprayectiva (sobre), si el plano $z = z_1$ (donde z_1 toma todos los valores correspondientes al codominio), corta la gráfica de f, por lo menos, en un punto.

Las ideas anteriores pueden generalizarse para ser aplicadas a funciones cuyo dominio sea E^n; siendo su codominio E^1:

$f: E^n \to E^1$; $(n \in N)$

Igualdad de dos funciones. Si dos funciones f y g están definidas en el mismo dominio D, y se tiene además que para toda $a \in D$, $f(a) = g(a)$, se afirma que las funciones son iguales: $f = g$.

Ejemplos:

a) Sea $f: R^1 \to R^1$; $g: R^1 \to R^1$

$f(x) = x^2$; $g(y) = y^2 \Rightarrow f = g$

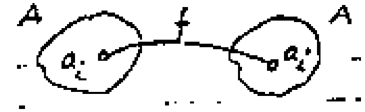
b) Sea $f: R^1 \to R^1$; $g: C \to C$

$f(x) = x^2$; $g(y) = y^2$

$f \neq g$ porque no tienen el mismo dominio

$g(i) = -1$; $f(i)$ no existe

Función identidad. Sea la función $f: A \to A$ definida por la fórmula $f(x) = x$; o sea, se hace corresponder a un elemento a, el mismo elemento a como su imagen.



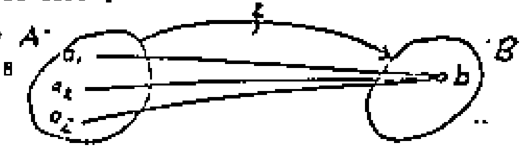
Notación:

$f = I$ ó $f = I_A$

$A \xrightarrow{I_A} A$

Si $f: R^1 \to R^1$; $R^1 \xrightarrow{I} R^1$

Función constante. Se dice que $f: A \to B$ es una función "constante", si todo $a_i \in A$ tiene la misma imagen $b \in B$



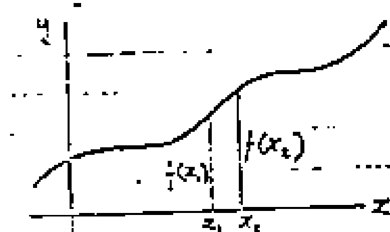
Ejemplo: Sea $f(x) = 2$; siendo $f(x) = x^2 + \text{sen}x$

$\Rightarrow x^2 + \text{sen}x = 2$

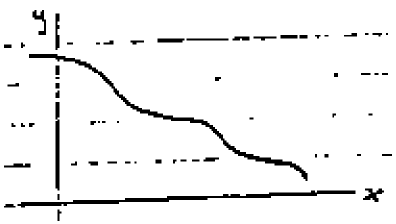
Función monótona creciente. Se dice que una función f es "monótona creciente" en un intervalo, si para dos puntos cualesquiera x_1 y x_2 del intervalo, tales que $x_1 < x_2$ se verifica que $f(x_1) \leq f(x_2)$.

Si $f(x_1) < f(x_2)$ la función se denomina "estrictamente creciente".

Función monótona decreciente. Su definición es semejante a la ante



FUNCION MONOTONA CRECIENTE



FUNCION MONOTONA DECRECIENTE

rior.

Funciones cóncavas y funciones convexas. Sea $f(x)$ una función definida en un intervalo (a,b). Si para dos puntos cualesquiera $(a \leq x_1 < x_2 \leq b)$ x_1 y x_2 se verifica: $\forall\, x_1, x_2 \in (a,b)$.

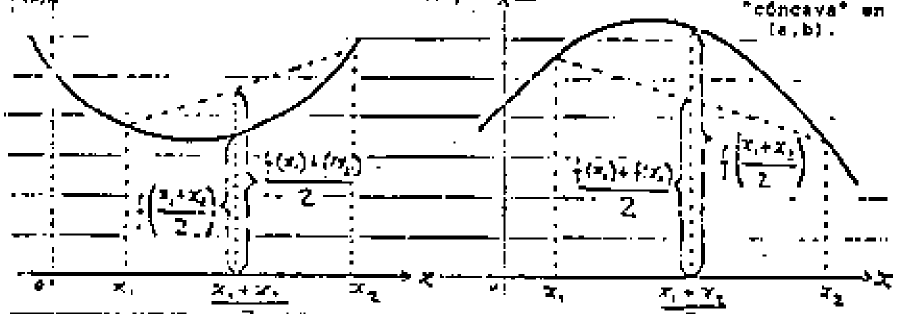
$$f\left(\frac{x_1+x_2}{2}\right) < \frac{f(x_1)+f(x_2)}{2}$$

se dice que la función es "convexa" en (a,b)

Si

$$f\left(\frac{x_1+x_2}{2}\right) > \frac{f(x_1)+f(x_2)}{2}$$

La función se denomina "cóncava" en (a,b).



Función convexa   Función cóncava

Función inversa. Si $f: A \rightarrow B$ es una función biyectiva (biunívoca y sobre), entonces existe la función inversa f^{-1}; la cual se define:

$f^{-1}: B \rightarrow A$

dado que para cada elemento $b \in B$ existe uno, y sólo uno de los elementos de A, tales que:

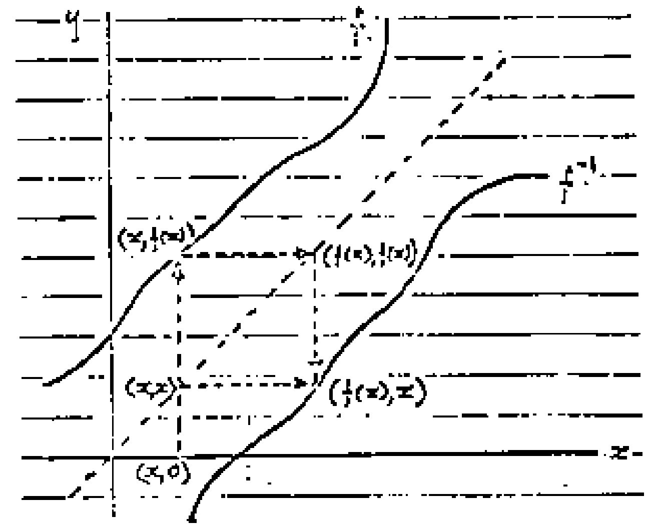
$f^{-1}(b) \in A$

$f^{-1} = \{(f(a), a) \mid a \in A\}; \quad A = D_f$

Es interesante observar que, en término del lenguaje de las relaciones estudiadas anteriormente, podemos decir que las funciones f y f^{-1}

consideradas conjuntamente, constituyen una "relación simétrica" en la cual f no contiene dos pares ordenados con sus primeros elementos iguales, y por lo tanto, f^{-1} tampoco; asimismo, ni f ni f^{-1} deben tener pares ordenados cuyos segundos elementos sean iguales.

Ejemplo 1)



Nótese que la función inversa f^{-1} tiene como dominio el conjunto B; siendo su codominio el conjunto A.

Ejemplo 2) Sea $f: R^1 \rightarrow R^1$

tal que

$f(x) = x^3$

por ser f biyectiva, existe f^{-1}

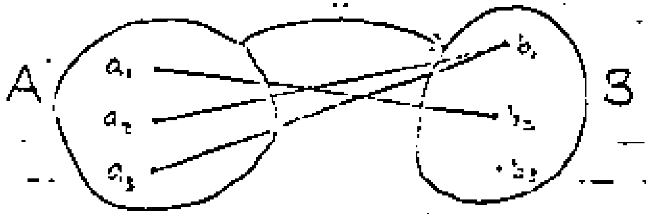
$f^{-1}(x) = \sqrt[3]{x}$

Inverso de una función. Sea $f: A \rightarrow B$

Se define como "inverso" de f al siguiente conjunto:

$$f^{-1}(b) = \{a \mid a \in A,\ f(a) = b\}$$

Ejemplo:



Resulta:

$$f^{-1}(b_1) = \{a_2, a_3\}$$

$$f^{-1}(b_2) = \{a_1\}$$

$$f^{-1}(b_3) = \emptyset$$

Observaciones: $f^{-1}(b) \subseteq A$

En este ejemplo la función inversa $f^{-1}: B \to A$ no existe, puesto que f no es biyectiva.

Operaciones con funciones. Definiremos, para las funciones, operaciones semejantes a las conocidas operaciones para los números reales. Estudiaremos para funciones de valores reales (funciones reales) tres operaciones básicas, a saber: suma, multiplicación, y composición, así como las propiedades de éstas.

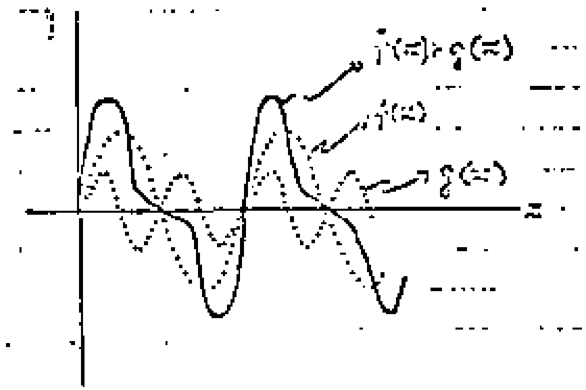
Suma (o adición) de funciones. Sean f y g dos funciones reales, cuyos dominios designaremos como D_f y D_g respectivamente. Se define la suma f + g como:

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

o sea:

$$f + g = \{(x,\ f(x) + g(x)) \mid x \in D_f \cap D_g\}$$

Geométricamente:



Análogamente: $(f - g)x = f(x) - g(x)$

Ejemplo 1) Sean:

$$f = \{(1,1),\ (2,3),\ (3,3),\ (4,2),\ (5,4)\}$$

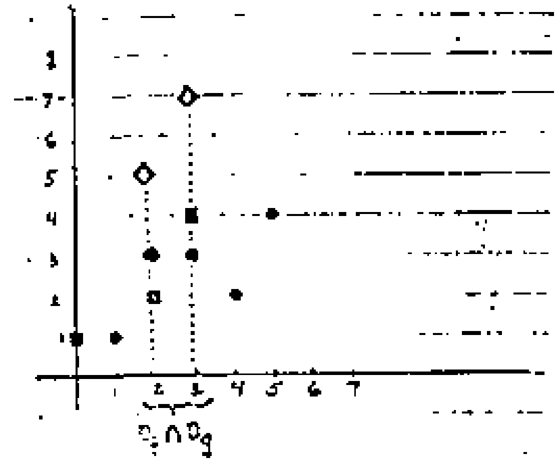
$$g = \{(0,1),\ (2,2),\ (3,4)\}$$

Hallar f + g

Solución: $D_f \cap D_g = \{2,3\}$

$$\Longrightarrow f + g = \{(2,5),\ (3,7)\}$$

Gráficamente:



Obsérvese que la adición de funciones no es lo mismo que la suma de pares ordenados (vectores en el plano).

Ejemplo 2) Resolver la siguiente inecuación:

$$|x - a| < |x - 3a| \qquad \ldots(1)$$

$a \in R^{+}$; $-3a < x < 4a$; $a > 0$

Solución:

de (1): $|x - a| - |x - 3a| < 0 \qquad \ldots(2)$

Haciendo:

$$\begin{cases} |x - a| - |x - 3a| = y \\ \left.\begin{array}{l} |x - a| = y_1 \\ |x - 3a| = y_2 \end{array}\right\} \Rightarrow y_1 - y_2 = y \end{cases}$$

de la definición de valor absoluto:

$$y_1 = |x - a| = \begin{cases} x - a, & \text{si } x - a \geq 0 \\ -x + a, & \text{si } x - a < 0 \end{cases}$$

$$y_2 = |x - 3a| = \begin{cases} x - 3a, & \text{si } x - 3a \geq 0 \\ -x + 3a, & \text{si } x - 3a < 0 \end{cases}$$

gráficamente:

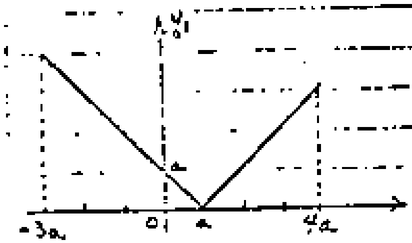


Figura (3)

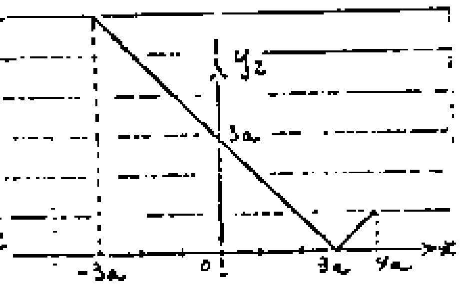


Figura (4)

Restando las funciones y_1 y y_2 gráficamente, mediante la superposición de los diagramas en las figuras (3) y (4):

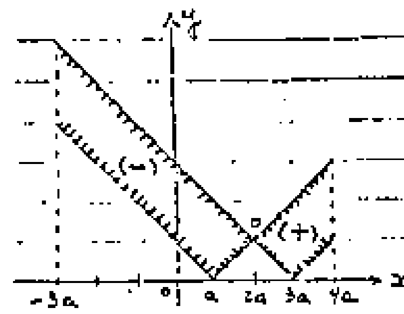


Figura (5)

ó bien:

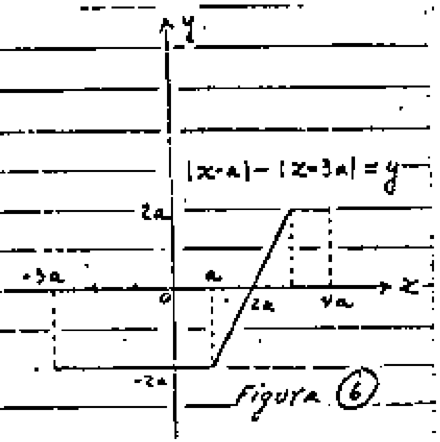


Figura (6)

De la figura (6) se puede obtener la solución de la inecuación (1):

Dado que: si $x = 0 \Rightarrow$ de 1

$|0 - a| - |0 - 3a| = |-a| - |-3a| = a - 3a = -2a = y_1 < 0$

$\Rightarrow$ de (2), $x = 0$ satisface (2)

Observando de la figura (6) que las ordenadas son negativas ($y < 0$) en el intervalo: $-3a < x < 2a$; serán éstos los valores de x que satisfarán la desigualdad (1); ésto es, la solución de la inecuación (1) resulta:

$$-3a < x < 2a, \in (-3a, 4a)$$

Ejemplo 3) Si: $f = x\,|x-2|$; $g = (x-2)\,|x|$

Hallar todos los valores de x para los cuales se verifica que $f = g$; $f = -g$

Solución:

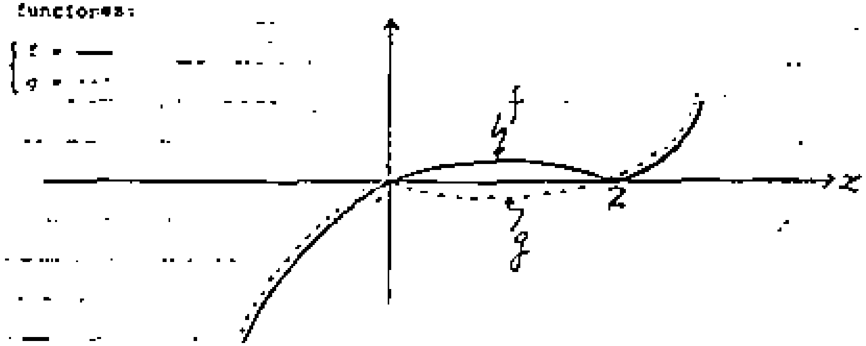
$$f = \begin{cases} x^2 - 2x; & \text{si } x \geq 2 \\ -x^2 + 2x; & \text{si } x < 2 \end{cases} \qquad g = \begin{cases} x^2 - 2x; & \text{si } x \geq 0 \\ -x^2 + 2x; & \text{si } x < 0 \end{cases}$$

Por lo tanto: $f = g$ en los siguientes intervalos:

$(-\infty, 0] \cup [2, \infty)$

$f = -g \in (0,2)$

Lo anterior se visualiza fácilmente observando las gráficas de ambas funciones:



Estudio de las concavidades:

$$\text{Si: } \begin{cases} u = x^2 - 2x \Rightarrow u' = 2x - 2 \Rightarrow u'' = 2 > 0 \Rightarrow \smile \\ v = -x^2 + 2x \Rightarrow v' = -2x + 2 \Rightarrow v'' = -2 < 0 \Rightarrow \frown \end{cases}$$

Ejemplo 4) Resolución gráfica de la ecuación general de Tercer grado con coeficientes $\in R^1$:

$$x^3 + \;x^2 + bx + c = 0 \quad (1)\ ;\ a,b,c \in R^1$$

Solución:

Observaciones preliminares:

1.- Aun cuando puede encontrarse la solución analítica empleando las fórmulas de Cardano, en la práctica se encuentra que este método presenta inconveniencias, como las que se citan a continuación:

i) La sustitución numérica en las fórmulas conduce a cálculos engorrosos.

ii) Si la ecuación tiene raíces racionales (o enteras) éstas se obtienen en forma de suma de números irracionales.

iii) Aun siendo las tres raíces reales, éstas se presentan como sumas de números complejos.

2.- Siempre es posible, mediante un cambio de variable, eliminar de la ecuación (1), el término en x^2; en efecto:

Haciendo $x = z + \alpha$ (2) en (1):

$$(z^3 + 3\alpha z^2 + 3\alpha^2 z + \alpha^3) + a(z^2 + 2\alpha z + \alpha^2) + b(z + \alpha) + c =$$
$$= z^3 + z^2(3\alpha + a) + z(3\alpha^2 + 2a\alpha + b) + (\alpha^3 + a\alpha^2 + b\alpha + c) = 0$$

Si:

$$(3)\begin{cases} 3\alpha + a = 0 \Rightarrow \alpha = -\frac{a}{3} \\ 3\alpha^2 + 2a\alpha + b = p \\ \alpha^3 + a\alpha^2 + b\alpha + c = q \end{cases}$$

Considerando (2) y (3) en (1), resulta:

$$z^3 + pz + q = 0$$

$\therefore$ la solución de (4) conduce a la solución de (1):

de (4): $pz + q = -z^3$ (5)

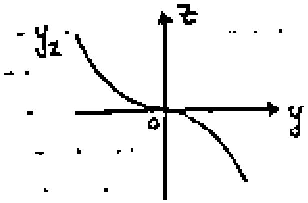
Si se hacen:

$pz + q = y$; $-z^3 = y$; (6)

queda resuelta la ecuación (¢) si $y_1 = y_2$; pero: $pz + q = y_1$ tiene por gráfica una línea recta, de pendiente p; con ordenada al origen q, siendo $-z^3 = y_2$ gráficamente una parábola cúbica fija.

Los hechos anteriores sugieren el siguiente procedimiento gráfico.

I) Dibújese, con la mayor precisión posible la curva $y_2 = -z^3$ en papel milimétrico. La forma general es:



II) A la misma escala, y sobre la gráfica que se dibujó en I), empleando los mismos ejes (y,z) trácese la recta de parámetros p = pendiente, q = ordenada al origen. Se sugiere que $y_2 = -z^3$ se trace con tinta; en tanto la recta y puede dibujarse con lápiz; de modo que la hoja milimétrica pueda servir muchas veces, con sólo borrar la recta en cada caso, el diagrama queda listo para usarse con nuevos parámetros p,q

III) Las raíces buscadas serán las abcisas de los puntos de intersección de la curva y_2 y la recta y_1

Si la recta corta a la curva en tres puntos, entonces la ecuación (1) tendrá tres raíces reales.

Obsérvese que la recta tiene que cortar a la curva por lo menos en un punto, puesto que y_2 es función biyectiva:

$y_2 : R^1 \to R^1$

Por tener 1 coeficientes $a,b,c \in R^1$, las raíces complejas tienen que ser pares conjugados; o sea, si $r + si$ es raíz también tiene que ser raíz $r - si$. $(i = \sqrt{-1})$

Multiplicación de funciones. Sean f y g dos funciones de valores reales, con dominios D_f y D_g respectivamente. Se define la multiplicación f g como sigue:

$(fg)(x) = f(x)\cdot g(x)$

o sea:

$fg = \{(x, f(x)\cdot g(x)) \mid x \in D_f \cap D_g\}$

Ejemplo: Sean f y g las mismas funciones definidas en el último ejemplo.

$\Rightarrow \quad fg = \{(2,6),\ (3,12)\}$

Propiedades de las operaciones de suma y multiplicación de funciones.

Sea S el conjunto de funciones reales de variable real: $f: R^1 \to R^1$

A_1) Cerradura: Si f y g $\in S \Rightarrow f + g \in S$

A_2) Conmutatividad: $f + g = g + f$; f y g $\in S$

A_3) Asociatividad: $(f + g) + h = f + (g + h)$; $f,g,h \in S$

A_4) Existe un elemento neutro único tal que:

para todo $f \in S$, $f + 0 = f$

M_1) Cerradura: Si f y g $\in S \Rightarrow f\,g \in S$

M_2) Conmutatividad: $fg = gf$; f y g $\in S$

M_3) Asociatividad: $(fg)h = f(gh)$; $f,g,h \in S$

M_4) Existe un elemento neutro único en S tal que:

para todo $f \in S$, $f1 = f$

D) Distributividad: $f(g + h) = fg + fh$; $f, g, h \in S$

Observaciones: $\emptyset \in S$, pues de otro modo no se cumplirían A_1 y M_1

Las funciones reales, de variable real poseen todas las propiedades de un campo, excepto la existencia de aditivo inverso único y de recíproco único (A_4 y A_7).

Si el dominio de f no es todo el conjunto de los reales, entonces no existe función g tal que $f + g = 0$ o $fg = 1$, dado que el dominio de las funciones constantes 0 y 1 es R^1; pero D_{f+g} y D_{fg} no puede ser R^1.

<u>Composición de funciones</u>. Algunos autores consideran la composición de funciones como un producto (o multiplicación) de funciones.

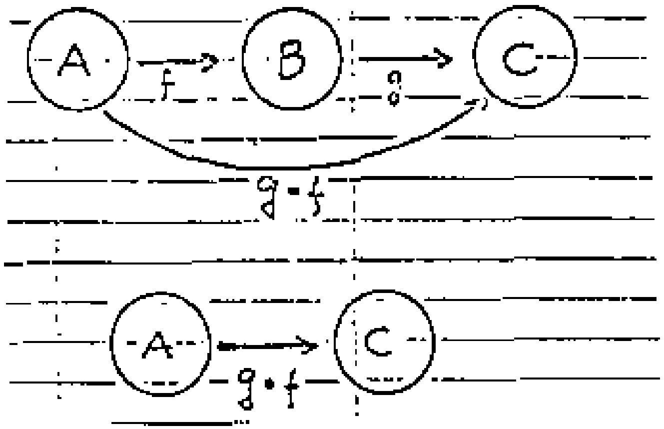
Definición: Sean $f: A \to B$; $g: B \to C$. $g \circ f$ denominada g composición f es una función cuyo dominio son los elementos $a \in A$ tales que $f(a) \in B$

$$(g \circ f)(a) = g(f(x))$$

$(g \circ f): A \to C$

Obsérvese que para que quede definida $g \circ f$ no es necesario que f y g sean funciones reales de variable real.

Gráficamente:



Ejemplos: Determinar $g \circ f$ y $f \circ g$ en cada uno de los siguientes ejemplos:

1) $f = \{(1,2), (2,3), (3,5), (4,7)\}$

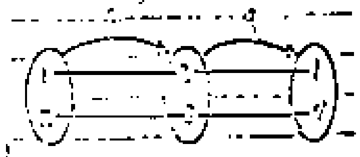
$g = \{(0,3), (1,2), (2,1), (3,4)\}$

a) Cálculo de $g \circ f$:

Determinación del dominio $D_{g \circ f}$:

$1 \in D_{g \circ f}$ puesto que $f(1) = 2 \in D_g$

$2 \in D_{g \circ f}$ puesto que $f(2) = 3 \in D_g$



$g \circ f = \{(1,1), (2,4)\}$

b) Cálculo de $f \circ g$

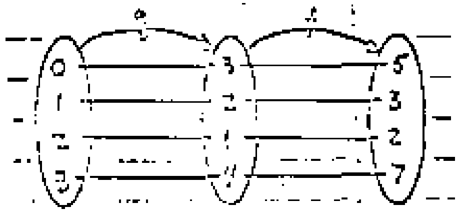
Determinación del dominio $D_{f \circ g}$:

$0 \in D_{f \circ g}$ porque $f(0) \in D_f$

$1 \in D_{f \circ g}$ porque $f(1) \in D_f$

$2 \in D_{f \circ g}$ porque $f(2) \in D_f$

$3 \in D_{f \circ g}$ porque $f(3) \in D_f$



$f \circ g = \{(0,5), (1,3), (2,2), (3,7)\}$

2) $f(x) = x^3$; $g(x) = x + 3$

Solución:

$(g \circ f)(x) = g(f(x)) = g(x^3) = x^3 + 3$

$(f \circ g)(x) = f(g(x)) = f(x + 3) = (x + 3)^3$

Vemos que, en general, la composición de funciones no es una opera - ción conmutativa.

3) $f: E^1 \to E^2$; donde $f(t) = (1 + t,\ 2 - t)$

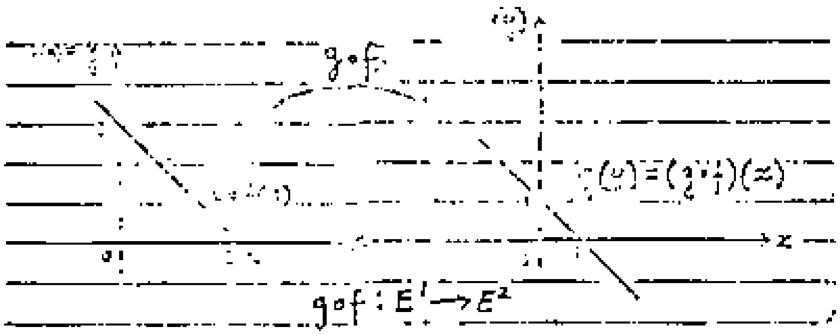
$g: E^2 \to E^2$; siendo $g(x,y) = (x - 1,\ y - 1)$

Solución:

$(g \circ f)(x) = g(f(t)) = g(1 + t,\ 2 - t) = (t,\ 1 - t)$

Nótese que $(f \circ g)(t)$ no existe, puesto que el dominio de f no es igual al codominio de g.

Geométricamente: $f(t) = (1 + t,\ 2 - t)$ representa una recta, cuyas ecuaciones paramétricas son: $x = 1 + t$, $y = 2 - t$; siendo la recta $x + y = 3$; en tanto que $(g \circ f)(t) = (t, 1 - t)$ nos da las ecuaciones paramétricas $x = t$; $y = 1 - t$, que corresponden a la recta $x + y = 1$



4) $f(x) = \sqrt{x}$; $g(x) = 3x + 1$

Solución:

$(g \circ f)(x) = g(f(x)) = g(\sqrt{x}) = 3\sqrt{x} + 1$

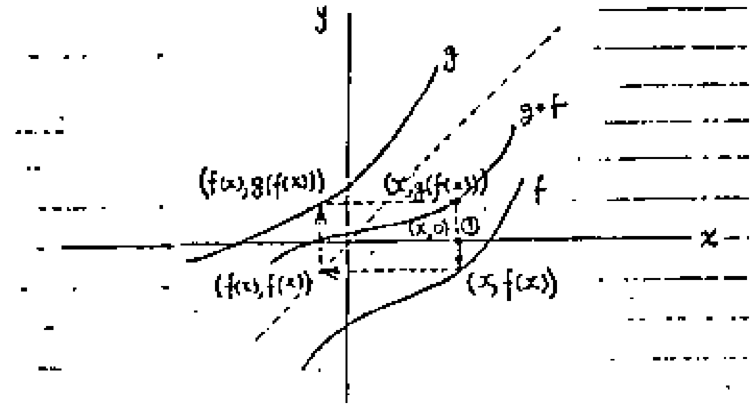
por lo cual $D_{g \circ f} = [0, \infty)$

$(f \circ g)(x) = f(g(x)) = f(3x + 1) = \sqrt{3x + 1}$

por lo cual $D_{f \circ g} = [-\frac{1}{3}, \infty)$

Construcción de la gráfica de la función compuesta $g \circ f$ a partir de las gráficas de f y g.

Sean f y g ambas funciones reales de variable real; siendo sus gráficas las curvas que se muestran en la siguiente figura:



Principiese a construir la gráfica de $g \circ f$ en el punto (1) de coordenadas $(x,0)$ de f.

Observación:

Si f es una función monótona, y además $f = f^{-1} \Rightarrow$ una de dos posibilidades:

$$\begin{cases} f(x) = x \text{ para toda } x; \text{ o} \\ f(x) = -x \text{ para toda } x \end{cases}$$

Demostración:

Aun cuando la intuición geométrica parece corroborar la afirmación, ésta se justificará mediante el siguiente razonamiento, ya que, como se comprobará en un ejemplo después de éste, la intuición geométrica puede fallar:

1o. Supóngase que f es monótona creciente:

a) Si $f(x) < x \Rightarrow f(f(x)) < f(x) < x$

$\Rightarrow f(f(x)) < x$; pero $f(f^{-1}(x)) = f(f(x)) = x$

por hipótesis; de ahí que $f(f(x)) < x$ es imposible.

b) si $f(x) > x \Rightarrow f(f(x)) > f(x) > x$

$\Rightarrow f(f(x)) > x$; lo que por la misma razón anterior es imposible.

Por lo tanto se concluye: Si f es monótona creciente; siendo $f = f^{-1} \Rightarrow f(x) = x$

2o. Considérese que f es monótona decreciente;

Razonamientos similares a los anteriores nos llevan a la conclusión que: Si f es monótona decreciente; teniéndose además que $f = f^{-1} \Rightarrow f(x) = -x$ para toda x.

(puede hacerse $f = -f$ en el caso 1o. para demostrar el 2o.)

Otra observación importante:

Aun cuando la intuición sugiere que sólo las funciones estrictamente monótonas, $f: A \to B$; donde $A \subset E^1$; $B \subset E^1$, pueden tener inversa $f^{-1}: B \to A$; el siguiente ejemplo nos demuestra que existen funciones que sin ser monótonas, tienen inversa; en efecto:

Sea:

$$f(x) = \begin{cases} x, & \text{si } x \in Q \\ -x, & \text{si } x \in Q' \end{cases}$$

Demostrar:

i) f no es monótona en ningún intervalo

ii) $f = f^{-1}$

Solución:

Sea I cualquier intervalo con más de un punto, escójanse

$$\begin{cases} x_1, x_2 \in Q \ni x_1 < x_2 \\ x_3, x_4 \in Q' \ni x_3 < x_4 \end{cases}$$

$$\Rightarrow \begin{cases} f(x_1) - f(x_2) = x_1 - x_2 < 0 \to \text{creciente} \\ f(x_3) - f(x_4) = -x_3 + x_4 > 0 \to \text{decreciente} \end{cases}$$

$\Rightarrow$ f no es monótona $\in$ I

ii) si $x \in Q \Rightarrow f(f(x)) = f(x) = x \Rightarrow f = f^{-1}$

si $x \in Q' \Rightarrow f(f(x)) = f(-x) = -(-x) = x \Rightarrow f = f^{-1}$

$\Rightarrow$ Para toda x: $f(f(x)) = x$

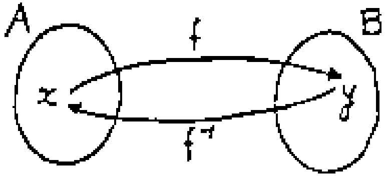
$\Rightarrow f = f^{-1}$

(Del ejemplo anterior, puede concluirse que la f de este ejemplo no puede ser monótona, ya que f no es tal que $f(x) = x$; ni $f(x) = -x$ para toda x)

Teorema. Sea f una función biyectiva tal que: $f: A \to B$; $f^{-1}: B \to A$; se afirma que:

$$f \circ f^{-1} = f^{-1} \circ f = I$$

Demostración:



Sea $x \in A$; siendo $f(x) = y \in B \Longrightarrow f^{-1}(y) = x$

$\Longrightarrow f^{-1}(y) = f^{-1}(f(x)) = (f^{-1} \circ f)\ (x) = x$

$\Longrightarrow f^{-1} \circ f = I$

Asimismo:

si:

$f(x) = y;\ x = f^{-1}(y) \Longrightarrow f(f^{-1}(y)) = y$

$\Longrightarrow f \circ f^{-1} = I$

$\therefore f^{-1} \circ f = f \circ f^{-1} = I$

Ejemplo:

Sabiendo que si: $f(x) = e^x$; $f^{-1}(x) = \ln x$; con base en el teorema anterior, demostrar que:

$\ln e^x = e^{\ln x} = x$

Solución: del teorema anterior:

$(f \circ f^{-1})(x) = x = f(f^{-1}(x)) = f(\ln x) = e^{\ln x}$

$(f^{-1} \circ f)(x) = x = f^{-1}(f(x)) = f^{-1}\ (e^x) = \ln e^x$

$\therefore x = e^{\ln x} = \ln e^x$

<u>Teorema</u>. Sea S el conjunto de funciones reales de variable real. Se verifica lo siguiente: Para toda $f, g, h \in S$

C_1) Si f y $g \in S$, $f \circ g \in S$

C_2) En general, $f \circ g \neq g \circ f$

C_3) $(f \circ g) \circ h = f \circ (g \circ h)$

C_4) Existe un único elemento $I \in S$ tal que:

$f \circ I = I \circ f = f$

D_1) $(f + g) \circ h = f \circ h + g \circ h$

pero $f \circ (g + h) \neq f \circ g + f \circ h$, en general.

D_2) $(fg) \circ h = (f \circ h)(g \circ h)$

pero $f \circ (gh) \neq (f \circ g)\ (f \circ h)$, en general.

Demostraciones:

C_1) Es una consecuencia de la definición de composición de funcio - nes.

C_2) Esto quedó evidenciado en los ejemplos resueltos anteriormente.

C_3) Demostraremos primero que los dominios son iguales:

$$D_{(fog)oh} = D_{fo(goh)}$$

$$\begin{aligned} D_{(fog)oh} &= \{x \mid x \in D_h \text{ y } h(x) \in D_{fog}\} \\ &= \{x \mid x \in D_h,\ h(x) \in D_g \text{ y } g(h(x)) \in D_f\} \\ &= \{x \mid x \in D_{goh} \text{ y } (goh)(x) \in D_f\} \\ &= D_{fo(goh)} \end{aligned}$$

En seguida se demostrará C_3. En efecto, para toda x en el dominio común de las funciones se tiene:

$$\begin{aligned} ((fog)oh)(x) &= (fog)(h(x)) \\ &= f(g(h(x))) \\ &= f((goh)(x)) \\ &= (fo(goh))(x) \end{aligned}$$

Con lo cual queda demostrada la asociatividad en la composición de funciones.

Obsérvese que en la demostración de C_3 no ha sido necesario estipular que las funciones f,g y h sean funciones reales de variable real; por consiguiente, la ley asociativa en la composición de funciones es cierta para funciones en general.

C_4) La función identidad I es la función neutra con respecto a la composición de funciones. Es necesario demostrar tanto que $f \circ I = f$ como también que $I \circ f = f$; puesto que no es válida la conmutatividad para la composición de funciones.

Los dominios de $f \circ I$ e $I \circ f$ son el mismo; siendo en ambos casos D_f.

Para toda $x \in D_f$ se tiene:

$$(foI)(x) = f(I(x)) = f(x)$$

$$(Iof)(x) = I(f(x)) = f(x)$$

$$\Rightarrow foI = Iof = f$$

En seguida se demostrará que I es único en S.

Supongamos que existe otro elemento neutro I':

$I' = I'oI = I$ Luego I es único

D_1) Demostraremos primero que los dominios son los mismos:

$$D_{(f+g)oh} = D_{foh + goh}$$

$$\begin{aligned} D_{(f+g)oh} &= \{x \mid x \in D_h \text{ y } h(x) \in D_{f+g}\} \\ &= \{x \mid x \in D_h \text{ y } h(x) \in D_f \cap D_g\} \\ &= \{x \mid x \in D_h \text{ y } h(x) \in D_f\} \cap \{x \mid x \in D_h \text{ y } h(x) \in D_g\} \\ &= D_{foh + goh} \end{aligned}$$

Establecido lo anterior, procedemos a demostrar D_1.

Sea x en el dominio común de las funciones:

$$\begin{aligned} ((f+g)oh)(x) &= (f+g)(h(x)) \\ &= f(h(x)) + g(h(x)) \\ &= (foh)(x) + (goh)(x) \\ &= (foh + goh)(x) \end{aligned}$$

D_2) La demostración de esta segunda ley distributiva se efectúa de manera semejante a la seguida en la demostración de la primera ley distributiva D_1).

Ejemplo en que qud se hace ver que:

$$f \circ (g + h) \neq fog + foh$$

$$fo(gh) \neq (fog)(foh)$$

Sea $f = 2$, $g = h = I$

$\Rightarrow 2 = 2o(I + I)$, en tanto que

$\Rightarrow 4 = 2oI + 2oI = 2 + 2$

por otro lado:

$2 = 2o(II)$; pero,

$4 = (2oI)(2oI) = 2\cdot 2$

En cambio:

$(2 + I)oI = 2 + I$

$2oI + IoI = 2 + I$

$(2I)oI = 2I$

$(2oI)(IoI) = (2)(I) = 2I$

La composición de funciones no sólo se circunscribe a funciones rea les de variable real; de ahí resulta, por ejemplo, que:

Si $f: A \to B \Rightarrow$

a) $I_B \circ f = foI_A = f$

b) Si f es biyectiva, existe f^{-1}

$\Rightarrow f^{-1} of = I_A$; $fof^{-1} = I_B$

siendo asimismo válida la proposición recíproca:

c) Si $f: A \to B$ y $g: B \to A$, satisfacen:

$gof = I_A$ y $fog = I_B$

Entonces existe $f^{-1}: B \to A$ y $g = f^{-1}$

d) Si $f: A \to B$, $g: B \to C$ y $h: C \to D$

se verifica la ley asociativa:

$(hog)of = ho(gof)$

Demostración:

Para toda $a \in A$:

$((hog)of)(a) = (hog)[f(a)] = h(g(f(a)))$

$(ho(gof))(a) = h((gof)(a)) = h(g(f(a)))$

$\Rightarrow (hog)of = ho(gof)$

(independientemente de que A y B sean iguales a R^1)

e) Si $f: A \to B$ y $g: B \to C$ son funciones biyectivas, entonces:

$(gof)^{-1} = f^{-1}og^{-1}$; o sea:

$(gof)^{-1}: C \to A$ existe, y es igual a

$f^{-1}og^{-1}: C \to A$

Demostración: de c) haremos ver que:

$(f^{-1} o\ g^{-1})o(gof) = I_A$

además:

$(gof)o(f^{-1} o\ g^{-1}) = I_B$

En efecto, de d):

$$\begin{aligned}(f^{-1} o\ g^{-1})o(gof) &= f^{-1}o(g^{-1}o(gof)) \\ &= f^{-1}o((g^{-1}og)of) \\ &= f^{-1}o(Iof) \\ &= f^{-1}of \\ &= I_A\end{aligned}$$

Asimismo:

$$\begin{aligned}(gof)o(f^{-1}og^{-1}) &= go(fo(f^{-1}og^{-1})) \\ &= go((fof^{-1})og^{-1}) \\ &= go(Iog^{-1}) \\ &= gog^{-1} \\ &= I_B\end{aligned}$$

f) Generalización del resultado e) anterior:

$(f_1 o\ f_2 o \ldots o f_n)^{-1} = f_n^{-1} o \ldots o f_2^{-1} o\ f_1^{-1}$

Demostración: por inducción matemática:

i) vale para $n = 2$ $(f_1 \circ f_2)^{-1} = f_2^{-1} \circ f_1^{-1}$, según problema e) anterior.

ii) supongamos que vale para $n = k$:

$$(f_1 \circ f_2 \circ \dots \circ f_k)^{-1} = f_k^{-1} \circ \dots \circ f_2^{-1} \circ f_1^{-1}$$

Sea ℓ el sucesor de k; debemos hacer ver que la siguiente igualdad es válida:

$$(f_1 \circ f_2 \circ \dots \circ f_k \circ f_\ell)^{-1} = f_\ell^{-1} \circ f_k^{-1} \circ \dots \circ f_2^{-1} \circ f_1^{-1}$$

En efecto: por asociatividad:

$$(f_1 \circ f_2 \circ \dots f_k \circ f_\ell)^{-1} = [(f_1 \circ f_2 \dots f_k) \circ f_\ell]^{-1} \quad \text{-----}(1)$$

de i):

$$[(f_1 \circ f_2 \circ \dots \circ f_k) \circ f_\ell]^{-1} = f_\ell^{-1} \circ (f_1 \circ f_2 \circ \dots \circ f_k)^{-1} \quad \text{-----}(2)$$

de ii):

$$f_\ell^{-1} \circ (f_1 \circ f_2 \circ \dots \circ f_k)^{-1} = f_\ell^{-1} \circ f_k^{-1} \circ \dots \circ f_2^{-1} \circ f_1^{-1} \quad \text{-----}(3)$$

de (1), (2) y (3):

$$(f_1 \circ f_2 \circ \dots \circ f_k \circ f_\ell)^{-1} = f_\ell^{-1} \circ f_k^{-1} \circ \dots \circ f_2^{-1} \circ f_1^{-1}$$

$\Longrightarrow$ Por el principio de inducción matemática:

$$(f_1 \circ f_2 \circ \dots \circ f_n)^{-1} = f_n^{-1} \circ \dots \circ f_2^{-1} \circ f_1^{-1}$$

Teorema: Sea $f: X \to Y$; si $A \subseteq X$, $B \subseteq X$, se verifican las siguientes expresiones:

1) $f(A \cup B) = f(A) \cup f(B)$

2) $f(A \cap B) \subset f(A) \cap f(B)$

3) $f(A - B) \supset f(A) - f(B)$

4) $A \subset B \Longrightarrow f(A) \subset f(B)$

Si $A \subseteq Y$, $B \subseteq Y$:

5) $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$

6) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$

7) $f^{-1}(A - B) = f^{-1}(A) - f^{-1}(B)$

8) $A \subset B \Longrightarrow f^{-1}(A) \subset f^{-1}(B)$

9) $f^{-1}(A') = (f^{-1}(A))'$, si $A \subset Y$

Si $A \subset X$, $B \subset Y$:

10) $A \subset (f^{-1} \circ f)(A)$

11) $B = (f \circ f^{-1})(B)$

Demostraciones:

1) Haremos ver que $f(A \cup B) \subset f(A) \cup f(B)$ y que $f(A) \cup f(B) \subset f(A \cup B)$

Solución:

a) Sea $y \in f(A \cup B) \Longrightarrow$ existe $x \in A \cup B$ tal que $f(x) = y \Longrightarrow$

$\Longrightarrow x \in A$ o $x \in B$, pero:

$x \in A \Longrightarrow f(x) = y \in f(A)$

o

$x \in B \Longrightarrow f(x) = y \in f(B)$

En cualquiera de los dos casos:

$y \in f(A) \cup f(B)$

b) Sea $y \in f(A) \cup f(B) \Longrightarrow y \in f(A)$ ó $y \in f(B)$ pero:

$y \in f(A) \Longrightarrow$ existe $x \in A$ tal que $f(x) = y$

$y \in f(B) \Longrightarrow$ existe $x \in B$ tal que $f(x) = y$

En cualquiera de los dos casos:

$y = f(x)$ con $x \in A\cup B \Longrightarrow y \in f(A\cup B)$

$\Longrightarrow f(A\cup B) = f(A)\cup f(B)$

2) Sea $y \in f(A\cap B) \Longrightarrow$ para alguna $x \in A\cap B$, $(x,y) \in f$; pero

$x \in A \cap B \Rightarrow x \in A$ y $x \in B$

$$\begin{cases} x \in A \text{ y } (x,y) \in f \Longrightarrow y \in f(A) \\ x \in B \text{ y } (x,y) \in f \Longrightarrow y \in f(B) \end{cases}$$

$\Longrightarrow y \in f(A)\cap f(B)$

$\Longrightarrow f(A\cap B) \subset f(A)\cap f(B)$

Para hacer ver que, en general

$f(A\cap B) \neq f(A)\cap f(B)$ se dará un ejemplo concreto:

Sea $X = \{a,b,c\}$; $Y = \{d,e\}$

Hagamos $A = \{a,b\}$; $B = \{b,c\}$

Definamos $f: X \to Y$ del siguiente modo:

$f(a) = d$; $f(b) = e$; $f(c) = d$

$\Longrightarrow f(A\cap B) = \{e\}$; en tanto que:

$f(A)\cap f(B) = \{d,e\}$

3) Sea $y \in f(A) - f(B) \Longrightarrow$ existe $x \in A$ tal que $f(x) = y$; pero

$y \notin \{f(x) \mid x \in B\}$

$\Longrightarrow x \notin B$ ó $x \in B - A \Rightarrow y \in f(A - B)$

$\Longrightarrow f(A - B) \supseteq f(A) - f(B)$

4) Sea $A \subset B \Rightarrow A\cup B = B \Longrightarrow$

$f(A\cup B) = f(B)$; pero de 1)

$f(A)\cup f(B) = f(B)$

$\Longrightarrow f(A) \subset f(B)$

Obsérvese que si $f(A\cup B) = f(B) \not\Rightarrow A\cup B = B$

5) a) Sea $x \in f^{-1}(A\cup B) \Rightarrow f(x) \in A\cup B$

$\Longrightarrow f(x) \in A$ o $f(x) \in B$

$\Longrightarrow x \in f^{-1}(A)$ o $x \in f^{-1}(B)$

$\Longrightarrow x \in f^{-1}(A)\cup f^{-1}(B)$

$\Longrightarrow f^{-1}(A\cup B) \subset f^{-1}(A)\cup f^{-1}(B)$

b) Sea $x \in f^{-1}(A)\cup f^{-1}(B)$

$\Longrightarrow x \in f^{-1}(A)$ o $x \in f^{-1}(B)$

$\Longrightarrow f(x) \in A$ o $f(x) \in B$

$\Longrightarrow f(x) \in A\cup B$

$\Longrightarrow x \in f^{-1}(A\cup B)$

$\Longrightarrow f^{-1}(A)\cup f^{-1}(B) \subset f^{-1}(A\cup B)$

de a) y b)

$f^{-1}(A\cup B) = f^{-1}(A)\cup f^{-1}(B)$

6) a) Sea $x \in f^{-1}(A\cap B) \Rightarrow f(x) \in A\cap B \Longrightarrow$

$f(x) \in A$ y $f(x) \in B \Rightarrow x \in f^{-1}(A)$ y $x \in f^{-1}(B)$

$\Longrightarrow x \in f^{-1}(A)\cap f^{-1}(B)$

b) Sea $x \in f^{-1}(A)\cap f^{-1}(B) \Rightarrow f(x) \in A$ y $f(x) \in B$

$\Longrightarrow f(x) \in A\cap B \Rightarrow x \in f^{-1}(A\cap B)$

$\Longrightarrow f^{-1}(A\cap B) = f^{-1}(A)\cap f^{-1}(B)$

7) a) Sea $x \in f^{-1}(A - B) \Rightarrow f(x) \in A - B \Rightarrow$

$f(x) \in A$ y $f(x) \notin B \Rightarrow x \in f^{-1}(A)$, pero

$x \notin f^{-1}(B) \Rightarrow x \in f^{-1}(A) - f^{-1}(B)$

b) Sea $x \in f^{-1}(A) - f^{-1}(B) \Rightarrow f(x) \in A$ pero

$f(x) \notin B \Rightarrow f(x) \in A - B \Rightarrow x \in f^{-1}(A - B)$

$\Rightarrow f^{-1}(A - B) = f^{-1}(A) - f^{-1}(B)$

8) Sea $A \subset B \Rightarrow A \cup B = B \Rightarrow$

$f(A \cup B) = f(B)$

Por ser f biyectiva existe $f^{-1} \Rightarrow$

$f^{-1}(A \cup B) = f^{-1}(B)$

de 5) $f^{-1}(A) \cup f^{-1}(B) = f^{-1}(B)$

$\Rightarrow f^{-1}(A) \subset f^{-1}(B)$

9) a) Sea $x \in f^{-1}(A') \Rightarrow f(x) \in A' \Rightarrow$

$f(x) \notin A \Rightarrow x \notin f^{-1}(A) \Rightarrow x \in [f^{-1}(A)]'$

$\Rightarrow f^{-1}(A') \subset (f^{-1}(A))'$

b) Sea $x \in (f^{-1}(A))' \Rightarrow x \notin f^{-1}(A) \Rightarrow f(x) \notin A$

$\Rightarrow f(x) \in A' \Rightarrow x \in f^{-1}(A')$

$\Rightarrow (f^{-1}(A))' \subset f^{-1}(A')$

$\Rightarrow f^{-1}(A') = (f^{-1}(A))'$

10) Sea $x \in A$; para alguna $y \in Y$ se tiene $(x,y) \in f \Rightarrow$

Si $x \in A$ y $(x,y) \in f \Rightarrow y \in f(A)$; pero,

$y \in f(A)$ y $(x,y) \in f \Rightarrow x \in f^{-1}(f(A))$

$\Rightarrow A \subset (f^{-1} \circ f)(A)$

Para hacer ver que, en general,

$A \neq f^{-1}(f(A))$ nos referiremos al siguiente caso concreto:

Sea $X = \{x_1, x_2, x_3, x_4\}$; $Y = \{y_1, y_2, y_3\}$

definamos $f: X \to Y$ como sigue:

$f(x_1) = y_1$; $f(x_2) = y_1$; $f(x_3) = y_1$; $f(x_4) = y_2$

Sea $A = \{x_1, x_2\} \Rightarrow f(A) = \{y_1\}$; en tanto

que $f^{-1}(f(A)) = f^{-1}(\{y_1\}) = \{x_1, x_2, x_3\} \neq A$

11) a) Sea $y \in B$; para alguna $x \in X$ se tiene $(x,y) \in f \Rightarrow$

si $(x,y) \in f$ y $y \in B \Rightarrow x \in f^{-1}(B)$;

pero $f(x) \in f(f^{-1}(B)) \Rightarrow y \in f(f^{-1}(B))$

$\Rightarrow B \subset f(f^{-1}(B))$

b) Si $f(x) = y \in f(f^{-1}(B)) \Rightarrow x \in f^{-1}(B) \Rightarrow f(x) \in B$

$\Rightarrow f(f^{-1}(B)) \subset B$

$\Rightarrow B = (f \circ f^{-1})(B)$

Algunos libros de referencia:

1.- "Fundamental Concepts of Analysis"
A. H. Smith, W. A. Albrecht
(Edit. Prentice Hall) 1966

2.- "Introduction to Analysis" Vol. I
N. B. Haaser, J. P. LaSalle, J. A. Sullivan
(Edit. Blaisdell) 1959

3.- "General Topology"
S. Lipschutz
(Edit. Schaum) 1965

4.- "Advanced Calculus"
R. C. Buck
(Edit. McGraw Hill) 1965

5.- "Set theory and related topics"
S. Lipschutz
(Edit. Schaum) 1964

6.- "Notas sobre conjuntos y números"
J. Salazar R.
(Edit. C. P. E.) 1967

TRANSFORMACIONES LINEALES

VI. 1. TRANSFORMACIONES.

Como hemos visto, una función f de A en B (donde A y B son dos conjuntos no vacíos cualesquiera) es una regla o criterio que asocia a cada elemento de A, uno y sólo un elemento de B, lo cual denotamos mediante

$$f: A \rightarrow B$$

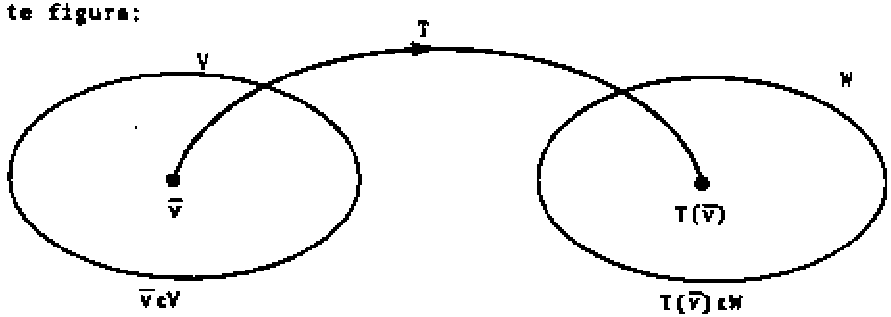
Existen también funciones entre espacios vectoriales que, en forma similar, denotaremos por

$$T: V \rightarrow W$$

donde V y W son espacios vectoriales definidos sobre el mismo campo K, y T es la regla de correspondencia que asigna a cada vector $\bar{v}$ de V uno y sólo un vector de W, al que llamaremos IMAGEN de $\bar{v}$ y representaremos con $T(\bar{v})$.

A este tipo de funciones les daremos el nombre de TRANSFORMACIONES. (Algunos autores emplean los términos operador, aplicación o mapeo en lugar del de transformación).

Podemos ilustrar las ideas anteriores mediante la siguiente figura:



donde V y W representan dos espacios vectoriales.

Ejemplo VI. 1.

a) Consideremos una transformación T que consiste en multiplicar por 3 un vector cualquiera (x, y) de R^2.

Examinemos algunos casos:

La imagen del vector (1, 4) es 3(1, 4)=(3, 12), lo que describimos como: $T(1, 4) = (3, 12)$

En forma análoga:

$$T(6,-2) = (18,-6)$$
$$T(0, 0) = (0, 0) \quad \text{etc.}$$

Es evidente que se trata de una transformación de R^2 en R^2, es decir: $T: R^2 \rightarrow R^2$

y la imagen de un vector cualquiera (x, y) está dada por la regla

$$T(x, y) = (3x, 3y) .$$

b) Sea ahora la transformación $S: R^2 \rightarrow R^2$

definida por la regla $S(x, y) = (x^2-1, 2xy)$.

Obtengamos, por ejemplo $S(1, 2) = (0, 4)$

que en palabras dice: La imagen del vector (1, 2) según la transformación S es el vector (0, 4); o bien: S aplica al vector (1, 2) en el vector (0, 4).

De igual forma se obtendrían:

$$S(-9, 3) = (80, -54)$$

$$S(0, 0) = (-1, 0) \quad \text{etc.}$$

- -

Definición.

Sean V y W dos espacios vectoriales definidos sobre un mismo campo K:

Una transformación de V en W es una función que asocia a cada vector de V uno y sólo un vector de W.

Ejemplo VI. 2.

Sea la transformación $T: R^3 \to R^2$

definida por la regla $T(x, y, z) = (x, y)$

Es claro que $T(1, 2, 3) = (1, 2)$

$T(1, 2, -14) = (1, 2)$

$T(6, 0, 1) = (6, 0)$ etc.

Una interpretación geométrica de esta transformación sería la siguiente:

Si (x, y, z) representa un segmento dirigido en el espacio de tres dimensiones, su imagen (x, y) representa la proyección de dicho segmento en el plano XY.

- -

Ejemplo VI. 3.

El Operador derivada (D).

Sea F el espacio vectorial formado por todas las funciones de R en R, derivables en un intervalo (a, b). La transformación según la cual la imagen de una función $f \in F$ es su derivada f' se llama operador derivada y se denota con D.

Tenemos por tanto: $D: F \to G$

donde G es el espacio que contiene a todas las derivadas.

Por ejemplo, $D(3x^2+1) = 6x$

(La imagen de $3x^2+1$ según D es $6x$).

$D(\text{sen}x) = \cos x$

$D(12) = 0$ etc.

En general: $D(f(x)) = f'(x)$

- -

Transformaciones lineales.

Regresemos a la transformación $T: R^2 \to R^2$ del ejemplo VI.1

dada por: $T(x, y) = (3x, 3y)$

Sabemos que $T(1, 4) = (3, 12)$

y $T(6, -2) = (18, -6)$

Podemos preguntarnos: ¿Cuál será la imagen del vector que se obtiene sumando los vectores $(1, 4)$ y $(6, -2)$?

Calculémosla: $(1, 4)+(6, -2) = (7, 2)$

Aplicando la regla de transformación al vector obtenido:

$T(7, 2) = (21, 6)$

Observemos que el mismo resultado se habría obtenido sumando simplemente las imágenes de los vectores involucrados, esto es:

$$T((1, 4)+(6,-2)) = T(1, 4) + T(6, -2)$$
$$= (3,12) + (18, -6)$$
$$= (21, 6) .$$

Cabe ahora preguntarse: ¿Se cumple lo anterior para cualquier par de vectores (x_1, y_1) y (x_2, y_2) de R^2? Es decir, ¿es siempre cierto que

$$T((x_1, y_1)+(x_2, y_2)) = T(x_1, y_1) + T(x_2, y_2) \ ?$$

Investiguémoslo:

$$T((x_1, y_1)+(x_2, y_2)) = T(x_1+x_2, y_1+y_2)$$

aplicando la regla:

$$T((x_1, y_1)+(x_2, y_2)) = (3(x_1+x_2), 3(y_1+y_2))$$
$$T((x_1, y_1)+(x_2, y_2)) = (3x_1+3x_2, 3y_1+3y_2)$$

Por otra parte, aplicando la regla tenemos:

$$T(x_1, y_1)+T(x_2, y_2) = (3x_1, 3y_1)+(3x_2, 3y_2)$$

o sea que:

$T(x_1, y_1)+T(x_2, y_2)= (3x_1+3x_2, 3y_1+3y_2)$

Por tanto: $T((x_1, y_1)+(x_2, y_2))= T(x_1, y_1)+T(x_2, y_2)$

Vemos que para esta transformación:

"La imagen de la suma de dos vectores es igual a la suma de las imágenes de cada uno de ellos".

En forma similar, podemos ver que ocurre cuando multiplicamos un vector por un escalar y obtenemos la imagen del vector resultante.

Tomemos por ejemplo el vector $5(6, -2) = (30, -10)$

Aplicando la regla de transformación al vector obtenido:

$$T(30, -10) = (90, -30)$$

Nuevamente, habría bastado con multiplicar por 5 la imagen del vector $(6, -2)$:

$$\begin{aligned} T(5(6, -2)) &= 5\,T(6, -2) \\ &= 5\,(18, -6) \\ &= (90, -30). \end{aligned}$$

Es fácil comprobar que esto se cumple para cualquier vector (x, y) y cualquier escalar $\alpha \in R$, o sea:

$$T(\alpha(x, y)) = \alpha\,T(x, y)$$

En efecto:

$$T(\alpha(x, y)) = T(\alpha x, \alpha y) = (3\alpha x, 3\alpha y)$$

$$\alpha\,T(x, y) = \alpha(3x, 3y) = (3\alpha x, 3\alpha y)$$

Concluímos que para esta transformación:

"La imagen del producto de un escalar por un vector es igual al escalar multiplicado por la imagen del vector".

A las transformaciones que tienen estas dos propiedades, se les llama TRANSFORMACIONES LINEALES.

Hagamos un estudio similar con la transformación $S: R^2 \to R^2$ del ejemplo VI. 1, definida por:

$$S(x, y) = (x^2-1, 2xy)$$

Sabemos que: $S(-9, 3) = (80, -54)$

$S(1, 2) = (0, 4)$

Obtengamos

$$S((-9, 3)+(1, 2)) = S(-8, 5) = (63, -80)$$

En este caso, sumar las imágenes de cada uno de los vectores hubiera conducido a un resultado distinto, pues:

$$S(-9, 3)+ S(1, 2)= (80, -54) + (0, 4)= (80, -50)$$

Dicho de otra forma:

$$S((-9, 3)+(1, 2)) \neq S(-9, 3) + S(1, 2).$$

En lo que concierne al producto por un escalar, obtengamos por ejemplo: $S(2(1, 2)) = S(2, 4) = (3, 16)$.

Este resultado difiere del que se obtiene multiplicando por 2 la imagen de $(1, 2)$, ya que:

$$2S(1, 2) = 2(0, 4) = (0, 8)$$

es decir: $S(2(1, 2)) \neq 2\,S(1, 2)$

Observamos que la transformación S no posee las propiedades que tiene T, por lo cual se dice que S NO es lineal.

<u>Definición.</u>

La transformación

$$T: V \to W$$

es una transformación lineal si:

1.- $T(\bar{v}_1+\bar{v}_2) = T(\bar{v}_1) + T(\bar{v}_2)$, $\forall\ \bar{v}_1, \bar{v}_2 \in V$

2.- $T(\alpha\bar{v}) = \alpha\,T(\bar{v})$, $\forall\ \bar{v} \in V$ y $\alpha \in K$

Se dice entonces que T preserva las operaciones de suma y producto por un escalar.

(En el estudio de sistemas físicos es común llamar a 1 y 2 propiedades de "superposición" y "homogeneidad" respectivamente). Estas dos propiedades pueden combinarse en una sola que establece que:

> T es lineal si:
>
> $$T(\alpha\bar{v}_1+\bar{v}_2)=\alpha T(\bar{v}_1)+T(\bar{v}_2)$$
>
> para todo par de vectores $\bar{v}_1$ y $\bar{v}_2$ de V y todo escalar α.

De entre todas las transformaciones de un espacio vectorial en otro, son especialmente útiles aquellas que son lineales pues se presentan en diversas ramas de la matemática y, frecuentemente, las propiedades de transformaciones más generales se obtienen aproximándolas mediante transformaciones lineales. Al estudio de estas últimas se dedicará el resto del capítulo.

Ya hemos demostrado que la transformación T del ejemplo VI. 1 es lineal, mientras que S no lo es. Veamos otros casos:

Ejemplo VI. 4.

Decir si las siguientes transformaciones son lineales:

a) $T: R^3 \to R^2$, $T(x, y, z) = (x+y, x+z)$

b) $T: R^2 \to R^2$, $T(x, y) = (x+1, y)$

Solución:

a) $T: R^3 \to R^2$, $T(x, y, z) = (x+y, x+z)$

Veamos si cumple con las propiedades de linealidad:

1) Sean (x_1, y_1, z_1) y (x_2, y_2, z_2) dos vectores arbitrarios de R^3:

$T((x_1, y_1, z_1)+(x_2, y_2, z_2)) = T(x_1+x_2, y_1+y_2, z_1+z_2)$

aplicando la regla

$T((x_1, y_1, z_1)+(x_2, y_2, z_2)) = (x_1+x_2+y_1+y_2, x_1+x_2+z_1+z_2)$

Además:

aplicando la regla

$T(x_1, y_1, z_1)+T(x_2, y_2, z_2)=(x_1+y_1, x_1+z_1)+(x_2+y_2, x_2+z_2)$

o sea que

$T(x_1, y_1, z_1)+T(x_2, y_2, z_2)=(x_1+y_1+x_2+y_2, x_1+z_1+x_2+z_2)$.

$\therefore$ Se cumple que $T((x_1, y_1, z_1)+(x_2, y_2, z_2))=T(x_1,y_1,z_1)+$
$+T(x_2,y_2, z_2)$.

2) Sea (x, y, z) un vector R^3 y α un número real:

$T(\alpha(x, y, z)) = T(\alpha x, \alpha y, \alpha z)$

$= (\alpha x+\alpha y, \alpha x+\alpha z)$

Además:

$\alpha T(x, y, z) = \alpha(x+y, x+z)$

$= (\alpha x+\alpha y, \alpha x+\alpha z)$

$\therefore$ Se cumple que $T(\alpha(x, y, z)) = \alpha T(x, y, z)$.

En conclusión, T es una transformación lineal.

b) $T: R^2 \to R^2$, $T(x, y) = (x+1, y)$

1) Sean (x_1, y_1) y (x_2, y_2) dos vectores cualesquiera de R^2:

$T((x_1, y_1)+(x_2, y_2)) = T(x_1+x_2, y_1+y_2)$

$= (x_1+x_2+1, y_1+y_2)$

Por otra parte:

$T(x_1, y_1)+T(x_2, y_2)= (x_1+1, y_1)+(x_2+1, y_2)$

$= (x_1+x_2+2, y_1+y_2)$

$\therefore$ $T((x_1, y_1)+(x_2, y_2)) \neq T(x_1, y_1)+T(x_2, y_2)$

Puesto que no se cumple la primera propiedad, la transformación que nos ocupa NO es lineal.

- -

Es evidente que, según la definición, basta con que no se cumpla alguna de las dos propiedades para que la transformación no sea lineal.

Se deja al estudiante demostrar que la transformación del ejemplo VI. 2 es lineal.

Ejemplo VI. 5.

El operador D es lineal.

En efecto, de cursos de matemáticas sabemos que:

"La derivada de la suma de dos funciones es igual a la suma de las derivadas de cada una de ellas", es decir:

$D(f_1+f_2) = D(f_1)+D(f_2) \quad \forall f_1, f_2 \in F.$

También sabemos que "La derivada de una constante por una función es igual a la constante por la derivada de la función", o sea: $D(\alpha f) = \alpha D(f) \quad \forall \alpha \in R$ y $f \in F$.

Como vemos, lo que hemos enunciado son precisamente las características de linealidad del operador D.

- -

Ejemplo VI. 6.

Transformación identidad.

La transformación identidad $I_V: V \to V$ definida por $I_V(\bar{v}) = \bar{v}$ para todo $\bar{v}$ de V, es lineal.

- -

Ejemplo VI. 7.

Transformación cero.

La transformación cero $0: V \to W$ definida por $0(\bar{v}) = \bar{0}_W$ para todo $\bar{v}$ de V, es lineal.

- -

VI. 2 DOMINIO, RECORRIDO Y NUCLEO.

Consideremos la transformación lineal $T: R^3 \to R^2$, dada por la regla $T(x, y, z) = (x, 2x)$.

Es claro que ésta es una regla que se puede aplicar a cualquier vector de R^3, por lo que llamaremos a R^3 DOMINIO de la transformación.

Veamos ahora que ocurre con los vectores imagen por medio de algunos ejemplos:

$T(2, 6, -8) = (2, 4)$

$T(1, 3, 0) = (1, 2)$

$T(-5, 6, -9) = (-5, -10)$

Es importante notar que, aunque las imágenes están en R^2, no todos los vectores de R^2 son imagen de algún vector del dominio. Esto se debe a que, como se ve de la regla de transformación, todos los vectores imagen tienen la forma $(x, 2x)$, es decir, su segunda componente es igual a dos veces la primera.

Surge por tanto la necesidad de dar un nombre al conjunto de vectores imagen. Nosotros lo llamaremos RECORRIDO de la transformación y lo representaremos por $T(R^3)$. Entonces:

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

Es claro que los vectores $(1, 3)$, $(2, -8)$ y $(6, 0)$, entre otros, no son imagen de ningún vector del dominio, lo que es equivalente a decir que no están en el recorrido de la transformación.

Analicemos algunos casos adicionales:

$T(0, 3, 4) = (0, 0)$

$T(0, 0, 0) = (0, 0)$

$T(0, -9, 2) = (0, 0)$

Fácilmente se deduce que todos aquellos vectores del dominio cuya primera componente es nula tienen como imagen el vector cero de R^2. A este conjunto de vectores se le da también un nombre: NUCLEO de la transformación y se le representa por $N(T)$.

En este caso: $N(T) = \{(0, y, z) \mid y, z \in R\}$.

Por lo anterior: $(0, 3, 4) \in N(T)$

$(0, 0, 0) \in N(T)$

$(1, 0, 0) \notin N(T)$ etc.

Pasaremos a formalizar las ideas anteriores con la siguiente:

Definición:

Sea la transformación lineal

$T: V \to W$

DOMINIO: Es el conjunto V de vectores sobre los cuales actúa la transformación.

RECORRIDO: Es el conjunto formado por las imágenes de los vectores del dominio. Lo representaremos con $T(V)$, es decir:

$T(V) = \{\bar{w} \mid \bar{w} = T(\bar{v}),\ \bar{v} \in V\}$

NUCLEO: Es el conjunto de vectores del dominio cuya imagen es el vector cero de W. Lo representaremos con $N(T)$, es decir:

$N(T) = \{\bar{v} \mid \bar{v} \in V,\ T(\bar{v}) = \bar{0}_W\}$

Ejemplo VI. 8.

Para el operador derivada D ya discutido anteriormente (ver ejemplo VI. 3) tenemos:

Dominio - Es el conjunto de todas las funciones derivables en el intervalo.

Recorrido - Es el conjunto de todas las funciones derivadas.

Núcleo - Es el conjunto de todas las funciones que son constantes en el intervalo (ya que al derivar una constante se obtiene la función cero).

- -

En general, para encontrar el núcleo de una transformación lineal, igualamos la imagen al vector cero y vemos que condiciones resultan de ello para los vectores del dominio.

Por ejemplo, para el caso discutido al principio de esta sección, donde:

$$T(x, y, z) = (x, 2x)$$

para obtener el núcleo de T haríamos

$$(x, 2x) = (0, 0)$$

lo que implica $x = 0$, por lo cual

$$N(T) = \{(0, y, z) \mid y, z \in R\}$$

como se tenía.

Hemos visto que, en una transformación lineal $T: V \to W$, el recorrido T(V) y el núcleo N(T) son subconjuntos de W y V, respectivamente. Demostraremos a continuación que no sólo son subconjuntos, sino también subespacios.

Teorema VI. 1.

Sea la transformación lineal $T: V \to W$.

El recorrido de la transformación es un subespacio de W.

Demostración.

De acuerdo con el Teorema V.4, basta con demostrar que el recorrido, T(V), es cerrado para la suma y el producto por un escalar.

1) Sean $T(\bar{v}_1)$ y $T(\bar{v}_2)$ dos vectores cualesquiera de T (V).

Por linealidad: $T(\bar{v}_1) + T(\bar{v}_2) = T(\bar{v}_1 + \bar{v}_2) \in T(V)$

ya que, como $\bar{v}_1, \bar{v}_2 \in V$ entonces $\bar{v}_1 + \bar{v}_2 \in V$.

Es decir, si $T(\bar{v}_1)$ y $T(\bar{v}_2)$ son elementos del recorrido, $T(\bar{v}_1) + T(\bar{v}_2)$ también lo es, cumpliéndose la cerradura para la suma.

2) Sean $T(\bar{v})$ un elemento cualquiera de T (V) y α un escalar del campo.

Por linealidad: $\alpha T(\bar{v}) = T(\alpha\bar{v}) \in T(V)$

Ya que, como $\bar{v} \in V$, entonces $\alpha\bar{v} \in V$.

Es decir, $\alpha T(\bar{v})$ es también un elemento del recorrido y hemos demostrado la cerradura para el producto por un escalar.

De 1) y 2) queda demostrado el teorema.

<u>Imagen del vector cero.</u>

Un resultado importante que se desprende de las propiedades de linealidad es el siguiente:

En una transformación lineal $T: V \to W$, la imagen del vector cero de V es el vector cero de W, o sea:

$$T(\bar{0}_V) = \bar{0}_W$$

Una demostración inmediata es la siguiente:

Haciendo $\alpha = 0$ en

$$T(\alpha\bar{v}) = \alpha T(\bar{v})$$

se obtiene

$$T(\bar{0}_V) = \bar{0}_W, \text{ como se quería.}$$

Por lo que, el núcleo siempre contiene al vector cero si T es lineal.

Teorema VI. 2.

Sea la transformación lineal $T: V \to W$

El núcleo de la transformación es un subespacio vectorial de V.

Demostración.

1) Sean $\bar{v}_1$ y $\bar{v}_2$ dos vectores de N (T),

entonces: $T(\bar{v}_1) = \bar{0}_W$ y $T(\bar{v}_2) = \bar{0}_W$

Por linealidad: $T(\bar{v}_1+\bar{v}_2) = T(\bar{v}_1) + T(\bar{v}_2)$

Sustituyendo: $T(\bar{v}_1+\bar{v}_2) = \bar{0}_W + \bar{0}_W$

$T(\bar{v}_1+\bar{v}_2) = \bar{0}_W$

por lo que, $\bar{v}_1+\bar{v}_2$ es también un vector de N (T).

2) Sea $\bar{v}$ un vector de N (T), entonces $T(\bar{v}) = \bar{0}_W$, y sea α un escalar del campo.

Por linealidad: $T(\alpha\bar{v}) = \alpha\, T(\bar{v})$

Sustituyendo: $T(\alpha\bar{v}) = \alpha\, \bar{0}_W$

$T(\alpha\bar{v}) = \bar{0}_W$

$\therefore$ $\alpha\bar{v}$ también está en N (T).

Esto completa la demostración.

Consideremos nuevamente la transformación $T: R^3 \to R^2$ dada por

$$T(x, y, z) = (x, 2x)$$

Puesto que el núcleo y el recorrido son a su vez espacios vectoriales, tienen base y dimensión.

Como se vió anteriormente, $T(R^3) = \{(x, 2x) \mid x \in R\}$ es el recorrido de la transformación, y el estudiante puede verificar fácilmente que el conjunto $B_R = \{(1, 2)\}$ es una base de dicho espacio, por lo que: $\dim T(R^3) = 1$.

En forma similar, para el núcleo $N(T) = \{(0,y,z) \mid y,z \in R\}$, una base es $B_N = \{(0,1,0), (0,0,1)\}$ y por ello: $\dim N(T) = 2$.

En cuanto al dominio R^3, sabemos que es de dimensión tres: $\dim R^3 = 3$.

Observamos que la dimensión del dominio es la suma de las dimensiones del recorrido y del núcleo. Este es un resultado muy importante que, como veremos a continuación, se tiene siempre que se trabaja con espacios de dimensión finita.

Teorema VI. 3.

Sea V un espacio vectorial de dimensión finita y sea

$$T: V \to W$$

una transformación lineal.

Entonces

$$\dim V = \dim T(V) + \dim N(T)$$

Demostración.

Sea $\{\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_p\}$ una base del recorrido T (V), es decir,

$\dim T(V) = p$

Sea $\{\bar{n}_1, \bar{n}_2, \ldots, \bar{n}_q\}$ una base del núcleo N (T), es decir,

$\dim N(T) = q$

Puesto que los vectores $\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_p$ pertenecen a T (V), deben existir p vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_p \in V$ tales que:

$$\left.\begin{array}{l} T(\bar{v}_1) = \bar{r}_1 \\ T(\bar{v}_2) = \bar{r}_2 \\ \cdots\cdots \\ T(\bar{v}_p) = \bar{r}_p \end{array}\right\} \quad \ldots\ldots (1)$$

Si demostramos que el conjunto

$$B = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_p, \bar{n}_1, \bar{n}_2, \ldots \bar{n}_q\}$$

es una base de V, habremos demostrado el teorema, pues eso probará que dim V = p+q.

1) B genera al espacio V. :

Sea $\bar{v}$ un vector cualquiera de V.

Ya que T ($\bar{v}$) está en el recorrido, podemos expresarlo como combinación lineal de los elementos de la base ($\bar{r}_1$, $\bar{r}_2$, ..., $\bar{r}_p$):

$$T(\bar{v}) = \alpha_1\bar{r}_1+\alpha_2\bar{r}_2+ \ldots + \alpha_p\,\bar{r}_p$$

Sustituyendo (1):

$$T(\bar{v}) = \alpha_1T(\bar{v}_1)+\alpha_2T(\bar{v}_2)+ \ldots +\alpha_pT\,(\bar{v}_p)$$

Por linealidad:

$$T(\bar{v}) = T\,(\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_p\bar{v}_p)$$

de donde: $T(\bar{v}) - T.(\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_p\bar{v}_p)=\bar{0}_W$

Por linealidad:

$$T(\bar{v}-\alpha_1\bar{v}_1-\alpha_2\bar{v}_2-\ldots-\alpha_p\bar{v}_p)=\bar{0}_W$$

Por tanto, el vector ($\bar{v}-\alpha_1\bar{v}_1-\alpha_2\bar{v}_2-\ldots-\alpha_p\bar{v}_p$) está en el núcleo (ya que su imagen es el cero) y lo podemos expresar como combinación lineal de los elementos de la base ($\bar{n}_1$, $\bar{n}_2$, ..., $\bar{n}_q$), es decir:

$$\bar{v} - \alpha_1\bar{v} - \alpha_2\bar{v}_2- \ldots - \alpha_p\bar{v}_p= \beta_1\bar{n}_1+\beta_2\bar{n}_2+\ldots+\beta_q\bar{n}_q$$

Finalmente:

$$\bar{v} = \alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_p\bar{v}_p+\beta_1\bar{n}_1+\beta_2\bar{n}_2+\ldots+\beta_q\bar{n}_q$$

Hemos demostrado que cualquier vector de V se puede expresar en términos de los vectores de B, por lo que éste es un conjunto generador.

2) B es linealmente independiente.

Formemos la combinación lineal

$$\lambda_1\bar{v}_1+\lambda_2\bar{v}_2+\ldots+\lambda_p\bar{v}_p+\gamma_1\bar{n}_1+\gamma_2\bar{n}_2+\ldots+\gamma_q\bar{n}_q=\bar{0}_V \quad \ldots . (2)$$

Transformando ambos miembros de (2):

$$T(\lambda_1\bar{v}_1+\lambda_2\bar{v}_2+\ldots+\lambda_p\bar{v}_p+\gamma_1\bar{n}_1+\gamma_2\bar{n}_2+\ldots+\gamma_q\bar{n}_q)=T(\bar{0}_V)$$

Por linealidad:

$$\lambda_1T(\bar{v}_1)+\ldots+\lambda_pT(\bar{v}_p)+\gamma_1T(\bar{n}_1)+\ldots+\gamma_qT(\bar{n}_q)=\bar{0}_W \quad \ldots . (3)$$

Puesto que los vectores $\bar{n}_1$, $\bar{n}_2$, ..., $\bar{n}_q$ están en el núcleo:

$$T(\bar{n}_1) = T(\bar{n}_2) = \ldots = T(\bar{n}_q) = \bar{0}_W \quad \ldots . (4)$$

Sustituyendo (1) y (4) en (3):

$$\lambda_1\bar{r}_1+\lambda_2\bar{r}_2+\ldots+\lambda_p\bar{r}_p = \bar{0}_W \quad \ldots . (5)$$

Por hipótesis, (r_1, r_2, ..., r_p) es una base, por lo que (5) implica que:

$$\lambda_1 = \lambda_2 = \ldots = \lambda_p = 0 \quad \ldots . (6)$$

Sustituyendo (6) en (2):

$$\gamma_1\bar{n}_1+\gamma_2\bar{n}_2+ \ldots +\gamma_q\,\bar{n}_q = \bar{0}_V \quad \ldots . (7)$$

Análogamente, por ser {$\bar{n}_1$, $\bar{n}_2$, ..., $\bar{n}_q$} una base, (7) implica que:

$$\gamma_1= \gamma_2 = \ldots = 0 \quad \ldots . (8)$$

Hemos demostrado que (2) implica que

$$\lambda_1= \lambda_2 =\ldots= \lambda_p = \gamma_1 = \gamma_2 = \ldots = \gamma_q = 0$$

por lo que B es linealmente independiente.

De 1) y 2), B es una base de V y por tanto:

dim V = p + q

quedando demostrado el teorema.

Ejemplo VI. 9.

Consideremos la transformación lineal $T: R^3 \rightarrow R^3$ dada por la regla $T\,(x, y, z) = (3x + y,\ 6x - z,\ 2y + z)$.

Determinemos el núcleo haciendo

$$(3x + y,\ 6x - z,\ 2y + z) = (0, 0, 0)$$

de donde se obtiene el sistema

$$3x + y = 0$$
$$6x - z = 0$$
$$2y + z = 0$$

El estudiante puede verificar que la solución general de este sistema homogéneo es:

$$y = -3x$$
$$z = 6x$$

por lo que:

$$N(T) = \{(x, -3x, 6x) \mid x \in R\}$$

Una base del núcleo es el conjunto $B_N = \{(1, -3, 6)\}$ y en consecuencia $\dim N(T) = 1$.

Es claro que la dimensión del dominio es $\dim R^3 = 3$.

Aplicando el teorema VI.3, concluimos que la dimensión del recorrido es $\dim T(R^3) = 2$.

Sabemos que los vectores del recorrido tienen tres componentes, pero el hecho de que la dimensión de éste sea dos nos obliga a pensar que únicamente dos de esas componentes son independientes. Hallemos la relación de dependencia entre las componentes de un vector

$$(a, b, c) \in T(R^3) \quad \ldots\ldots (1)$$

Para ello, podemos plantear la ecuación

$$\alpha_1 a + \alpha_2 b + \alpha_3 c = 0 \quad \ldots\ldots (2)$$

Según la regla de transformación:

$$(a, b, c) = (3x + y,\ 6x - z,\ 2y + z)$$

y sustituyendo en (2)

$$\alpha_1(3x+y)+\alpha_2(6x-z)+\alpha_3(2y+z)=0$$

Reacomodando:

$$(3\alpha_1+6\alpha_2)x+(\alpha_1+2\alpha_3)y+(-\alpha_2+\alpha_3)z=0$$

esta ecuación se verifica para cualquier valor de x, y, z si:

$$3\alpha_1+6\alpha_2 = 0$$
$$\alpha_1+2\alpha_3 = 0$$
$$-\alpha_2+\alpha_3 = 0$$

La solución general de este sistema es:

$$\alpha_1=-2\alpha_3$$
$$\alpha_2=\alpha_3$$

que sustituyéndola en (2) nos dice:

$$-2\alpha_3 a+\alpha_3 b+\alpha_3 c = 0$$

o bien $\quad -2a + b + c = 0$

$$\therefore \quad c = 2a-b$$

y el recorrido se puede describir como

$$T(R^3) = \{(a, b, 2a-b) \mid a, b \in R\}$$

Finalmente, una base del recorrido es el conjunto

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

donde se ve claramente que $\dim T(R^3) = 2$.

- -

Ejemplo VI. 10.

Para la transformación $T: R^2 \rightarrow R^2$ del ejemplo VI. 1 dada por $T(x, y) = (3x, 3y)$, obtengamos el núcleo:

$$(3x, 3y)=(0, 0)$$

$$\therefore \quad 3x=0$$
$$3y=0$$

La única solución de este sistema es

$x = 0$

$y = 0$

de donde:

$N(T) = \{(0, 0)\}$

y, por definición, $\dim N(T) = 0$

Además, la dimensión del dominio es $\dim R^2 = 2$.

Por tanto, del teorema VI. 3, $\dim T(R^2) = 2$, lo que implica que el recorrido es todo R^2 para este caso. Es decir: $T(R^2) = R^2$.

- -

Obsérvese que, como consecuencia del teorema VI. 3, la dimensión del recorrido es cuando más igual a la del dominio, pero no puede ser mayor.

Ejemplo VI. 11.

Sea V el espacio vectorial de matrices de orden 2x3, y sea W el espacio vectorial de matrices de orden 2x2, ambas sobre R.

Consideremos la transformación $T: V \to W$

definida por $T(X) = XA$, para todo X de V donde:

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

Es claro que T es lineal pues, para todo $X, Y \in V$:

$$T(X + Y) = (X + Y)A$$
$$= XA + YA$$
$$= T(X) + T(Y)$$

en forma análoga

$$T(\alpha X) = (\alpha X) A$$
$$= \alpha (XA)$$
$$= \alpha T(X), \quad \forall \alpha \in R.$$

El dominio de T es $V = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \middle| a, b, c, d, e, f \in R \right\}$

Una base del dominio es:

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

$\therefore \quad \dim V = 6$

Para encontrar el núcleo necesitamos conocer la imagen de una matriz cualquiera $X = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$:

$$T(X) = XA$$

$$T(X) = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 0 \\ 6 & 0 \end{bmatrix} = \begin{bmatrix} a+3b+6c & 2a \\ d+3e+6f & 2d \end{bmatrix}$$

Para que X esté en el núcleo:

$$\begin{bmatrix} a + 3b + 6c & 2a \\ d + 3e + 6f & 2d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

$\therefore \quad a+3b+6c = 0$

$2a = 0$

$d+3e+6f = 0$

$2d = 0$

La solución general de este sistema de ecuaciones es

$a = 0$

$b = -2c$

$d = 0$

$e = -2f$

$$\therefore \quad N(T) = \left\{ \begin{bmatrix} 0 & -2c & c \\ 0 & -2f & f \end{bmatrix} \;\middle|\; c, f \in R \right\}$$

Una base del núcleo es

$$B_N = \left\{ \begin{bmatrix} 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 1 \end{bmatrix} \right\}$$

$\therefore$ dim $N(T) = 2$.

Por último, según el teorema VI. 3, la dimensión del recorrido es dim $T(V) = 4$.

Esto implica que

$$T(V) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \;\middle|\; a, b, c, d, \in R \right\} = W$$

es decir, el recorrido de la transformación esta constituido por todas las matrices cuadradas de orden dos.

Una base de $T(V)$ es

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

- -

VI. 3. REPRESENTACION MATRICIAL DE UNA TRANSFORMACION LINEAL.

Consideremos el siguiente caso:

De una transformación lineal

$$T: R^2 \to R^2$$

se conoce únicamente que:

$$T(2, 1) = (1, 3)$$
y
$$T(7, 3) = (0, 8)$$

¿Es posible encontrar $T(1, 0)$?

Veámoslo:

El vector $(1, 0)$ se puede expresar como combinación lineal de $(2, 1)$ y $(7, 3)$. En efecto:

$$(1,0) = (7, 3) - 3(2,1)$$

Por ello $T(1,0) = T((7,3)-3(2,1))$

y, puesto que T es lineal:

$$T(1,0) = T(7,3) - 3T(2,1)$$
$$T(1,0) = (0,8) - 3(1,3)$$
$$T(1,0) = (-3, -1)$$

Es fácil comprobar que el conjunto $\{(2,1), (7,3)\}$ es una base de R^2, por lo que cualquier vector (x,y) de R^2 puede obtenerse a partir de dicho conjunto. Para este caso:

$$(x,y) = (-3x+7y)(2,1)+(x-2y)(7,3) \qquad , \forall\ (x,y) \in R^2.$$

Esto nos permite obtener la regla de transformación, ya que:

$$T(x,y) = T((-3x+7y)(2,1)+(x-2y)(7,3))$$

y como T es lineal:

$$T(x,y) = (-3x+7y)T(2,1)+(x-2y)T(7,3)$$

de donde:

$$T(x,y)= (-3x+7y)(1,3)+(x-2y)(0,8)$$

Finalmente:

$$T(x,y)= (-3x+7y,\ -x+5y)$$

En general, si se conocen las imágenes de los vectores de una base del dominio, se tiene completamente definida una transformación lineal. Formalizaremos este resultado mediante el siguiente:

Teorema VI. 4

Sea $B= \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ una base de un espacio vectorial V, y sean $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n$ n vectores cualesquiera de otro espacio W.

Entonces, existe una única transformación lineal

$$T:\ V \to W$$

tal que:

$$\left.\begin{aligned} T(\bar{v}_1)&=\bar{w}_1\\ T(\bar{v}_2)&=\bar{w}_2\\ &\vdots\\ T(\bar{v}_n)&=\bar{w}_n \end{aligned}\right\} \quad \ldots\ldots (1)$$

Demostración.

Si $\bar{v}$ es un vector cualquiera de V, podrá expresarse unívocamente en términos de la base B como:

$$\bar{v}=\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_n\bar{v}_n \qquad \ldots\ldots (2)$$

Definamos una transformación T por:

$$T(\bar{v})=\alpha_1\bar{w}_1+\alpha_2\bar{w}_2+\ldots+\alpha_n\bar{w}_n \qquad \ldots\ldots (3)$$

y demostremos que es lineal.

En efecto, si tomamos otro vector $\bar{u}$ de V tal que:

$\bar{u}=\beta_1\bar{v}_1+\beta_2\bar{v}_2+\ldots+\beta_n\bar{v}_n$,

se tiene que $\bar{v}+\bar{u}=(\alpha_1+\beta_1)\bar{v}_1+(\alpha_2+\beta_2)\bar{v}_2+\ldots+(\alpha_n+\beta_n)\bar{v}_n$

En consecuencia:

$$\begin{aligned} T(\bar{v}+\bar{u})&=(\alpha_1+\beta_1)\bar{w}_1+(\alpha_2+\beta_2)\bar{w}_2+\ldots+(\alpha_n+\beta_n)\bar{w}_n\\ &=\alpha_1\bar{w}_1+\alpha_2\bar{w}_2+\ldots+\alpha_n\bar{w}_n+\beta_1\bar{w}_1+\beta_2\bar{w}_2+\ldots+\beta_n\bar{w}_n\\ &=T(\bar{v})+T(\bar{u}) \end{aligned}$$

Además, si c es un escalar,

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

Por tanto

$$\begin{aligned} T(c\bar{v})&= c\alpha_1\bar{w}_1+c\alpha_2\bar{w}_2+\ldots+c\alpha_n\bar{w}_n\\ &=c(\alpha_1\bar{w}_1+\alpha_2\bar{w}_2+\ldots+\alpha_n\bar{w}_n)\\ &=cT(\bar{v}) \end{aligned}$$

Vemos que T, definida por (3) es lineal. Además, haciendo $\bar{v}=\bar{v}_1$, de (2) se obtiene que

$\alpha_1=1 \quad$ y $\alpha_2=\alpha_3=\ldots=\alpha_n=0$

Sustituyendo en (3), $T(\bar{v}_1)=\bar{w}_1$

Procediendo en forma similar, $T(\bar{v}_2)=\bar{w}_2$

$$\vdots$$

$$T(\bar{v}_n)=\bar{w}_n$$

Hemos demostrado que existe una transformación lineal que satisface (1). Resta sólo demostrar que dicha transformación es única.

Para ello supondremos que existen dos transformaciones lineales, T y T', y demostraremos que son iguales. Tomemos el vector $\bar{v}$ dado por (2). Entonces:

$$T'(\bar{v})=T'(\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_n\bar{v}_n)$$

Por linealidad: $T'(\bar{v})=\alpha_1T'(\bar{v}_1)+\alpha_2T'(\bar{v}_2)+\ldots+\alpha_nT'(\bar{v}_n)$

Por hipótesis, $T'(\overline{v}_1)=\overline{w}_1$

$T'(\overline{v}_2)=\overline{w}_2$

$\vdots$

$T'(\overline{v}_n)=\overline{w}_n$

Por tanto, $T'(\overline{v})=\alpha_1\overline{w}_1+\alpha_2\overline{w}_2+\ldots+\alpha_n\overline{w}_n$

de donde $T=T'$, quedando demostrado el teorema.

Veremos ahora que las transformaciones lineales pueden representarse por medio de matrices, para lo cual estudiemos el siguiente caso:

Tenemos una transformación lineal $T: R^3 \rightarrow R^2$ dada por la regla $T(x, y, z)=(x+y, x+z)$ . . . . . (1)

Es claro que $T(4, 2, 3)=(6, 7)$

Si consideramos la matriz:

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

y premultiplicamos el vector columna $\begin{bmatrix}4\\2\\3\end{bmatrix}$ por dicha matriz:

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

se obtiene la imagen del vector (4, 2, 3) según la transformación.

De hecho, para cualquier vector (x, y, z) de R^3:

$$\begin{bmatrix}1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+y\\x+z\end{bmatrix}$$

que es la regla de transformación dada por (1).

Podemos por tanto escribir:

$$T(\overline{v})=M\overline{v} \qquad \forall\ \overline{v}\in R^3$$

donde $\overline{v}$ es un vector columna.

La matriz M se obtiene a partir de las imágenes de los vectores de la base canónica de R^3:

$T(1, 0, 0)=(1, 1)$

$T(0, 1, 0)=(1, 0)$

$T(0, 0, 1)=(0, 1)$

las cuales, como puede verse, son las columnas de la matriz M, que recibe el nombre de "matriz asociada a la transformación, referida a las bases canónicas".

Ejemplo VI. 12.

Para la transformación $T: R^2 \rightarrow R^2$ definida por

$T(x,y)=(3y,2x)$

obtener su matriz asociada referida a la base canónica de R^2 y calcular $T(4, 2)$ empleándola.

Solución.

De la regla de transformación

$T(1, 0)=(0, 2)$

$T(0, 1)=(3, 0)$

y la matriz pedida es

$$\begin{bmatrix}0 & 3\\ 2 & 0\end{bmatrix}$$

para obtener $T(4, 2)$ efectuamos el producto

$$\begin{bmatrix}0 & 3\\ 2 & 0\end{bmatrix}\begin{bmatrix}4\\2\end{bmatrix}=\begin{bmatrix}6\\8\end{bmatrix}$$

por lo que $T(4, 2)=(6, 8)$.

- -

A partir de ahora trabajaremos exclusivamente con vectores columna, ya que esto facilita grandemente el manejo de las transformaciones lineales cuando se emplean las matrices asociadas.

Matriz asociada referida a dos bases cualesquiera.

Analicemos ahora el caso de una transformación lineal $T: R^3 \to R^2$, sabiendo que $A = \{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ es una base de R^3 y $B = \{\bar{w}_1, \bar{w}_2\}$ es una base de R^2. Se conoce además que:

$$\left.\begin{array}{l} T(\bar{v}_1) = 2\bar{w}_1 + 3\bar{w}_2 \\ T(\bar{v}_2) = -6\bar{w}_1 + 7\bar{w}_2 \\ T(\bar{v}_3) = \bar{w}_1 - 5\bar{w}_2 \end{array}\right\} \quad (1)$$

lo cual define completamente la transformación T, según el teorema VI. 4.

Supongamos que se tiene un vector $\bar{v} \in R^3$ tal que

$$\bar{v} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + \alpha_3\bar{v}_3$$

Entonces, como T es lineal:

$$T(\bar{v}) = \alpha_1 T(\bar{v}_1) + \alpha_2 T(\bar{v}_2) + \alpha_3 T(\bar{v}_3)$$

Utilizando las expresiones dadas por (1):

$$T(\bar{v}) = \alpha_1(2\bar{w}_1 + 3\bar{w}_2) + \alpha_2(-6\bar{w}_1 + 7\bar{w}_2) + \alpha_3(\bar{w}_1 - 5\bar{w}_2)$$

que puede escribirse como

$$T(\bar{v}) = (2\alpha_1 - 6\alpha_2 + \alpha_3)\bar{w}_1 + (3\alpha_1 + 7\alpha_2 - 5\alpha_3)\bar{w}_2$$

por lo que, el vector de coordenadas de $T(\bar{v})$ en la base B es:

$$(T(\bar{v}))_B = \begin{bmatrix} 2\alpha_1 - 6\alpha_2 + \alpha_3 \\ 3\alpha_1 + 7\alpha_2 - 5\alpha_3 \end{bmatrix}$$

Este vector puede obtenerse a partir del vector de coordenadas de $\bar{v}$ en la base A;

$$(\bar{v})_A = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

mediante el producto

$$(T(\bar{v}))_B = \begin{bmatrix} 2 & -6 & 1 \\ 3 & 7 & -5 \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

que representamos mediante

$$(T(\bar{v}))_B = M_B^A \, (\bar{v})_A$$

donde $M_B^A = \begin{bmatrix} 2 & -6 & 1 \\ 3 & 7 & -5 \end{bmatrix}$ es la matriz asociada a la transformación, referida a las bases A y B.

Es importante observar dos cosas:

1.- La matriz M_B^A transforma un vector de coordenadas en A, en un vector de coordenadas en B.

2.- Las columnas de M_B^A son las coordenadas de los vectores imagen de la base A en la base B.

Formalizaremos a continuación el estudio hecho anteriormente:

Teorema VI. 5

Sean V y W dos espacios vectoriales tales que 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, entonces existe una única matriz M_B^A, de orden mxn, tal que

$T(\bar{v})_B = M_B^A\ (\bar{v})_A$ para todo $\bar{v}$ de V, donde, si

$T(\bar{v}_i)=a_{1i}\bar{w}_1+a_{2i}\bar{w}_2+\ldots+a_{mi}\bar{w}_m$

entonces, la iésima columna de M_B^A es

$$\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} = (T(\bar{v}_i))_B$$

Este teorema nos garantiza que, para encontrar las coordenadas de la imagen de un vector $\bar{v}$ referidas a la base B de W, bastará con premultiplicar al vector de coordenadas de $\bar{v}$ en la base A por la matriz M_B^A, a la que llamaremos "matriz de la transformación referida a las bases A y B".

Demostración.

Si $\bar{v}$ es un vector cualquiera de V, sabemos que se puede expresar unívocamente en términos de la base A como:

$\bar{v}=\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_n\bar{v}_n$ . . . . . (1)

$\therefore\ T(\bar{v})=T(\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_n\bar{v}_n)$

Por linealidad:

$T(\bar{v})=\alpha_1T(\bar{v}_1)+\alpha_2T(\bar{v}_2)+\ldots+\alpha_nT(\bar{v}_n)$ . . . . . (2)

Por otra parte, los vectores imagen $T(\bar{v}_1), \ldots, T(\bar{v}_n)$ se pueden obtener en forma única a partir de la base B:

$T(\bar{v}_1)=a_{11}\bar{w}_1+a_{21}\bar{w}_2+\ldots+a_{m1}\bar{w}_m$

$T(\bar{v}_2)=a_{12}\bar{w}_1+a_{22}\bar{w}_2+\ldots+a_{m2}\bar{w}_m$

. . . .

$T(\bar{v}_n)=a_{1n}\bar{w}_1+a_{2n}\bar{w}_2+\ldots+a_{mn}\bar{w}_m$

Llevando estas expresiones a (2):

$T(\bar{v})=\alpha_1(a_{11}\bar{w}_1+a_{21}\bar{w}_2+\ldots+a_{m1}\bar{w}_m)+\alpha_2(a_{12}\bar{w}_1+a_{22}\bar{w}_2+\ldots+a_{m2}\bar{w}_m)+$
$\quad+\ldots+\alpha_n(a_{1n}\bar{w}_1+a_{2n}\bar{w}_2+\ldots+a_{mn}\bar{w}_m)$.

Agrupando:

$T(\bar{v})=(\alpha_1a_{11}+\alpha_2a_{12}+\ldots+\alpha_na_{1n})\bar{w}_1+(\alpha_1a_{21}+\alpha_2a_{22}+\ldots+\alpha_na_{2n})\bar{w}_2+$
$\quad+\ldots+(\alpha_1a_{m1}+\alpha_2a_{m2}+\ldots+\alpha_na_{mn})\bar{w}_m$

y, por definición, las sumas encerradas entre paréntesis son las coordenadas de $T(\bar{v})$ referidas a la base B, por lo que podemos escribir:

$$(T(\bar{v}))_B=\begin{bmatrix} \alpha_1a_{11}+\alpha_2a_{12}+\ldots+\alpha_na_{1n} \\ \alpha_1a_{21}+\alpha_2a_{22}+\ldots+\alpha_na_{2n} \\ \vdots \\ \alpha_1a_{m1}+\alpha_2a_{m2}+\ldots+\alpha_na_{mn} \end{bmatrix}$$

o, en forma matricial:

$$(T(\bar{v}))_B=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

y, por (1):
$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = (\bar{v})_A$$

es decir:

$$(T(\bar{v}))_B = M_B^A \, (v)_A \qquad \ldots . (3)$$

donde: $$M_B^A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

como se quería demostrar.

La prueba de que esta matriz es única es inmediata, ya que si suponemos que existe N_B^A tal que

$$(T(\bar{v}))_B = N_B^A \, (\bar{v})_A , \quad \forall \, \bar{v} \in V \qquad \ldots . (4)$$

las ecuaciones (3) y (4) claramente implican que:

$$N_B^A = M_B^A$$

Esto completa la demostración.

Ejemplo VI. 13.

Sea la transformación lineal $T: R^3 \to R^2$

definida por: $T(x, y, z) = (z, x+y)$

Fácilmente se comprueba que $A=\{(1, 2, 3), (0, -2, -1), (1, 1, 1)\}$ y $B=\{(1, 2), (0, 3)\}$ son bases de R^3 y R^2, respectivamente.

Calculemos la representación matricial de T en estas dos bases.

Es claro que: $T(1,2,3) = (3,3) = 3(1,2)-1(0,3)$

$T(0,-2,-1) = (-1,-2) = -1(1,2)+0(0,3)$

$T(1,1,1) = (1,2) = 1(1,2)+0(0,3)$

de donde

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

Utilicemos esta matriz para obtener la imagen del vector $\bar{v}=(-1,4,6)$, cuyo vector de coordenadas en la base A es:

$$(\bar{v})_A = \begin{bmatrix} 3 \\ -1 \\ -4 \end{bmatrix}$$

Entonces: $$(T(-1,4,6))_B = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ -4 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$$

El vector (6, -3) es el vector de coordenadas de $T(\bar{v})$ en la base B.

- -

Ejemplo VI. 14.

Volvamos a la transformación lineal del ejemplo VI.12

$$T: R^2 \to R^2$$

definida por $T(x,y) = (3y, 2x)$, y consideremos la base canónica $E = \{(1,0), (0,1)\}$.

Dado que $T(1,0) = (0,2) = 0(1,0)+2(0,1)$

$T(0,1) = (3,0) = 3(1,0)+0(0,1)$

tenemos $$M_E^E = \begin{bmatrix} 0 & 3 \\ 2 & 0 \end{bmatrix}$$

Por ello $(T(\bar{v}))_E = M_E^E \, (\bar{v})_E$ para todo $\bar{v}$ de R^2.

Cuando trabajemos con matrices referidas a bases canónicas, omitiremos los índices.

Para este caso: $M_B^E = M$

$\therefore$ $T(\overline{v}) = M\overline{v}$ para todo $\overline{v}$ de R^2, como se obtuvo en la página 31.

- - - - - - - - - - - - - - - - - - - -

Ejemplo VI. 15.

Sean P el espacio vectorial de polinomios de grado menor o igual que tres y Q el espacio vectorial de polinomios de grado menor o igual que dos.

Como sabemos, una base de P es $B = \{1, x, x^2, x^3\}$ y una de Q es $C = \{1, x, x^2\}$.

Consideremos el operador derivada $D: P \to Q$ y encontremos su representación matricial en las bases B y C.

Calculemos, para ello, las imágenes de los vectores de la base B refiriéndolas a la base C:

$$D(1) = 0 = 0(1) + 0(x) + 0(x^2)$$
$$D(x) = 1 = 1(1) + 0(x) + 0(x^2)$$
$$D(x^2) = 2x = 0(1) + 2(x) + 0(x^2)$$
$$D(x^3) = 3x^2 = 0(1) + 0(x) + 3(x^2)$$

Por tanto:

$$M_C^B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Calculemos $D(6x^2+9x^3-1)$ utilizando esta matriz; sabemos que:

$$(6x^2+9x^3-1)_B = \begin{bmatrix} -1 \\ 0 \\ 6 \\ 9 \end{bmatrix}$$

$$\therefore \quad (D(6x^2+9x^3-1))_C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 12 \\ 27 \end{bmatrix}$$

$$\therefore \quad D(6x^2+9x^3-1) = 0(1) + 12(x) + 27(x^2)$$

$$D(6x^2+9x^3-1) = 12x + 27x^2$$

que se comprueba con facilidad derivando directamente.

Para hacer énfasis en el hecho de que se trabaja con bases ordenadas, consideremos ahora como base de P a

$F = \{x^2, x, x^3, 1\}$

En este caso, la representación matricial de D es

$$M_C^F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}$$

- -

<u>Rango de la matriz de transformación.</u>

Sea la transformación $T: R^3 \to R^2$ definida por $T(x, y, z) = (x, 2x)$

Encontremos la matriz de la transformación referida a las bases canónicas:

$$T(1,0,0) = (1,2)$$
$$T(0,1,0) = (0,0)$$
$$T(0,0,1) = (0,0)$$

$$\therefore \quad M = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

Es claro que el rango de la matriz M es uno:

$$R(M) = 1$$

que es precisamente la dimensión del recorrido de la transformación.

Se deja al estudiante la demostración del siguiente teorema.

Teorema VI. 6.

En una transformación lineal, la dimensión del recorrido es igual al rango de la matriz de la transformación referida a dos bases cualesquiera.

VI. 4. ALGEBRA DE LAS TRANSFORMACIONES LINEALES.

En las secciones anteriores hemos trabajado únicamente con una transformación a la vez, estudiando sus propiedades y características particulares.

En esta sección utilizaremos simultáneamente dos o más transformaciones lineales, con el objeto de estudiar las operaciones que con ellas pueden efectuarse.

Espacio vectorial de transformaciones lineales.

Consideremos el caso de dos transformaciones lineales T y S de R^3 en R^2 definidas por:

$$T(x,y,z)=(x,2x)$$
y
$$S(x,y,z)=(x+y,x+z)$$

y hallemos la imagen del vector $(4,2,3)$ según ambas transformaciones:

$$T(4,2,3)=(4,8)$$
$$S(4,2,3)=(6,7)$$

Podemos formar ahora una nueva transformación $T+S$ tal que, la imagen de un vector de R^3 según $T+S$ sea la suma de las imágenes de ese mismo vector según T y S.

Por ejemplo:

$$\begin{aligned}(T+S)(4,2,3)&=T(4,2,3)+S(4,2,3)\\&=(4,8)+(6,7)\\&=(10,15)\end{aligned}$$

De igual forma, para el vector $(1,5,-2)$:

$$\begin{aligned}(T+S)(1,5,-2)&=T(1,5,-2)+S(1,5,-2)\\&=(1,2)+(6,-1)\\&=(7,1)\end{aligned}$$

La regla de transformación de T+S puede obtenerse calculando la imagen de un vector cualquiera (x,y,z) de R^3:

$(T+S)(x,y,z)=T(x,y,z)+S(x,y,z)$

$(T+S)(x,y,z)=(x,2x)+(x+y,\ x+z)$

$(T+S)(x,y,z)=(2x+y,\ 3x+z)$ . . . . (1)

Es fácil verificar que utilizando esta regla se obtienen los mismos resultados anteriores.

Un estudio similar se puede llevar a cabo para una operación consistente en multiplicar una transformación lineal por un escalar.

Por ejemplo, a partir de la transformación S podemos definir la transformación 5S en la siguiente forma:

$(5S)(\bar{v})=5S(\bar{v})\quad \forall\ \bar{v}$ de R^3,

Calculemos algunos casos:

$(5S)(4,2,3)=5S(4,2,3)$

$=5\ (6,7)$

$=(30,35)$

Análogamente, para el vector $(1,5,-2)$:

$(5S)(1,5,-2)=5S(1,5,-2)$

$=5\ (6,-1)$

$=(30,-5)$

y, para obtener la regla de transformación:

$(5S)(x,y,z)=5S(x,y,z)$

$(5S)(x,y,z)=5\ (x+y,\ x+z)$

$(5S)(x,y,z)=(5x+5y,\ 5x+5z)$ . . . . (2)

El estudiante puede demostrar que las transformaciones T+S y 5S, definidas por (1) y (2) respectivamente, son lineales.

A continuación definiremos la suma de transformaciones y el producto por un escalar para el caso general:

Definición.

Sean V y W dos espacios vectoriales sobre un mismo campo K, y sean T y S dos transformaciones de V en W:

1) La suma T+S es una transformación tal que:

$(T+S)(\bar{v})=T(\bar{v})+S(\bar{v}),\quad \forall\ \bar{v}\in V.$

2) El producto αT es una transformación tal que:

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

Teorema VI. 7.

Si T y S son transformaciones lineales, entonces T+S y αT son también lineales.

Demostración.

a) T+S es lineal.

Sean $\bar{v}_1$ y $\bar{v}_2$ dos vectores de V, y c un escalar de K:

$(T+S)(c\bar{v}_1+\bar{v}_2)=T(c\bar{v}_1+\bar{v}_2)+S(c\bar{v}_1+\bar{v}_2)$	Por definición de T+S.
$=cT(\bar{v}_1)+T(\bar{v}_2)+cS(\bar{v}_1)+S(\bar{v}_2)$	Por linealidad de T y S.
$=cT(\bar{v}_1)+cS(\bar{v}_1)+T(\bar{v}_2)+S(\bar{v}_2)$	
$=c(T(\bar{v}_1)+S(\bar{v}_1))+T(\bar{v}_2)+S(\bar{v}_2)$	
$=c(T+S)(\bar{v}_1)+(T+S)(\bar{v}_2)$	Por definición de T + S.

quedando demostrado.

b) αT es lineal.

En efecto, sean $\bar{v}_1$ y $\bar{v}_2$ dos vectores de V, y c un escalar:

$(\alpha T)(c\bar{v}_1+\bar{v}_2)=\alpha T(c\bar{v}_1+\bar{v}_2)$	Por definición de αT.
$=\alpha(cT(\bar{v}_1)+T(\bar{v}_2))$	Por linealidad de T.
$=\alpha(cT(\bar{v}_1))+\alpha T(\bar{v}_2)$	
$=c(\alpha T(\bar{v}_1))+\alpha T(\bar{v}_2)$	
$=c(\alpha T)(\bar{v}_1)+(\alpha T)(\bar{v}_2)$	Por definición de αT.

quedando demostrado.

Teorema VI. 8.

Sean V y W dos espacios vectoriales y sea L(V,W) el conjunto formado por todas las transformaciones lineales de V en W.

L(V,W) es un espacio vectorial.

De las definiciones dadas para la suma de transformaciones y producto por un escalar, es fácil demostrar el teorema anterior. Unicamente enumeraremos las propiedades de L(V,W), dejando como ejercicio al estudiante su demostración:

Sean S,T y U transformaciones lineales de V en W, es decir: $S, T, U \in L(V,W)$.

1) Cerradura.

Ya demostramos que la suma de transformaciones lineales es lineal, esto es, $S+T \in L(V,W)$.

También demostramos que el producto de una transformación lineal por un escalar es lineal, esto es, $\alpha T \in L(V, W)$.

2) $S+(T+U)=(S+T)+U$.

3) Transformación cero:

$\exists 0 \in L(V,W)$ tal que, $\forall T \in L(V,W)$:

$T+0=0+T=T$.

4) $T+(-T)=0$, donde $-T=(-1)T \in L(V,W)$.

5) $T+S=S+T$.

Sean α y β dos escalares cualesquiera:

6) $\alpha(\beta T)=(\alpha\beta)T$.

7) $\alpha(T+S)=\alpha T+\alpha S$.

8) $(\alpha+\beta)T=\alpha T+\beta T$.

9) $1 \cdot T=T$.

- -

Volvamos ahora a las transformaciones T y S definidas por:

$T(x,y,z)=(x,2x)$

y

$S(x,y,z)=(x+y,\ x+z)$

para las cuales se obtuvo:

$(T+S)(x,y,z)=(2x+y,\ 3x+z)$ . . . . (1)

La suma de estas transformaciones se puede llevar a cabo por medio de las matrices asociadas, referidas a las bases canónicas. Representaremos con M(T), M(S) y M(T+S) a las matrices asociadas a T, S y T+S respectivamente; dichas matrices son:

$$M(T)=\begin{bmatrix}1 & 0 & 0\\ 2 & 0 & 0\end{bmatrix} \qquad M(S)=\begin{bmatrix}1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$$

$$\text{y} \quad M(T+S)=\begin{bmatrix}2 & 1 & 0\\ 3 & 0 & 1\end{bmatrix}$$

Como es fácil comprobar, esta última matriz puede también obtenerse sumando las matrices M(T) y M(S). Es decir:

$M(T+S) = M(T) + M(S)$.

Recordemos que al inicio de esta sección, también se obtuvo la transformación 5S cuya regla de correspondencia es:

$(5S)(x,y,z)=(5x+5y,\ 5x+5z)$ . . . . (2)

y en forma análoga, vemos que el multiplicar por el escalar 5 la matriz M(S) da como resultado M(5S):

$$5M(S)=5\begin{bmatrix}1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}=\begin{bmatrix}5 & 5 & 0\\ 5 & 0 & 5\end{bmatrix}=M(5S).$$

En general, si V y W son dos espacios vectoriales y S y T dos transformaciones lineales cualesquiera de L(V,W), se tiene:

para la suma:	para el producto por un escalar:
si $M(T) \leftrightarrow T$	si $M(T) \leftrightarrow T$
y $M(S) \leftrightarrow S$	y α es cualquier escalar,
entonces:	entonces:
$M(T)+M(S) \leftrightarrow T+S$	$\alpha M(T) \leftrightarrow \alpha T$.
o sea:	o sea:
$M(T)+M(S)=M(T+S)$	$\alpha M(T) = M(\alpha T)$.

Por lo que, sumar transformaciones o multiplicarlas por un escalar es equivalente a sumar sus matrices asociadas o multiplicarlas por un escalar.

Se dice que $L(V,W)$ es isomorfo al conjunto de las matrices de orden $m\times n$ (donde $m=\dim W$ y $n=\dim V$) para las operaciones de suma y producto por un escalar.

Teorema VI. 9.

(Teorema de Isomorfismo)

Para cualesquiera T y S de $L(V,W)$ y todos los escalares α se tiene que:

$$M(T+S) = M(T)+M(S)$$

y

$$M(\alpha T) = \alpha M(T)$$

Demostraremos este teorema únicamente para transformaciones de R^n a R^m, empleando las bases canónicas. (El caso general se puede demostrar en forma análoga, teniendo en cuenta que las matrices que se van a operar deben estar referidas a las mismas bases)

Sean $T: R^n \to R^m$ y $S: R^n \to R^m$ dos transformaciones lineales y sean $M(T)$ y $M(S)$ sus respectivas matrices asociadas, en las bases canónicas.

Sabemos que
$$\left.\begin{array}{l} T(\bar{v})=M(T)\bar{v} \\ \text{y} \quad S(\bar{v})=M(S)\bar{v}, \quad \forall\, \bar{v} \text{ de } R^n \end{array}\right\} \quad \ldots\ldots (1)$$

a) Para la suma:

por definición de $T+S$:

$$(T+S)(\bar{v})=T(\bar{v})+S(\bar{v})$$

por (1):

$$(T+S)(\bar{v})=M(T)\bar{v}+M(S)\bar{v}$$

por la distributividad del producto sobre la suma de matrices:

$$(T+S)(\bar{v})=(M(T)+M(S))\bar{v} \quad \ldots\ldots (2)$$

por otra parte, de (1):

$$(T+S)(\bar{v})=M(T+S)\bar{v} \quad \ldots\ldots (3)$$

de (2) y (3) se sigue que

$$M(T+S)=M(T)+M(S).$$

b) Para el producto por un escalar:

por definición de αT:

$$(\alpha T)(\bar{v})=\alpha T(\bar{v})$$

por (1):

$$(\alpha T)(\bar{v})=\alpha(M(T)\bar{v})$$

por las propiedades de las matrices para el producto por un escalar:

$$(\alpha T)(\bar{v})=(\alpha M(T))\bar{v} \quad \ldots\ldots (4)$$

por otra parte, de (1):

$$(\alpha T)(\bar{v})=M(\alpha T)\bar{v} \quad \ldots\ldots (5)$$

de (4) y (5) se sigue que

$$M(\alpha T)=\alpha M(T).$$

Esto completa la demostración.

Composición de transformaciones lineales.

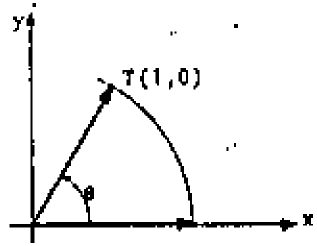
Una operación bastante más interesante con transformaciones lineales es la composición.

Consideremos el problema siguiente:

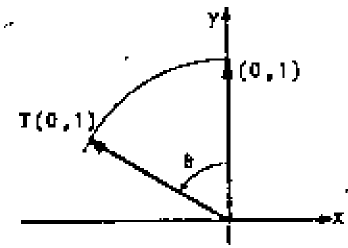
Se tiene un segmento dirigido en el espacio euclidiano bidimensional y se desea girarlo un ángulo de 60° (sentido antihorario), y una vez girado, se desea aumentar ocho veces su magnitud.

Resolvamos este problema en dos partes:

Primero, obtengamos la transformación lineal T según la cual un vector del plano es girado un ángulo de 60°. Por el teorema VI. 4, bastará con hallar la imagen de los vectores de la base canónica para poder definir T. Tenemos pues, para un ángulo cualquiera θ:



$T(1,0) = (\cos\theta,\ \operatorname{sen}\theta)$



$T(0,1) = (-\operatorname{sen}\theta,\ \cos\theta)$

Por tanto, la matriz de la transformación en la base canónica es:

$$M(T) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \ldots\ldots (1)$$

Haciendo $\theta = 60°$:

$$M(T) = \begin{bmatrix} \cos 60° & -\operatorname{sen} 60° \\ \operatorname{sen} 60° & \cos 60° \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \quad \ldots\ldots (2)$$

En segundo lugar, encontramos la transformación lineal S que multiplica por 8 la magnitud de un vector.

Este es un caso inmediato: $S(1,0) = (8,0)$

$S(0,1) = (0,8)$

$$\therefore \quad M(S) = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} \quad \ldots\ldots (3)$$

Ya podemos resolver nuestro problema:

Por ejemplo, para el vector (2,4), primero lo giramos 60°:

$$T(2,4) = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1-2\sqrt{3} \\ \sqrt{3}+2 \end{bmatrix}$$

$T(2,4) = (1-2\sqrt{3},\ \sqrt{3}+2)$

y después lo multiplicamos por 8:

$$S(1-2\sqrt{3},\ \sqrt{3}+2) = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 1-2\sqrt{3} \\ \sqrt{3}+2 \end{bmatrix} = \begin{bmatrix} 8-16\sqrt{3} \\ 8\sqrt{3}+16 \end{bmatrix}$$

$S(1-2\sqrt{3},\ \sqrt{3}+2) = (8-16\sqrt{3},\ 8\sqrt{3}+16)$

Vemos que el proceso aquí seguido se puede resumir en la expresión: $S(T(2,4)) = (8-16\sqrt{3},\ 8\sqrt{3}+16)$

que es la composición de las transformaciones S y T, simbolizada con $S \circ T$. (Léase: "S composición T").

Esto es, $S \circ T$ es una transformación que se define como:

$(S \circ T)(\bar{v}) = S(T(\bar{v})) \quad \forall\ \bar{v}$ de $\mathbb{R}^2$

Por ello podemos escribir

$(S \circ T)(2,4) = (8-16\sqrt{3},\ 8\sqrt{3}+16)$

donde $S \circ T$ es justamente la transformación que al mismo tiempo gira un vector 60° y lo multiplica por 8.

Si queremos obtener la matriz asociada a $S \circ T$ podemos hacer lo siguiente:

$$(S\circ T)(x,y)=S(T(x,y))$$

$$(S\circ T)(x,y)=S\left[\begin{bmatrix}\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}\right] \qquad \text{Utilizando } M(T)$$

$$(S\circ T)(x,y)=\begin{bmatrix}8 & 0\\ 0 & 8\end{bmatrix}\begin{bmatrix}\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} \qquad \text{Utilizando } M(S)$$

$$(S\circ T)(x,y)=\begin{bmatrix}4 & -4\sqrt{3}\\ 4\sqrt{3} & 4\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

$$\therefore\; M(S\circ T)=\begin{bmatrix}4 & -4\sqrt{3}\\ 4\sqrt{3} & 4\end{bmatrix}$$

Observamos que la matriz asociada a $S\circ T$ es el producto de las matrices asociadas a S y T, esto es: $M(S\circ T)=M(S)M(T)$.

Utilizando esta matriz, podemos obtener directamente

$$(S\circ T)(2,4)=\begin{bmatrix}4 & -4\sqrt{3}\\ 4\sqrt{3} & 4\end{bmatrix}\begin{bmatrix}2\\ 4\end{bmatrix}=\begin{bmatrix}8-16\sqrt{3}\\ 8\sqrt{3}+16\end{bmatrix}$$

como se tenía.

En resumen, $S\circ T$ es una transformación lineal que "reune los efectos" de S y T, y cuya matriz asociada es el producto de las matrices $M(S)$ y $M(T)$ (en ese orden).

Veamos otro caso:

Sean dos transformaciones lineales

$G: \mathbb{R}^4 \to \mathbb{R}^3$ dada por $G(x,y,z,w)=(2x,y,z+w)$

$F: \mathbb{R}^3 \to \mathbb{R}^2$ dada por $F(x,y,z)=(3x+y,z)$

El estudiante puede verificar fácilmente que

$$M(G)=\begin{bmatrix}2 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\end{bmatrix} \quad \text{y} \quad M(F)=\begin{bmatrix}3 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

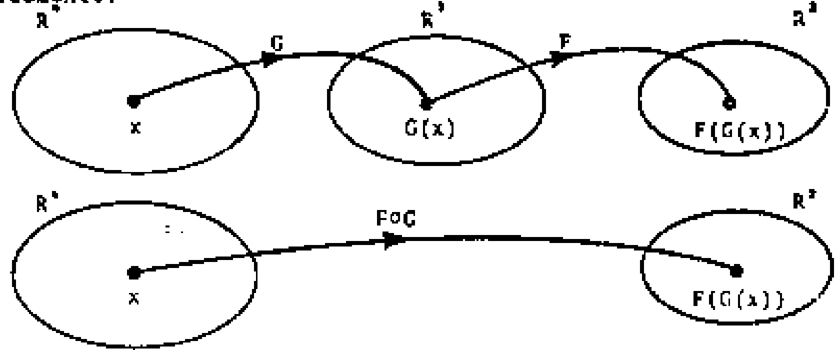
son las matrices asociadas a G y F, en las bases canónicas.

Encontremos la matriz asociada a $F\circ G$:

$$M(F\circ G)=M(F)M(G)=\begin{bmatrix}3 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\end{bmatrix}$$

$$\therefore\; M(F\circ G)=\begin{bmatrix}6 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\end{bmatrix}$$

Es claro que $F\circ G$ es una transformación de $\mathbb{R}^4$ en $\mathbb{R}^2$. Gráficamente:



Encontremos por ejemplo la imagen del vector $(-1,4,2,8)$ según $F\circ G$:

$$(F\circ G)(-1,4,2,8)=\begin{bmatrix}6 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\end{bmatrix}\begin{bmatrix}-1\\ 4\\ 2\\ 8\end{bmatrix}=\begin{bmatrix}-2\\ 10\end{bmatrix}$$

Este resultado se puede verificar fácilmente si observamos que $G(-1,4,2,8)=(-2,4,10)$

y $F(-2,4,10)=(-2,10)$

es decir, $(F\circ G)(-1,4,2,8)=F(G(-1,4,2,8))$

o, en general:

$$(F\circ G)(\overline{v})=F(G(\overline{v})) \quad \forall\ \overline{v} \text{ de } \mathbb{R}^4,$$

que es la definición de $F\circ G$.

Daremos ahora la definición para el caso general:

Definición.

Sean las transformaciones $T: U \to V$ y $S: V \to W$.

Definimos la transformación

$S \circ T: U \to W$

como

$(S \circ T)(\bar{u}) = S(T(\bar{u}))$ $\forall\ \bar{u}$ de U.

Podemos representar gráficamente la composición de transformaciones lineales mediante la siguiente figura:

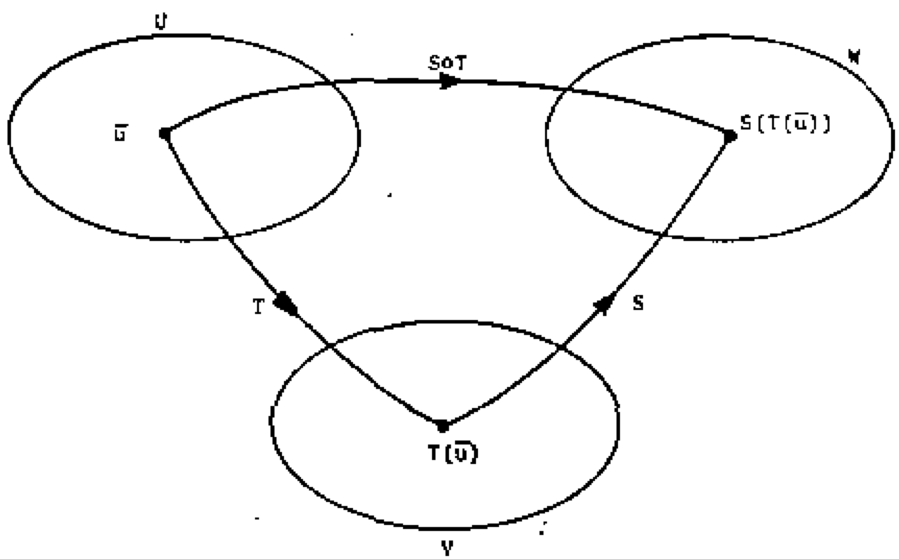


Figura VI. 2.

Teorema VI. 10

Si $T: U \to V$ y $S: V \to W$ son dos transformaciones lineales, entonces

$S \circ T: U \to W$

también es lineal.

Demostración.

Sean $\bar{u}_1$ y $\bar{u}_2$ dos vectores de U, y sea α un escalar; entonces:

$(S \circ T)(\alpha\bar{u}_1 + \bar{u}_2) = S(T(\alpha\bar{u}_1 + \bar{u}_2))$	Por definición de $S \circ T$.
$= S(\alpha T(\bar{u}_1) + T(\bar{u}_2))$	Por linealidad de T.
$= \alpha S(T(\bar{u}_1)) + S(T(\bar{u}_2))$	Por linealidad de S.
$= \alpha(S \circ T)(\bar{u}_1) + (S \circ T)(\bar{u}_2)$	Por definición de $S \circ T$.

quedando demostrado.

Hemos mencionado el hecho de que la composición de transformaciones lineales corresponde a la multiplicación de matrices. Esto se puede generalizar con el siguiente:

Teorema VI. 11.

Sean las transformaciones lineales $T: U \to V$ y $S: V \to W$.

Sean A, B y C bases de U, V y W, respectivamente. Si $M_B^A(T)$ es la matriz asociada a T, y $M_C^B(S)$ es la matriz asociada a S, entonces

$M_C^B(S) M_B^A(T) = M_C^A(S \circ T)$

es la matriz asociada a $S \circ T$.

La demostración de este teorema puede consultarse en la referencia 2, pág. 127.

Ejemplo VI. 16.

Teníamos (ver ejemplo VI. 13) una transformación lineal $T: R^3 \to R^2$ dada por $T(x,y,z) = (z, x+y)$.

Siendo $A = \{(1,2,3), (0,-2,-1), (1,1,1)\}$ y $B = \{(1,2), (0,3)\}$ dos bases de R^3 y R^2 respectivamente, habíamos obtenido

$$M_B^A(T) = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 0 & 0 \end{bmatrix}$$

Si ahora consideramos la transformación lineal $S: R^2 \to R^2$ dada por

$$S(1,2)=(1,2)$$
y
$$S(0,3)=(0,3)$$

y utilizamos la base canónica $E=\{(1,0),(0,1)\}$ se tendrá:

$$M_E^B(S)=\begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix}$$

Por tanto, la matriz asociada a $S\circ T$ es:

$$M_E^B(S)\,M_B^A(T)=\begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix}\begin{bmatrix}3 & -1 & 1\\ -1 & 0 & 0\end{bmatrix}$$

$$=\begin{bmatrix}3 & -1 & 1\\ 3 & -2 & 2\end{bmatrix}=M_E^A(S\circ T)$$

Encontremos $(S\circ T)(\bar{v})$, donde $\bar{v}=(-1,4,6)$.

Sabemos que: $(\bar{v})_A=\begin{bmatrix}3\\ -1\\ -4\end{bmatrix}$

$$((S\circ T)(\bar{v}))_E=M_E^A(S\circ T)\;(\bar{v})_A$$

$$(S\circ T)(-1,4,6)=\begin{bmatrix}3 & -1 & 1\\ 3 & -2 & 2\end{bmatrix}\begin{bmatrix}3\\ -1\\ -4\end{bmatrix}=\begin{bmatrix}6\\ 3\end{bmatrix}$$

$$(S\circ T)(-1,4,6)=(6,3)$$

Vemos que en este caso la transformación S equivale simplemente a un cambio de base en R^2 (ver resultado del ejemplo VI.13).

- -

<u>Algunas propiedades de las operaciones con transformaciones lineales</u>

Sean U, V y W tres espacios vectoriales sobre un campo K.

1.- Si $F: U \to V$

$S: V \to W$

$T: V \to W$

son transformaciones lineales, entonces

$$(S+T)\circ F=S\circ F+T\circ F.$$

2.- Si $F: U \to V$

y $T: V \to W$

son transformaciones lineales y α es un escalar, entonces

$$(\alpha T)\circ F=\alpha(T\circ F)$$

3.- Si $F: U \to V$

$G: U \to V$

y $T: V \to W$

son transformaciones lineales, entonces

$$T\circ(F+G)=T\circ F+T\circ G$$

4.- Si X es un espacio vectorial y

$F: U \to V$

$S: V \to W$

$T: W \to X$

son transformaciones lineales, entonces

$$T\circ(S\circ F)=(T\circ S)\circ F.$$

5.- Si $T: V \to W$ es una transformación lineal, entonces:

$$T\circ I_V=T$$
y
$$I_W\circ T=T$$

donde I_V e I_W son las transformaciones identidad en los espacios V y W, respectivamente (ver definición en el ejemplo VI. 6).

Observe que estas propiedades son equivalentes a las ya obtenidas para operaciones con matrices, en el capítulo IV.

Nótese que, salvo casos muy particulares, la composición de transformaciones lineales NO ES CONMUTATIVA (resultado análogo al de multiplicación de matrices).

Transformación inversa.

Veamos la transformación lineal $G: R^2 \rightarrow R^2$ definida por $G(x,y)=(3y, y)$, y consideremos el problema de determinar, dado un vector del recorrido, de que vector del dominio es imagen.

Por ejemplo, sabiendo que $G(\overline{v})=(6,2)$, queremos determinar $\overline{v} \in R^2$.

Surge sin embargo una dificultad: varios vectores del dominio tienen a (6,2) como imagen. Mencionemos por ejemplo

$G(-3,2)=(6,2)$

$G(\ 8,2)=(6,2)$

$G(\ 0,2)=(6,2)$

Ya que un vector dado del recorrido es imagen de más de un vector del dominio, observamos que el problema que nos planteamos NO tiene solución única.

Analicemos otro caso:

Como vimos en la página 48, la transformación $T: R^2 \rightarrow R^2$ cuya matriz asociada es

$$M(T)=\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

gira un vector del plano un ángulo de 60° (sentido antihorario).

Es fácil pensar que conociendo un vector imagen se puede determinar con certeza de que vector del dominio procede, ya que bastará simplemente con girarlo un ángulo de -60° (60° sentido horario) para obtener el vector original.

Por ejemplo, si $T(\overline{v}_1)=(1-2\sqrt{3}, \sqrt{3}+2)$, deseamos determinar $\overline{v}_1$ de R^2.

Obtengamos para ello la transformación $S: R^2 \rightarrow R^2$ que gira un vector un ángulo de -60°. Haciendo $\theta=-60°$ en la expresión (1) de la página 48, tenemos:

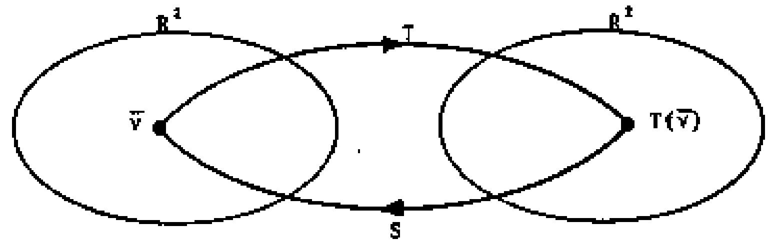
$$M(S)=\begin{bmatrix} \cos(-60°) & -\mathrm{sen}(-60°) \\ \mathrm{sen}(-60°) & \cos(-60°) \end{bmatrix}=\begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

Entonces:

$$S(1-2\sqrt{3}, \sqrt{3}+2)=\begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1-2\sqrt{3} \\ \sqrt{3}+2 \end{bmatrix}=\begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$\therefore \quad \overline{v}_1 = (2,4)$$

Hemos resuelto el problema, pues ya tenemos una transformación (S) que nos permite "regresar" cualquier vector del recorrido al vector del dominio del cual es imagen. Gráficamente:



A S le llamaremos transformación inversa de T y la representaremos mediante T^{-1}.

Para el caso que hemos analizado:

$T(2,4)=(1-2\sqrt{3}, \sqrt{3}+2)$

y $T^{-1}(1-2\sqrt{3}, \sqrt{3}+2)=(2,4)$

Esto es válido para todo vector $\overline{v}$ de R^2, es decir:

$T(\overline{v}) = \overline{u}$

y $T^{-1}(\overline{u})= \overline{v}$

De estas expresiones es claro que:

$$T^{-1}(T(\bar{v}))=\bar{v}$$
y
$$T(T^{-1}(\bar{u}))=\bar{u} \qquad \forall\ \bar{v},\ \bar{u} \in R^2$$

Definición.

Sea una transformación $T: V \to W$. A la transformación $T^{-1}: W \to V$ tal que

$$T^{-1}(T(\bar{v}))=\bar{v} \qquad \forall\ \bar{v} \in V \qquad \ldots\ldots (1)$$
y
$$T(T^{-1}(\bar{w}))=\bar{w} \qquad \forall\ \bar{w} \in W \qquad \ldots\ldots (2).$$

se le llama transformación inversa de T.

De (1) y (2) se tiene que:

$$(T^{-1}oT)(\bar{v}) = \bar{v} \qquad \forall\ \bar{v} \in V$$
y
$$(ToT^{-1})(\bar{w}) = \bar{w} \qquad \forall\ \bar{w} \in W$$

y, por definición de transformación identidad:

$$T^{-1}oT = I_V$$
$$ToT^{-1} = I_W$$

Estas expresiones, son equivalentes a las condiciones (1) y (2) de la definición.

Es lógico preguntarse ahora cuándo una transformación tiene inversa. Para responder a ésto, conviene comparar algunas características de las transformaciones G y T de los dos ejemplos anteriores:

Como vimos, para la transformación $G: R^2 \to R^2$ definida por $G(x,y)=(3y,\ y)$ existen vectores diferentes del dominio que tienen como imagen un mismo vector del recorrido. Entonces, no existe una transformación G^{-1} tal que $G^{-1}(G(\bar{v}))=\bar{v} \quad \forall\ \bar{v} \in R^2$, por lo que G no tiene inversa.

Por el contrario, para la transformación $T: R^2 \to R^2$ que gira un vector 60° (para la cual obtuvimos una transformación inversa), es claro que dos vectores diferentes del dominio tienen siempre imágenes distintas en el recorrido (se dice que T es uno a uno). Dicho de otra forma:

$$T(\bar{v}_1)=T(\bar{v}_2) \Rightarrow \bar{v}_1=\bar{v}_2 \qquad \forall\ \bar{v}_1,\ \bar{v}_2 \in R$$

Además, es fácil verificar que todos los vectores de R^2 son imagen de algún vector del dominio, o sea, el recorrido de la transformación $T: R^2 \to R^2$ es todo R^2 (se dice que T es sobre).

Estas diferencias básicas entre las transformaciones G y T son las que determinan la existencia de la transformación inversa, como se establece a continuación de la siguiente definición.

Definición.

Decimos que la transformación $T: V \to W$ es biyectiva si:

1) T es uno a uno, esto es:

$$T(\bar{v}_1)=T(\bar{v}_2) \Rightarrow \bar{v}_1=\bar{v}_2 \qquad \forall\ \bar{v}_1,\ \bar{v}_2 \text{ de } V. \qquad \text{y}$$

2) T es sobre, esto es:

$$T(V)=W.$$

Podemos ahora enunciar un teorema que establece la condición necesaria y suficiente para la existencia de la transformación inversa:

Teorema VI. 12.

Sea la transformación $T: V \to W$. Existe la inversa de T y es única si y sólo si T es biyectiva.

Demostración.

Sea $T: V \to W$ una transformación biyectiva:

Como T es sobre, $\forall\ \bar{w} \in W$ existe al menos un vector $\bar{v} \in V$

tal que $T(\bar{v})=\bar{w}$. Como T es uno a uno, dicho vector es único.

Existe entonces una única transformación $G: W \to V$ tal que:

$G(\bar{w})=\bar{v} \qquad \forall\ \bar{w} \in W$

Es claro que G es también biyectiva y que

$T(G(\bar{w}))=T(\bar{v})=\bar{w}$

y

$G(T(\bar{v}))=G(\bar{w})=\bar{v}$

por lo que $G=T^{-1}$.

Esto completa la demostración.

<u>Teorema</u> VI. 13.

Sea $T: V \to W$ una transformación lineal.

Su inversa T^{-1}, si existe, también es lineal.

Demostración.

Supongamos que existe $T^{-1}: W \to V$. Sean $\bar{w}_1$ y $\bar{w}_2$ dos vectores de W y sea α un escalar.

Necesitamos demostrar que

$T^{-1}(\alpha\bar{w}_1+\bar{w}_2)=\alpha T^{-1}(\bar{w}_1)+T^{-1}(\bar{w}_2)$

para lo cual, consideremos dos vectores $\bar{v}_1, \bar{v}_2 \in V$ tales que:

$\bar{v}_1=T^{-1}(\bar{w}_1) \quad$ y $\quad \bar{v}_2=T^{-1}(\bar{w}_2) \qquad \ldots\ldots (1)$

Por definición de T^{-1}:

$T(\bar{v}_1)=\bar{w}_1 \quad$ y $T(\bar{v}_2)=\bar{w}_2 \qquad \ldots\ldots (2)$

como T es lineal:

$T(\alpha\bar{v}_1+\bar{v}_2)=\alpha T(\bar{v}_1)+T(\bar{v}_2)$

de (2) se sigue que

$T(\alpha\bar{v}_1+\bar{v}_2)=\alpha\bar{w}_1+\bar{w}_2$

por definición de T^{-1}:

$\alpha\bar{v}_1+\bar{v}_2=T^{-1}(\alpha\bar{w}_1+\bar{w}_2)$

finalmente, de (1):

$\alpha T^{-1}(\bar{w}_1)+T^{-1}(\bar{w}_2)=T^{-1}(\alpha\bar{w}_1+\bar{w}_2)$

como queríamos demostrar.

Regresemos a las transformaciones G y T que estamos analizando.

Vimos que $G: R^2 \to R^2$ no es uno a uno. El lector puede verificar que el núcleo de G es:

$N(G)=\{(x,0) \mid x \in R\}$.

En cambio, el núcleo de T es:

$N(T)=\{(0,0)\}=\{\bar{0}_V\}$

y, como vimos, T es uno a uno.

Notemos que el núcleo de G contiene otros vectores además del vector cero (por ejemplo, (1,0), (6,0) etc.), mientras que el núcleo de T contiene únicamente al vector cero.

<u>Teorema</u> VI. 14

Una transformación lineal $T: V \to W$.

es uno a uno si y sólo si el núcleo contiene únicamente al vector cero, o sea:

$N(T) = \{\bar{0}_V\}$

Demostración.

1o. Si T es uno a uno, entonces $N(T)=\{\bar{0}_V\}$.

En efecto, sabemos que, por ser T lineal:

$T(\bar{0}_V)=\bar{0}_W$

Supongamos que $\bar{v}$ está en el núcleo, es decir

$T(\bar{v})=\bar{0}_W$

Por hipótesis, T es uno a uno, o sea que

$$T(\bar{v})=T(\bar{0}_V) \Rightarrow \bar{v}=\bar{0}_V$$

$\therefore \quad N(T)=\{\bar{0}_V\}$

2o. Si $N(T)=\{\bar{0}_V\}$, entonces T es uno a uno.

En efecto, supongamos que existen dos vectores $\bar{v}_1$, $\bar{v}_2$ de V tales que $T(\bar{v}_1)=T(\bar{v}_2)$.

Entonces $T(\bar{v}_1)-T(\bar{v}_2)=\bar{0}_W$

como T es lineal:

$T(\bar{v}_1-\bar{v}_2) = \bar{0}_W$

Es decir, $\bar{v}_1-\bar{v}_2$ está en el núcleo y, por hipótesis,

$\bar{v}_1-\bar{v}_2=\bar{0}_V$

o bien $\bar{v}_1=\bar{v}_2$

Por tanto: $T(\bar{v}_1)=T(\bar{v}_2) \Rightarrow \bar{v}_1=\bar{v}_2 \quad \forall\, \bar{v}_1, \bar{v}_2 \in V$

y T es uno a uno.

Queda demostrado el teorema.

Recordemos que en una transformación lineal $T: V \to W$:

$\dim V=\dim T(V)+\dim N(T)$.

Ahora bien, si $N(T)=\{\bar{0}_V\}$, entonces $\dim V=\dim T(v)$.

De este razonamiento y de los teoremas VI. 12 y VI. 14, se obtiene el siguiente importante resultado:

Corolario.

Sea un espacio V de dimensión finita y sea $T: V \to W$ una transformación lineal

T tiene inversa si y sólo si

1o. $\dim V= \dim W$

2o. $N(T) = \{\bar{0}_V\}$

Obsérvese que estas condiciones son equivalentes a afirmar que T TIENE INVERSA SI Y SOLO SI LA MATRIZ ASOCIADA A T ES CUADRADA Y NO SINGULAR.

Ahora ya podemos decidir en cualquier caso si una transformación dada tiene inversa. Nos falta únicamente encontrar una forma de calcular dicha inversa, cuando exista.

Para ello, regresemos a la transformación T que gira un vector del plano un ángulo de 60°, cuya matriz asociada es:

$$M(T)=\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

La transformación inversa de T, como se encontró anteriormente, tiene como matriz asociada:

$$M(T^{-1})=\begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

Es fácil ver que:

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

Es decir, $M(T^{-1})$ ES LA MATRIZ INVERSA DE $M(T)$:

$M(T^{-1}) = (M(T))^{-1}$

De hecho, en cualquier transformación lineal donde el dominio sea de dimensión finita, para encontrar la matriz asociada a la transformación inversa bastará con invertir la matriz asociada a la transformación, como establece el siguiente:

Teorema VI. 15.

Sea la transformación lineal $T: V \to W$ y sea $T^{-1}: W \to V$ su inversa.

Sean A y B las respectivas bases de V y W.

Si $M_B^A(T)$ es la matriz asociada a T y $M_A^B(T^{-1})$ es la matriz asociada a T^{-1}, entonces:

$$M_A^B(T^{-1}) = (M_B^A(T))^{-1}$$

Se deja al estudiante la demostración de este teorema.

Ejemplo VI. 17.

Sea la transformación lineal $T: R^3 \to R^3$ dada por la regla $T(x, y, z) = (3x+y,\ x+y+z,\ 2y+4z)$.

Investiguemos si tiene inversa:

Para que un vector esté en el núcleo

$(3x+y,\ x+y+z,\ 2y+4z) = (0,0,0)$

Es decir,
$$\left.\begin{aligned} 3x+y &= 0 \\ x+y+z &= 0 \\ 2y+4z &= 0 \end{aligned}\right\} \quad \ldots\ldots (1)$$

Puesto que el determinante $\begin{vmatrix} 3 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 4 \end{vmatrix} = 2$ es diferente de cero, el sistema (1) tiene por única solución:

$$x = 0$$
$$y = 0$$
$$z = 0$$

$\therefore\ N(T) = \{(0,0,0)\}$

Además, por ser la transformación de R^3 en R^3, podemos concluir que existe la inversa T^{-1}.

Es fácil verificar que la matriz asociada a T es:

$$M(T) = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 4 \end{bmatrix}$$

Note que en la obtención del núcleo se investigó implícitamente si M(T) es singular o no lo es, es decir, si se puede invertir, ya que $\begin{vmatrix} 3 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 4 \end{vmatrix}$ es precisamente el determinante de M(T).

La inversa de la matriz M(T) es

$$(M(T))^{-1} = \begin{bmatrix} 1 & -2 & \frac{1}{2} \\ -2 & 6 & -\frac{3}{2} \\ 1 & -3 & 1 \end{bmatrix} = M(T^{-1})$$

Podemos por último encontrar la regla de transformación para T^{-1}:

$$T^{-1}(x,y,z) = \begin{bmatrix} 1 & -2 & \frac{1}{2} \\ -2 & 6 & -\frac{3}{2} \\ 1 & -3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-2y+\frac{1}{2}z \\ -2x+6y-\frac{3}{2}z \\ x-3y+z \end{bmatrix}$$

$\therefore\ T^{-1}(x,y,z) = (x-2y+\frac{1}{2}z,\ -2x+6y-\frac{3}{2}z,\ x-3y+z).$

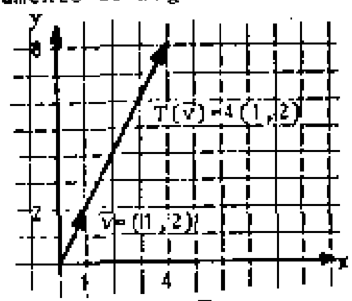
- -

VI. 5. VALORES Y VECTORES CARACTERISTICOS.

Consideremos la transformación lineal $T: R^2 \to R^2$ definida por: $T(x,y)=(2x+y,\ 6x+y)$ y encontremos la imagen del vector (1,2):

$T(1,2)=(4,8)=4(1,2)$

Si representamos al vector (1,2) como un segmento dirigido en plano x y, podemos ver que la transformación NO le cambió de dirección, sino únicamente de magnitud:



$T(\overline{v})=4\overline{v}$

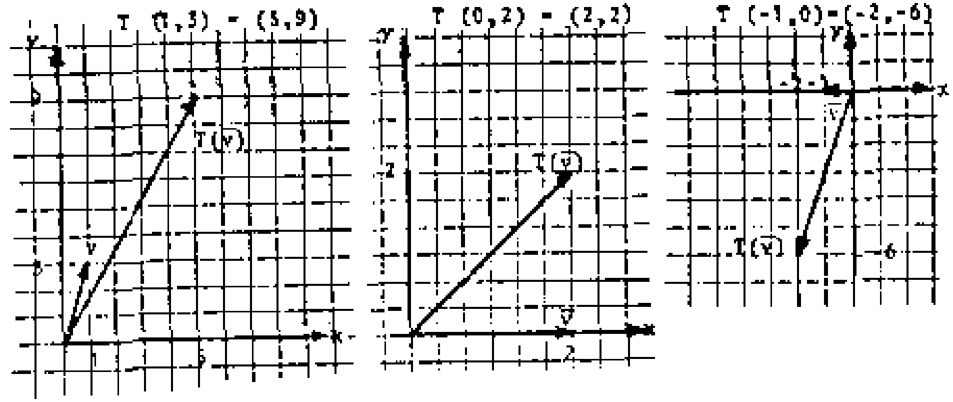
Si ahora tomamos el vector (3,6):

$T(3,6)=(12,24)=4(3,6)$

vemos que le ocurre exactamente lo mismo que al anterior (su magnitud aumenta cuatro veces pero no cambia de dirección al aplicarle la regla de transformación).

Otro caso similar: $T(-2,-4)=4(-2,-4)$

Es fácil ver, sin embargo, que los vectores a los que esto ocurre son casos muy particulares, y que en general la transformación modificará tanto la magnitud como la dirección del vector sobre el cual actúa. Por ejemplo:



Cabe entonces preguntarse:

1.- ¿Cómo deben ser los vectores para que la transformación no les cambie de dirección?

2.- Los vectores que según T no son modificados en dirección, ¿se multiplican siempre por el mismo escalar (4 en este caso) o existen algunos otros valores?

Para responder a lo anterior, planteemos el problema en una forma más general, es decir, veamos en que casos la imagen de un vector es simplemente el producto de un escalar por dicho vector, esto es

$T(\overline{v})=\lambda\overline{v}$ . . . . (1)

para algún escalar λ y algún vector $\overline{v}$ de R^2.

Sea $\overline{v}=(x,y)$ un vector cualquiera de R^2.

Requerimos, de (1), que

$T(x,y)=\lambda(x,y)$

Aplicando la regla de transformación:

$(2x+y,\ 6x+y)=(\lambda x,\ \lambda y)$

de donde

$$2x+y=\lambda x$$
$$6x+y=\lambda y$$

o bien: $\left.\begin{array}{l}(2-\lambda)x+y=0\\ 6x+(1-\lambda)y=0\end{array}\right\}$ . . . . (2)

Como se vió en el capítulo IV, las soluciones del sistema homogéneo (2) dependen del valor del determinante de la matriz de coeficientes:

$$\begin{vmatrix}2-\lambda & 1\\ 6 & 1-\lambda\end{vmatrix}=(2-\lambda)(1-\lambda)-6=\lambda^2-3\lambda-4$$

Si este determinante, llamado polinomio característico, es diferente de cero, tendremos la solución trivial

$$\bar{v}=(x,y)=(0,0)$$

que no nos interesa, pues es obvio que

$$T(\bar{0}_v)=\lambda\bar{0}_v \quad \text{para todo escalar } \lambda.$$

Por tanto, para obtener vectores no nulos, requerimos que el determinante sea cero, es decir:

$$\lambda^2-3\lambda-4=0 \qquad . . . . (3)$$

que se conoce como ecuación característica, y de ella se obtienen dos valores de λ, llamados valores característicos: $\lambda_1=-1$
$\lambda_2=4$

Con estos valores podemos regresar a las ecuaciones (2) y obtener soluciones diferentes de la trivial:

Para $\lambda_1=-1$, de (2): $\begin{array}{l}3x+y=0\\ 6x+2y=0\end{array}$

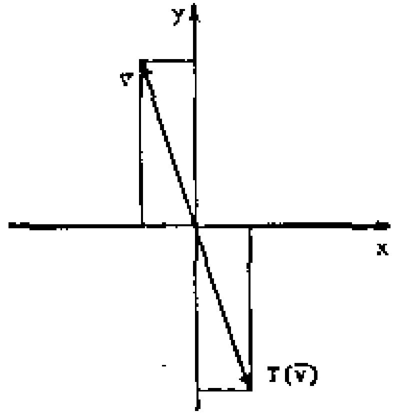
de donde $y=-3x$

Es decir, todos aquellos vectores de la forma $(x,-3x)$ no sufrirán cambio alguno en su dirección al aplicarles la regla de transformación y únicamente se multiplicarán por -1:

$$T(x,-3x)=-(x,-3x)$$

A $(x,-3x)$, con $x\neq0$, se le conoce como vector característico asociado al valor característico -1.

Gráficamente:



En forma similar, llevando a (2) el otro valor característico $\lambda_2=4$, se tiene

$$\begin{array}{l}-2x+y=0\\ 6x-3y=0\end{array}$$

cuya solución es $y=2x$.

Entonces, $(x,2x)$, con $x\neq0$, es el vector característico asociado al valor característico 4, es decir:

$$T(x,2x)=4(x,2x)$$

Observe que los vectores considerados al principio de esta sección son justamente de la forma $(x,2x)$; de ahí que al transformarlos su magnitud se multiplicara por 4 y no cambiaran su dirección.

Podemos asegurar, por otra parte, que los valores obtenidos (-1 y 4) son los únicos que nos permiten hallar vectores no nulos tales que

$$T(\bar{v})=\lambda\bar{v}, \qquad . . . . (1)$$

ya que cualquier otro valor de λ nos haría distinto de cero el determinante, quedando la trivial ($\bar{v}=\bar{0}_V$) como única solución de (1).

Daremos ahora la definición formal de estos conceptos.

Definición.

Sea V un espacio vectorial sobre un campo K y sea $T:V\to V$ una transformación lineal. Entonces, si

$$T(\bar{v})=\lambda\bar{v}$$

para algún vector no nulo $\bar{v}\in V$ y algún escalar $\lambda\in K$, λ es llamado valor característico de T y $\bar{v}$ es llamado vector característico asociado a λ.[1]

Obsérvese que si $\bar{v}$ es un vector característico asociado al valor λ y $k\in K$ es un escalar no nulo, entonces $k\bar{v}$ es también un vector característico asociado al valor λ.

En efecto, si $T(\bar{v})=\lambda\bar{v}$

entonces, como T es lineal: $T(k\bar{v}) = kT(\bar{v})$

$T(k\bar{v}) = k(\lambda\bar{v})$

Es decir: $T(k\bar{v}) = \lambda(k\bar{v})$ y, por definición, $k\bar{v}$ es también un vector característico.

Volvamos al problema de la transformación lineal $T:R^2\to R^2$ definida por $T(x,y)=(2x+y,\ 6x+y)$.

Es claro que $T(1,0)=(2,6)$

$T(0,1)=(1,1)$

por lo que la matriz asociada a la transformación en la base canónica es:

(1) Existen varios sinónimos de valor y vector característico: eigenvalor y eigenvector; valor propio y vector propio; autovalor y autovector; etc.

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

Fácilmente se comprueba que la matriz de coeficientes del sistema (2) (que nos permitió obtener los valores y vectores característicos de T) es $A-\lambda I$, siendo I la matriz identidad:

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

En general, si T es una transformación lineal de un espacio de dimensión finita en sí mismo, el problema de encontrar sus valores y vectores característicos se puede resolver por medio de la matriz asociada a T.

En efecto:

Sea V un espacio vectorial de dimensión finita sobre un campo K, y sea $T:V\to V$ una transformación lineal cuya matriz asociada es A.

Un vector no nulo $\bar{v}\in V$ es un vector característico de T (se dice también que es un vector característico de la matriz A) si:

$A\bar{v}=\lambda\bar{v}$ para algún $\lambda\in K$.

Es decir:

$A\bar{v}-\lambda\bar{v}=\bar{0}$

$A\bar{v}-\lambda I\bar{v}=\bar{0}$

$(A-\lambda I)\bar{v}=\bar{0}$

Al determinante $\det(A-\lambda I)$ se le conoce como polinomio característico, y al igualarlo a cero (para así obtener vectores no nulos) se obtiene la ecuación característica:

$$\det(A-\lambda I)=0$$

cuyas soluciones en K son los valores característicos de la transformación T (o de la matriz A).

Si λ_j es un valor característico, la solución general del sistema homogéneo

$$(A-\lambda_j I)\bar{v}=\bar{0}$$

nos da los vectores característicos asociados al valor característico λ_j.

Ejemplo VI. 18.

a) Ya hemos estudiado la transformación lineal $T:R^2 \to R^2$ que gira un vector un ángulo de 60°, dada por la regla de correspondencia

$$T(x,y)=(\tfrac{1}{2}x-\tfrac{\sqrt{3}}{2}y,\ \tfrac{\sqrt{3}}{2}x+\tfrac{1}{2}y) \quad \ldots\ldots (1)$$

y cuya matriz asociada es

$$A=\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \quad \ldots\ldots (2)$$

¿Tendrá esta transformación vectores característicos? Es claro que no, ya que ningún vector de R^2 diferente del cero conservará su misma dirección después de aplicarle la transformación, pues el efecto de esta transformación es precisamente cambiar la dirección del vector al girarlo 60°.

Empleando el proceso descrito anteriormente, se tendría:

$$A-\lambda I=\begin{bmatrix} \frac{1}{2}-\lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2}-\lambda \end{bmatrix}$$

El polinomio característico es

$$\det(A-\lambda I)=\begin{vmatrix} \frac{1}{2}-\lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2}-\lambda \end{vmatrix}=(\tfrac{1}{2}-\lambda)(\tfrac{1}{2}-\lambda)+\tfrac{\sqrt{3}}{2}\ \tfrac{\sqrt{3}}{2}$$

$$\det(A-\lambda I)=\lambda^2-\lambda+1$$

cuyas raíces son:
$$\left.\begin{array}{l} \lambda_1=\frac{1}{2}+\frac{\sqrt{3}}{2}i \\ \lambda_2=\frac{1}{2}-\frac{\sqrt{3}}{2}i \end{array}\right\} \quad \ldots\ldots (3)$$

por lo que no existen escalares $\lambda \in R$ tales que

$$T(\bar{v})=\lambda\bar{v},\quad \bar{v}\in R^2$$

Es decir, T no tiene valores característicos ni, en consecuencia, vectores característicos, como habíamos concluído anteriormente.

b) Consideremos ahora el problema de determinar los valores característicos de una transformación $T:C^2 \to C^2$ (1) dada por la misma regla:

$$T(x,y)=(\tfrac{1}{2}x-\tfrac{\sqrt{3}}{2}y,\tfrac{\sqrt{3}}{2}x+\tfrac{1}{2}y) \quad \ldots\ldots (1)$$

donde $x,\ y \in C$.

Si elegimos la base $B=\{(1,0),(0,1)\}$ de C^2, la matriz A asociada a T es la misma, y:

$$A-\lambda I=\begin{bmatrix} \frac{1}{2}-\lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2}-\lambda \end{bmatrix} \quad \ldots\ldots (2)$$

Entonces, los valores característicos de T son las raíces ya obtenidas del polinomio característico $\det(A-\lambda I)$:

$$\left.\begin{array}{l} \lambda_1=\frac{1}{2}+\frac{\sqrt{3}}{2}i \\ \lambda_2=\frac{1}{2}-\frac{\sqrt{3}}{2}i \end{array}\right\} \quad \ldots\ldots (3)$$

---

(1) C^2 es el espacio vectorial de pares ordenados de números complejos sobre el campo C.

por lo que existen escalares $\lambda \in C$ tales que

$$T(\bar{v})=\lambda\bar{v}, \quad \bar{v} \in C^2$$

Dado que existen valores característicos, podemos obtener sus correspondientes vectores característicos:

Llevando $\lambda_1 = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ a (2), el sistema $(A-\lambda I)\bar{v}=\bar{0}$ queda:

$$\begin{bmatrix} \frac{1}{2} - \frac{1}{2} - \frac{\sqrt{3}}{2}i & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} - \frac{1}{2} - \frac{\sqrt{3}}{2}i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{donde } x, y \in C$$

$$\begin{bmatrix} -\frac{\sqrt{3}}{2}i & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2}i \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Resolviendo se obtiene que: $y = -ix$, $\forall x \in C$,

$\therefore$ $(x, -ix)$ es un vector característico asociado al valor $\lambda_1 = \frac{1}{2} + \frac{\sqrt{3}}{2}i$, para cualquier número complejo $x \neq 0$.

En forma análoga se obtiene que (x, ix) es un vector característico asociado a $\lambda_2 = \frac{1}{2} - \frac{\sqrt{3}}{2}i$ para cualquier número complejo $x \neq 0$.

Este ejemplo pone de manifiesto que la existencia de valores característicos depende del campo sobre el cual se trabaja.

- -

Ejemplo VI. 19.

Se tiene una transformación lineal $T: R^3 \to R^3$, cuya matriz asociada es:

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

Encontremos sus valores característicos:

$$A-\lambda I = \begin{vmatrix} 2-\lambda & 3 & 3 \\ 0 & 4-\lambda & 4 \\ 0 & -5 & -5-\lambda \end{vmatrix} = \lambda\,(2-\lambda)(1+\lambda)$$

$$\therefore \quad \left.\begin{array}{l} \lambda_1 = 0 \\ \lambda_2 = 2 \\ \lambda_3 = -1 \end{array}\right\} \text{ Valores característicos.}$$

Obtengamos los vectores asociados, utilizando $(A-\lambda I)\bar{v}=\bar{0}$:

Para $\lambda_1 = 0$

$$\begin{bmatrix} 2 & 3 & 3 \\ 0 & 4 & 4 \\ 0 & -5 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{l} x=0 \\ y=k_1 \\ z=-k_1 \end{array}$$

$\therefore$ $(0, k_1, -k_1)$ es un vector asociado a $\lambda_1 = 0$, $\forall k_1 \neq 0$.

Para $\lambda_2 = 2$

$$\begin{bmatrix} 0 & 3 & 3 \\ 0 & 2 & 4 \\ 0 & -5 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{l} x=k_2 \\ y=0 \\ z=0 \end{array}$$

$\therefore$ $(k_2, 0, 0)$ es un vector asociado a $\lambda_2 = 2$, $\forall k_2 \neq 0$.

Para $\lambda_3 = -1$

$$\begin{bmatrix} 3 & 3 & 3 \\ 0 & 5 & 4 \\ 0 & -5 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{l} x=-k_3 \\ y=-4k_3 \\ z=5k_3 \end{array}$$

$\therefore$ $(-k_3, -4k_3, 5k_3)$ es un vector asociado a $\lambda_3 = -1$, $\forall k_3 \neq 0$.

- -

Observemos que, si bien la definición de valor y vector característico EXCLUYE AL VECTOR CERO, permite al escalar cero ser eventualmente un valor característico, como ocurrió en este último ejemplo. De hecho, en cualquier transformación lineal $T:V \to V$, con núcleo distinto del $\{\bar{0}_V\}$, todos los vectores no nulos del núcleo pertenecen al valor característico cero, pues $T(\bar{v})=\bar{0}_V=0(\bar{v})$ si $\bar{v}\in N(T)$.

Se deja al estudiante encontrar el núcleo de la transformación del ejemplo anterior.

Ejemplo VI. 20.

Sea F el espacio vectorial de todas las funciones $f: R\to R$ que admiten derivadas de cualquier orden en un intervalo [a,b], y consideremos el problema de obtener los valores y vectores característicos del operador derivada

$$D: F\to F$$

Por ser F un espacio de dimensión infinita, no existe una representación matricial para el operador $D: F\to F$, y no podemos proceder como en los ejemplos anteriores. Necesitamos entonces recurrir a la definición de valor y vector característico, según la cual:

Una función f, diferente de la función cero, es un vector característico del operador D si

$$Df(x)=\lambda f(x) \quad \text{para algún } \lambda\in R. \quad \ldots\ldots (1)$$

o bien

$$\frac{df(x)}{dx}=\lambda f(x)$$

lo cual implica que

$$\frac{df(x)}{f(x)}=\lambda dx$$

Integrando ambos miembros de esta expresión, podemos obtener la función f que buscamos:

$$\int\frac{df(x)}{f(x)}=\int\lambda dx$$

$$L(f(x))=\lambda x+c_1$$

$$f(x)=e^{\lambda x+c_1}=e^{\lambda x}\,e^{c_1}$$

finalmente, haciendo $e^{c_1}=c$:

$$f(x)=ce^{\lambda x} \quad \forall\, c,\lambda\in R,\ c\neq 0 \quad \ldots\ldots (2)$$

Derivando (2) se comprueba (1), ya que:

$$D(ce^{\lambda x})=\lambda(ce^{\lambda x}) \quad \forall\, c,\lambda\in R,\ c\neq 0.$$

En conclusión, $ce^{\lambda x}$ es un vector característico asociado al valor característico λ, $\forall\,\lambda\in R$, por lo que todos los números reales son valores característicos del operador $D: F\to F$

- -

<u>Vectores característicos linealmente independientes.</u>

Volvamos al ejemplo VI. 19. Teníamos que $(0,k_1,-k_1)$, $(k_2,0,0)$ y $(-k_3,-4k_3,5k_3)$ son los vectores característicos asociados a los valores característicos 0, 2 y -1, respectivamente.

Fácilmente se comprueba que estos vectores son linealmente independientes. En efecto:

$$\begin{vmatrix} 0 & k_2 & -k_3 \\ k_1 & 0 & -4k_3 \\ -k_1 & 0 & 5k_3 \end{vmatrix} = -k_1k_2k_3\neq 0 \quad (\text{pues } k_i\neq 0)$$

para el caso general tendremos que:

<u>Teorema</u> VI. 16

Vectores característicos asociados a valores característicos distintos, son linealmente independientes.

Demostración. (Por inducción matemática).

Sea $T: V\to V$ una transformación lineal con n valores característicos diferentes $\lambda_1, \lambda_2, \ldots, \lambda_n$, y sean $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$

los respectivos vectores característicos asociados.

Demostraremos que

$\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_n\bar{v}_n=\bar{0}_V \implies \alpha_i=0, \quad \forall i$ . . . . (1)

1o. Para $n=1$, $\alpha_1\bar{v}_1=\bar{0}_V \implies \alpha_1=0$, ya que $\bar{v}_1\neq\bar{0}_V$

2o. Supongamos válida para $n=k$, es decir

$\alpha_1\bar{v}_1+\alpha_2\bar{v}_2+\ldots+\alpha_k\bar{v}_k=\bar{0}_V \implies \alpha_i=0, \quad \forall i$ . . . . (2)

y verifiquemos para $n=k+1$:

$\beta_1\bar{v}_1+\beta_2\bar{v}_2+\ldots+\beta_k\bar{v}_k+\beta_{k+1}\bar{v}_{k+1}=\bar{0}_V$ . . . . (3)

Apliquemos la transformación lineal T en ambos miembros de (3) tomando en cuenta que

$T(\bar{v}_i)=\lambda_i\bar{v}_i$

Se tiene entonces que:

$\beta_1\lambda_1\bar{v}_1+\beta_2\lambda_2\bar{v}_2+\ldots+\beta_k\lambda_k\bar{v}_k+\beta_{k+1}\lambda_{k+1}\bar{v}_{k+1}=\bar{0}_V$ . . . . (3')

Multiplicando (3) por λ_{k+1}

$\beta_1\lambda_{k+1}\bar{v}_1+\beta_2\lambda_{k+1}\bar{v}_2+\ldots+\beta_k\lambda_{k+1}\bar{v}_k+\beta_{k+1}\lambda_{k+1}\bar{v}_{k+1}=\bar{0}_V$. . . . (3"

Restando (3") de (3'):

$\beta_1(\lambda_1-\lambda_{k+1})\bar{v}_1+\beta_2(\lambda_2-\lambda_{k+1})\bar{v}_2+\ldots+\beta_k(\lambda_k-\lambda_{k+1})\bar{v}_k=\bar{0}_V$

Pero por hipótesis (ecuación (2)), los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$, son linealmente independientes, es decir:

$$\left.\begin{aligned} \beta_1(\lambda_1-\lambda_{k+1}) &= 0 \\ \beta_2(\lambda_2-\lambda_{k+1}) &= 0 \\ &\vdots \\ \beta_k(\lambda_k-\lambda_{k+1}) &= 0 \end{aligned}\right\} \quad \ldots\ldots (4)$$

Además, todos los valores característicos son distintos, esto es, $\lambda_i-\lambda_{k+1}\neq 0$, $\forall i=1, 2, \ldots, k$. En consecuencia, de (4), $\beta_1=\beta_2=\ldots=\beta_k=0$.

Llevando estos valores a (3),

$\beta_{k+1}\,\bar{v}_{k+1}=\bar{0}_V$

Como $\bar{v}_{k+1}\neq\bar{0}_V$, entonces $\beta_{k+1}=0$

Por tanto, los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k, \bar{v}_{k+1}$ son linealmente independientes.

Con esto queda demostrado el teorema.

Es conveniente señalar que el recíproco del teorema VI.16 no es cierto. Esto es, si T tiene vectores característicos linealmente independientes $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$, entonces los correspondientes valores característicos $\lambda_1, \lambda_2, \ldots, \lambda_n$ no son necesariamente distintos.

Por ejemplo, para la transformación identidad $I:R^2 \to R^2$ definida por

$I(\bar{v})=\bar{v} \quad \forall\, \bar{v} \in R^2$

Tenemos:

$I(1,0)=(1,0)$

$I(0,1)=(0,1)$

Obviamente, (1,0) y (0,1) son vectores característicos linealmente independientes, ambos asociados al mismo valor característico $\lambda=1$.

- -

Espacios característicos.

Recordemos que para la transformación T del ejemplo VI.19 se obtuvieron los valores característicos: $\lambda_1=0$, $\lambda_2=2$, $\lambda_3=-1$.

El conjunto de todos los vectores característicos asociados al valor $\lambda_3=-1$ es

$\{(-k_3, -4k_3, 5k_3) \mid k_3 \in R \text{ y } k_3 \neq 0\}$

Si a este conjunto agregamos el vector cero, obtenemos el subespacio vectorial de R^3:

$$E(\lambda_3)=\{(-k_3, -4k_3, 5k_3) \mid k_3 \in R\}$$

llamado espacio característico asociado al valor característico λ_3.

En forma similar se pueden obtener los espacios característicos:

$$E(\lambda_1)=\{(0, k_1, -k_1) \mid k_1 \in R\}$$

y

$$E(\lambda_2)=\{(k_2, 0, 0) \mid k_2 \in R\}$$

Teorema VI. 17.

Sea $T:V \rightarrow V$ una transformación lineal y sea λ un valor característico de T. El conjunto

$$E(\lambda)=\{\bar{v} \mid \bar{v} \in V \ \text{ y } T(\bar{v})=\lambda\bar{v}\}$$

es un subespacio vectorial de V.

Demostración:

1) Sean $\bar{v}_1$, $\bar{v}_2 \in E(\lambda)$. Entonces

$$T(\bar{v}_1)=\lambda\bar{v}_1$$

y

$$T(\bar{v}_2)=\lambda\bar{v}_2$$

como T es lineal

$$T(\bar{v}_1+\bar{v}_2)=T(\bar{v}_1)+T(\bar{v}_2)=\lambda\bar{v}_1+\lambda\bar{v}_2=\lambda(\bar{v}_1+\bar{v}_2)$$

$\therefore$ $\bar{v}_1+\bar{v}_2 \in E(\lambda)$.

2) Sean $\bar{v}_1 \in E(\lambda)$ y $\alpha \in K$. Entonces

$$T(\bar{v}_1)=\lambda\bar{v}_1$$

y como T es lineal

$$T(\alpha\bar{v}_1)=\alpha T(\bar{v}_1)=\alpha(\lambda\bar{v}_1)=\lambda(\alpha\bar{v}_1)$$

$\therefore$ $\alpha\bar{v}_1 \in E(\lambda)$.

Esto demuestra el teorema.

Definición.

A $E(\lambda)$ se le conoce como espacio característico asociado al valor característico λ.

Ejemplo VI. 21.

a) Sea la matriz $A=\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$

Esta matriz define una transformación lineal $T: R^3 \rightarrow R^3$ dada por $T(\bar{v})=A\bar{v}$, $\forall\ \bar{v} \in R^3$

Encontremos los espacios característicos.

Para obtener los valores característicos, calculemos:

$$\det(A-\lambda I)=\begin{vmatrix} 3-\lambda & 1 & -1 \\ 2 & 2-\lambda & -1 \\ 2 & 2 & 0-\lambda \end{vmatrix}=-\lambda^3-5\lambda^2+8\lambda-4$$

Las raíces de este polinomio son: $\lambda_1=1$

$\lambda_2=2$

$\lambda_3=2$

Note que existen valores característicos repetidos. Encontremos los correspondientes vectores:

Para $\lambda_1=1$

$$(A-\lambda_1 I)\bar{v}=\begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

de donde $z = 2x$

$y = 0$

$\therefore$ $(k_1, 0, 2k_1)$ $\forall\ k_1 \neq 0$, es un vector asociado a $\lambda_1=1$ y $E(\lambda_1)=\{(k_1, 0, 2k_1) \mid k_1 \in R\}$.

Este subespacio es generado por el vector $(1,0,2)$ y, como es claro, $\dim E(\lambda_1)=1$.

Para $\lambda_2=\lambda_3=2$

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

de donde $z=2x$

$y=x$

$\therefore$ $(k_2,\ k_2,\ 2k_2)$ $\forall\ k_2 \neq 0$, es un vector asociado a $\lambda_2=\lambda_3=2$ y $E(\lambda_2)=E(\lambda_3)=\{(k_2,\ k_2,\ 2k_2)\ |\ k_2\ \in R\}$.

Este subespacio es generado por el vector $(1,1,2)$ y su dimensión es $\dim E(\lambda_2)=1$.

b) Consideremos otra transformación con valores característicos repetidos.

Sea $T:R^3 \to R^3$ la transformación lineal definida por

$T(\bar{v})=B\bar{v} \qquad \forall\ \bar{v} \in R^3$

donde $B=\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$

Entonces,

$$\det(B-\lambda I) = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 3 & 3 & 4-\lambda \end{vmatrix} = \lambda^3-9\lambda^2+15\lambda-7$$

y $\quad \lambda_1=7$

$\lambda_2=1$

$\lambda_3=1$

Para $\lambda_1=7$

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

de donde $y=2x$

$z=3x$

$\therefore$ $(k_1,\ 2k_1,\ 3k_1)$ $\forall\ k_1 \neq 0$, es un vector asociado a $\lambda_1=7$ y $E(\lambda_1)=\{(k_1,\ 2k_1,\ 3k_1)\ |\ k_1 \in R\}$.

En este caso, $\dim E(\lambda_1)=1$, ya que es generado por el vector $(1,2,3)$.

Para $\lambda_2=\lambda_3=1$,

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

es decir, $x+y+z=0$

o bien $z=-x-y$

$\therefore$ $(k_2,\ k_3,\ -k_2-k_3)$ $\forall k_2,\ k_3$, no ambos cero, es un vector característico asociado a $\lambda_2=\lambda_3=1$, y $E(\lambda_2)=E(\lambda_3)=\{(k_2,\ k_3,\ -k_2-k_3)\ |\ k_2,\ k_3\ \in R\}$.

Ahora la dimensión es dos, pues $E(\lambda_2)$ es generado por los vectores independientes $(1,0,-1)$ y $(0,1,-1)$:

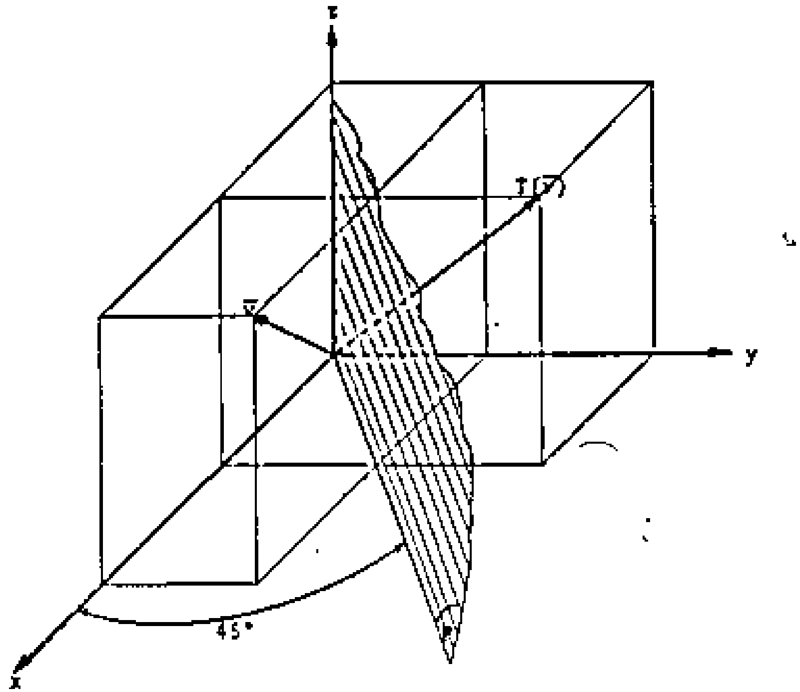
$E(\lambda_2)=E(\lambda_3)=\{k_2(1,0,-1)+k_3(0,1,-1)|\ k_2,\ k_3\ \in R\}$

Tenemos aquí otro caso de vectores característicos linealmente independientes asociados al mismo valor característico 1 - (ver página 79).

- -

Ejemplo VI.22

Sea T aquella que transforma un vector cualquiera del espacio euclidiano tridimensional en su simétrico respecto al plano P que contiene al eje z y forma un ángulo de 45° con el plano x-z, como se muestra en la figura:



La imagen de un vector cualquiera (x,y,z) de R^3 según T es (y,x,z) por lo que la regla de correspondencia es

$T(x,y,z) = (y,x,z)$ . . . . (1)

Como es fácil verificar, se trata de una transformación lineal, por lo que, para obtener la matriz asociada a T en la base canónica de R^3, calculamos

$T(1,0,0) = (0,1,0)$

$T(0,1,0) = (1,0,0)$

$T(0,0,1) = (0,0,1)$

Entonces

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

Obtengamos ahora los espacios característicos de T:

$$\det(M-\lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = \lambda^2(1-\lambda)-(1-\lambda) = -\lambda^3+\lambda^2+\lambda-1$$

haciendo $-\lambda^3+\lambda^2+\lambda-1=0$ se obtienen los valores característicos

$\lambda_1 = -1$

$\lambda_2 = 1$

$\lambda_3 = 1$

Del sistema homogéneo $(M-\lambda I)\bar{v}=\bar{0}$, con $\lambda=-1$:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

cuya solución general es

$x = -y$

$z = 0$

se obtiene el espacio característico

$E(\lambda_1) = \{(k_1, -k_1, 0) \mid k_1 \in R\}$

y del sistema homogéneo $(M-\lambda I)\bar{v}=\bar{0}$, con $\lambda = 1$

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

cuya solución general es

$x=y$

$z=z$

se obtiene el espacio característico

$E(\lambda_2)=\{(k_2, k_2, k_3) \mid k_2, k_3 \in R\}$.

Veamos ahora la interpretación geométrica de estos resultados:

El espacio característico $E(\lambda_1)$, de dimensión uno, representa al conjunto de vectores sobre la recta $y=-x$ alojada en el plano x-y. Estos vectores solamente cambian su sentido al aplicar T, pero conservan su magnitud y dirección. Es decir:

$$T(\bar{v})=-\bar{v}, \quad \forall\ \bar{v} \in E(\lambda_1).$$

El espacio característico $E(\lambda_2)$, de dimensión dos, representa al conjunto de vectores contenidos en el plano P. Estos vectores no sufren alteración alguna al aplicar T. Es decir:

$$T(\bar{v})=\bar{v}, \quad \forall\ \bar{v} \in E(\lambda_2).$$

- -

BIBLIOGRAFIA

1) Tom M. Apostol
Calculus Vols. I y II, 2a. Edición
Editorial Reverté S. A. 1972.

2) Serge Lang
Linear Algebra, 2nd. Edition.
Editorial Addison Wesley, 1972.

3) Kenneth Hoffman / Ray Kunze
Algebra Lineal
Ed. Prentice Hall, 1973.

SISTEMAS DE ECUACIONES LINEALES Y MATRICES

INTRODUCCION

Consideramos el siguiente ejemplo:

Una industria fabrica tres tipos de productos (llamémosles A, B y C) y cuenta con un total de 50 obreros que trabajan 8 hr. diarias. Es decir, dispone de un total de 400 horas-hombre al día.

Para fabricar un producto del tipo A, se requieren 20 horas-hombre ; para uno del tipo B, 100 y para uno del tipo C, 40.

En condiciones normales, no existen restricciones de materia prima ni de maquinaria y los obreros estan capacitados para intervenir en la fabricación de cualquiera de los productos.

Se quiere saber, que cantidad de productos A, B y C pueden fabricarse diariamente empleando la totalidad de horas-hombre disponible.

Podemos entonces plantear el modelo matemático siguiente:

Si x_1, x_2 y x_3 representan el número de productos A, B y C que se fabrican diariamente; entonces $20x_1$, $100x_2$ y $40x_3$ serán, el número de horas-hombre empleadas diariamente en fabricar todos los productos A, B y C.

Como se desean emplear las 400 horas-hombre disponibles, los valores de x_1, x_2 y x_3 deben ser tales que

$$20x_1+100x_2+40x_3=400 \quad . . . . . . (1)$$

Expresiones como ésta, reciben el nombre de ecuaciones lineales.

Una respuesta al problema podría ser fabricar 4 productos del tipo A, 2 del tipo B y 3 del tipo C, ya que, al sustituir estos valores en la expresión (1), se verifica la igualdad. Entonces, diremos que

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

es una solución de la ecuación (1).

Sin embargo, podemos notar que esta solución no es única ya que los valores

$$x_1=7, \quad x_2=1, \quad x_3=4,$$

también satisfacen la ecuación (1), por lo que constituyen otra solución.

Generalizando, la ecuación (1) es una expresión del tipo

$$a_1x_1+a_2x_2+\ldots+a_nx_n=b \quad . . . . . . (2)$$

a la que llamaremos ecuación lineal.

A las constantes $a_1, a_2, \ldots, a_n$ se les llama coeficientes, a b término independiente y a $x_1, x_2, \ldots, x_n$ incógnitas o variables. Vemos que las incógnitas aparecen todas elevadas a la primera potencia, de ahí el nombre de ecuación lineal.

IV.1 SISTEMAS DE ECUACIONES LINEALES.

Regresando al ejemplo anterior, supongamos que por razones de demanda los productos B y C deben fabricarse en cantidades iguales.

Entonces, se tiene la restricción adicional $x_2=x_3$, que expresada en la forma (2) queda

$$0x_1+x_2-x_3=0 \quad \ldots\ldots \quad (3)$$

Ahora el problema consiste en encontrar una solución que satisfaga, simultáneamente, las ecuaciones (1) y (3), por lo que las dos soluciones que antes encontramos no son útiles.

Si se fabrican 6 productos del tipo A y 2 de los tipos B y C, vemos que se satisfacen simultáneamente las ecuaciones (1) y (3). Entonces diremos que

$$x_1=6, \quad x_2=2, \quad x_3=2$$

es una solución del sistema

$$\left.\begin{array}{l} 20x_1+100x_2+40x_3=400 \\ 0x_1+x_2-x_3=0 \end{array}\right\} \quad \ldots\ldots \quad (4)$$

el cual consta de 2 ecuaciones con tres incógnitas.

Definición.

Un sistema de ecuaciones lineales es un conjunto de ecuaciones de la forma (2) que deben resolverse simultáneamente.

En general, un sistema de m ecuaciones con n incógnitas, definido sobre el campo de los números complejos, es de la forma

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2 \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m \end{array}\right\} \quad \ldots\ldots \quad (5)$$

donde $a_{ij} \in C$ son los coeficientes, x_i las incógnitas y $b_i \in C$ los términos independientes.

Definición.

Una solución del sistema de ecuaciones (5) es un conjunto ordenado de n valores que satisface simultáneamente a todas las ecuaciones.

Una forma de obtener soluciones.

Tal vez, la técnica fundamental para encontrar las soluciones de un sistema de ecuaciones lineales, sea la de eliminación. El proceso consiste en transformar el sistema original en un nuevo sistema de ecuaciones que puede resolverse fácilmente. El nuevo sistema, deberá tener las mismas soluciones que el original y en este caso diremos que ambos sistemas son equivalentes.

Para obtener el nuevo sistema se hace uso de las siguientes transformaciones, que reciben el nombre de "Transformaciones elementales" y consisten en:

1) Intercambiar dos ecuaciones.

2) Multiplicar una ecuación por un número $k\neq 0$

3) Multiplicar una ecuación por un número $k\neq 0$ y sumar el resultado a otra ecuación del sistema.

No es difícil aceptar que el sistema original y el nuevo sistema son equivalentes. Sin embargo, el estudiante puede consultar la referencia 1 pag. 414 para una demostración.

Podemos ilustrar la técnica mediante el siguiente ejemplo:

Ejemplo IV.1.

$$2x_2+2x_3=6$$
$$x_1-2x_2+x_3=1$$
$$-x_1+x_2+5x_3=10$$

si intercambiamos las dos primeras ecuaciones (con el objeto de que x_1 aparezca en la primera ecuación) obtenemos

$$x_1-2x_2+x_3=1$$
$$2x_2+2x_3=6$$
$$-x_1+x_2+5x_3=10$$

si sumamos la primera ecuación a la tercera obtenemos

$$x_1-2x_2+x_3=1$$
$$2x_2+2x_3=6$$
$$0-x_2+6x_3=11$$

si multiplicamos la segunda ecuación por $\frac{1}{2}$ y sumamos el resultado a la tercera obtenemos

$$x_1-2x_2+x_3=1$$
$$2x_2+2x_3=6$$
$$0+7x_3=14$$

dividiendo las ecuaciones dos y tres entre 2 y 7 respectivamente:

$$x_1-2x_2+x_3=1$$
$$x_2+x_3=3$$
$$x_3=2$$

De la tercera ecuación inmediatamente vemos que $x_3=2$ y sustituyendo este valor en la segunda ecuación se encuentra que $x_2=1$. Finalmente, al sustituir en la primera ecuación se encuentra que $x_1=1$, por lo que la solución del sistema es

$$x_1=1, \quad x_2=1, \quad x_3=2$$

Ahora bien, hay sistemas de ecuaciones que admiten más de una solución; un ejemplo lo tenemos en el sistema

$$\left.\begin{aligned} 20x_1+100x_2+40x_3&=400 \\ 0x_1+x_2-x_3&=0 \end{aligned}\right\} \ldots\ldots (4)$$

que empleamos como introducción. La solución que habíamos mencionado es

$$x_1=6, \quad x_2=2, \quad x_3=2$$

y el estudiante puede comprobar que

$$x_1=13, \quad x_2=1, \quad x_3=1$$

es otra solución. A este tipo de sistemas, que tienen más de una solución, se les llama indeterminados.

También hay sistemas que no admiten solución. Por ejemplo, resulta evidente que el sistema

$$x_1+x_2=1$$
$$x_1+x_2=3$$

no tiene solución, puesto que no existen dos números cuya suma sea 1 y 3 a la vez. A este tipo de sistemas se les llama incompatibles (o inconsistentes).

En general, de acuerdo con la existencia y tipos de solución, los sistemas de ecuaciones lineales se clasifican en:

$$\begin{cases} \text{Compatibles (tienen solución)} & \begin{cases} \text{Determinados (una solución)} \\ \text{Indeterminados (más de una solución)} \end{cases} \\ \text{Incompatibles (no tienen solución)} & \end{cases}$$

Más adelante, y con ayuda de otros conceptos que trataremos, podremos determinar si un sistema tiene solución o no la tiene, y si ésta es única o no lo es.

IV.2 MATRICES.

El estudio de los sistemas de ecuaciones realizado en la sección precedente, puede servir como una introducción natural al concepto de matriz. En el proceso de construcción de un sistema equivalente, puede advertirse que no es necesario escribir las incógnitas $x_1, x_2, \ldots, x_n$, ya que realmente sólo se opera con los coeficientes a_{ij} y con los términos independientes b_i.

Si analizamos el ejemplo IV.1, vemos que el sistema

$$\begin{cases} 2x_2+2x_3=6 \\ x_1-2x_2+x_3=1 \\ -x_1+x_2+5x_3=10 \end{cases}$$

queda completamente determinado al conocer el valor y la posición de cada uno de los coeficientes y términos independientes. Esta información se puede presentar convenientemente en el siguiente arreglo

$$\begin{bmatrix} 0 & 2 & 2 & 6 \\ 1 & -2 & 1 & 1 \\ -1 & 1 & 5 & 10 \end{bmatrix} \quad \ldots \quad (6)$$

al cual se le llama MATRIZ.

Este arreglo en particular, consta de 12 elementos (números) dispuestos en 3 renglones y 4 columnas, por lo que diremos que la matriz es de orden 3x4.

Definición.

Una matriz de orden mxn sobre el campo de los números complejos, es un arreglo rectangular

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \ldots\ldots (7)$$

con m renglones y n columnas, donde los $a_{ij} \in C$ se llaman sus elementos.

Comunmente, se representa a las matrices con letras mayúsculas y a sus elementos con letras minúsculas. En forma abreviada la matriz (7) puede expresarse como

$A=(a_{ij})$ donde $i=1,2,\ldots,m$ y $j=1,2,\ldots,n$

Los subíndices i, j indican, respectivamente, el renglón y la columna en que se encuentra el elemento a_{ij}. Así por ejemplo, a_{23} representa al elemento que se encuentra en el segundo renglón y tercera columna de la matriz A.

Dada la importancia de la posición que guardan los elementos en el arreglo, decimos que dos matrices son iguales si son del mismo orden y sus elementos correspondientes son iguales. A esto obedece la siguiente

Definición.

Sean $A=(a_{ij})$ y $B=(b_{ij})$ dos matrices del mismo orden, entonces:

$A=B \iff a_{ij}=b_{ij} \quad \forall\ i,j$

Haciendo referencia nuevamente al ejemplo IV. 1, podemos efectuar con los renglones de la matriz (6) transformaciones equivalentes a las efectuadas con las ecuaciones del sistema.

El proceso sería entonces el que se ilustra a continuación

$$M=\begin{bmatrix} 0 & 2 & 2 & 6 \\ 1 & -2 & 1 & 1 \\ -1 & 1 & 5 & 10 \end{bmatrix} \quad \ldots\ldots (6)$$

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

hemos intercambiado los dos primeros renglones.

$$M_{II}=\begin{bmatrix} 1 & -2 & 1 & 1 \\ 0 & 2 & 2 & 6 \\ 0 & -1 & 6 & 11 \end{bmatrix}$$

hemos sumado el primer renglón al tercero.

$$M_{III}=\begin{bmatrix} 1 & -2 & 1 & 1 \\ 0 & 2 & 2 & 6 \\ 0 & 0 & 7 & 14 \end{bmatrix}$$

multiplicando el segundo renglón por $\frac{1}{2}$, lo hemos sumado al tercero

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

hemos dividido el segundo y tercer renglón entre 2 y 7 respectivamente

La última matriz (M_{IV}) representa el sistema equivalente

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

cuya solución puede obtenerse fácilmente.

La matriz M_{IV}, se dice que está en forma escalonada o que es una matriz escalonada. En general, una matriz es escalonada si el primer elemento distinto de cero de cada renglón, es igual a 1 y el número de ceros anteriores a dicho elemento aumenta de renglón a renglón. Las siguientes matrices son escalonadas:

$$M=\begin{bmatrix}1&3&6\\0&1&5\\0&0&1\end{bmatrix}\qquad N=\begin{bmatrix}1&2&3&4\\0&0&1&1\\0&0&0&1\end{bmatrix}\qquad P=\begin{bmatrix}1&2&3&4&1\\0&0&1&5&2\\0&0&0&1&0\\0&0&0&0&0\end{bmatrix}$$

Transformaciones elementales por renglón.

Las transformaciones efectuadas con los renglones de la matriz M, para obtener finalmente la matriz M_{IV}, se llaman "transformaciones elementales por renglón" y como hemos visto, pueden ser de tres tipos:

1) Intercambio de dos renglones.

2) Multiplicación de un renglón por un número $k\neq 0$.

3) Multiplicación de un renglón por un número $k\neq 0$ y suma del resultado a otro renglón de la matriz.

Utilizando las transformaciones elementales por renglón, es posible transformar cualquier matriz en una matriz escalonada.

Definición.

Diremos que dos matrices son equivalentes, si cualquiera de ellas puede obtenerse a partir de la otra efectuando un número finito de transformaciones elementales por renglón.

En el ejemplo anterior tenemos que las matrices M, M_I, M_{II}, M_{III} y M_{IV} son equivalentes.

Ejemplo IV.2

Sea la matriz $A=\begin{bmatrix}2&4&2&8&-2\\4&8&4&16&-4\\3&-3&9&1&1\\1&-1&3&\frac{1}{3}&4\end{bmatrix}$

transformar a la matriz A en una matriz escalonada equivalente utilizando transformaciones elementales por renglón.

Solución:

Dividiendo entre 2 el primer renglón de A obtenemos

$$A_I=\begin{bmatrix}1&2&1&4&-1\\4&8&4&16&-4\\3&-3&9&1&1\\1&-1&3&\frac{1}{3}&4\end{bmatrix}$$

Multiplicando por -4 el primer renglón y sumando al 2o. renglón de A_I obtenemos

$$A_{II}=\begin{bmatrix}1&2&1&4&-1\\0&0&0&0&0\\3&-3&9&1&1\\1&-1&3&\frac{1}{3}&4\end{bmatrix}$$

Multiplicando por -3 el primer renglón y sumando al 3er. renglón de A_{II} obtenemos

$$A_{III}=\begin{bmatrix}1&2&1&4&-1\\0&0&0&0&0\\0&-9&6&-11&4\\1&-1&3&\frac{1}{3}&4\end{bmatrix}$$

Multiplicando por -1 el primer renglón y sumando al 4o. renglón de A_{III} obtenemos

$$A_{IV}=\begin{bmatrix}1&2&1&4&-1\\0&0&0&0&0\\0&-9&6&-11&4\\0&-3&2&\frac{-11}{3}&5\end{bmatrix}$$

Intercambiando el segundo renglón con el cuarto renglón de A_{IV} obtenemos

$$A_V=\begin{bmatrix}1&2&1&4&-1\\0&-3&2&\frac{-11}{3}&5\\0&-9&6&-11&4\\0&0&0&0&0\end{bmatrix}$$

Multiplicando por -3 el segundo renglón y sumando al tercer renglón de A_V obtenemos

$$A_{VI}=\begin{bmatrix} 1 & 2 & 1 & 4 & -1 \\ 0 & -3 & 2 & -\frac{11}{3} & 5 \\ 0 & 0 & 0 & 0 & -11 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Dividiendo entre -3 el segundo renglón de A_{VI} obtenemos

$$A_{VII}=\begin{bmatrix} 1 & 2 & 1 & 4 & -1 \\ 0 & 1 & -\frac{2}{3} & \frac{11}{9} & -\frac{5}{3} \\ 0 & 0 & 0 & 0 & -11 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Dividiendo entre -11 el tercer renglón de A_{VII} obtenemos

$$A_{VIII}=\begin{bmatrix} 1 & 2 & 1 & 4 & -1 \\ 0 & 1 & -\frac{2}{3} & \frac{11}{9} & -\frac{5}{3} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

La matriz A_{VIII} es una matriz escalonada.

- -

Al hablar de las transformaciones elementales, hemos hecho énfasis en el término "por renglón". Esto obedece a que existen transformaciones, análogas a las aquí descritas, efectuadas -- con las columnas de una matriz. En este capítulo no se justificará la existencia de dichas transformaciones y tampoco serán utilizadas. El estudiante interesado en saber más acerca de esto, puede consultar la referencia 3 pag. 143, una vez que haya terminado el capítulo.

Rango de una matriz.

Definición.

Si transformamos una matriz A en una matriz escalonada B, el número de renglones de la matriz B con al menos un elemento distinto de cero se llama RANGO DE LA MATRIZ A y se representa con -- R(A). El mismo rango se asigna a la matriz B.

De acuerdo con esta definición, cuando dos matrices son equivalentes ambas tienen el mismo rango. Las matrices A, $A_I, \ldots, A_{VIII}$ del ejemplo IV.2 son todas de rango 3.

El concepto de rango de una matriz, juega un papel muy - importante en la teoría de sistemas de ecuaciones lineales que veremos más adelante. El estudiante puede darse cuenta que la definición de rango, tal como se enuncia aquí, proporciona a la vez un método para obtenerlo.

IV.3. PRODUCTO DE MATRICES.

Consideremos nuevamente el sistema

$$\left.\begin{array}{l} 20x_1+100x_2+40x_3=400 \\ 0x_1+x_2-x_3=0 \end{array}\right\} \quad \ldots\ldots (4)$$

podemos formar las matrices

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

donde hemos reunido los coeficientes(en A), las incógnitas (en $\bar{x}$) y los términos independientes (en $\bar{b}$) que aparecen en el sistema.:

Con ayuda de estas matrices, podemos expresar el sistema (4) como

$$A\bar{x}=\bar{b} \quad \ldots\ldots (8)$$

siempre y cuando demos una definición adecuada para el producto $A\bar{x}$.

La matriz producto $A\bar{x}$, debe ser de orden 2x1 para poder establecer la igualdad con $\bar{b}$. Además, por igualdad de matrices, los elementos correspondientes de $A\bar{x}$ y $\bar{b}$ deben ser iguales.

Sabemos que la igualdad entre matrices

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

se satisface si y sólo si

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

que son precisamente las condiciones que establece el sistema (4). Por tanto

$$A\bar{x}=\begin{bmatrix} 20x_1+100x_2+40x_3 \\ 0x_1+x_2-x_3 \end{bmatrix} \quad \ldots\ldots (9)$$

A la matriz $A\bar{x}$ expresada en (9) le llamaremos "el producto de las matrices A y $\bar{x}$" (en ese orden). Veamos como puede obtenerse la matriz $A\bar{x}$, a partir de las matrices A y $\bar{x}$:

El primer elemento de $A\bar{x}$, es igual a la suma de los productos de los elementos del primer renglón de A por los elementos de la única columna de $\bar{x}$. En forma esquemática:

$$\begin{bmatrix} 20 & 100 & 40 \end{bmatrix} \qquad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{array}{l} \longrightarrow 20x_1 \\ \quad + \\ \longrightarrow 100x_2 \\ \quad + \\ \longrightarrow 40x_3 \\ \hline 20x_1+100x_2+40x_3 \end{array}$$

El segundo elemento de $A\bar{x}$, es igual a la suma de los productos de los elementos del segundo renglón de A por los elementos de la única columna de $\bar{x}$. En forma esquemática:

$$\begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \qquad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{array}{l} \longrightarrow 0x_1 \\ \quad + \\ \longrightarrow x_2 \\ \quad - \\ \longrightarrow x_3 \\ \hline 0x_1+x_2-x_3 \end{array}$$

En forma similar, obtengamos el producto de las matrices

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

con el objeto de establecer una definición general.

El producto será la matriz

$$AB=\begin{bmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} \end{bmatrix}$$

donde observamos que:

1o.) El elemento que se encuentra en el primer renglón primera columna de la matriz producto, es la suma:

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

2o.) El elemento que se encuentra en el segundo renglón primera columna de la matriz producto, es la suma:

$$a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31}=\sum_{k=1}^{3} a_{2k}b_{k1}$$

Generalizando, el elemento que se encuentra en el renglón i, columna j, de una matriz producto AB, es la suma:

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

donde n es el número de columnas de la matriz A y el número de renglones de la matriz B, que deberá ser el mismo para que pueda efectuarse el producto.

Definición.

Sean: $A=(a_{ij})$ $(i=1, 2, \ldots, m$ y $j=1, 2, \ldots, n)$

y $B=(b_{ij})$ $(i=1, 2, \ldots, n$ y $j=1, 2, \ldots, p)$

dos matrices de orden mxn y nxp respectivamente.

El producto AB es una matriz

$C=(c_{ij})$ $(i=1, 2, \ldots, m$ y $j=1, 2, \ldots, p)$

de orden mxp cuyos elementos están dados por

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

Ejemplo IV.3

Para ilustrar la definición, obtengamos el producto de las matrices

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

que es una matriz C de orden 2x4 tal que

$c_{11}= 2-9+1= -6$

$c_{12}= 0-3+3= 0$

$c_{13}= -4+0-4= -8$

$c_{14}= 2-3-3= -4$

$c_{21}= -1+12-1= 10$

$c_{22}= 0+4-3= 1$

$c_{23}= 2+0+4 = 6$

$c_{24}= -1+4+3= 6$

entonces, la matriz producto es

$$AB=\begin{bmatrix} -6 & 0 & -8 & -4 \\ 10 & 1 & 6 & 6 \end{bmatrix}$$

Obsérvese que el elemento que se encuentra en el renglón i columna j de la matriz producto AB, se obtiene efectuando el producto escalar del renglón i de A por la columna j de B.

Por ejemplo, el elemento c_{13} es el producto

$$\begin{bmatrix} 2 & -3 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix} = -4+0-4= -8=c_{13}$$

- -

Podemos concluir, en base a la definición, que podemos efectuar el producto AB sólo cuando el número de columnas de A es igual al número de renglones de B. En este caso, diremos que las matrices A y B son conformables para la multiplicación.

En general, la multiplicación de las matrices NO es conmutativa. Incluso, en muchas ocasiones se tiene que dos matrices A y B son conformables para multiplicarse en ese orden (es decir puede obtenerse el producto AB), mientras que no son conformables para multiplicarse en el orden contrario (no puede obtenerse el producto BA). Por tal motivo, es necesario precisar el orden en que las matrices se van a multiplicar. Para el caso del producto AB, diremos que A premultiplica a B, o bien que B postmultiplica a A

Ejemplo IV. 4.

Dada las matrices

$$R=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix},\quad S=\begin{bmatrix}0 & -1\\ 3 & -1\end{bmatrix},\quad T=\begin{bmatrix}1 & 0\\ 1 & 1\\ 2 & 3\end{bmatrix}$$

obtener:

a) RS

b) SR

c) TS

d) ST

Solución:

a) $RS=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}0 & -1\\ 3 & -1\end{bmatrix}=\begin{bmatrix}6 & -3\\ 12 & -7\end{bmatrix}$

(Las matrices R y S son de orden 2x2, por lo que les llamaremos matrices cuadradas y diremos simplemente que son de orden 2)

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

En este caso, pudieron obtenerse tanto el producto RS como el producto SR, debido a que ambas matrices son cuadradas y del mismo orden. Sin embargo, notamos que $RS\neq SR$ pues, como se mencionó, el producto de matrices no es una operación conmutativa.

c) $TS=\begin{bmatrix}1 & 0\\ 1 & 1\\ 2 & 3\end{bmatrix}\begin{bmatrix}0 & -1\\ 3 & -1\end{bmatrix}=\begin{bmatrix}0 & -1\\ 3 & -2\\ 9 & -5\end{bmatrix}$

d) $ST=\begin{bmatrix}0 & -1\\ 3 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 1\\ 2 & 3\end{bmatrix}$ no se puede efectuar, ya que el número de columnas de S (2) es diferente del número de renglones de T (3). En otras palabras, S y T no son conformables para la multiplicación.

Obsérvese que, aunque pudo obtenerse el producto TS, no fué posible obtener ST, lo cual resalta la importancia de especificar claramente el orden en que se desea multiplicar dos matrices.

\- -

Definición.

Si una matriz A es de orden nxn, diremos que A es una matriz cuadrada de orden n.

Ejemplo IV. 5.

Para las matrices

$$M=\begin{bmatrix}i\\ -i\\ 1\end{bmatrix}\qquad N=\begin{bmatrix}-i, & 1, & i\end{bmatrix}\qquad P=\begin{bmatrix}i+\alpha & i & -1\\ p_{21} & -i & 1\\ -i & 1+\beta & p_{33}\end{bmatrix}$$

encontrar los valores de α, β , p_{21}, p_{33} de tal forma que se verifique la igualdad

$$MN = P$$

Solución:

Primero obtenemos

$$MN = \begin{bmatrix} 1 & i & -1 \\ -1 & -i & 1 \\ -i & 1 & i \end{bmatrix}$$

Por otra parte, como $MN=P$,

$$\begin{bmatrix} 1 & i & -1 \\ -1 & -i & 1 \\ -i & 1 & i \end{bmatrix} = \begin{bmatrix} i+\alpha & i & -1 \\ p_{21} & -i & 1 \\ -i & 1+\beta & p_{33} \end{bmatrix}$$

$i+\alpha=1$ ∴ $\underline{\alpha=1-i}$

$\underline{p_{21}=-1}$

$1+\beta=1$ ∴ $\underline{\beta=0}$

$\underline{p_{33}=i}$

- -

Teorema IV.1

La multiplicación de matrices (cuando puede efectuarse) es asociativa.

Demostración.

Sean $A=(a_{ij})_{mxn}$, $B=(b_{jk})_{nxp}$ y $C=(c_{kr})_{pxq}$ tres matrices cualesquiera de orden mxn, nxp, pxq respectivamente. Queremos probar que

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

De la definición de producto:

$$AB = \left(\sum_{j=1}^{n} a_{ij}\ b_{jk}\right)_{mxp}$$

donde los índices libres son $i=1,2,...,m$ y $k=1,2,...,p$, y la matriz es de orden mxp. (Los índices libres son los que indican el renglón y la columna del nuevo elemento).

Aplicando ahora la definición de producto a las matrices $(AB)_{mxp}$ y C_{pxq}:

$$(AB)C = \left(\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}\ b_{jk}\right) c_{kr}\right)_{mxq}$$

multiplicando C_{kr} por cada una de las sumas del paréntesis, obtenemos:

$$(AB)C = \left(\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}\ b_{jk}\ c_{kr}\right)\right)_{mxq}$$

donde los índices libres son $i=1,2,...,m$ y $r=1,2,...,q$ y la matriz es de orden mxq.

En forma análoga obtenemos:

$$A(BC) = \left(\sum_{j=1}^{n}\left(\sum_{k=1}^{p} a_{ij}\ b_{jk}\ c_{kr}\right)\right)_{mxq}$$

Dado que el orden de la suma es arbitrario,

$$\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}\ b_{jk}\ c_{kr}\right) = \sum_{j=1}^{n}\left(\sum_{k=1}^{p} a_{ij}\ b_{jk}\ c_{kr}\right)\ \forall\ i,\ r$$

Por lo tanto queda demostrado que

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

Se recomienda al estudiante consultar el apéndice I de la referencia 3 para obtener habilidad en el manejo y comprensión de los símbolos de Σ(suma) y $\Sigma\Sigma$ (doble suma), y hacer la demostración completa de este teorema para el caso particular de las matrices cuadradas de orden dos.

Ejemplo IV. 6.

Verificar la propiedad asociativa para el producto, con las matrices.

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

Solución. Debemos verificar que

$(AB)C=A(BC)$

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

$$(AB)C=\begin{bmatrix}2 & 0 & 0\\ 2 & 1 & 1-i\end{bmatrix}\begin{bmatrix}1 & 0\\ 2 & 1\\ 3 & \frac{1}{2}\end{bmatrix}=\begin{bmatrix}2 & 0\\ 7-3i & \frac{3}{2}-\frac{1}{2}i\end{bmatrix}$$

Por otro lado:

$$(BC)=\begin{bmatrix}-1 & 0 & 0\\ 3 & 1 & 1-i\end{bmatrix}\begin{bmatrix}1 & 0\\ 2 & 1\\ 3 & \frac{1}{2}\end{bmatrix}=\begin{bmatrix}-1 & 0\\ 8-3i & \frac{3}{2}-\frac{1}{2}i\end{bmatrix}$$

$$A(BC)=\begin{bmatrix}-2 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}-1 & 0\\ 8-3i & \frac{3}{2}-\frac{1}{2}i\end{bmatrix}=\begin{bmatrix}2 & 0\\ 7-3i & \frac{3}{2}-\frac{1}{2}i\end{bmatrix}$$

por lo que

$(AB)C=A(BC)$

- -

Matriz Identidad.

Si con las matrices

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

efectuamos el producto IB vemos que

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

Si con la misma matriz I efectuamos el producto IA, donde A es una matriz con 3 renglones y cualquier número de columnas, obtenemos siempre que $IA=A$. A la matriz I le llamamos matriz identidad de orden 3 y la representamos mediante I_3.

En general, a la matriz cuadrada

$$I_n=\begin{bmatrix}1 & 0 & 0 & \dots & 0\\ 0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \dots & 1\end{bmatrix}_{n\times n}$$

le llamaremos matriz identidad de orden n.

Esta matriz puede expresarse en forma abreviada como

$$I_n=(\delta_{ij})\qquad \text{donde}\quad \begin{cases}\delta_{ij}=1 & \text{si}\quad i=j\\ \delta_{ij}=0 & \text{si}\quad i\neq j\end{cases}$$

Teorema IV.2

Para toda matriz A de orden mxn se tiene que:

$I_mA=A$

$A\,I_n=A$

Demostración

Sean $I_m=(\delta_{ij})_{m\times m}$ y $A=(a_{jk})_{m\times n}$

De la definición de producto:

$$I_mA=\left(\sum_{j=1}^{m}\delta_{ij}a_{jk}\right)_{m\times n}$$

donde los índices libres son $i=1,2,\dots,m$ y $k=1,2,\dots,n$.

Como $\delta_{ij}=0\ \forall\ i\neq j$:

$$I_mA=\left(\sum_{j=1}^{m}\delta_{jj}a_{jk}\right)_{m\times n}$$

donde ahora, los índices libres son $j=1,2,\ldots,m$ y $k=1,2,\ldots,n$.

Como $\delta_{jj}=1 \quad \forall j$:

$$I_m A = \left(\sum_{j=1}^{m} a_{jk}\right)_{m\times n}$$

Como j es un índice libre

$$I_m A = (a_{jk})_{m\times n} = A$$

La segunda parte del teorema se demuestra en forma similar.

Matrices elementales.

El resultado de efectuar un número finito de transformaciones elementales por renglón a una matriz A de orden mxn, puede obtenerse también si premultiplicamos A por una cierta matriz cuadrada de orden m.

Para mostrar lo anterior, consideremos por separado cada una de las tres transformaciones elementales por renglón.

1) Intercambio de renglones.

Supongamos que tenemos una matriz A de mxn y que efectuamos el intercambio de sus renglones i y j. Llamemos a la matriz así obtenida matriz B.

Si por otro lado, tomamos la matriz identidad de orden m (I_m) e intercambiamos sus renglones i y j, obtendremos una nueva matriz que llamaremos $I_m^{(i,j)}$.

Si ahora efectuamos el producto $I_m^{(i,j)} A$, se obtiene la matriz B.

Ejemplo IV. 7.

Sea $A=\begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 5 & 6 \end{bmatrix}$

intercambiamos sus renglones 1 y 3 para obtener la matriz

$$B=\begin{bmatrix} 5 & 6 \\ 1 & 0 \\ 1 & 2 \end{bmatrix}$$

Por otro lado, consideremos la matriz

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

e intercambiemos sus renglones 1 y 3 para obtener

$$I_3^{(1,3)}=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Efectuando el producto $I_3^{(1,3)}A$, tenemos:

$$I_3^{(1,3)}A=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 1 & 0 \\ 1 & 2 \end{bmatrix}$$

vemos entonces que, $I_3^{(1,3)} A$ es igual a B.

- -

2) Multiplicación de un renglón por un número $k\neq 0$.

Supongamos que tenemos una matriz A de mxn y que multiplicamos su renglón i por el escalar $k\neq 0$. Llamemos a la matriz así obtenida matriz B.

Si por otro lado, tomamos la matriz identidad de orden m (I_m) y multiplicamos su renglón i por el escalar $k\neq 0$ obtendremos una nueva matriz que llamaremos $I_m^{k(i)}$.

Si ahora efectuamos el producto $I_m^{k(i)}A$, se obtiene la ma triz B.

Ejemplo IV. 8.

Sea $A=\begin{bmatrix}1 & 2 & 1\\ 5 & 10 & 20\end{bmatrix}$

multipliquemos el segundo renglón por 3 para obtener la matriz

$$B=\begin{bmatrix}1 & 2 & 1\\ 15 & 30 & 60\end{bmatrix}$$

Por otro lado, consideremos la matriz

$$I_2=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

y multipliquemos su segundo renglón por 3 para obtener

$$I_2^{3(2)}=\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}$$

Efectuando el producto $I_2^{3(2)}A$, obtenemos:

$$I_2^{3(2)}A=\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\ 5 & 10 & 20\end{bmatrix}=\begin{bmatrix}1 & 2 & 1\\ 15 & 30 & 60\end{bmatrix}$$

vemos entonces que, $I_2^{3(2)}A$ es igual a B.

- -

3) Multiplicación de un renglón por un número $k\neq 0$, sumando el resultado a otro renglón diferente.

Supongamos que tenemos una matriz A de mxn en la que multiplicamos su renglón i por el escalar $k\neq 0$ y el resultado lo sumamos al renglón $j\neq i$. Llamemos a la matriz así obtenida, matriz B.

Si por otro lado, tomamos la matriz identidad de orden m (I_m), multiplicamos su renglón i por el escalar $k\neq 0$ y el resultado lo sumamos al renglón j, obtendremos una nueva matriz que llamaremos $I_m^{k(i,j)}$.

Si ahora efectuamos el producto $I_m^{k(i,j)}A$, se obtiene la matriz B.

Ejemplo IV. 9

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

multipliquemos el primer renglón por 2 y sumemos al cuarto renglón para obtener la matriz

$$B=\begin{bmatrix}1 & 2\\ 3 & 4\\ 0 & 1\\ 7 & 14\end{bmatrix}$$

Por otro lado consideremos la matriz

$$I_4=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

multiplicando el primer renglón por 2 y sumando al cuarto renglón obtenemos

$$I_4^{2(1,4)}=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 2 & 0 & 0 & 1\end{bmatrix}$$

Efectuando el producto $I_4^{2(1,4)}A$, obtenemos:

$$I_4^{2(1,4)}A=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 2 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 2\\ 3 & 4\\ 0 & 1\\ 5 & 10\end{bmatrix}=\begin{bmatrix}1 & 2\\ 3 & 4\\ 0 & 1\\ 7 & 14\end{bmatrix}$$

vemos entonces que, $I_3^{k(1,3)}A$ es igual a B.

Definición.

A las matrices $I_m^{(i,j)}$, $I_m^{k(i)}$, $I_m^{k(i,j)}$ se les llama matrices elementales correspondientes a las transformaciones elementales 1, 2, 3, respectivamente.

Vemos ahora que, cada transformación elemental puede ser llevada a cabo premultiplicando, la matriz dada, por la matriz elemental que se obtiene efectuando en I, la misma transformación elemental.

Ejemplo IV. 10.

Hallar la matriz P, de orden 3, tal que transforme a la matriz A en una matriz escalonada (es decir, que el producto PA sea una matriz escalonada).

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

Solución.

En la siguiente tabla aparece una secuencia de transformaciones elementales que transforma a la matriz A en una matriz escalonada, las correspondientes matrices elementales y el resultado de efectuar estas transformaciones sobre A.

Transformación.	Matriz elemental correspondiente.	$\begin{bmatrix} 0 & 1 & -2 \\ \frac{1}{2} & 6 & 1 \\ 2 & \frac{1}{7} & 0 \end{bmatrix}=A$
Intercambio de los renglones 1 y 2	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{2} & 6 & 1 \\ 0 & 1 & -2 \\ 2 & \frac{1}{7} & 0 \end{bmatrix}$
Multiplicación del primer renglón por 2	$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 12 & 2 \\ 0 & 1 & -2 \\ 2 & \frac{1}{7} & 0 \end{bmatrix}$
Multiplicación del renglón 1 por -2 y sumar al renglón 3.	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 12 & 2 \\ 0 & 1 & -2 \\ 0 & -\frac{167}{7} & -4 \end{bmatrix}$
Multiplicación del renglón 2 por $\frac{167}{7}$ y sumar al renglón 3.	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{167}{7} & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 12 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & -\frac{362}{7} \end{bmatrix}$
Multiplicación del renglón 3 por $-\frac{7}{362}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{7}{362} \end{bmatrix}$	$\begin{bmatrix} 1 & 12 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ (matriz escalonada)

El mismo resultado que se obtiene al aplicar la secuencia de transformaciones, puede obtenerse premultiplicando por las respectivas matrices elementales.

Efectuando estos productos con la secuencia que marca la tabla y dado que la multiplicación es asociativa, podemos escribir:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&\frac{-7}{362}\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&\frac{167}{7}&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\-2&0&1\end{bmatrix}\begin{bmatrix}2&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\underbrace{\begin{bmatrix}0&1&-2\\\frac{1}{2}&6&1\\2&\frac{1}{7}&0\end{bmatrix}}_{A}=\underbrace{\begin{bmatrix}1&12&2\\0&1&-2\\0&0&1\end{bmatrix}}_{\text{Matriz escalón}}$$

La matriz P buscada, será el producto de las cinco matrices elementales.

Sin embargo, para obtener la matriz P no es necesario multiplicar las cinco matrices elementales; bastará con efectuar la misma secuencia de transformaciones elementales sobre la matriz I (referencia 1, pag. 452).

Las siguientes transformaciones sobre I conducen a la matriz P (las transformaciones son las de la tabla).

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\rightarrow\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\rightarrow\begin{bmatrix}0&2&0\\1&0&0\\0&0&1\end{bmatrix}\rightarrow\begin{bmatrix}0&2&0\\1&0&0\\0&-4&1\end{bmatrix}\rightarrow\begin{bmatrix}0&2&0\\1&0&0\\\frac{167}{7}&-4&1\end{bmatrix}\rightarrow\begin{bmatrix}0&2&0\\1&0&0\\-\frac{167}{362}&\frac{14}{181}&\frac{-7}{362}\end{bmatrix}=P$$

Para comprobar, efectuemos el producto PA

$$PA=\begin{bmatrix}0&2&0\\1&0&0\\-\frac{167}{362}&\frac{14}{181}&\frac{-7}{362}\end{bmatrix}\begin{bmatrix}0&1&-2\\\frac{1}{2}&6&1\\2&\frac{1}{7}&0\end{bmatrix}=\begin{bmatrix}1&12&2\\0&1&-2\\0&0&1\end{bmatrix}$$

- -

Una situación interesante, se presenta en el caso de las matrices cuadradas de orden n cuyo rango es también n, las cuales pueden transformarse en la matriz identidad I_n.

Estas matrices son equivalentes a una matriz escalonada de la forma

$$\begin{bmatrix}1&a_{12}&a_{13}&\cdots&a_{1,n-1}&a_{1n}\\0&1&a_{23}&\cdots&a_{2,n-1}&a_{2n}\\0&0&1&a_{34}\cdots&a_{3,n-1}&a_{3n}\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&0\cdots&1&a_{n-1\,n}\\0&0&0&0\cdots&0&1\end{bmatrix}\qquad(10)$$

donde todos los elementos a_{ij} tales que $i=j$ (a los que se llama elementos de la diagonal principal) son iguales a 1. Es decir, $a_{11}=a_{22}=a_{33}=\ldots=a_{nn}=1$.

Mostraremos ahora que, una matriz de este tipo, se puede transformar en una matriz identidad I_n mediante una secuencia de transformaciones elementales.

En efecto, si el último renglón de la matriz (10) lo multiplicamos por $-a_{1n}$ y lo sumamos al primer renglón, luego lo multiplicamos por $-a_{2n}$ y lo sumamos al segundo renglón, etc., obtenemos la matriz.

$$\begin{bmatrix}1&a_{12}&a_{13}&\cdots&a_{1,n-1}&0\\0&1&a_{23}&\cdots&a_{2,n-1}&0\\0&0&1&\cdots&a_{3,n-1}&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1&0\\0&0&0&\cdots&0&1\end{bmatrix}$$

Si ahora, el penúltimo renglón de esta nueva matriz lo multiplicamos por $-a_{1,n-1}$ y lo sumamos al primer renglón, luego lo multiplicamos por $-a_{2,n-1}$ y lo sumamos al segundo renglón, etc., obtenemos

$$\begin{bmatrix} 1 & a_{12} & a_{13} & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 1 & a_{23} & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 1 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix}.$$

Al final de este proceso se obtendrá la matriz identidad I_n.

Ejemplo IV.-1).

Usando transformaciones elementales, transformaremos la matriz escalón

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

en una matriz identidad:

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

Este procedimiento puede aplicarse a cualquier matriz escalonada arbitraria.

- -

Sin embargo, es importante notar que una matriz cuadrada de orden n, en forma escalonada y cuyo rango sea menor que n, no podrá nunca transformarse en una matriz identidad. Esto se debe a que una matriz de este tipo, tendrá siempre en su último renglón únicamente ceros.

Así mismo, una matriz no cuadrada tampoco podrá transformarse en una matriz identidad (ya que ésta es una matriz cuadrada).

<u>Matriz Inversa</u>.

Con lo que hemos visto hasta ahora, sabemos que si se tiene una matriz cuadrada A de orden n y rango n, es posible (mediante una serie de transformaciones elementales) transformarla primero en una matriz escalonada y luego en la matriz identidad I_n.

También se vió anteriormente, que el efecto de una sucesión finita de transformaciones elementales sobre cualquier matriz A, puede obtenerse premultiplicando A por una cierta matriz P. Por lo que, el efecto de toda la secuencia de transformaciones utilizadas para llevar la matriz A a la matriz I_n, se puede obtener premultiplicando A por una cierta matriz P.

Podemos entonces enunciar el siguiente

<u>Teorema</u> IV. 3.

Si A es una matriz cuadrada de orden n y rango n, existe una matriz P tal que

$$PA = I_n$$

No es difícil demostrar[1] que dicha matriz P, cumple también con

$$AP = I_n$$

aunque la multiplicación de matrices no sea conmutativa.

Puesto que la matriz I_n, es el elemento idéntico para la

(1) El estudiante puede obtener una demostración, a partir de la demostración del teorema 10-4·9 de la referencia I (pags. 454 y 455).

multiplicación en el conjunto de las matrices cuadradas de orden n, resulta adecuada la siguiente

Definición.

Si A y P son dos matrices cuadradas de orden n tales que

$$PA=AP=I_n$$

a la matriz P le llamaremos matriz inversa de A.

Si una matriz A tiene inversa, diremos que es no singular y a su inversa la representaremos con A^{-1}. Se puede demostrar que la inversa A^{-1} de una matriz A es única y también que (referencia 1 pag. 444) el producto de dos matrices no singulares es una matriz no singular.

Dado que una matriz cuadrada de orden n y rango menor que n no puede transformarse en una matriz identidad, no existe ninguna matriz P tal que

$$PA = I_n$$

En consecuencia, estas matrices no tienen inversa. A las matrices que no tienen inversa las llamaremos matrices singulares.

Con los conceptos tratados hasta ahora, el estudiante debe poder demostrar el siguiente

Teorema IV. 4.

Si A es una matriz de orden mxn, existe su inversa A^{-1} si y solo si

$$m=n= R(A).$$

Ejemplo IV. 12.

Investigue si la siguiente matriz tiene inversa

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

Solución

Vamos a transformar A en una matriz escalón

$$\begin{bmatrix} 1 & -2 & 3 & 0 \\ 3 & -1 & \frac{7}{2} & 1 \\ 4 & 2 & 1 & 0 \\ 2 & 0 & \frac{8}{5} & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 5 & -\frac{11}{2} & 1 \\ 4 & 2 & 1 & 0 \\ 2 & 0 & \frac{8}{5} & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 5 & -\frac{11}{2} & 1 \\ 0 & 10 & -11 & 0 \\ 2 & 0 & \frac{8}{5} & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 5 & -\frac{11}{2} & 1 \\ 0 & 10 & -11 & 0 \\ 0 & 4 & -\frac{22}{5} & 1 \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & -\frac{11}{10} & \frac{1}{5} \\ 0 & 10 & -11 & 0 \\ 0 & 4 & -\frac{22}{5} & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & -\frac{11}{10} & \frac{1}{5} \\ 0 & 0 & 0 & -2 \\ 0 & 4 & -\frac{22}{5} & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & -\frac{11}{10} & \frac{1}{5} \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & \frac{1}{5} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & -\frac{11}{10} & \frac{1}{5} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \frac{1}{5} \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & -\frac{11}{10} & \frac{1}{5} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

vemos que la matriz escalonada tiene tres renglones diferentes de cero, en consecuencia R(A)=3. Entonces, A es una matriz singular (no existe A^{-1}).

- -

Para obtener la matriz inversa de una matriz no singular, haremos uso del siguiente teorema.

Teorema IV. 5.

La inversa de una matriz A no singular se puede obtener si aplicamos a I, la misma secuencia de transformaciones elementales por renglón que se utilizan para transformar la matriz A en la matriz I.

Demostración

Sabemos que podemos transformar la matriz A en I mediante una cierta secuencia de transformaciones elementales. Entonces, existe una secuencia de matrices elementales $E_1, E_2, \ldots, E_{k-1}, E_k$ tales que

$$E_k E_{k-1} \ldots E_2 E_1 A = I$$

si llamamos P al producto de las matrices elementales

$$P = E_k E_{k-1} \ldots E_2 E_1$$

podemos escribir

$$PA = I$$

además, sabemos que

$$AP = I$$

Es decir; P es la inversa de A.

Como I es el idéntico para el producto

$$P = PI$$

entonces

$$P = E_k E_{k-1} \ldots E_2 E_1 I$$

lo cual nos indica que P se obtiene a partir de I, efectuando en orden las transformaciones elementales que corresponden a $E_1, E_2, \ldots E_{k-1}$ y E_k

Ejemplo IV. 13.

Investigar si la siguiente matriz tiene inversa y en caso afirmativo hallarla.

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

Solución:

Aplicaremos el proceso descrito en el teorema IV.5. La primera columna de la siguiente tabla, describe la secuencia de transformaciones elementales por renglón, utilizada para transformar la matriz A en la matriz I. La segunda columna, muestra las matrices obtenidas a partir de A con la aplicación de estas transformaciones. La tercera columna, muestra las matrices obtenidas a partir de I con la aplicación de las mismas transformaciones.

Transformaciones		
	$\begin{bmatrix} 0 & 2 & 2 \\ 1 & -2 & 1 \\ -1 & 1 & 5 \end{bmatrix} = A$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$
Intercambio del primero y segundo renglones	$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & 2 \\ -1 & 1 & 5 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Suma del primer renglón al tercero	$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & 2 \\ 0 & -1 & 6 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$
Multiplicación del segundo renglón por $\frac{1}{2}$	$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & 6 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$
Suma del segundo renglón al tercero	$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 7 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 1 & 1 \end{bmatrix}$
Multiplicación del ter		

cer renglón por $\frac{1}{7}$

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{7} & 0 & 0 \\ \frac{1}{14} & \frac{1}{7} & \frac{1}{7} \end{bmatrix}$$

Hasta ahora hemos transformado a la matriz A en una matriz escalonada. Vemos que el rango es 3, e igual al orden de la matriz, por lo tanto la matriz A tiene inversa.

Continuando el proceso:

Multiplicación del tercer renglón por -1 y sumar al segundo renglón

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 \\ \frac{6}{14} & -\frac{1}{7} & -\frac{1}{7} \\ \frac{1}{14} & \frac{1}{7} & \frac{1}{7} \end{bmatrix}$$

Multiplicación del tercer renglón por -1 y sumar al primer renglón

$$\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} -\frac{1}{14} & \frac{6}{7} & -\frac{1}{7} \\ \frac{6}{14} & -\frac{1}{7} & -\frac{1}{7} \\ \frac{1}{14} & \frac{1}{7} & \frac{1}{7} \end{bmatrix}$$

Multiplicación del segundo renglón por 2 y sumar al primer renglón

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I \qquad \begin{bmatrix} \frac{11}{14} & \frac{4}{7} & -\frac{3}{7} \\ \frac{6}{14} & -\frac{1}{7} & -\frac{1}{7} \\ \frac{1}{14} & \frac{1}{7} & \frac{1}{7} \end{bmatrix} = A^{-1}$$

La matriz que se encuentra al final de la segunda columna, es la matriz identidad, en consecuencia, por el teorema IV. 5, la matriz que se encuentra al final de la tercera columna es la inversa de A. Por lo que:

$$A^{-1} = \begin{bmatrix} \frac{11}{14} & \frac{4}{7} & -\frac{3}{7} \\ \frac{6}{14} & -\frac{1}{7} & -\frac{1}{7} \\ \frac{1}{14} & \frac{1}{7} & \frac{1}{7} \end{bmatrix}$$

Podemos comprobar el resultado efectuando el producto:

$$A^{-1} A = \begin{bmatrix} \frac{11}{14} & \frac{4}{7} & -\frac{3}{7} \\ \frac{6}{14} & -\frac{1}{7} & -\frac{1}{7} \\ \frac{1}{14} & \frac{1}{7} & \frac{1}{7} \end{bmatrix} \begin{bmatrix} 0 & 2 & 2 \\ 1 & -2 & 1 \\ -1 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

- -

En ocasiones, en lugar de hacer una tabla como la del ejemplo anterior, se trabaja con un arreglo que contiene a la matriz A en el lado izquierdo y a la matriz I en el lado derecho. Se efectúan las mismas transformaciones en ambas matrices, hasta que se obtenga la matriz I en el lado izquierdo. La matriz que resulta en el lado derecho es A^{-1}.

En forma esquemática:

$$\left[A \;\vdots\; I \right] \rightarrow \ldots \rightarrow \left[I \;\vdots\; A^{-1} \right]$$

IV. 4. SUMA DE MATRICES.

Definición.

Sean $A=(a_{ij})$ y $B=(b_{ij})$ dos matrices del mismo orden mxn. La adición o suma A+B de dichas matrices es una nueva matriz $C=(c_{ij})$ de orden mxn tal que

$$c_{ij}=a_{ij}+b_{ij}$$

Es decir, los elementos de la matriz C son las sumas de los elementos correspondientes de A y B.

Ejemplo IV. 14.

Dadas las matrices

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

a) Obtener A+B

b) Obtener C+D

Solución

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

b) No puede efectuarse la suma C+D dado que las matrices no son del mismo orden, en estos casos se dice que las matrices no son conformables para la suma.

- -

Propiedades de la suma de matrices.

El estudiante puede demostrar las siguientes propiedades.

Sean A, B y C tres matrices del mismo orden (mxn), se cumple siempre que:

1.- $(A+B)+C=A+(B+C)$ la suma es asociativa

2.- $A+B = B+A$ la suma es conmutativa

3.- Existe una matriz $0=(c_{ij})$ (donde $c_{ij}=0\ \forall\ i, j$) de orden mxn, a la que llamaremos matriz nula, tal que

$$A+0=0+A=A$$

4.- Para toda matriz $A=(a_{ij})$ de orden mxn, existe una matriz a la que llamaremos simétrica de A y representaremos con -A, tal que

$$A+(-A)=(-A)+A=0$$

(fácilmente se puede comprobar que la simétrica de $A=(a_{ij})$ es $-A=(-a_{ij})$)

De la definición de suma y las propiedades anteriores, vemos que el conjunto de las matrices del mismo orden forman un grupo abeliano.

Un caso particular, es el de las matrices cuadradas de orden n. El conjunto (M) de estas matrices, con las operaciones de suma y producto, forma un anillo con unidad (no conmutativo), ya que,

$\forall A, B, C \in (M)$ se cumple siempre que:

1) $A + B \in (M)$

$A\ B \in (M)$

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

$(AB)C= A(BC)$

3) $\exists\ 0 \in (M)$ tal que $A+0= 0+A=A$

$\exists\ I_n \in (M)$ tal que $A\ I_n=I_n\ A=A$

4) $\forall$ $A \in (M)$, $\exists$ $-A \in (M)$ tal que $A+(-A) = -A+A = 0$

5) $A+B = B+A$

6) $A(B+C) = AB+AC$ y $(B+C)A = BA + CA$

Las propiedades 1 a 5, se han tratado ya para el caso general de matrices de orden $m \times n$. La propiedad distributiva (6) la demostraremos a continuación.

Demostración.

Sean $A=(a_{ij})$, $B=(b_{ij})$ y $C=(c_{ij})$ tres matrices cualesquiera de orden n.

De la definición de suma

$$B+C= (b_{ij}+c_{ij})$$

De la definición de producto

$$A(B+C)=(a_{ij})(b_{ij}+c_{ij})$$

$$A(B+C)=\sum_{j=1}^{n} a_{ij}(b_{ij}+c_{ij})$$

Como la multiplicación es distributiva sobre la suma en C y por las propiedades de la suma

$$A(B+C)=\sum_{j=1}^{n} a_{ij}b_{ij}+\sum_{j=1}^{n} a_{ij}c_{ij}$$

De la definición de producto: $A(B+C) = AB+AC$.

La segunda parte se demuestra en forma análoga.

La propiedad distributiva del producto sobre la suma de matrices, tanto por la izquierda como por la derecha, puede generalizarse (referencia 3 pág. 26) de la siguiente forma:

> $A(B+C)=AB+AC$ y $(B+C)A=BA+CA$
>
> siempre y cuando las operaciones indicadas puedan ser efectuadas.

Definiremos la resta o sustracción de matrices, a partir de la suma, como

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

lo cual equivale simplemente a restar de los elementos de A, los correspondientes de B. Resulta claro según la definición que, para que dos matrices sean conformables para la resta deben serlo para la suma.

Ejemplo IV. 15.

Para las matrices

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

a) Obtener A-B

b) Obtener A-C

Solución

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

b) A-C no puede efectuarse.

- -

Producto por un escalar.

> Definición.
>
> Sean: $A=(a_{ij})$ una matriz de orden $m \times n$ y $\alpha \in C$ un escalar.
>
> El producto α por A, que representaremos mediante αA, es la matriz
>
> $\alpha A=(\alpha a_{ij})$

Ejemplo IV. 16.

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

el producto αA es la matriz

$$\alpha A=\begin{bmatrix} 0 & -3i \\ -3 & 9i \\ 3i & -3+6i \end{bmatrix}$$

- -

El estudiante puede fácilmente demostrar que esta operación tiene las siguientes propiedades.

Sean A y B dos matrices del mismo orden y α, β dos números complejos, se cumple siempre que:

1.- $(\alpha+\beta)A=\alpha A+\beta A$

2.- $\alpha(A+B)=\alpha A+\alpha B$

3.- $\alpha(\beta A)=(\alpha\beta)A$

4.- $1\cdot A=A$

Transpuesta de una matriz.

Definición.

Sea la matriz A de orden mxn. Llamaremos "transpuesta de A" y la representaremos mediante A^T, a la matriz de orden nxm cuyos renglones son las columnas de A y cuyas columnas son los renglones de A.

es decir:

si $A=(a_{ij})$ entonces $A^T=(a_{ji})$

Se puede demostrar (referencia 3 pag. 33) que

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

Ejemplo IV. 17.

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

obtener $(AB)^T$ y $B^T A^T$.

Solución

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

$$(AB)^T=\begin{bmatrix} 5 & 5 & 0 \\ 10 & 10 & 0 \end{bmatrix}$$

$$B^T A^T=\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 3 & 5 & 2 \end{bmatrix}=\begin{bmatrix} 5 & 5 & 0 \\ 10 & 10 & 0 \end{bmatrix}$$ $\therefore$ se

comprueba que $(AB)^T=B^T A^T$

- -

Ecuaciones matriciales.

Ejemplo IV. 18.

Dadas las matrices

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

obtener una matriz x tal que

$Ax - B^T = C + Dx$

Solución

En este caso tenemos una ecuación entre matrices, donde la incógnita es la matriz x. Este tipo de ecuaciones pueden resolverse empleando las propiedades de las operaciones que hemos definido, siempre y cuando la ecuación haya sido planteada correc tamente.

Para el caso que nos ocupa, x debe ser tal que

$$Ax-B^T=C+Dx$$

sumando en ambos miembros	$-Dx$:
$-Dx+(Ax-B^T)=-Dx+(C+Dx)$	
$-Dx+(Ax-B^T)=-Dx+(Dx+C)$	La suma es conmutativa.
$-Dx+Ax-B^T=(-Dx+Dx)+C$	La suma es asociativa.
$-Dx+Ax-B^T=0+C$	$-Dx$ es el inverso para la suma de Dx.
$-Dx+Ax-B^T=C$	La matriz nula es el idéntico para la suma.
$(-Dx+Ax-B^T)+B^T=C+B^T$	Sumando en ambos miembros B^T.
$-Dx+Ax+(-B^T+B^T)=C+B^T$	La suma es asociativa.
$-Dx+Ax+0=C+B^T$	$-B^T$ es el inverso para la suma de B^T.
$-Dx+Ax=C+B^T$	La matriz nula es el idéntico para la suma.
$(-D+A)x=C+B^T$	La multiplicación es distributiva sobre la suma.
$(-D+A)^{-1}((-D+A)x)=(-D+A)^{-1}(C+B^T)$	Suponiendo que existe $(-D+A)^{-1}$ premultiplicamos ambos miembros por dicha matriz.
$((-D+A)^{-1}(-D+A))x=(-D+A)^{-1}(C+B^T)$	Por asociatividad del producto.
$I_2x=(-D+A)^{-1}(C+B^T)$	Por la definición de matriz inversa.

Finalmente, como I_2 es el idéntico para el producto y por la conmutatividad de la suma

$$x=(A-D)^{-1}(C+B^T)$$

Lo que hemos hecho es despejar la incógnita x de la ecuación matricial, justificando los pasos de dicho despeje. Pueden omitirse estas justificaciones y algunos pasos que se consideren -- obvios, con el objeto de hacer mas corto el proceso.

Sin embargo, el estudiante debe tener cuidado al despejar una incógnita de una ecuación matricial, ya que existen diferencias entre las propiedades de las operaciones con matrices y con números reales.

Efectuemos las operaciones indicadas en la última expresión:

$$C+B^T=\begin{bmatrix}2 & -1 & 4\\ 0 & -3 & -3\end{bmatrix}+\begin{bmatrix}0 & 3 & -5\\ -1 & 2 & 4\end{bmatrix}=\begin{bmatrix}2 & 2 & -1\\ -1 & -1 & 1\end{bmatrix}$$

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

Veamos si efectivamente existe $(A-D)^{-1}$:

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

$$\therefore\quad (A-D)^{-1}=\begin{bmatrix}-\frac{1}{2} & \frac{1}{4}\\ 0 & -\frac{1}{2}\end{bmatrix}=\frac{1}{4}\begin{bmatrix}-2 & 1\\ 0 & -2\end{bmatrix}$$

Por lo que, la incógnita x puede obtenerse como:

$$x=(A-D)^{-1}(C+B^T)=\frac{1}{4}\begin{bmatrix}-2 & 1\\ 0 & -2\end{bmatrix}\begin{bmatrix}2 & 2 & -1\\ -1 & -1 & 1\end{bmatrix}=$$

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

Finalmente, la matriz pedida es

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

- -

La principal diferencia entre las matrices y los números reales es que, mientras que podemos sumar o multiplicar dos números cualesquiera, no siempre podemos hacerlo con las matrices. Suponiendo que pueden efectuarse las operaciones indicadas, enlistamos a continuación las principales diferencias entre las propiedades de las operaciones con números y con matrices:

1) La multiplicación de números es conmutativa; la multiplicación de matrices no lo es.

2) Si definimos $A^2=AA$, el desarrollo matricial $(A+B)^2=A^2+2AB+B^2$ es en general falso (como consecuencia de 1). El desarrollo correcto es $(A+B)^2=A^2+AB+BA+B^2$.

3) El producto de dos números diferentes de cero nunca es cero, pero el producto de dos matrices diferentes de la matriz nula puede ser igual a cero (la matriz nula).

Por ejemplo, si

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

se tiene que $AB=0$

4) La ley cancelativa para el producto se verifica en los números, pero no en las matrices.

Esto es, si $a, b \in \mathbb{R}$ y $a \neq 0$

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

pero en las matrices

$$(AB = AC) \not\Rightarrow (B=C)$$

Por ejemplo, con las matrices A y B del ejemplo anterior tenemos que

$AB = A0$ pero $B\neq0$, por lo que no podemos cancelar A.

Nota. Aunque $(AB=AC) \not\Rightarrow (B=C)$, sí se cumple que

$$(B=C) \implies (AB=AC) \quad \forall A.$$

Antes de pasar a la sección siguiente, es conveniente aclarar que aunque hemos definido matriz como un arreglo de números (reales o complejos), el concepto de matriz puede generalizarse como "un arreglo de números, funciones u operadores". En el siguiente capítulo y en cursos posteriores se verá la utilidad de dicha extensión.

IV. 5. SOLUCION DE SISTEMAS DE ECUACIONES.

Como vimos al inicio de la sección IV. 3, podemos representar el sistema

$$\left.\begin{aligned} 20x_1+100x_2+40x_3&=400 \\ 0x_1+\ \ x_2-\ \ x_3&=0 \end{aligned}\right\}\ldots\ldots(4)$$

mediante la expresión matricial

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

A esta ecuación matricial, donde $\bar{x}$ es la matriz incógnita o indeterminada, se le conoce como "forma matricial del sistema de ecuaciones (4)".

La matriz

$$A=\begin{bmatrix} 20 & 100 & 40 \\ 0 & 1 & -1 \end{bmatrix}$$

recibe el nombre de "matriz de coeficientes del sistema", y a las matrices

$$\bar{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\ \text{y}\qquad \bar{b}=\begin{bmatrix} 400 \\ 0 \end{bmatrix}$$

se les llama "vector de incógnitas" y "vector de términos independientes" respectivamente.

En general, la forma matricial del sistema

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

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

donde la matriz de coeficientes es la matriz de mxn:

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

el vector de incógnitas es la matriz de nx1:

$$\bar{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

y el vector de términos independientes es la matriz de mx1:

$$\bar{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Es claro que resolver la ecuación matricial $A\bar{x}=\bar{b}$ es equivalente a resolver el sistema, por lo que una solución de dicho sistema será una matriz columna de n renglones

$$\bar{k}=\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}\qquad (k_i\ \varepsilon\ C)$$

tal que $A\bar{k}=\bar{b}$.

Definiremos una matriz más para el sistema, a la que llamaremos "matriz ampliada" del sistema, la cual juega un papel importante en el estudio teórico de los sistemas de ecuaciones.

Definición.

La matriz ampliada de un sistema de ecuaciones, que representaremos con $(A, \bar{b})$, es la matriz de orden $m\times(n+1)$ que resulta de aumentar, a la matriz de coeficientes, el vector de términos independientes.

Entonces, la matriz ampliada del sistema es:

$$(A,\bar{b})=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}$$

Sistemas incompatibles.

Si analizamos el sistema

$$x_1+x_2+x_3=6$$
$$x_1+2x_2-x_3=1$$
$$x_1+x_2+x_3=-5$$

vemos que las ecuaciones 1a. y 3a. no pueden satisfacerse simultáneamente, ya que no existen tres números x_1, x_2, x_3 cuya suma sea igual a 6 y -5 a la vez. El sistema es entonces incompatible.

Veamos que sucede con los rangos de las matrices

$$A=\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \quad y \quad (A,\bar{b})=\begin{bmatrix} 1 & 1 & 1 & 6 \\ 1 & 2 & -1 & 1 \\ 1 & 1 & 1 & -5 \end{bmatrix}$$

Al efectuar solamente transformaciones por renglón, es posible determinar a la vez los rangos de las matrices A y $(A, \bar{b})$ operando sobre la matriz ampliada, la cual contiene a la matriz A.

Transformaremos entonces en una matriz escalón el siguiente arreglo.

$$\underbrace{\left[\begin{array}{ccc:c} \multicolumn{3}{c}{\overbrace{\hphantom{1 \quad 1 \quad 1}}^{A}} \\ 1 & 1 & 1 & 6 \\ 1 & 2 & -1 & 1 \\ 1 & 1 & 1 & -5 \end{array}\right]}_{(A,\bar{b})} \rightarrow \left[\begin{array}{ccc:c} 1 & 1 & 1 & 6 \\ 1 & 2 & -1 & 1 \\ 0 & 0 & 0 & -11 \end{array}\right] \rightarrow \left[\begin{array}{ccc:c} 1 & 1 & 1 & 6 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

De la última matriz vemos que

$R(A) = 2$

y $R(A, \bar{b}) = 3$

es decir $R(A) < R(A, \bar{b})$, que es una característica de todos los sistemas incompatibles.

Sabemos que el sistema es incompatible porque las ecuaciones 1a. y 3a. no admiten solución simultánea. Daremos otra prueba de la incompatibilidad de dicho sistema:

La última matriz

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

representa el siguiente sistema equivalente

$$x_1+x_2+x_3=6$$
$$0x_1+x_2-2x_3=-5$$
$$0x_1+0x_2+0x_3=1$$

donde vemos que no existen valores para x_1, x_2, x_3 que satisfagan la 3a. ecuación. Esto implica que el sistema equivalente no tiene solución y en consecuencia el sistema original tampoco.

Conviene hacer la siguiente observación.

El rango de la matriz ampliada de un sistema de ecuaciones lineales, es igual al rango de la matriz de coeficientes o mayor en una unidad, por lo que:

$$R(A) \leq R(A, \bar{b})$$

En efecto, sea un sistema de m ecuaciones con n incógnitas

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

cuya matriz de coeficientes es:

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

y cuya matriz ampliada es:

$$(A,\bar{b})=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix} \ \dots(13)$$

Sea el rango de la matriz de coeficientes $R(A)=r$. Efectuando en $(A, \bar{b})$ las mismas transformaciones que se efectúan para llevar la matriz A a su forma escalonada, obtendremos la matriz:

$$\begin{array}{c} r \\ \text{renglones} \end{array}\left\{\begin{bmatrix} 1 & c_{12} & \dots & c_{1r} & c_{1,r+1} & \dots & c_{1n} & d_1 \\ 0 & c_{22} & \dots & c_{2r} & c_{2,r+1} & \dots & c_{2n} & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{rr} & c_{r,r+1} & \dots & c_{rn} & d_r \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & d_{r+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & d_m \end{bmatrix}\right.$$

continuando con el proceso, hasta transformar esta matriz en una matriz escalonada obtendremos:

$$\begin{bmatrix} 1 & c_{12} & \dots & c_{1r} & c_{1,r+1} & \dots & c_{1n} & d_1 \\ 0 & c_{22} & \dots & c_{2r} & c_{2,r+1} & \dots & c_{2n} & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{rr} & c_{r,r+1} & \dots & c_{rn} & d_r \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & e \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \end{bmatrix} \ \dots(14)$$

$$\text{donde} \left\{\begin{array}{lll} c_{ij}=1 & \text{si} & j=k \\ c_{ij}=0 & \text{si} & j<k \end{array}\right\} \text{para algún } k \geq i.$$

Existen entonces sólo dos posibilidades:

a) Si $e=0$, $R(A, \bar{b})=r=R(A)$

b) Si $e\neq 0$, podemos dividir el $r+1$ ésimo renglón entre e, obteniendo:

$$\begin{bmatrix} 1 & c_{12} & \dots & c_{1r} & c_{1,r+1} & \dots & c_{1n} & d_1 \\ 0 & c_{22} & \dots & c_{2r} & c_{2,r+1} & \dots & c_{2n} & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{rr} & c_{r,r+1} & \dots & c_{rn} & d_r \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \end{bmatrix} \ \dots(15)$$

de donde $R(A,\bar{b}) = r+1 = R(A)+1$

Entonces $R(A) < R(A, \bar{b})$.

Por tanto, hemos demostrado que $R(A) \leq R(A, \bar{b})$.

<u>Teorema</u> IV. 6.

Si en un sistema de ecuaciones lineales

$$R(A) < R(A,\bar{b})$$

entonces el sistema es incompatible.

Demostración

Si $R(A) < R(A, \bar{b})$, la matriz ampliada del sistema (11) puede transformarse en la matriz (15). En esta última, el r+1 ésimo renglón representa la ecuación

$0x_1+0x_2+ \dots +0x_n=1$

de un sistema equivalente.

Obviamente, esta ecuación no se satisface para ningún conjunto de valores $x_1, x_2, \dots, x_n$, por lo que el sistema es incompatible.

Ejemplo IV. 19.

Investigue si el siguiente sistema de ecuaciones tiene solución

$2x_1+3x_2+5x_3=2$

$6x_1+9x_2+15x_3=1$

Transformemos la matriz de coeficientes en una matriz escalonada

$$\begin{bmatrix}2 & 3 & 5\\ 6 & 9 & 15\end{bmatrix} \to \begin{bmatrix}1 & \frac{3}{2} & \frac{5}{2}\\ 6 & 9 & 15\end{bmatrix} \to \begin{bmatrix}1 & \frac{3}{2} & \frac{5}{2}\\ 0 & 0 & 0\end{bmatrix}$$

$\therefore \ R(A)=1$

Transformemos ahora la matriz ampliada en una matriz escalonada

$$\begin{bmatrix}2 & 3 & 5 & 2\\ 6 & 9 & 15 & 1\end{bmatrix} \to \begin{bmatrix}1 & \frac{3}{2} & \frac{5}{2} & 1\\ 6 & 9 & 15 & 1\end{bmatrix} \to \begin{bmatrix}1 & \frac{3}{2} & \frac{5}{2} & 1\\ 0 & 0 & 0 & -5\end{bmatrix} \to$$

$$\to \begin{bmatrix}1 & \frac{3}{2} & \frac{5}{2} & 1\\ 0 & 0 & 0 & 1\end{bmatrix}$$

$R(A,\bar{b})=2$

Como $R(A) < R(A,\bar{b})$, el sistema no tiene solución.

Obsérvese que se hubiera llegado a la misma conclusión de haber trabajado directamente con la matriz ampliada.

- -

Sistemas compatibles determinados.

Analicemos primeramente el caso particular de los sistemas de n ecuaciones con n incógnitas, en los cuales la matriz de coeficientes es de rango n.

Un sistema de n ecuaciones con n incógnitas tiene la forma

$$\begin{matrix} a_{11} & x_1 & + \dots + & a_{1n} & x_n & = b_1 \\ a_{21} & x_1 & + \dots + & a_{2n} & x_n & = b_2 \\ \dots & & & & & \\ \dots & & & & & \\ a_{n1} & x_1 & + \dots + & a_{nn} & x_n & = b_n \end{matrix}$$

En este caso, la matriz de coeficientes es la matriz cuadrada de orden n

$$A=\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

y si además $R(A)=n$, por el teorema IV. 4 existe su inversa A^{-1}

El sistema puede escribirse como

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

Premultiplicando ambos miembros por A^{-1}

$$A^{-1}(A\bar{x})= A^{-1}\bar{b}$$
$$(A^{-1}A)\bar{x}= A^{-1}\bar{b}$$
$$I\bar{x}= A^{-1}\bar{b}$$
$$\bar{x}= A^{-1}\bar{b}$$

Entonces, la matriz columna $\bar{x}$ que satisface la ecuación $A\bar{x}=\bar{b}$ puede ser obtenida con sólo efectuar el producto $A^{-1}\bar{b}$.

Puesto que el producto $A^{-1}\bar{b}$ define una única solución $\bar{x}$, podemos concluir que

> Un sistema de n ecuaciones lineales con n incógnitas, tal que la matriz de coeficientes tiene rango n, es compatible determinado y su solución puede obtenerse como:
>
> $$\bar{x}=A^{-1}\bar{b}$$

Ejemplo IV. 20.

Resolver el siguiente sistema de ecuaciones

$$3x_3-5+2x_2=-x_1$$
$$2x_1+4x_3= 11+x_2$$
$$-x_2+ x_3= 3$$

Solución

Ordenando el sistema tenemos

$$x_1+2x_2+3x_3=5$$
$$2x_1-x_2+4x_3=11$$
$$-x_2+ x_3=3$$

que en forma matricial puede expresarse como:

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

donde: $A=\begin{bmatrix}1 & 2 & 3\\ 2 & -1 & 4\\ 0 & -1 & 1\end{bmatrix}$, $\bar{x}=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$, $\bar{b}=\begin{bmatrix}5\\ 11\\ 3\end{bmatrix}$

Investiguemos si existe la inversa de A:

$$\underbrace{\begin{bmatrix}1 & 2 & 3\\ 2 & -1 & 4\\ 0 & -1 & 1\end{bmatrix}}_{A}\begin{matrix}\vdots\\ \vdots\\ \vdots\end{matrix}\underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}}_{I} \rightarrow \left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & -5 & -2 & -2 & 1 & 0\\ 0 & -1 & 1 & 0 & 0 & 1\end{array}\right] \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & 1 & \frac{2}{5} & \frac{2}{5} & -\frac{1}{5} & 0\\ 0 & -1 & 1 & 0 & 0 & 1\end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & 1 & \frac{2}{5} & \frac{2}{5} & -\frac{1}{5} & 0\\ 0 & 0 & \frac{7}{5} & \frac{2}{5} & -\frac{1}{5} & 1\end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & 1 & \frac{2}{5} & \frac{2}{5} & -\frac{1}{5} & 0\\ 0 & 0 & 1 & \frac{2}{7} & -\frac{1}{7} & \frac{5}{7}\end{array}\right]$$

En este paso, vemos que $R(A)=3$, por lo que existe A^{-1}

Continuando con el proceso

$$\left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & 1 & 0 & \frac{2}{7} & -\frac{1}{7} & -\frac{2}{7}\\ 0 & 0 & 1 & \frac{2}{7} & -\frac{1}{7} & \frac{5}{7}\end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc}1 & 2 & 0 & \frac{1}{7} & \frac{3}{7} & -\frac{15}{7}\\ 0 & 1 & 0 & \frac{2}{7} & -\frac{1}{7} & -\frac{2}{7}\\ 0 & 0 & 1 & \frac{2}{7} & -\frac{1}{7} & \frac{5}{7}\end{array}\right] \rightarrow \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}}_{I}\begin{matrix}\vdots\\ \vdots\\ \vdots\end{matrix}\underbrace{\begin{bmatrix}-\frac{3}{7} & \frac{5}{7} & -\frac{11}{7}\\ \frac{2}{7} & -\frac{1}{7} & -\frac{2}{7}\\ \frac{2}{7} & -\frac{1}{7} & \frac{5}{7}\end{bmatrix}}_{A^{-1}}$$

Entonces

$$A^{-1}=\begin{bmatrix}-\frac{3}{7} & \frac{5}{7} & -\frac{11}{7}\\ \frac{2}{7} & -\frac{1}{7} & -\frac{2}{7}\\ \frac{2}{7} & -\frac{1}{7} & \frac{5}{7}\end{bmatrix} = -\frac{1}{7}\begin{bmatrix}3 & -5 & 11\\ -2 & 1 & 2\\ -2 & 1 & -5\end{bmatrix}$$

En consecuencia, la solución es $\bar{x}=A^{-1}\bar{b}$

de donde

$$\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = -\frac{1}{7}\begin{bmatrix}3 & -5 & 11\\ -2 & 1 & 2\\ -2 & 1 & -5\end{bmatrix}\begin{bmatrix}5\\ 11\\ 3\end{bmatrix} = -\frac{1}{7}\begin{bmatrix}-7\\ 7\\ -14\end{bmatrix} = \begin{bmatrix}1\\ -1\\ 2\end{bmatrix}$$

o sea $x_1=1$, $x_2=-1$, $x_3=2$

- -

Veamos ahora cuales son los rangos de A y $A,\bar{b}$ para el sistema del ejemplo anterior:

$$\overbrace{\left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 2 & -1 & 4 & 11 \\ 0 & -1 & 1 & 3 \end{array}\right]}^{A}_{(A,\bar{b})} \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 0 & -5 & -2 & 1 \\ 0 & -1 & 1 & 3 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 0 & -1 & 1 & 3 \\ 0 & -5 & -2 & 1 \end{array}\right] \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & -7 & -14 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

De la última matriz vemos que

$R(A)=3$

$R(A,\bar{b})=3$

es decir $R(A)=R(A,\bar{b})$

Además, $R(A)=R(A,\bar{b})=n$, que es una característica de los sistemas compatibles determinados.

<u>Teorema</u> IV. 7.

Si en un sistema de m ecuaciones lineales con n incógnitas

$R(A)=R(A,\bar{b})=n$

entonces el sistema es compatible determinado.

Demostración.

Si $R(A)=R(A,\bar{b})=n$, la matriz ampliada del sistema (11) puede transformarse en la matriz

$$\begin{array}{c} n \\ \text{renglones} \end{array}\left\{\begin{array}{l} \\ \\ \\ \\ \\ \end{array}\right. \left[\begin{array}{ccccc} 1 & c_{12} & \cdots\cdots & c_{1n} & d_1 \\ 0 & 1 & \cdots\cdots & c_{2n} & d_2 \\ \cdots & \cdots & \cdots\cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots\cdots & \cdots & \cdots \\ 0 & 0 & \cdots\cdots & 1 & d_n \\ 0 & 0 & \cdots\cdots & 0 & 0 \\ \cdots & \cdots & \cdots\cdots & \cdots & \cdots \\ 0 & 0 & \cdots\cdots & 0 & 0 \end{array}\right]$$

que representa al sistema equivalente

$$\begin{aligned} x_1+c_{12}x_2+\ \dots\ +c_{1n}x_n&=d_1 \\ x_2+\ \dots\ +c_{2n}x_n&=d_2 \\ \dots\dots\dots\dots& \\ \dots\dots\dots\dots& \\ x_n&=d_n \end{aligned}$$

ya que las ecuaciones representadas por los renglones $n+1, \dots, m$ son todas de la forma

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

y se satisfacen para cualquier conjunto de valores $x_1, x_2, \dots, x_n$.

El sistema equivalente tiene n ecuaciones con n incógnitas, y el rango de su matriz de coeficientes es también n, por lo que es compatible determinado.

Ejemplo IV. 21.

Resolver el sistema

$3x_1+2x_2=6$

$x_1-\ x_2=1$

$4x_1+\ x_2=7$

Solución

Calculemos los rangos de A y $(A,\bar{b})$

$$\left[\begin{array}{cc|c} 3 & 2 & 6 \\ 1 & -1 & 1 \\ 4 & 1 & 7 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{2}{3} & 2 \\ 1 & -1 & 1 \\ 4 & 1 & 7 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{2}{3} & 2 \\ 0 & -\frac{5}{3} & -1 \\ 0 & -\frac{5}{3} & -1 \end{array}\right] \rightarrow$$

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

vemos que $R(A)=R(A,\bar{b})=n=2$, por lo que el sistema es compatible determinado.

Resolviendo el sistema equivalente

$$x_1+\frac{2}{3}x_2=2$$
$$x_2=\frac{3}{5}$$

obtenemos la solución

$$x_1=\frac{8}{5}$$
$$x_2=\frac{3}{5}$$

- -

<u>Sistemas compatibles indeterminados.</u>

Volvamos nuevamente al sistema

$$\left.\begin{array}{l} 20x_1+100x_2+40x_3=400 \\ 0x_1+\ \ x_2-\ \ x_3=\ 0 \end{array}\right\} \ldots\ldots(4)$$

para el que, al inicio del capítulo, dimos las soluciones

a) $x_1=6,\quad x_2=2,\quad x_3=2$

y b) $x_1=13,\quad x_2=1,\quad x_3=1$

Veamos ahora cuales son los rangos de A y $(A,\bar{b})$:

Dividiendo entre 20 el primer renglón de la matriz ampliada

$$\left[\begin{array}{ccc|c} 20 & 100 & 40 & 400 \\ 0 & 1 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 5 & 2 & 20 \\ 0 & 1 & -1 & 0 \end{array}\right]$$

En consecuencia

$R(A)=R(A,\bar{b})=2$

si en el sistema equivalente

$$x_1+5x_2+2x_3=20$$
$$x_2-x_3=0$$

damos a x_3 el valor de $x_3=2$, tendremos

$$x_1+5x_2+4=20$$
$$x_2-2=0$$

De la segunda ecuación

$$x_2=2$$

y, sustituyendo en la primera ecuación

$$x_1=6$$

con lo que hemos obtenido la solución a).

Si en el sistema equivalente hacemos ahora $x_3=1$, obtenemos la solución b).

En general, haciendo $x_3=k$ (una constante arbitraria) en el sistema equivalente, obtenemos

$$x_1+5x_2+2k=20$$
$$x_2-k=0$$

De la segunda ecuación

$$x_2=k$$

y, sustituyendo en la primera

$$x_1+5k+2k=20 \Longrightarrow x_1=20-7k.$$

con lo que podemos dejar la solución en función de k como

$$\left.\begin{array}{l} x_1=20-7k \\ x_2=k \\ x_3=k \end{array}\right\} \ldots\ldots(16)$$

A la expresión (16) le llamaremos solución general del sistema (4), ya que, para cualquier valor de k, los valores de x_1, x_2, x_3 dados por dicha expresión constituyen una solución del sistema.

Recordemos que el sistema (4) nos representa el problema de encontrar el número (x_1, x_2, x_3) de objetos de diferentes tipos (A, B, C) que se pueden fabricar, por lo que la solución está restringida a valores enteros no negativos (0, 1, 2 ...) de dichas incógnitas. Entonces, de (16), es fácil observar que k puede tomar únicamente los valores 0, 1 y 2 ($k \geq 3$ haría x_1 negativo).

Tenemos por tanto un sistema indeterminado con tres soluciones (para $k=0, 1, 2$).

Sin embargo, si consideramos un problema en el que las incógnitas x_1, x_2 y x_3 pueden tomar cualquier valor, tendremos un sistema indeterminado con un número infinito de soluciones (una para cada valor real de k).

Considerando nuevamente los rangos de A y $(A, \bar{b})$ y el número de incógnitas n, vemos que, para el sistema del ejemplo anterior

$$R(A) = R(A, \bar{b}) < n$$

que es una característica de los sistemas compatibles indeterminados.

Teorema IV. 6

Si en un sistema de m ecuaciones lineales con n incógnitas

$$R(A) = R(A, \bar{b}) = r$$

y $\quad r < n$

entonces el sistema es compatible indeterminado

Demostración

Si $R(A) = R(A, \bar{b}) = r$ y $r < n$

la matriz ampliada del sistema (11) puede reducirse a la forma (14), donde $c=0$.

Los primeros r renglones de dicha matriz son tales que su primer elemento distinto de cero es igual a uno y el número de ceros anteriores a dicho elemento aumenta de renglón a renglón.

Si llamamos $z_1, z_2, \ldots, z_r$ a las incógnitas cuyos coeficientes son los elementos mencionados, podemos ordenar el sistema equivalente en la forma

$$
\begin{aligned}
z_1 + p_{12}z_2 + \cdots + p_{1r}z_r &= d_1 - p_{1,r+1}z_{r+1} - \cdots - p_{1n}z_n \\
z_2 + \cdots + p_{2r}z_r &= d_2 - p_{2,r+1}z_{r+1} - \cdots - p_{2n}z_n \\
&\cdots\cdots\cdots\cdots \\
z_r &= d_r - p_{r,r+1}z_{r+1} - \cdots - p_{rn}z_n
\end{aligned}
$$

Si asignamos valores arbitrarios a las incógnitas $z_{r+1}, z_{r+2}, \ldots, z_n$ trasladadas a los segundos miembros, el sistema es compatible determinado en las incógnitas $z_1, z_2, \ldots, z_r$.

Ahora bien, como a $z_{r+1}, z_{r+2}, \ldots, z_n$ podemos asignarles un número infinito de valores arbitrarios, el sistema además de ser compatible es indeterminado.

Ejemplo IV. 22

Resolver el siguiente sistema de ecuaciones:

$$x_1 + x_2 + x_3 - 2x_4 - x_5 = 1$$

$$3x_1 + 3x_2 - x_3 + x_4 + 4x_5 = 4$$

$$x_1 + x_2 + 5x_3 - 9x_4 - 8x_5 = 0$$

Solución.

Calculamos primero los rangos de A y $(A, \bar{b})$

$$
\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & -1 & 1 \\ 3 & 3 & -1 & 1 & 4 & 4 \\ 1 & 1 & 5 & -9 & -8 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & -1 & 1 \\ 0 & 0 & -4 & 7 & 7 & 1 \\ 0 & 0 & 4 & -7 & -7 & -1 \end{array}\right] \rightarrow
$$

$$
\rightarrow \left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & -1 & 1 \\ 0 & 0 & -4 & 7 & 7 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & -1 & 1 \\ 0 & 0 & 1 & -\frac{7}{4} & -\frac{7}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]
$$

Como $R(A)=R(A,\ \bar{b})=2=r$ el sistema es compatible.

Como $r < n$, $(2 < 5)$, el sistema es indeterminado.

Tenemos ahora el sistema equivalente

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

Las dos incógnitas que dejamos en el sistema equivalente son x_1 y x_3, ya que sus coeficientes son los primeros elementos distintos de cero en los renglones de la última matriz. En consecuencia, trasladamos las tres incógnitas restantes a los segundos miembros de las ecuaciones.

Asignando a x_2, x_4, x_5 los valores $x_2=k_1$, $x_4=k_2$ y $x_5=k_3$ tenemos el sistema

$$
\begin{aligned} x_1+x_3 &= 1-k_1+2k_2+k_3 \\ x_3 &= -\tfrac{1}{4}+\tfrac{7}{4}k_2+\tfrac{7}{4}k_3 \end{aligned}
$$

Sustituyendo x_3 en la primera ecuación obtenemos finalmente

$$
\begin{aligned} x_1 &= \tfrac{1}{4}(5-4k_1+k_2-3k_3) \\ x_2 &= k_1 \\ x_3 &= \tfrac{1}{4}(-1+7k_2+7k_3) \\ x_4 &= k_2 \\ x_5 &= k_3 \end{aligned}
$$

que es la solución general del sistema.

Si por ejemplo hacemos

$$k_1=1,\quad k_2=1,\quad k_3=0$$

obtenemos la solución particular.

$$x_1=\tfrac{1}{2},\ x_2=1,\ x_3=\tfrac{3}{2},\ x_4=1,\ x_5=0$$

Otra solución particular sería

$$x_1=\tfrac{5}{4},\ x_2=0,\ x_3=-\tfrac{1}{4},\ x_4=0,\ x_5=0$$

que se obtuvo para los valores

$$k_1=k_2=k_3=0$$

- -

Los resultados obtenidos en esta sección, pueden resumirse en el siguiente cuadro

$$
\text{Sistemas de ecuaciones lineales } A\bar{x}=\bar{b}\ \begin{cases} \text{Compatibles } R(A)=R(A,\bar{b})=r & \begin{cases} \text{Determinados } r=n \\ \text{Indeterminados } r<n \end{cases} \\ \text{Incompatibles } R(A)<R(A,\bar{b}) \end{cases}
$$

Las condiciones que aparecen en el cuadro, fueron enunciadas en los teoremas IV. 6, IV. 7 y IV. 8 como condiciones suficientes. No es difícil probar que dichas condiciones son también necesarias.

De los ejemplos utilizados en el desarrollo de esta sección, se ve que un proceso conveniente para obtener soluciones de sistemas de ecuaciones lineales es el siguiente:

1) Se obtienen los rangos de A y $(A,\ \bar{b})$ trabajando directamente con la matriz ampliada.
2) Se clasifica el sistema en base al cuadro anterior.
3) De ser compatible, se trabaja con el sistema equivalente representado por la matriz escalonada que se obtuvo al calcular los rangos.
4) Se trasladan $n-r$ incógnitas a los segundos miembros de las --

ecuaciones, de tal forma que se obtenga un sistema compatible determinado en r incógnitas.

5) Se resuelve el sistema obtenido en 4 por cualquier método (matriz inversa, sustitución, etc.)

Ejemplo IV. 23.

Resolver el siguiente sistema

$$-4x_1+x_2-4x_3-10x_4+22x_5+4x_6=-9$$
$$x_1+2x_3+4x_4-5x_5-x_6=2$$
$$-2x_1+2x_2+4x_3+5x_4+13x_5+4x_6=-3$$
$$3x_1-2x_2-2x_3+3x_4-21x_5+3x_6=15$$
$$5x_1+10x_3+20x_4-25x_5-5x_6=10$$
$$x_1-x_2-2x_3-x_4-6x_5+x_6=2$$

obtengamos los rangos:

$$\left[\begin{array}{cccccc|c} -4 & 1 & -4 & -10 & 22 & 4 & -9 \\ 1 & 0 & 2 & 4 & -5 & -1 & 2 \\ -2 & 2 & 4 & 5 & 13 & 4 & -3 \\ 3 & -2 & -2 & 3 & -21 & 3 & 15 \\ 5 & 0 & 10 & 20 & -25 & -5 & 10 \\ 1 & -1 & -2 & -1 & -6 & 1 & 2 \end{array}\right] \to \left[\begin{array}{cccccc|c} 1 & 0 & 2 & 4 & -5 & -1 & 2 \\ -4 & 1 & -4 & -10 & 22 & 4 & -9 \\ -2 & 2 & 4 & 5 & 13 & 4 & -3 \\ 3 & -2 & -2 & 3 & -21 & 3 & 15 \\ 5 & 0 & 10 & 20 & -25 & -5 & 10 \\ 1 & -1 & -2 & -1 & -6 & 1 & 2 \end{array}\right] \to$$

$$\to \left[\begin{array}{cccccc|c} 1 & 0 & 2 & 4 & -5 & -1 & 2 \\ 0 & 1 & 4 & 6 & 2 & 0 & -1 \\ 0 & 2 & 8 & 13 & 3 & 2 & 1 \\ 0 & -2 & -8 & -9 & -6 & 6 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -4 & -5 & -1 & 2 & 0 \end{array}\right] \to \left[\begin{array}{cccccc|c} 1 & 0 & 2 & 4 & -5 & -1 & 2 \\ 0 & 1 & 4 & 6 & 2 & 0 & -1 \\ 0 & 2 & 8 & 13 & 3 & 2 & 1 \\ 0 & -2 & -8 & -9 & -6 & 6 & 9 \\ 0 & -1 & -4 & -5 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \to$$

$$\to \left[\begin{array}{cccccc|c} 1 & 0 & 2 & 4 & -5 & -1 & 2 \\ 0 & 1 & 4 & 6 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 2 & 3 \\ 0 & 0 & 0 & 3 & -2 & 6 & 7 \\ 0 & 0 & 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cccccc|c} 1 & 0 & 2 & 4 & -5 & -1 & 2 \\ 0 & 1 & 4 & 6 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 2 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \to$$

$$\to \left[\begin{array}{cccccc|c} 1 & 0 & 2 & 4 & -5 & -1 & 2 \\ 0 & 1 & 4 & 6 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

De esta matriz podemos observar que

$R(A)=R(A,\bar{b})=4$ (sistema compatible).

Tenemos entonces $r=4$

y $n=6$ (número de incógnitas), por lo que el sistema en cuestión es compatible indeterminado.

El sistema equivalente que representa la última matriz es:

$$x_1+2x_3+4x_4-5x_5-x_6=2$$
$$x_2+4x_3+6x_4+2x_5=-1$$
$$x_4-x_5+2x_6=3$$
$$x_5=-2$$

Como se dijo anteriormente, debemos despejar $n-r=2$ incógnitas, dejando del lado izquierdo aquellas cuyo coeficiente es el primer uno de cada renglón de la matriz escalonada (x_1, x_2, x_4, x_5):

$$x_1+4x_4-5x_5=2-2x_3+x_6$$
$$x_2+6x_4+2x_5=-1-4x_3$$
$$x_4-x_5=3-2x_6$$
$$x_5=-2$$

Dando valores arbitrarios a x_3 y x_6;

$x_3 = k_1$

$x_6 = k_2$

obtenemos el sistema

$$\begin{aligned} x_1 \quad + 4x_4 - 5x_5 &= 2 - 2k_1 + k_2 \\ x_2 + 6x_4 + 2x_5 &= -1 - 4k_1 \\ x_4 - x_5 &= 3 - 2k_2 \\ x_5 &= -2 \end{aligned}$$

Este sistema es de la forma

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

donde

$$A = \begin{bmatrix} 1 & 0 & 4 & -5 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}_{4\times 4} \qquad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix}_{4\times 1} \qquad \bar{b} = \begin{bmatrix} 2 - 2k_1 + k_2 \\ -1 - 4k_1 \\ 3 - 2k_2 \\ -2 \end{bmatrix}_{4\times 1}$$

y podemos resolverlo utilizando la matriz inversa de A, como

$\bar{x} = A^{-1}\bar{b}$

Como es fácil verificar:

$$A^{-1} = \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 1 & -6 & -8 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Por tanto

$$\begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 1 & -6 & -8 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 - 2k_1 + k_2 \\ -1 - 4k_1 \\ 3 - 2k_2 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 - 2k_1 + k_2 - 4(3 - 2k_2) - 2 \\ -1 - 4k_1 - 6(3 - 2k_2) - 8(-2) \\ 3 - 2k_2 - 2 \\ -2 \end{bmatrix}$$

finalmente

$$\begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -12 - 2k_1 + 9k_2 \\ -3 - 4k_1 + 12k_2 \\ 1 - 2k_2 \\ -2 \end{bmatrix}$$

Entonces, la solución general del sistema original es:

$x_1 = -12 - 2k_1 + 9k_2$

$x_2 = -3 - 4k_1 + 12k_2$

$x_3 = k_1$

$x_4 = 1 - 2k_2$

$x_5 = -2$

$x_6 = k_2$

$\forall\ k_1,\ k_2 \in R.$

- -

Ejemplo IV. 24.

Sea el sistema

$$\begin{aligned} x + y + z &= 6 \\ x - 2z &= -4 \\ 3x + 2y &= 8 \\ 5x + 2y + \beta z &= 0 \end{aligned}$$

¿Qué valores puede tomar β para que la solución del sistema sea única? Dando a β uno de esos valores resuélvase el sistema.

Solución.

La solución del sistema es única para aquellos valores de β que hagan $R(A) = R(A, \bar{b}) = n$

Reduciendo la matriz ampliada a su forma escalonada:

$$\begin{bmatrix} 1 & 1 & 1 & | & 6 \\ 1 & 0 & -2 & | & -4 \\ 3 & 2 & 0 & | & 8 \\ 5 & 2 & \beta & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & | & 6 \\ 0 & -1 & -3 & | & -10 \\ 0 & -1 & -3 & | & -10 \\ 0 & -3 & -5+\beta & | & -30 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & | & 6 \\ 0 & 1 & 3 & | & 10 \\ 0 & 0 & 0 & | & 0 \\ 0 & -3 & -5+\beta & | & -30 \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 1 & | & 6 \\ 0 & 1 & 3 & | & 10 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 4+\beta & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & | & 6 \\ 0 & 1 & 3 & | & 10 \\ 0 & 0 & 4+\beta & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

De la última matriz vemos que

Si $4+\beta=0$, $R(A)=R(A,\bar{b})=2$ (compatible indeterminado)

Si $4+\beta\neq 0$, $R(A)=R(A,\bar{b})=3$ (compatible determinado)

por tanto, el sistema tiene solución única $\forall\ \beta \in R,\ \beta \neq -4$.

Si $\beta\neq -4$, de la última matriz se tiene el sistema equivalente

$$x+y+z=6$$
$$y+3z=10$$
$$(4+\beta)z=0$$

cuya solución es

$$x=-4,\quad y=10,\quad z=0$$

- -

Sistemas homogéneos.

A los sistemas de ecuaciones lineales de la forma

$$\begin{array}{l} a_{11}\ x_1 + a_{12}\ x_2 + \dots + a_{1n}\ x_n = 0 \\ a_{21}\ x_1 + a_{22}\ x_2 + \dots + a_{2n}\ x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}\ x_1 + a_{m2}\ x_2 + \dots + a_{mn}\ x_n = 0 \end{array}$$

Se les llama sistemas homogéneos.

La representación matricial de un sistema homogéneo es

$A\bar{x}=0$ (donde 0 es la matriz nula de orden $m\times 1$)

Estos sistemas son siempre compatibles puesto que admiten la solución

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

llamada solución trivial. Además, en un sistema homogéneo se tiene siempre que

$$R(A)=R(A,\bar{b})$$

puesto que agregando una columna de ceros no puede elevarse el rango de una matriz.

Si $R(A)=n$, el sistema solo admite la solución trivial ya que la solución de la ecuación

$$A\bar{x}=0$$

es $\quad x=A^{-1}0=0.$

Si $R(A)< n$, el sistema es indeterminado y admite otras soluciones además de la trivial, las cuales pueden obtenerse mediante cualquiera de los procedimientos vistos anteriormente.

IV. 6. DETERMINANTES.

Consideremos el sistema de ecuaciones

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2=b_1 \\ a_{21}x_1+a_{22}x_2=b_2 \end{array}\right\} \ldots (17)$$

cuya matriz de coeficientes es

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

Para eliminar x_2 por sustracción multipliquemos la primera ecuación por a_{22} y la segunda por a_{12}, con lo que resulta el sistema equivalente.

$$a_{11}a_{22}x_1+a_{12}a_{22}x_2=a_{22}b_1$$

$$a_{12}a_{21}x_1+a_{12}a_{22}x_2=a_{12}b_2$$

restando la segunda ecuación de la primera

$$(a_{11}a_{22}-a_{12}a_{21})x_1=a_{22}b_1-a_{12}b_2$$

si $a_{11}a_{22}-a_{12}a_{21}\neq 0$

$$x_1=\frac{b_1a_{22}-a_{12}b_2}{a_{11}a_{22}-a_{12}a_{21}} \quad \ldots (18)$$

En forma similar podemos obtener

$$x_2=\frac{a_{11}b_2-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}} \quad \ldots (19)$$

Sustituyendo en las ecuaciones (17), los valores de x_1 y x_2 dados por las expresiones (18) y (19), puede comprobarse fácilmente que se trata de una solución.

El común denominador de las expresiones (18) y (19) está expresado en términos de los elementos de la matriz A de coeficientes. A este número se le conoce como determinante de la matriz A y se le representa con det (A). Se dice que es un determinante de segundo orden, por ser la matriz de orden 2.

Para designar el determinante de la matriz A se emplea la siguiente notación

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21} \ldots\ldots (20)$$

donde hemos reemplazado los paréntesis rectangulares por barras verticales para distinguir al determinante de la matriz. Es importante subrayar que, mientras que la matriz es un arreglo de números, el determinante es un número perfectamente definido por la matriz cuadrada.

Por ejemplo

$$\begin{vmatrix} -2 & 3 \\ -1 & 5 \end{vmatrix}=(-2)(5)-(3)(-1)=-10+3=-7$$

Los numeradores de las expresiones (18) y (19) tienen la misma forma que el denominador, o sea, también son determinantes de segundo orden.

De acuerdo con la expresión (20) que define al determinante de segundo orden, podemos escribir

$$x_1=\frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

$$x_2=\frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

El numerador de la primera expresión es el determinante de una matriz que se obtiene a partir de la matriz A, sustituyendo la columna de coeficientes de x_1 por el vector de términos independientes. El numerador de la segunda expresión es el determinante de otra matriz, que se obtiene también a partir de A, reemplazando ahora la columna de coeficientes de x_2 por el vector de términos independientes.

A la regla sugerida por estas expresiones para resolver un sistema de ecuaciones se le conoce como regla de Cramer y obviamente es aplicable siempre y cuando det (A) ≠ 0

Consideremos ahora el sistema de 3 ecuaciones con 3 incógnitas

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2 \\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3 \end{array}\right\} \ldots\ldots \quad (21)$$

cuya matriz de coeficientes es

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

Mediante un proceso similar al empleado para el sistema (17), encontramos que los valores de las incógnitas que satisfacen el sistema son

$$x_1=\frac{b_1a_{22}a_{33}+a_{13}b_2a_{32}+a_{12}a_{23}b_3-a_{13}a_{22}b_3-b_1a_{23}a_{32}-a_{12}b_2a_{33}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}} \quad (22)$$

$$x_2=\frac{a_{11}b_2a_{33}+a_{13}a_{21}b_3+b_1a_{23}a_{31}-a_{13}b_2a_{31}-a_{11}a_{23}b_3-b_1a_{21}a_{33}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}} \quad (23)$$

$$x_3=\frac{a_{11}a_{22}b_3+b_1a_{21}a_{32}+a_{12}b_2a_{31}-b_1a_{22}a_{31}-a_{11}b_2a_{32}-a_{12}a_{21}b_3}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}} \quad (24)$$

Al común denominador de las expresiones (22), (23) y (24) se le conoce como determinante de la matriz A y es un determinante de tercer orden. La expresión que define al determinante de tercer orden es;

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}. \quad (25)$$

Por ejemplo

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

De acuerdo con la expresión (25), podemos escribir la solución del sistema como sigue

$$x_1=\frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

$$x_2=\frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

$$x_3=\frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

con lo que resulta lógico tratar de generalizar la regla de Cramer para el caso de un sistema de n ecuaciones con n incógnitas. Sin embargo, requerimos de una definición para el determinante de orden n.

Definición de determinante.

Queremos generalizar la definición de determinante para un n arbitrario, en base a las expresiones (20) y (25) que definen a los determinantes de orden 2 y 3, respectivamente.

Sin embargo, no es posible hacer esto del mismo modo (es decir, resolviendo en forma general un sistema de ecuaciones lineales), pues a medida que aumenta n, los cálculos se hacen más complicados y, siendo n arbitrario, son prácticamente irrealizables.

Estableceremos una ley general examinando los determinantes de segundo y tercer orden, y tomaremos esta como definición para el determinante de orden n. Antes, definiremos los conceptos de permutación y clase de una permutación que necesitaremos para ello.

Las permutaciones de los n números del conjunto {1,2,3, ...,n} son las diferentes maneras en que pueden ser arreglados.

Por ejemplo, las permutaciones de los números 1,2,3, son los arreglos

1, 3, 2
2, 3, 1
3, 2, 1
1, 2, 3
2, 1, 3
3, 1, 2

El conjunto de todas las permutaciones de n números, está formado por n! arreglos o permutaciones y se representa por S_n.

Llamaremos permutación principal a aquella en la que los números aparecen en el orden natural.

En el ejemplo anterior tenemos 3!=6 permutaciones, donde la permutación principal es 1, 2, 3.

Consideremos una permutación arbitraria $\alpha_1, \alpha_2, \ldots, \alpha_n$ de los números 1, 2, ..., n. Diremos que dicha permutación es de clase par o impar según exista un número par o impar de inversiones en el orden natural, es decir; parejas (α_1, α_k) tales que α_1 precede a α_k en la permutación y $\alpha_k < \alpha_1$. A la permutación principal le asignaremos la clase par.

Para las permutaciones del ejemplo tenemos que:

1, 3, 2 es de clase impar ya que existe una pareja, la pareja (3, 2), tal que 3 precede a 2 en la permutación y 2<3.

2, 3, 1 es de clase par ya que existen dos parejas, (2,1) y (3,1), tales que 2 precede a 1 y 1<2; 3 precede a 1 y 1<3

3, 2, 1 es de clase impar ya que existen tres parejas, (3,2), (3,1) y (2,1), tales que 3 precede a 2 y 2<3; 3 precede a 1 y 1<3; 2 precede a 1 y 1<2.

1, 2, 3 es de clase par (por definición).

Procediendo en forma análoga vemos que:

2, 1, 3 es de clase impar.

3, 1, 2 es de clase par.

Recordando la expresión que define al determinante de 2o. orden

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

observamos que:

a) El determinante es la suma algebraica de dos (2!) productos.

b) Cada producto tiene dos factores.

c) En cada producto hay elementos de todos los renglones (y uno sólo de cada renglón).

d) En cada producto hay elementos de todas las columnas (y uno sólo de cada columna).

e) Si los elementos se escriben de tal manera que los primeros índices formen una permutación principal, se antepone un signo + al producto en que los segundos índices forman una permutación de clase par y un signo - al producto en que los segundos índices forman una permutación de clase impar.

Recordando la expresión que define al determinante de 3er. orden

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \qquad \ldots\ldots (25)$$

observamos que:

a) El determinante es la suma algebraica de seis (3!) productos.

b) Cada producto tiene 3 factores

c) En cada producto hay elementos de todos los renglones (y uno sólo de cada renglón).

d) En cada producto hay elementos de todas las columnas (y uno sólo en cada columna).

e) Si los elementos se escriben de tal manera que los primeros índices formen una permutación principal, se antepone un signo + a los productos en que los segundos índices forman una permutación de clase par y un signo - a los productos en que los segundos índices forman una permutación de clase impar:

Generalizando, para el determinante de orden n se tendrá:

a) El determinante es la suma algebraica de n! productos.

b) Cada producto tiene n factores.

c) En cada producto hay elementos de todos renglones (y uno sólo de cada renglón).

d) En cada producto hay elementos de todas las columnas (y uno sólo de cada columna).

e) Si los elementos se escriben de tal manera que los primeros índices formen una permutación principal, se antepone un signo + a los productos en que los segundos índices forman una permutación de clase par y un signo - a los productos en que los segundos índices forman una permutación de clase impar.

Establezcamos ahora la ley que define al determinante de orden n:

Consideremos la matriz cuadrada de orden n

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

y un producto de n de sus elementos

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

tomados de tal manera que haya elementos de todos sus renglones (y uno sólo de cada renglón), y que haya elementos de todas sus columnas (y uno sólo de cada columna).

Los factores de este producto están ordenados de tal modo que los primeros índices forman una permutación principal. La sucesión de los segundos índices forman una permutación.

$\alpha_1, \alpha_2, \ldots, \alpha_n$

de los números 1, 2, ... n. Para esta permutación definiremos

$\varepsilon=+1$ si la permutación es de clase par

$\varepsilon=-1$ si la permutación es de clase impar

Formemos el producto provisto de signo

$$\varepsilon\, a_{1\alpha_1}\ a_{1\alpha_2}\ \ldots\ a_{n\alpha_n} \qquad (26)$$

Como el conjunto S_n de todas las permutaciones de n números está formada por n! arreglos, podemos formar n! productos del tipo (26).

El determinante de la matriz A se define como la suma de estos n! productos dotados de signo (a los cuales se les llama términos del determinante).

Definición.

$$\det(A)=\sum_{S_n} \varepsilon\ a_{1\alpha_1}\ a_{2\alpha_2}\ \ldots\ a_{n\alpha_n} \quad \ldots\ldots(27)$$

De la definición anterior se concluye que:

1) Para obtener todos los términos del desarrollo de un determinante, basta con escribir el término principal (multiplicando los elementos sobre la diagonal principal) y a partir de éste obtener todos los demás dejando fijos los primeros índices y permutando de todas las maneras posibles los segundos índices.

2) Como de los segundos índices hay n! permutaciones, la mitad de clase par y la mitad de clase impar, habrá n! términos en el desarrollo del determinante, la mitad con signo + y la mitad con signo -.

3) Los signos + y - se asignan según la permutación de los segundos índices sea de clase par o impar.

Ejemplo IV. 25.

Obtener el desarrollo del determinante de 2o. orden a partir de la definición

$$\det(A)=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

De acuerdo con lo anterior, el término principal será

$a_{11}\ a_{22}$

Dejamos fijos los primeros índices y permutamos los segundos de todas las maneras posibles:

12 clase par + +

21 clase impar + -

El desarrollo será entonces:

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

- -

Ejemplo IV. 26.

Obtener el desarrollo del determinante de 3er. orden, a partir de la definición

$$\det(A)=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Término principal: $a_{11}\ a_{22}\ a_{33}$.

Dejando fijos los primeros índices, las permutaciones de los segundos índices son:

1 2 3	clase par	+	+
1 3 2	clase impar	+	-
2 1 3	clase impar	+	-
2 3 1	clase par	+	+
3 1 2	clase par	+	+
3 2 1	clase impar	+	-

el desarrollo será:

$$\det(A)=a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-$$
$$-a_{13}a_{22}a_{31}$$

\- -

<u>Cálculo de determinantes.</u>

Si (M) representa al conjunto de las matrices cuadradas de orden n con elementos en C, podemos definir la función

$$\det: (M) \to C$$

que asigna a la matriz $A \in (M)$ el escalar específico $\det(A) \in C$.

Tal función queda definida por la expresión (27) que representa una suma de $n!$ productos de elementos de A.

El empleo de dicha expresión para el cálculo de determinantes no se acostumbra en la práctica por resultar demasiado laborioso, a cambio se han desarrollado métodos más sencillos que conducen a los mismos resultados. Tres de dichos métodos trataremos en esta sección.

Regla de Sarrus.

La regla de Sarrus indica que para obtener el valor de un determinante de 2o. orden, al producto de los elementos de la diagonal principal (líneas llenas) se resta el producto de los elementos de la "diagonal secundaria" (líneas punteadas). Así tendremos que

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}$$

que coincide con el desarrollo según la definición.

por ejemplo

$$\begin{vmatrix} 3 & -5 \\ 4 & -3 \end{vmatrix}= (3)(-3)-(-5)(4)=-9+20=11$$

La regla de Sarrus indica que para obtener el valor de un determinante de 3er. orden, a los productos de los elementos de la diagonal principal y paralelas a ella (líneas llenas) se restan los productos de los elementos de la diagonal secundaria y paralelas a ella (líneas punteadas). Para observar mejor éstos productos, se pueden escribir el primero y segundo renglón inmediatamente después del tercero. Así tenemos que:

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

que coincide con el desarrollo según la definición.

Por ejemplo:

$$\begin{vmatrix} 3 & 2 & 4 \\ 1 & 5 & 2 \\ -3 & -5 & 4 \end{vmatrix}= (3)(5)(4)+(1)(-5)(4)+(-3)(2)(2)-(4)(5)(-3)-$$
$$-(2)(1)(4)-(3)(-5)(2)=110$$

Vale la pena subrayar que la regla de Sarrus se ha enunciado exclusivamente para los determinantes de 2o. y 3er. orden.

Es frecuente tender a generalizar esta regla para calcular determinantes de orden mayor, sin embargo; el estudiante puede fácilmente demostrar que al aplicar la regla de Sarrus a un determinante de orden superior al tercero, se obtiene un desarrollo que no coincide con la definición. El método que veremos a continuación es aplicable a un determinante de cualquier orden.

Desarrollo por cofactores.

Volviendo al ejemplo IV. 26, el resultado obtenido para el determinante de 3er. orden según la definición es:

$$\det(A)=a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+$$
$$+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}$$

factorizando los elementos del primer renglón tenemos:

$$\det(A)=a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{12}(a_{21}a_{33}-a_{23}a_{31})$$
$$+a_{13}(a_{21}a_{32}-a_{22}a_{31})$$

de acuerdo con la definición de determinante de 2o. orden podemos escribir:

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

para determinar los signos de los términos hacemos:

$$\det(A)=(-1)^{1+1}a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}+(-1)^{1+2}a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}+$$
$$+(-1)^{1+3}a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \quad \ldots\ldots(27)$$

donde los exponentes de (-1) son la suma de los índices del elemento del primer renglón.

A la expresión (27) se le llama desarrollo por cofactores del determinante respecto a su primer renglón.

Se recomienda al estudiante hacer el desarrollo del mismo determinante respecto a alguno de los renglones o columnas restantes. Es obvio que el resultado de ambos desarrollos es numéricamente igual a det(A).

En la expresión (27) puede observarse que, los determinantes que multiplican a los elementos del primer renglón, pueden obtenerse eliminando en el determinante original, el renglón y la columna donde se encuentra el elemento correspondiente.

Por ejemplo, el determinante que multiplica al elemento a_{12} puede obtenerse suprimiendo el primer renglón y la segunda columna del determinante original

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

a este determinante se le llama el menor de a_{12} y lo representaremos con M_{12}.

Generalizando, daremos la siguiente

> **Definición.**
>
> Se llama menor del elemento a_{ij} de un determinante de orden n, al determinante de orden n-1 que se obtiene al suprimir, en el determinante original, el renglón i y la columna j.

Al menor de a_{ij} lo representaremos con M_{ij}.

De acuerdo con esta notación, la expresión (27) puede escribirse como

$$\det(A)=(-1)^{1+1}a_{11}M_{11}+(-1)^{1+2}a_{12}M_{12}+(-1)^{1+3}a_{13}M_{13}$$

o bien, en forma condensada

$$\det(A)=\sum_{j=1}^{3}(-1)^{1+j}\,a_{1j}\,M_{1j}$$

Si elegimos el renglón k, para desarrollar por cofactores el determinante de 3er. orden, dicho desarrollo está expresado por

$$\det(A)=\sum_{j=1}^{3}(-1)^{k+j}a_{kj}\,M_{kj}$$

y si elegimos la columna k, por

$$\det(A)=\sum_{i=1}^{n}(-1)^{i+k}a_{ik}M_{ik}$$

En general. Si A es una matriz cuadrada de orden n:

$$\det(A)=\sum_{j=1}^{n}(-1)^{k+j}a_{kj}M_{kj} \quad \ldots\ldots \quad (28)$$

$$\det(A)=\sum_{i=1}^{n}(-1)^{i+k}a_{ik}M_{ik} \quad \ldots\ldots \quad (29)$$

donde $k \in N$ y $1<k<n$

Definición.

Llamamos cofactor del elemento a_{ij} al determinante $(-1)^{i+j}M_{ij}$, al que representamos con A_{ij}.

Con ayuda de esta definición, el método sugerido por las expresiones (28) y (29) (al que se llama "método de desarrollo por cofactores") puede enunciarse como sigue:

El valor de un determinante puede obtenerse efectuando la suma de los productos de los elementos de una cualquiera de sus líneas (renglón o columna) por sus respectivos cofactores.

Ejemplo IV. 27

Calcular el determinante de la matriz A por el método de desarrollo por cofactores.

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

Solución.

Primero, debemos elegir un renglón o una columna para efectuar el desarrollo. Analizando la matriz, vemos que es conveniente elegir alguna de las siguientes líneas:

2o. Renglón

3er. Renglón

1a. Columna

3a. Columna

ya que cada una de ellas tiene un elemento nulo y el producto de dicho elemento por su respectivo cofactor será igual a cero. Esto simplifica el trabajo al cálculo de sólo tres determinantes de 3er. orden.

Desarrollando por cofactores según el 2o. renglón:

$\det(A)=1A_{21}-3A_{22}+0A_{23}-6A_{24}$

Los cofactores A_{ij} se obtienen como

$A_{ij}=(-1)^{i+j}M_{ij}$

Desarrollando por la regla de Sarrus los menores M_{21}, M_{22} y M_{24} obtenemos:

$$M_{21}=\begin{vmatrix}1 & -5 & 1\\ 2 & -1 & 2\\ 4 & -7 & 6\end{vmatrix}=-6-40-14+4+60+14=18 \qquad \therefore\ A_{21}=-18$$

$$M_{22}=\begin{vmatrix}2 & -5 & 1\\ 0 & -1 & 2\\ 1 & -7 & 6\end{vmatrix}=-12-10+1+28=7 \qquad \therefore\ A_{22}=7$$

$$M_{24}=\begin{vmatrix}2 & 1 & -5\\ 0 & 2 & -1\\ 1 & 4 & -7\end{vmatrix}=-28-1+10+8=-11 \qquad \therefore\ A_{24}=-11$$

finalmente

$\det(A)=-18-3(7)-6(-11)=27$

. . . .

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 éstos 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, los cuales pueden resolverse empleando la regla de Sarrus.

Propiedades elementales.

Los determinantes tienen ciertas propiedades que es útil conocer. Son de interés principalmente las condiciones bajo las cuales un determinante es nulo; así como las transformaciones que, efectuadas en la matriz, no alteran el valor del determinante o le producen una alteración fácilmente calculable.

Estas propiedades, llamadas propiedades elementales, se demuestran a partir de la definición(2) y son las que a continuación se enlistan. En lo que sigue, por una línea deberá entenderse un renglón o una columna.

1) Si una línea está constituida por ceros, el determinante es nulo.
2) Si dos líneas paralelas son proporcionales, el determinante es nulo. (En particular, si dos líneas paralelas son iguales, el determinante es nulo).
3) Si se intercambian dos líneas paralelas cualesquiera, el determinante sólo cambia de signo.
4) Si se multiplican todos los elementos de una línea por un número k, el valor del determinante queda multiplicado por k.
5) El valor del determinante no cambia, si se intercambian renglones por columnas y viceversa. Es decir:

$$\det(A)=\det(A^T)$$

6) El valor de un determinante no cambia, si a los elementos de una de sus líneas, se suman los elementos correspondientes a otra línea paralela multiplicados por un número k.

(2) El estudiante puede consultar la demostración en la referencia 2 pags. 35 a 39.

Un método de condensación.

El método de desarrollo por cofactores, aunque aplicable a determinantes de cualquier orden, puede resultar en ocasiones demasiado laborioso. Si regresamos al ejemplo IV. 27, vemos que al elegir el 2o. renglón en lugar del 1o., nos hemos evitado el cálculo de un determinante de 3er. orden.

Si en una línea cualquiera de un determinante, todos los elementos excepto uno fueran nulos, sin duda escogeríamos dicha línea para efectuar el desarrollo. El método que propondremos a continuación se basa en esta idea y consiste en:

a) Elegir la línea que contenga el mayor número de ceros.
b) Aplicar reiteradamente la propiedad (6) hasta reducir a cero todos los elementos de dicha línea excepto uno. (Si todos los elementos se reducen a cero, el determinante es nulo por la propiedad (1)).
c) Desarrollar por cofactores según dicha línea.
d) Repetir el proceso hasta obtener un determinante de segundo o tercer orden.

Ejemplo IV. 28

Calcular el determinante de la matriz A empleando el método de condensación.

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

La adjunta es:

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

desarrollando los determinantes obtenemos finalmente:

$$A^* = \begin{bmatrix} -7 & 4 & -8 \\ 7 & -14 & -7 \\ -7 & -1 & 2 \end{bmatrix}$$

- -

Se puede demostrar que la inversa de una matriz A, puede obtenerse mediante la relación

$$A^{-1} = \frac{1}{\det(A)} A^* \quad \ldots\ldots(30)$$

de esta relación vemos que

$$\exists\, A^{-1} \Longleftrightarrow \det(A) \neq 0.$$

En consecuencia:

1) Si A es una matriz cuadrada de orden n, no singular, las siguientes afirmaciones son equivalentes:

   a) $\exists\, A^{-1}$

   b) $R(A) = n$

   c) $\det(A) \neq 0$

2) Si A es una matriz cuadrada de orden n, singular, las siguientes afirmaciones son equivalentes:

   a) $\nexists\, A^{-1}$

   b) $R(A) < n$

   c) $\det(A) = 0$

Empleando la relación (30), podemos obtener la inversa de la matriz

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

del ejemplo IV. 29

En efecto, como

$$\det(A) = \begin{vmatrix} 1 & 0 & 4 \\ -1 & 2 & 3 \\ 3 & 1 & -2 \end{vmatrix} = -35 \neq 0$$

existe A^{-1}.

Del resultado del ejemplo IV. 29

$$A^{-1} = \frac{1}{\det(A)} A^* = -\frac{1}{35} \begin{bmatrix} -7 & 4 & -8 \\ 7 & -14 & -7 \\ -7 & -1 & 2 \end{bmatrix}$$

Para comprobar efectuemos el producto $A\,A^{-1}$:

$$A\,A^{-1} = -\frac{1}{35} \begin{bmatrix} 1 & 0 & 4 \\ -1 & 2 & 3 \\ 3 & 1 & -2 \end{bmatrix} \begin{bmatrix} -7 & 4 & -8 \\ 7 & -14 & -7 \\ -7 & -1 & 2 \end{bmatrix} = -\frac{1}{35} \begin{bmatrix} -35 & 0 & 0 \\ 0 & -35 & 0 \\ 0 & 0 & -35 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- -

<u>Regla de Cramer.</u>

Otra de las aplicaciones de los determinantes, la encontramos en la resolución de sistemas de ecuaciones lineales. La definición de determinante dada por la expresión (27), permite generalizar la regla de Cramer [3] enunciada al principio de este -

(3) La demostración de esta afirmación puede consultarse en la referencia 2 pags. 50 a 54.

$$x_1+x_2+x_3=3$$
$$-x_1+x_2+2x_3=4$$

Solución

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ -1 & 1 & 2 & 4 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 2 & 3 & 7 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & \frac{3}{2} & \frac{7}{2} \end{array}\right]$$

$R(A)=R(A,\bar{b})=2=r$ el sistema es compatible

$2<3$, $(r<n)$ el sistema es indeterminado

Trasladando a los segundos miembros la incógnita x_3 y - - asignando el valor de k obtenemos

$$x_1+x_2=3-k$$
$$-x_1+x_2=4-2k$$

$$\det(A)=\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix}=2$$

$$\det(A_1)=\begin{vmatrix} 3-k & 1 \\ 4-2k & 1 \end{vmatrix}=-1+k$$

$$\det(A_2)=\begin{vmatrix} 1 & 3-k \\ -1 & 4-2k \end{vmatrix}=7-3k$$

$$x_1=\frac{-1+k}{2}$$

$$x_2=\frac{7-3k}{2}$$

$$x_3=k \qquad \forall\, k \in R.$$

Una solución particular es por ejemplo

$$x_1=-\frac{1}{2}$$

$$x_2=\frac{7}{2}$$

$$x_3=0$$

la cual se obtuvo haciendo k=0

- -

BIBLIOGRAFIA

1) Pierce Beaumont.
The algebraic foundations of Mathematics.
Editorial Addison Wesley. 1963.

2) A. G. Kurosch
Curso de Algebra Superior
Editorial MIR. Moscú, 1968

3) Franz E. Hohn
Algebra de Matrices
Editorial Trillas, S. A. 1970