

*METODOS EXACTOS EN SISTEMAS  
DE ECUACIONES LINEALES*

**BIBLIOTECA  
DE LA  
UNIVERSIDAD DEL VALLE DE GUATEMALA**

Te.  
M28  
1983

*UNIVERSIDAD DEL VALLE DE GUATEMALA*  
*Facultad de Ciencias y Humanidades*

*METODOS EXACTOS EN SISTEMAS*  
*DE ECUACIONES LINEALES*

*JOSE MATA ESTRADA*

*Trabajo de investigación para*  
*optar al grado académico de*  
*Licenciado en Matemáticas*

*Guatemala*

*1983*

*Vo. Bo. :*

*(f)* \_\_\_\_\_

*Asesor: Ingeniero Jacinto Quan*

*Tribunal:*

*(f)* \_\_\_\_\_

*Doctor Jorge Antillón Matta*

*(f)* \_\_\_\_\_

*Ingeniero Jacinto Quan*

*(f)* \_\_\_\_\_

*Licenciado Leonel Morales*

*Fecha de aprobación: 29 de marzo de 1983*

*A mis padres y a mi her-  
mana Guadalupe.*

## CONTENIDO

	<i>páginas</i>
<i>RESUMEN</i>	<i>xiii</i>
<i>INTRODUCCION</i>	1
<i>CAPITULO I</i>	7
<i>Espacios vectoriales de dimensión finita.</i>	
1.1 <i>Primeras definiciones</i>	7
D-1 <i>Espacio vectorial</i>	7
D-2 <i>Combinación lineal</i>	8
D-3 <i>Subespacio</i>	8
D-4 <i>Independencia lineal</i>	8
D-5 <i>Base de un espacio vectorial</i>	
D-6 <i>Matrices</i>	9
D-7 <i>Operaciones en matrices</i>	10
D-8 <i>Transformación lineal (operador lineal)</i>	12
D-9 <i>Isomorfismo</i>	12
D-10 <i>Kernel de una transformación lineal</i>	12
D-11 <i>Operador lineal no-singular</i>	12
D-12 <i>Funcional lineal</i>	13
1.2 <i>Primeros teoremas</i>	13
T-1 <i>Subespacio y combinación lineal</i>	14
T-2 <i>Unicidad de componentes</i>	14
T-3 <i>Transformación de componentes</i>	15
T-4 <i>Propiedades algebraicas de las matrices</i>	16
T-5 <i>Espacio vectorial de las transformaciones lineales</i>	17
T-6 <i>Kernel y subespacio</i>	18
T-7 <i>Kernel y transformación lineal uno-uno</i>	18
T-8 <i>Espacio vectorial de los funcionales lineales</i>	19
1.3 <i>Otros teoremas y definiciones</i>	20
T-9 <i>Unicidad del número de elementos de una base</i>	

D-13	Dimensión y espacio de dimensión finita	21
T-10	Dimensión y subespacio	22
T-11	Completitud de base	22
D-14	Producto interior: espacio Euclideo y espacio de Hermite	22
D-15	Conjuntos ortogonales y ortonormales	23
T-12	Propiedades del producto interior	24
T-13	Invariancia del producto interior	24
D-16	Rango de una matriz	25
D-17	El espacio vectorial $M_m^2$	26
D-18	Inversa de una matriz en $M_m^2$	26
D-19	El espacio $T_m^2$	26
T-14	Isomorfismo entre $T_m^2$ y $M_m^2$	27
T-15	Dimensión y espacios isomorfos	30
T-16	Dimensión de $T_m^2$ y $M_m^2$	30
1.4	Teoría de determinantes	30
A)	Definición de determinantes	31
B)	Propiedades básicas de un determinante	32
C)	Determinantes y linealidad	32
D)	Unicidad del determinante y representación cuadrada	34
E)	Expansión del determinante en terminos de una columna	35
F)	Determinante del producto de matrices	39
G)	Traspuesta de A y $D(A) = D(A^t)$	40

CAPITULO II 41

Teoría de sistemas de ecuaciones lineales

2.1	Definición de un sistema lineal	41
2.2	Interpretaciones equivalentes del sistema $L_{mn}$	42
2.3	Una solución teórica del sistema $L_{mn}$	47
2.4	El sistema homogéneo, $H(L_{mn})$ , y una solución vectorial de $L_{mn}$	49
2.5	Kernel y el sistema $H(L_{mn})$	54
2.6	El sistema homogéneo traspuesto, $T(L_{mn})$ , a $L_{mn}$	55
2.7	Teorema de dualidad	58
2.8	Condiciones de existencia	60
2.9	Determinantes y rango de una matriz	62
2.10	Ecuaciones principales y nuevas condiciones de existencia	64
2.11	La solución de Cramer	69
	<b>CAPITULO III</b>	75
	<b>Solución numérica de sistemas lineales</b>	
3.1	Solución numérica de sistemas lineales y algunos problemas	75
3.2	Matrices y transformaciones no-singulares	77
	D-1 Matriz no-singular	77
	T-1 Operador lineal no-singular e independencia lineal	78
	Corolario T-1	79
	T-2 Matriz no-singular, $H(L_m^2)$ , determinante, inversa	80
	T-3 Lema de Gersgorin	82
3.3	Operatoria numérica sobre matrices	83
	D-2 Operaciones entre filas	83

D-3	Matriz elemental	84
D-4	Matriz diagonal	84
D-5	Matriz triangular	84
D-6	Matriz equivalente por filas	84
T-4	Matriz identidad y operaciones entre filas	85
	Corolario T-4	89
D-7	Relaciones y matrices equivalentes entre filas	90
T-5	Matrices equivalentes definen una relación de equivalencia	90
T-6	Matrices, clases de equivalencia, rango y matriz no-singular	92
3.4	Rango, condiciones de existencia y solución de $L_{mn}$	92
T-7	$L_{mn}$ y matrices equivalentes por filas	92
D-8	Matriz en la forma normal de Hermite	95
T-8	Matrices y la forma normal de Hermite	96
	Corolario T-8	103
3.5	Eliminación de Gauss	108
3.6	Cálculo de determinantes y matriz inversa	113
	a) Determinantes	113
	b) Cálculo de la inversa de una matriz	117
	CONCLUSIONES	119
	BIBLIOGRAFIA	123

*RESUMEN DE LA TESIS*  
*Métodos exactos en sistemas de*  
*ecuaciones lineales*

*por*

*José Mata Estrada*

*Licenciado en Matemáticas*

*Universidad del Valle de Guatemala, 1983*

*Asesor, Jacinto Quan*

*Este trabajo desarrolla en forma más o menos completa el problema de los sistemas de ecuaciones lineales, aprovechando la teoría del álgebra lineal.*

*En el capítulo I se expone, en forma resumida, la teoría elemental del álgebra lineal. Teoría que será necesaria para desarrollar los sistemas lineales y que además se puede utilizar en una gran variedad de problemas.*

*En el capítulo II se desarrolla lo que se podría llamar las soluciones teóricas de los sistemas de ecuaciones lineales: la solución de Cramer se incluye en este capítulo.*

*En el capítulo III, finalmente y como principal objetivo, se desarrolla la solución numérica de los sistemas lineales: solución numérica que se orienta por los llamados métodos exactos. En este capítulo se plantean ciertos problemas numéricos de los sistemas lineales y del análisis numérico; sin embargo, no se tratan a profundidad estos problemas.*

*La conexión entre los tres capítulos, aparte de su secuencia lógica, intenta contrastar un enfoque teórico y un enfoque numérico; es en este aspecto donde el trabajo se aparta de los enfoques tradicionales y donde pueda tener algún mérito.*

## I. INTRODUCCION

El presente trabajo consiste en desarrollar un enfoque numérico del álgebra lineal contrapuesto al enfoque teórico tradicional. Sin embargo, el objetivo mencionado no es así de simple como se plantea. Enfoque numérico del álgebra lineal quiere decir, al menos aquí, orientar el álgebra lineal hacia el problema de los sistemas lineales y sus asociados: determinantes e inversas de matrices.

La razón de reducir el enfoque numérico hacia estos tres problemas radica, nada más, en la frecuencia con que se demanda la solución de cada uno; además de la actualidad e insistencia con que se trabaja en ellos. Un matemático, tomado por sorpresa, podría creer que en este tipo de problemas ya todo está hecho: falso, no todo está hecho, más aún, lo que queda por hacer parece de gran interés y de complejidad extremada. Si apelamos a la autoridad, de los que tienen prestigio en la materia, aquí tenemos un par de opiniones. En primer lugar, Scheid (1972:334) nos dice:

"La solución de sistemas lineales puede muy bien ser la tarea más destacada del análisis numérico."

Por otro lado, Faddeev, en American Mathematical Society (1968:39) dice:

"Historically, the first task of linear algebra is that of solving a system of linear equations."

*The simplest case of this problem is treated in a school course on elementary algebra. The problem of finding methods for the simplest possible and least laborious numerical solution of systems for large  $n$  still attracts the close attention of many researchers, because the numerical solution of such systems enters as an important constituent part into many calculations and investigations."*

*Este trabajo no intenta abordar los problemas no resueltos de los sistemas lineales, sino dirigir el álgebra lineal hacia la solución conocida de los sistemas lineales, determinantes e inversas de matrices.*

*Qué se puede lograr con un enfoque de este tipo?. En primer lugar una mejor comprensión del tema: en este caso el álgebra lineal. Cómo se puede lograr lo anterior?. La mayor parte de libros exponen el álgebra lineal desde un punto de vista esencialmente teórico, tratando de mostrar todos los alcances y bellezas del tema; por supuesto, evitando los embarazosos cálculos de los aspectos numéricos, los cuales se mencionan sólo de paso. Otros libros, menos populares, dejan la teoría un tanto de lado y se sumergen en la amplitud de cálculos numéricos que presenta el álgebra lineal. Es inevitable, existen enfoques marcadamente opuestos: teoría y enfoque aplicado o numérico. Debe quedar*

claro que, por un lado no puede existir análisis numérico sin teoría. Por otro la teoría se torna sin sentido, o en el mejor de los casos un tanto "despiada", si pierde de vista sus aspectos numéricos. Deberá existir una relación entre ambas tendencias si esperamos que las cosas caminen; y este enfoque se desenvuelve dentro del sentido de esa dependencia. El contenido no va dirigido exclusivamente ni al matemático, ni al ingeniero: es básico y provechoso para la formación de ambos.

La mitad del trabajo, o más, es casi sólo teoría que puede ramificarse o desembocar en muchos aspectos, pero aquí dicha teoría sólo se usa para abarcar en forma más o menos completa, el aparentemente simple problema de los sistemas lineales y sus asociados o afines.

Un objetivo de este trabajo, siempre presente, es no evadir la teoría necesaria, aunque a veces sí el rigor lógico, para mostrar la cantidad de razonamiento teórico necesario para los sistemas lineales. Luego, al desembocar en el aspecto numérico (capítulo III), la teoría tiene nuevos elementos para seguir adelante. Algo así como que; la solución numérica de un problema hace significativa la investigación teórica por varios senderos. Por ejemplo, las operaciones entre filas en una matriz, recurso de índole numérica para resolver un sistema lineal, inducen una partición en clases de equivalencia sobre un espacio vectorial de matrices. Es-

tas clases de equivalencia, generadas por un procedimiento numérico, ofrecen interesantes aspectos teóricos y prácticos: una de esas clases reúne a las matrices no-singulares, cuyo campo de aplicación parece ser ilimitado.

La solución numérica de los sistemas lineales ha sido limitada, en este trabajo, sólo a los llamados métodos exactos (directos); no por ser mejores sino por que son más sencillos y no existen métodos de mayor confiabilidad. Los métodos exactos son buenos en ciertos casos, malos en otros; así están las cosas.

El objetivo central del trabajo se alcanza en el teorema 8 del capítulo III. Es ahí donde, por fin, hay solución numérica completa para los sistemas lineales en general, por lo menos para principiar. Qué queremos decir con principiar? En primer lugar; decir solución numérica completa lo único que indica es: que el método resuelve un sistema lineal en forma completa; desde el rango de la matriz, condiciones de existencia hasta las soluciones si las hay. Pero el método esbozado en el teorema 8 puede, en muchos casos, ser sólo el comienzo a causa de las cosas numéricas extrañas que pueden suceder; éstas son, entre otras: errores de acumulación, cantidad de operaciones, inestabilidad del sistema, gran tamaño del sistema, etc. Lo que llamamos cosas extrañas son problemas de índole marcadamente numérica y en gran parte independientes o ajenos al álgebra lineal

*en general. Dichas cosas extrañas no se desarrollan en este trabajo, sólo se menciona y se esboza su importancia, pues se consideran problemas especializados dentro del problema.*

*Las dificultades de un enfoque, como el que aquí se presenta, consisten en dar una estructura lógica flexible (marco teórico), luego conectar dicho marco con el problema numérico y, finalmente, mostrar los resultados y nuevas dificultades: eso es lo que intenta este trabajo.*

## CAPITULO I

### 1.1 Primeras definiciones

**D-1 Espacio Vectorial.** Un espacio vectorial se define en un conjunto  $E$  arbitrario y un campo  $F$  (los reales  $R$ , o si no los complejos,  $C$ ). Luego dos operaciones: la suma definida en  $E$  y el producto de un escalar por un vector; donde escalar se llama un elemento arbitrario de  $F$  y vector un elemento cualquiera de  $E$ .

Para la suma, definida en  $E$ , debe cumplirse que:

- i.  $X, Y \in E$ , entonces  $X + Y \in E$ . (cerradura)
- ii.  $X + (Y + Z) = (X + Y) + Z$ . (asociatividad)
- iii.  $X + Y = Y + X$ . (conmutatividad)
- iv.  $\exists 0 \in E$ , tal que  $X + 0 = X$ . (existencia de neutro)
- v.  $X \in E$ , entonces  $\exists -X \in E$  tal que  $X + (-X) = 0$ .  
(existencia de simétricos)

donde  $X, Y$  y  $Z$  son elementos cualesquiera del conjunto  $E$ .

Para el producto de un escalar por un vector tenemos:

- i.  $\alpha \in F$  &  $X \in E$ , entonces  $\alpha X \in E$   
(“cerradura”)
- ii.  $\alpha (\beta X) = (\alpha \beta) X$ . (”asociatividad”)
- iii.  $\alpha (X + Y) = \alpha X + \alpha Y$ . (”distributividad”)
- iv.  $(\alpha + \beta) X = \alpha X + \beta X$ .
- v.  $1X = X$ .

donde  $\alpha, \beta$  son cualesquiera escalares de  $F$ , mientras que  $X$ ,

$Y$  son vectores arbitrarios.

D-2 Combinación lineal. Dados  $n$  vectores  $X_i$ ,  $i = 1, 2, \dots, n$ , de un espacio vectorial  $E$ , decimos que una combinación lineal de ellos es cualquiera de las formas

$$\sum_{i=1}^n \alpha^i X_i,$$

en donde los  $\alpha^i$ ,  $i = 1, 2, \dots, n$  son escalares arbitrarios.

D-3 Subespacio. Sea  $E$  un espacio vectorial y  $V \subset E$ , decimos que  $V$  es subespacio de  $E$  si cumple con:

- i.  $X, Y \in V$ , entonces  $X + Y \in V$ . (cerradura de la suma)
- ii.  $X \in V$  y  $\alpha \in F$ , entonces  $\alpha X \in V$ . (cerradura del producto)

D-4 Independencia lineal. Dados  $n$  vectores  $X_i$ , que pertenecen al espacio  $E$ , decimos que esos  $n$  vectores son linealmente independientes si para cualquier combinación lineal

$$\sum_{i=1}^n \alpha^i X_i = 0,$$

se deduce que  $\alpha^i = 0$ ,  $i = 1, 2, \dots, n$ .

Si los vectores no son linealmente independientes se llaman linealmente dependientes o, sencillamente dependientes.

D-5 Base de un espacio vectorial. Decimos que un conjunto de  $n$  vectores linealmente independientes,  $\{X_i\}_{i=1}^n$ , es base del espacio  $E$  si dichos vectores cumplen con que para cualquier  $X \in E$ ,  $X$  es combinación lineal de los  $X_i$ . Es decir:

$$X = \sum_{i=1}^n x^i X_i.$$

Los escalares  $x^i$  en la combinación lineal anterior se llaman las componentes del vector  $X$  en la base dada. En el teorema 2 se mostrará que estas componentes son únicas.

D-6 Matrices. Un arreglo rectangular de escalares, ya sean reales o complejos, que presente la siguiente forma

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix} \begin{array}{l} m \text{ filas} \\ n \text{ columnas} \end{array}$$

se llama una matriz.

Nota: Observe que una matriz, llamémosle  $A$ , se puede considerar como un arreglo de números en filas y columnas. Al elemento que aparece en la fila  $j$ -ésima, columna  $i$ -ésima, lo llamamos elemento general  $j, i$ -ésimo y lo representamos así  $a_i^j$ . La matriz  $A$  se puede representar en base al elemento general así:  $A = [a_i^j]$ .

La matriz que definimos tiene  $m$  filas y  $n$  columnas, por ello se llama una matriz  $m$  por  $n$ . Una matriz  $m$  por  $m$  se llama matriz cuadrada de orden  $m$ .

En este trabajo se seguirá el convenio de que el superíndice representa la variable de filas, mientras que el subíndice representa la variable de columnas.

índice la variable de columnas. La mayor parte de las veces se busca que la variable  $j$  sea supraíndice (para filas) y la variable  $i$  subíndice (para columnas). Si embargo, esto no siempre será así; debe quedar claro que lo importante no es el nombre de la variable sino su posición. De ahí que la variable  $i$ , en ciertos casos, representará filas por ser su praíndice.

D-7 Operaciones en matrices. Sea  $E$  el conjunto de todas las matrices  $m$  por  $n$ , cuyos elementos pertenecen al campo real o complejo, vamos a definir en  $E$  dos operaciones de la manera siguiente.

a) Dadas las matrices  $A = [a_i^j]$  y  $B = [b_i^j]$  tenemos por suma de  $A$  con  $B$ ,  $A + B$ , a la nueva matriz dada así:

$$A + B = [(a + b)_i^j] = [a_i^j + b_i^j].$$

El producto de un escalar  $\alpha$  por una matriz  $A$ ,  $\alpha A$ , lo definimos como:

$$\alpha A = [\alpha a_i^j], \text{ en donde } A = [a_i^j].$$

En la representación rectangular, empleada para matrices, las dos definiciones anteriores aparecen como:

$$A + B = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix} + \begin{pmatrix} b_1^1 & b_2^1 & \dots & b_n^1 \\ b_1^2 & b_2^2 & \dots & b_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ b_1^m & b_2^m & \dots & b_n^m \end{pmatrix} =$$

$$A+B = \begin{pmatrix} a_1^1 + b_1^1, & a_2^1 + b_2^1, & \dots, & a_n^1 + b_n^1 \\ a_1^2 + b_1^2, & a_2^2 + b_2^2, & \dots, & a_n^2 + b_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m + b_1^m, & a_2^m + b_2^m, & \dots, & a_n^m + b_n^m \end{pmatrix}.$$

$$\alpha A = \alpha \begin{pmatrix} a_1^1, & a_2^1, & \dots, & a_n^1 \\ a_1^2, & a_2^2, & \dots, & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m, & a_2^m, & \dots, & a_n^m \end{pmatrix} = \begin{pmatrix} \alpha a_1^1, & \alpha a_2^1, & \dots, & \alpha a_n^1 \\ \alpha a_1^2, & \alpha a_2^2, & \dots, & \alpha a_n^2 \\ \vdots & \vdots & & \vdots \\ \alpha a_1^m, & \alpha a_2^m, & \dots, & \alpha a_n^m \end{pmatrix}.$$

b) La operación llamada producto de matrices se define en el conjunto  $E$  de todas las matrices  $m$  por  $n$  y el conjunto  $V$  de todas las matrices  $n$  por  $r$ . Sea  $A = [a_j^i] \in E$  y  $B = [b_k^j] \in V$ ; en donde  $i = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ ;  $k = 1, 2, \dots, r$ ; entonces el producto  $AB = C$  viene dado por:

$$C = [c_k^i] = \sum_{j=1}^n a_j^i b_k^j, \quad \begin{matrix} i = 1, 2, \dots, m \\ k = 1, 2, \dots, r \end{matrix}.$$

El elemento general de la matriz  $AB$ ,  $(c_k^i)$ , también es posible representarlo como  $(ab)_k^i$ .

**Nota 1.** Observe que el resultado de multiplicar una matriz  $m$  por  $n$  con una  $n$  por  $r$  da una nueva matriz  $m$  por  $r$ .

**Nota 2.** Para el caso de una matriz  $A$ ,  $m$  por  $n$ , y una matriz  $X$ ,  $n$  por  $1$  (se llama a la matriz  $X$  un vector columna, en caso que  $X$  fuera una matriz  $1$  por  $n$  se llamaría vector

fila.) el producto viene representado así:

$$AX = \sum_{j=1}^n a_{ij} x^j = C, \quad i=1, 2, \dots, m.$$

Debe quedar claro que el resultado de este producto es un vector columna  $m$  por  $1$ .

D-8 Transformación lineal (operador lineal). Sean  $E$  y  $V$  dos espacios vectoriales con un mismo campo. Una función  $T$  de  $E$  en  $V$  es llamada transformación lineal, u operador lineal, si cumple con:

a)  $T(X+Y) = T(X) + T(Y)$ , con  $X, Y \in E$ .

b)  $T(\alpha X) = \alpha T(X)$ , con  $\alpha \in F$ .

Un caso importante de operador lineal consiste cuando  $E = V$ .

D-9 Isomorfismo. Una transformación lineal  $T$  de  $E$  en  $V$  ( $E$  y  $V$  espacios vectoriales con un mismo campo.) se llama isomorfismo si  $T$  es inyectiva (uno-uno) y sobreyectiva (sobre). Cuando existe un isomorfismo entre dos espacios  $E$  y  $V$ , estos se llaman isomorfos.

D-10 Kernel de una transformación lineal. Dada una transformación lineal de el espacio  $E$  en el espacio  $V$ , decimos que el conjunto  $K(T)$  es el kernel de la transformación lineal,  $T$ , cuando dicho conjunto se define como:

$$K(T) = \{ X \in E / T(X) = 0 \} .$$

D-11 Operador lineal no-singular. Dada una transformación lineal de  $E$  en  $E$ , decimos que  $T$  es un operador lineal no-singular si la transformación  $T$  es sobre

yectiva (sobre). Un operador que no es no-singular se dice que es singular.

D-12 Funcional lineal. Decimos que una función  $f$  de un espacio vectorial  $E$  en  $F$  (con  $F$  campo de  $E$ .) es un funcional lineal si cumple que:

- a)  $f(X+Y) = f(X) + f(Y)$ , con  $X, Y \in E$ .
- b)  $f(\alpha X) = \alpha f(X)$ , con  $\alpha \in F$  y  $X \in E$ .

Las primeras cinco definiciones, dadas en esta sección, atañen a la teoría general del álgebra lineal; las definiciones 6 y 7 van sobre el caso particular de matrices, que serán el medio directo para trabajar los problemas desde un punto de vista numérico; finalmente, las definiciones 8-12 van, esencialmente, sobre el caso de las transformaciones lineales. Son las transformaciones lineales las que, según Halmos (1974:55), hacen interesante o valedero el estudio de los espacios vectoriales.

Los teoremas que a continuación desarrollaremos están orientados de la siguiente manera: los tres primeros son propiedades básicas de la teoría del álgebra lineal. El teorema 4 no es más que el carácter de espacio vectorial del conjunto de las matrices, y las propiedades de la operación producto de matrices. Finalmente, los últimos cuatro teoremas señalan propiedades elementales de las transformaciones lineales.

## 1.2 Primeros teoremas

**T-1** Subespacio y combinación lineal. Dado un conjunto de vectores  $\{X_i\}_{i=1}^n$  que pertenecen a un espacio vectorial  $E$ , el conjunto de todas las combinaciones lineales de estos  $X_i$  es un subespacio vectorial de  $E$ . Dicho subespacio vectorial se llama el subespacio generado por el conjunto  $\{X_i\}_{i=1}^n$ , y lo representamos por  $S(\{X_i\}_{i=1}^n)$ .

*Demostración.*

De acuerdo a D-3, hay que demostrar que:

- i)  $X, Y \in S(\{X_i\}_{i=1}^n)$ , entonces  $X+Y \in S(\{X_i\}_{i=1}^n)$ .  
 ii)  $X \in S(\{X_i\}_{i=1}^n)$  y  $\alpha \in F$ , entonces  $\alpha X \in S(\{X_i\}_{i=1}^n)$ .

Para demostrar i) tenemos: si

$$X = \sum_{i=1}^n x^i X_i \quad y \quad Y = \sum_{i=1}^n y^i X_i, \quad \text{entonces}$$

$$X+Y = \sum_{i=1}^n x^i X_i + \sum_{i=1}^n y^i X_i = \sum_{i=1}^n (x^i + y^i) X_i,$$

de donde se concluye que  $X+Y \in S(\{X_i\}_{i=1}^n)$ .

Para demostrar ii) tenemos: si

$$\alpha \in F, \quad \text{entonces} \quad \alpha X = \alpha \sum_{i=1}^n x^i X_i = \sum_{i=1}^n (\alpha x^i) X_i,$$

por lo que  $\alpha X \in S(\{X_i\}_{i=1}^n)$ . #

**T-2** Unicidad de componentes. Dada una base  $\{e_i\}_{i=1}^m$  del espacio vectorial  $E$ , las componentes  $x^i$ ,  $i=1, 2, \dots, m$  tales que  $X = \sum_{i=1}^m x^i e_i$  son únicas.

*Demostración.*

Supongamos que  $X = \sum_{i=1}^m x^i e_i = \sum_{i=1}^m y^i e_i$ , entonces

$$\sum_{i=1}^m (x^i - y^i) e_i = 0.$$

Por la independencia lineal de los  $e_i$  tenemos que

$$x^i - y^i = 0, \quad \text{para } i=1, 2, \dots, m;$$

y de ahí, finalmente,

$$x^i = y^i, \text{ para } i=1, 2, \dots, m.$$

**T-3 Transformación de componentes.** Dadas dos bases

$\{e_i\}_{i=1}^m$  y  $\{e'_j\}_{j=1}^m$  del espacio vectorial  $E$ , entonces existen dos matrices  $A = [a_i^j]$  y  $A' = [d_j^i]$  tal que, para  $X \in E$  con  $X = \sum_{i=1}^m x^i e_i = \sum_{j=1}^m x'^j e'_j$ , se cumple que:

$$X = A'X' \text{ \& } X' = AX.$$

$X$ , en la ecuación anterior, es el vector columna de componentes de  $X$  en la base  $\{e_i\}_{i=1}^m$ , mientras que  $X'$  es el vector columna de componentes de  $X$  en la base  $\{e'_j\}_{j=1}^m$ . Ver nota 2, definición 7.

Demostración.

Por cuanto que  $\{e'_j\}_{j=1}^m$  es una base tenemos que existen los  $a_i^j$  tal que  $e_i = \sum_{j=1}^m a_i^j e'_j$ ,  $i=1, 2, \dots, m$ . Luego hacemos  $A = [a_i^j]$ . Ya con lo anterior resulta que:

$$X = \sum_{j=1}^m x'^j e'_j = \sum_{i=1}^m x^i \left( \sum_{j=1}^m a_i^j e'_j \right),$$

$$\sum_{j=1}^m x'^j e'_j = \sum_{j=1}^m \left( \sum_{i=1}^m x^i a_i^j \right) e'_j.$$

Finalmente, por el teorema 2, la ecuación anterior se reduce a:

$$x'^j = \sum_{i=1}^m x^i a_i^j, \quad j=1, 2, \dots, m;$$

es decir,  $X' = AX$ .

De forma similar se demuestra  $X = A'X'$ , en donde  $A' = [d_j^i]$  dada por la ecuación  $e'_j = \sum_{i=1}^m d_j^i e_i$ ,  $j=1, 2, \dots, m$ .

Nota: las matrices  $A$  y  $A'$  se llaman matrices de transformación de componentes.

#

**T-5 Propiedades algebraicas de las matrices.** a) Da-

do el conjunto  $E$  de todas las matrices  $m$  por  $n$  y las dos operaciones definidas en D-7 (suma de matrices y producto de un escalar por una matriz.); entonces el conjunto  $E$  es un espacio vectorial de dimensión  $nm$ .

b) El producto de matrices, también definido en D-7, cumple con las propiedades siguientes:

i.  $A(cB) = c(AB)$ , con  $c$  escalar.

ii.  $C(BA) = (CB)A$ . (asociatividad)

iii.  $(B+C)A = BA + CA$ . (distributividad)

iv.  $A(B+C) = AB + AC$ .

c) Adición y multiplicación definen una estructura de anillo en el conjunto de todas las matrices  $m$  por  $m$  (matrices cuadradas).

**Demostración.**

a) La prueba es muy simple y consiste en verificar que se satisfacen las propiedades, para suma y producto, dadas en D-1. Aquí sólo diremos que la matriz cero es la que tiene cero como elemento general, y que una base de este espacio son los  $mn$  matrices dadas así:

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \dots \begin{pmatrix} 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

$$\begin{array}{c}
 \left( \begin{array}{cccc} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{array} \right) \left( \begin{array}{cccc} 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{array} \right) \dots \left( \begin{array}{cccc} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{array} \right) \\
 \vdots \\
 \left( \begin{array}{cccc} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \dots & 0 \end{array} \right) \left( \begin{array}{cccc} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 1 & \dots & 0 \end{array} \right) \dots \left( \begin{array}{cccc} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{array} \right) .
 \end{array}$$

b) La demostración está basada en la definición del producto de matrices, y en las propiedades de la suma y producto para un campo.

c) Esto se sigue por lo demostrado en a) y b), junto con la definición de anillo: un anillo es un conjunto y dos operaciones, llamadas suma y producto; la suma debe ser un grupo abeliano, mientras que el producto debe ser asociativo; y finalmente el producto debe ser distributivo respecto a la suma.

**T-5** Espacio vectorial de las transformaciones lineales. El conjunto de todas las transformaciones lineales de un espacio  $E$  en  $V$ , llamémosle  $\mathcal{T}$ , es de nuevo un espacio vectorial.

*Demostración.*

Primero se define la suma de dos transformaciones lineales así:

$$(T_1 + T_2)(X) = T_1(X) + T_2(X), \quad X \in E.$$

Luego, se define el producto de un escalar por una transformación lineal así:

$$(\alpha T_1)(X) = \alpha T_1(X), \quad \alpha \in F \text{ y } X \in E.$$

Ya definidas las dos operaciones, la demostración sólo consiste en verificar las propiedades dadas en D-1. El vector cero en este espacio  $\mathcal{L}$  es la transformación nula definida como  $T(X) = 0$  para toda  $X \in E$ . #

**T-6 Kernel y subespacio.** Dada una transformación lineal  $T$  de  $E$  en  $V$ , entonces el  $K(T)$  es un subespacio.

*Demostración.*

Hay que probar primero que si  $X$  y  $Y$ , vectores arbitrarios, pertenecen a  $K(T)$  entonces  $X+Y$  pertenece a  $K(T)$ . Esto se demuestra porque  $T(X+Y) = T(X) + T(Y) = 0 + 0 = 0$ .

Luego, hay que probar que si  $X$  pertenece a  $K(T)$ , entonces  $\alpha X$  pertenece a  $K(T)$ , para  $\alpha$  un escalar arbitrario. Esto está demostrado a causa de que  $T(\alpha X) = \alpha T(X) = \alpha 0 = 0$ . #

**T-7 Kernel y transformación lineal uno-uno.** Dada una transformación lineal  $T$  de  $E$  en  $V$ , entonces tenemos que:  $K(T) = 0$  si y sólo si  $T$  es inyectiva (uno-uno).

*Demostración.*

i. Si  $K(T) = 0$ , entonces hay que demostrar que  $T(X) = T(Y)$

implica que  $X=Y$ . Esto se prueba así:

de  $T(X) = T(Y)$  se deduce que  
 $T(X) - T(Y) = 0$ , luego  
 $T(X - Y) = 0$ ; y como  $K(T) = 0$ , entonces  
 $X - Y = 0$ , es decir  
 $X = Y$ .

ii. Ahora hay que probar que  $T$  uno-uno implica  $K(T) = 0$ .

$T(X) = T(Y)$  implica  $X = Y$ , es equivalente con,  
 $X \neq Y$  implica  $T(X) \neq T(Y)$ ; y por lo tanto sólo  
 $T(0) = 0$ . Es decir,  $K(T) = 0$ . #

#### T-8 Espacio vectorial de los funcionales lineales.

Dado un espacio vectorial  $E$ , el conjunto de todos los funcionales lineales definidos en  $E$  es un espacio vectorial. Dicho espacio vectorial se llama espacio dual.

##### Demostración.

La demostración aquí, como en el caso de probar que cualquier conjunto es un espacio vectorial (con sus operaciones de suma y producto), consiste en demostrar las propiedades de la suma y producto dadas en D-1. Primero, pues, hay que contar con la suma y el producto de escalar por vector, definidos sobre el conjunto de todos los funcionales lineales. Estas operaciones se definen así:

$$(f + g)(X) = f(X) + g(X), \text{ con } X \in E.$$

$$(\alpha f)(X) = \alpha f(X), \text{ con } \alpha \in F.$$

El resto de la demostración es rutina. #

### 1.3 Otros teoremas y definiciones

a) Sobre la teoría general del álgebra lineal.

#### T-9 Unicidad del número de elementos de una base

Sea  $E$  un espacio vectorial y una base de  $m$  elementos  $\{e_i\}_{i=1}^m$ , entonces cualquier otra base de  $E$  tiene también  $m$  elementos.

*Demostración.*

Supongamos, en primer lugar, que existiera otra base  $\{é_j\}_{j=1}^p$ ,  $p > m$ . Dado que  $\{e_i\}_{i=1}^m$  es base de  $E$  entonces los  $e_1, e_2, \dots, e_m$  son linealmente independientes, luego los  $é_1, e_1, e_2, \dots, e_m$  son linealmente dependientes y generan al espacio  $E$ . Por lo tanto, existe en el conjunto de los  $é_1, e_1, e_2, \dots, e_m$  un elemento que es combinación lineal de los precedentes. Dicho elemento no puede ser  $é_1$ , porque los  $é_j$  son linealmente independientes, así pues debe ser algún  $e_j$  de los  $\{e_i\}_{i=1}^m$ . Si del conjunto de  $m+1$  elementos

$$é_1, e_1, e_2, \dots, e_m$$

quitamos el elemento  $e_j$  formamos un conjunto de  $m$  elementos dado por

$$é_1, e_1, e_2, \dots, e_{j-1}, e_{j+1}, \dots, e_m.$$

Este conjunto genera a  $E$ ; y si en él agregamos, como anteriormente, el siguiente elemento del conjunto  $\{é_j\}_{j=1}^p$ , se forma nuevamente un conjunto linealmente dependiente, de  $m+1$  elementos, generador de  $E$ :

$$é_1, é_2, e_{k_1}, e_{k_2}, \dots, e_{k_m}, \text{ en donde ya no aparece}$$

el  $e_j$  de los  $\{e_i\}_{i=1}^m$ . En este último conjunto de nuevo existe un elemento  $e_{k_j}$  de los  $\{e_{k_i}\}_{k_i=1}^{m-1}$  que es combinación lineal de los precedentes.  $e_1, e_2$ , se excluyen debido a que son linealmente independientes. Si quitamos  $e_{k_j}$  tendremos un conjunto de  $m$  elementos generador de  $E$ , dado así:

$$e_1, e_2, e_{k_2}, \dots, e_{k_{m-2}}.$$

Como  $p > m$ , después de  $m$  pasos, de repetir el proceso indicado anteriormente, tendríamos que el conjunto

$$e_1, e_2, \dots, e_m$$

genera  $E$ ; de donde, entonces, se debe tener que el conjunto

$$e_1, e_2, \dots, e_m, e_{m+1}$$

es linealmente independiente. Este resultado contradice, abiertamente, el hecho de que  $\{e_j\}_{j=1}^p$  es una base. Por lo tanto hay que concluir que  $p \leq m$ .

Si, por otro lado, suponemos que  $p < m$  y efectuamos los reemplazos anteriores de la base  $\{e_i\}_{i=1}^m$  en la base  $\{e_j\}_{j=1}^p$  concluiremos que  $p > m$ . Y es así como con  $p \leq m$  y  $p > m$  se concluye que  $p = m$ .

**D-13 Dimensión y espacio de dimensión finita.** Decimos que un espacio vectorial  $E$  es de dimensión finita cuando tiene una base finita; si este es el caso, el número de elementos de la base, que por el teorema anterior es único, se llama la dimensión del espacio.

**Notación:** Si un espacio vectorial  $E$  tiene dimensión finita  $m$ , denotamos la dimensión del espacio así:  $\dim(E) = m$ ; y al

y al espacio  $E$  como  $E_m$ .

T-10 Dimensión y subespacio. Sea  $V$  un subespacio de  $E_m$ , entonces la dimensión de  $V$  es menor o igual que la de  $E_m$ , es decir:  $\dim(V) \leq m$ .

*Demostración.*

La demostración es un corolario del teorema nueve y la definición 3 de subespacio.

T-11 Completitud de una base. Sea  $E_m$  un espacio vectorial, por lo acordado en D-13, de dimensión  $m$ , si tenemos un conjunto de  $p$  vectores linealmente independientes, con  $p < m$ , entonces dicho conjunto se puede completar en una base de  $E$ .

*Demostración.*

Este teorema es un corolario, por el método de demostración empleado, del teorema 9.

D-14 Producto interior, espacio Euclideo y espacio de Hermite. Sea  $(E_m, R)$  un espacio vectorial con campo  $R$  y  $(V_m, C)$  un espacio vectorial con campo  $C$ . En cada uno de estos dos espacios vectoriales, diferenciados por sus campos escalares, definiremos un producto llamado, en general, producto interior.

a) Para el caso de  $(E_m, R)$ , una base  $\{e_i\}_{i=1}^m$  de  $E_m$ , tendremos que si  $X, Y \in E_m$ , cuyas componentes en la base  $\{e_i\}_{i=1}^m$  son  $x^1, x^2, \dots, x^m$  &  $y^1, y^2, \dots, y^m$ , respectivamente, el producto interior queda definido así:

$$XY = (X, Y) = \sum_{i=1}^m x^i y^i.$$

b) Para el caso de  $(V_m, C)$ , y una base  $\{e_i\}_{i=1}^m$  de  $V_m$ , tendremos que el producto de dos vectores  $X, Y \in V_m$ , con componentes  $x^1, x^2, \dots, x^m$  &  $y^1, y^2, \dots, y^m$ , respectivamente, el producto interior queda definido así:

$$XY = (X, Y) = \sum_{i=1}^m x^i \bar{y}^i.$$

$\bar{y}^i$  es el conjugado del número complejo  $y^i$ . El producto definido en  $(V_m, C)$  también se llama producto de Hermite.

c) Llamamos un espacio real Euclideo,  $m$  dimensional, a  $(E_m, R)$  con el producto interior definido como en el caso a). De manera semejante, al espacio  $(V_m, C)$  con el producto de Hermite le llamamos un espacio  $m$  dimensional de Hermite.

D-15 Conjuntos ortogonales y ortonormales. Sea un espacio Euclideo  $(E_m, R)$  o un espacio de Hermite  $(V_m, C)$  decimos, para ambos casos, que un conjunto de vectores  $\{X_i\}_{i=1}^m$  es ortogonal si el producto interior correspondiente cumple con:

$$X_i X_j = (X_i, X_j) = 0,$$

para cualquier  $(i \neq j) = 1, 2, \dots, m$ .

Se dice que el conjunto de los  $\{X_i\}_{i=1}^m$  es ortonormal si:

$$(X_i, X_j) = X_i X_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} = \delta_{ij}^i.$$

Además, al caso particular de  $(X, X) = XX = N(X)$  se le llama la norma del vector  $X$ . Es claro que para vectores de una base ortonormal su norma es uno, por lo cual se dice

que son unitarios.

**T-12** Propiedades del producto interior. a) Las propiedades que cumple el producto interior en un espacio euclideo  $(E_m, R >$ , con una base dada, son las siguientes:

- i.  $X(Y+Z) = XY + XZ.$  (distributividad)
- ii.  $XY = YX.$  (conmutatividad)
- iii.  $X(\alpha Y) = \alpha(XY).$  }
- iv.  $(\alpha X)Y = \alpha(XY).$  } ("asociatividad")

Para  $X, Y \& Z \in E_m$ , con  $\alpha$  un escalar arbitrario de  $R$ .

b) Las propiedades que cumple el producto interior definido en un espacio de Hermite  $(V_m, C >$ , con una base dada son las siguientes:

- i.  $X(Y+Z) = XY + XZ.$  (distributividad)
- ii.  $XY = \overline{YX}.$  (conmutatividad)
- iii.  $(\alpha X)Y = \alpha(XY).$  }
- iv.  $X(\alpha Y) = \overline{\alpha}(XY).$  } ("asociatividad")

*Demostración.*

La prueba es rutina, basada en la definición 14, y por lo tanto se omite. #

**T-13** Invariancia del producto interior. Sea un espacio Euclideo  $(E_m, R >$  y  $\{e_i\}_{i=1}^m$  cualquier base ortonormal del espacio  $E_m$ , entonces el producto interior, definido respecto de una base  $\{e_i\}_{i=1}^m$ , queda invariante en su definición, para la base ortonormal  $\{e_i\}_{i=1}^m$ .

BIBLIOTECA  
DE LA  
UNIVERSIDAD DEL VALLE DE GUATEMALA

*Demostración.*

Sean  $X, Y$  vectores de  $E_m$  cuyas componentes en  $\{e_i\}_{i=1}^m$  son  $x^1, x^2, \dots, x^m$  &  $y^1, y^2, \dots, y^m$  respectivamente; mientras que  $\hat{x}^1, \hat{x}^2, \dots, \hat{x}^m$  &  $\hat{y}^1, \hat{y}^2, \dots, \hat{y}^m$  son las componentes, en la base  $\{\hat{e}_i\}_{i=1}^m$ , de  $X, Y$  respectivamente. Lo que hay que probar es que:

$$XY = \sum_{i=1}^m \hat{x}^i \hat{y}^i.$$

Como  $X = \sum_{i=1}^m \hat{x}^i \hat{e}_i$ ,  $Y = \sum_{j=1}^m \hat{y}^j \hat{e}_j$  tenemos que

$$XY = \left( \sum_{i=1}^m \hat{x}^i \hat{e}_i \right) \left( \sum_{j=1}^m \hat{y}^j \hat{e}_j \right).$$

Por la propiedad distributiva en a), del teorema T-12, así como por asociatividad, la ecuación anterior se convierte en

$$XY = \sum_{i=1}^m \sum_{j=1}^m \hat{x}^i \hat{y}^j \hat{e}_i \hat{e}_j.$$

Pero como la base  $\{\hat{e}_i\}_{i=1}^m$  es ortonormal la ecuación anterior, finalmente, se reduce a:

$$XY = \sum_{i=1}^m \hat{x}^i \hat{y}^i.$$

b) *Sobre matrices y transformaciones lineales.*

D-16 Rango de una matriz. Dada una matriz  $m$  por  $n$ , con elementos reales o complejos, para definir el rango de dicha matriz recurriremos a un espacio vectorial  $E_m$  arbitrario, y una base  $\{e_i\}_{i=1}^m$  del espacio  $E_m$ .

De la matriz  $A = [a_i^j]$  formaremos  $n$  vectores  $X_i$  así:

$$X_i = \sum_{j=1}^m a_i^j e_j, \quad i = 1, 2, \dots, n.$$

Es decir, los escalares de la columna  $i$ -ésima de la matriz  $m$  por  $n$  forman las componentes del vector  $X_i$  en la base

$$\{e_i\}_{i=1}^m.$$

Ya con estos vectores  $X_i$ , decimos que el rango de la matriz es el máximo número de vectores linealmente independientes que tenga el conjunto  $\{X_i\}_{i=1}^m$ .

D-17 El espacio vectorial  $M_m^2$ . Llamamos  $M_m^2$  al conjunto de todas las matrices cuadradas de orden  $m$ . Por el teorema 4 y el teorema 9 es claro que  $(M_m^2, R > 0)$   $(M_m^2, C > 0)$  es un espacio vectorial, con dimensión  $m^2$ .

D-18 Inversa de una matriz en  $M_m^2$ . Dada una matriz  $A \in M_m^2$  decimos que dicha matriz tiene inversa, representada por  $A^{-1}$ , si existe una matriz,  $A^{-1}$ , que cumple con:

$$AA^{-1} = I.$$

$I$  es la matriz identidad: matriz con ceros en todos sus elementos, excepto en la diagonal que tiene unos.

D-19 El espacio  $\mathcal{T}_m^2$ . Definimos a  $\mathcal{T}_m^2$  como el conjunto de todas las transformaciones lineales de un espacio  $E_m$  en  $E_m$ . El teorema 5 nos garantiza que dicho conjunto es un espacio vectorial.

De manera parecida al caso de matrices cuadradas, decimos que  $T^{-1}$  es la transformación lineal inversa de  $T$  si y sólo si  $TT^{-1} = I$ , con  $I$  la transformación lineal identidad:  $I(X) = X$ , para toda  $X \in E_m$ . Mostrar que  $I$  es transformación lineal resulta de la definición de transformación lineal y de la definición de  $I$ ; por otro lado, mostrar que  $T^{-1}$  es transformación lineal se prueba por la definición de  $T^{-1}$ ,

el caracter lineal de  $T$  y la condición  $TT^{-1} = I$ .

T-14 Isomorfismo entre  $\mathcal{T}_m$  y  $M_m$ . Sean los conjuntos:

$$\mathcal{T}_m = \left\{ T/T: E_m \rightarrow E_m, \text{ con } T \text{ transformación lineal.} \right\}$$

$$M_m = \left\{ A/A \text{ es matriz cuadrada de orden } m. \right\},$$

en donde el campo de escalares de  $E_m$  y de  $M_m$  es el mismo.

Los conjuntos  $\mathcal{T}_m$  y  $M_m$  son isomorfos.

*Demostración.*

Tomemos una base  $\{e_i\}_{i=1}^m$  de  $E_m$ , y luego para cada  $X \in E_m$  y  $T \in \mathcal{T}_m$  tendremos que:

$$1) \quad T(X) = T\left(\sum_{i=1}^m x^i e_i\right) = \sum_{i=1}^m x^i T(e_i),$$

en donde los  $x^i$  son las componentes de  $X$  en la base  $\{e_i\}_{i=1}^m$ .

De la ecuación 1) se deduce que, dada una base  $\{e_i\}_{i=1}^m$ , cada transformación de  $\mathcal{T}_m$  está caracterizada por los  $m$  vectores  $T(e_i)$ . Como cada uno de estos  $T(e_i)$  pertenecen al espacio  $E_m$ , entonces se pueden representar en la base  $\{e_i\}_{i=1}^m$  así:

$$2) \quad T(e_i) = \sum_{j=1}^m a_i^j e_j, \quad i=1, 2, \dots, m.$$

La ecuación 2) facilita que a cada  $T \in \mathcal{T}_m$  le podamos asociar una matriz, denominada  $A_T$ , dada por los  $a_i^j$ .

En forma explícita, la correspondencia entre  $\mathcal{T}_m$  y  $M_m$  que llamaremos  $\mathcal{L}$  viene definida así:

$$\mathcal{L}: \mathcal{T}_m \rightarrow M_m \text{ tal que } \mathcal{L}(T) = A_T,$$

con  $A_T = [a_i^j]$ , y los  $a_i^j$  dados vía la ecuación 2).

Vamos a caracterizar por medio de dicha matriz  $A_T$  la

correspondencia que hace la transformación  $T(X) = Y$ : por medio de  $A_T$  vamos a pasar de  $X$  a  $Y$ . Para lograr eso, tomemos a  $x^1, x^2, \dots, x^m$  &  $y^1, y^2, \dots, y^m$  como las componentes, en la base  $\{e_i\}_{i=1}^m$ , de los vectores  $X, Y$  respectivamente. Luego, basados en la ecuación 1) y 2)

$$T(X) = Y = \sum_{i=1}^m x^i T(e_i) = \sum_{i=1}^m x^i \left( \sum_{j=1}^m a_i^j e_j \right).$$

De la ecuación anterior se desprende que,

$$Y = \sum_{j=1}^m y^j e_j = \sum_{j=1}^m \left( \sum_{i=1}^m x^i a_i^j \right) e_j,$$

y de ahí finalmente:

$$3) \quad y^j = \sum_{i=1}^m x^i a_i^j, \text{ con } j=1, 2, \dots, m.$$

La ecuación 3) no es más que la ecuación de matrices:

$$4) \quad Y = A_T X,$$

en donde  $X, Y$  son vectores columna, cuyos elementos son las componentes de  $X, Y$  en la base  $\{e_i\}_{i=1}^m$ .

Los razonamientos anteriores conducen al importante resultado siguiente: Sea  $T$  transformación lineal de  $E_m$  en  $E_m$ , y  $\{e_i\}_{i=1}^m$  una base de  $E_m$ ; si  $A_T$  es la matriz asociada a  $T$  en la base  $\{e_i\}_{i=1}^m$ , entonces tenemos que

$$5) \quad T(X) = Y \text{ es equivalente a } A_T X = Y,$$

en donde  $X, Y \in E_m$  y también  $X, Y$  son vectores columna de componentes de  $X, Y$  en la base  $\{e_i\}_{i=1}^m$ .

Para probar el isomorfismo  $\mathcal{L}$ , lo haremos en tres pasos que denotaremos a), b) y c).

a) Aquí probaremos que la correspondencia es inyectiva (uno-uno). La demostración de esto es inmediata dado

que si  $A_{T_1}$  y  $A_{T_2}$ , matrices correspondientes a las transformaciones  $T_1$  y  $T_2$ , fueran iguales, entonces por el enunciado 5) resulta evidente que  $T_1 = T_2$ . Por lo tanto  $\mathcal{L}$  es uno-uno.

b)  $\mathcal{L}$  es sobreyectiva ya que a una matriz arbitraria  $A = [a_i^j]$  le podemos asociar una transformación  $T$  mediante la ecuación 1) y 2). Empero, faltaría probar que la  $T$  definida vía la ecuación 1) y 2) es, en realidad, una transformación lineal. Esto queda probado así:

$$\begin{aligned} T(X+Y) &= T\left(\sum_{i=1}^m (x^i + y^i) e_i\right), \\ &= \sum_{i=1}^m (x^i + y^i) T(e_i), \\ &= \sum_{i=1}^m x^i T(e_i) + \sum_{i=1}^m y^i T(e_i), \\ &= T(X) + T(Y). \end{aligned}$$

$$\begin{aligned} T(\alpha X) &= T\left(\alpha \sum_{i=1}^m x^i e_i\right), \\ &= T\left(\sum_{i=1}^m \alpha x^i e_i\right), \end{aligned}$$

por definición de  $T$  tenemos:

$$\begin{aligned} T(\alpha X) &= \sum_{i=1}^m \alpha x^i T(e_i), \\ &= \alpha \sum_{i=1}^m x^i T(e_i), \\ &= \alpha T(X). \end{aligned}$$

c) El isomorfismo estará completo al probar que la correspondencia  $\mathcal{L}$  es transformación lineal. Lo que hay que probar es que para dos transformaciones lineales arbitrarias  $T_1$  y  $T_2$  se cumple que

$$\mathcal{L}(T_1 + T_2) = \mathcal{L}(T_1) + \mathcal{L}(T_2).$$

$$\mathcal{L}(\alpha T_1) = \alpha \mathcal{L}(T_1).$$

La demostración es rutina, basada en la definición de  $\mathcal{L}$ .

T-15 Dimensión y espacios isomorfos. Dos espacios isomorfos, de dimensión finita, tienen la misma dimensión.

*Demostración.*

Sea  $T$  un isomorfismo de  $E$  en  $V$ , y  $\{e_i\}_{i=1}^m$  una base de  $E$ , o sea que la  $\dim(E) = m$ . Tenemos que demostrar que  $\dim(V) = m$ . Dado que  $T$  es sobreyectiva, para un  $Y$  arbitrario de  $V$  existe un  $X \in E$  tal que:

$$T(X) = Y.$$

Si tomamos las componentes  $x^1, x^2, \dots, x^m$  de  $X$  en la base  $\{e_i\}$  tendremos que la ecuación anterior se convierte en:

$$Y = T\left(\sum_{i=1}^m x^i e_i\right) = \sum_{i=1}^m x^i T(e_i).$$

Esta última ecuación muestra, claramente, que los  $m$ ,  $T(e_i)$ , generan al espacio  $V$ . Para que la demostración sea completa sólo nos falta mostrar que estos  $m$   $T(e_i)$  son linealmente in dependientes. Tomemos:

$$\sum_{i=1}^m \alpha^i T(e_i) = 0,$$

por ser  $T$  lineal  $T\left(\sum_{i=1}^m \alpha^i e_i\right) = 0$

dado que  $T$  es uno-uno,  $\sum_{i=1}^m \alpha^i e_i = 0$ ,

por ser  $\{e_i\}_{i=1}^m$  base de  $E$ ,  $\alpha^i = 0$ , con  $i = 1, 2, \dots, m$ . #

T-16 Dimensión de  $T_m^2$  y  $M_m^2$ . La dimensión de los espacios  $T_m^2$  y  $M_m^2$  es  $m^2$ .

*Demostración.*

Esta demostración es un corolario del T-4 y el T-15.

#### 1.4 Teoría de determinantes

La teoría de determinantes es importante; Shilov(1971: 5) hace el comentario de que la teoría de determinantes es la herramienta básica para estudiar sistemas lineales.

A. Definición de determinante. Dado un espacio vectorial  $E$ , con base  $\{e_i\}_{i=1}^m$ , y  $m$  vectores  $A_1, A_2, \dots, A_m$ . Postulamos una función  $D$

$$D: (A_1, A_2, \dots, A_m) \rightarrow F,$$

que cumpla con las propiedades siguientes:

- i.  $D(A_1, A_2, \dots, A_i, \dots, A_j, \dots, A_m) = -D(A_1, \dots, A_j, \dots, A_i, \dots, A_m)$ .
- ii.  $D(\alpha A_1, A_2, \dots, A_m) = \alpha D(A_1, A_2, \dots, A_m)$ , con  $\alpha$  escalar.
- iii.  $D(B+C, A_2, \dots, A_m) = D(B, A_2, \dots, A_m) + D(C, A_2, \dots, A_m)$ .
- iv.  $D(e_1, e_2, \dots, e_m) = 1$ .

Esta función  $D$ , la llamaremos determinante. Observe que por el momento  $D$ , aparentemente, depende de la base escogida para el espacio  $E_m$ . También es claro que hay que probar que dicha función determinante existe y es única. La unicidad la probaremos en el inciso D., mientras que la existencia no la demostraremos, dado que su prueba no conduce a nuevos conocimientos que se vayan a requerir en nuestra futura teoría. Dicha demostración se puede encontrar en Lichnerowicz (1967:26-29).

Por ahora, lo primero que haremos sera enunciar unas propiedades básicas de los determinantes, en el supuesto de que dicha función existe y es única. La mayoría de las de-

mostraciones, de dichos enunciados, son muy sencillas y por lo tanto se omiten.

B. Propiedades básicas del determinante.

$$P-1 \quad D(A_1, A_2, \dots, \alpha A_j, \dots, A_m) = \alpha D(A_1, A_2, \dots, A_j, \dots, A_m).$$

$$P-2 \quad D(A_1, A_2, \dots, \sum_{i=1}^m \alpha^i A_j^i, \dots, A_m) = \sum_{i=1}^m D(A_1, A_2, \dots, A_j^i, \dots, A_m).$$

$$P-3 \quad D\left(\sum_{v_1=1}^m \alpha_{v_1} A_1^{v_1}, \dots, \sum_{v_m=1}^m \alpha_{v_m} A_m^{v_m}\right) = \sum_{v_m=1}^m \dots \sum_{v_1=1}^m \alpha_{v_1} \dots \alpha_{v_m} D(A_1, \dots, A_m).$$

$$P-4 \quad D(A_1, A_2, \dots, A_i, \dots, A_i, \dots, A_m) = 0.$$

$$P-5 \quad D(A_1, A_2, \dots, 0, \dots, A_m) = 0.$$

$$P-6 \quad D(A_1, A_2, \dots, A_j, \dots, A_m) = D(A_1, A_2, \dots, B_j, \dots, A_m),$$

$$\text{en donde } B_j = A_j + \sum_{(i \neq j)=1}^m \alpha^i A_i.$$

C. Determinantes y linealidad. El teorema de esta sección busca la relación esencial que existe entre determinantes e independencia lineal. Este teorema viene a ser:

Teorema sobre determinantes y linealidad.

i. Si los vectores  $A_1, A_2, \dots, A_m$  son linealmente dependientes entonces  $D(A_1, A_2, \dots, A_m) = 0$ .

ii. Los vectores  $A_1, A_2, \dots, A_m$  son linealmente independientes si y sólo si  $D(A_1, A_2, \dots, A_m) \neq 0$ .

*Demostración.*

Para demostrar i. procedemos así: dado que  $A_1, A_2, \dots, A_m$  son linealmente dependientes existe  $A_j$  tal que:

$$A_j = \sum_{(i \neq j)=1}^m \alpha^i A_i; \text{ de donde}$$

$$D(A_1, A_2, \dots, A_j, \dots, A_m) = D(A_1, A_2, \dots, \sum_{(i \neq j)=1}^m \alpha^i A_i, \dots, A_m) \\ = \sum_{(i \neq j)=1}^m \alpha^i D(A_1, A_2, \dots, A_i, \dots, A_m).$$

Finalmente, por la propiedad 4, la ecuación anterior nos lleva a que:

$$D(A_1, A_2, \dots, A_m) = 0.$$

La demostración de ii. sólo requiere probar, a causa de lo ya mostrado en i., que si los  $A_1, A_2, \dots, A_m$  son linealmente independientes entonces  $D(A_1, A_2, \dots, A_m) \neq 0$ . Dado que los  $A_1, A_2, \dots, A_m$  son linealmente independientes tenemos que existen los  $a_i^{k_i}$  tal que:

$$e_i = \sum_{k_i=1}^m a_i^{k_i} A_{k_i}, \quad k_i = 1, 2, \dots, m.$$

Luego formamos el determinante,

$D(e_1, e_2, \dots, e_m) = D\left(\sum_{k_1=1}^m a_1^{k_1} A_{k_1}, \sum_{k_2=1}^m a_2^{k_2} A_{k_2}, \dots, \sum_{k_m=1}^m a_m^{k_m} A_{k_m}\right)$ , y por la propiedad 3, la ecuación anterior implica:

$$6) \quad D(e_1, e_2, \dots, e_m) = \sum_{k_1=1}^m \sum_{k_2=1}^m \dots \sum_{k_m=1}^m a_1^{k_1} a_2^{k_2} \dots a_m^{k_m} D(A_{k_1}, A_{k_2}, \dots, A_{k_m}).$$

En esta última ecuación tenemos que cada determinante,  $D(A_{k_1}, A_{k_2}, \dots, A_{k_m})$ , en el cual los subíndices no sean una permutación de la sucesión  $1, 2, \dots, m$ , es igual a cero, por la propiedad 4. Además, por la propiedad i., que define a los determinantes, se tiene que los determinantes en la ecuación 6) cumplen con que  $D(A_{k_1}, A_{k_2}, \dots, A_{k_m}) = \pm D(A_1, A_2, \dots, A_m)$ , dado que los que no forman permutación de  $1, 2, \dots, m$  son cero.

Ya con esto está claro que si se diera el caso de que  $D(A_1, A_2, \dots, A_m) = 0$ , entonces tendríamos que:

$$D(e_1, e_2, \dots, e_m) = 0,$$

lo cual está en contradicción con la propiedad iv. #

D. Unicidad del determinante & representación cuadrada.

Sean  $m$  vectores  $A_j$ ,  $j=1, 2, \dots, m$ , expresados en términos de sus componentes en la base  $\{e_i\}_{i=1}^m$ , del espacio  $E_m$ , como:

$$A_j = \sum_{k_i=1}^m a_j^{k_i} e_{k_i}, \quad j=1, 2, \dots, m.$$

El determinante de los  $m$  vectores  $A_j$  viene dado así:

$D(A_1, A_2, \dots, A_m) = D\left(\sum_{k_1=1}^m a_1^{k_1} e_{k_1}, \sum_{k_2=1}^m a_2^{k_2} e_{k_2}, \dots, \sum_{k_m=1}^m a_m^{k_m} e_{k_m}\right)$ , y por la propiedad P-3 la ecuación anterior se puede expresar

$$D(A_1, A_2, \dots, A_m) = \sum_{k_m=1}^m \dots \sum_{k_1=1}^m a_1^{k_1} a_2^{k_2} \dots a_m^{k_m} D(e_{k_1}, e_{k_2}, \dots, e_{k_m}).$$

Por la propiedad P-4, tenemos que los  $D(e_{k_1}, e_{k_2}, \dots, e_{k_m})$  que no son iguales a cero son aquellos en donde  $k_1, k_2, \dots, k_m$  es una permutación de  $1, 2, \dots, m$ ; es decir todos los  $k_1, k_2, \dots, k_m$  son diferentes entre sí. A causa del razonamiento anterior podemos plantear la siguiente ecuación:

$$7) \quad D(A_1, A_2, \dots, A_m) = \sum_{k_m=1}^m \dots \sum_{k_1=1}^m a_1^{k_1} a_2^{k_2} \dots a_m^{k_m} (-1)^{I(k_1, \dots, k_m)} D(e_1, e_2, \dots, e_m).$$

$I(k_1, k_2, \dots, k_m)$  es el número total de inversiones de  $k_1, k_2, \dots, k_m$  (dos términos de una sucesión forman una inversión si el más grande aparece a la izquierda del más pequeño.).

El papel de las inversiones en la ecuación 7) conduce, junto con la propiedad i. de determinantes, a que se pueda pasar de  $D(e_{k_1}, e_{k_2}, \dots, e_{k_m})$  a  $(-1)^{I(k_1, k_2, \dots, k_m)} D(e_1, \dots, e_m)$ .

Luego, por la propiedad iv., llegamos de la ecuación 7) a la siguiente ecuación:

$$8) \quad D(A_1, \dots, A_m) = \sum_{k_1=1}^m \dots \sum_{k_m=1}^m a_1^{k_1} a_2^{k_2} \dots a_m^{k_m} (-1)^{I(k_1, k_2, \dots, k_m)}$$

Esta ecuación 8) nos prueba que si el determinante existe, es decir satisface las cuatro propiedades que lo caracterizan, entonces es una función única, definida por la ecuación 8).

El teorema de existencia, que no se demostrará, consiste en probar que la función, definida por medio de la ecuación 8), satisface las cuatro propiedades dadas para caracterizar un determinante.

La ecuación 8) también nos muestra que un determinante sólo depende de los escalares  $a_i^{k_i}$ , en productos de  $m$  factores cada uno. Es por ésto que un determinante se puede representar como un arreglo cuadrado de números. Explícitamente la representación del determinante, dado por la ecuación 8), viene así:

$$\Delta = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_m^1 \\ a_1^2 & a_2^2 & \dots & a_m^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_m^m \end{vmatrix} .$$

En esta representación cuadrada, los  $m$  escalares de cada columna, por ejemplo la  $j$ , corresponden a las componentes del vector  $A_j$ , en la base  $\{e_i\}_{i=1}^m$ .

E. Expansión del determinante en términos de una columna

El resultado de esta sección es necesario para el procedimiento eficaz de evaluación numérica de un determinante, procedimiento que se desarrollará en forma completa en el capítulo III. La ecuación 8), de la sección anterior, nos da un método para evaluar numéricamente un determinante. La gran desventaja, desde el punto de vista numérico, es la enorme cantidad de operaciones que emplea para calcular el determinante.

Dados los  $m$  vectores  $A_k$ ,  $k = 1, 2, \dots, m$ , su determinante lo calcularemos apoyados sobre una columna dada, por ejemplo la columna  $j$ .

$$9) \quad D(A_1, A_2, \dots, A_m) = D(A_1, A_2, \dots, A_j, \dots, A_m) \\ + \sum_{i=1}^m a_j^i D(A_1, A_2, \dots, A_{j-1}, e_i, A_{j+1}, \dots, A_m).$$

En la ecuación anterior el vector  $A_j = \sum_{i=1}^m a_j^i e_i$ , y el vector  $e_i$  aparecen en la posición  $j$ -ésima.

Cada uno de los  $m$  determinantes que aparecen en la ecuación 9) lo llamaremos un cofactor de  $a_j^i$ , y lo representaremos así:

$$10) \quad \alpha_i^j = D(A_1, A_2, \dots, A_{j-1}, e_i, A_{j+1}, \dots, A_m).$$

La expansión del determinante en términos de la columna  $j$ -ésima viene dada como:

$$11) \quad D(A_1, A_2, \dots, A_m) = \sum_{i=1}^m \alpha_i^j a_j^i.$$

Nuestro interés será, ahora, calcular los cofactores  $\alpha_i^j$ ,  $i = 1, 2, \dots, m$ , en términos de las componentes de los  $m-1$  vectores  $A_k$ , ( $k \neq j$ ) =  $1, 2, \dots, m$ .

Pasar el determinante  $D(A_1, A_2, \dots, A_{j-1}, e_i, A_{j+1}, \dots, A_m)$ , que aparece en la ecuación 10), al determinante

$$D(e_i, A_1, A_2, \dots, A_{j-1}, A_{j+1}, \dots, A_m),$$

requiere  $j-1$  cambios de signos de  $+$  a  $-$ , de acuerdo a la propiedad  $i.$ , al desplazarse  $e_i$ , paso a paso, desde la posición  $j$ -ésima a la primera posición. O sea que, de acuerdo a lo anterior tenemos la siguiente ecuación:

$$12) \quad D(A_1, A_2, \dots, A_{j-1}, A_{j+1}, \dots, A_m) = (-1)^{j-1} D(e_i, A_1, A_2, \dots, A_{j-1}, A_{j+1}, \dots, A_m).$$

La representación cuadrada del determinante

$$D(e_i, A_1, A_2, \dots, A_{j-1}, A_{j+1}, \dots, A_m)$$

viene dada así:

$$\begin{vmatrix} 0 & a_1^1 & a_2^1 & \dots & a_{j-1}^1 & a_{j+1}^1 & \dots & a_m^1 \\ 0 & a_1^2 & a_2^2 & \dots & a_{j-1}^2 & a_{j+1}^2 & \dots & a_m^2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a_1^i & a_2^i & \dots & a_{j-1}^i & a_{j+1}^i & \dots & a_m^i \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_1^m & a_2^m & \dots & a_{j-1}^m & a_{j+1}^m & \dots & a_m^m \end{vmatrix} .$$

Observe, en el determinante anterior, que en la primera columna el único escalar distinto de cero e igual a uno es el que aparece en la  $i$ -ésima fila.

Si nos guiamos por la ecuación 8), de la sección anterior, tenemos que el determinante de la ecuación 12), en el

lado derecho queda así:

$$D(e_i, A_1, A_2, \dots, A_{j-1}, A_{j+1}, \dots, A_m) = \sum_{k_m=1}^m \dots \sum_{k_1=1}^m \sum_{k_2=1}^m e_i^{k_1} a_1^{k_2} a_2^{k_3} \dots a_{j-1}^{k_j} a_{j+1}^{k_{j+1}} \dots a_m^{k_m} (-1)^{I(k_1, k_2, \dots, k_m)}.$$

En la ecuación anterior desaparecen todos los términos en donde  $k_1 \neq i$ , es decir el lado derecho de dicha ecuación, queda como:

$$\sum_{k_m=1}^m \dots \sum_{k_2=1}^m e_i^{k_1} a_1^{k_2} a_2^{k_3} \dots a_{j-1}^{k_j} a_{j+1}^{k_{j+1}} \dots a_m^{k_m} (-1)^{I(i, k_2, \dots, k_m)}.$$

Dado que  $e_i^i = 1$ , lo anterior se reduce a:

$$13) \sum_{k_m=1}^m \dots \sum_{k_2=1}^m a_1^{k_2} a_2^{k_3} \dots a_{j-1}^{k_j} a_{j+1}^{k_{j+1}} \dots a_m^{k_m} (-1)^{I(i, k_2, \dots, k_m)}.$$

Observe que en 13) no aparecen los escalares de la columna  $j$ -ésima, ni los de la fila  $i$ -ésima.

La siguiente ecuación es clara de por sí:

$$I(i, k_2, \dots, k_m) = (i-1) + I(k_2, k_3, \dots, k_m).$$

Con el resultado anterior 13) pasa a la forma:

$$(-1)^{i-1} \sum_{k_m=1}^m \dots \sum_{k_2=1}^m a_1^{k_2} a_2^{k_3} \dots a_{j-1}^{k_j} a_{j+1}^{k_{j+1}} \dots a_m^{k_m} (-1)^{I(k_2, k_3, \dots, k_m)}.$$

Y de ahí:

$$14) D(e_i, A_1, A_2, \dots, A_{j-1}, A_{j+1}, \dots, A_m) = (-1)^{i-1} \sum_{k_m=1}^m \dots \sum_{k_2=1}^m a_1^{k_2} a_2^{k_3} \dots a_{j-1}^{k_j} a_{j+1}^{k_{j+1}} \dots a_m^{k_m} (-1)^{I(k_2, k_3, \dots, k_m)}.$$

En la ecuación 14) llamaremos menor de  $a_j^i$ , representado por  $D_i^j$ , al determinante:

$$\sum_{k_m=1}^m \dots \sum_{k_2=1}^m a_1^{k_2} a_2^{k_3} \dots a_{j-1}^{k_j} a_{j+1}^{k_{j+1}} \dots a_m^{k_m} (-1)^{I(k_2, k_3, \dots, k_m)}.$$

Nótese que en el menor  $D_i^j$  se ha eliminado la fila  $i$ -ésima y la columna  $j$ -ésima. Tomando en cuenta el resultado de la ecuación 14) y la notación introducida para el menor, tendremos que la ecuación 12) queda como:

$$\begin{aligned}
 D(A_1, A_2, \dots, A_{j-1}, A_{j+1}, \dots, A_m) &= (-1)^{j-1} (-1)^{i-1} D_i^j, \\
 &= (-1)^{i+j} (-1)^{-2} D_i^j, \\
 &= (-1)^{i+j} D_i^j.
 \end{aligned}$$

Por la ecuación 10), que define a los cofactores, tenemos que estos cofactores vienen dados por:

$$\alpha_i^j = (-1)^{i+j} D_i^j.$$

La ecuación anterior realiza el objetivo de calcular los cofactores en términos de las componentes  $A_k$ , ( $k \neq j$ ) = 1, 2, ..., m. Además, y más importante, la expansión del determinante original  $D(A_1, A_2, \dots, A_m)$ , en términos de la  $j$ -ésima columna ha quedado finalmente como:

$$15) \quad D(A_1, A_2, \dots, A_m) = \sum_{i=1}^m (-1)^{i+j} a_{ij}^i D_i^j,$$

para cualquier columna  $j = 1, 2, \dots, m$ .

El resultado de la ecuación 15) será muy importante para la evaluación numérica de un determinante, la cual se verá en el capítulo III.

Nótese que en esta sección trabajamos con la variable  $j$  como variable de columna, mientras que la  $i$  como variable de fila.

F. Determinante del producto de matrices. Sean  $A, B$  dos matrices cuadradas,  $m$  por  $m$ , entonces tenemos que:

$$D(AB) = D(A)D(B).$$

*Demostración.*

La demostración la omitiremos; sin embargo, no ofrece ningún problema particular. Puede consultarse en Lichner-

wicz (1967:61).

G. Traspuesta de A y  $D(A) = D(A^t)$ . En general si tenemos una matriz  $A = [a_i^j]$ , m por n, entonces llamamos matriz traspuesta de A, denotada por  $A^t$ , a la nueva matriz n por m definida así:

$$A^t = [{}^t a_j^i], \text{ con } {}^t a_j^i = a_i^j, \quad \begin{matrix} j = 1, 2, \dots, m \\ i = 1, 2, \dots, n \end{matrix}.$$

La propiedad  $D(A) = D(A^t)$  se demuestra basados en la definición de traspuesta de una matriz y la definición de determinante, como en la ecuación 8).

## CAPITULO II

### 2.1 Definición de un sistema lineal

La representación usual de un sistema de ecuaciones lineales, que denotaremos por  $L_{mn}$ , la conocemos como el siguiente arreglo de ecuaciones:

$$\begin{array}{l} a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n = b^1 \\ a_1^2 x^1 + a_2^2 x^2 + \dots + a_n^2 x^n = b^2 \\ 1) \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_1^m x^1 + a_2^m x^2 + \dots + a_n^m x^n = b^m \end{array}$$

Le llamamos  $L_{mn}$  por tener  $m$  ecuaciones y  $n$  columnas con coeficientes  $x^i$ ; los coeficientes  $x^i$  comúnmente se llaman incógnitas.

El sistema  $L_{mn}$  ofrece tres problemas que clasificaremos y definiremos así:

Problema 1. Se dan los coeficientes  $a_i^j$ , los cuales forman una matriz  $A = [a_i^j]$  asociada al sistema  $L_{mn}$ ; y se dan los coeficientes  $b^j$ , entonces hay que encontrar las llamadas incógnitas  $x^1, x^2, \dots, x^n$ . Este es el problema más frecuente asociado a un sistema  $L_{mn}$ . Cuando, de aquí en adelante, no se especifique qué se entiende por resolver un sistema  $L_{mn}$  deberá asumirse que se refiere a este problema 1. Además es este problema el que presenta interés para su resolución numérica, la cual se trabajará en el capítulo III.

Problema 2. En este caso se da la matriz asociada  $A = [a_i^j]$ , y los coeficientes  $x^1, x^2, \dots, x^n$ , entonces hay que encontrar los  $b^1, b^2, \dots, b^m$ . La solución numérica para este caso es obvia.

Problema 3. Para el arreglo 1), se dan  $m$  conjuntos  $x^1, x^2, \dots, x^n$  y  $m$  conjuntos  $b^1, b^2, \dots, b^m$ , entonces hay que encontrar la matriz  $A = [a_i^j]$ . Este caso plantea un problema numérico reducible al problema 1. Un ejemplo, de este caso, lo tendremos en el problema de encontrar la inversa de una matriz (capítulo III).

## 2.2 Interpretaciones equivalentes de un sistema $L_{mn}$

Un aspecto que le da importancia a un sistema  $L_{mn}$  es la variedad de formas equivalentes de interpretación, por un lado, mientras que por otro la relativa simplicidad de resolución numérica. En esta sección veremos algunas formas de interpretación de un sistema  $L_{mn}$ .

a) Interpretación matricial. Si basados en 1) definimos las siguientes matrices:

$$A = [a_i^j], \quad X = \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} b^1 \\ \vdots \\ b^m \end{bmatrix},$$

tendremos que la ecuación,

$$2) \quad AX = B.$$

es equivalente al sistema  $L_{mn}$ . Equivalente quiere decir, en este caso, que reproduce las mismas ecuaciones, Más claro,

tanto 1) como 2) quedan definidas por la simple ecuación:

$$3) \quad \sum_{i=1}^n a_i^j x^i = b^j, \quad j=1,2,\dots,m.$$

El sistema  $L_{mn}$ , visto por medio de la ecuación 2) o 3), tendrá las siguientes ventajas:

- i. Brevedad de notación que conduce a rápido manejo operacional.
- ii. El sistema  $L_{mn}$  tiene un acceso más directo para una máquina computadora: se le da a la máquina, según sea el caso la matriz  $A$ ,  $B$  o  $X$ .
- iii. La más importante, quizá, de todas las ventajas de la ecuación 2) o 3) es su íntima vinculación con una transformación lineal. Esta vinculación la resumiremos en el siguiente teorema.

#### Teorema 2.2

- i. Todo sistema  $L_{mn}$  induce a una transformación lineal de un espacio  $E_n$  en  $E_m$ .
- ii. Toda transformación lineal  $T$  de  $E_n$  en  $E_m$  induce a un sistema  $L_{mn}$ .

También podemos decir que una transformación lineal  $T$  funciona como un sistema  $L_{mn}$ , vía ecuación 2) o 3); y a la inversa, un sistema  $L_{mn}$  como una transformación  $T$ .

#### Demostración.

Para el caso de que el sistema  $L_{mn}$  tenga asociada una matriz cuadrada, este teorema no es más, excepto por nuevos comentarios, que el isomorfismo dado en el teorema 14, capí

tulo I.

i. Supongamos que tenemos un sistema  $L_{mn}$  dado por 2),

$$4) \quad AX=Y,$$

en donde, para mayor claridad de lo que se a exponer, hemos tomado  $Y$  en lugar de  $B$  en la ecuación 2). La ecuación 4) funcionará como el problema 2), dado en la sección 2.1: se dispone de la matriz  $A$  y  $X$ , entonces podemos decir que si entra un vector columna  $X$ , por medio de la matriz  $A$ , obtendremos un vector columna  $Y$ ; similar a que si entraran los escalares  $x^1, x^2, \dots, x^n$ , entonces, por medio de la matriz  $A$ , obtenemos los  $y^1, y^2, \dots, y^m$ .

Reducir la ecuación 4) a que funcione como una transformación  $T$  de  $E_n$  en  $E_m$ , necesita que introduzcamos dos espacios arbitrarios  $E_n$  y  $E_m$  y dos bases  $\{e_i\}_{i=1}^n$ ,  $\{e_j\}_{j=1}^m$  de  $E_n$  y  $E_m$  respectivamente. Luego, la transformación  $T$  viene definida de la siguiente manera:

$$5) \quad T(e_i) = \sum_{j=1}^m a_{ij} e_j, \text{ y si } X = \sum_{i=1}^n x^i e_i \text{ entonces} \\ T(X) = \sum_{i=1}^n x^i T(e_i), \text{ para cada } X \in E_n.$$

La función  $T$ , dada en 5), está bien definida ya que conocemos o disponemos de la matriz  $A$ . Demostrar que  $T$  es, de hecho, una transformación lineal procede en la misma forma que se hizo la demostración de linealidad de  $T$  en el teorema 14, capítulo I.

Lo único que quedaría por mostrar es que  $T$  funciona como la ecuación 4), es decir, que  $T$  asocia a los escalares

$x^1, x^2, \dots, x^n$  los escalares  $y^1, y^2, \dots, y^m$ . Esto se prueba basados en la ecuación 5) así:

$$T(X) = \sum_{i=1}^n x^i T(e_i) = \sum_{i=1}^n x^i \left( \sum_{j=1}^m a_{ij}^j e_j \right),$$

$$\sum_{j=1}^m \left( \sum_{i=1}^n a_{ij}^j x^i \right) e_j.$$

Suponiendo que  $T(X) = \sum_{j=1}^m z^j e_j$ , tenemos que la ecuación anterior se reduce a

$$6) \quad z^j = \sum_{i=1}^n a_{ij}^j x^i, \quad j=1, 2, \dots, m$$

La ecuación 6) no es más que  $Z=AX$  con  $Z = \begin{bmatrix} z^1 \\ \vdots \\ z^m \end{bmatrix}$ . Comparando 6) con 4) se deduce que:

$$Z = Y.$$

De esta forma la transformación lineal pasa de los  $x^1, x^2, \dots, x^n$  a los  $y^1, y^2, \dots, y^m$ .

ii. Demostrar que una transformación  $T$  de  $E_n$  en  $E_m$  induce a un sistema  $L_{mn}$  que funciona como  $T$  se logra, sencillamente, al definir la matriz  $A$  por medio de 5). Luego  $T$  nos llevará a la ecuación 6) que no es más que  $AX=Y$ . #

Es importante notar que la identificación de un sistema  $L_{mn}$  con una transformación lineal  $T$  de  $E_n$  en  $E_m$  descansa, en su esencia, sobre la disponibilidad de una matriz  $A$ ,  $m$  por  $n$ ; y un par de bases para  $E_n$  y  $E_m$ . Ya con  $A$  y las bases de  $E_n$  y  $E_m$ , el teorema 2.2 se puede expresar así:

$$7) \quad AX=Y \text{ si y sólo si } T(X)=Y.$$

b) Interpretación vectorial. Dado el sistema  $L_{mn}$ , como lo presenta la ecuación 1), si introducimos un espacio vectorial arbitrario  $E_m$ , con base  $\{e_i\}_{i=1}^m$ , entonces

ces podemos formar  $n$  vectores  $T_i \in E_m$ , dados así:

$$T_i = \sum_{j=1}^m a_i^j e_j, \quad i=1, 2, \dots, n.$$

Las componentes de  $T_i$  son los elementos de la columna  $i$ -ésima de la matriz  $A$  asociada al sistema  $L_{mn}$ . Ya con ésto, 1) es equivalente a la siguiente ecuación vectorial:

$$8) \quad \sum_{i=1}^n x^i T_i = B, \quad \text{con } B = \sum_{j=1}^m b^j e_j.$$

La equivalencia de la ecuación 8) con la 1) se muestra basados en la definición de los  $T_i$  y el vector  $B$ . Explicitamente  $\sum_{j=1}^m b^j e_j = B = \sum_{i=1}^n x^i T_i = \sum_{i=1}^n x^i \left( \sum_{j=1}^m a_i^j e_j \right) = \sum_{i=1}^n \left( \sum_{j=1}^m x^i a_i^j \right) e_j$ , de donde se concluye que:

$$b^j = \sum_{i=1}^n x^i a_i^j, \quad j=1, 2, \dots, m.$$

La ventaja de la interpretación vectorial consiste en poder desarrollar la teoría del sistema  $L_{mn}$  de una forma natural. Observe que la interpretación vectorial conduce de nuevo al teorema 2.2 con el simple hecho de definir los  $T_i = T(e_i)$ .

c) Interpretación con funcionales lineales. Dado un sistema  $L_{mn}$ , con su matriz asociada  $A$ , y un espacio  $E_n$  arbitrario, con base  $\{e_i\}_{i=1}^n$ . Definimos  $m$  funcionales lineales así:

$$9) \quad \begin{aligned} F^j(e_i) &= a_i^j, \quad j=1, 2, \dots, m \text{ \& } i=1, 2, \dots, n; \\ \text{si } X &= \sum_{i=1}^n x^i e_i, \text{ entonces } F^j(X) = \sum_{i=1}^n x^i F^j(e_i), \\ &\text{para cada } X \in E_n \end{aligned}$$

Demostrar que cada  $F^j$  es funcional lineal es pura rutina. En esta interpretación, resolver el sistema  $L_{mn}$  consis-

te en encontrar aquellos  $X \in E_n$  tal que:

$$F^j(X) = b^j, \text{ con } j = 1, 2, \dots, m.$$

Lo anterior queda demostrado por el hecho de que si

$$F^j(X) = b^j, \quad j = 1, 2, \dots, m \text{ entonces}$$

$$F^j\left(\sum_{i=1}^n x^i e_i\right) = b^j, \quad j = 1, 2, \dots, m,$$

$$\sum_{i=1}^n x^i F^j(e_i) = b^j, \quad j = 1, 2, \dots, m \text{ y tal como}$$

están definidos los  $F^j$  en la ecuación 9) tenemos finalmente que:

$$\sum_{i=1}^n x^i a_i^j = b^j, \quad j = 1, 2, \dots, m.$$

La interpretación de funcionales lineales, tal y como se planteó aquí, se dirige específicamente al problema 1, dado en la sección 2.1.

### 2.3 Una solución teórica del sistema $L_{mn}$

Trabajaremos, para un sistema  $L_{mn}$ , el problema 1 de la sección 2.1. Vamos a partir de la ecuación 8), es decir tomaremos  $L_{mn}$  desde el punto de vista de la interpretación vectorial.

Tomemos un espacio  $E_m$  arbitrario con una base  $\{e_j\}_{j=1}^m$ . Luego, definimos los vectores  $T_i = \sum_{j=1}^m a_i^j e_j$ ,  $i = 1, 2, \dots, n$ , tal y como se hizo en la interpretación vectorial.

Llamaremos  $r$  al rango de  $L_{mn}$  (rango es el mayor número de vectores linealmente independientes del conjunto  $\{T_i\}_{i=1}^n$ ); y mediante un simple reordenamiento de subíndices podemos hacer que los primeros  $r$   $T_1, T_2, \dots, T_r$  sean los vectores li-

nealmente independientes del conjunto  $\{T_i\}_{i=1}^n$ . Como consecuencia de lo anterior tenemos la siguiente ecuación:

existen los  $\alpha_{r+\varepsilon}^h$  tal que,

$$10) \quad T_{r+\varepsilon} = \sum_{h=1}^r \alpha_{r+\varepsilon}^h T_h, \text{ con } \varepsilon = 1, 2, \dots, n-r.$$

Es decir, los  $T_{r+1}, T_{r+2}, \dots, T_n$  son combinación lineal de los vectores linealmente independientes  $T_1, T_2, \dots, T_r$ .

Llamaremos  $E_r$  al subespacio de  $E_n$  generado por los  $T_1, T_2, \dots, T_r$ ; y, es claro, que el sistema  $L_{mn}$  tiene solución si  $B$  está en  $E_r$ . Lo anterior se explica porque la ecuación 8) puede expresarse como combinación lineal de los  $r$   $T_i$  linealmente independientes. En resumen, si existe solución para  $L_{mn}$   $B$  pertenece a  $E_r$ , y entonces existen los  $x^1, x^2, \dots, x^n$  que satisfacen la ecuación 8). La siguiente ecuación es consecuencia inmediata de que  $B \in E_r$ :

$$11) \quad \text{existen } r \text{ } b^h \text{ tal que } B = \sum_{h=1}^r b^h T_h.$$

Para la solución teórica que buscamos combinamos las ecuaciones 8), 10) y 11) de la siguiente manera: de la ecuación 8),  $B = \sum_{i=1}^n x^i T_i$ , separamos los linealmente independientes de los que no lo son, y tenemos

$$B = \sum_{h=1}^r x^h T_h + \sum_{\varepsilon=1}^{n-r} x^{r+\varepsilon} T_{r+\varepsilon}.$$

Luego sustituimos, en la ecuación anterior, las ecuaciones 10) y 11) para obtener:

$$\begin{aligned} \sum_{h=1}^r b^h T_h &= \sum_{h=1}^r x^h T_h + \sum_{\varepsilon=1}^{n-r} x^{r+\varepsilon} \left( \sum_{h=1}^r \alpha_{r+\varepsilon}^h \right) T_h, \\ \sum_{h=1}^r b^h T_h &= \sum_{h=1}^r \left( x^h + \sum_{\varepsilon=1}^{n-r} x^{r+\varepsilon} \alpha_{r+\varepsilon}^h \right) T_h, \text{ y luego} \\ b^h &= x^h + \sum_{\varepsilon=1}^{n-r} x^{r+\varepsilon} \alpha_{r+\varepsilon}^h, \quad h = 1, 2, \dots, r. \end{aligned}$$

De la ecuación anterior, al despejar  $x^h$ , llegamos, finalmente, a lo que buscábamos:

$$12) \quad x^h = \delta^h - \sum_{\epsilon=1}^{n-r} x^{r+\epsilon} \alpha_{r+\epsilon}^h, \quad h=1, 2, \dots, r.$$

La solución teórica dada en 12) nos indica que de los  $n$   $x^i$ , los últimos  $n-r$ ,  $x^{r+1}, x^{r+2}, \dots, x^n$ , se pueden escoger arbitrariamente; y esta selección determina los primeros  $r$ ,  $x^1, x^2, \dots, x^r$ .

La gran desventaja de la solución 12), desde el punto de vista numérico, radica en que hay que determinar, primero, los  $r$  vectores linealmente independientes de los  $T_i$ , y también encontrar los escalares  $\alpha_{r+\epsilon}^h$  y  $\delta^h$ ; todo esto entorpece y descarta a 12) como solución numérica. Sin embargo, la ecuación 12) es el punto de partida para importantes resultados teóricos que, a la postre, serán de gran utilidad para la resolución numérica.

#### 2.4 El sistema homogéneo, $H(L_{mn})$ , y solución vectorial de $L_{mn}$ .

La introducción de un sistema lineal homogéneo, asociado al sistema  $L_{mn}$  y denotado por  $H(L_{mn})$ , tiene el propósito, en esta sección, de caracterizar una solución vectorial del sistema  $L_{mn}$ . Este objetivo se realizará basados en la ecuación 12) de la sección anterior.

El sistema homogéneo,  $H(L_{mn})$ , asociado a  $L_{mn}$ , lo definiremos como:

$$13) \quad \sum_{i=1}^n x^i T_i = 0,$$

en donde  $L_{mn}$  viene considerado según la ecuación 8). Si el sistema  $L_{mn}$  viniera considerado según la ecuación 3), entonces el sistema homogéneo quedaría definido como:

$$\sum_{i=1}^n x^i a_i^j = 0, \quad j=1,2,\dots,m.$$

Es claro que el sistema homogéneo siempre tiene, al menos una solución, a saber: cuando los  $n$   $x^i$  resultan ser todos iguales a cero. Esta solución la llamamos solución trivial.

Para el caso de un sistema homogéneo tenemos que los  $r$  vectores  $b^h$ , de la ecuación 11), son todos cero, a causa de que  $B=0$ , y que los  $T_1, T_2, \dots, T_r$  son linealmente independientes. Considerando lo anterior, la ecuación 12), en el caso homogéneo, pasa a ser:

$$14) \quad x^h = -\sum_{\epsilon=1}^{n-r} x^{r+\epsilon} \alpha_{r+\epsilon}^h, \quad h=1,2,\dots,r.$$

Caracterizar una solución vectorial del sistema  $L_{mn}$  requiere introducir un espacio vectorial  $E_n$  arbitrario y una base  $\{\epsilon_i\}_{i=1}^n$  de  $E_n$ . Con dicha base de  $E_n$  podemos definir lo que consideraremos un vector solución así:

$$15) \quad X = \sum_{i=1}^n x^i \epsilon_i \in E_n \text{ es vector solución si y sólo si } x^1, x^2, \dots, x^n \text{ es solución de } L_{mn}.$$

Ahora, el primer objetivo, para caracterizar una solución vectorial de  $L_{mn}$ , es encontrar todos los vectores solución del sistema homogéneo. Este primer objetivo lo lograremos considerando la ecuación 14) y la siguiente matriz:



Teorema 2.4.a

El subespacio vectorial  $E_{n-r}$  de  $E_n$  generado por los  $X_p$ ,  $p=1,2,\dots,n-r$  es el conjunto de todos los vectores solución del sistema homogéneo  $H(L_{mn})$ .

Por lo tanto la dimensión del conjunto solución de un sistema homogéneo es  $n-r$ , con  $r$  rango del sistema  $L_{mn}$ . Recuerde que los  $n-r$   $X_p$  son linealmente independientes.

*Demostración.*

Dado que  $E_{n-r} = S(\{X_p\}_{p=1}^{n-r})$ , si llamamos a  $E$  el conjunto de todos los vectores solución de  $H(L_{mn})$ , entonces lo que hay que demostrar es que  $E = E_{n-r}$ .

i. Primero demostraremos que  $E \subset E_{n-r}$ . Si  $X \in E$ , con  $X = \sum_{i=1}^n x^i e_i$ , entonces  $X$  es vector solución, o igual, los  $n$   $x^i$  son solución de  $H(L_{mn})$ , a causa de 15). Para demostrar que  $E \subset E_{n-r}$  es suficiente expresar  $X$  como una combinación lineal de los  $n-r$   $X_p$ ; es decir hay que encontrar  $n-r$   $c^p$  tal que:

$$X = \sum_{p=1}^{n-r} c^p X_p.$$

Si escogemos  $c^1 = x^{r+1}$ ,  $c^2 = x^{r+2}$ , ...,  $c^{n-r} = x^{n-r}$ , entonces, a causa de la forma como se construyeron los  $X_p$  dados en 16) y a la ecuación 14), se cumple finalmente que

$$X = \sum_{p=1}^{n-r} c^p X_p.$$

ii. Para demostrar que  $E_{n-r} \subset E$  tomemos un  $X = \sum_{p=1}^{n-r} x^p X_p \in E_{n-r}$ , y lo que hay que probar es que  $X$  es un vector solución, o lo que es igual según 15), que sus componentes, en la base dada, son solución de  $H(L_{mn})$ . Esto se demuestra apoyados en

las propiedades sobre el producto de matrices, dada en el teorema 4, así:

$$AX = A\left(\sum_{p=1}^{n-r} c^p X_p\right) = \sum_{p=1}^{n-r} A(c^p X_p) = \sum_{p=1}^{n-r} c^p (AX_p);$$

y como  $AX_p = 0$  para  $p = 1, 2, \dots, n-r$ , finalmente llegamos a lo que queríamos probar

$$AX = 0.$$

Ahora estamos listos para el objetivo principal de esta sección: una solución vectorial de  $L_{mn}$ . Esta solución está resumida en el teorema siguiente:

Teorema 2.4.b

Si  $X$  es vector solución de  $L_{mn}$ , entonces  $X$  es la suma de una solución particular de  $L_{mn}$  y algún vector del espacio solución del sistema homogéneo; y, a la inversa, una solución particular del sistema  $L_{mn}$  sumada a cualquier solución del sistema homogéneo es solución del sistema  $L_{mn}$ .

*Demostración.*

Sea  $E_{n-r}$  el espacio solución del sistema homogéneo, tal y como se trabajó en el teorema 2.4.a; además sea  $F$  una solución particular del sistema  $L_{mn}$  y  $H$  un elemento arbitrario de  $E_{n-r}$ , entonces demostraremos:

i.  $F \dagger H$  es solución de  $L_{mn}$ . La demostración es clara a causa del producto de matrices y sus propiedades.

$$A(F \dagger H) = AF \dagger AH.$$

Como  $AH = 0$  porque  $H \in E_{n-r}$ , y  $AF = B$  dado que  $F$  es solución particular, entonces  $A(F \dagger H) = B$ .

ii. Si  $X$  es solución de  $L_{mn}$  entonces  $X$  lo podemos representar como la suma de una solución particular,  $F$ , más algún elemento  $H \in E_{n-r}$ . Esto se demuestra escogiendo  $H = X - F$ , en donde es claro que  $X - F \in E_{n-r}$ , ya que  $A(X - F) = AX - AF = B - B = 0$ .

### 2.5 $K(T)$ y el sistema $H(L_{mn})$

Sea  $T$  una transformación lineal de  $E_n$  en  $E_m$ , con  $\{e_i\}_{i=1}^n$  y  $\{e_j\}_{j=1}^m$  bases para  $E_n$  y  $E_m$  respectivamente. Los teoremas del capítulo 1, el T-5 y el T-6, nos dicen respectivamente que  $K(T)$  es un subespacio vectorial de  $E_n$  y que  $K(T) = 0$  si y sólo si  $T$  es uno-uno. En esta sección veremos como el problema numérico de encontrar  $K(T)$  y ver si  $T$  es uno-uno se reduce a la determinación del espacio solución de un sistema  $H(L_{mn})$ .

La prueba es muy simple a causa del teorema 2.2 que dice que para  $T$  transformación lineal, con  $T(X)$  Y existe el sistema lineal  $L_{mn}$  dado por:

$$\sum_{i=1}^n x^i T_i = Y.$$

$x^1, x^2, \dots, x^n$  son componentes de  $X$  en la base  $\{e_i\}_{i=1}^n$ ;  $Y \in E_m$  con componentes  $y^1, y^2, \dots, y^m$  en la base  $\{e_j\}_{j=1}^m$ . Además los vectores  $T_i \in E_m$ , y sus componentes son las de  $T(e_i)$  en la base  $\{e_j\}_{j=1}^m$ . De lo anterior es claro que, como los  $X \in K(T)$  satisfacen que  $T(X) = 0$ , entonces estos  $X$  son los que satisfacen:

$$\sum_{i=1}^n x^i T_i = 0.$$

Por lo tanto el  $K(T)$  es el mismo subespacio solución de  $H(L_{mn})$ . Además la transformación  $T$  es uno-uno cuando el sistema  $H(L_{mn})$  sólo tiene como solución la trivial.

Hemos mostrado que el  $K(T)$  está caracterizado por el subespacio solución de  $H(L_{mn})$  asociado a la transformación  $T$ .

## 2.6 El sistema homogéneo traspuesto, $T(L_{mn})$ , a $L_{mn}$

Nuestro trabajo con el sistema homogéneo traspuesto de  $L_{mn}$  lleva como primer objetivo buscar un resultado muy importante llamado, a veces, teorema de dualidad. Este teorema se presentará en la sección 2.7, mientras en esta sección se definirá el sistema homogéneo traspuesto y daremos una solución teórica de éste.

Dado un sistema lineal  $L_{mn}$ ,

$$\sum_{i=1}^n a_i^j x^i = b^j, \quad j=1, 2, \dots, m,$$

llamamos sistema homogéneo traspuesto de  $L_{mn}$ , denotado por

$T(L_{mn})$ , al sistema:

$$17) \quad \sum_{j=1}^m a_i^j y_j = 0, \quad i=1, 2, \dots, n; \text{ o bien}$$

$$\sum_{j=1}^m {}^t a_j^i y_j = 0, \quad i=1, 2, \dots, n; \quad {}^t a_j^i = a_i^j.$$

Es importante comentar respecto de la matriz asociada a  $T(L_{mn})$  y su relación con la matriz asociada a  $L_{mn}$ .

$A^t$ , traspuesta de  $A$ , es la matriz asociada a  $T(L_{mn})$ . Recordemos que  $A^t$  se definió en el capítulo 1, sección 1.4, así:

$$18) \quad A^t = [\overset{t}{a}_j^i], \text{ con } \overset{t}{a}_j^i = a_i^j.$$

Con 18) y como se definió  $T(L_{mn})$  en 17) se prueba que la matriz asociada a  $T(L_{mn})$  viene a ser  $A^t$ , traspuesta de  $A$ .  $A^t$  justifica el nombre del sistema  $T(L_{mn})$ . Nótese que la matriz  $A^t$  se obtiene pasando los columnas de  $A$  a filas en  $A^t$ , conservando el orden de las columnas en  $A$  como el correspondiente en las filas de  $A^t$ .

Para el propósito de encontrar soluciones a  $T(L_{mn})$  introducimos una base  $\{e_i\}_{i=1}^m$  en el espacio  $E_m$ , tal y como se hizo en 2.3. Recordemos que los vectores  $T_i$ , también definidos en 2.3, pertenecen a  $E_m$ , con componentes  $a_i^j$  en la base  $\{e_i\}_{i=1}^m$ . De estos  $T_i$  supusimos que el rango era  $r$ , y arreglamos los primeros vectores de  $T_i$ ,  $i=1, 2, \dots, r$ , como los linealmente independientes.

Rango de un sistema  $L_{mn}$  se llama al rango de los  $\{T_i\}_{i=1}^n$ . De tal forma que si  $r$  es el rango de los  $T_i$  entonces  $r$  es el rango de  $L_{mn}$ , y así a la inversa. También se dice que el rango de  $L_{mn}$  es el rango de la matriz  $A$ , asociada a  $L_{mn}$ ; lo cual es consecuente con la definición 16 dada en el capítulo I.

Para encontrar soluciones a  $T(L_{mn})$  formaremos una nueva base de  $E_m$ , denotada por  $\{é_i\}_{i=1}^m$ , cuyos primeros  $r$  vectores son los  $r$  primeros vectores de  $\{T_i\}_{i=1}^n$ ; es decir,

$$é_1 = T_1, \quad é_2 = T_2, \quad \dots, \quad é_r = T_r, \quad é_{r+1}, \quad \dots, \quad é_m.$$

Los últimos  $m-r$  vectores,  $é_{r+1}$  hasta  $é_m$ , son los que comple

tan la base y sabemos que existen por el teorema de completitud de base, dado en el capítulo I como teorema 11.

Las ecuaciones de transformación de base entre  $\{e_i\}_{i=1}^m$  y  $\{é_j\}_{j=1}^m$ , guiados por el teorema 3 del capítulo I, las podemos representar así:

$$19) \quad e_j = \sum_{h=1}^m s_j^h é_h, \quad y$$

$$20) \quad é_i = \sum_{j=1}^m u_i^j e_j.$$

O sea que  $S = [s_j^h]$  viene en lugar de  $A$ , y  $U = [u_i^j]$  en lugar de  $A'$ , en el teorema 3 del capítulo I. De este teorema concluimos las dos siguientes ecuaciones:

$$21) \quad \begin{aligned} X' &= SX, \quad X, \text{ componentes en la base } \{e_i\}_{i=1}^m \text{ de un} \\ &\quad \text{vector en } E_m. \\ X &= UX', \quad X', \text{ componentes del mismo vector en} \\ &\quad \text{la base } \{é_j\}_{j=1}^m. \end{aligned}$$

Dado que los vectores  $T_i$ , en la base  $\{e_i\}_{i=1}^m$ , sus componentes vienen dadas por las columnas de la matriz  $A = [a_i^j]$ , entonces usando la ecuación 21 de transformación de componentes tenemos:

$$T_i' = ST_i, \quad i = 1, 2, \dots, n \quad \text{y con}$$

$$a_i^1, a_i^2, \dots, a_i^m, \quad \text{componentes de } T_i \text{ en } \{e_i\}_{i=1}^m \text{ y}$$

$$d_i^1, d_i^2, \dots, d_i^m, \quad \text{componentes de } T_i \text{ en } \{é_j\}_{j=1}^m.$$

$T_i$  como vector columna es  $T_i = (a_i^1, a_i^2, \dots, a_i^m)$ , mientras que

$T_i'$  como vector columna es  $T_i' = (d_i^1, d_i^2, \dots, d_i^m)$ . La ecuación

$T_i' = ST_i$  es equivalente a la ecuación:

$$d_i^h = \sum_{j=1}^m s_j^h a_i^j, \quad \text{con } \begin{matrix} h = 1, 2, \dots, m \\ i = 1, 2, \dots, n \end{matrix},$$

o bien a la ecuación:

$$22) \quad d_i^h = \sum_{j=1}^m s_j^h a_j^i .$$

Por el hecho de que los primeros  $r$   $\{T_i\}_{i=1}^n$  son también los primeros elementos de la nueva base  $\{e_j\}_{j=1}^m$  concluimos que las últimas  $m-r$  componentes, de cualquiera de los primeros  $r$   $T_i$ , en la base nueva son cero. Lo mismo se puede decir de los  $T_{r+1}$  hasta  $T_n$ , ya que éstos son linealmente dependientes de los primeros  $r$   $T_i$ . Estos hechos justifican las ecuaciones:

$$\begin{aligned} 0 &= d_i^{r+\epsilon} = \sum_{j=1}^m s_j^{r+\epsilon} a_j^i & i &= 1, 2, \dots, n \\ 0 &= \sum_{j=1}^m a_j^i s_j^{r+\epsilon} & \epsilon &= 1, 2, \dots, m-r \end{aligned}$$

finalmente, la ecuación:

$$23) \quad 0 = \sum_{j=1}^m a_j^i s_j^{r+\epsilon} .$$

La ecuación 23) nos indica que los  $s_1^{r+\epsilon}, s_2^{r+\epsilon}, \dots, s_m^{r+\epsilon}$  forman una solución del sistema  $T(L_{mn})$ , para cada  $\epsilon = 1, 2, \dots, m-r$ . Por lo tanto, tenemos:  $m-r$  conjuntos de soluciones que podemos representar así:

$$C^\epsilon = (s_1^{r+\epsilon}, s_2^{r+\epsilon}, \dots, s_m^{r+\epsilon}), \quad \epsilon = 1, 2, \dots, m-r.$$

Los  $C^\epsilon$  eran las soluciones que buscábamos para el sistema  $T(L_{mn})$ .

## 2.7 Teorema de dualidad

Previo al teorema de dualidad, objetivo principal de esta sección, es necesario mostrar que los  $m-r$   $C$  son linealmente independientes. Para ésto, primero probaremos que:

$$SU = US = I.$$

Para demostrar la ecuación anterior es suficiente con llegar a  $\sum_{j=1}^m s_j^h u_i^j = S_i^h$ , en donde  $S_i^h$  es 1 si  $h=i$  y  $S_i^h = 0$  si  $h \neq i$ . De las ecuaciones 19) y 20) tenemos que:

$$\begin{aligned} \epsilon_i &= \sum_{j=1}^m u_i^j e_j = \sum_{j=1}^m u_i^j \left( \sum_{h=1}^m s_j^h \epsilon_h \right), \text{ o sea que,} \\ \epsilon_i &= \sum_{h=1}^m \left( \sum_{j=1}^m u_i^j s_j^h \right) \epsilon_h; \text{ de donde finalmente lo} \end{aligned}$$

que queríamos:

$$24) \quad \sum_{j=1}^m s_j^h u_i^j = S_i^h.$$

Para demostrar, ahora, que los  $C^\epsilon = (s_1^{r+\epsilon}, s_2^{r+\epsilon}, \dots, s_m^{r+\epsilon})$ ,  $\epsilon = 1, 2, \dots, m-r$ , son linealmente independientes es suficiente que se cumpla con:

Si  $\sum_{h=1}^m \alpha_h s_j^h = 0$ , entonces  $\alpha_h = 0$ ,  $h = 1, 2, \dots, m$ . De  $\sum_{h=1}^m \alpha_h s_j^h = 0$ ,  $j = 1, 2, \dots, m$  se deduce que  $\sum_{h=1}^m \alpha_h u_i^j s_j^h = 0$ ,  $j = 1, 2, \dots, m$  &  $i = 1, 2, \dots, m$ ; de ahí  $\sum_{j=1}^m \left( \sum_{h=1}^m \alpha_h u_i^j s_j^h \right) = 0$ ,  $i = 1, 2, \dots, m$ . Luego,  $\sum_{h=1}^m \alpha_h \left( \sum_{j=1}^m s_j^h u_i^j \right) = 0$ ,  $i = 1, 2, \dots, m$ ; pero por 24),  $\sum_{h=1}^m \alpha_h S_i^h = 0$ ,  $i = 1, 2, \dots, m$ , de ahí finalmente lo que queríamos  $\alpha_i = 0$ ,  $i = 1, 2, \dots, m$ .

Con lo anterior queda probado que los  $m-r$   $C^\epsilon$  son linealmente independientes. Ahora, estamos listos para el teorema de dualidad.

Teorema de dualidad. El sistema homogéneo  $H(L_{mn})$  y el traspuesto homogéneo  $T(L_{mn})$ , ambos de un sistema  $L_{mn}$ , tienen el mismo rango. En otras palabras, la matriz asociada,  $A$ , y su traspuesta,  $A^t$ , tienen el mismo rango, o más general, una matriz arbitraria y su traspuesta tienen el mismo rango.

*Demostración.*

Por el teorema 2.4.a sabemos que si el sistema  $H(L_{mn})$  tiene rango  $r$ , entonces dicho sistema tiene, a lo sumo,  $n-r$  soluciones linealmente independientes. Por lo anterior si el sistema  $T(L_{mn})$  tuviera rango  $r'$ , entonces a lo sumo tendría  $m-r'$  soluciones linealmente independientes. Ya encontramos, para  $T(L_{mn})$ ,  $m-r$  soluciones linealmente independientes, a saber los  $m-r$   $C^E$ , por lo que hay que concluir:

$$m-r \leq m-r', \text{ y de ahí que,}$$

$$r' \leq r.$$

Luego, como el traspuesto homogéneo de  $T(L_{mn})$  coincide con  $H(L_{mn})$  tendríamos que hay  $n-r'$  soluciones linealmente independientes de  $H(L_{mn})$  y que el sistema  $H(L_{mn})$  tiene a lo sumo  $n-r$  soluciones linealmente independientes, por lo que hay que concluir que:

$$n-r' \leq n-r, \text{ y de ahí que:}$$

$$r \leq r'.$$

Con los dos resultados,  $r' \leq r$  y  $r \leq r'$ , se concluye que

$$r = r'.$$

La última ecuación demuestra el teorema de dualidad. #

## 2.8 Condiciones de existencia

Sabemos que una condición suficiente y necesaria para que el sistema  $L_{mn}$ , dado por,

$$\sum_{i=1}^n a_i^j x^i = b^j, \quad j=1,2,\dots,m,$$

tenga solución es que  $B = (b^1, b^2, \dots, b^m)$  pertenezca al subespacio generado por los  $n$   $T_i = (a_i^1, a_i^2, \dots, a_i^m)$ .

Un caso particular importante ocurre cuando el sistema es de orden cuadrado,  $m$  por  $m$ : si los  $m$   $T(e_i)$ , en este caso, son linealmente independientes, es decir el rango es  $m$ , entonces siempre existe solución y, por supuesto, ésta es única.

Si el rango de  $L_{mn}$  es  $r$  entonces la condición necesaria y suficiente para que exista solución es que  $B \in E_r$ ; en donde  $E_r$  es el subespacio generado por los  $r$   $T_i$  linealmente independientes de  $\{T_i\}_{i=1}^n$ .

Si para el vector  $B \in E_m$  con componentes  $b^1, b^2, \dots, b^m$ , en la base  $\{e_i\}_{i=1}^m$ , utilizamos las ecuaciones de transformación de componentes, dadas en 21), tenemos que:

$$25) \quad b^h = \sum_{j=1}^m s_j^h b^j, \quad h=1, 2, \dots, m;$$

con  $b^h$  componentes de  $B$  en  $\{e_j\}_{j=1}^m$ .

De 25), las últimas  $m-r$  componentes de  $B$ , en la base  $\{e_j\}_{j=1}^m$ , las podemos expresar así:

$$b^{r+\epsilon} = \sum_{j=1}^m s_j^{r+\epsilon} b^j, \quad \epsilon=1, 2, \dots, m-r.$$

La ecuación anterior, junto con que  $T_1 = e_1, T_2 = e_2, \dots, T_r = e_r$ , nos lleva a plantear el siguiente enunciado:

$$26) \quad \sum_{i=1}^n a_i^j x^i = b^j, \quad j=1, 2, \dots, m \text{ tiene solución,}$$

si y sólo si  $\sum_{j=1}^m s_j^{r+\epsilon} b^j = 0, \epsilon=1, 2, \dots, m-r.$

Como  $C^\epsilon = (s_1^{r+\epsilon}, s_2^{r+\epsilon}, \dots, s_m^{r+\epsilon})$ ,  $\epsilon=1, 2, \dots, m-r$ , entonces

26) pasa a quedar así:

$$27) \quad \sum_{i=1}^n a_i^j x^i = b^j, \quad j=1,2,\dots,m \text{ tiene solución si}$$

$$\text{y sólo si } C^\varepsilon \cdot B = 0, \quad \varepsilon=1,2,\dots,m-r.$$

Pero como los  $m-r$   $C^\varepsilon$ , ver sección 2.6, son base del espacio solución de  $T(L_{mn})$ , podemos decir finalmente que

27) se reduce a:

$$28) \quad \sum_{i=1}^n a_i^j x^i = b^j, \quad j=1,2,\dots,m \text{ tiene solución, si}$$

$$\text{y sólo si, } B \text{ es ortogonal al espacio solución}$$

$$\text{de } T(L_{mn}).$$

Es claro que  $L_{mn}$  tiene solución suficiente con que  $B$  sea ortogonal a  $m-r$  soluciones linealmente independientes del sistema  $T(L_{mn})$ .

## 2.9 Determinantes y rango de una matriz

En esta sección veremos un teorema que relaciona determinantes y rango de matrices. Es importante que se recuerde la definición de rango de una matriz, dada en el capítulo I, como D-16. Además se expone un método para calcular numéricamente el rango de una matriz, vía saber calcular un determinante. Los resultados de esta sección no son los mejores, desde el punto de vista numérico, para calcular el rango de una matriz. El método numérico adecuado para calcular el rango de una matriz lo expondremos en el capítulo III; mostrando sus ventajas sobre el aquí expuesto.

El resultado de esta sección viene en el siguiente teorema.

Teorema 2.9. El rango  $r$  de una matriz  $A$ ,  $m$  por  $n$ , es igual al número  $p$ , tal que el mayor determinante diferente de cero, sacado de la matriz  $A$ , es de orden  $p$ .

*Demostración.*

Sea la matriz  $A$ , representada en su forma rectangular,

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix}.$$

Supongamos que existe en  $A$  un determinante de orden  $p$  diferente de cero. Reordenando filas y columnas en  $A$ , podemos llegar a la submatriz de  $A$ , que llamaremos  $L$ , dada así:

$$L = \begin{pmatrix} a_{r_1}^1 & a_{r_2}^1 & \dots & a_{r_p}^1 \\ a_{r_1}^2 & a_{r_2}^2 & \dots & a_{r_p}^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{r_1}^m & a_{r_2}^m & \dots & a_{r_p}^m \end{pmatrix}$$

en donde las primeras  $p$  filas de esta matriz  $L$  forman el determinante de orden  $p$  diferente de cero. A causa del teorema de dualidad, y al teorema sobre determinantes y linealidad, dado en el capítulo I, todos los vectores columna de la matriz  $L$  son linealmente independientes. Como estos  $p$  vectores columna de  $L$  son también vectores columna de  $A$ , tenemos que:

$$29) \quad p \leq r.$$

Por otro lado si  $r$  es el rango de  $A$  entonces existen en  $A$   $r$  vectores columna que son linealmente independientes. Reordenando columnas en  $A$  podemos dar una submatriz de  $A$  así:

$$Q = \begin{pmatrix} a_{q_1}^1 & a_{q_2}^1 & \dots & a_{q_r}^1 \\ a_{q_1}^2 & a_{q_2}^2 & \dots & a_{q_r}^2 \\ \vdots & \vdots & & \vdots \\ a_{q_1}^m & a_{q_2}^m & \dots & a_{q_r}^m \end{pmatrix}$$

en donde  $r$  es el rango de  $Q$ .

Según el teorema de dualidad, deben existir en  $Q$   $r$  vectores filas que son linealmente independientes. Estas  $r$  filas, de  $r$  componentes, forman un determinante de orden  $r$  diferente de cero. Por lo tanto, tomando en cuenta la definición de  $p$ , tenemos que:

$$30) \quad r \leq p.$$

Es claro que de 29) y 30) concluimos que:

$$r = p.$$

### 2.10 Ecuaciones principales y nuevas condiciones de existencia.

De el sistema  $L_{mn}$ , con rango  $r$ , dado por,

$$\sum_{i=1}^n a_{ij}^j x^i = b^j, \quad j=1, 2, \dots, m;$$

reordenando filas y columnas podemos pasar al mismo sistema  $L_{mn}$  expresado así:





$$35) \quad \Delta^\epsilon = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_r^1 & b^1 \\ a_1^2 & a_2^2 & \dots & a_r^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^r & a_2^r & \dots & a_r^r & b^r \\ a_1^{r+\epsilon} & a_2^{r+\epsilon} & \dots & a_r^{r+\epsilon} & b^{r+\epsilon} \end{vmatrix},$$

con  $\epsilon = 1, 2, \dots, m-r$ .

Las nuevas condiciones de existencia, ofrecidas para esta sección, están contenidas en el teorema siguiente.

Teorema 2.10.b Dado un sistema  $L_{mn}$ , con rango  $r$ , el sistema  $L_{mn}$  tiene solución si y sólo si los  $m-r$  determinantes característicos de  $L_{mn}$ ,  $\Delta^\epsilon$ , son iguales a cero.

*Demostración.*

El sistema  $L_{mn}$  tiene solución si y sólo si el conjunto de los  $n$  vectores columna, de la matriz  $A$  asociada al sistema  $L_{mn}$ , tiene el mismo rango que la matriz  $A'$ , formada por la matriz  $A$  más una nueva columna constituida por el vector columna  $(b^1, b^2, \dots, b^m)$ . Lo anterior lo podemos expresar más claro si representamos  $A$  y  $A'$  de la siguiente manera:

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix}$$

$$A' = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m & b^m \end{pmatrix} .$$

El sistema  $L_{mn}$ , con matriz asociada  $A$ , tiene solución si y sólo si el rango de  $A$  es igual al rango de  $A'$ .  $A$  y  $A'$  como en 36).

Supongamos que el rango de  $A$  y  $A'$  es  $r$ ; según el teorema 2.9, podemos concluir que todo determinante de orden  $r+1$ , sacado de  $A'$ , es cero. En particular serán cero todos los determinantes característicos  $\Delta^\epsilon$ , ya que provienen de  $A'$  y son de orden  $r+1$ .

Sólo nos resta probar la condición suficiente. Si los  $\Delta^\epsilon$  son cero, entonces existe una dependencia lineal entre los vectores columna de  $\Delta^\epsilon$ ; es decir existen  $v^i(\epsilon)$  y  $u(\epsilon)$  para cada  $\epsilon=1,2,\dots,m-r$ , tal que alguno de los  $v^i(\epsilon)$  o  $u(\epsilon)$  no son cero en las ecuaciones:

$$37) \quad \sum_{i=1}^r a_i^j v^i(\epsilon) + u(\epsilon) b^j = 0$$

$$\sum_{i=1}^r a_i^{r+\epsilon} v^i(\epsilon) + u(\epsilon) b^{r+\epsilon} = 0,$$

con  $j=1,2,\dots,r$  y  $\epsilon=1,2,\dots,m-r$ .

En las ecuaciones 37) tenemos que  $u(\epsilon) \neq 0$ , ya que si no fuera así, entonces tendríamos que  $\Delta=0$ , lo cual es contradictorio. Dado que  $u(\epsilon) \neq 0$  podemos pasar de 37) a las siguientes ecuaciones, despejando  $b^j$  y  $b^{r+\epsilon}$ :

$$38) \quad b^j = \sum_{i=1}^r a_i^j w^i(\varepsilon)$$

$$b^{r+\varepsilon} = \sum_{i=1}^r a_i^{r+\varepsilon} w^i(\varepsilon),$$

con  $j=1, 2, \dots, r$  y  $\varepsilon=1, 2, \dots, m-r$ . En las ecuaciones 38) es claro que  $w^i(\varepsilon) = \frac{v^i(\varepsilon)}{u(\varepsilon)}$ .

La primera ecuación en 38) nos indica que los  $w^i(\varepsilon)$ ,  $i=1, 2, \dots, r$ , constituyen una solución de un sistema  $L_{rr}$ , es decir  $L_r^2$ ; formado con las  $r$  primeras filas y columnas de la matriz  $A$ . Dado que la solución de los  $w^i(\varepsilon)$ , para el sistema  $L_r^2$ , es única, a causa de que el rango de  $L_r^2$  es  $r$ , concluimos que los  $w^i(\varepsilon)$  no dependen de  $\varepsilon$ , y por lo tanto 38) se convierte en:

$$39) \quad b^j = \sum_{i=1}^r a_i^j w^i$$

$$b^{r+\varepsilon} = \sum_{i=1}^r a_i^{r+\varepsilon} w^i,$$

con  $j=1, 2, \dots, r$  y  $\varepsilon=1, 2, \dots, m-r$ . De las ecuaciones en 39), es claro que la solución que buscamos para el sistema  $L_{mn}$  viene dada así:  $x^1 = w^1, x^2 = w^2, \dots, x^r = w^r$ ; mientras que las  $n-r$  incógnitas que se pueden escoger arbitrariamente, dado que el sistema tiene rango  $r$ , las podemos hacer igual a cero: es decir,  $x^{r+1} = 0, x^{r+2} = 0, \dots, x^n = 0$ . #

### 2.11. La solución de Cramer

Ahora veremos un procedimiento numérico para solucionar un sistema  $L_{mn}$ ; sin embargo, su defecto consiste en la cantidad abundante de operaciones que emplea. La solución de Cramer, que es la que se presentará en esta sección, in-

volucra el cálculo de determinantes, los cuales resueltos mediante la ecuación 2), sección 1.4.d, capítulo I, son los que acrecientan desmesuradamente el número de operaciones. Será en el capítulo III donde veremos la diferencia en la cantidad de operaciones para el método de Cramer comparado con otros métodos; así como también porque es importante la cantidad de operaciones a pesar de disponer de máquinas calculadoras de alta velocidad.

En esta sección presentaremos dos casos del método de Cramer; en realidad el mismo método, pero empleado en un caso particular y en el caso general.

Caso 1. Sea el sistema  $L_{mn}$  de orden  $m$ , es decir,  $L_{mn} = L_m^2$  con rango  $m$ . Dicho sistema viene dado así:

$$40) \quad \sum_{i=1}^m a_i^j x^i = b^j, \quad j = 1, 2, \dots, m.$$

Como los  $m$  vectores columna de la matriz  $A$ , asociada a 40), son linealmente independientes se concluye que existe solución y esta es única. El determinante del sistema 40) viene dado como:

$$41) \quad \Delta = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_m^1 \\ a_1^2 & a_2^2 & \dots & a_m^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_m^m \end{vmatrix} \neq 0$$

Para buscar la solución de Cramer multiplicamos la ecuación 40) por el cofactor  $\alpha_j^k$  con  $k = 1, 2, \dots, m$ . Dichos co-

factores vienen definidos como en la sección 1.4.e, capítulo I. Ya con esto, la ecuación 40) queda como:

$$\sum_{i=1}^m \alpha_j^k a_i^j x^i = \alpha_j^k b^j, \quad \begin{array}{l} j=1,2,\dots,m \\ k=1,2,\dots,m. \end{array}$$

Luego sumamos la ecuación anterior sobre el subíndice  $j$  y obtenemos

$$\sum_{j=1}^m \left( \sum_{i=1}^m \alpha_j^k a_i^j x^i \right) = \sum_{j=1}^m \alpha_j^k b^j, \quad k=1,2,\dots,m,$$

o bien:

$$42) \quad \sum_{i=1}^m \left( \sum_{j=1}^m \alpha_j^k a_i^j \right) x^i = \sum_{j=1}^m \alpha_j^k b^j, \quad k=1,2,\dots,m.$$

En la ecuación 42) se demuestra:

- i.  $\sum_{j=1}^m \alpha_j^k a_i^j = \Delta$  si  $k=i$
- ii.  $\sum_{j=1}^m \alpha_j^k a_i^j = 0$  si  $k \neq i$ .

i. está probado por ser la expansión del determinante  $\Delta$ , dado en 41), por la  $k$ -ésima columna. ii. está probado porque corresponde a la expansión, por la  $k$ -ésima columna, de un determinante de la forma de 41) pero con la misma columna en la  $k$ -ésima posición y la  $i$ -ésima; es decir, es un determinante con dos columnas iguales y por lo tanto cero.

Los asertos i. y ii. convierten a 42) en

$$\Delta x^k = \sum_{j=1}^m \alpha_j^k b^j, \quad k=1,2,\dots,m;$$

de donde, finalmente, obtenemos la llamada solución de Cramer:

$$43) \quad x^k = \frac{1}{\Delta} \sum_{j=1}^m \alpha_j^k b^j, \quad k=1,2,\dots,m$$

Caso 2. Aquí tendremos un sistema  $L_{mn}$ , de rango  $r$  y con solución; en donde las primeras  $r$  filas y colum-

nas de la matriz  $A$  asociada a  $L_{mn}$  forman un determinante, de orden  $r$  y diferente de cero. Es decir:

$$44) \quad \Delta = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_r^1 \\ a_1^2 & a_2^2 & \dots & a_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^r & a_2^r & \dots & a_r^r \end{vmatrix} \neq 0$$

Las primeras  $r$  filas son las ecuaciones principales de  $L_{mn}$ , que según el teorema 2.10.a, dan las soluciones de  $L_{mn}$  y por lo tanto trabajaremos sólo con ellas. En este caso el procedimiento es similar al caso 1, y se desarrolla así: del sistema  $\sum_{i=1}^n a_i^j x^i = b^j$ ,  $j=1,2,\dots,r$ , pasamos a las ecuaciones principales.

$$\sum_{h=1}^r a_h^j x^h + \sum_{\epsilon=1}^{n-r} a_{r+\epsilon}^j x^{r+\epsilon} = b^j, \quad j=1,2,\dots,r.$$

Luego, seguimos el caso 1 de la siguiente manera,

$$\sum_{h=1}^r \alpha_j^k a_h^j x^h + \sum_{\epsilon=1}^{n-r} \alpha_j^k a_{r+\epsilon}^j x^{r+\epsilon} = \alpha_j^k b^j,$$

con  $j=1,2,\dots,r$  y  $k=1,2,\dots,r$ . En la ecuación anterior

sumamos sobre el subíndice  $j$  y obtenemos:

$$\sum_{h=1}^r \left( \sum_{j=1}^r \alpha_j^k a_h^j \right) x^h + \sum_{\epsilon=1}^{n-r} \left( \sum_{j=1}^r a_{r+\epsilon}^j x^{r+\epsilon} \right) \alpha_j^k = \sum_{j=1}^r \alpha_j^k b^j,$$

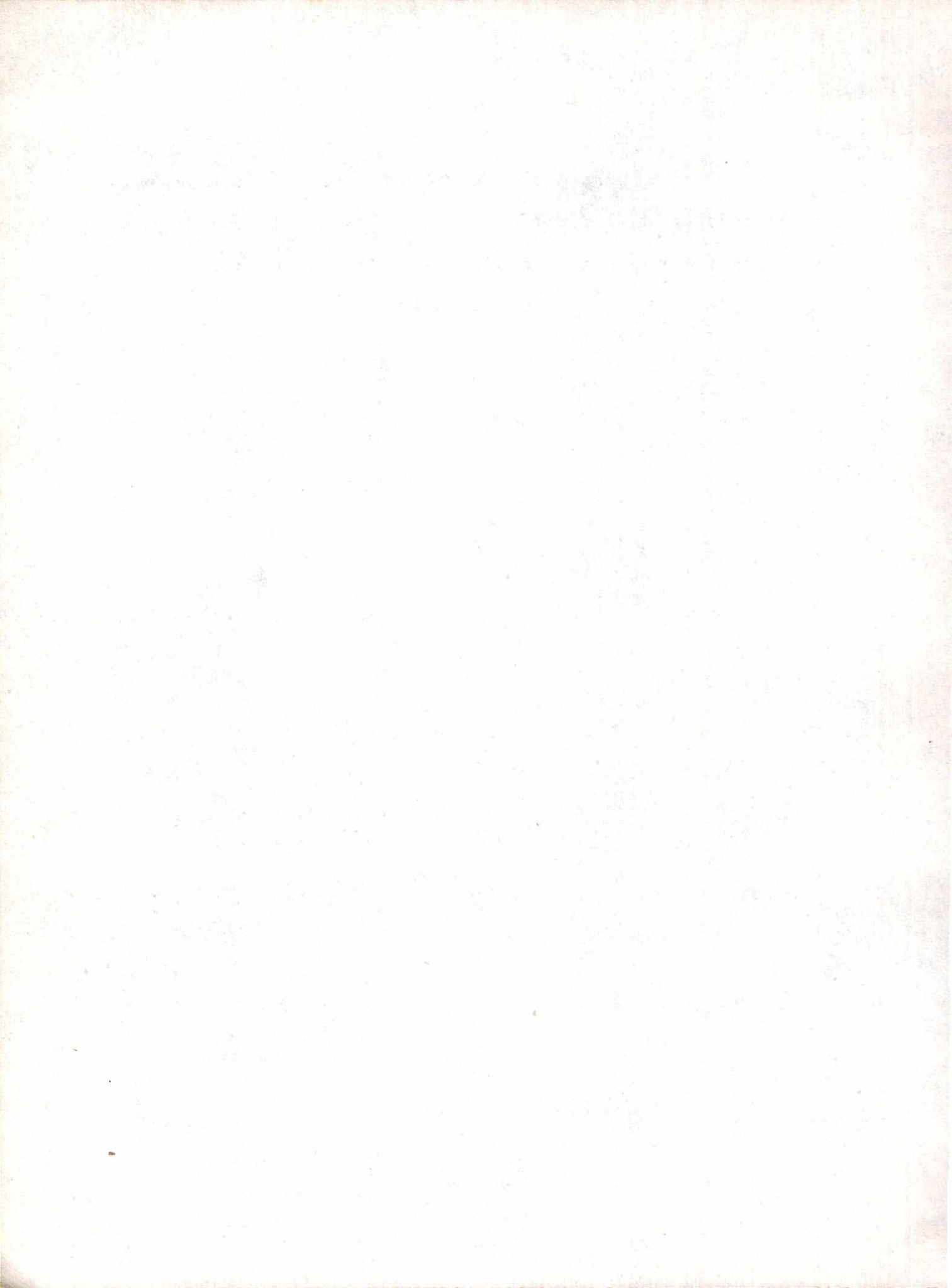
$$\Delta x^k + \sum_{j=1}^r \left( \sum_{\epsilon=1}^{n-r} a_{r+\epsilon}^j x^{r+\epsilon} \right) \alpha_j^k = \sum_{j=1}^r \alpha_j^k b^j, \quad \text{con}$$

$k=1,2,\dots,r$ . Note en esta última ecuación que  $\Delta$  es justamente el determinante dado en 44). Ya con ésto la solución

de Cramer, al despejar  $x^k$ , resulta ser:

$$45) \quad x^k = \frac{1}{\Delta} \left[ \sum_{j=1}^r \alpha_j^k (b^j - \sum_{\epsilon=1}^{n-r} a_{r+\epsilon}^j x^{r+\epsilon}) \right], \quad k=1,2,\dots,r.$$

Observe el gran parecido de esta solución con la solución dada en 43); también note que las  $n-r$  incógnitas  $x^{r+\epsilon}$  se pueden escoger arbitrariamente, lo cual coincide con la solución teórica dada en 2.3, ecuación 12).



### CAPITULO III

#### 3.1. Solución numérica de sistemas lineales y algunos problemas

La solución numérica de un sistema lineal consiste en cualquier método que de un algoritmo para encontrar, si es que existen, las incógnitas. Sin embargo, el problema se complica, por ejemplo, a causa de que ciertos métodos emplean una cantidad de operaciones tan grande que la solución se vuelve impracticable.

La cantidad de operaciones pareciera no ser un problema si se dispone de un computador, sin embargo aún así el problema subsiste. En ciertos casos, no muy comunes, pero no por eso carentes de importancia, la cantidad de operaciones es el factor decisivo para decidir si el problema tiene o no solución. Esto porque la cantidad de iteraciones es proporcional, más o menos, al tiempo de trabajo que requiere la máquina para resolver un problema.

Algunos problemas dependen del tiempo que la máquina emplee en dar solución, y aún más, otros tienen solución según el tiempo que tarde la máquina en resolver las operaciones. Por ejemplo, en el caso de un misil se necesita una cantidad mínima de tiempo para procesar información y reaccionar, si aún se espera que el misil de en el blanco. También, en cierto grado, el número de operaciones es proporcional al costo, por el mayor o menor tiempo que se use la

máquina. Así pues, en un caso el tiempo en que la máquina resuelve es vital para que ésta pueda dar la solución deseable al problema; en otro caso, tal vez el tiempo de solución no sea vital pero incrementa gastos, dependiendo, indirectamente, de la cantidad de operaciones. Finalmente, siempre subsiste el problema en sí, teórico, de plantear soluciones que requieran el mínimo de iteraciones.

En este capítulo, no entraremos en detalles al respecto de la cantidad de operaciones, nos limitaremos a indicar la cantidad de operaciones que requiere un método u otro. También, hay que notar que el tamaño de la matriz asociada a  $L_{mn}$  puede generar problemas serios para su computación: tanto por incrementar la cantidad de operaciones como por errores de redondeo. De acuerdo a Scheid (1968:338) la acumulación de error de redondeo es aproximadamente dependiente de la raíz cuadrada del número de operaciones.

En este capítulo, sección 3.2, trabajaremos el importante caso de las matrices no-singulares, mientras que en 3.3. veremos las técnicas numéricas para trabajar con matrices. Luego en las secciones 3.4 y 3.5 desarrollaremos unos métodos exactos para resolver un sistema lineal  $L_{mn}$ . En la sección final, 3.6, se desarrolla la solución de los problemas asociados a un sistema lineal: determinantes y cálculo de la inversa.

Llamaremos métodos exactos (directos) para resolver un

sistema  $L_{mn}$ , aquellos que emplean un número finito de operaciones elementales (+, -, ×, ÷) que conducen a la solución exacta; en el supuesto de que las máquinas no hicieran redondeos. Los métodos indirectos trabajan, esencialmente, por criterios de convergencia, es decir, aproximar la solución mediante procesos de convergencia. Para una explicación más amplia de los diversos métodos consultar Faddeeva (1959:63-65).

### 3.2. Matrices y transformaciones no-singulares

D-1 Matriz no-singular. Dada una matriz  $A = [a_{ij}^j]$ ,  $m$  por  $m$ , decimos que es no-singular si al asociarle su correspondiente operador lineal  $T_A$ , de un espacio arbitrario  $E_m$  en el mismo  $E_m$ , sobre una base dada  $\{e_i\}_{i=1}^m$ , resulta que  $T_A$  es no-singular.

Recordemos la definición de un operador lineal no-singular, dada en D-11 capítulo I:  $T: E_m \rightarrow E_m$  es no-singular si y sólo si la transformación lineal  $T$  es sobreyectiva.

En esta sección será de utilidad tener a la vista los siguientes teoremas:

Teorema de determinantes y linealidad; Cap. I.

A)

Para una matriz  $A$ ,  $m$  por  $m$ , los vectores columna son linealmente independientes si y sólo si  $d(A) \neq 0$ .

Teorema 2.1; Cap. I.

B) Si  $T$  es transformación lineal, de  $E_m$  en  $E_m$ , y  $A$  su matriz asociada sobre una base  $\{e_i\}_{i=1}^m$ , entonces se cumple que  $T(X) = Y$  si y sólo si  $AX = Y$ .

T-1 Operador lineal no-singular e independencia lineal. Dada una transformación lineal  $T$ , definida de  $E_m$  en  $E_m$ , y una base arbitraria  $\{e_i\}_{i=1}^m$  de  $E_m$ , entonces  $T$  es no-singular si y sólo si los  $m$  vectores  $T(e_i)$ ,  $i = 1, 2, \dots, m$ , son linealmente independientes. Otra forma equivalente del enunciado anterior viene a ser que  $T$  es no-singular si y sólo si el determinante de la matriz asociada a  $T$ , sobre una base arbitraria, es diferente de cero.

*Demostración.*

i. Si  $T$  es no-singular entonces para cualquier  $Y \in E_m$ , existe  $X \in E_m$  tal que  $T(X) = Y$ . Lo anterior por definición de que  $T$  es no-singular.

Sea  $\{e_j\}_{j=1}^m$  una base arbitraria de  $E_m$ , entonces tendremos que existen los  $x^j$  tal que  $X = \sum_{j=1}^m x^j e_j$ . Dado que  $T(X) = Y$  y  $T$  es lineal, entonces  $Y = \sum_{j=1}^m x^j T(e_j)$ . Lo anterior indica que los  $m$   $T(e_j)$  generan al espacio  $E_m$ . Finalmente, como  $E_m$  es de dimensión  $m$ , entonces los  $m$   $T(e_j)$  son linealmente independientes.

ii. Si los  $m$   $T(e_j)$  son linealmente independientes, entonces generan al espacio  $E_m$  de dimensión  $m$ ; por lo tanto para cualquier  $Y \in E_m$  existen los  $x^j$  tal que  $Y = \sum_{j=1}^m x^j T(e_j)$ .

Como  $T$  es transformación lineal tenemos que la ecuación anterior se convierte en  $Y = T\left(\sum_{j=1}^m x^j e_j\right)$ . Si hacemos  $X = \sum_{j=1}^m x^j e_j$ , entonces hemos probado que  $T$  es sobreyectiva, por lo tanto no-singular.

*Nota:* La forma equivalente referente al determinante se demuestra basados en el enunciado A), de determinantes y linealidad dado al principio de esta sección, junto con el hecho de que las componentes de  $T(e_j)$  forman las columnas de la matriz asociada a  $T$ .

Corolario T-1. i. Cualquier transformación no-singular es uno-uno. ii. Cualquier transformación no-singular es un isomorfismo en el espacio en el cual está definida.

*Demostración:*

Sea  $T$  de  $E_m$  en  $E_m$  no-singular, dado que la matriz asociada a  $T$ , llamémosle  $A$ , tiene por columnas las componentes de  $T(e_j)$  sobre una base  $\{e_j\}_{j=1}^m$  de  $E_m$ , resulta, por T-1, que el determinante de  $A$  es diferente de cero. Como el determinante de  $A$  es diferente de cero el rango de  $A$  es  $m$ ; entonces el sistema lineal  $AX=Y$ , para cada  $Y \in E_m$ , tiene una solución única. Es decir,  $T$  es uno-uno; ver capítulo II, sección 2.8.

ii. Esto es claro por la definición de transformación lineal no-singular, el resultado i. del corolario dado aquí y la definición de isomorfismo. #

T-2 Matrix no-singular,  $H(L_m^2)$ , determinante, inversa.

Para una matriz  $A = [a_{ij}^1]$  las siguientes proposiciones son todas equivalente, con  $A$  matriz  $m$  por  $m$ .

- i.  $A$  es no-singular.
- ii.  $AX=0$  tiene como única solución la trivial.
- iii. Determinante de  $A$  es diferente de cero.
- iv. Existe inversa de  $A$ .

*Demostración.*

Lo que hay que probar son las siguientes implicaciones:

$i. \Rightarrow ii.$ ;  $ii. \Rightarrow iii.$ ;  $iii. \Rightarrow iv.$ ;  $iv. \Rightarrow i.$

$i. \Rightarrow ii.$

Dado que  $A$  es no-singular entonces, por el teorema 1, el determinante de  $A$  es diferente de cero y de ahí que el rango de  $A$  es  $m$ . En este caso el sistema lineal  $AX=Y$ , para cada  $Y \in E_m$ , tiene solución única. Finalmente,  $AX=0$  tiene entonces sólo una solución para  $X$ : ésta necesariamente es  $X=0$ .

Lo anterior se puede resumir así:

$$AX=0 \text{ implica que } X=0.$$

La última ecuación prueba que en este caso el sistema homogéneo  $H(L_m^2)$  tiene como única solución la trivial.

$ii. \Rightarrow iii.$

Si  $AX=0$  implica que  $X=0$  se deduce, por teorema 6 capítulo I, que  $AX=Y$  tiene solución única de  $X$  para cada  $Y$ . Por lo tanto los vectores columna de  $A$  son linealmente inde

pendientes y, por el enunciado A), dado al principio de la sección, tenemos que el determinante de  $A$  es diferente de cero.

iii.  $\Rightarrow$  iv.

Si el determinante de  $A$  es diferente de cero entonces los vectores columna de  $A$  son linealmente independientes, por el enunciado A), a causa de ésto resulta que el sistema lineal  $AX=Y$  tiene solución única para cada vector columna  $Y$ . De lo anterior, junto con el enunciado B) dado al principio de la sección, tenemos que  $T_A$  asociada a  $A$  es uno-uno y sobreyectiva. Por este último resultado concluimos que  $T_A$  tiene transformación lineal inversa, que llamaremos  $T_A^{-1}$ , la cual cumple con:

$$T_A T_A^{-1} = I.$$

La matriz inversa de  $A$  viene entonces dada como la matriz asociada a  $T_A^{-1}$  y cumplirá, por supuesto, con que  $AA^{-1} = I$ .

iv.  $\Rightarrow$  i.

Lo que demostraremos es que dado  $A$ , entonces la transformación lineal asociada  $T_A$  es sobreyectiva. Si tomamos cualquier vector columna  $Y$  tenemos que, como  $AA^{-1} = I$ , entonces:

$$AA^{-1}Y = Y.$$

De la ecuación anterior es claro que para cualquier  $Y$  existe un  $X$  tal que  $AX=Y$ ; explícitamente  $X=A^{-1}Y$ . Con ésto queda probado que  $T_A$  es sobreyectiva. #

Corolario T-2. Sean  $A$  y  $B$  dos matrices no-singulares entonces el producto de ellas es no-singular.

*Demostración.*

Por la propiedad de determinantes, dada en el capítulo I, que plantea que  $d(AB) = d(A)d(B)$ , tendremos que ya que  $d(A) \neq 0$  y  $d(B) \neq 0$ , por ser  $A$  y  $B$  no-singulares,  $d(AB) \neq 0$ . Si  $d(AB) \neq 0$  entonces por el teorema 1 resulta que  $AB$  es no-singular.

T-3 Lema de Gersgorin. Dada una matriz cuadrada  $A$ , de orden  $m$ ,  $A = [a_i^j]$ ; si  $|a_j^j| > \sum_{i \neq j} |a_i^j|$   $j = 1, 2, \dots, m$ , entonces  $A$  es no-singular.

*Demostración.*

Si  $A$  es singular entonces existe un vector columna  $X \neq 0$  tal que  $AX = 0$ . Es decir, la solución del sistema homogéneo  $AX = 0$  no es sólo la trivial. Existe, pues

$$X = (x^1, x^2, \dots, x^m) \neq 0.$$

De dicho vector  $X$  tomemos la componente de mayor valor absoluto, es decir:  $|x^j| = \max |x^i|$ ,  $i = 1, 2, \dots, m$ . Es claro que  $x^j \neq 0$ ; de ahí que  $\sum_{i=1}^m a_i^j x^i = 0$ , implica  $x^j a_j^j = - \sum_{(i \neq j)=1}^m a_i^j x^i$ . De la última implicación se deduce, por simple álgebra, que

$$a_j^j = \sum_{(i \neq j)=1}^m - \frac{a_i^j x^i}{x^j}$$

Tomando el módulo a ambos lados de la ecuación anterior, luego aplicando la desigualdad triangular y, finalmente, con el hecho de que  $|x^j| = \max |x^i|$ ,  $i = 1, 2, \dots, m$  tenemos

las siguientes expresiones:

$$\left( |a_j^j| = \left| \sum_{(i \neq j)=1}^m \frac{a_i^j x^i}{x^j} \right| \right) \leq \sum_{(i \neq j)=1}^m \frac{|a_i^j| |x^i|}{|x^j|} \\ \leq \sum_{(i \neq j)=1}^m |a_i^j|.$$

Finalmente queda demostrado, de las expresiones anteriores,

que  $|a_j^j| = \sum_{(i \neq j)=1}^m |a_i^j|$ .

El lema de Gersgorin, aparentemente, sería el método numérico ideal para verificar si una matriz cuadrada es no-singular. La desventaja, obvia, del lema radica en que sólo plantea una condición suficiente para que una matriz sea no-singular. En otras palabras, si se satisface la hipótesis del lema estamos seguros que la matriz es no-singular, pero si la hipótesis no se satisface, no sabemos si la matriz es singular o no-singular. De todas formas el lema de Gersgorin es una buena alternativa numérica para el análisis de singularidad de una matriz.

### 3.3. Operatoria numérica sobre matrices

D-2 Operaciones entre filas. Dada una matriz  $A$ ,  $m$  por  $m$ , definimos como operaciones entre filas, en la matriz  $A$ , cualquiera de las siguientes tres operaciones:

- i)  $(i) \leftrightarrow (j)$ , Intercambiar la fila  $i$ -ésima con la fila  $j$ -ésima.
- ii)  $(i) \rightarrow c(i)$ , Sustituir la fila  $i$ -ésima por el producto de un número  $c \neq 0$  con la  $i$ -ésima fila.

iii)  $(i) \rightarrow (i) + c(j)$

Sustituir la fila  $i$ -ésima por la suma de la fila  $i$ -ésima con el producto de  $c \neq 0$  a la  $j$ -ésima fila.

D-3 Matriz elemental. Sea  $I_m$  la matriz identidad  $m$  por  $m$ , llamamos matriz elemental a cualquier matriz derivada de  $I_m$  por medio de aplicar, sólo una vez, cualquiera de las operaciones entre filas.

D-4 Matriz diagonal. Dada una matriz  $A$ ,  $m$  por  $m$ , llamamos matriz diagonal aquella que tiene ceros en cualquiera de sus elementos, excepto quizá, en la diagonal. Es decir,  $a_i^j = 0$  si  $i \neq j$ .

D-5 Matriz triangular. Dada una matriz  $A$ ,  $m$  por  $m$ , llamamos matriz triangular aquella que tiene ceros en todos los elementos que estén por debajo de la diagonal. Es decir,  $a_i^j = 0$ , si  $i < j$ .

D-6 Matriz equivalente por filas. Decimos que una matriz  $B$ ,  $m$  por  $n$ , es equivalente por filas a la matriz  $A$ ,  $m$  por  $n$ , si  $B$  se obtiene de la matriz  $A$  por medio de aplicar una sucesión finita de operaciones entre filas comenzado con  $A$  y continuando con las que genera  $A$ .

Ejemplo:

$B = \begin{pmatrix} 12 & 16 \\ 2 & 4 \end{pmatrix}$  es equivalente por filas a

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

dado que  $B$  se obtiene de  $A$  mediante las operaciones entre filas siguientes:

$$(1) \rightarrow 2(1)$$

$$(2) \rightarrow 4(2)$$

$$(1) \leftrightarrow (2).$$

T-4 Matriz identidad y operaciones entre filas. Da-

das dos matrices  $A, B$ , ambas  $m$  por  $n$ , el teorema que demostraremos aquí establece lo siguiente:

a) Si  $B$  es equivalente a  $A$ , por medio de una sólo operación entre filas, entonces existe una matriz elemental  $E$  tal que:

$$B = EA.$$

En dicha ecuación, la matriz elemental  $E$  se obtiene de la identidad aplicando la misma operación entre filas que se aplicó en  $A$  para pasar a  $B$ .

b) Finalmente, demostraremos que cualquier matriz elemental es no-singular.

*Demostración.*

Sea  $A$  una matriz  $m$  por  $n$  con elemento general  $a_{k}^p$ , e  $I_m$  la matriz identidad con elemento general  $e_p^r$ . Para la matriz  $B$ , equivalente a  $A$ , llamaremos  $b_k^p$  a su elemento general; mientras que el elemento general de la matriz elemental  $E$  será  $e_p^r$ .

a) Lo que hay que demostrar es que si se aplica alguna de las operaciones entre filas, dadas en D-2, en  $A$  para obte-

ner  $B$ , y la misma operación entre filas en  $I_m$  para obtener  $E$ , entonces  $B=EA$ . Como tenemos tres operaciones entre filas hay que probar que  $B=EA$  para cada una de las operaciones entre filas.

Al usar la operación  $(i) \leftrightarrow (j)$ ,  $B$  se obtiene de  $A$  intercambiando la fila  $i$ -ésima por la  $j$ -ésima, mientras que  $E$  se obtiene de  $I_m$  haciendo lo mismo. Las siguientes ecuaciones plantean el intercambio de filas explicado en el párrafo anterior:

$$2) \quad b_k^p = a_k^p, \text{ para toda } p \neq i \ \& \ p \neq j, \ k = 1, 2, \dots, n$$

$$b_k^i = a_k^j \ \& \ b_k^j = a_k^i, \ k = 1, 2, \dots, n.$$

$$3) \quad e_p^r = e_p^r, \text{ para toda } r \neq i \ \& \ r \neq j, \ p = 1, 2, \dots, m$$

$$e_p^i = e_p^j \ \& \ e_p^j = e_p^i, \ p = 1, 2, \dots, m.$$

Si tomamos el producto  $EA$ , con elemento general  $r$ -ésimo,  $k$ -ésimo,  $r \neq i \ \& \ r \neq j$ , apoyados en 2) y 3) tendremos

$$\sum_{p=1}^m e_p^r a_k^p = \sum_{p=1}^m e_p^r a_k^p = \sum_{p=1}^m \delta_p^r a_k^p = a_k^r = b_k^r, \ r \neq i,$$

$r \neq j$ . Hemos probado que  $B=EA$  para todas las filas excepto la  $i$ -ésima y la  $j$ -ésima. Recordemos que  $\delta_p^r = e_p^r = \begin{cases} 1, & \text{si } r=p \\ 0, & \text{si } r \neq p \end{cases}$ .

Las siguientes ecuaciones, basadas también en 2) y 3), prueban que  $B=EA$  para los casos que faltaban de la  $i$ -ésima y  $j$ -ésima fila.

$$\sum_{p=1}^m e_p^i a_k^p = \sum_{p=1}^m e_p^j a_k^p = \sum_{p=1}^m \delta_p^j a_k^p = a_k^j = b_k^i.$$

$$\sum_{p=1}^m e_p^j a_k^p = \sum_{p=1}^m e_p^i a_k^p = \sum_{p=1}^m \delta_p^i a_k^p = a_k^i = b_k^j.$$

Al usar la operación  $(i) \rightarrow c(i)$ , con  $c \neq 0$ , tendremos las siguientes ecuaciones, después de formar las matrices  $B$  y  $E$ , equivalentes por filas respectivamente de  $A$  e  $I_m$  por medio de la operación mencionada.

$$4) \quad \begin{aligned} b_k^p &= a_k^p \quad \text{para toda } p \neq i \quad \& \quad k=1,2,\dots,n \\ b_k^i &= ca_k^i, \quad k=1,2,\dots,n. \end{aligned}$$

$$5) \quad \begin{aligned} e_p^r &= e_p^r, \quad \text{para toda } r \neq i \quad \& \quad p=1,2,\dots,m \\ e_p^i &= ce_p^i, \quad p=1,2,\dots,m. \end{aligned}$$

De las ecuaciones 4) y 5) tenemos que el elemento general de  $EA$ , excepto para la fila  $i$ , viene dado así:

$$\sum_{p=1}^m e_p^r a_k^p = \sum_{p=1}^m e_p^r a_k^p = \sum_{p=1}^m S_p^r a_k^p = a_k^r = b_k^r, \quad r \neq i, \quad k=1,2,\dots,n. \quad \text{Para la fila } i \text{ tenemos:}$$

$$\sum_{p=1}^m e_p^i a_k^p = \sum_{p=1}^m ce_p^i a_k^p = c \sum_{p=1}^m S_p^i a_k^p = ca_k^i = b_k^i, \quad k=1,2,\dots,n. \quad \text{Así pues, de nuevo se cumple para esta operación que } EA = B.$$

Al usar  $(i) \rightarrow (i) + c(j)$  tendremos las siguientes ecuaciones:

$$\begin{aligned} b_k^p &= a_k^p, \quad \text{para toda } p \neq i \quad \& \quad k=1,2,\dots,n. \\ b_k^i &= a_k^i + ca_k^j, \quad k=1,2,\dots,n. \\ e_p^r &= e_p^r, \quad \text{para toda } r \neq i \quad \& \quad p=1,2,\dots,m. \\ e_p^i &= e_p^i + ce_p^j, \quad p=1,2,\dots,m. \end{aligned}$$

De las ecuaciones anteriores obtenemos que:

$$\sum_{p=1}^m e_p^r a_k^p = \sum_{p=1}^m e_p^r a_k^p = \sum_{p=1}^m S_p^r a_k^p = a_k^r = b_k^r, \quad r \neq i,$$

$k = 1, 2, \dots, n$ . Para el caso de la fila  $i$ -ésima tenemos que:

$$\sum_{p=1}^m e_p^i a_k^p = \sum_{p=1}^m (e_p^i + c e_p^j) a_k^p = \sum_{p=1}^m e_p^i a_k^p + c \sum_{p=1}^m e_p^j a_k^p =$$

$$\sum_{p=1}^m S_p^i a_k^p + c \sum_{p=1}^m S_p^j a_k^p = a_k^i + c a_k^j = b_k^i, \quad k = 1, 2, \dots, n.$$

Con ésto queda probado que  $EA = B$ .

Para concluir el teorema nos falta probar b), es decir, cualquier matriz elemental es no-singular.

b) Sea  $E$  la matriz elemental obtenida de  $I_m$ , intercambiando la fila  $i$ -ésima por la fila  $j$ -ésima. Para demostrar que  $E$  es no-singular, de acuerdo al teorema 2, basta con encontrar una matriz  $E^{-1}$  tal que  $EE^{-1} = I_m$ . Para el caso de la operación que intercambia filas basta hacer  $E^{-1} = E$  y tendremos que  $EE^{-1} = I_m$ . La demostración se logra, sencillamente, desarrollando el producto  $EE^{-1}$ . Para el caso de  $E$  obtenida de  $I_m$  mediante la operación  $(i) \rightarrow c(i)$ , con  $c$  constante, la inversa de  $E$ ,  $E^{-1}$ , viene dada por la matriz que se obtiene de  $I_m$  al aplicarle la operación entre filas  $(i) \rightarrow 1/c(i)$ . Desarrollando  $EE^{-1}$  comprobaremos que de hecho  $E^{-1}$  es la inversa de  $E$ . Finalmente, si  $E$  es la matriz elemental obtenida de  $I_m$  mediante la operación  $(i) \rightarrow (i) + c(j)$ , entonces la inversa se forma al aplicar sobre  $I_m$  la operación  $(i) \rightarrow (i) - c(j)$ .

Con lo anterior queda probado que cualquier matriz elemental es no-singular.

Corolario T-4. Sea  $B$  una matriz equivalente por filas a  $A$ , (mediante una sucesión finita de operaciones entre filas) existe una matriz no-singular  $E$  tal que:

$$B = EA.$$

$E$  se obtiene aplicando sobre  $I_m$  la misma sucesión de operaciones entre filas que se aplicó en  $A$  para obtener  $B$ .

*Demostración:*

Basados en el teorema 4 podemos generar el siguiente proceso. Si  $B$  es equivalente por filas a  $A$  entonces, como primer paso, se debió aplicar una operación entre filas sobre  $A$ . De lo anterior se concluye que existe una matriz elemental  $E_1$  que reproduce el primer paso sobre la matriz  $A$  para llegar a  $B$ . Lo anterior lo podemos representar así:

$$6) \quad B_1 = E_1 A.$$

Luego, sobre  $B_1$  siempre para llegar a  $B$ , se desarrolla otra operación entre filas que genera  $B_2$ , tal que:

$$7) \quad B_2 = E_2 B_1, \text{ con } E_2 \text{ matriz elemental.}$$

Finalmente, después de un número de  $n$  pasos, similares a los de 6) y 7), por medio de operaciones entre filas sucesivas, llegamos a  $B$ .

$$8) \quad B = E_n B_{n-1}, \text{ con } E_n \text{ matriz elemental.}$$

De acuerdo a las ecuaciones 6), 7) tenemos que, sustituyendo sucesivamente  $B_{n-1}, \dots, B_2, B_1$  en 8), llegamos a:

$$9) \quad B = E_n E_{n-1} \dots E_1 A.$$

Por el corolario del teorema 2, producto de matrices no-singulares es no-singular, tenemos que  $E_n E_{n-1} \dots E_1$  es una matriz no-singular. Es claro que si  $E = E_n E_{n-1} \dots E_1$ , entonces la ecuación 9) queda como  $B = EA$ , que es lo que queríamos demostrar. El procedimiento indica que la matriz  $E$  se obtiene de  $I_m$  por las mismas operaciones que se aplicaron sobre  $A$  para llegar a  $B$ . #

D-7 Relación y matrices equivalentes por filas. Dado el conjunto  $M_m^2$ , de todas las matrices  $m$  por  $m$  sobre un campo escalar, definimos una relación  $R$  sobre  $M_m^2$  de la siguiente manera:

$A \sim B$  (  $A$  relacionado con  $B$  ) si y sólo si  $B$  es equivalente por filas con  $A$ .

T-5 Matrices equivalentes definen una relación de equivalencia. Este teorema justifica el hecho de llamar a dos matrices, deducible una de la otra por medio de operaciones entre filas, matrices equivalentes por filas.

El teorema dice: la relación de matrices equivalentes por filas, definida en D-7, es una relación de equivalencia sobre el conjunto de todas las matrices cuadradas  $M_m^2$ .

Demostración.

Lo que hay que probar es lo siguiente:

- i.  $A \sim A$  (reflexividad)
- ii.  $A \sim B$  entonces  $B \sim A$ . (simetría)
- iii.  $A \sim B$  &  $B \sim C$  entonces  $A \sim C$  (transitividad).

La demostración de estos tres incisos es bastante clara, siempre y cuando no se desarrolle una prueba rigurosa, la cual es un ejercicio engorroso de notación.

i. Probar que  $A \sim A$  queda establecido sencillamente al aplicarse sobre  $A$  la operación entre filas  $(i) \rightarrow 1(i)$ . La cual genera de nuevo la matriz  $A$ .

ii. Probar que  $A \sim B$  implica que  $B \sim A$  procede así: para regresar de  $B$  a  $A$  sólo se necesita invertir las operaciones que generaron  $B$  a partir de  $A$ . Por ejemplo, si en un paso, para ir de  $A$  a  $B$ , se usó la operación  $(i) \leftrightarrow (j)$  en el regreso de  $B$  a  $A$  se usa la operación  $(j) \leftrightarrow (i)$ ; si en otro paso, para ir de  $A$  a  $B$ , se usó la operación  $(i) \rightarrow c(i)$  entonces para regresar de  $B$  a  $A$  se usa la operación  $(i) \rightarrow 1/c(i)$ ; finalmente, si se usó la operación  $(i) \rightarrow (i) + c(j)$ , para ir de  $A$  a  $B$ , la operación por usar para ir de  $B$  a  $A$  será  $(i) \rightarrow (i) - c(j)$ .

iii. Probar que  $A \sim B$  &  $B \sim C$  implica que  $A \sim C$  se establece porque de  $A$  a  $B$  se llega por  $n$  operaciones entre filas, luego de  $B$  a  $C$  se llega por  $m$  operaciones: finalmente, de  $A$  a  $C$  se llega con las mismas  $n+m$  operaciones entre filas de  $A$  a  $B$  y de  $B$  a  $C$  respectivamente.

Dado que una relación de equivalencia induce una partición de clases disjuntas en el conjunto sobre el cual fue definida, con la característica de que los elementos de cada clase están relacionados entre sí; tendremos que  $M_m^2$  que

da partido en clases de equivalencia. La clase de equivalencia correspondiente a la identidad  $I_m$  vendrá a ser el conjunto de todas las matrices no-singulares. Esta última afirmación viene a ser parte del siguiente teorema.

T-6 Matrices, clases de equivalencia, rango y matriz no-singular. Dos matrices equivalentes por filas tienen el mismo rango. La clase de equivalencia a la que pertenece la matriz identidad es el conjunto de todas las matrices no-singulares.

*Demostración:*

La demostración de este teorema la expondremos luego de desarrollar el teorema 7. La razón de ésto obedece a que el teorema 7 servirá para demostrar este teorema 6.

### 3.4. Rango, condiciones de existencia y solución de $L_{mn}$

T-7  $L_{mn}$  y matrices equivalentes por filas. Sea un sistema lineal  $L_{mn}$  dado por  $AX=Y$ . Si  $B$  es equivalente por filas a  $A$ , entonces el sistema  $AX=Y$  es equivalente al sistema  $BX=PY$ ; en donde  $B=PA$ , con  $P$  la matriz no-singular obtenida de  $I_m$  mediante las mismas operaciones entre filas que llevaron a obtener  $B$  a partir de  $A$ . Además, la matriz aumentada  $B|PY$  es equivalente por filas a la matriz aumentada  $A|Y$ , mediante las mismas operaciones entre filas que llevaron a obtener  $B$  a partir de  $A$ .

*Demostración.*

$AX=Y$  es equivalente a  $BX=PY$  quiere decir que cualquier solución de un sistema también es solución del otro.

Sea  $Z$  solución de  $AX=Y$ . Ya que  $B=PA$  tenemos que

$$BZ = (PA)Z,$$

de donde por asociatividad del producto de matrices

$$BZ = P(AZ).$$

Luego, ya que  $Z$  es solución de  $AX=Y$ ,

$$10) \quad BZ = PY.$$

La última ecuación nos muestra que  $Z$  es solución de  $BX=PY$ .

Sea  $Z$  solución de  $BX=PY$ , entonces tendremos que,

$$BZ = PY.$$

Luego, por ser  $P$  no-singular,

$$P^{-1}BZ = Y.$$

De ahí, ya que  $B=PA$  tendremos

$$P^{-1}PAZ = Y;$$

de donde, finalmente,

$$11) \quad AZ = Y.$$

La última ecuación muestra que  $Z$  es solución de  $AX=Y$ .

Las ecuaciones 10) y 11) prueban que  $AX=Y$  y  $BX=PY$  son equivalentes.

Finalmente para demostrar que  $(A|Y) \sim (B|PY)$ , mediante las mismas operaciones entre filas que generaron  $B$  a partir de  $A$ , sólo hay que probar que:

$$B|PY = P(A|Y).$$

La ecuación anterior es clara, ya que el producto  $P(A|Y)$  da

como resultado para las primeras  $n$  columnas la matriz  $PA$ , pero  $PA = B$ , y esas  $n$  columnas son las mismas primeras  $n$  columnas de  $B|PY$ . La última columna de  $P(A|Y)$  viene a ser  $PY$  que es también la última columna de  $B|PY$ .

Este teorema 7 nos garantiza que las operaciones entre filas son un procedimiento numérico válido para transformar un sistema  $L_{mn}$  sin cambiar sus soluciones. De hecho las operaciones entre filas, como veremos en esta sección y la siguiente, son la técnica numérica que ofrecerá un resultado idóneo para resolver un sistema  $AX = Y$ .

Demostración teorema 6. Sea  $A$ ,  $m$  por  $n$ , matriz equivalente a  $B$ ,  $m$  por  $n$ . De acuerdo al teorema 2.4.a del capítulo II, si  $r_1$  es el rango de  $A$  y  $r_2$  el rango de  $B$ , entonces la dimensión del conjunto solución de  $AX = 0$  es  $n - r_1$  y la dimensión del conjunto solución de  $BX = 0$  es  $n - r_2$ . Por el teorema 7, el conjunto solución de  $AX = 0$  y  $BX = 0$  es el mismo, ya que  $A$  es equivalente por filas a  $B$ . Por lo tanto, la dimensión de dichos conjuntos es idéntica, es decir,

$$n - r_1 = n - r_2 .$$

De la ecuación anterior es claro que  $r_1 = r_2$ : el rango de  $A$  y  $B$  es el mismo.

La clase de equivalencia de  $I_m$  la forman todas las matrices equivalentes por filas a la identidad; por el resultado anterior de que dos matrices equivalentes tienen el mismo rango tenemos que todas las matrices de la clase  $I_m$

tienen el rango de  $I_m$ : rango que es justamente  $m$ . Por el enunciado B) y el teorema 2, de la sección 3.2, dicha clase es la de todas las matrices no-singulares. #

D-8 Matrix en la forma normal de Hermite. El tipo de matriz que aquí definiremos también recibe el nombre de matriz escalonada en la forma canónica, o bien matriz en la forma normal echelon por filas. Este tipo de matriz será de utilidad clave para determinar, numéricamente, el rango de una matriz, condiciones de existencia de un sistema  $L_{mn}$ , y también soluciones de un sistema  $L_{mn}$ .

Definimos como una matriz en la forma normal de Hermite a cualquier matriz que cumpla las siguientes tres propiedades:

- 1) Para cada fila de la matriz el primer elemento que aparezca diferente de cero, si es que existe, debe ser uno y se llama elemento distinguido.
- 2) Conforme bajamos una fila en la matriz, el elemento distinguido, si es que existe, debe aparecer a la derecha del elemento distinguido de la fila precedente.
- 3) En la columna donde aparece un elemento distinguido, todos los demás elementos de la columna deben ser cero.

Las mismas propiedades se pueden definir, mediante símbolos de la siguiente manera:

Sea  $A = [a_i^j]$  una matriz  $m$  por  $n$ , decimos que dicha matriz está en la forma normal de Hermite si existen  $r$  elementos

$$a_{c_1}^1, a_{c_2}^2, \dots, a_{c_r}^r, \quad 0 \leq r \leq m;$$

llamados elementos distinguidos que cumplen con:

- i)  $c_1 < c_2 < \dots < c_r$
- ii)  $a_{c_k}^k = 1, k = 1, 2, \dots, r$
- iii)  $a_i^j = 0$ , para  $i < c_k$  &  $j = k, \dots, m$
- iv)  $a_{c_k}^j = 0$ , para toda  $j \neq k$  &  $k = 1, 2, \dots, r$ .

Un ejemplo de este tipo de matriz es el siguiente:

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Observe, en el ejemplo anterior, que los elementos distinguidos ocurren en las columnas  $c_1 = 3, c_2 = 4, c_3 = 6, c_4 = 7$ .

Por la definición, la matriz cero es una matriz que ya está en la forma normal de Hermite.

T-8 Matrices y forma normal de Hermite. A) Cualquier matriz  $A$ ,  $m$  por  $n$ , se puede reducir aplicándole operaciones entre filas a una forma normal de Hermite.

B) La forma normal de Hermite, para una matriz dada, es única.

*Demostración.*

A) Lo demostración de que cualquier matriz  $A$  se puede

reducir a una forma de Hermite, es interesante porque plantea un algoritmo numérico que ayudará a resolver el rango de una matriz, las condiciones de existencia de un sistema  $L_{mn}$  y también dará soluciones a un sistema  $L_{mn}$ . Estos resultados, consecuencia de la demostración de este inciso A), serán presentados en un corolario a este teorema.

Dada la matriz  $A$  de elemento general  $a_i^j$ , buscamos la primera columna donde no todos los elementos sean cero. Llamamos a dicha columna  $c_1$ , y es claro, que por definición cumple con que:

$$12) \quad a_i^j = 0, \text{ si } i < c_1 \text{ \& } j = 1, 2, \dots, m; c_1 \leq n.$$

Si dicha columna no existe entonces  $A = 0$ , la cual ya está en la forma normal de Hermite.

Con la columna  $c_1$ , escogemos al elemento de mayor valor absoluto; supongamos que fuera  $a_{c_1}^p$ , entonces intercambiamos la fila  $p$ -ésima con la fila uno. Lo anterior lo podemos resumir diciendo que formamos una matriz equivalente a  $A$  haciendo la operación entre filas  $(p) \leftrightarrow (1)$ . Esta matriz equivalente a  $A$ , por medio de la mencionada operación, la llamaremos nuevamente  $A$  renombrando (redefiniendo) los elementos de la fila 1 por la fila  $p$ , y los de la fila  $p$  por la uno.

La matriz que buscamos,  $A_1 = \begin{bmatrix} 1 & j \\ a_i^j \end{bmatrix}$ , equivalente a la matriz  $A$  la construimos así:

$$13) \quad a_i^1 = \frac{a_i^p}{a_{c_1}^p}, \quad i > c_1.$$

La ecuación 13) corresponde a la operación entre filas,

$$(1) \rightarrow \frac{1}{a_{c_1}^1}(1),$$

operación que define la nueva fila 1. Luego, las demás filas de  $A_1$  las construimos así:

$$14) \quad a_i^j = a_i^j - a_{c_1}^j a_i^1, \quad i \geq c_1 \text{ \& } j \neq 1.$$

La ecuación anterior es la operación entre filas,

$$(j) \rightarrow (j) - a_{c_1}^j (1), \quad j \neq 1,$$

que define a todas las filas de la nueva matriz  $A_1$ , excepto la fila 1 que está definida por la ecuación 13).

De la ecuación 13) verificamos que el elemento distinguido efectivamente cumple con que:

$$15) \quad a_{c_1}^1 = 1.$$

También se concluye que,

$$16) \quad a_i^j = 0, \text{ para } i < c_1 \text{ \& } j = 1, 2, \dots, m.$$

La ecuación anterior está justificada por la definición de la columna  $c_1$ .

La ecuación 14) nos conduce a que,

$$a_{c_1}^j = 0, \quad j \neq 1.$$

Con la última ecuación se completan los requerimientos de las propiedades ii), iii) y iv), de la definición D-8, para el caso de  $c_1$ .

Las ecuaciones 13), 14) y 16) son las que definen la matriz  $A_1$ , equivalente a la matriz  $A$ . Un esquema del proceso

anterior, para observar parcialmente lo que se ha hecho, lo presenta el siguiente ejemplo; en donde los # representan números que pueden ser o no cero.

$$\begin{pmatrix} 1 & \# & \# & \# & \# & \# \\ 0 & \# & \# & \# & \# & \# \\ 0 & \# & \# & \# & \# & \# \\ 0 & \# & \# & \# & \# & \# \end{pmatrix} \text{ En este ejemplo } c_1 = 1.$$

Para formar la matriz  $A_2$  equivalente a la  $A_1$  repetimos los mismos pasos que se hicieron para alcanzar la  $A_1$ , sólo que sobre la siguiente fila. Es decir, pasamos a la fila 2 y buscamos la primera columna,  $c_2$ , debajo de la primera fila, donde no todos los elementos sean cero. Es obvio, a causa de 16), que  $c_1 < c_2$ . Nuevamente hay que escoger en la columna  $c_2$ , variando de la fila 2 hasta la  $m$ , el elemento que sea mayor en valor absoluto. Ya encontrado este elemento se intercambia la fila donde está dicho elemento con la fila 2. Un esquema del paso anterior es el siguiente ejemplo, donde los # representan números que pueden ser o no cero.

$$\begin{pmatrix} 1 & \# & 0 & \# & \# & \# \\ 0 & 0 & 1 & \# & \# & \# \\ 0 & 0 & 0 & \# & \# & \# \\ 0 & 0 & 0 & \# & \# & \# \end{pmatrix} \text{ En este ejemplo } c_1 = 1 \text{ y } c_2 = 3.$$

Hay que notar que conforme se avanza un paso, las columnas anteriores a donde está el elemento distinguido, ya no

cambian.

El proceso se repite de la misma manera, siempre bajando una fila en las nuevas matrices equivalentes que se van obteniendo, hasta que suceda alguno de los siguientes casos: exista algún  $c_r$  tal que,

a)  $r = m$

b)  $c_r = n$

c) Todos los elementos de las columnas por debajo de la fila  $r$  son igual a cero.

Para un  $k$ -ésimo paso intermedio, es decir,  $1 < k < r$  tendremos que la matriz  $A_k$  viene definida así:

$$17) \quad a_i^{kk} = \frac{a_i^{k-1k}}{c_k}, \quad i \geq c_k$$

La ecuación anterior define la nueva fila  $k$  mediante la operación entre filas  $(k) \rightarrow \frac{1}{c_k} (k)$ . Luego tenemos la siguiente ecuación que define las restantes filas, para  $i \geq c_k$

$$18) \quad a_i^{kj} = a_i^{k-1j} - \frac{a_i^{k-1k} a_i^{kj}}{c_k}, \quad i \geq c_k \quad \& \quad j \neq k.$$

Las columnas anteriores a  $c_k$  se conservan igual que la matriz  $A_{k-1}$ ; es decir

$$19) \quad a_i^{kj} = a_i^{k-1j} \quad \text{si} \quad i < c_k.$$

Debe quedar claro que en cualquier paso  $k$ -ésimo  $1 < k < r$ , los elementos de la matriz  $A_k$ ,  $a_i^{kj}$ , con  $i < c_k$  &  $j = 1, 2, \dots, m$ , permanecen igual que los elementos correspondientes en

la matriz  $A_{k-1}$ . Lo recién dicho, es parte del contenido de la ecuación 19).

Supongamos que en  $r$  pasos llegamos a alguno de los casos a), b) o c), entonces lo que queremos demostrar es que la matriz  $A_r$ , equivalente a  $A$ , está en la forma normal de Hermite. Es decir, cumple i), ii), iii) y iv) dados en la definición 8 de este capítulo.

i) Esto queda demostrado por la forma sucesiva de construcción de la matriz equivalente  $A_k$ .

ii) Que  $\frac{r^k}{c_k} = 1$ ,  $k=1,2,\dots,r$  lo demuestra la ecuación 17) y la ecuación 19).

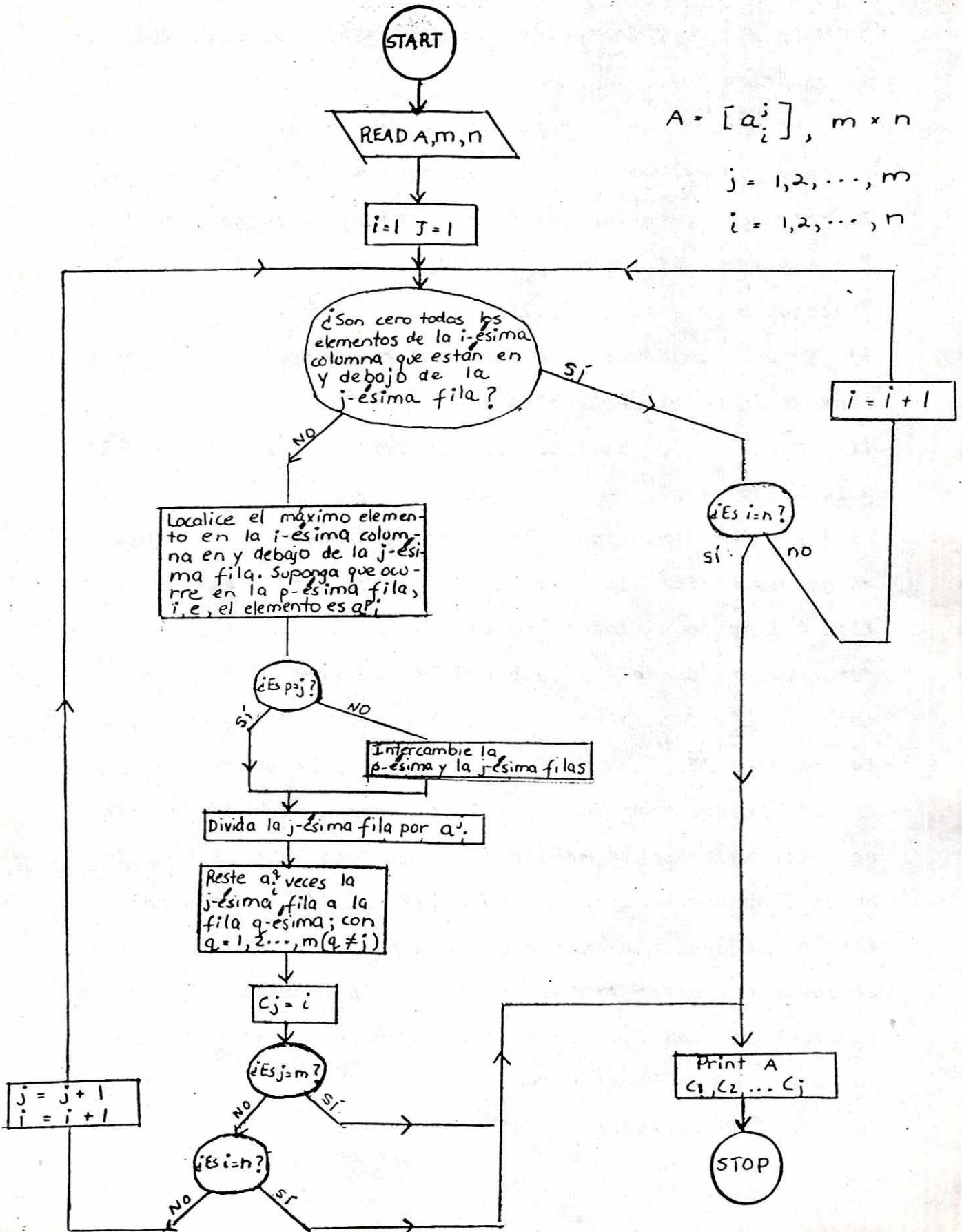
iii) La demostración de este inciso proviene de la forma en que se definen los  $c_k$ : la primera columna debajo de la fila  $k-1$  donde no todos los elementos sean cero. Además la forma recursiva de 19) completa la demostración.

iv) Que  $\frac{r^j}{c_k} = 0$  para toda  $j \neq k$ ,  $k=1,2,\dots,r$  lo demuestra la ecuación 18), cuando  $k=r$  &  $i=c_r$ , y la ecuación 19).

B) La demostración de que la forma normal de Hermite es única no presenta mayores dificultades, sin embargo la omitiremos por no aportar mayor información al problema numérico que aquí nos interesa. Esta parte de la demostración se puede consultar en Noble(1969:82). A continuación presentamos un diagrama de flujo, para reducir una matriz a la forma normal de Hermite, que resume toda la información anterior, tomado, también, de Noble(1969:83)

Diagrama de flujo para la forma normal de Hermite

102



Corolario T-8. 1) *Rango de una matriz.* El rango de una matriz  $A$  corresponde al número de filas de la matriz menos aquellas filas en donde no todos los elementos sean cero, cuando  $A$  ha sido reducida a la forma normal de Hermite. También, por supuesto, el rango de la matriz corresponde al número de elementos distinguidos.

*Demostración.*

La demostración se basa en el teorema 8 y el aspecto que presenta la forma normal de Hermite. Como la matriz  $A$  tiene a su forma normal de Hermite como una matriz equivalente, entonces el teorema 6 nos garantiza que ambas matrices tienen el mismo rango. Luego, viendo el aspecto de una forma de Hermite, notamos que cada fila que tenga todos sus elementos cero es linealmente dependiente con las otras filas; mientras que las filas que no son cero son linealmente independientes a causa de que en cada columna donde aparece un elemento distinguido, todos los demás elementos de dicha columna son cero. Es claro, que el número de elementos distinguidos corresponde al rango de la matriz porque hay elemento distinguido en cada fila donde no todos los elementos sean cero.

ii) *Condiciones de existencia de  $L_{mn}$ .* Sea un sistema lineal  $L_{mn}$ , dado por  $AX=Y$ , dicho sistema tiene solución si y sólo si el número de elementos distinguidos de la forma de Hermite de  $A$ , es el mismo número de elementos distinguidos

de la forma de Hermite de  $A|Y$ . Una forma equivalente, de decir lo anterior, consiste en: para la forma de Hermite de la matriz aumentada  $A|Y$ , en la última columna de dicha matriz no aparece ningún elemento distinguido.

*Demostración.*

Debe estar claro que; en la forma de Hermite de  $A|Y$  al quitarle la última columna nos da la forma de Hermite de la matriz  $A$ . La demostración se basa en la condición de existencia dada en la ecuación 36), capítulo II:  $AX = Y$  tiene solución si y sólo si el rango de  $A$  es igual al rango de la matriz aumentada  $A|Y$ . Con este teorema, el corolario i) dado aquí, y el contenido expresado al principio de esta demostración, tenemos que la prueba está completa.

La forma equivalente para este teorema se prueba porque; el único caso en que puede diferir el rango de las formas de Hermite  $A$  y  $A|Y$  consiste cuando un elemento distinguido está en la última columna de la forma de Hermite de  $A|Y$ .

iii) Soluciones de  $L_{mn}$ . Si tenemos un sistema  $AX = Y$ , entonces las soluciones, si es que existen, vendrán directamente de la forma de Hermite para la matriz aumentada  $A|Y$ .

*Demostración.*

Las ecuaciones principales, ver sección 2.10 capítulo II, corresponden a las filas donde aparezcan elementos distinguidos; esto porque el número de elementos distinguidos coincide con el rango de la matriz y porque las únicas fi-

las donde no todos los elementos son cero son las de elementos distinguidos. Si el rango del sistema es  $r$  entonces  $r$  son las ecuaciones principales; y si hay solución, las incógnitas que se pueden seleccionar arbitrariamente son  $n-r$ : ver sección 2.3, capítulo II. Dado que en las columnas donde aparecen elementos distinguidos, excepto dicho elemento, todos los demás son cero, tenemos que las incógnitas que se seleccionan arbitrariamente son las que corresponden a las columnas donde no hay elemento distinguido, mientras que las incógnitas de columnas de elementos distinguidos quedan determinadas. #

Los siguientes ejemplos corresponden a sistemas  $L_{mn}$ , en donde  $A|Y$ , la matriz aumentada del sistema  $L_{mn}$ , ha sido reducida a la forma normal de Hermite, que es la que se presenta en cada caso.

a) 
$$\left( \begin{array}{cccc|c} 1 & 2 & 0 & 4 & 3 \\ 0 & 0 & 1 & -5 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad \begin{array}{l} - \text{ Sistema } L_{mn}, 4 \text{ por } 4. \\ - \text{ Rango del sistema } 2. \\ - \text{ Tiene infinitas soluciones; ver ii) del corolario anterior.} \end{array}$$

Soluciones: guiados por iii), del corolario anterior, tenemos que si  $x_2 = t_1$ ,  $x_4 = t_2$ , incógnitas que se pueden seleccionar arbitrariamente, entonces  $x_1 = -2t_1 - 4t_2 + 3$  &  $x_3 = 5t_2 + 2$ .

b) El siguiente ejemplo corresponde a un sistema  $L_{mn}$ , 4 por 4. El rango del sistema es 2 y no existe solución del

sistema :ver ii) del corolario anterior.

$$\left( \begin{array}{cccc|c} 1 & 0 & 2 & 4 & 0 \\ 0 & 1 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

c) 
$$\left( \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -4 \end{array} \right) \quad \begin{array}{l} - \text{ Sistema } L_{mn}, 3 \text{ por } 3. \\ - \text{ Rango del sistema } 3. \\ - \text{ Hay solución y es única.} \end{array}$$

Soluciones: guiados por iii), del corolario anterior, tenemos que,  $x_1=2$ ,  $x_2=3$ ,  $x_3=-4$ .

d) 
$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 2 & 4 \\ 0 & 1 & 0 & -3 & 5 \\ 0 & 0 & 1 & -1 & -\frac{1}{2} \end{array} \right) \quad \begin{array}{l} - \text{ Sistema } L_{mn}, 3 \text{ por } 4. \\ - \text{ Rango del sistema } 3. \\ - \text{ Hay infinitas soluciones.} \end{array}$$

Soluciones: La única incógnita arbitraria, en este caso, es  $x_4=t$ . Luego, las otras quedan determinadas así:  $x_1=-2t+4$ ,  $x_2=3t+5$ ,  $x_3=t-\frac{1}{2}$ .

e) 
$$\left( \begin{array}{cccc|c} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & -1/3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad \begin{array}{l} - \text{ Sistema } 5 \text{ por } 4. \\ - \text{ Rango } 3. \\ - \text{ Hay infinitas soluciones.} \end{array}$$

Soluciones: Para este caso, sólo hay una incógnita arbitra-

ria, y es  $x_2 = t$ . Las otras incógnitas quedan determinadas como  $x_1 = -2t + 1$ ,  $x_3 = -2$ ,  $x_4 = -1/3$ .

f) 
$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} - \text{ Sistema 5 por 4.} \\ - \text{ Rango 4.} \\ - \text{ La solución es única.} \end{array}$$

Soluciones:  $x_1 = 6$ ,  $x_2 = -\frac{1}{2}$ ,  $x_3 = 2$ ,  $x_4 = 1$ .

Nota: Observe que en un sistema  $L_{mn}$  si tenemos que  $n > m$ , más incógnitas que ecuaciones, sólo pueden suceder dos cosas:

- a) El sistema no tiene solución.
- b) Si el sistema tiene solución, entonces hay infinitas soluciones.

La razón de b) estriba en que el rango, a lo sumo, puede ser  $r = m$  (teorema de dualidad), por lo que las incógnitas arbitrarias son  $n - m$ ; y de ahí que se generan infinitas soluciones.

Para el caso de más ecuaciones que incógnitas ( $m > n$ ), la solución puede ser única, ver ejemplo f; pueden existir infinitas soluciones, ver ejemplo e); o puede no existir solución.

### 3.5. Eliminación de Gauss

La eliminación de Gauss es un método para resolver un sistema lineal  $L_{mn}$ . Esencialmente, es lo mismo que el procedimiento usado para llegar a la forma normal de Hermite, excepto por las columnas donde están los elementos distinguidos. En la forma de Hermite, hacemos cero a todos los elementos de una columna donde hay elemento distinguido, con la excepción, por supuesto, del mismo elemento distinguido; en la eliminación de Gauss, hacemos cero sólo los elementos que están por debajo del elemento distinguido. Aparte de esa diferencia, todo el método para llegar a la eliminación de Gauss es el mismo que el de la forma de Hermite. Si vemos la definición D-8, para tener una forma de Hermite, comprobamos que la eliminación de Gauss, tal y como se definió, debe cumplir con todas las condiciones, excepto la condición iv); que como dijimos sólo vale para los elementos por debajo del elemento distinguido. Dado que la eliminación de Gauss es un caso particular de la forma normal de Hermite, a causa de los razonamientos anteriores, resulta que las condiciones de existencia, rango y solución se deducen igual en la eliminación de Gauss y en la forma de Hermite.

Veamos un ejemplo, desarrollado en forma completa, para comparar la eliminación de Gauss y la forma normal de Hermite.

Ejemplo.

Sea el sistema 3 por 3

$$2x - 3y + 4z = 1$$

$$20) \quad x + y - z = 2$$

$$3x - y - 2z = 3.$$

Si llamamos

$$A = \begin{pmatrix} 2 & -3 & 4 \\ 1 & 1 & -1 \\ 3 & -1 & -2 \end{pmatrix} \quad Y = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

entonces la forma matricial del sistema viene dada por  $AX=Y$ .  
Lo único que necesitamos para trabajar la forma de Hermite o la eliminación de Gauss es la matriz aumentada

$$A | Y = \left( \begin{array}{ccc|c} 2 & -3 & 4 & 1 \\ 1 & 1 & -1 & 2 \\ 3 & -1 & -2 & 3 \end{array} \right).$$

Desarrollaremos sobre la matriz  $A | Y$  la forma de Hermite y la eliminación de Gauss, simultáneamente, por el método delineado en el diagrama de flujo de la página 102.

Si hacemos la operación entre filas  $(3) \leftrightarrow (1)$  y, luego,  $(1) \rightarrow \frac{1}{a_1} (1)$ , entonces tendremos sucesivamente las siguientes matrices.

$$\left( \begin{array}{ccc|c} 3 & -1 & -2 & 3 \\ 1 & 1 & -1 & 2 \\ 2 & -3 & 4 & 1 \end{array} \right); \left( \begin{array}{ccc|c} 1 & -1/3 & -2/3 & 1 \\ 1 & 1 & -1 & 2 \\ 2 & -3 & 4 & 1 \end{array} \right).$$

Luego, procedemos con las operaciones entre filas  $(2) \rightarrow (2) - a_1^2(1)$ , y  $(3) \rightarrow (3) - a_1^3(1)$ , las cuales conducen a la siguiente matriz

$$\left( \begin{array}{ccc|c} 1 & -1/3 & -2/3 & 1 \\ 0 & 4/3 & -1/3 & 1 \\ 0 & -7/3 & 16/3 & -1 \end{array} \right).$$

Con la última matriz hemos desarrollado tanto la eliminación de Gauss como la forma de Hermite para la primera columna, que es donde apareció el primer elemento distinguido:  $c_1 = 1$ .

Como el siguiente elemento distinguido está en la segunda columna  $c_2 = 2$ , hacemos las siguientes operaciones entre filas:

$$(2) \leftrightarrow (3) \quad \& \quad (2) \rightarrow \frac{1}{a_2} (2).$$

Estas operaciones conducen a las matrices

$$\left( \begin{array}{ccc|c} 1 & -1/3 & -2/3 & 1 \\ 0 & -7/3 & 16/3 & -1 \\ 0 & 4/3 & -1/3 & 1 \end{array} \right); \left( \begin{array}{ccc|c} 1 & -1/3 & -2/3 & 1 \\ 0 & 1 & -16/7 & 3/7 \\ 0 & 4/3 & -1/3 & 1 \end{array} \right).$$

Luego, para la eliminación de Gauss sólo tenemos que hacer la operación  $(3) \rightarrow (3) - a_2^3(2)$ ; mientras que para la forma de Hermite hacemos  $(3) \rightarrow (3) - a_2^3(2)$  &  $(1) \rightarrow (1) - a_2^1(2)$ . Las matrices que obtenemos, al hacer las operaciones anteriores, son las siguientes

$$\left( \begin{array}{ccc|c} 1 & -1/3 & -2/3 & 1 \\ 0 & 1 & -16/7 & 3/7 \\ 0 & 0 & 19/7 & 3/7 \end{array} \right); \left( \begin{array}{ccc|c} 1 & 0 & -10/7 & 8/7 \\ 0 & 1 & -16/7 & 3/7 \\ 0 & 0 & 19/7 & 3/7 \end{array} \right).$$

Para eliminación de Gauss. Para la forma de Hermite.

Finalmente, el último elemento distinguido estará en la

tercer columna,  $c_3 = 3$ . Por lo tanto para empezar hacemos, en ambas matrices de arriba, la operación  $(3) \rightarrow \frac{1}{a_3} (3)$ . Con dicha operación obtenemos

$$21) \left( \begin{array}{ccc|c} 1 & -1/3 & -2/3 & 1 \\ 0 & 1 & -16/7 & 3/7 \\ 0 & 0 & 1 & 3/19 \end{array} \right); \left( \begin{array}{ccc|c} 1 & 0 & -10/7 & 8/7 \\ 0 & 1 & -16/7 & 3/7 \\ 0 & 0 & 1 & 3/19 \end{array} \right).$$

Para eliminación de Gauss.      Para la forma de Hermite.

La eliminación de Gauss ya está completa, mientras que para la forma de Hermite hay que hacer cero toda la tercer columna, excepto donde está el elemento distinguido. Lo anterior se logra con las operaciones  $(2) \rightarrow (2) - a_3^2(3)$  y  $(1) \rightarrow (1) - a_3^1(3)$ . Dichas operaciones conducen a la forma de Hermite

$$22) \left( \begin{array}{ccc|c} 1 & 0 & 0 & 26/19 \\ 0 & 1 & 0 & 15/19 \\ 0 & 0 & 1 & 3/19 \end{array} \right) \quad \text{Forma de Hermite.}$$

Nuestro objetivo es, ahora, encontrar las soluciones del sistema original, ecuaciones 20), a partir de 21) (eliminación de Gauss) y de la forma de Hermite, 22). Ambas matrices, las de la ecuación 21) y 22), nos dan tres elementos distinguidos; por lo tanto el rango es 3, y como el sistema es 3 por 3, la solución es única.

Directamente de la forma normal de Hermite, ecuación 22), tenemos que las soluciones son

$$x = 26/19, \quad y = 15/19, \quad z = 3/19.$$

Para encontrar las soluciones de 20), mediante la eliminación de Gauss, hay que hacer un proceso de retroceso, puesto que sólo tenemos que  $x = 3/19$ . El cálculo de retroceso consiste en encontrar  $y$  a partir de  $x$ , y luego  $x$  a partir de  $x$  &  $y$ .

$$24) \quad y = 16/7(3/19) + 3/7 = 15/19$$

$$x = 1/3(15/19) + (2/3)(3/19) + 1 = 26/19.$$

Es claro, las soluciones dadas en 23), 24) son las mismas, y de hecho son soluciones, lo cual se puede comprobar si las sustituimos en la ecuación 20) que da el sistema original.

La nueva pregunta por hacer es cuál de los dos métodos emplea menos operaciones para llegar al resultado. Para fines de una máquina calculadora, las operaciones que requieren tiempo significativo son las multiplicaciones y las divisiones; por lo tanto sólo contaremos dichas operaciones. Tomemos un sistema lineal  $L_n^2$  y supongamos que existe solución única: en este caso es cuando más operaciones hay que hacer. Bajo estas suposiciones, la cantidad de operaciones que hay que hacer, en la eliminación de Gauss, para llegar a la matriz triangular es de

$$1^2 + 2^2 + \dots + n^2 = 1/3 n^3 + \frac{1}{2}n^2 + 1/6 n.$$

Luego, para el proceso de retroceso, la cantidad de operaciones es de

$$1 + 2 + \dots + (n-1) = n(n-1)/2 .$$

Sumando las operaciones para llegar a la matriz triangular con las del proceso de retroceso tenemos que Gauss emplea un total de operaciones de

$$25) \quad O_G = 1/3 n^3 + n^2 - 1/3 n.$$

Por otro lado el total de operaciones que emplea la forma de Hermite es de

$$26) \quad O_H = n(1 + 2 + \dots + (n-1) + n) = n^3/2 + n^2/2.$$

Los resultados de 25) y 26) se logran por simple inspección del método para lograr la forma de Hermite y la eliminación de Gauss.

Para un sistema 3 por 3, como el del ejemplo, la forma de Hermite requiere 18 productos y divisiones; el método de Gauss requiere 16. Para un tamaño pequeño, como el del ejemplo, la diferencia en la cantidad de operaciones empleada por los dos métodos es mínima. Pero si el sistema fuera 50 por 50, el método de Gauss requeriría, de acuerdo a 25), 44150 productos y divisiones, mientras que el de Hermite requeriría 63750. En este caso la diferencia, 19600, ya empieza a ser un poco significativa. Lo anterior es la razón por la cual entre Gauss y Hermite, casi siempre se trabaja con Gauss; y los libros, también, casi siempre sólo exponen el método de Gauss.

### 3.6. Cálculo de determinantes y matriz inversa

a) Determinantes. Dado un determinante de orden  $n$

$$\Delta = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{vmatrix} .$$

Para calcular dicho determinante nos apoyaremos en las propiedades de determinantes

- i) P-1  $D(A_1, \dots, \alpha A_j, \dots, A_n) = \alpha D(A_1, \dots, A_j, \dots, A_n)$   
 ii) P-6  $D(A_1, A_2, \dots, A_j, \dots, A_n) = D(A_1, A_2, \dots, B_j, \dots, A_n)$ ,  
 con  $B_j = A_j + \sum_{i \neq j}^n \alpha_i A_i$ .  
 iii)  $D(A) = D(A^t)$ .

Todas las propiedades anteriores fueron dadas en el capítulo I. Además, usaremos la expansión de un determinante por una columna, dada en el capítulo I, sección 1.4.

Las propiedades i) y ii) nos permiten trabajar operaciones entre filas, en el determinante, de forma parecida a como se hace en una matriz. Al final la expansión la haremos por la primera columna. Veamos como se procede

$$\Delta = a_1^1 \begin{vmatrix} 1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \dots & a_n^n \end{vmatrix}$$

El determinante anterior se obtuvo aplicando la propiedad i) y iii); en donde la primera fila se trabajó como si

hubiera actuado la operación entre filas (1)  $\rightarrow \frac{1}{a_1}(1)$ .

Luego, si usamos ii) tendremos

$$\Delta = a_1^1 \begin{vmatrix} 1 & \frac{1^1}{a_2} & \dots & \frac{1^1}{a_n} \\ 0 & \frac{1^2}{a_2} & \dots & \frac{1^2}{a_n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & \frac{1^n}{a_2} & \dots & \frac{1^n}{a_n} \end{vmatrix}$$

en donde en todas las filas  $i \neq 1$  es como si hubiera actuado la operación entre filas  $(i) \rightarrow (i) - a_1^i(1)$ . Ahora, si expandemos el determinante anterior por la primera columna, obtenemos

$$27) \quad \Delta = a_1^1 \begin{vmatrix} \frac{1^2}{a_2} & \frac{1^2}{a_3} & \dots & \frac{1^2}{a_n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \frac{1^n}{a_2} & \frac{1^n}{a_3} & \dots & \frac{1^n}{a_n} \end{vmatrix} .$$

Si repetimos el mismo proceso sobre el determinante de orden  $n-1$  que quedó en la ecuación 27) tendremos

$$\Delta = a_1^1 a_2^{1^2} \begin{vmatrix} \frac{2^3}{a_3} & \frac{2^3}{a_4} & \dots & \frac{2^3}{a_n} \\ \frac{2^4}{a_3} & \frac{2^4}{a_4} & \dots & \frac{2^4}{a_n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \frac{2^n}{a_3} & \frac{2^n}{a_4} & \dots & \frac{2^n}{a_n} \end{vmatrix}$$

Si continuamos de esta forma sobre los nuevos determinantes cada vez de un orden menos, llegamos finalmente a que

$$28) \quad \Delta = a_1^1 a_2^{1^2} a_3^{2^3} \dots a_n^{n-1^n} .$$

El determinante original  $\Delta$  es el producto de estos  $n$  factores; cada uno de los cuales es un paso que se alcanza en el proceso de eliminación de Gauss. De ahí, es claro el parentesco operatorio en la eliminación de Gauss y el cálculo de un determinante.

Recordemos la propiedad 1) de determinantes, capítulo I sección 1.4: un intercambio de filas cambia el signo del determinante. Por ejemplo, si deseamos calcular el determinante de la matriz asociada 3 por 3, al sistema lineal dado en el ejemplo de la sección anterior tendremos

$$\Delta = \begin{vmatrix} 2 & -3 & 4 \\ 1 & 1 & -1 \\ 3 & -1 & -2 \end{vmatrix} = (-3)(7/3)(19/7) = -19 ;$$

en donde  $a_1^1 = 3$ ,  $a_2^2 = -7/3$ ,  $a_3^3 = 19/7$ .  $a_1^1$  y  $a_2^2$  cambian signo por el intercambio de filas que ocurre para buscar el elemento de mayor valor absoluto en la columna donde hay elemento distinguido. El mismo determinante, calculado mediante una expansión por la primera columna conduce a

$$2 \begin{vmatrix} 1 & -1 \\ -1 & -2 \end{vmatrix} - \begin{vmatrix} -3 & 4 \\ -1 & -2 \end{vmatrix} + 3 \begin{vmatrix} -3 & 4 \\ 1 & -1 \end{vmatrix} = -19 .$$

Para concluir con el tema de determinantes calcularemos la cantidad de productos y divisiones que se emplean para resolver un determinante de orden  $n$  por el método expuesto anteriormente. Dicha cantidad de operaciones viene dada así

$$s = [(1 \times 2) + (2 \times 3) + (3 \times 4) + \dots + (n-1)n] + (n-1).$$

El término  $(1 \times 2) + (2 \times 3) + (3 \times 4) + \dots + (n-1)n$  corresponde sólo a la reducción del determinante, siguiendo las pautas de la eliminación de Gauss; falta hacer el producto de los  $n$  términos dados en 26), los cuales son los  $n-1$  productos del último sumando de  $s$ . De la ecuación

$(1 \times 2) + (2 \times 3) + \dots + (n-1)n = 1/3 n^3 - 1/3 n$ , se concluye que la cantidad total de productos y divisiones para calcular un determinante es

$$29) \quad O_D = s = 1/3 n^3 - 1/3 n + (n-1) = 1/3(n^3 + 2n - 3).$$

Si comparamos este método de evaluar un determinante con el método que se puede seguir por la ecuación 2), capítulo I, el cual requiere  $n!$  productos, tendremos: para un determinante 20 por 20, el método, siguiendo los lineamientos de Gauss requiere 2679 productos y divisiones; mientras que siguiendo el procedimiento del capítulo I se requieren  $2.4329 \times 10^{18}$  productos. La diferencia, en la cantidad de operaciones, es obvia.

b) Cálculo de la inversa de una matriz. Dada una matriz

$A = [a_{ij}^j]$ ,  $n$  por  $n$ , para la cual existe inversa; vamos a encontrar dicha inversa  $A^{-1}$ . Llamaremos  $\bar{a}_{ij}^{lk}$  los elementos de la matriz  $A^{-1}$ . A causa de que  $AA^{-1} = I$  entonces el elemento  $j, i$ -ésimo del producto  $AA^{-1}$  viene dado por la ecuación

$$30) \quad \sum_{k=1}^n a_k^j \bar{a}_i^{lk} = \delta_i^j .$$

Dicha ecuación genera un sistema lineal, de orden  $n$ , si va-

riamos  $j=1,2,\dots,n$ . En dicho sistema la matriz asociada es  $A$  y las incógnitas  $\bar{a}_i^{11}, \bar{a}_i^{12}, \dots, \bar{a}_i^{1n}$ . La ecuación 30), junto a los razonamientos anteriores, prueba que para encontrar  $A^{-1}$  hay que resolver  $n$  sistemas lineales de la forma

$$\sum_{k=1}^n a_k^j \bar{a}_i^{1k} = \delta_i^j, \quad j=1,2,\dots,n.$$

Cada sistema se obtiene variando la  $i$  de 1 hasta  $n$ . Si  $i$  es uno obtenemos la primera columna de  $A^{-1}$ , para  $i=2$  la segunda, y así sucesivamente.

Lo anterior prueba que el problema de encontrar la inversa de una matriz se reduce a resolver  $n$  sistemas lineales de orden  $n$ . De acuerdo a lo anterior, el número de productos y divisiones para resolver la inversa de una matriz resulta ser de

$$31) \quad O_I = n(1/3 n^3 + n^2 - 1/3 n).$$

La ecuación 31) está basada en la ecuación 25), es decir se resuelven los  $n$  sistemas lineales bajo el método que llamamos de eliminación de Gauss.

#### IV. CONCLUSIONES

Las siguientes conclusiones son los aspectos principales que se han desarrollado en este trabajo: pretenden resumir lo más importante, no todo el trabajo. Recuérdese que muchos aspectos del álgebra lineal, dados aquí, pueden orientarse a una gran variedad de direcciones, no precisamente sólo hacia sistemas de ecuaciones lineales.

A. Los sistemas de ecuaciones lineales son "Matemáticamente iguales" a las transformaciones lineales, las cuales están trabajadas ampliamente dentro de la teoría de los espacios vectoriales de dimensión finita. La definición de transformación lineal, mediante sistemas de ecuaciones lineales, es muy importante para el enfoque numérico.

B. Las transformaciones lineales no-singulares (operadores) son un caso particular de transformaciones lineales; tienen bastante importancia por su frecuencia de aparición y aspectos teóricos. Los operadores se reflejan en sistemas lineales en el caso de sistemas de ecuaciones de orden cuadrado.

C. El problema general de resolver un sistema de ecuaciones se puede reducir, tanto en el procedimiento de resolución como para interpretación, al caso de sistemas de orden cuadrado. Sin embargo, introducir el caso general no repercute, en complejidad considerable, respecto de un sistema de orden cuadrado; y sí es, por otro lado, bas-

tante más instructivo el caso general que el particular de orden cuadrado.

D. Las soluciones teóricas, dentro del álgebra lineal, para los sistemas de ecuaciones están basados sobre teoremas de existencia. De hecho, se reducen a la existencia de coordenadas sobre una base o, en forma equivalente, al rango de un sistema lineal.

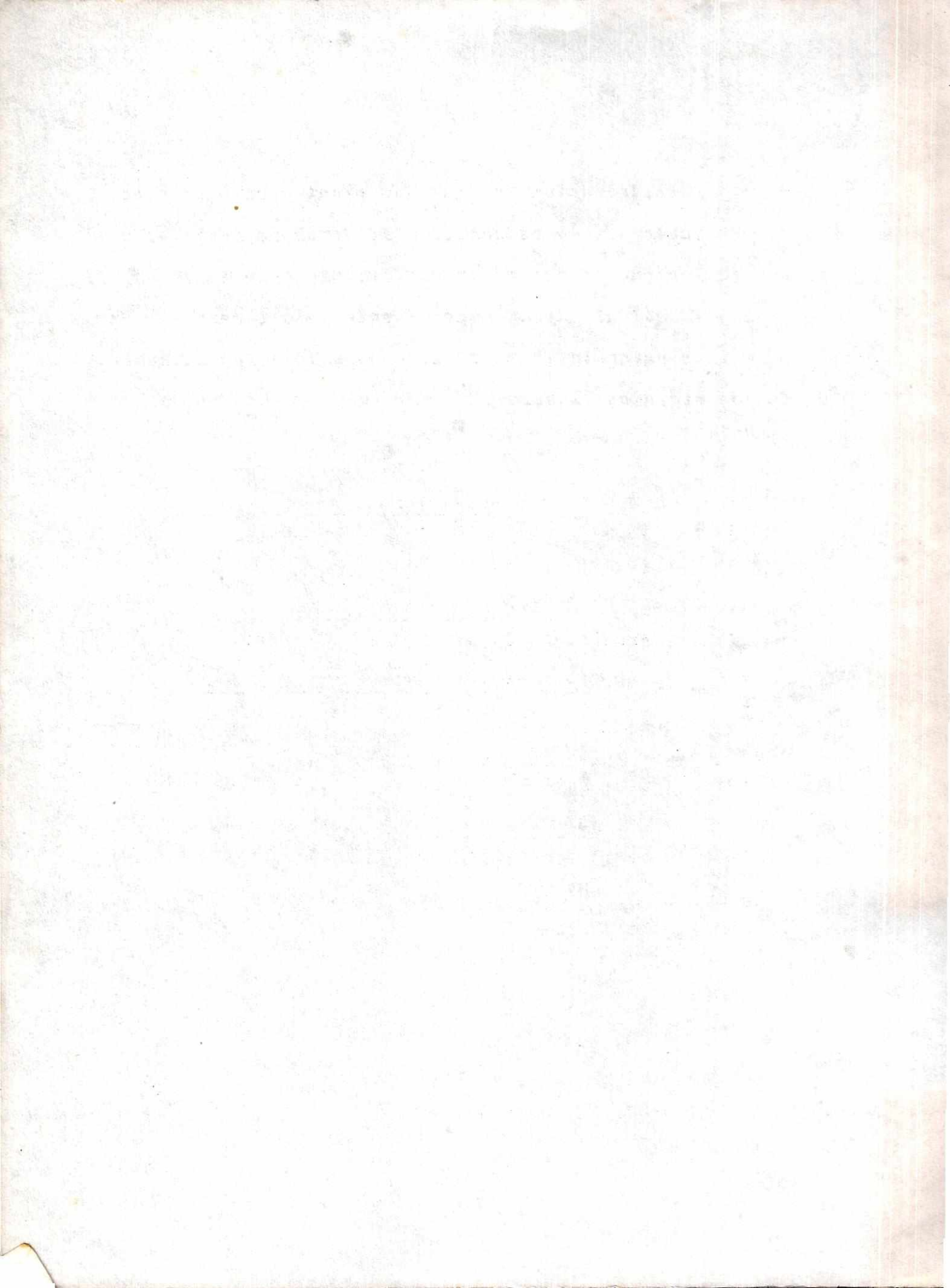
E. El problema numérico de los sistemas lineales queda resuelto en forma completa, como una buena alternativa, por medio de la reducción a la forma normal de Hermite. La forma normal de Hermite es un procedimiento típico que caracteriza los llamados métodos exactos.

F. Los métodos exactos resuelven los sistemas lineales pero requieren amplia fundamentación teórica del álgebra lineal. Esta fundamentación, no sólo es necesaria para el rigor lógico, sino también ofrece un puente para reducir innumerables problemas dentro de los sistemas lineales.

G. Existen problemas numéricos de los sistemas lineales que atañen al tamaño del sistema lineal, los cuales no se desarrollan con detalle en este trabajo, pero sí se mencionan: acumulación de errores de redondeo, cantidad de operaciones, etc. Otros problemas no necesariamente están relacionados con el tamaño del sistema, por ejemplo, estabilidad de un sistema lineal.

H. La eliminación de Gauss se plantea como un caso particular de la reducción a la forma de Hermite; caso particular que ofrece mejor rendimiento numérico.

I. El cálculo de un determinante y la inversa de una matriz están íntimamente ligados a la solución numérica de los sistemas lineales.



## BIBLIOGRAFIA

- American Mathematical Society. Mathematical of the  
1968 decision sciences. 2a. ed. Rhode Island.  
2 v.
- Faddeeva, V. N. Computational method of linear alge-  
1959 bra. New York, Dover publications, Inc.  
252 pp.
- Halmos, P. R. Finite-dimensional vector spaces. New  
1974 York, Heidelberg, Berlin; Springer-Verlag.  
199 pp.
- Lichnerowicz, A. Linear algebra and analysis. San  
1967 Francisco, California, Holden-day. 304 pp.
- M.I.T. Press. Mathematics, its content, method,  
1981 meaning. 3rd. ed. Cambridge, Mass. 3 v.
- Noble, B. Applied linear algebra. New Jersey, Pren-  
1969 tice Hall, Inc. 523 pp.
- Scheid, F. Análisis numérico. Boston, Mc Graw-Hill,  
1972 serie Schaum. 422 pp.
- Shilov, G. E. Linear Algebra. New Jersey, Prentice  
1971 Hall, Inc. 387 pp.

