TESIS

Fornes Peral
Pedro Alonso

UNA PROPUESTA PARA EL CURSO de Algebra lineal I

I N D I C E

C A P I T U L O 0

I N T R O D U C C I O N

Nuestro propósito fundamental al escribir esta tesis es presentar una propuesta acerca del material que debe cubrirse en un curso de Algebra Lineal I, el cual es común a todas las carreras de Ciencias e Ingeniería de la Universidad de Sonora.

Creemos que uno de los problemas mas graves dentro del Departamento de Matemáticas es la falta de programas actualizados para cada una de las materias que se ofrecen en las diversas áreas, problema que se traduce en la no uniformidad de criterios a la hora de impartir un mismo curso. Como consecuencia inmediata de ésto, pudiéramos citar el hecho de que al presentarse los alumnos a cursar una materia seriada, existe material que un profesor sí cubrió y que otros no lo hicieron. Así se añaden dificultades para el maestro de la nueva materia.

Además de lo anteriormente expuesto, hemos de mencionar el hecho de que una gran mayoría de nuestros alumnos no utiliza los libros de texto que gran ayuda prestan para llevar a buen término el aprendizaje de un curso. Algunas veces porque no existen en el mercado local, otras por no tener medios para adquirirlos, otras porque las bibliotecas no los tienen o hay pocos en existencia, y otras más porque no hay uno adecuado, pero al fin no se usan.

Soluciones a las deficiencias mencionadas pensamos que existen muchas, y consideramos que las personas más adecuadas para buscarlas y ponerlas en práctica son aquéllas que tienen una cierta experiencia en el área.

Animadas por esta idea y aprovechando las observaciones que a lo largo de algunos semestres de impartir el curso de Algebra Lineal I hemos hecho, decidimos implementar como material de nuestra tesis estas notas, mismas que deseamos que con la colaboración de las instancias adecuadas puedan convertirse en un futuro próximo en un libro de texto que pudiéramos usar alumnos y maestros de esta Universidad de Sonora.

Para hacer una propuesta de modificación curricular pueden aplicarse varios criterios:

a) Un primer camino sería la realización de una encuesta entre los distintos jefes de carrera del área de C.I. de la institución, para informarse cuáles son los conceptos de Algebra Lineal I que sus respectivos estudiantes necesitan manejar.
b) Un segundo camino a seguir sería atender el material que necesariamente debe dominar el alumno al cursar el resto de las materias del tronco común (Cálculo Diferencial e Integral III, Ecuaciones Diferenciales I, Electromagne---

tismo, Análisis Numérico I, Probabilidad y Estadística).
c) Otro más sería atender las ideas obtenidas mediante la -- experiencia propia en el aula, tanto como alumnas y como profesoras de la materia.

Los únicos criterios que aplicamos fueron los señalados en los incisos b) y c).

También consideramos, aunque hasta el momento no se ha - mencionado, que todo el contenido de la propuesta pueda ser - manejado sin problemas con los conocimientos que el alumno -- posee, adquiridos en los cursos anteriores al que nos ocupa, (Algebra Superior I, Mecánica, Cálculo Diferencial e Integral I y Geometría Analítica).

Entre los objetivos generales de nuestro trabajo pudié-- ramos citar:

a) Exposición clara y sencilla de todos los temas tratados.
b) Lograr un buen manejo de teoría para cursos posteriores.
c) Ilustrar la teoría con aplicaciones que no sean trivia--- les.

Inicialmente también estaba contemplado analizar el as-- pecto computacional de los métodos desarrollados. Sin embargo debimos abandonar esta idea porque realmente el material a -- tratar era bastante extenso. Queda entonces sin tocar un as-- pecto que es verdaderamente interesante.

Los temas que se incluyeron, en su orden correspondiente son :

CAPITULO I.- Vectores en R^n
CAPITULO II.- Sistemas de Ecuaciones Lineales (SEL)
CAPITULO III.- Matrices
CAPITULO IV.- Espacios Vectoriales
CAPITULO V.- Determinantes
CAPITULO VI.-Transformaciones Lineales
Apéndice.

En términos generales, a cada capítulo se trató de darle la estructura siguiente:

a) Una pequeña introducción en la que se mencionaran los as- pectos más relevantes que se tratan ahí.
b) Un bosquejo histórico, citando a los personajes más im--- portantes que intervinieron en el desarrollo de tal o --- cual tema.
c) Una sección de problemas aplicados en distintas ramas: E- conomía, Ciencias Sociales, Química, Física,(sólo la pre- sentación).
d) El mínimo de teoría que se creyó conveniente.
e) Se buscó intercalar las soluciones de los problemas con--

Se asegura que todos los problemas que se incluyen se -- pueden resolver con la teoría vista.

La sección a) se incluye pretendiendo dar un panorama -- general del material que se trata a lo largo del capítulo, -- buscando situar al alumnado.

Las secciones b) y c) se introdujeron con el afán de motivar a los estudiantes, sobre todo a los de las distintas -- ingenierías, que sienten poca atracción hacia la cuestión --- teórica, cuando ésta no vá acompañada de una buena dosis de - aplicaciones.

En lo correspondiente a la sección d), toda la exposi--- ción busca ir introduciendo al alumno de manera natural en -- conceptos que van aumentando en grado de dificultad y donde - es posible se les motiva mediante ejemplos muy sencillos.

¿Qué relación guarda esta propuesta con respecto al programa actual que se maneja en la materia?

De entrada, podemos decir que no hay una diferencia fundamental en lo referente al contenido del material que se cubre. Básicamente se trata del mismo, aunque difiere sustan--- cialmente en el orden en el que se presentan los temas.

¿Por qué decidimos seguir este orden? Trataremos de ex-- plicarlo lo más ampliamente posible.

Elegimos empezar con Vectores en R^n, para aprovechar que se mantienen relativamente frescos en la mente de los alumnos conceptos que se manejaron en Geometria Analítica y en Mecá-- nica, ejemplo: la noción de vector en los espacios de dos y - de tres dimensiones y aspectos relacionados, normas, ángulo, distancia, etc. De hecho, cada vez que ésto es posible, se -- les menciona que tal o cual cosa ya la conocen, remitiéndolos al curso donde lo vieron.

No perdamos de vista, sin embargo, que todas estas no--- ciones están generalizadas al espacio de n dimensiones porque es muy importante desde el principio tratar de mostrar que no tenemos por qué limitarnos a casos particulares, aquí serían los espacios de dos y tres dimensiones, cuando "se siente", que sin mayor problema puede pasarse al de n dimensiones y -- que ello no implica el perderse en abstracciones innecesarias.

Otra argumentación a favor de la introducción de Vecto-- res en R^n como Capítulo I, sería decir que la estructura de - un vector es mucho más sencilla que la de una matriz y por -- ende, mas fácilmente asimilable.

En la presentación de SEL como Capítulo II, lo importante sería el mostrar distintos métodos de resolverlos, que es

la tesis en sí. Colocamos primero el material de SEL antes -- que otros temas porque consideramos que el hacerlo así, con-- tribuye a despertar el interés de los alumnos de las carreras de Ciencias e Ingeniería, al ver un material al cual le pue-- den encontrar aplicación inmediata tanto en el resto de las - materias que cursan como en problemas reales específicos de - sus carreras.

Otro detalle que deseamos aclarar es el siguiente:

¿Por qué creemos que se necesita toda una teoría para -- resolver sistemas de ecuaciones lineales? Son varias las res- puestas que podemos dar :

a) Aparecen problemas reales, planteados en términos de SEL, y que requieren solución.
b) Hasta el momento de iniciarse el curso de Algebra Lineal I el alumno sólo sabe resolver sistemas cuadrados. La --- gran mayoría de los problemas reales no son cuadrados.
c) Los sistemas cuadrados que se saben resolver son de orden a lo más 3 y existen problemas reales que plantean siste- temas de orden mayor que 3.
d) Existen problemas reales que involucran sistemas con va-- rias soluciones (y son la gran mayoría de ellos) . Se - debe tener un método para encontrarlas.
e) En sistemas de gran tamaño se requiere de procedimientos que sean confiables para encontrar su solución.
f) En el desarrollo de conceptos teóricos como conjuntos de vectores linealmente independientes (L.I.) y linealmente dependiente (L.D.)y escribir un vector como combinación - lineal de otros, por cuestiones prácticas llegamos a re-- solver un SEL (estos son conceptos correspondientes al - capítulo de Espacios Vectoriales).

Los SEL nos brindan, además de un amplio campo de apli-- caciones en todas las áreas, una manera que nosotros sentimos muy natural de definir a las matrices. Esto es, mostramos cómo las matrices aparecen como una forma de agilizar la reso-- lución de un SEL. Se justifica entonces su aparición y el po- der retomarlas ya como objetos matemáticos con los cuales po- demos operar y que tienen importancia y utilización por sí -- mismas, independientemente de su relación con los SEL. Todo - ello conforma y justifica su aparición como Capítulo III.

Ya que se ha trabajado suficientemente con vectores y -- con matrices, resulta que estos dos son ejemplos clásicos de lo que es un Espacio Vectorial, tema que introducimos como -- Capítulo IV. El trabajar con espacios vectoriales siempre nos remite (se apuntaba líneas arriba) a resolver un SEL.

Recuérdese la técnica que se sigue para demostrar cuándo un conjunto de vectores es LI o LD.

sustento teórico al Capítulo II. Recordemos también que entre las personas que cursan la materia están los futuros Licen--- ciados en Matemáticas y Licenciados en Física, en cuya formación este tema tiene especial importancia, dado que en su currícula aparece Algebra Lineal II, Cálculo Diferencial e Integral III, Cálculo Diferencial e Integral IV, etc., que son materias que necesitan de este tópico.

Resaltaremos que en el tema de Espacios Vectoriales se ha incluído algo que es totalmente novedoso, que hasta la fecha creemos que nunca se ha incluído: ¿Qué hacer cuando tenemos un SEL sin solución? . ¿Cuál es la mejor aproximación a - lo que pudiera llamarse " solución "?.

Pasamos después a conformar el Capítulo V mediante el - estudio de los Determinantes. Aprovechando la noción de vec-- tor que ya se maneja, hacemos una presentación que consideramos media entre la formal que involucra permutaciones y la un tanto trivial, que sólo los define como números asociados a matrices. Se muestra además otro método de resolver SEL, que si bien operacionalmente es inferior a los que ya se mostraron en el Capítulo II, histórica y teóricamente es importante: la Regla de Cramer.

Se incluye otro método para el cálculo de la inversa de una matriz, distinto al que se presenta en el Capítulo III.

Finalmente aparece como Capítulo VI el de Transformaciones Lineales. Consta sólo de una breve exposición y decidimos no extenderlo, básicamente por cuestiones de tiempo, pensando en que el material que se tiene es más que suficiente para -- cubrir un curso de un semestre.

Se queda en el entendido de que puede retomarse posteriormente en Algebra Lineal II.

En el apéndice se incluyeron algunas demostraciones y -- detalles sin los cuales el curso puede llegar a buen término, pero que tal vez alguna persona interesada en profundizar --- quisiera ampliar.

No se deja de reconocer que éste es un trabajo incompleto que puede perfeccionarse bastante, y mas que considerar -- ésto nocivo, lo pensamos favorable, ya que deja la puerta --- abierta a que los compañeros maestros participen en su enri-- quecimiento.

C A P I T U L O I

VECTORES EN R^n

1.1 I N T R O D U C C I O N.

Ya en cursos anteriores, como en Mecánica y Geometría -- Analítica, se ha tenido contacto con estos entes matemáticos llamados vectores, aunque sólo ha sido posible hacerlo en los espacios de dos y tres dimensiones. Sabemos que hablando en el mundo físico, manejamos un vector como un objeto que tiene magnitud, dirección y sentido. En Geometría se maneja simplemente con la noción de punto, ya sea en el plano o en el - espacio. En este capítulo aprenderemos que el concepto de --- vector es muchísimo más amplio, es decir, que se puede gene-- ralizar y hablar sin mucho problema de vectores de n compo--- nentes (donde n es un natural arbitrario), y que el hacer - ésto no significa divagar en abstracciones inútiles, dado que existen muchos problemas, tanto prácticos como teóricos, que pueden resolverse echando mano de ellos.

Veremos también que los vectores pueden considerarse como objetos matemáticos con los cuales es posible operar (ha-- cer sumas, multiplicaciones, etc.) y que todos estos detalles no serán más que la generalización de lo que se mencionaba - líneas arriba ya se vió en otros cursos.

Analizar como Capítulo I Vectores en R^n, nos será de --- gran utilidad posteriormente, ya que capítulos adelante estudiaremos estructuras que son un poco más complicadas, pero -- que al fin de cuentas tienen gran similitud con los vectores: las matrices.

Afortunadamente, existe una amplísima gama de áreas en - donde encontramos inmiscuídos a los vectores, lo cual nos ha permitido anexar una serie de problemas de aplicación que es bastante diversa.

1.2 R E S E Ñ A H I S T O R I C A

La idea de un paralelogramo de entes físicos resultó de gran influencia dentro de la historia antigua de los concep-- tos vectoriales.

Aunque este concepto no necesariamente involucra la idea de lo que es un vector, proporcionó a los antiguos vectoris-- tas un área en la que se ilustró de modo convincente la uti-- lidad de los métodos vectoriales elementales.

Ciertas cantidades físicas, tales como velocidades y --- fuerzas, pueden expresarse de tal manera que representen un - paralelogramo. La suma de ellas estará determinada por la --

Pertenece a Gottfried Wilhelm Leibniz (1646-1716) el crédito de haber visualizado la naturaleza de un sistema para el análisis del espacio de tres dimensiones. Hizo además muchas otras contribuciones matemáticas, entre las cuales está su concepto de una geometría de situación. Aquí Leibniz discutió la posibilidad de crear un sistema que sirviera como método directo del análisis del espacio; decía:

> "Es necesario un tipo de análisis que sea indistintamente geométrico o lineal y que exprese directamente situación (situs) como el Algebra expresa directamente magnitud, de modo que podamos representar figuras y aún máquinas y movimientos por caracteres, como el Algebra representa números o magnitudes. He descubierto ciertos elementos de una nueva característica que es enteramente diferente del Algebra y que tendrán grandes ventajas en su representación para la imaginación, exactamente y de una manera fiel a su naturaleza, aún sin figuras, todo lo cual depende del sentido de percepción. Esta nueva característica, que sigue de las figuras visuales, no puede fallar para dar la solución, construcción y demostración geométrica, todas a la vez, y en una forma natural y en un análisis, es decir, mediante un procedimiento determinado".

Aunque el término Análisis Vectorial se usa ahora principalmente en referencia a sistemas matemáticos que pueden aplicarse en el espacio tridimensional, no debe olvidarse el sistema de los números complejos, el cual puede considerarse legítimamente como un sistema vectorial. El sistema vectorial bidimensional ciertamente no es tan útil como los sistemas vectoriales tridimensionales, pero en él se incluye el sistema de los números complejos. Cabe señalar que se descubrieron los cuaternios durante la búsqueda de un espacio tridimensional análogo al sistema de los números complejos.

El deseo de encontrar una manera de representar ciertos objetos en los espacios de dos y tres dimensiones llevaron al matemático irlandés William Rowan Hamilton (1805-1865) a la idea de los cuaternios, que desembocaron posteriormente en lo que hoy conocemos como vectores. En aquel entonces se discutió mucho sobre la utilidad de los vectores, llegando incluso el físico inglés Kelvin (1824 -1907) a opinar que eran algo totalmente inútil. Obviamente estaba equivocado, ya que gran parte de la Física Clásica y Moderna son representadas por medio de vectores. También se usan con bastante frecuencia en Ciencias Biológicas y Sociales.

1.3 P R O B L E M A S D I V E R S O S

A continuación enunciaremos algunos problemas que involucran vectores dentro de diversas áreas y cuya solución se

Problema 1.- Las dimensiones de un cuarto son 10 mts. X 12 - mts. X 14 mts. Un insecto que sale de una esquina llega a la esquina diagonalmente opuesta, siguiendo la trayectoria de -- esa diagonal.

a) ¿Cuál es la magnitud de su desplazamiento ?
b) ¿Podría ser la longitud de su trayectoria menor, mayor o igual que esta distancia ?
c) Escoja un sistema de coordenadas adecuado y determine en sus componentes el vector desplazamiento.

Problema 2.- Un avión tiene una velocidad de vuelo máxima de 500 km./h. Si vuela a su velocidad máxima con una dirección de 30 grados al Oeste del Norte y el viento está soplando de Norte a Sur a una velocidad de 40 km./h., determine :

a) La velocidad con respecto a Tierra.
b) La dirección del viaje sobre la Tierra.
c) Si una ciudad A está situada a 1500 km. en dirección 30 - grados Oeste con relación al Norte de la ciudad B y el avión se demora 3 horas para volar de B a A bajo el supuesto de que la velocidad del viento es constante, ¿cuál es el mínimo ---- tiempo posible para el viaje de regreso ?

Problema 3.- Movimiento de Proyectiles.- Una partícula se mueve siguiendo una trayectoria curva en el plano XY. Se -- conoce su posición R, su velocidad V y su aceleración A en el tiempo t. Consideremos el caso especial del movimiento en - un plano con aceleración constante. En general, la partícula recorrerá una trayectoria curva en el plano.

Como ya es sabido, las ecuaciones de movimiento son

$$V = Vo + A\,t\ ,$$

donde Vo = velocidad inicial.

Análogamente, podemos determinar la posición R de la --- partícula mediante

$$R = Ro + Vo\,t + 1/2\ \ A\,t^2\ ,$$

con Ro = posición inicial.

Un ejemplo de este tipo de movimiento, es el movimiento de proyectiles. Se trata del movimiento en el plano de una - partícula disparada oblicuamente en el aire. Suponemos des-- preciable el posible efecto del aire en el movimiento.

El movimiento de un proyectil es de aceleración constante g dirigida hacia abajo.

Este ejemplo muestra claramente y en forma muy sencilla

instante pueden considerarse en términos de operaciones con - vectores (sumas vectoriales).

Después de mostrar ésto, lo podemos aplicar así :

Un bombardero lleva una velocidad horizontal constante - de 1500 km./h. y vuela a una altura de 16000 metros hacia un punto situado directamente sobre su blanco. ¿Que ángulo θ debe formar la visual con respecto al blanco en el momento de - dejar caer la bomba para que oé precisamente en el blanco ?

Problema 4.- Vector Receta y Vector Precio.- Para la preparación de un exquisito pay helado se requieren los siguientes ingredientes : 1/4 de litro de Leche del Clavel, 1/4 de litro de Leche Nestlé, 7 limones, 200 gramos de Galletas Marías, 50 gramos de mantequilla derretida, 30 gramos de azúcar y 1/2 - kg. de piña picada. La receta dá para 6 porciones.

a) Escriba el vector receta A que proporcione el número de - unidades requeridas de cada ingrediente para una porción.
b) Escriba el vector de precios totales P de los ingredien-- tes de la receta si los precios son los siguientes :

1 litro de Leche del Clavel	$ 182.00
1 litro de Leche Nestlé	$ 460.00
Un limón	$ 7.00
1 kg. de Galletas Marías	$ 400.00
1 kg. de mantequilla	$ 900.00
1 kg. de azúcar	$ 182.00
1 kg. de piña	$ 140.00

c) Encuentre el producto punto de A y P y explique su significado.

Problema 5.- Trabajo.- Consideremos una partícula sobre la cual obra una fuerza constante F que provoca un desplazamiento a lo largo de una línea recta en una dirección determina-- da. En tal caso, definimos el trabajo hecho por la fuerza F sobre la partícula como el producto de la magnitud de la ---- fuerza F por la distancia d que se mueve la partícula.

Si la fuerza no se aplica en la misma dirección en que - se mueve la partícula, el trabajo realizado se define como el producto de la componente de la fuerza F en la dirección del movimiento por la distancia d que se mueve el cuerpo a lo --- largo de esa línea. Así, podemos expresar ésto como

$$W = F \cdot d \ ,$$

donde W representa dicho trabajo.

El trabajo, tal como lo hemos definido, resulta ser un - concepto muy útil en la Física. Ejemplo :

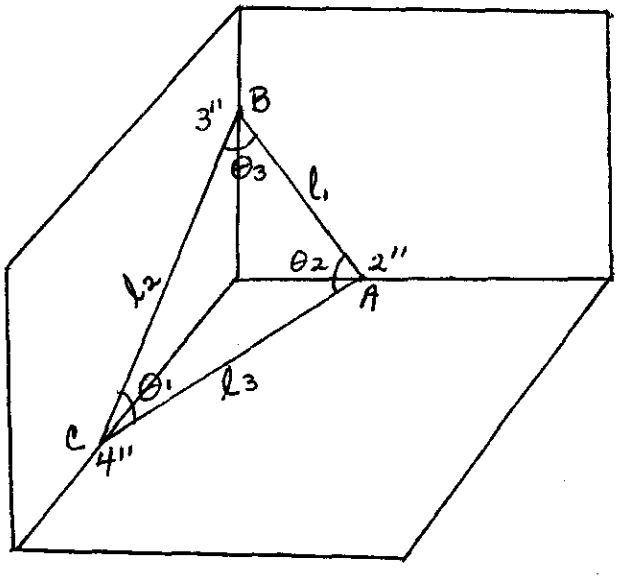
Una fuerza de 2 Nt. actúa en la dirección del vector

realizado al mover un objeto del punto (1,2,3) al punto (2,8, 11), donde las distancias se miden en metros.

Problema 6.- Encuentre el cuarto vértice de un tetraedro regular que tiene tres de sus vértices en los puntos A (4,0,0), B(-4,0,0), C(0,4 5,0).

Problema 7.- Muestre que las diagonales de un rombo son perpendiculares.

Problema 8.- Un hombre desea recortar una tabla triangular para que se acomode en una esquina, tal como se muestra en la figura. Determine los ángulos de las esquinas y las longitudes de las aristas en la cara exterior de la tabla.



1.4 V E C T O R E S : Definición y ejemplos.

Podríamos tratar de dar algunas definiciones de vector, por ejemplo : en el plano (R^2), desde el punto de vista físico, lo tenemos que considerar como un objeto que tiene magnitud y dirección ; desde el punto de vista geométrico un vector en R^2 es el conjunto de todos los segmentos de recta equivalentes a un segmento de recta dado. (Dos segmentos de recta dirigidos son equivalentes si tienen la misma magnitud y dirección).

La definición algebraica sería que un vector en el plano XY es un par ordenado de números reales (a,b).

tor. Esta última definición, que es la que manejaremos, la - podemos generalizar al espacio de n dimensiones de la si----- guiente manera :

Definición 1.4.1.- Un vector V, de n componentes rea--- les, es un arreglo ordenado de n números reales. Llamaremos R^n al conjunto de todos los vectores de n componentes reales.

Ejemplos :

Cuando n=2 : (1,2), (6,10000), (85,4), etc.

Cuando n=3 : $(1/2, 2/4.\sqrt{2})$, $(\pi, 6, 0)$, etc.

Cuando n=4 : $(1, 1, 0, -3)$, $(6, \sqrt{3}, -\sqrt{2}, 6)$, etc.

Notación.- Acostumbraremos denotar a los vectores con - letras mayúsculas y a sus componentes con minúsculas.

Ejemplo : $A=(a_1, a_2, \ldots, a_n)$; $V=(v_1, v_2, \ldots, v_n)$.

El hecho de que los arreglos sean ordenados permite com- parar dos vectores A y B, siendo iguales si y sólo si tienen las mismas componentes en el mismo orden, ésto es :

Definición 1.4.2.- Sean A y B dos vectores de n compo-- nentes reales tales que $A=(a_1, a_2, \ldots, a_n)$, $B=(b_1, b_2, \ldots, b_n)$, - entonces A=B si y sólo si $a_1=b_1$, $a_2=b_2$, ..., $a_n=b_n$.

Observación.- A pesar de que ya hemos dado la defini--- ción de vector que consideramos conveniente para los propó-- sitos de esta sección, es necesario remarcar que hay muchas - maneras de abordar este concepto.

Se habla de vectores coordenados, vectores geométricos,- etc., pero sin embargo éstos no son más que ejemplos de la -- amplísima gama de situaciones en las que la idea abstracta de vector se puede utilizar. Con ésto lo que queremos dar a en- tender es que hay una generalización que vá mucho más allá de todo lo que se ha manejado hasta ahora y que esperamos acla-- rar al momento de estudiar el tema de Espacios Vectoriales. - Veremos que el hacer ésto, nos proveerá de herramientas úti-- les, pues cualquier proposición que sea válida para espacios vectoriales lo será en particular para R^n.

1.5.- O P E R A C I O N E S E N T R E V E C T O R E S : Propiedades y ejemplos.

DEFINICION 1.5.1.- Sean V, W $\in R^n$ tales que
$V=(v_1, v_2, \ldots, v_n)$, $W=(w_1, w_2, \ldots, w_n)$.
Definimos la suma de dos vectores, denotada V + W, como el -- vector de componentes $(v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$, ésto es
$V + W = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$.

simplemente el producto por un escalar, denotado λV, como el vector

$$\lambda V = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n),$$

donde $\lambda \in R$.

Ejemplos :

Sean $V = (3,6,-2)$, $W = (5,\sqrt{2},-1)$. Calcule :

$V + W = (3 + 5, 6 + \sqrt{2}, -2-1) = (8, 6 + \sqrt{2}, -3)$.

$5V = (5 \times 3, 5 \times 6, 5 \times (-2)) = (15, 30, -10)$.

$3W = (3 \times 5, 3 \times \sqrt{2}, 3 \times (-1)) = (15, 3\sqrt{2}, -3)$.

$5V + 3W = (30, 30 + 3\sqrt{2}, -13)$.

Observación.- $V-W = V + (-W) = V + (-1)W$.

Propiedades de la suma de vectores.-

Sean V, W y $Z \in R^n$, $\alpha, \lambda \in R$.

1) $V + W$ es otro vector en R^n (cerradura)

2) $V + W = W + V$ (conmutatividad).

3) $(V + W) + Z = V + (W + Z)$ (asociatividad).

4) $V + \theta = V$, donde $\theta = (0,0,\ldots,0)$ (elemento neutro).

5) $V + (-V) = \theta$, donde $-V = (-1)V$ (inverso aditivo).

Propiedades de la multiplicación por un escalar.

1) λV es un vector en R^n (cerradura), y si $\lambda = 1$ entonces $1 \cdot V = V$.

2) $\alpha(\lambda V) = (\alpha\lambda)V$ (pseudo-asociatividad)

3) $\lambda(V + W) = \lambda V + \lambda W$ (distributividad).

Nota.- Como se verá algunos capítulos adelante, esta serie de propiedades clasifican a un conjunto como algo más especial llamado Espacio Vectorial.

Demostrar en base a las propiedades de los números reales las propiedades de suma de dos vectores y multiplicación de un escalar por un vector es bastante sencillo utilizando la definición. Pero,

1.- ¿Qué efecto produce el multiplicar un vector de R^2 o R^3 por un escalar λ, siendo :

a) $\lambda > 1$

b) $0 \leq \lambda < 1$

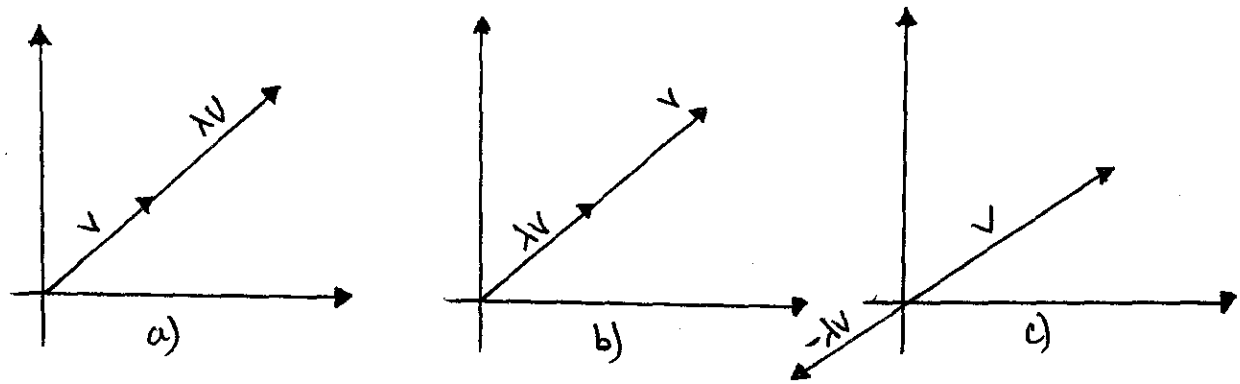
c) $\lambda < 0$?

vectores de dos y tres componentes ?

Veamos : según la definición 1.3.1 tenemos que

$$\lambda V = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n), \quad \lambda \in R, \ V \in R^n .$$

Para a) y b), si $\lambda > 0$ entonces λV será un vector con -- magnitud igual a λ veces la magnitud de V, con su misma di--- rección y sentido. Para c), $\lambda < 0$ provoca cambio de sentido, - altera la magnitud (igual que en a) y b)) y conserva la di--- rección; gráficamente, en R^2:

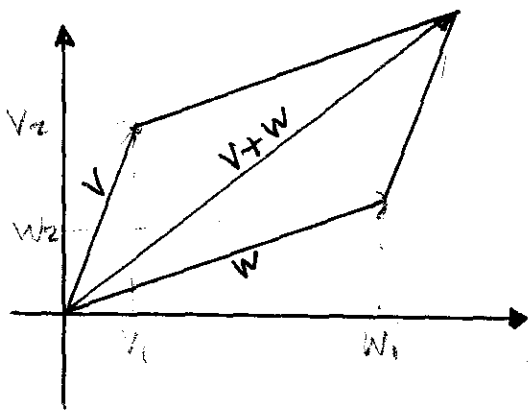


6

En 2, también por la definición 1.3.1, tenemos que si -- $V, W \in R^2$ entonces

$$V + W = (v_1 + w_1, v_2 + w_2).$$

Y ésto no es otra cosa más que la conocidísima Ley del Paralelogramo.



Producto interior, punto o escalar de vectores.

Definición 1.5.2.- Sean $V, W \in R^n$, definimos el producto interior, punto o escalar de V por W, denotado $V \cdot W$, como

$$V \cdot W = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n .$$

Ejemplos en R^3 :

1) $V = (4, 5, 4)$ $W = (-1, 4, 0)$ Entonces

2) $Z = (-1/2, 3, 4)$, $M = (-4, 6, 8)$. Entonces
$Z \cdot M = (-1/2) \times (-4) + 3 \times 6 + 4 \times 8 = 52$.

Observaciones:
- El producto interior sólo es posible entre vectores de igual número de componentes.
- El producto interior de dos vectores es un número real (escalar).

Propiedades:
Sean $V, W, Z, \in R^n$, $\lambda \in R$, entonces:

1) $V \cdot W = W \cdot V$ (conmutatividad).

2) $V \cdot (W + Z) = V \cdot W + V \cdot Z$ (distributividad)

3) $(\lambda V) \cdot W = \lambda (V \cdot W)$.

4) Si $V = 0$ entonces $V \cdot V = 0$.

5) Si $V \neq 0$ entonces $V \cdot V \neq 0$.

OBS $V \cdot V = (V_1, V_2, \ldots, V_n) \cdot (V_1, V_2, \ldots, V_n) = V_1^2 + V_2^2 + \ldots + V_n^2$.

Estas propiedades son bastante fáciles de demostrar mediante la definición.

Antes de continuar debemos advertir que la definición de producto escalar que se acaba de dar nos será de bastante utilidad pues, como veremos páginas más adelante, construiremos a partir de él toda una geometría para los vectores.

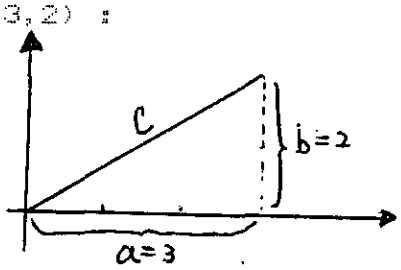
Pasaremos entonces a estudiar otros conceptos relacionados con los vectores.

1.6.- N O R M A O L O N G I T U D D E U N V E C T O R .

Definición y propiedades. Solución de problemas aplicados.

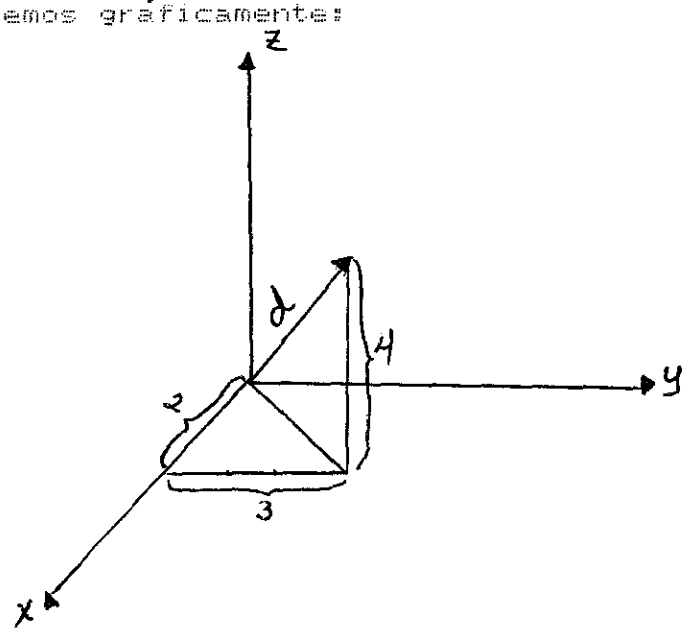
En el caso de R^2 el encontrar la norma o longitud de un vector (a,b) es equivalente a encontrar la longitud de la hipotenusa del triángulo rectángulo cuyos catetos están determinados por los números a y b, problema que se resuelve mediante una sencilla aplicación del Teorema de Pitágoras.

Por ejemplo, si queremos calcular la longitud del vector (3,2) :

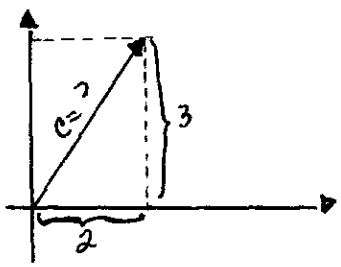


$$c = \sqrt{a^2 + b^2} = \sqrt{13}$$

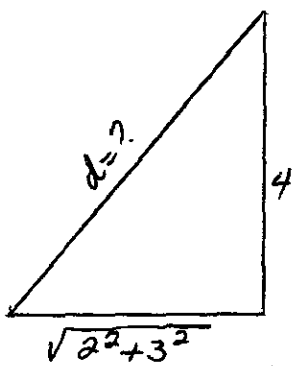
ejemplo, si deseáramos calcular la longitud del vector --- (2,3,4), tenemos gráficamente:



Esto lo podemos separar en dos triángulos rectángulos en $\mathbb{R}^2$ y trabajarlos de manera análoga al caso anterior, con lo -- que obtenemos:



$$c = \sqrt{2^2 + 3^2} = \sqrt{13}$$



$$d = \text{norma} = \sqrt{2^2 + 3^2 + 4^2}$$

Esta idea nos permite generalizar el concepto de norma de la siguiente manera:

Definición 1.6.1. Sea $V \in \mathbb{R}^n$. Definimos la norma o magnitud de V, denotada $\| V \|$, como:

$$|| V || = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2} = \sqrt{V \cdot V}.$$

Ejemplo: Sea $V \in R^5$ dado por $(5,-3,6,-1,0)$. Entonces:

$$|| V || = \sqrt{25 + 9 + 36 + 1 + 0} = \sqrt{71}.$$

Propiedades de la norma.

Sean $V, W \in R^n$:

1) $|| V || \geq 0$.

2) $|| V || = || -V ||$.

3) $|| \lambda V || = | \lambda | \, || V ||$, con $\lambda \in R$

4) $|| V || = 0$ si y sólo si $V = \theta$.

5) $|| V + W || \leq || V || + || W ||$. (Desigualdad del Triángulo para la norma).

Todas estas propiedades son bastante sencillas de demostrar utilizando la definición, excepto la última, en la cual se hace uso de uno de los resultados (llamado Desigualdad de Cauchy-Schwarz) que se contemplan al final de este capítulo.

Para demostrarla trabajaremos en base a los cuadrados de ambos miembros (ésto es válido ya que ambos son elementos no negativos), ésto es:

$$|| V + W ||^2 \leq (|| V || + || W ||)^2$$

Pero:

$$|| V + W ||^2 = (V + W) \cdot (V + W).$$

$$= V \cdot V + V \cdot W + W \cdot V + W \cdot W$$

$$= V \cdot V + 2\, V \cdot W + W \cdot W \quad (\text{pues } V \cdot W = W \cdot V)$$

$$= || V ||^2 + 2\, V \cdot W + || W ||^2$$

$$\leq || V ||^2 + 2 || V || \, || W || + || W ||^2$$

$$= (|| V || + || W ||)^2.$$

$V \cdot W \leq \|V\| \|W\|$

$V \cdot W = \|V\| \|W\| \cos\theta$

$\cos\theta = \dfrac{V \cdot W}{\|V\| \|W\|}$

De donde, $|| V + W || \leq || V || + || W ||$.

Una pregunta que nos podríamos hacer es la siguiente: dado cualquier vector V en R^n, ¿existirán otros vectores que posean su misma dirección, sentido igual u opuesto, pero que tengan una cierta norma arbitraria L ?

Este problema se presenta al momento de querer modelar -

vector dado posee una magnitud menor que L significa que de-- alguna manera necesitamos alargarlo; si por el contrario, posee una magnitud mayor que L tendremos que acortarlo; si se - quiere que sean de sentido opuesto, deberemos invertir el --- sentido original ($k<0$). En todos los casos el problema se -- resuelve multiplicando el vector original por alguna constante k. Necesitamos, pues, encontrar el valor de esta constante.

Sabemos que se debe satisfacer que el vector kV sea tal que $\| kV \| = L$.

Pero por la propiedad 3 de la norma tenemos que

$$\| kV \| = |k| \, \| V \| \Rightarrow L/\| V \| = |K|$$

Ejemplos:

1.- Dado el vector $V = (5,-1,4,2)$, ¿cuál será el vector de -- longitud 1/2 que está en la misma dirección y sentido de V?

Solución.- Necesitamos determinar el valor de k.

$$k = L/\| V \| = (1/2)/\sqrt{46} = 1/2\sqrt{46}.$$

Por lo tanto,

$$kV = 1/2\sqrt{46}\,(5,-1,4,2) = (5/2\sqrt{46}, -1/2\sqrt{46}, 2/\sqrt{46}, 1/\sqrt{46}).$$

Comprobación:

$$\| kV \| = \sqrt{25/(4\times46) + 1/(4\times46) + 16/(4\times46) + 4/(4\times46)}$$

$$= \sqrt{1/4} = 1/2.$$

2.- Dado el vector $V = (2,8,-3)$, ¿cuál es el vector de longitud 1 que se encuentra en la misma dirección de V pero sentido opuesto?

Solución:

$|k| = 1/\| V \| = 1/\sqrt{4 + 64 + 9} = 1/\sqrt{77}$, entonces,

$$k = -1/\sqrt{77}.$$

$$kV = -1/\sqrt{77}\,(2,8,-3) = (-2/\sqrt{77}, -8/\sqrt{77}, 3/\sqrt{77}).$$

Comprobación:

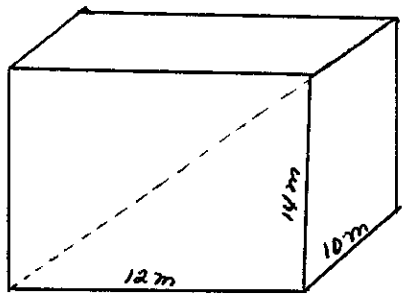
$$\| kV \| = \sqrt{4/77 + 64/77 + 9/77} = 1.$$

Observaciones :

- A aquellos vectores $V \in R^n$ que satisfacen que su norma es 1 - se les conoce como vectores unitarios.
- El proceso de encontrar un vector unitario en la dirección

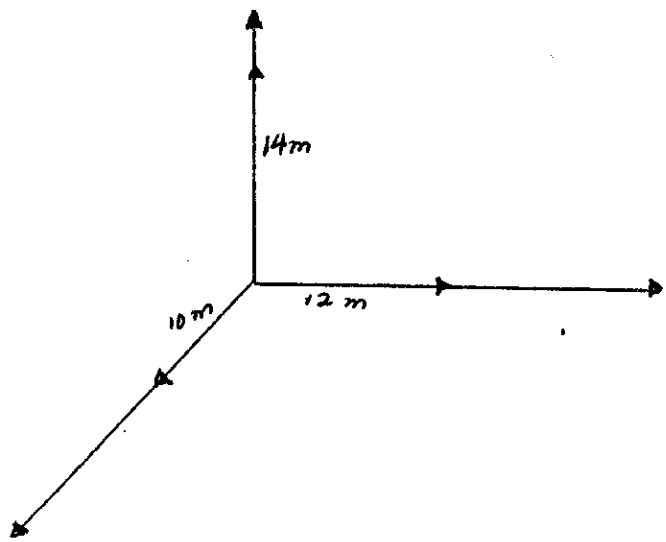
En este momento ya tenemos armas suficientes para resolver algunos de los problemas planteados en la introducción.

Problema 1.- Gráficamente interpretamos este problema como sigue



Si V denota al vector desplazamiento, ¿cuánto vale || V ||?, (es decir, ¿cuánto mide el vector V?).

a) Situando los lados del cuarto como vectores en un cierto sistema, tenemos:



De donde resulta que todos son vectores en R^3 de componentes de (10,0,0), (0,12,0), (0,0,14), respectivamente.

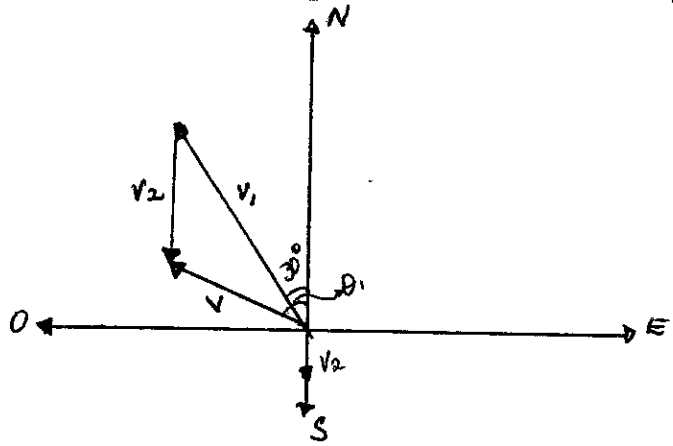
Si V_1 = (10,0,0), V_2 = (0,12,0), V_3 = (0,0,14), entonces $V = V_1 + V_2 + V_3$ = (10,12,14).

De donde:

b) Ya que por propiedad de la norma se tiene que si $V = V_1 + V_2 + V_3$ entonces $\|V\| \leq \|V_1\| + \|V_2\| + \|V_3\|$, se tiene que esta distancia recorrida en el desplazamiento -- puede ser mayor, pero no menor.

c) Con este sistema de coordenadas seleccionado, el vector V $V = (10, 12, 14)$.

Problema 2.- Una visión geométrica de este problema es - la siguiente:



Los datos son:

V_1 = velocidad de vuelo máxima, entonces $\|V_1\| = 500$.

V_2 = velocidad del viento, entonces $\|V_2\| = 40$.

Como la fuerza del viento afecta tanto a la velocidad del --- avión como a la dirección que sigue el mismo durante el via-- je, podemos trasladar el vector V_2 al extremo de V_1 y definir $V = V_1 + V_2$ como el vector "velocidad real" y θ la dirección - real del avión. Con esto, las preguntas son : $\|V\| = ?$ y - $\theta = ?$

Obteniendo las componentes en X e Y, para V_1 y V_2, tenemos que :

$V_1 = (-\|V_1\| \text{sen } 30 \text{ grados}, \|V_1\| \text{cos } 30 \text{ grados})$

$= (-500 \times 0.5, 500 \times 0.866)$

$V_1 = (-250, 433.0127)$

Por otro lado :

$V_2 = (0, -\|V\|) = (0, -40)$.

Entonces :

$V = (-250 + 0, 433.0127 - 40)$

$= (-250, 393.0127)$

De aquí que :

$$\| V \| = \sqrt{(250)^2 + (393.0127)^2} = \sqrt{216958.98} = 465.78856$$

De donde la velocidad con respecto a Tierra es de

465.78856 km./h.

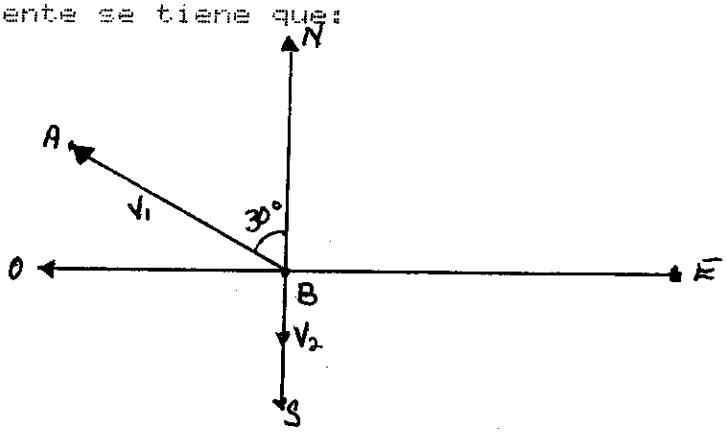
b) Como θ = arc tan (y/x), entonces

$$\theta = \text{arc tan}\left(\frac{393.0127}{-250}\right) = -57.539 \text{ grados.}$$

De donde θ = 32.46 grados.

Por lo tanto, la dirección del viaje sobre la Tierra es de 32.46 grados.

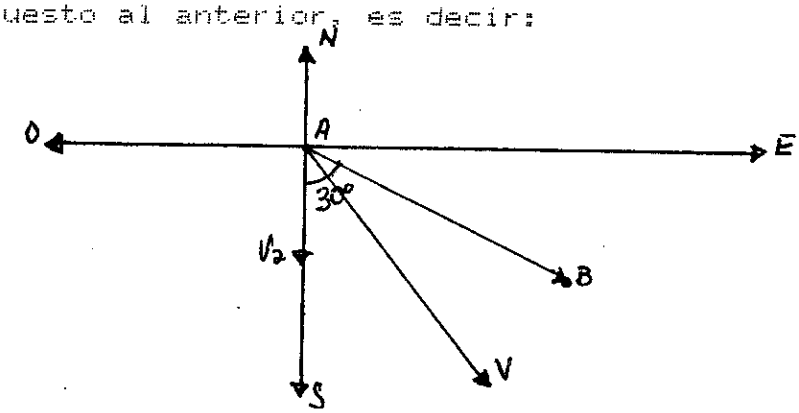
c) Gráficamente se tiene que:



La velocidad con respecto a tierra en el viaje de ida es de 1500/3 = 500 km/h., de manera que si V_1 = velocidad con -- respecto a tierra, entonces $\| V_1 \| = 500$.

Además, $\| V_2 \| = 40$, con V_2= velocidad del viento.

En el viaje de regreso suponemos que la velocidad del -- viento permanece constante (40 Km/h.), pero ahora existirá -- un nuevo vector velocidad con respecto a tierra V_1 que tendrá sentido opuesto al anterior, es decir:



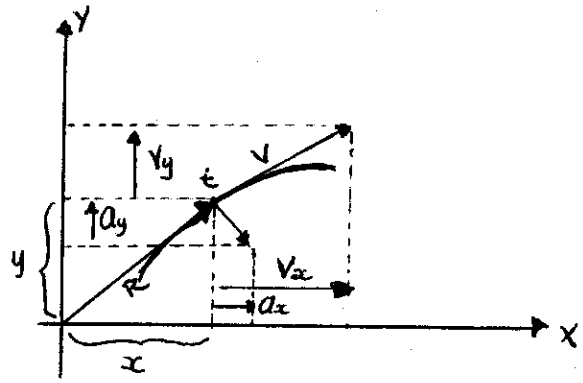
$V_1 = (\| V_1 \| \text{Sen } 30^\circ \ , -\| V_1 \| \text{Cos } 30^\circ) = (250, -433.0127)$.

Por otro lado, $V_2 = (0, -40)$.

Así, $V = (250, 393.0127)$ y $\| V \| = 465.78856$

Con esto, el tiempo para el viaje de regreso es de ----- 500/465.78856 = 1.073 horas.

Problema 3.- Movimiento de Proyectiles.- En la gráfica se --- muestran la posición R, la velocidad V y la aceleración A de la partícula en el tiempo t, así como sus respectivos vecto-- res componentes.



Estos tres vectores se interrelacionan y pueden expre--- sarse en funcion de sus componentes, al igual que en los pro- blemas anteriores. Así, $R = (x,y)$, $V = (v_x, v_y)$, y ---------- $A = (a_x, a_y)$.

Ya que las ecuaciones del movimiento son $V = V_o + a\,t$, con V_o = velocidad inicial, podemos expresarlas en términos de vectores como

$$V = (v_x, v_y) = (v_{x_o} + a_x t, v_{y_o} + a_y t)$$

$$= (v_{x_o}, v_{y_o}) + (a_x t, a_y t)$$

$$= (v_{x_o}, v_{y_o}) + (a_x, a_y)\, t$$

$$= V_o + A\, t.$$

Con esto observamos que la velocidad V en el tiempo t es la suma de la velocidad inicial V_o que tendría la partícula - (si no hubiera aceleración) más el cambio vectorial de la ve- locidad, At, adquirido durante el tiempo t bajo la acelera--- ción constante A.

Algo análogo sucede con el vector posición R.

En lo concerniente a movimiento de proyectiles, éste es de aceleración constante g dirigida hacia abajo, es decir, -- sus componentes son $a_x = 0$, $a_y = -g$. Si seleccionamos como ---

$x_o = y_o = 0$. La velocidad inicial en t = 0 (que es cuando el proyectil comienza a desplazarse) es V_o y forma un ángulo θ_o con el sentido positivo del eje de las X. Así, las componentes x_o e y_o de V_o son:

$$V_{x_o} = ||V_o||\text{Cos}\,\theta_o\,, \quad V_{y_o} = ||V_o||\text{Sen}\,\theta_o\,.$$

Ahora bien, como $a_x = 0$, entonces V_x conserva su valor -- inicial durante todo el movimiento, es decir,

$$V_x = V_{x_o} + \cancel{a_x t}^{\,0} = ||V_o||\text{Cos}\,\theta_o\,.$$

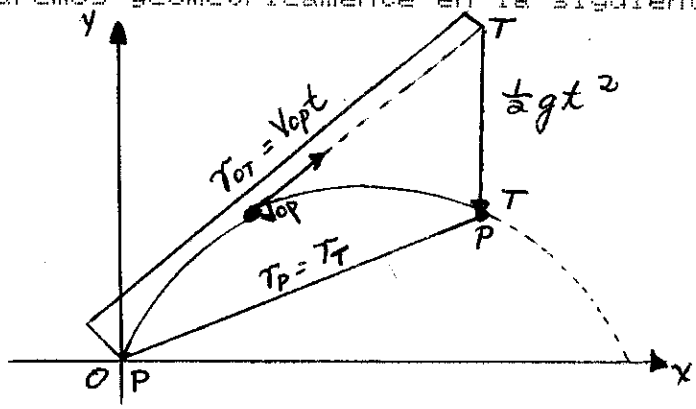
Además, como $a_y = -g$ y $V_{y_o} = ||V_o||\text{Sen}\,\theta_o$, entonces -- $V_y = V_{y_o} + a_y t = ||V_o||\text{Sen}\,\theta_o - gt$.

Por otro lado, la posición de la partícula en cualquier momento t con $x_o = 0$, $a_x = 0$ y $V_{x_o} = ||V_o||\text{Cos}\,\theta_o$ es:

$$x = (||V_o||\text{Cos}\,\theta_o)t\,, \quad y = (||V_o||\text{Sen}\,\theta_o)t - 1/2\,gt^2$$

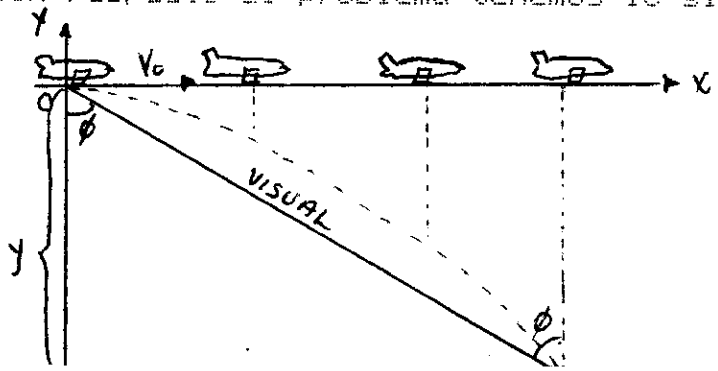
Con ésto, la trayectoria seguida por el proyectil al --- dispararse es parabólica.

Lo visualizaremos geométricamente en la siguiente forma:



En el movimiento de un proyectil, su desplazamiento a -- partir del origen en un tiempo cualquiera t puede considerarse como la suma de dos vectores : $V_{op}\,t$, en la dirección y -- sentido de V_{op} , y $1/2\,gt^2$ dirigido hacia abajo.

Con respecto al problema tenemos lo siguiente :



$g = 9.81\ m/seg.^2$
$||V_o|| = 15000\ km/h.$
$= 416.66667\ m/seg.$
$Y = -16000\ m.$
$\theta_o = 0.$

Si ϕ es el ángulo formado por la visual y el blanco B, queremos determinar su valor. Esto lo podemos hacer mediante la formula ϕ = Arc Tan x/|y|, (con |y|, por ser y negativa).

Sabemos que
$y = ||V_0|| \text{ Sen } \theta_0 t - 1/2\, gt^2$,

$-16000 = (416.66667)(\text{Sen } 0)t - 1/2(9.81)t^2$.

$$t = \sqrt{\frac{(2)(-16000)}{-9.81}} = 57.113725 \text{ seg.}$$

Esto nos sirve para calcular la componente x de B:
$x = (||V_0|| \text{ Cos } \theta_0)\, t = (416.66667)(\text{Cos } 0)(57.113725)$
$x = 23797.386$ metros.

De donde, sustituyendo en el ángulo ϕ :
$$\phi = \text{Arc Tan}\left(\frac{23797.386}{16000}\right) = 56 \text{ grados.}$$

Problema 4.- Vector Receta y Vector Precio.
a) Los ingredientes que se requieren para preparar un pay (equivalente a 6 porciones) podemos representarlos por un vector B de elementos: B = (1/4,1/4,7,.2,.05,.03,1/2).

El número de unidades de cada ingrediente para una sola porción estará dado por un vector A definido como: A = 1/6 B, ésto es, A = (1/24,1/24,7/6,.033,.0083,.005,1/12).

b) Lo mismo podemos hacer con lo que corresponde a los precios de los ingredientes. Sea P este vector, entonces
P = (182,460,7,400,900,182,140).
c) Por definición:
A•P = (1/24,1/24,7/6,.033,.0083,.005,1/2) (182,460,7,400,900, 182,140) = 68.1633, que representa el precio por cada porción.

Observación.- Este problema es un ejemplo algo burdo que representa un presupuesto en pequeño.

Problema 5.- Trabajo.- Se tiene una fuerza F tal que su magnitud es de 3 Nt., y que está actuando en la dirección del vector de componentes $(\sqrt{6}/2, \sqrt{3}, 3\sqrt{2}/2)$. Como la norma de dicho vector es 3, resulta que este vector es la fuerza F.

Ahora bien, el objeto al moverse lo hace del punto A(1,2,3) al B(2,8,11) con lo que podemos determinar un vector distancia d = AB de componentes (1,6,8), (pues $\overrightarrow{AB} = B - A$).

Además, como W = F•d , entonces $W = \sqrt{6}/2 + 6\sqrt{3} + 12\sqrt{2}$
$W \cong 28.59$ joules.

Observación.- Si la fuerza considerada no es constante,-

el cálculo del trabajo requiere conocimientos de integración, pues se reduce a una integral definida, conceptos propios de Cálculo.

1.7 D I S T A N C I A E N T R E V E C T O R E S. Definición, propiedades y ejemplos.

Esta definición, para el caso particular de vectores en R^2, coincide con la visión geométrica que se tiene del problema del cálculo de la distancia entre dos puntos del plano.

Supongamos que queremos calcular la distancia existente entre los vectores V = (3,1) y W = (4,5). En base a la definición tenemos que:

$$d(V,W) = \sqrt{(3-4)^2 + (1-5)^2} = \sqrt{17}.$$

En general tenemos:

Definición 1.7.1 .- Sean $V,W \in R^n$, definimos la distancia de V a W como:

$$d(V,W) = ||V - W||.$$

La distancia satisface las siguientes propiedades:

Sean $V,W,Z, \in R^n$,

1) $d(V,W) \geq 0$.

2) Si $d(V,W) = 0$ entonces $V = W$.

3) $d(V,W) = d(W,V)$.

4) $d(V,W) \leq d(V,Z) + d(Z,W)$. (Desigualdad del triángulo para la distancia).

Estas propiedades se demuestran en base a la definición en forma análoga a como se hace con la norma.

Ejemplo:

Problema 6.- Un tetraedro regular con tres de sus vértices en los puntos A(4,0,0), B(-4,0,0) y $C(0,4\sqrt{5},0)$ tiene el cuarto vértice desconocido. Llamemos P(x,y,z) a este vértice. Tenemos que:

$$d(A,P) = \sqrt{(4-x)^2 + y^2 + z^2}$$

$$d(B,P) = \sqrt{(-4-x)^2 + y^2 + z^2}$$

$$d(C,P) = \sqrt{x^2 + (4\sqrt{5}-y)^2 + z^2}$$

Ahora bien, como es un tetraedro regular se tiene que $d(A,P) = d(B,P)$. Esto es,

$(4-x)^2 + y^2 + z^2 = (-4-x)^2 + y^2 + z^2$

$16 - 8x + x^2 = 16 + 8x + x^2$, de donde $-16x = 0$ y por lo tanto $x = 0$.

Por otro lado, $d(A,P) = d(C,P)$.

$(4-x)^2 + y^2 + z^2 = x^2 + (4\sqrt{5}-y)^2 + z^2$

$16 - 8x = 80 - 8\sqrt{5}\,y$. Pero como $x = 0$, $y = 8/\sqrt{5}$.

Si calculamos $d(A,B) = 8$.

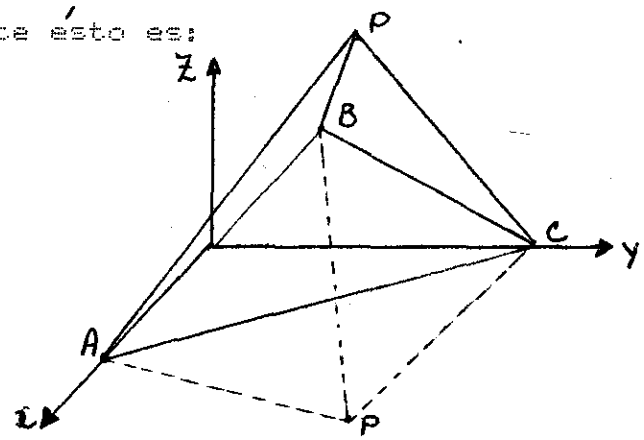
Como además $d(A,B) = d(A,P)$ con $x = 0$, $y = 8/\sqrt{5}$, entonces $d(A,P) = \sqrt{4^2 + (8/\sqrt{5})^2 + z^2} = 8$. Despejando el valor de z llegamos a que:

$z = \pm\sqrt{176/5}$.

Por lo tanto, el vértice buscado está en:

$(0,8/\sqrt{5},\sqrt{176/5})$ o en $(0,8/\sqrt{5},-\sqrt{176/5})$.

Gráficamente ésto es:



1.8 A N G U L O E N T R E D O S V E C T O R E S

Vectores perpendiculares. Definición del ángulo entre dos vectores. Proyecciones. Ejemplos.

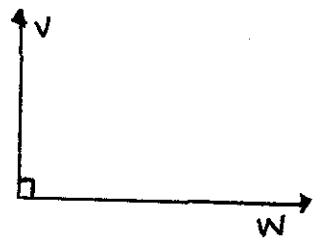
Aprovechando lo visto expondremos ahora la noción de proyección para después, en base a ésto, dar una definición para el caso general de lo que es el ángulo entre dos vectores y cómo determinar su valor.

Si tenemos dos vectores V,W la proyección del vector V sobre el vector W será un nuevo vector colocado sobre la recta determinada por W y obtenido al trazar la perpendicular desde el extremo de V a esta recta.

El papel que juega la proyección de un vector sobre

es muy importante.

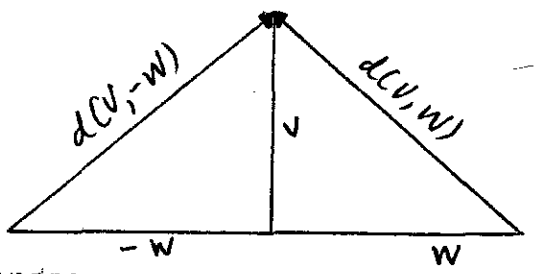
Si resultara que los vectores son perpendiculares entonces la proyección de V sobre W no es un vector:



En todo caso, la proyección corresponde solamente a un punto, con lo que podemos decir que dicha proyección como vector es nula.

La definición de distancia entre dos vectores que se dió en la sección anterior nos proporciona una forma de definir la perpendicularidad entre dos vectores. ¿Por qué ?

Geométricamente (fácil de imaginar en R^2), dos vectores son perpendiculares si y sólo si $d(V,W) = d(V,-W)$.



Desarrollando:

$d(V,W) = ||V - W||$; $d(V,-W) = ||V -(-W)|| = ||V + W||$.

$||V -W|| = ||V + W||$.

Por un lado:

$||V - W||^2 = (V - W)\bullet(V - W) = V\bullet V - 2V\bullet W + W\bullet W$.

Por otro:

$||V + W||^2 = (V + W)\bullet(V + W) = V\bullet V + 2V\bullet W + W\bullet W$.

Igualando:

$V\bullet V - 2V\bullet W + W\bullet W = V\bullet V + 2\ V\bullet W + W\bullet W$.

$4V\bullet W = 0$, entonces $V\bullet W = 0$.

En el otro sentido hay que demostrar que si $V \cdot W = 0$ entonces $d(V,W) = d(V,-W)$.

Como $d(V,W) = ||V - W||$, entonces

$$||V - W||^2 = (V - W) \cdot (V - W)$$
$$= V \cdot V - 2V \cdot W + W \cdot W$$
$$= V \cdot V + W \cdot W .$$

Además,

$d(V,-W) = ||V + W||$, entonces

$$||V + W||^2 = (V + W) \cdot (V + W)$$
$$= V \cdot V + 2V \cdot W + W \cdot W$$
$$= V \cdot V + W \cdot W .$$

De aquí que $||V - W||^2 = ||V + W||^2$ y por lo tanto $||V - W|| = ||V + W||$, o sea $d(V,W) = d(V,-W)$.

Observación.- Podemos aprovechar el desarrollo anterior para enunciar el concepto de perpendicularidad como sigue:

Definición 1.8.1.- Sean $V,W \in R^n$. Decimos que V y W son perpendiculares si y sólo si $V \cdot W = 0$.

Observaciones:

- Fijémonos que la demostración que se acaba de hacer es la primera del tipo $\Leftrightarrow$ "si y sólo si" que se presenta. Este conectivo lógico se interpreta como una doble implicación y significa que las proposiciones que está uniendo son equivalentes. Para demostrarse debe de verse tanto que $p \Rightarrow q$ como que $q \Rightarrow p$, donde p y q son proposiciones arbitrarias. En el caso que nos ocupa: p: $d(V,W) = d(V,-W)$, (o lo que es lo mismo, V y W son perpendiculares) y q: $V W = 0$.

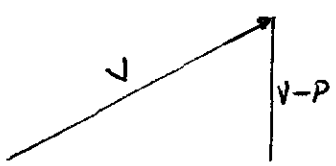
- Notemos además que si dos vectores son perpendiculares también lo son cualquier pareja de múltiplos de ellos, lo cual es fácil de demostrar. Hay que verificar entonces que si V y W son perpendiculares entre sí, entonces αV y λW también lo son, con $\alpha, \lambda \in R$; en otras palabras, que $(\alpha V) \cdot (\lambda W) = 0$.

Por hipótesis $V \cdot W = 0$, y como $(\alpha V) \cdot (\lambda W) = \alpha (V \cdot (\lambda W))$
$= (\alpha \lambda)(V W) = 0$.

Por lo tanto αV y λW son perpendiculares.

Retomando lo comentado páginas atrás, al buscar la proyección de V sobre W, lo que buscamos es un vector P que satisfaga que $V - P$ sea perpendicular a W y tal que se pueda expresar como $P = cW$, para algún $c \in R$. Obviamente aquí ya consideramos que entre V y W existe un ángulo distinto de $\pi/2$.

La representación geométrica en R^2 sería:



En lo general,
$(V - P)\cdot W = 0 = (V - cW)\cdot W = u.$
$V\cdot W - cW\cdot W = 0$, y así se deberá tener que:

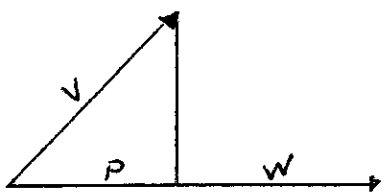
$$c = (V\cdot W)/(W\cdot W) = (V\cdot W)/(\| W \|)^2.$$

Esta constante c determina el tipo de ángulo que existe entre V y W y recibe el nombre de "componente de V a lo largo de W".

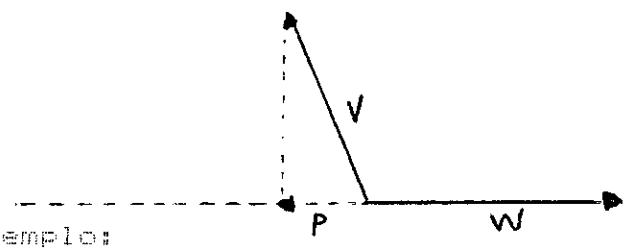
Entonces $P = cW = ((V\cdot W)/(W\cdot W))\ W.$
Estamos en condiciones de formalmente establecer la

Definición 1.8.2.- Dados dos vectores $V,W \neq 0 \in R^n$ la proyección vectorial de V sobre W es el vector
$P = cW = ((V\cdot W)/(W\cdot W))\ W.$

Observación.- Se ha afirmado que c determina el tipo de ángulo que existe entre V y W. En efecto, si éste fuera agudo, la dirección del vector proyección coincide con la dirección de W. Esto, claro, si la constante de proyección es positiva.



Si el ángulo fuera obtuso, la dirección de W y el vector proyección son contrarias, aquí la constante de proyección debe ser negativa. Ahora bien, el signo de c depende exclusivamente de $V\cdot W$, dado que el denominador es siempre positivo. Entonces, a fin de cuentas quien nos determina el tipo de ángulo es $V\cdot W$.



Ejemplo:
Si tenemos los vectores V = (-1,3,5), W = (6,8,4), se desea calcular la proyección de V a lo largo de W. Además dígase cuál es el tipo de ángulo que existe entre ellos.

Solución.- Obtenemos primero el valor de c.
$c = V\cdot W / W\cdot W = 38/116.$

La proyección vectorial de V sobre W es:

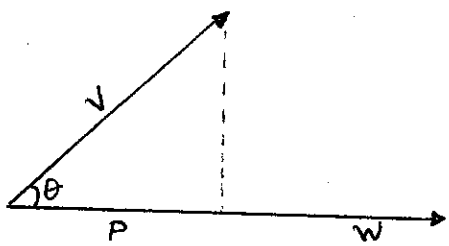
Como el valor de c es positivo entonces el ángulo es --- agudo.

Observación.- La proyección vectorial en el caso en que los vectores V y W forman un ángulo de 0 grados (es decir, V y W son colineales con la misma dirección y sentido) consiste en lo siguiente: como P = ((V•W)/(W•W)) W, entonces ---------- V - P = V - ((V•W)/(W•W)) W = 0 si y sólo si se cumple que -- V = ((V•W)/(W•W)) W = P. Es decir, la proyección vectorial - de V sobre W es el mismo vector V. Análogamente, si V y W - son colineales de sentidos opuestos, entonces:
P = - ((V•W)/(W•W)) W. (V•W < 0 puesto que se tiene un án-- gulo mayor que $\pi/2$).

Así, V - P = V + ((V•W)/(W•W)) W = 0 si y sólo si ------ V = -((V•W)/(W•W)) W = P.

De nueva cuenta, la proyección vectorial de V sobre W -- resulta ser el mismo vector V.

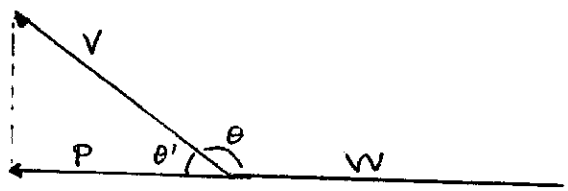
Finalmente, definiremos el ángulo entre dos vectores dados. Para esto, separaremos en dos casos:
Primer caso.- Obsérvese en la figura que P = cW y c>0.



Aquí, Cos θ = ||P||/||V|| = ||cW||/||V||
= |c| ||W||/||V|| = c ||W||/||V||.
Pero
c = V•W/||W||^2, de donde:

Cos θ = ((V•W/||W||^2) ||W||)/||V|| = V•W/||V|| ||W||.

Segundo caso.-



P = cW, con c<0. Por ser ángulos suplementarios, tenemos que Cos θ = - Cos θ'. Cos θ' = ||P||/||V||.

Cos θ = - ||P||/||V|| = - ||cW||/||V||
= - |c| ||W||/||V|| = -(-c) ||W||/||V||

Con ésto concluimos la siguiente:

Definición 1.8.3.- Sean $V,W \in R^n$, $V \neq 0$ y $W \neq 0$. Definimos el ángulo θ entre V y W al número tal que

$$\text{Cos } \theta = V \cdot W \,/\, ||V||\,||W||.$$

Observaciones:
- Si $V \cdot W > 0$ entonces $0 \leq \theta < \pi/2$.
- Si $V \cdot W < 0$ entonces $\pi/2 < \theta \leq \pi$.
- Si $V \cdot W = 0$ entonces $\theta = \pi/2$.
- Existe otra manera de definir el producto punto:

Definición 1.8.4.- Si $V,W \in R^n$, y θ es el ángulo comprendido entre ellos, definimos el producto escalar o punto como:

$$V \cdot W = ||V||\,||W|| \text{ Cos } \theta.$$

- Otra forma de calcular la constante c (valor de la compo--- nente de V a lo largo de W) es en base al ángulo:

$$c = ||V|| \text{ Cos } \theta / \,||W||.$$

Si resumimos las principales expresiones algebraicas obtenidas hasta el momento tenemos:

a) Norma de un vector $V \in R^n$: $||V|| = \sqrt{V \cdot V}$.

b) Distancia entre $V, W \in R^n$: $d(V,W) = ||V-W||$.

c) Perpendicularidad entre $V, W \in R^n$: V y W son perpendicu-- lares si y sólo si $V \cdot W = 0$.

d) Angulo formado por V y W : $\text{Cos } \theta = V \cdot W / ||V||\,||W||$.

e) Proyección de V sobre W:

$$cW = ((V \cdot W)/||W||^2)\, W.$$

Observamos que en todas ellas el producto punto juega un papel de gran importancia. He ahí una justificación al hecho de que se haya definido un producto de vectores de tal forma.

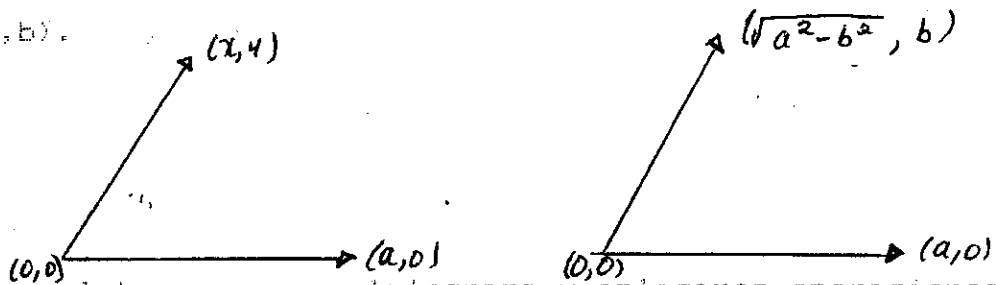
Pasaremos ahora a resolver los últimos problemas plan--- teados en la introducción.

Problema 7.- Este problema pide que se muestre que las diago- nales de un rombo son perpendiculares.

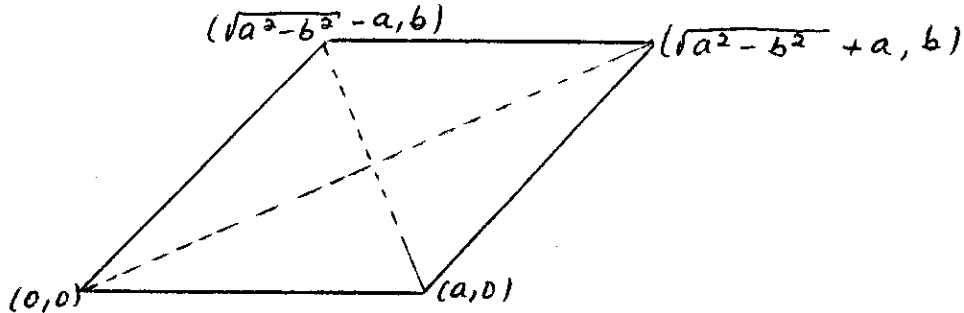
Sean dos vectores A y B tales que A = (a,0) y B = (x,y) y además que tienen la misma magnitud. Es decir, $||A|| = ||B||$. Con ésto:

$$\sqrt{x^2 + y^2} = \sqrt{a^2} \qquad x = \sqrt{a^2 - y^2}\,, \quad \text{Si } y = b \text{ entonces --}$$

$x = \sqrt{a^2 - b^2}$. Así, el vector B tiene como componentes - $(\sqrt{a^2 - b^2}, b)$.



Si completamos un paralelogramo y aplicamos operaciones entre vectores, tendremos que las diagonales del paralelogramo tendrán como componentes:



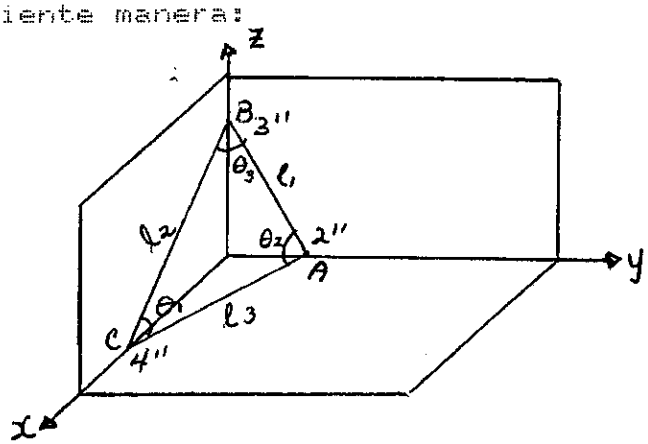
Para comprobar que estas diagonales digamos D_1 y D_2, - son perpendiculares es necesario verificar que su producto -- punto es igual a cero.

$$D_1 \cdot D_2 = (\sqrt{a^2 - b^2} - a,\ b)\ (\sqrt{a^2 - b^2} + a,\ b)$$

$$= (\sqrt{a^2 - b^2} - a)\ (\sqrt{a^2 - b^2} + a) + b^2$$

$$= (\sqrt{a^2 - b^2})^2 - a^2 + b^2 = 0,$$ que es exactamente lo que queríamos demostrar.

Problema 8.- Situaremos la tabla en un sistema de referencia de la siguiente manera:



Con ésto: A = (0,2,0), B = (0,0,3), C = (4,0,0).

Así:

l_1 = AB = (0,-2,3); || AB || = $\sqrt{13}$.

l_2 = BC = (4,0,-3); || BC || = 5.

Si llamamos θ_1 a el ángulo entre CA y CB,
Cos θ_1 = CA•CB / ||CA|| ||CB|| = 0.7155, entonces el ángulo θ_1 = 44.3 grados.

Si llamamos θ_2 a el ángulo entre AC y AB,
Cos θ_2 = AC•AB/ || AC || || AB || = 0.248, entonces el ángulo θ_2 = 75.6 grados.

Si llamamos θ_3 a el ángulo entre BC y AB,
Cos θ_3 = BC•BA/ ||BC|| ||BA|| = 0.4992, entonces el ángulo θ_3 = 60 grados.

Cerraremos este capítulo con algunos :

1.9 R E S U L T A D O S G E N E R A L E S

En base a todo lo expuesto en el transcurso de este capítulo, podemos hacer uso de este material y generalizar algunos resultados a vectores cualesquiera. Veamos lo que se conoce como:

TEOREMA DE PITAGORAS GENERALIZADO

Proposición 1.9.1.- Si V, W $\in R^n$ y son perpendiculares, entonces $||V + W||^2 = ||V||^2 + ||W||^2$.

Desarrollando:

$$\begin{aligned} ||V + W||^2 &= (V + W) \cdot (V + W) \\ &= V \cdot V + 2\, V \cdot W + W \cdot W \\ &= V \cdot V + W \cdot W \\ &= ||V||^2 + ||W||^2 . \end{aligned}$$

Por último veamos un importante resultado conocido como:

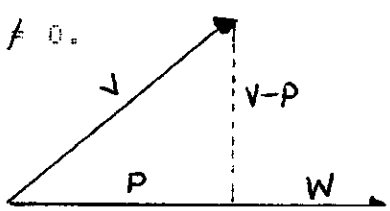
DESIGUALDAD DE CAUCHY-SCHWARZ

Proposición 1.9.2.- Sean V,W$\in R^n$, entonces

$$|V \cdot W| \leq ||V||\, ||W|| .$$

a) Si V = 0 ó W = 0, la afirmación es cierta.

b) Supongamos V,W $\neq$ 0.



Aplicando el Teorema de Pitágoras Generalizado con los vectores V y V - P:

$$||V||^2 = ||V-P||^2 + ||P||^2$$

$$||P||^2 \leq ||V||^2$$

$$||P|| \leq ||V||.$$

Pero $P = cW = (V \cdot W / ||W||^2) W.$

Entonces, $||P|| = || (V \cdot W / ||W||^2) W || = |V \cdot W| / ||W||.$

Con lo que $|V \cdot W| / ||W|| \leq ||V||$ $\quad |V \cdot W| \leq ||V|| \, ||W||$

Esto mismo puede verse trabajando la definición para el producto punto de vectores en base al ángulo formado entre -- ellos:

$V \cdot W = ||V|| \, ||W|| \cos \theta$ $\quad \cos \theta = V \cdot W / ||V|| \, ||W||.$

Pero $-1 \leq \cos \theta \leq 1$, entonces $-1 \leq V \cdot W / ||V|| \, ||W|| \leq 1$

$- ||V|| \, ||W|| \leq V \cdot W \leq ||V|| \, ||W||$

$|V \cdot W| \leq ||V|| \, ||W||.$

C A P I T U L O 2

SISTEMAS DE ECUACIONES LINEALES

2.1 I N T R O D U C C I O N .

Al introducirnos a este capítulo, estamos penetrando a - lo que será la parte medular del curso : la resolución de --- Sistemas de Ecuaciones Lineales (SEL). Desarrollaremos aquí dos de los métodos más utilizados para ello : el llamado Método de Eliminación de Gauss y el Método de Gauss-Jordan.

Estamos seguras de que el contenido de este capítulo a-- traerá poderosamente la atención del alumnado, puesto que verán un material que tiene aplicaciones muy variadas e interesantes, ya sea en el área de estudio específico de sus carreras, como en el resto de los cursos que llevan.

Además, este material ya les es un tanto conocido, puesto que desde la secundaria se tiene contacto con los SEL, aunque de tamano pequeño, a lo más de tres ecuaciones con tres - incógnitas.

Aparte de las aplicaciones, ya de por sí interesantes, - los conceptos y algoritmos que aquí se trabajen serán de gran utilidad posterior.

2.2 B R E V E R E S E Ñ A H I S T O R I C A.

Hablar de la historia de los Sistemas de Ecuaciones Li-- neales significa hablar primero de lo que es una ecuación lineal.

El problema de resolver una de las más sencillas de las ecuaciones lineales, la que tiene la forma ax = b, es algo -- que ya ocupaba a los hombres antiguos. Los egipcios incluso ya sabían resolverla, pues se les había aparecido como resultado de dificultades a la hora de realizar ciertas medicio--- nes.

La aparición de los SEL se debe más que nada a la nece-- sidad de resolver ciertos problemas prácticos.

Francois Viete (1540-1603) y Rene Descartes (1596-1650) fueron los que establecieron la notación que actualmente se - utiliza. Anteriormente todo se escribió en latín y no exis-- tía una notación que fuera práctica.

Carl Gustav Jacob Jacobi (1804-1851) también planteó la necesidad de resolver un SEL obtenido de sus estudios en Fí-- sica, aunque nunca llegó a solucionarlo.

Colin Mac Laurin (1698-1746) en un tratado de Algebra --

senta la solución habitual de las ecuaciones simultáneas por eliminación sucesiva de incógnitas. En uno de sus capítulos presenta además otra solución mediante lo que se conoce como determinante, e incluso llega a mencionar la famosa regla a-- tribuida hoy a Gabriel Cramer.

Gabriel Cramer (1704-1752) publicó en 1750 la misma re-- gla que dos años antes había aparecido en el tratado de Mac - Laurin. Lo más probable es que esta regla sea conocida como la Regla de Cramer debido a la superioridad de la notación.

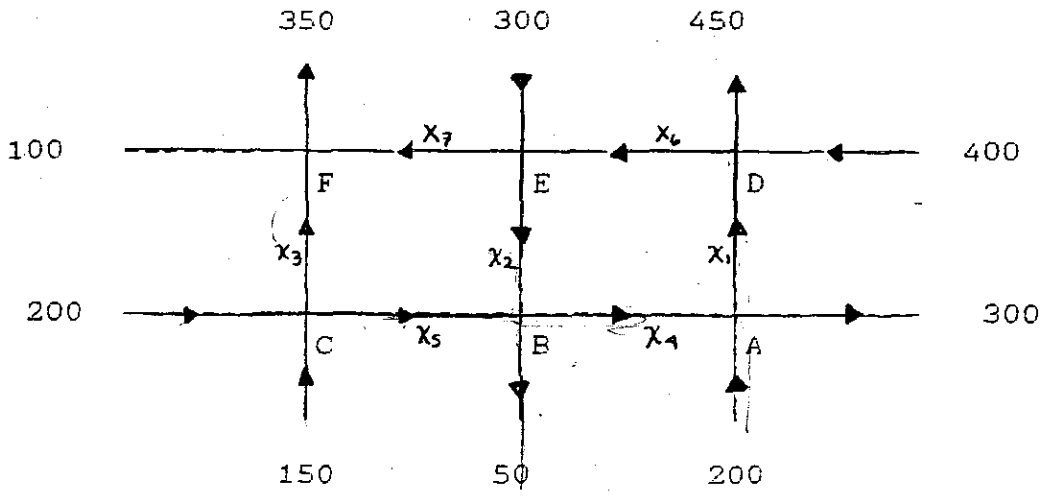
Posteriormente, Karl Friedrich Gauss (1777-1855) dá un método para resolver el caso general del sistema Ax = b, co-- nocido con el nombre de Método de Eliminación de Gauss y cu-- yas ideas fundamentales se remontan a la época de los babilo- nios.

En la actualidad, el problema de resolver un SEL se ha - visto reforzado con el empleo de la computadora y ha alcanza- do grados de desarrollo que son en verdad asombrosos. Ha da- do lugar a la aparición de ramas de las Matemáticas como son el Análisis Numérico y la Programación Lineal que en nuestros días son de utilidad insospechada.

2.3 P R O B L E M A S A P L I C A D O S .

Problema 1.- Un modelo sobre flujo de tránsito.- Lo que ve- remos a continuación es cómo un problema de flujo de tránsito citadino puede modelarse mediante un SEL.

En el diagrama siguiente se muestra una sección de la -- ciudad de Hermosillo, Sonora.



Las flechas nos indican el sentido en el que viajan los vehículos y los números que aparecen representan la cantidad de vehículos que entran y que salen en un período de una ho-

Esta medición se hizo colocando contadores en cada una - de las intersecciones en una de las hora "pico" del tráfico - citadino (de 8:00 a 9:00 a.m.). Se han designado mediante -- letras mayúsculas las intersecciones de las calles y mediante variables con subíndice el flujo de tráfico entre intersec--- ciones adyacentes.

La calle que vá de B a A estará en etapa de bacheo. Los trabajos se realizarán a partir de las 8:00 a.m. Se requiere por lo tanto que el tráfico sea mínimo, pero sin causar embotellamientos en otras calles.

¿Qué recomendación daría usted al Departamento de Tránsito, encargado de resolver este problema ?

Problema 2.- Teoría de Circuitos Eléctricos.- Un circuito - eléctrico es una serie de dispositivos eléctricos conectados entre sí de tal manera que existe un flujo de corriente entre ellos y se produce un efecto determinado.

Un circuito eléctrico consta de dos tipos de elementos : activos y pasivos. Entre los elementos activos se tienen las fuentes de voltaje o fuentes de corriente ; entre los pasivos están las resistencias, transformadores, inductancias, capa-- citancias, etc.

Un circuito eléctrico puede dividirse en mallas o redes para desglosar la información que pueda tenerse de él. Una - malla o red es una interconexión de dispositivos eléctricos - como transistores, baterías, resistencias, capacitores, etc.

En un circuito se tienen ciertos puntos en los cuales -- coinciden (o hay conexión entre) varios conductores. Este -- punto de unión recibe el nombre de nodo.

Debido a la naturaleza propia de los elementos que se -- cuentan entre los del tipo pasivo, resulta que los más sencillos son las resistencias.

Dentro de la teoría de circuitos eléctricos la hipótesis fundamental de linealidad se conoce como la Ley de Ohm. Esta ley establece que :

> "La intensidad de corriente eléctrica es directamente -- proporcional al voltaje e inversamente proporcional a - la resistencia".

Pero también se hace presente una propiedad de aditivi-- dad, ya que :

"Si la misma corriente pasa por una red con solamente -- dos (o más) resistencias conectadas en serie, la caída de --- voltaje a través de la red es la suma de las caídas indivi---

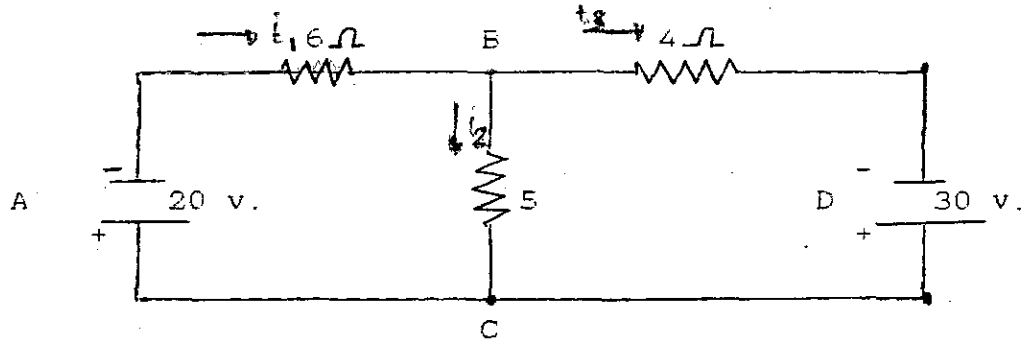
Esto se conoce como la Segunda Ley de Kirchhoff.

Pero además se tiene la Primera Ley de Kirchhoff que dice que :

"La corriente que llega a cierto nodo en un circuito --- eléctrico es igual a la corriente que sale de él".

Así, la teoría de circuitos lineales ofrece un excelente ejemplo de SEL utilizado en Ingeniería, como se vé a conti--- nuación:

Considere el siguiente circuito. Determine los valores de la corriente en cada resistencia.



Problema 3.- Modelos de Leontief (*).- Este es un modelo -- usado frecuentemente en Economía y consiste en lo siguiente:

Supongamos que un sistema económico tiene n sectores --- económicos. Cada sector tiene dos tipos de demanda : exter-- na, que viene de afuera del sistema, e interna, que es la hecha a un sector por otro sector del mismo sistema.

Tenemos e_i para representar la demanda externa hecha al i-ésimo sector y a_{ij} para representar la demanda interna hecha al i-ésimo sector por el j-ésimo sector. Más precisamente, a_{ij} representa el número de unidades del producto del -- i-ésimo sector necesario para producir una unidad del producto del sector j. Sea x_i la producción del sector i. Supon-- gamos que la producción de cada sector es igual a su demanda (ésto es, no hay sobreproducción). La demanda total es igual a la suma de las demandas externa e interna. Para calcular - la demanda interna del sector i notemos que $a_{i1}\ x_1$ es la - demanda hecha al sector i por el sector 1; así, la demanda -- interna total del sector i es

$$a_{i1}\ x_1 + a_{i2}\ x_2 + \ldots + a_{in}\ x_n .$$

(*) Llamado así debido al economista americano Wassily W.---- Leontief.

Llegamos así al siguiente sistema de ecuaciones obtenido de igualar la demanda total con la producción de cada sector:

$$\begin{aligned} a_{11}\, x_1 + a_{12}\, x_2 + \ldots + a_{1n}\, x_n + e_1 &= x_1 \\ a_{21}\, x_1 + a_{22}\, x_2 + \ldots + a_{2n}\, x_n + e_n &= x_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}\, x_1 + a_{n2}\, x_2 + \ldots + a_{nn}\, x_n + e_n &= x_n . \end{aligned}$$

De donde

$$\begin{aligned} (1 - a_{11})x_1 - a_{12}\, x_2 - \ldots - a_{1n}\, x_n &= e_1 \\ - a_{21}\, x_1 + (1 - a_{22})x_2 - \ldots - a_{2n}\, x_n &= e_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ - a_{n1}\, x_1 - a_{n2}\, x_2 - \ldots + (1 - a_{nn})x_n &= e_n \end{aligned}$$

Este sistema de n ecuaciones con n incógnitas es muy importante en el análisis económico. Ejemplo :

La economía de un cierto estado está sostenida por tres sectores económicos que dependen entre sí, pero que no dependen de sectores del resto del país. Estos sectores son : agricultura, ganadería y pesca. En la siguiente tabla se muestra la porción de cada producto consumido por cada uno de los sectores :

		Producción		
		Agricultura	Ganadería	Pesca
Consumo	Agricultura	3/6	4/7	2/5
	Ganadería	2/6	3/14	1/5
	Pesca	1/6	3/14	2/5

Este arreglo podemos interpretarlo como sigue : la agricultura consume 3/6 de su propia producción, la ganadería consume 1/5 de la producción del sector de la pesca, etc.

Supongamos que los ingresos de cada sector económico están dados por I_2, I_2 e I_3, respectivamente.

Bajo el supuesto de que el ingreso ocasionado por las ventas del producto es igual al gasto ocasionado por la producción, ésto es la producción de cada sector es igual a su consumo (es decir, no hay sobreproducción), determine los ingresos de cada uno de los miembros del sistema económico.

Observación.- El tipo de modelo que se manejó es muy utilizado en Economía y recibe el nombre de Modelo Cerrado de Insumo Producto de Leontief.

Problema 4.- Ecuaciones Químicas.- Las reacciones químicas pueden expresarse mediante representaciones simbólicas llamadas Ecuaciones Químicas, por su semejanza con las ecuaciones matemáticas. La igualdad química es una expresión de la ley

tar simbólicamente una reacción química, debemos tomar en --- cuenta que :

i) En el primer miembro de la ecuación se tienen las fórmu-- las o símbolos de las sustancias que reaccionan.
ii) En el segundo miembro se tienen los símbolos o fórmulas de las sustancias que se forman en la reacción.
iii) Se equilibra la ecuación con respecto a los átomos, éste es, que el número de átomos que intervienen en el primer - miembro aparezca en el segundo. A este proceso se le conoce como Balanceo de Ecuaciones.

Ejemplo :
Balancee la ecuación que representa a la reacción química : Hidróxido de Sodio con Acido Sulfúrico se transforma en Sulfato de Sodio más Agua.

$$Na\ OH + H_2\ SO_4 \longrightarrow Na_2\ SO_4 + H_2O\ .$$

2.4 S I S T E M A S D E E C U A C I O N E S L I N E A- L E S .- Definicion, clasificación y ejemplos.

Históricamente, el Algebra Lineal evolucionó a partir de la idea de formular un método general para investigar la so-- lución de SEL.

Las técnicas que actualmente se manejan para resolver un SEL son algo antiguas, sin embargo han resultado altamente -- provechosas a partir del adelanto de las computadoras elec--- trónicas, mismas que realizan un gran número de operaciones - con enorme rapidez.

Tanto la teoría de las matrices como la de los determi-- nantes dan un procedimiento para la solución de sistemas de - ecuaciones lineales, aunque ésto se verá posteriormente.

A continuación daremos ciertos conocimientos básicos necesarios para poder tratar los métodos de solución de SEL.

Definición 2.4.1.- Una ecuación lineal de n incógnitas es una expresión del tipo $a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b$, --- donde x_i son las incógnitas, a_i los coeficientes, $i = 1, 2, \ldots, n$ y b es el término constante. Salvo que se indique --- otra cosa, entenderemos que trabajamos con $a_i , x_i , b \in R$.

Encontrar valores $x_1, x_2, \ldots, x_n$ que satisfagan la igualdad pedida significa encontrar una solución de la ecuación.

Ya que tanto hemos hablado de la solución de SEL, debe-- mos definir formalmente lo que son los SEL. Para ello tene-- mos :

ecuaciones lineales de la forma

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

se dice que tenemos un sistema de m ecuaciones lineales con n incógnitas. Cada a_{ij} , $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ recibe el nombre de coeficiente, cada x_i es una incógnita y cada b_i es llamado término independiente.

Observaciones :
-Remarcamos que estamos trabajando con a_{ij} , x_i , $b_i \in R$.
-El doble subíndice de cada coeficiente es de bastante utilidad, porque el primero de ellos nos indica la ecuación a la cual pertenece la constante y el segundo de ellos indica la incógnita a la cual acompañan.
-Resolver un sistema de este tipo significa encontrar los conjuntos de valores $x_1, x_2, \ldots, x_n$ que satisfagan las m ecuaciones. Usualmente se habla de cada uno de esos conjuntos de valores simplemente como un vector solución.

Definición 2.4.3.- Cuando cada $b_i = 0$, $i = 1, 2, \ldots, m$ se dice que el sistema es HOMOGENEO. Cuando m = n, el sistema se llama CUADRADO.

Para la solución de un sistema de ecuaciones, se pueden presentar dos posibilidades :
a) Que no exista solución, en cuyo caso se dice que el sistema es incompatible o inconsistente.
b) Que exista solución ; se trata entonces de un sistema consistente, pero aún aquí se abren otras dos poibilidades :
b') Que la solución sea única (sistema determinado)
b") Que existan infinidad de soluciones (sistema indeterminado).

Esta clasificación que hemos hecho de los SEL se justificará con la teoría que veremos más adelante. Esto es, como resultado de la teoría se concluirá que si acaso un SEL tiene solución, ésta o es única o son una infinidad. En los sistemas pequeños de 2 x 2 y de 3 x 3, estas posibilidades se pueden analizar geométricamente.

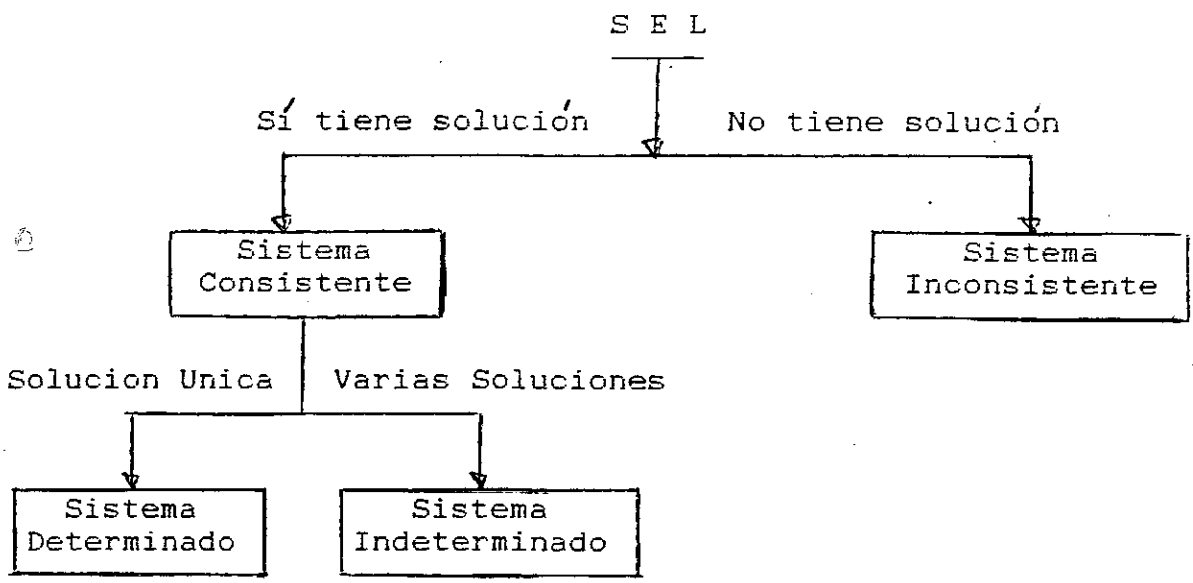
Por ejemplo, en sistemas de 2 x 2, la gráfica de cada una de las ecuaciones es una recta. La solución será el conjunto de puntos donde las rectas se corten. El sistema incompatible quedará determinado por un par de líneas paralelas ; el sistema determinado por un par de rectas que se intersecan y el sistema indeterminado por un par de rectas coincidentes.

En el caso de los sistemas de 3 x 3, cada una de las e--

un plano (éso cuando al menos uno de los coeficientes sea --- distinto de cero).

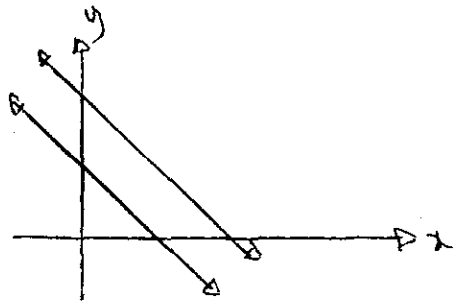
Los sistemas incompatibles de 3 x 3 quedarán graficados como tres planos de los cuales al menos dos son paralelos. Los determinados son tres planos que se cortan en un solo --- punto y los indeterminados los podemos graficar de varias maneras, dado que la intersección de los tres planos, aparte de ser un punto (caso anterior), puede ser una recta o un plano.

Resumiremos todo ésto en un cuadro sinóptico de la si--- guiente manera :



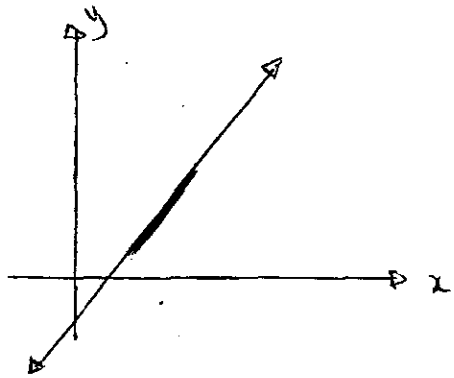
SEL (2x2)

Sistema Inconsistente

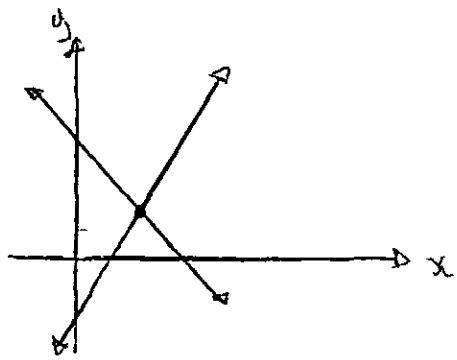


Sistema Consistente

Sistema Indeterminado

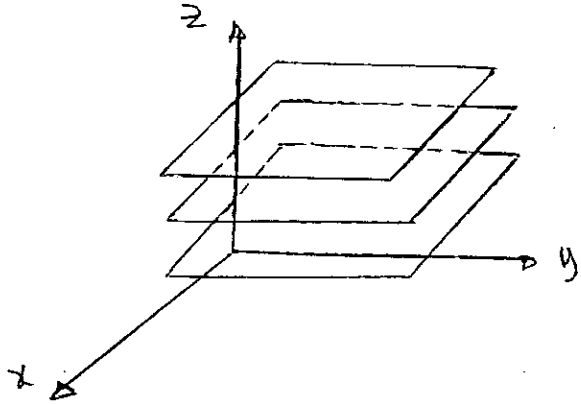


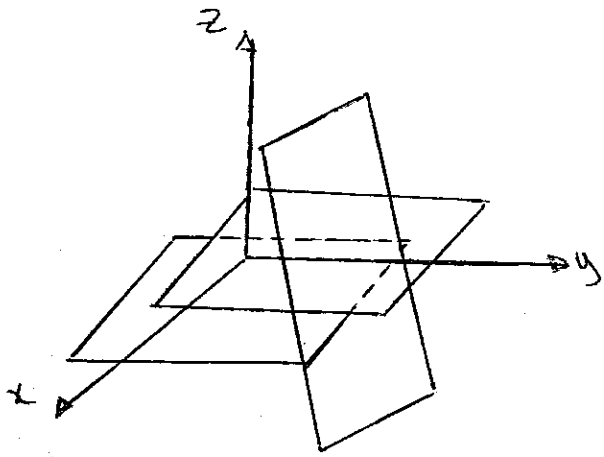
Sistema Determinado

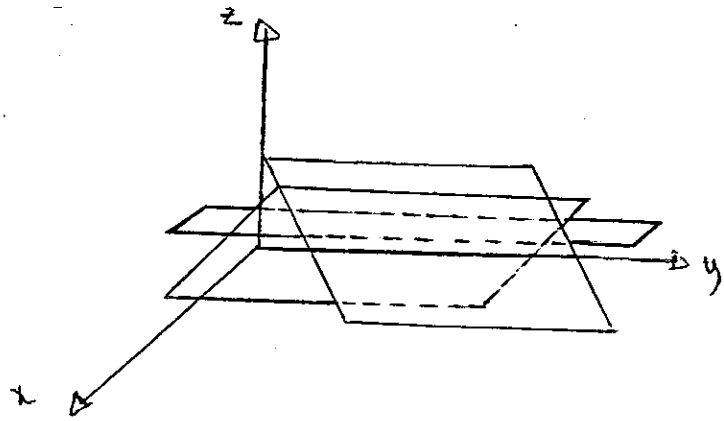


S E L
(3 X 3)

Sistemas Inconsistentes

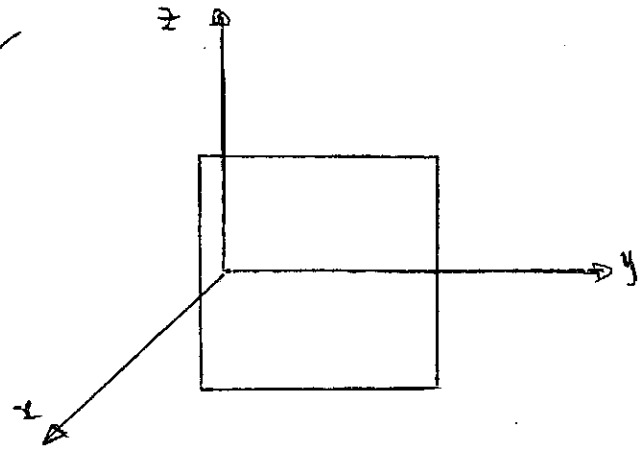


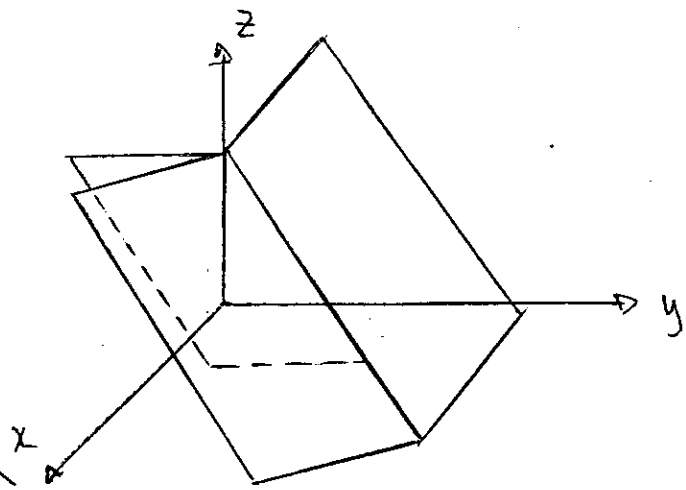




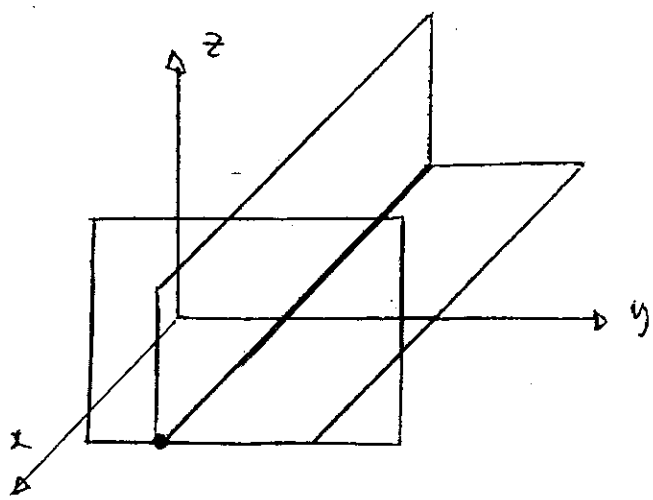
S i s t e m a s C o n s i s t e n t e s

Sistemas Indeterminados.





Sistema Determinado.



Antes de adentrarnos en los métodos de solución de los SEL, y dado que ya conocemos la forma que éstos tienen, vamos a tratar de encontrar cuáles de ellos nos representan los problemas que se enunciaron en la sección anterior.

Problema 1.- En el problema de flujo de tráfico se tienen las siguientes características :
A, B, C, D, E, F designan las intersecciones de las calles.
x_1, x_2, x_3, x_4, x_5, x_6, x_7 designan el flujo de tránsito entre intersecciones adyacentes.

Como la calle que está en reparación es la que vá de B a A, el problema es encontrar entonces el valor mínimo de x_4, que estará también sujeto a ciertas restricciones, dadas por el mismo problema.

El argumento fundamental a utilizarse es que, como las calles son de un solo sentido, entonces el número de carros que llegue a una intersección debe ser igual al número de carros que sale.

De acuerdo con ésto :
Por ejemplo en A, entran $x_4 + 200$ vehículos y salen $x_1 + 300$, de ahí que

$$x_1 + 300 = x_4 + 200.$$

Entonces

$$x_1 - x_4 = -100$$

y obtenemos una primera ecuación.

Análogamente, para cada una de las intersecciones :
En B, $x_2 + x_5 = 50 + x_4$, entonces $x_2 - x_4 + x_5 = 50$.

En C, $150 + 200 = x_3 + x_5$, entonces $x_3 + x_5 = 350$.

En D, $400 + x_1 = x_6 + 450$, entonces $x_1 - x_6 = 50$.

En E, $x_6 + 300 = x_2 + x_7$, entonces $x_2 - x_6 + x_7 = 300$.

En F, $x_3 + x_7 = 350 + 100$, entonces $x_3 + x_7 = 450$.

Resumiendo, el sistema por resolver es :

$$\begin{aligned}
x_1 \qquad\quad - x_4 \qquad\qquad\quad &= -100 \\
x_2 - x_4 + x_5 \qquad\quad &= 50 \\
x_3 \quad + x_5 \qquad\quad &= 350 \\
x_1 \qquad\quad - x_6 \qquad\quad &= 50 \\
x_2 \qquad\quad - x_6 + x_7 &= 300. \\
x_3 \qquad\qquad + x_7 &= 450.
\end{aligned}$$

Se trata de un sistema de 6 ecuaciones con 7 incógnitas cuya resolución, tal como se planteará más adelante, resultará bastante ilustrativa e interesante.

para el uso de las leyes mencionadas (Ohm - Kirchhoff) procedemos como sigue:

1) Divida el circuito en tantas mallas como sea posible.
2) [illegible] una dirección [illegible] cada malla.
3) [illegible] Ohm [illegible] tancia d resulta negativa.
3. Utilice la Primera Ley de Kirchhoff [illegible] ciones [illegible] puntos de unión (nodos) [illegible]
4) Utilice la Ley de Ohm para expresar [illegible] je en cada resistencia.
5) Use la Segunda Ley de Kirchhoff.
6) Resuelva el SEL que resulte.

En el problema dado en la sección anterior se tiene:

Malla AB: V_1 [illegible] Ley de
V_2 [illegible] Ohm.

$i_1 + 5 i_2 = 20$ Segunda Ley de Kirchhoff

Malla BDC: $V_2 = 5 i_2$
$V_3 = 4 i_3$

$5 i_2 + 4 i_3 = 30$

Malla ACDB: $V_1 = -6 i_1$ pues i_1 tiene sentido contrario al de la corriente I en esa malla.
$V_2 = 4 i_3$

$-6 i_1 + 4 i_3 = 10$ (después de algebra [illegible] el enunciado).

Nodo B: $i_1 + i_3 = i_2$ Primera Ley de Kirchhoff } ecuaciones iguales
Nodo C: $i_1 + i_3 = i_2$
Entonces: $i_1 - i_2 + i_3 = 0$

Por lo tanto (ordenando un poco las ecuaciones), el sistema a resolver es:

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ i_1 + 5 i_2 &= 20 \\ 5 i_2 + 4 i_3 &= 30 \\ -6 i_1 + 4 i_3 &= 10 \end{aligned}$$

Problema 3.- En el Modelo de Leontief se han designado como I_1, I_2 e I_3 los ingresos de la agricultura, ganadería y pesca, respectivamente. Se supone también que Ingreso por Ventas = Gasto por consumo. Entonces :

$$\begin{aligned} I_1 &= 3/6\ I_1 + 4/7\ I_2 + 2/5\ I_3 \\ I_2 &= 2/6\ I_1 + 3/14\ I_2 + 1/5\ I_3 \\ I_3 &= 1/6\ I_1 + 3/14\ I_2 + 2/5\ I_3 \end{aligned}$$

que en su forma homogénea es :

$$\begin{aligned} 3/6\ I_1 - 4/7\ I_2 - 2/5\ I_3 &= 0 \\ -1/3\ I_1 + 11/14\ I_2 - 1/5\ I_3 &= 0 \\ -1/6\ I_1 - 3/14\ I_2 + 3/5\ I_3 &= 0 \end{aligned}$$

donde I_1, I_2, I_3 son las incógnitas del sistema.

Dado que el número de ecuaciones es igual al número de incógnitas, tenemos un sistema homogéneo cuadrado.

Problema 4.- Este tipo de problemas (balanceo de ecuaciones químicas) pueden modelarse de tal forma que de la única ecuación química a balancear se obtenga un SEL. Esto podemos hacerlo de la siguiente forma : como para que una ecuación química esté balanceada es necesario que el número de átomos --- (moléculas) existentes en el lado izquierdo de la ecuación aparezca también en el derecho, entonces a cada sustancia --- (tanto a las que reaccionan como a las que se forman en la -- reacción) le asociamos una incógnita que nos represente dicha cantidad de cada una de ellas. Así, podemos escribir

$$x_1\ NaOH + x_2\ H_2SO_4 \longrightarrow x_3\ Na_2SO_4 + x_4\ H_2O$$

Después procedemos a separar (desglosar) la ecuación --- completa en ecuaciones para cada uno de los elementos que intervienen en la reacción. De aquí se obtiene un sistema de ecuaciones a resolver que es :

Sodio (Na) $\quad x_1 = 2x_3$, entonces $\quad x_1 - 2x_3 = 0$

Oxígeno (O) $\quad x_1 + 4x_2 = 4x_3 + x_4$, entonces

$$x_1 + 4x_2 - 4x_3 - x_4 = 0.$$

Hidrógeno (H) $\quad x_1 + 2x_2 = 2x_4$, entonces $\quad x_1 + 2x_2 - 2x_4 = 0$

Azufre (S) $\quad x_2 = x_3$, entonces $\quad x_2 - x_3 = 0.$

2.5 METODOS DE SOLUCION DE SEL: Método de Gauss, Operaciones Elementales, Ejemplos.

En la actualidad existen varios métodos para la solución numérica de SEL que todavía se siguen perfeccionando. Estos métodos podemos separarlos en dos categorías : Métodos Exactos y Métodos Iterativos. Los métodos exactos son aquéllos que dan una solución del problema usando un número finito de operaciones aritméticas elementales. Si tanto los datos como las cálculos hechos son exactos, la solución también es exacta. En este tipo de métodos el número de operaciones de cálculo necesarias depende sólo del tipo de esquema utilizado y del tamaño del sistema asociado al problema dado.

El primer método de este tipo se basa en la idea de la eliminación de incógnitas ; recibe el nombre de Método de Eliminación de Gauss (*) y consiste en una serie de eliminaciones sucesivas por medio de las cuales el sistema original se reduce a una forma lo suficientemente sencilla como para poder resolverse a simple vista. Actualmente existen muchos esquemas para aplicación de este método ; entre ellos están los esquemas compactos que requieren un mínimo de notación.

Hay otros métodos, como el del límite y el de inversión de matrices, aunque para poder entenderlos es necesario primero introducirnos en la teoría de matrices. Unicamente como comentario mencionaremos que, por ejemplo el método del límite considera una matriz dada como un conjunto de matrices encajadas sucesivamente, iniciando con una matriz de orden uno. Ciertas variaciones de este método se fundamentan en operaciones con matrices.

Los métodos iterativos proporcionan un medio para determinar soluciones aproximadas de un SEL, y además, un límite para las aproximaciones sucesivas obtenidas mediante algún proceso.

En estos métodos son esenciales tanto la convergencia de estas aproximaciones sucesivas como también la velocidad de convergencia. Es importante también una preparación preliminar del sistema para sustituir el sistema original por uno equivalente que converja tan rápidamente como sea posible, por lo que también son importantes los métodos de aceleración.

Los métodos exactos se utilizan cuando el número de ecuaciones no es muy grande (del orden de 40 o menos).

Los métodos iterativos son más apropiados cuando aparece un gran número de ecuaciones simultáneas (100 o más) las cuales a menudo poseen otras características especiales.

(*) Llamado así en honor al célebre matemático alemán Karl Friedrich Gauss (1777-1855)

A continuación describiremos los métodos más simples - - - pertenecientes a la categoría de métodos exactos. Iniciaremos con la siguiente definición :

Definición 2.5.1.- Dados dos sistemas de ecuaciones lineales, éstos se dicen equivalentes si cualquier solución del primer sistema es también solución del segundo, y, recíprocamente, cualquier solución del segundo sistema lo es también - del primero.

Por ejemplo, el sistema

$$\begin{aligned} 5x - 3y &= -1 \\ x + 2y &= 5 \end{aligned}$$

es equivalente al sistema

$$\begin{aligned} 7x + y &= 9 \\ -3x - 6y &= -15 \end{aligned}$$

En ambos casos la solución es el punto (1,2).

El cambio de orden con que se presentan las ecuaciones en un sistema no influye en la determinación de las soluciones en caso de que existan. Así, dos sistemas que difieren apenas en cuanto al orden en que se disponen sus ecuaciones son equivalentes. Existen dos formas particularmente útiles de transformar un SEL en un sistema equivalente:

1.- Dos sistemas son equivalentes si uno de ellos difiere del otro solamente en el hecho de que una ecuación es un múltiplo escalar no nulo de una de las ecuaciones del otro - sistema, siendo idénticas las ecuaciones restantes en los dos sistemas. Ejemplo:

$$\begin{aligned} 3x - 4y &= -6 \\ x + 2y &= 8 \end{aligned} \qquad \text{y} \qquad \begin{aligned} 3x - 4y &= -6 \\ 3x + 6y &= 24 \end{aligned}$$

son equivalentes.

2.- Dos SEL son equivalentes si uno difiere del otro sólo por el hecho de que una de las ecuaciones del primer sistema se sustituye en el segundo sistema por la suma de esa e - cuación con otra ecuación también perteneciente al primer sistema, siendo todas las demás ecuaciones las mismas en ambos sistemas. Ejemplo:

$$\begin{aligned} 2x - 3y &= -5 \\ 3x + y &= 9 \end{aligned} \qquad \text{es equivalente al sistema} \qquad \begin{aligned} 2x - 3y &= -5 \\ 5x - 2y &= 4 \end{aligned}$$

Usando repetidamente estos dos hechos podemos obtener -- SEL mucho más simples equivalentes a los sistemas considera-- dos originalmente ; ésto se verá en detalle más adelante.

M E T Ó D O D E G A U S S .

Lo introduciremos mediante un ejemplo. Consideremos el problema de calcular el valor de la corriente en cada resis-- tencia (es decir, i_1, i_2 e i_3) en el circuito dado anterior-- mente .

Como ya lo hemos mencionado, esto se reduce a resolver el sistema

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 6i_1 + 5i_2 &= 20 \\ 5i_2 + 4i_3 &= 30 \\ -6i_1 + 4i_3 &= 10 \end{aligned}$$

y llegar al final a encontrar dichos valores que satisfacen todas las ecuaciones simultáneamente. Para esto utilizaremos el Método de Reducción de Gauss (o de eliminación), cuyos rudimentos se remontan a la época de los babilonios. El método consiste en eliminar una de las incógnitas, combinando las -- ecuaciones a través de operaciones de los dos tipos anterio-- res. De esto resulta un sistema más simple y equivalente al original.

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 6i_1 + 5i_2 &= 20 \\ 5i_2 + 4i_3 &= 30 \\ -6i_1 + 4i_3 &= 10. \end{aligned}$$

Multiplicamos la segunda ecuación por (-1) y la sumamos a la tercera :

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 6i_1 + 5i_2 &= 20 \\ -6i_1 + 4i_3 &= 10 \\ -6i_1 + 4i_3 &= 10 \end{aligned}$$

Multiplicamos la tercera ecuación por (-1) y la sumamos a la cuarta :

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 6i_1 + 5i_2 &= 20 \\ -6i_1 + 4i_3 &= 10 \\ 0 &= 0. \end{aligned}$$

Multiplicamos por 5 la primera ecuación y la sumamos a la segunda :

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 11i_1 + 5i_3 &= 20 \\ -6i_1 + 4i_3 &= 10 \\ 0 &= 0. \end{aligned}$$

Multiplicamos por (-1) la tercera ecuación y la sumamos a la segunda :

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 17i_1 + i_3 &= 10 \\ -6i_1 + 4i_3 &= 10 \\ 0 &= 0. \end{aligned}$$

Finalmente, multiplicamos por (-4) la segunda ecuación y la - sumamos a la tercera :

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 17i_1 + i_3 &= 10 \\ -74i_1 &= -30 \\ 0 &= 0. \end{aligned}$$

Despejando i_1 de la tercera ecuación obtenemos :

$$i_1 = 30/74$$

Sustituyendo el valor de i_1 en la segunda ecuación y despe--- jando i_3 :

$$\begin{aligned} i_3 &= 10 - 17i_1 \\ &= 10 - 17 \ 30/74 \\ i_3 &= 230/74 \end{aligned}$$

Despejando i_2 en la primera ecuación y sustituyendo i_1 e i_3 :

$$\begin{aligned} i_2 &= i_1 + i_3 = 30/74 + 230/74 \\ &= 260/74 \end{aligned}$$

De donde, la solución es

$$i_1 = 30/74 \text{ amperios}$$

$$i_2 = 260/74 \text{ amperios}$$

$$i_3 = 230/74 \text{ amperios.}$$

Observamos que dicha solución es única. Como a su vez - podemos retornar al sistema original por operaciones combinadas de los dos mismos tipos anteriores, se tiene que 30/74 , 260/74, 230/74 es una solución para él.

Antes de pasar a resolver otros problemas, introduciremos una notación que haga más fácil escribir cada paso en nuestro procedimiento. Si mentalmente se lleva un registro de la colocación de los signos (+) o (-) de las incógnitas y de los - signos (=) de cada ecuación, un sistema de m ecuaciones lineales con n incógnitas se puede abreviar escribiendo únicamente el arreglo rectangular de números

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{array}\right]$$

Este arreglo se conoce como la Matriz Aumentada del Sistema. (El término matriz se emplea en Matemáticas para denotar un arreglo rectangular de números. Posteriormente se abordará su estudio más detalladamente).

En cuanto a la matriz

$$\left[\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right]$$

es la Matriz de los Coeficientes del Sistema. (Obsérvese que sus elementos componentes son exactamente los coeficientes del sistema general de m ecuaciones con n incógnitas).

O P E R A C I O N E S E L E M E N T A L E S .

El método básico para resolver un SEL consiste en reemplazar el sistema dado por un sistema nuevo que tenga el mismo conjunto solución, pero que sea más facil de resolver. Por lo general, este nuevo sistema se obtiene en una serie de etapas, aplicando los siguientes tres tipos de operaciones elementales :

1.- Intercambio de dos renglones R_i y R_j , indicada por R_{ij} .
2.- Multiplicación del i-ésimo renglón R_i por un escalar no nulo k, indicada por kR_i .
3.- Sustitución de un renglón R_i por su suma con un múltiplo escalar de otro renglón R_j , indicada por $kR_j + R_i$.

Estas operaciones o transformaciones elementales corresponden a las siguientes operaciones en el sistema asociado de ecuaciones :

1.- Cambio de orden en las ecuaciones i y j del sistema.
2.- Multiplicacion de cada término de la ecuación i por la misma constante no nula.
3.- Sustitución de la i-ésima ecuación por su suma con un múltiplo escalar de otra ecuación (la j-ésima).

El proceso de efectuar estas operaciones elementales de renglón para simplificar una matriz aumentada se llama Reduc-

ción por Renglones.

Como ya vimos, la aplicación de estas operaciones produce un nuevo SEL que es equivalente al sistema original, en el sentido de que ambos sistemas tienen la(s) misma(s) solución-(es).

A continuación veremos ejemplos que ilustran la forma en que se pueden emplear estas operaciones para resolver SEL.

Retomaremos el problema del circuito eléctrico. Apli---cando ahora estos tres tipos de operaciones encontraremos la solución del problema (trabajando simultáneamente el sistema y la matriz asociada a él).

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 6i_1 + 5i_2 &= 20 \\ 5i_2 + 4i_3 &= 30 \\ -6i_1 + 4i_3 &= 10 \end{aligned} \qquad \left[\begin{array}{cccc} 1 & -1 & 1 & 0 \\ 6 & 5 & 0 & 20 \\ 0 & 5 & 4 & 30 \\ -6 & 0 & 4 & 10 \end{array}\right]$$

Multiplicamos por (-6) la primera ecuación y la sumamos a la segunda :

$$\downarrow \; -6R_1 + R_2$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 11i_2 - 6i_3 &= 20 \\ 5i_2 + 4i_3 &= 30 \\ -6i_1 + 4i_3 &= 10 \end{aligned} \qquad \left[\begin{array}{cccc} 1 & -1 & 1 & 0 \\ 0 & 11 & -6 & 20 \\ 0 & 5 & 4 & 30 \\ -6 & 0 & 4 & 10 \end{array}\right]$$

Multiplicamos por 6 la primera ecuación y la sumamos a la cuarta:

$$\downarrow \; 6R_1 + R_4$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 11i_2 - 6i_3 &= 20 \\ 5i_2 + 4i_3 &= 30 \\ -6i_2 + 10i_3 &= 10 \end{aligned} \qquad \left[\begin{array}{cccc} 1 & -1 & 1 & 0 \\ 0 & 11 & -6 & 20 \\ 0 & 5 & 4 & 30 \\ 0 & -6 & 10 & 10 \end{array}\right]$$

Intercambiamos las ecuaciones segunda y tercera :

$$\downarrow \; R_{23}$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ 5i_2 + 4i_3 &= 30 \\ 11i_2 - 6i_3 &= 20 \\ -6i_2 + 10i_3 &= 10 \end{aligned} \qquad \left[\begin{array}{cccc} 1 & -1 & 1 & 0 \\ 0 & 5 & 4 & 30 \\ 0 & 11 & -6 & 20 \\ 0 & -6 & 10 & 10 \end{array}\right]$$

Multiplicamos por 1/5 la segunda ecuación :

$$\downarrow \; 1/5\, R_2$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ i_2 + 4/5\, i_3 &= 6 \\ 11 i_2 - 6\, i_3 &= 20 \\ -6 i_2 + 10\, i_3 &= 10 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & 4/5 & 6 \\ 0 & 11 & -6 & 20 \\ 0 & -6 & 10 & 10 \end{array}\right]$$

Multiplicamos por (-11) la segunda ecuación y la sumamos a la tercera :

$$\downarrow \; -11\, R_2 + R_3$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ i_2 + 4/5\, i_3 &= 6 \\ -74/5\, i_3 &= -46 \\ -6 i_2 + 10\, i_3 &= 10 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & 4/5 & 6 \\ 0 & 0 & -74/5 & -46 \\ 0 & -6 & 10 & 10 \end{array}\right]$$

Multiplicamos por 6 la segunda ecuación y la sumamos a la cuarta :

$$\downarrow \; 6\, R_2 + R_4$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ i_2 + 4/5\, i_3 &= 6 \\ -74/5\, i_3 &= -46 \\ 74/5\, i_3 &= 46 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & 4/5 & 6 \\ 0 & 0 & -74/5 & -46 \\ 0 & 0 & 74/5 & 46 \end{array}\right]$$

Intercambiamos tercera y cuarta ecuaciones :

$$\downarrow \; R_{3\,4}$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ i_2 + 4/5\, i_3 &= 6 \\ 74/5\, i_3 &= 46 \\ -74/5\, i_3 &= -46 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & 4/5 & 6 \\ 0 & 0 & 74/5 & 46 \\ 0 & 0 & -74/5 & -46 \end{array}\right]$$

Sumamos tercera y cuarta ecuaciones :

$$\downarrow \; R_3 + R_4$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ i_2 + 4/5\, i_3 &= 6 \\ 74/5\, i_3 &= 46 \\ 0 &= 0 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & 4/5 & 6 \\ 0 & 0 & 74/5 & 46 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Multiplicamos por (5/74) la tercera ecuación :

$$\downarrow \; 5/74\, R_3$$

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ i_2 + 4/5\, i_3 &= 6 \\ i_3 &= 230/74 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & 4/5 & 6 \\ 0 & 0 & 1 & 230/74 \end{array}\right]$$

Nos devolvemos a sustituir valores en las ecuaciones anteriores y, mediante despejes, concluimos que

$$i_1 = 30/74$$
$$i_2 = 260/74$$
$$i_3 = 230/74$$

Observaciones :
- En cada paso las operaciones se refieren al sistema de ecuaciones equivalente obtenido en el paso inmediato anterior.
- Estas operaciones se realizan en secuencia y no simultáneamente.
- Con la práctica se puede realizar más de una operación en cada caso (sin olvidar la observación anterior).
- Si puede obtenerse una matriz D mediante una secuencia de operaciones elementales efectuadas sobre los renglones de una matriz A, se dice que A es equivalente por renglones a D.

El proceso de encontrar el valor de una primer incógnita se llama CURSO HACIA ADELANTE y el proceso de encontrar la solución completa se llama CURSO DE REGRESO.

Dentro de todos los esquemas computacionales diferentes correspondientes al Método de Gauss, éste corresponde al llamado ESQUEMA DE DIVISION SIMPLE.

Ejemplo.- Vamos ahora a resolver el sistema obtenido de el problema 1 de flujo de tráfico.

Habíamos llegado al sistema

$$\begin{aligned}
x_1 \qquad\quad - x_4 \qquad\qquad\qquad\quad &= -100 \\
x_2 \qquad - x_4 + x_5 \qquad\qquad &= 50 \\
x_3 \qquad\quad + x_5 \qquad\qquad &= 350 \\
x_1 \qquad\qquad\quad - x_6 \qquad &= 50 \\
x_2 \qquad\qquad\quad - x_6 + x_7 &= 300 \\
x_3 \qquad\qquad\qquad\quad + x_7 &= 450
\end{aligned}$$

Utilizando la notación matricial :

$$\left[\begin{array}{ccccccc|c}
1 & 0 & 0 & -1 & 0 & 0 & 0 & -100 \\
0 & 1 & 0 & -1 & 1 & 0 & 0 & 50 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 350 \\
1 & 0 & 0 & 0 & 0 & -1 & 0 & 50 \\
0 & 1 & 0 & 0 & 0 & -1 & 1 & 300 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 450
\end{array}\right]
\xrightarrow{-R_1 + R_4}
\left[\begin{array}{ccccccc|c}
1 & 0 & 0 & -1 & 0 & 0 & 0 & -100 \\
0 & 1 & 0 & -1 & 1 & 0 & 0 & 50 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 350 \\
0 & 0 & 0 & 1 & 0 & -1 & 0 & 150 \\
0 & 1 & 0 & 0 & 0 & -1 & 1 & 300 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 450
\end{array}\right]$$

$$\xrightarrow{-R_2+R_5}\begin{bmatrix}1&0&0&-1&0&0&0&|&-100\\0&1&0&-1&1&0&0&|&50\\0&0&1&0&1&0&0&|&350\\0&0&0&1&0&-1&0&|&150\\0&0&0&1&-1&-1&1&|&250\\0&0&1&0&0&0&1&|&450\end{bmatrix}\xrightarrow{-R_3+R_6}\begin{bmatrix}1&0&0&-1&0&0&0&|&-100\\0&1&0&-1&1&0&0&|&50\\0&0&1&0&0&0&0&|&350\\0&0&0&1&0&-1&0&|&150\\0&0&0&1&-1&-1&1&|&250\\0&0&0&0&-1&0&1&|&100\end{bmatrix}$$

$$\xrightarrow{-R_4+R_5}\begin{bmatrix}1&0&0&-1&0&-1&0&|&-100\\0&1&0&-1&1&0&0&|&50\\0&0&1&0&1&0&0&|&350\\0&0&0&1&0&-1&0&|&150\\0&0&0&0&-1&0&1&|&100\\0&0&0&0&-1&0&1&|&100\end{bmatrix}\xrightarrow[R_5+R_6]{-R_5}\begin{bmatrix}1&0&0&-1&0&0&0&|&-100\\0&1&0&-1&1&0&0&|&50\\0&0&1&0&1&0&0&|&350\\0&0&0&1&0&-1&0&|&150\\0&0&0&0&1&0&-1&|&-100\\0&0&0&0&0&0&0&|&0\end{bmatrix}$$

Como el último renglón está formado por ceros, entonces podemos eliminarlo.

Este arreglo corresponde al sistema :

$$\begin{aligned}x_1 \qquad\quad - x_4 \qquad\quad &= -100\\ x_2 \quad - x_4 + x_5 \quad &= 50\\ x_3 + x_5 \quad &= 350\\ x_4 \quad - x_6 \quad &= 150\\ x_5 \quad - x_7 \quad &= -100\end{aligned}$$

Despejando en todas las ecuaciones obtenemos :

$$\begin{aligned}x_1 &= x_4 - 100\\ x_2 &= x_4 - x_5 + 50\\ x_3 &= - x_5 + 350\\ x_4 &= x_6 + 150\\ x_5 &= x_7 - 100.\end{aligned}$$

Sustituyendo x_4 y x_5 en las expresiones para x_1, x_2 y x_3 llegamos a :

$$\begin{aligned}x_1 &= x_6 + 50\\ x_2 &= x_6 - x_7 + 300\\ x_3 &= - x_7 + 450\\ x_4 &= x_6 + 150\\ x_5 &= x_7 - 100.\end{aligned}$$

Recordemos que el problema es minimizar x_4 .

Como x_6 no puede ser negativo ($x_6 \geq 0$, pues las calles son en sentido único), entonces $x_4 \geq 150$.

Por otro lado, si $x_4 = 150$ entonces $x_6 = 0$, por lo que, para evitar riesgos, deberá cerrarse la circulación en la calle que vá de D a E.

De las ecuaciones para x_2 , x_3 y x_5 tenemos :

$$x_2 = - x_7 + 300$$
$$x_3 = - x_7 + 450$$
$$x_5 = x_7 - 100.$$

De aquí :

$$x_7 \leq 300$$
$$x_7 \leq 450$$
$$x_7 \geq 100$$

con lo cual $100 \leq x_7 \leq 300$.

Por lo tanto

$$0 \leq 300 - x_7 \leq 200$$
$$150 \leq 450 - x_7 \leq 350 \quad ,$$
$$0 \leq x_7 - 100 \leq 200$$

ésto es

$$0 \leq x_2 \leq 200$$
$$150 \leq x_3 \leq 350$$
$$0 \leq x_5 \leq 200 \quad .$$

Resumiendo, la solución del problema es

$$x_1 = 50$$
$$0 \leq x_2 \leq 200$$
$$150 \leq x_3 \leq 350$$
$$x_4 = 150$$
$$0 \leq x_5 \leq 200$$
$$x_6 = 0$$
$$100 \leq x_7 \leq 300 \quad .$$

Este método y algunas modificaciones a él que funcionan en base a las operaciones elementales enunciadas anteriormente se justifican por el siguiente teorema :

Teorema 2.5.2.- Si un SEL se transforma en otro SEL a través de operaciones elementales, los sistemas son equiva--- lentes.

Demostración.- Sea un sistema arbitrario de ecuaciones lineales

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n = b_1$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{i1} x_1 + a_{i2} x_2 + \ldots + a_{in} x_n = b_i$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{j1} x + a_{j2} x_2 + \ldots + a_{jn} x_n = b_j$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n = b_m \, ,$$

en donde $m > 1$. El conjunto de valores

$$x_1 = c_1, \quad x_2 = c_2 \, , \ldots \, , x_n = c_n$$

es una solución del sistema si y sólo si sustituyendo estos valores en el sistema se obtienen m identidades. Debemos --- mostrar que la existencia de esas m identidades no se altera por las tres operaciones elementales :

la existencia de las m identidades cuando sustituimos los valores de $x_1, x_2, \ldots, x_n$ en el SEL.

2.- La multiplicación de una ecuación del SEL por un escalar k no nulo no altera la existencia de las m identidades cuando sustituimos $x_1, x_2, \ldots, x_n$ en el SEL, pues

$$k(a_{i1}c_1 + a_{i2}c_2 + \ldots + a_{in}c_n) = kb_i \quad , \quad k \neq 0$$

si y sólo si

$$a_{i1}c_1 + a_{i2}c_2 + \ldots + a_{in}c_n = b_i \ .$$

3.- Consideremos la i-ésima y la j-ésima ecuaciones del SEL, $(i \neq j)$. Sumemos k veces los términos de la j-ésima ecuación a los términos correspondintes de la i-ésima ecuación. El sistema resultante puede expresarse como

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n + k(a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n) = b_i + kb_j \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n = b_j \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m
\end{aligned}$$

La i-ésima ecuación fue la única ecuación modificada en el nuevo sistema. Luego, sustituyendo los valores de $x_1, x_2, \ldots, x_n$ en este nuevo sistema, tenemos que

$$a_{i1}c_1 + a_{i2}c_2 + \ldots + a_{in}c_n + k(a_{j1}c_1 + a_{j2}c_2 + \ldots + a_{jn}c_n) = b_i + kb_j$$

y

$$k(a_{j1}c_1 + a_{j2}c_2 + \ldots + a_{jn}c_n) = kb_j$$

si y solamente si

$$a_{i1}c_1 + a_{i2}c_2 + \ldots + a_{in}c_n = b_i \qquad \text{y}$$
$$a_{j1}c_1 + a_{j2}c_2 + \ldots + a_{jn}c_n = b_j \qquad \text{L.Q.Q.D.}$$

El proceso de reducción de Gauss aplicado al sistema general de m ecuaciones con n incógnitas se puede describir como sigue :

Sea el sistema

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = f_1 \\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = f_2 \qquad (1) \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = f_m
\end{aligned}$$

Suponemos $a_{11} \neq 0$. Dividimos los coeficientes de la primera ecuación por el coeficiente a_{11}, al que llamaremos el elemento principal (en el primer paso), y escribimos

$$b_{1j} = \frac{a_{1j}}{a_{11}} \quad (j > 1) \quad , \quad g_1 = \frac{f_1}{a_{11}} \ . \qquad (2)$$

Con esto, obtenemos la nueva ecuación

$$x_1 + b_{12}x_2 + \ldots + b_{1n}x_n = g_1 \qquad (3)$$

Ahora, eliminamos la incógnita x_1 de todas las ecuaciones restantes del sistema (1), comenzando con la segunda, restando de cada una de ellas el producto de la ecuación (3)

ciones transformadas serán :

$$\begin{aligned} a_{22\cdot1}\, x_2 + \ldots + a_{2n\cdot1}\, x_n &= f_{2\cdot1} \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{m2\cdot1}\, x_2 + \ldots + a_{mn\cdot1}\, x_n &= f_{m\cdot1} \end{aligned} \qquad (4)$$

donde

$$a_{ij\cdot1} = a_{ij} - a_{i1}\, b_{1j} \quad (i, j > 2), \qquad (5)$$

$$f_{i1} = f_i - a_{i1}\, g_1 .$$

Ahora dividimos los coeficientes de la primera de las ecuaciones transformadas por el elemento principal del segundo paso $a_{22\cdot1}$, que suponemos que es diferente de cero. Obtenemos la ecuación $x_2 + b_{23} x_3 + \ldots + b_{2n} x_n = g_2$, donde

$$b_{2j} = \frac{a_{2j\cdot1}}{a_{22\cdot1}} \quad , \quad g_2 = \frac{f_{2\cdot1}}{a_{22\cdot1}} \quad .$$

Eliminando las incógnitas x_2 de las ecuaciones (4), comenzando con la segunda, llegamos a las ecuaciones :

$$\begin{aligned} a_{33\cdot2}\, x_3 + \ldots + a_{3n\cdot2}\, x_n &= f_{3\cdot2} \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{m3\cdot2}\, x_3 + \ldots + a_{mn\cdot2}\, x_n &= f_{m\cdot2} \end{aligned}$$

donde

$$a_{ij\cdot2} = a_{ij\cdot1} - a_{i2\cdot1}\, b_{2j} \quad (i,j > 3) ,$$

$$f_{i\cdot2} = f_{i\cdot1} - a_{i2\cdot1}\, g_2 .$$

Continuamos el proceso de esta forma. En el r-ésimo paso obtenemos las ecuaciones :

$$\begin{aligned} x_r + b_{r,r+1}\, x_{r+1} + \ldots + b_{rn}\, x_n &= g_r \\ a_{r+1,r+1\cdot r}\, x_{r+1} + \ldots + a_{r+1,n\cdot r}\, x_n &= f_{r+1\cdot r} \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{mr+1\cdot r}\, x_{r+1} + \ldots + a_{mn\cdot r}\, x_n &= f_{m\cdot r} \end{aligned} \qquad (6)$$

donde

$$b_{rj} = \frac{a_{rj\cdot r-1}}{a_{rr\cdot r-1}} \quad (j > r+1) \quad , \quad g_r = \frac{f_{r,r-1}}{a_{rr\cdot r-1}} \quad ,$$

$$a_{ij\cdot r} = a_{ij\cdot r-1} - a_{ir\cdot r-1}\, b_{rj} \quad , \quad f_{ir} = f_{i\cdot r-1} - a_{ir\cdot r-1}\, g_r \qquad (7)$$

Combinando todas las primeras ecuaciones de cada paso obtenemos el sistema

$$\begin{aligned} x_1 + b_{12} x_2 \qquad\qquad + \ldots + b_{1n} x_n &= g_1 \\ x_2 + b_{23} x_3 \qquad\qquad + \ldots + b_{2n} x_n &= g_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ x_k + b_{k,k+1}\, x_{k+1} + \ldots + b_{kn} x_n &= g_k \end{aligned} \qquad (8)$$

Señalemos que $k \leq m$, puesto que, posiblemente, alguna(s) de las ecuaciones se hayan eliminado y además, evidentemente, $k \leq n$.

Si $k = n$, el sistema es de la forma

$$\begin{aligned} x_1 + b_{12} x_2 + b_{13} x_3 + \ldots + b_{1n} x_n &= g_1 \\ x_2 + b_{23} x_3 + \ldots + b_{2n} x_n &= g_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \end{aligned} \qquad (8')$$

En cualquiera de los dos casos, el sistema al que se --- llega al final es equivalente al original.

En (8'), de la última ecuación se obtiene un valor perfectamente determinado para la incógnita x_n. Determinamos los valores de las incógnitas restantes, de x_n a x_1, por la fórmula

$$x_m = g_m - b_{m,m+1}\ x_{m+1} - \ldots - b_{mn}\ x_n .$$

Por último llegamos a que el sistema final, y por lo --- tanto el original, poseen solución única.

En otro caso, si en (8') resulta que $x_n = 0$ y $g_n =$ constante $\neq 0$, se tiene que $0 =$ constante $\neq 0$, con lo que --- (8') no tiene solución, y por lo tanto, el sistema original- tampoco la tiene.

Si $k < n$ (sistema (8')), asignamos valores numéricos arbitrarios para las variables $x_{K+1}, \ldots, x_n$ y aplicamos después el proceso de regreso. Por tanto, como estos valores arbi -- trarios los podemos elegir de muchas maneras distintas, el -- sistema tendrá una infinidad de soluciones.

Debemos hacer notar que el proceso descrito es posible - sólo bajo la condición de que los elementos principales ----- $a_{rr,r-1}$ son diferentes de cero, ya que dividimos por ellos -- durante el proceso. Si alguno de estos elementos es cero podemos entonces intercambiar las ecuaciones, de manera que el nuevo $a_{rr,r-1}$ sea distinto de cero. Hemos de señalar que, -- bajo esta suposición, el SEL toma una forma "triangular"(8') o "trapezoidal"(8). En el caso general, el SEL a que llegaremos después de aplicar hasta el fin el proceso de eliminación de las incógnitas tomará una de estas dos formas sólo -- después de un cambio adecuado de la numeración de las incógnitas.

Así, para resolver el sistema dado por el esquema de división simple, primero construimos un sistema triangular auxiliar (o trapezoidal), y después procedemos a resolverlo.

La matriz aumentada correspondiente a este nuevo sistema es de la forma

$$\left[\begin{array}{ccccc|c} 1 & b_{12} & b_{13} & \ldots & b_{1n} & g_1 \\ 0 & 1 & b_{23} & \ldots & b_{2n} & g_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & g_n \end{array}\right]$$

y se dice que está en la forma escalonada ; las incógnitas x_i que no aparecen al principio de alguna ecuación ($i \neq 1$) se -- llaman variables libres.

Para resolver un sistema de este tipo, se aplica el siguiente teorema :

Teorema 2.5.3.- Hay dos casos para la solución del sistema en la forma escalonada :

i) $r = n$, ésto es, hay tantas ecuaciones como incógnitas. -- Entonces el sistema tiene una solución única.
ii) $r < n$, es decir, hay menos ecuaciones que incógnitas. - Entonces podemos asignar arbitrariamente valores a las $n - r$ variables libres y obtener una solución del sistema.

Demostración.- Consideremos el sistema en forma escalonada que sigue :

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n &= f_1 \\ a_{22\cdot1} x_2 + \ldots + a_{2n\cdot1} x_n &= f_{2\cdot1} \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{tr+1\cdot r} x_{r+1} + \ldots + a_{tn\cdot r} x_n &= f_{tn\cdot r}, \end{aligned}$$

donde $a_{11} \neq 0$, $a_{22\cdot1} \neq 0$, ... , $a_{tr+1\cdot r} \neq 0$. Nuestro sistema - contiene ahora t ecuaciones, $t < m$, puesto que, posiblemente, algunas de las ecuaciones hayan sido suprimidas. La demostración es por inducción sobre el número r de ecuaciones del sistema. Si $r = 1$, tenemos una sola ecuación lineal

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = f_1 ,$$

donde $a_1 \neq 0$. Las variables libres (o independientes) son x_2 , ... , x_n . Asignemos valores arbitrarios a las variables libres : $x_2 = k_2$, ... , $x_n = k_n$. Sustituyendo en la - ecuación obtenemos

$$x_1 = \frac{1}{a_1}(f - a_2 k_2 - \ldots - a_n k_n)$$

que es una solución de la ecuación. Así :

$$a_1 \left[\frac{1}{a_1} (f - a_2 k_2 - \ldots - a_n k_n)\right] + a_2 k_2 + \ldots + a_n k_n = f$$

Si $r = n = 1$, entonces se tiene una ecuación lineal de - la forma $ax = f$, donde $a \neq 0$. Entonces $x = \frac{f}{a}$ es solución. Si también $x = k$ es solución, es decir, si $ak = f$, entonces $k = \frac{f}{a}$. Por lo tanto, la solución es única.

Supongamos ahora $r > 1$ y que el teorema es válido para un sistema de $r-1$ ecuaciones. Consideremos las $r-1$ ecuaciones del sistema

$$\begin{aligned} a_{22\cdot1} x_2 + a_{23\cdot1} x_3 + \ldots + a_{2n\cdot1} x_n &= f_{2\cdot1} \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{rr\cdot r-1} x_r + \ldots + a_{rn\cdot r-1} x_n &= f_{r\cdot r-1} \end{aligned}$$

Obsérvese que este sistema está en la forma escalonada.

asignar valores arbitrarios a las (n-2+1) - (r-1) = n-r-2+2 variables libres en el sistema reducido para obtener una solución (por ejemplo, $x_2 = k_2$, $x_3 = k_3$, ... , $x_n = k_n$) de la primera ecuación con

$$x_1 = \frac{1}{a_{11}} \quad (f - a_{12}\, k_2 - \ldots - a_{1n}\, k_n) \quad .$$

De donde obtenemos una solución única del subsistema y, en consecuencia, una solución del sistema total. L.Q.Q.D.

La siguiente tabla presenta un ejemplo para el esquema de división simple aplicado en la solución de un sistema de 4 ecuaciones con 3 incógnitas. Específicamente, es la solución al sistema obtenido del circuito eléctrico.

Tabla 2.1.- Esquema de División Simple.

a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	1	-1	1	0	1
a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	6	5	0	20	31
a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	0	5	4	30	39
a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	-6	0	4	10	8
1	b_{12}	b_{13}	b_{14}	b_{15}	1	-1	1	0	1
	$a_{22\cdot1}$	$a_{23\cdot1}$	$a_{24\cdot1}$	$a_{25\cdot1}$		11	-6	20	25
	$a_{32\cdot1}$	$a_{33\cdot1}$	$a_{34\cdot1}$	$a_{35\cdot1}$		5	4	30	39
	$a_{42\cdot1}$	$a_{43\cdot1}$	$a_{44\cdot1}$	$a_{45\cdot1}$		-6	10	10	14
	1	b_{23}	b_{24}	b_{25}			-6/11	20/11	25/11
		$a_{33\cdot2}$	$a_{34\cdot2}$	$a_{35\cdot2}$			74/11	230/11	304/11
		$a_{43\cdot2}$	$a_{44\cdot2}$	$a_{45\cdot2}$			74/11	230/11	304/11
		1	b_{34}	b_{35}			1	230/74	304/74
			$a_{44\cdot3}$	$a_{45\cdot3}$				0	0
			0	0				0	0
		1	x_3	$\bar{x}_3$			1	230/74	304/74
	1		x_2	$\bar{x}_2$		1		260/74	334/74
1			x_1	$\bar{x}_1$	1			30/74	104/74

En la tabla, las primeras tres columnas corresponden a los coeficientes de cada incógnita del sistema, la cuarta columna representa a los términos constantes y la quinta se usa para control. Este control se basa en la siguiente idea : si en el sistema dado tomamos $\bar{x}_i = x_i + 1$, entonces para determinar $\bar{x}_i$ obtenemos un sistema con los mismos coeficientes anteriores pero con los términos constantes iguales a las sumas de los elementos en los renglones de la matriz aumentada. Así, si incluimos esta columna y realizamos en ella todas las operaciones efectuadas en las demás, entonces en ausencia de errores de cálculo deberemos obtener resultados iguales a las sumas de los elementos de los renglones recién construidos. Por lo tanto, al terminar el curso de regreso, los números $\bar{x}_i$ obtenidos deberán ser iguales a $x_i + 1$. Con esto, sólo compa-

con los resultados de las operaciones en la columna de con--- trol.

Por ejemplo, en la tabla :

$a_{22\cdot1} = a_{22} - a_{21}\, b_{12} = 5 - (6)(-1) = 11$

$a_{43\cdot1} = a_{43} - a_{41}\, b_{13} = 4 - (-6)(1) = 4 + 6 = 10$

$b_{23} = \dfrac{a_{23\cdot1}}{a_{22\cdot1}} = \dfrac{-6}{11}$

$a_{34\cdot2} = a_{34\cdot1} - a_{32\cdot1}\, b_{24} = 30 - (5)(20/11) = 230/11$

Al llegar al valor de x_3 (que nos representa al valor de la corriente i_3) aplicamos el proceso de regreso.

El número de operaciones necesarias para encontrar la solución de un sistema de n ecuaciones por este esquema es -- $n/3\ (n^2 + 3n - 1)$. Este número se reduce significativamente si la matriz original es casi triangular. Además, si $a_{ij} = a_{ji}$ en la matriz original, obviamente $a_{ij\cdot k} = a_{ji\cdot k}$ y puede re-- ducirse la notación.

Observación.- Este tipo de matriz en la cual $a_{ij} = a_{ji}$ recibe el nombre de Simetrica.

Si es necesario resolver varios sistemas que tengan una misma matriz de coeficientes dada, es natural buscar todas -- las soluciones simultáneamente, escribiendo fuera los térmi-- nos constantes en columnas vecinas. La columna control está formada por la suma de los elementos en los renglones de la matriz aumentada. Este esquema se representa en la siguiente tabla.

Tabla 2.2.- Esquema de División Simple para Varios Sistemas de Ecuaciones Lineales.

1	-1	1	0	0	1
6	5	0	20	25	56
0	5	4	30	40	79
-6	0	4	10	15	23
1	-1	1	0	0	1
	11	-6	20	25	50
	5	4	30	40	79
	-6	10	10	15	29
	1	-6/11	20/11	25/11	50/11
		74/11	230/11	315/11	619/11
		74/11	230/11	315/11	619/11
		1	230/74	315/74	619/74
			0	0	0
			0	0	0
		1	230/74	315/74	619/74
	1		260/74	170/37	674/74
1			30/74	25/74	129/74

En el esquema de división y sustracción todas las ecuaciones se dividen en cada paso por el coeficiente de la incógnita que vá a eliminarse y ésto se efectuá sustrayendo una ecuación del resto de ellas.

En el esquema de multiplicación y sustracción eliminamos la incógnita x_i de la i-ésima ecuación en el primer paso multiplicando esta ecuación por a_{11} y sustrayendo la primer ecuación multiplicada por a_{i1} .

El mismo procedimiento continúa en los siguientes pasos, de forma que los coeficientes de las ecuaciones auxiliares $\tilde{a}_{ij\cdot m}$ se calculan en el r-ésimo paso mediante las fórmulas

$$\begin{aligned} \tilde{a}_{ij\cdot r} &= a_{rr\cdot r-1}\ \tilde{a}_{ij\cdot r-1} - \tilde{a}_{ir\cdot r-1}\ \tilde{a}_{rj\cdot r-1} \\ \tilde{f}_{ir} &= \tilde{a}_{rr\cdot r-1}\ f_{ir-1} - \tilde{a}_{ir\cdot r-1}\ f_{r\cdot r-1} . \end{aligned}$$

En particular, es posible aplicar cada uno de los esquemas descritos, con el orden para eliminación de incógnitas descrito o escogiendo el orden de eliminación durante el proceso.

Un refinamiento del método de reducción de Gauss lo es el Método de Gauss-Jordan (*). Este método no efectúa el curso de regreso de Gauss para determinar la solución del sistema, sino que, en lugar de efectuar sustituciones sucesivas, hace operaciones elementales y elimina un mayor número de incógnitas en cada ecuación, de tal manera que la matriz asociada a este sistema la reduce todavía más en el valor de sus componentes ; es decir, en base a los elementos principales de cada ecuación elimina la incógnita correspondiente en todas las ecuaciones restantes del sistema y con esto llega a una matriz que, además de ser escalonada al igual que con Gauss, tiene todavía más componentes nulas. Este tipo de matrices las describe la siguiente

Definición 2.5.4.- Una matriz está en forma escalonada reducida si cumple las siguientes propiedades :
1) Todos los renglones que consisten únicamente de ceros (si existen) aparecen en la parte inferior de la matriz.
2) En cualquier renglón que no consista únicamente de ceros, el primer término distinto de cero (empezando por la izquierda) es uno.
3) El número de ceros al inicio de un renglón aumenta a medida que descendemos.
4) Todos los demás elementos de la columna donde aparece el primer 1 de un renglón, son ceros.

En base a esta definición vemos que para resolver un SEL por el Método de Gauss-Jordan simplemente tomamos la matriz -

(*) Carl Friedrich Gauss (1777-1855) y Camille Jordan (1838-1922)

del sistema (la que más convenga, dependiendo del tipo de --- sistema : homogéneo o no homogéneo) y realizamos las opera--- ciones elementales necesarias para obtener una matriz de la - forma escalonada reducioda. Así, el conjunto solución del -- sistema puede obtenerse a simple vista o, en el peor de los - casos, después de unas cuantas etapas más.

Observaciones :
- La diferencia entre la forma escalonada y escalonada redu-- cida de matrices es clara. En la forma escalonada todos los números que están abajo del primer 1 son ceros. En la forma escalonada reducida todos los números que están arriba y aba- jo del primer 1 de un renglón son ceros. Así, la forma esca- lonada reducida es más exclusiva, ésto es : cualquier matriz en forma escalonada reducida está en forma escalonada, pero - no inversamente.
- Siempre podemos reducir una matriz a la forma escalonada -- reducida o a la forma escalonada realizando operaciones ele-- mentales de renglón. Esto se ha visto en los ejemplos.

Como se observa, existe una fuerte conexión entre la --- forma escalonada reducida de una matriz y la existencia de -- una única solución al sistema. Como la forma escalonada re-- ducida de la matriz de coeficientes tiene un 1 en cada ren--- glón, la solución existe y es única. En otras ocasiones, la forma escalonada reducida de la matriz de coeficientes tiene un renglón de ceros y el sistema no tiene solución o tiene un número infinito de soluciones. Esto siempre es cierto en --- cualquier sistema con el mismo número de ecuaciones que de -- incógnitas.

Tenemos, por tanto, dos métodos para resolver nuestro -- sistema de ecuaciones:

1.- Reducir la matriz de coeficientes a la forma escalonada reducida (Eliminación de Gauss-Jordan).
2.- Reducir la matriz de coeficientes a la forma escalonada, resolver para la última incógnita y luego aplicar el curso de regreso (sustitucion) para resolver para las otras incógnitas (Eliminación de Gauss).

¿Cuál de estos dos métodos es mas útil?. Si se resuel-- ven sistemas de ecuaciones en una computadora, la eliminación de Gauss es el método más usado, ya que requiere menos opera- ciones elementales de reducción por renglón. Por otro lado, hay ocasiones en las que es necesario obtener la forma esca-- lonada reducida de una matriz. Estos procesos se efectúan -- hasta que ocurra una de las tres situaciones siguientes :

i) La última ecuación diferente de cero queda $x_n = c$, para -- alguna constante c. Entonces hay una solución única o hay un número infinito de soluciones para el sistema.
ii) La última ecuación diferente de cero queda

para alguna constante c donde al menos dos de las a's son diferentes de cero. Esto es, la última es una ecuación lineal en dos o mas variables. Entonces existe un número infinito de soluciones.
iii) La última ecuación queda $0 = c$, con $c \neq 0$. Entonces no existe solución.

Como se verá después de resolver varios problemas, los cálculos se pueden convertir en algo muy complicado. Es una buena medida usar una calculadora cuando las fracciones se vuelven muy complicadas ; sin embargo, es necesario hacer notar que si los cálculos se efectúan en calculadora o computadora, se pueden introducir errores por "redondeo".

2.6 SISTEMAS HOMOGENEOS DE ECUACIONES LINEALES.

Como ya se ha señalado, todo SEL tiene tres posibilidades de solución, aunque en ocasiones lo que interesa es saber el número de soluciones del sistema. En lo que respecta a SEL homogéneos podemos afirmar una de dos cosas :
i) El sistema tiene solamente la solución trivial.
ii) El sistema tiene un número infinito de soluciones no triviales, ademas de la trivial.

Un caso especial en el que es posible asegurar la segunda afirmacion es cuando el sistema tiene más incógnitas que ecuaciones. Para ver ésto, analizaremos el siguiente ejemplo.

Refirámonos ahora al sistema del Modelo de Leontief.

$$\begin{aligned} 3/6\ I_1 - 4/7\ I_2 - 2/5\ I_3 &= 0 \\ -1/3\ I_1 + 11/14\ I_2 - 1/5\ I_3 &= 0 \\ -1/6\ I_1 - 3/14\ I_2 + 3/5\ I_3 &= 0 \ . \end{aligned}$$

Trabajando mediante notación matricial y con eliminación gaussiana :

$$\begin{bmatrix} 3/6 & -4/7 & -2/5 \\ -2/6 & 11/14 & -1/5 \\ -1/6 & -3/14 & 3/5 \end{bmatrix} = \begin{bmatrix} 1/2 & -4/7 & -2/5 \\ -1/3 & 11/14 & -1/5 \\ -1/6 & -3/14 & 3/5 \end{bmatrix}$$

$$\xrightarrow[1/3\ R_1 + R_2 \\ 1/6\ R_1 + R_3]{2R_1} \begin{bmatrix} 1 & -8/7 & -4/5 \\ 0 & 17/42 & -7/15 \\ 0 & -17/42 & 7/15 \end{bmatrix} \xrightarrow{R_2 + R_3} \begin{bmatrix} 1 & -8/7 & -4/5 \\ 0 & 17/42 & -7/15 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{42/17\ R_2}\begin{bmatrix} 1 & -8/7 & -4/5 \\ 0 & 1 & -98/85 \\ 0 & 0 & 0 \end{bmatrix} .$$

De aquí que

$$\begin{aligned} I_1 - 8/7\ I_2 - 4/5\ I_3 &= 0 \\ I_2 - 98/85\ I_3 &= 0 \\ 0 &= 0. \end{aligned}$$

Hay una variable libre y por lo tanto una infinidad de soluciones. La variable libre es I_3 .

Despejando I_2 e I_1 :

$$I_2 = 98/85\ I_3 \quad ;$$
$$I_1 = 8/7\ I_2 + 4/5\ I_3 = 8/7\ (98/85)\ I_3 + 4/5\ I_3$$
$$= 252/119\ I_3 .$$

La forma general de la solución es

$$(252/119\ I_3\ ,\ 98/85\ I_3\ ,\ I_3) .$$

De ahí que si, por ejemplo, el ingreso del sector pesca es de $ 9'265,000.00, el ingreso del sector ganadero sera

$$I_2 = 98/85\ I_3 = \$\ 10'682,000.00$$

y el de la agricultura

$I_1 = 252/119\ I_3 = \$\ 19'620,000.00.$

Con lo que una solución particular será (19 620 000 , 10 682 000 , 9 265 000).

Resolvamos ahora el problema del balanceo de la ecuación química. El sistema a resolver es :

$$\begin{aligned} x_1 \qquad - 2x_3 \qquad &= 0 \\ x_1 + 4x_2 - 4x_3 - x_4 &= 0 \\ x_1 + 2x_2 \qquad - 2x_4 &= 0 \\ x_2 - x_3 \qquad &= 0. \end{aligned}$$

Utilizaremos de nueva cuenta el método de Gauss, con matrices.

$$\left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & -1 \\ 1 & 4 & -4 & -1 & 0 \\ 1 & 2 & 0 & -2 & 1 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right] \begin{matrix} -R_1 + R_2 \\ \longrightarrow \\ -R_1 + R_3 \end{matrix} \left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & -1 \\ 0 & 4 & -2 & -1 & 1 \\ 0 & 2 & 2 & -2 & 2 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right] \xrightarrow{R_{24}}$$

$$\left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 2 & 2 & -2 & 2 \\ 0 & 4 & -2 & -1 & 1 \end{array}\right] \begin{array}{l} -2R_2 + R_3 \\ -4R_2 + R_4 \end{array} \left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 4 & -2 & 2 \\ 0 & 0 & 2 & -1 & 1 \end{array}\right] \xrightarrow{1/4\ R_3}$$

$$\left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1/2 & 1/2 \\ 0 & 0 & 2 & -1 & 1 \end{array}\right] \xrightarrow{-2R_3 + R_4} \left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Debemos aclarar que la última columna corresponde al control, y no a las constantes. De aquí que el sistema asociado a la última matriz es :

$$\begin{aligned} x_1 - 2x_3 &= 0 \\ x_2 - x_3 &= 0 \\ x_3 - 1/2\ x_4 &= 0 \end{aligned}$$

Entonces

$$\begin{aligned} x_1 &= 2x_3 = x_4 \\ x_2 &= x_3 = 1/2\ x_4 \\ x_3 &= 1/2\ x_4 \end{aligned}$$

Como se ve existe una variable libre, x_4 , por lo que se tienen muchas soluciones. Esto es porque, según la Química, dependiendo de las cantidades de cada sustancia a reaccionar que se mezclen, se van a obtener diferentes cantidades de sustancias después de la reacción. Podemos representar el conjunto solución como :

$$S = \{(k,\ 1/2\ k,\ 1/2\ k,\ k) \ /\ k > 0\} \ ,$$

donde k es un parámetro que nos representa la cantidad de agua.

Si k=2, entonces (2,1,1,2) es una solución particular. Y efectivamente la ecuación

$$2\ NaOH + H_2SO_4 \longrightarrow Na_2SO_4 + 2\ H_2O$$

está balanceada.

Es obvio que si una cierta matriz B tiene una columna de ceros, ésta no se altera si se ejecutan operaciones en los renglones de la matriz (esta es la razón por la cual en estos casos trabajamos con la matriz de coeficientes del sistema y no con la aumentada). Por lo tanto, el último sistema también será homogéneo. Además, como se notó en los ejemplos, el número de ecuaciones en el sistema simplificado es menor o igual que el número de ecuaciones en el sistema original

Suponga que determinado sistema homogéneo de m ecuaciones con n incógnitas es tal que $m < n$ y que además, después de que se llevó el sistema a la forma escalonada, las variables principales en el sistema de ecuaciones correspondiente son x_{k_1}, x_{k_2}, ... , x_{k_r} , $r < n$. Así, el sistema tiene la forma

$$\begin{aligned} \ldots \; x_{k_1} \qquad\qquad + \sum (\quad) &= 0 \\ x_{k_2} \qquad\qquad + \sum (\quad) &= 0 \\ \vdots \qquad\qquad\qquad\quad \vdots & \\ x_{k_r} + \sum (\quad) &= 0 \quad , \end{aligned}$$

donde $\sum (\quad)$ indica sumas en donde aparecen algunas de las n-r variables libres. Resolviendo para las variables principales tenemos :

$$\begin{aligned} x_{k_1} &= - \sum (\quad) \\ &\vdots \\ x_{k_r} &= - \sum (\quad) \; . \end{aligned}$$

Donde, como se vió en el ejemplo, es posible asignar valores arbitrarios a las variables del lado derecho de las ecuaciones del sistema y, por tanto, obtener un número infinito de soluciones para el sistema. En resumen, tenemos que :

Teorema 2.6.1.- Un SEL homogéneo que tiene más incógnitas que ecuaciones siempre tiene un número infinito de soluciones.

El siguiente teorema proporciona una propiedad característica de los sistemas homogéneos.

Teorema 2.6.2.- Cualquier "combinación lineal" (*) de dos soluciones de un SEL homogéneo es también una solución del sistema. Esto es, si $\tau = (c_1, c_2, \ldots, c_n)$ y $\sigma = (d_1, d_2, \ldots, d_n)$ son soluciones de un sistema de ecuaciones homogéneo, entonces la n-upla

$$\alpha\tau + \beta\sigma = (\alpha c_1 + \beta d_1, \alpha c_2 + \beta d_2, \ldots, \alpha c_n + \beta d_n)$$

en donde α, β son números cualesquiera también será una solución del sistema. Por lo tanto, la suma de dos soluciones cualesquiera es una solución, y un múltiplo de cualquier solución también es una solución del sistema de ecuaciones homogéneo.

(*) Nota : Aunque en este teorema se nos haga referencia al término combinación lineal y su representación, este concepto se definirá formalmente en Espacios Vectoriales.

Demostración.- Considere el sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0 \ . \end{aligned}$$

Para escribir este sistema en forma más compacta, denotemos por $L_1(x_1, x_2, \ldots, x_n), \ldots, L_m(x_1, x_2, \ldots, x_n)$ a los miembros izquierdos del sistema, respectivamente, y por $x = (x_1, x_2, \ldots, x_n)$ a la n-upla de las incógnitas. Así,

$$\begin{aligned} L_1(x) &= 0 \\ L_2(x) &= 0 \\ \ldots\ldots & \\ L_m(x) &= 0 \ . \end{aligned}$$

Sean $x = (x_1, x_2, \ldots, x_n)$, y $y = (y_1, y_2, \ldots, y_n)$ (no importa si son o no soluciones del sistema). Entonces

$$\begin{aligned} L_1(\alpha x + \beta y) &= \alpha L_1(x) + \beta L_1(y) \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ L_m(\alpha x + \beta y) &= \alpha L_m(x) + \beta L_m(y) \end{aligned}$$

para cualesquier par de números α y β. Ahora tomemos x e y tales que sean soluciones para el sistema y además $x = \tau$, $y = \sigma$. Entonces

$$\begin{aligned} L_1(\alpha\tau + \beta\sigma) &= \alpha L_1(\tau) + \beta L_1(\sigma) = \alpha 0 + \beta 0 = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ L_m(\alpha\tau + \beta\sigma) &= \alpha L_m(\tau) + \beta L_m(\sigma) = \alpha 0 + \beta 0 = 0 \ . \end{aligned}$$

Con ésto queda demostrada la primera afirmación del teorema. Ahora bien, como suma y múltiplo son simplemente casos especiales de combinaciones lineales (como se verá en Espacios Vectoriales), entonces se tiene probada también la segunda afirmación del teorema :

$$\begin{aligned} \tau + \sigma &= \alpha\tau + \beta\sigma \ , \text{ con } \alpha = \beta = 1, \\ \alpha\tau &= \alpha\tau + \beta\sigma \ , \text{ con } \beta = 0 \ . \end{aligned}$$

El caso en que la n-upla cero es siempre una solución se sigue de la verificación .

Observación.- Básicamente, este teorema clasifica al conjunto solución de un SEL homogéneo como un espacio vectorial. En su momento se dará esta definición.

Para poder apreciar en todo lo que vale la fuerza de este teorema, es necesario saber en qué circunstancias son falsas sus afirmaciones.

Proposición 2.6.3.- Dado un SEL no homogéneo, para cada dos de sus soluciones $\tau = (c_1, c_2, \ldots, c_n)$, $\sigma = (d_1, d_2, \ldots, d_n)$, la suma $\sigma + \tau$ no es una solución. Ningún múltiplo de una solución, $\alpha\sigma$, es una solución, a menos que $\alpha = 1$. La n-upla cero nunca es una solución.

Demostración.- Sea el sistema

$$L_1(x) = b_1$$
$$\cdots\cdots\cdots$$
$$L_m(x) = b_m .$$

Si $\hat{c} + \sigma$ fuera solución, entonces
$L_i(\hat{c} + \sigma) = L_i(\hat{c}) + L_i(\sigma) = b_i + b_i = 2b_i \neq b_i$, para toda $i = i, 2, \ldots, m$ y esto es posible sólo cuando todos los b_i son iguales con cero, lo cual contradice la hipótesis de que el sistema es no homogéneo.

Si $\alpha\sigma$ es solución, entonces
$L_i(\alpha\sigma) = \alpha L_i(\sigma) = \alpha b_i = b_i$, para toda $i = 1, 2, \ldots, m$. Como algún b_i es diferente de cero, entonces α tiene que ser igual a 1.

Finalmente, analizaremos la relación existente entre las soluciones de un SEL y su sistema homogéneo asociado (el sistema obtenido al sustituir las constantes por ceros). Esto es :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \quad (1) \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x &= b_m \end{aligned} \qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{mn}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{mn}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \quad (2) \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x &= 0 \end{aligned}$$

Esta relación queda determinada mediante el siguiente :

Teorema 2.6.4.- Sea u una solución particular de (1) y W la solución general de (2). Entonces la suma de cualquier solución de (1) con cualquier solución de (2) es nuevamente solución de (1).

Demostración.- Por ser $u = (u_1, u_2, \ldots, u_n)$ solución de (1) se debe de cumplir para toda $i = 1, 2, \ldots, m$ que

$$L_i(u) = b_i .$$

Ahora, tomemos a $w \in W$, donde $w = (w_1, w_2, \ldots, w_n)$. Por ser w solución de (2), se debe cumplir para toda $i = 1, 2, \ldots, m$ que

$$L_i(w) = 0 .$$

De estas dos afirmaciones tenemos que, para toda $i = 1, 2, \ldots, m$

$$\begin{aligned} L_i(u + w) &= a_{i1}(u_1 + w_1) + a_{i2}(u_2 + w_2) + \ldots + a_{in}(u_n + w_n) \\ &= a_{i1}u_1 + a_{i1}w_1 + a_{i2}u_2 + a_{i2}w_2 + \ldots + a_{in}u_n + a_{in}w \\ &= (a_{i1}u_1 + a_{i2}u_2 + \ldots + a_{in}u_n) + (a_{i1}w_1 + a_{i2}w_2 + \ldots a_{in}w_n) \\ &= L_i(u) + L_i(w) = b_i + 0 = b . \end{aligned}$$

Por lo tanto, u + w es solución de (1).

Veamos algunos ejemplos. Tomemos el sistema del problema 4 ; éste es

$$\begin{aligned} x_1 \quad & - 2x_3 \quad = 0 \\ x_1 + 4x_2 & - 4x_3 - x_4 = 0 \\ x_1 + 2x_2 & \quad - 2x_4 = 0 \\ x_2 - x_3 & \quad = 0 \end{aligned}$$

y su solución general es

$$\begin{aligned} x_1 &= x_4 \\ x_2 &= 1/2\ x_4 \\ x_3 &= 1/2\ x_4 . \end{aligned}$$

Sean $\tau = (2,1,1,2)$ la solución que se obtiene con $x_4 = 2$ y $\sigma = (4,2,2,4)$ la solución obtenida con $x_4 = 4$. Tomemos $\alpha = \beta = 2$. Entonces $\alpha\tau = (4,2,2,4)$ y $\beta\sigma = (8,4,4,8)$, de donde $\alpha\tau + \beta\sigma = (12,6,6,12)$, que efectivamente satisface el sistema y es, por lo tanto, solución de él, tal y como lo afirma el teorema 2.6.2.

Por otro lado, observemos el sistema del problema (1), cuyas soluciones son

$$\begin{aligned} x_1 &= 50 \\ 0 \leq x_2 &\leq 200 \\ 150 \leq x_3 &\leq 350 \\ x_4 &= 150 \\ 0 \leq x_5 &\leq 200 \\ x_6 &= 0 \\ 100 \leq x &\leq 300. \end{aligned}$$

Sean τ y σ dos soluciones tales que
$\tau = (50, 100, 200, 150, 150, 0, 150)$
$\sigma = (50, 150, 300, 150, 200, 0, 200)$.

De aquí que
$\tau + \sigma = (100, 250, 500, 300, 350, 0, 350)$,
la cual claramente no es solución del sistema (teorema 2.6.3)

Para ilustrar el último teorema tomemos el sistema

$$\begin{aligned} x_1 - x_2 + 3x_3 &= 1 \\ 3x_1 - 3x_2 + 9x_3 &= 3 \qquad (I) \\ -2x_1 + 2x_2 - 6x_3 &= -2 \end{aligned}$$

Resolviendolo tenemos :

$$\left[\begin{array}{ccc|c} 1 & -1 & 3 & 1 \\ 3 & -3 & 9 & 3 \\ -2 & 2 & -6 & -2 \end{array}\right] \xrightarrow[1/2\ R_3]{-1/3\ R_2} \left[\begin{array}{ccc|c} 1 & -1 & 3 & 1 \\ -1 & 1 & -3 & -1 \\ -1 & 1 & -3 & -1 \end{array}\right] \xrightarrow[R_1 + R_3]{R_1 + R_2}$$

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

De aquí :

$$x_1 - x_2 + 3x_3 = 1 \ ,$$

por lo tanto

$$x_1 = x_2 - 3x_3 + 1 \ ,$$

de donde la solución general está dada por
$S = \{(x_2 - 3x_3 + 1, x_2, x_3) / x_2, x_3 \in R\}$.

Una solución particular es, si $x_3 = x_2 = 1$, entonces $x_1 = -1$, de donde $(-1,1,1)$ es solución. Sea $u = (-1,1,1)$.

Tomemos ahora el sistema homogéneo asociado a (I) :

$$\begin{bmatrix} 1 & -1 & 3 \\ 3 & -3 & 9 \\ -2 & 2 & -6 \end{bmatrix} \xrightarrow[1/2\ R_3]{-1/3\ R_2} \begin{bmatrix} 1 & -1 & 3 \\ -1 & 1 & -3 \\ -1 & 1 & -3 \end{bmatrix} \xrightarrow[R_1 + R_3]{R_1 + R_2} \begin{bmatrix} 1 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ,$$

de donde $x_1 - x_2 + 3x_3 = 0$. Entonces $x_1 = x_2 - 3x_3$. La solución general en este caso es

$$S = \{(x_2 - 3x_3 , x_2 , x_3) / x_2 , x_3 \in R\} \ .$$

Una solución particular es, si $x_2 = x_3 = 1$, entonces $x_1 = -2$. Sea w esta solución, ésto es, $w = (-2,1,1)$.

Entonces

$$u + w = (-3,2,2) \ ,$$

que no es difícil verificar que efectivamente también es solución de (I).

C A P I T U L O 3

MATRICES

3.1 I N T R O D U C C I O N .

En el capítulo anterior se introdujo el concepto de ma-- triz y se empleó como una herramienta auxiliar muy importante para el estudio y solución de los SEL. A raíz de esto, como- muchos de los problemas prácticos pueden plantearse como SEL, resulta que dichos problemas pueden ser planteados en térmi-- nos de matrices y resueltos a través de operaciones matricia- les y sus propiedades.

En la actualidad, las matrices son instrumentos muy efi- cientes en varias ramas de la Matemática Aplicada pues utili- zan métodos matriciales. Definiremos ahora algunos conceptos y desarrollaremos un álgebra para matrices de manera que re-- almente proporcionen aplicaciones útiles. Veremos cómo en el algebra matricial se definen operaciones de modo que el sis-- tema matemático resultante es un espacio vectorial.

Debemos hacer hincapié en que lo más importante de este capítulo es el estudio del papel que juegan las matrices den- tro de la solución de SEL representativos de situaciones rea- les y el dejar lo suficientemente clara la relación que exis- te entre unas y otros.

3.2 R E S E Ñ A H I S T O R I C A .

Aunque el concepto de matriz precede lógicamente al de - determinante, es éste último el que se desarrolló primero. No fué hasta después cuando los matemáticos se interesaron -- específicamente en el arreglo rectangular de los números que aparecen en el determinante, aunque varias propiedades rela-- cionadas con esta teoría ya se habían desarrollado.

Los conceptos de matriz aumentada y matriz de los coefi- cientes fueron introducidos por Henry J. S. Smith (1826-1883) con ocasión de la solución de sistemas de m ecuaciones con n incógnitas.

Se considera a Arthur Cayley (1821-1895) el fundador de la teoría de matrices pues fué el primero en publicar una -- serie de tratados sobre el tema.

Muchos matemáticos pretendieron extender y desarrollar - la idea de Cayley al morir éste; los que prestaron un mayor - número de contribuciones fueron Frobenius, Smith, Clebsch, -- Buchheim, Hensel, Jordan, Metzler y Taber.

3.3 P R O B L E M A S D I V E R S O S.

Algunos problemas típicos acerca de los cuales se hace mención arriba son:

Problema 1.-Contactos de primero y segundo orden de una enfermedad contagiosa.- Suponga que en un cierto poblado tres personas han contraído una enfermedad infecciosa mucho muy contagiosa. Existe el interés por parte de las autoridades sanitarias por conocer mediante un estudio el número de contactos habidos entre personas contagiadas por otras igualmente contagiadas y por las inicialmente enfermas. ¿Cómo se podría determinar ésto?

Problema 2.- Cadenas de Markov.- Suponga que en una cierta actividad laboral cada persona está clasificada en uno sólo de los tres estados siguientes:
Estado 1: Trabajador Profesional (TP)
Estado 2: Trabajador No Clasificado (N) Calificados
Estado 3: Trabajador Calificado (C).

Suponga además que cada persona tiene un hijo y que la matriz

$$\begin{matrix} & TP & N & C \end{matrix}$$
$$\begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.3 & 0.2 & 0.5 \end{bmatrix}$$

dá la probabilidad de que la próxima generación pase de un estado a otro. Si el 20% de los trabajadores son trabajadores profesionales, el 50% son trabajadores no calificados y el 30% restante son trabajadores calificados, determine las probabilidades respectivas de que el nieto de un trabajador sea profesional, trabajador no calificado o trabajador calificado.

Problema 3.- Se combinan tres compuestos diferentes A, B,C, para producir tres tipos de fertilizantes. Una unidad de fertilizante de tipo I requiere 15 kg. del compuesto A, 15 kg. del B y 10 kg. del C; una unidad del tipo II requiere de 10 kg. del compuesto A y 20 kg. del C; una unidad del tipo III necesita 18 kg. del compuesto A, 12 kg. del B y 16 kg. del C. Si se dispone de 1800 kg. del compuesto A, 1800 kg. del B y 1200 kg del C en total, ¿cuantas unidades de cada fertilizante se producirán al utilizar todo el material disponible?

Problema 4.- Alimentación y Requerimientos dietéticos.- Suponga que un doctor ordena que la dieta de un paciente contenga un requerimiento mínimo diario de una unidad de tiamine, dos unidades de riboflavina y tres unidades de hierro. Si seleccionamos tres menús diferentes que contienen a estas

vitaminas en las siguientes cantidades:

	M_1	M_2	M_3
Tiamina	1	0	1
Rivoflabina	0	2	3
Hierro	4	0	1

¿Qué porción de cada receta debería servirse al paciente para que cubra los requerimientos mínimos?

Problema 5.- Modelos de Leontief.- Suponga que la agricultura, la manufactura y la construcción son los tres sectores o compañías que interactúan en el estado de Sonora, mientras los ahorros es el sector abierto. Cada " compañía " fabrica su producto de tal manera que surte a las otras dos --- compañías y crean un producto final (llamado inversión), ---- mientras que los fondos invertidos, llamados ahorros, son --- consumidos solamente por el sector construcción. Estos son - los datos en billones de pesos:

		Consumo			
		Agr.	Man.	Cons.	Inv.	Total
	Agr.	4	6	3	7	20
	Man.	8	18	1	3	30
	Cons.	8	6	4	2	20
	Ahorros	0	0	10	0	12
	Total	20	30	20	12	82

Convierta este problema en un modelo de Leontief de 3 -- sectores y prediga la producción necesaria para satisfacer la inversión dada.

3.4. M A T R I C E S : Definición y operaciones entre ma--- trices.

Introduciremos la noción de matriz mediante ejemplos para posteriormente formalizar su definición.

Ejemplo 1.- Supóngase que un fabricante de juguetes utiliza 4 materiales diferentes en su construcción. Estos materiales son caucho, madera, vidrio y plástico, y las cantidades de cada uno de ellos están dadas por el vector $\begin{bmatrix} 4 \\ 8 \\ 13 \\ 17 \end{bmatrix}$.

Supongamos también que hay otros dos fabricantes de juguetes que trabajan en base a los mismos materiales en cantidades $\begin{bmatrix} 2 \\ 10 \\ 15 \\ 19 \end{bmatrix}$ y $\begin{bmatrix} 3 \\ 9 \\ 15 \\ 17 \end{bmatrix}$, respectivamente.

Conjuntando toda la información anterior podemos almacenarla en alguna forma, es decir, arreglarla de manera que podamos interpretar este arreglo de acuerdo al significado de los datos. En este caso, juntando las órdenes de los cuatro materiales para los tres fabricantes, el arreglo

$$\begin{array}{l} \\ \text{Caucho} \\ \text{Madera} \\ \text{Vidrio} \\ \text{Plástico} \end{array} \begin{array}{c} \begin{array}{ccc} \text{fab. 1} & \text{fab.2} & \text{fab.3} \end{array} \\ \begin{bmatrix} 4 & 2 & 3 \\ 8 & 10 & 9 \\ 13 & 16 & 15 \\ 17 & 13 & 12 \end{bmatrix} \end{array}$$

representa dichas cantidades. Al interpretarlo vemos que este arreglo nos proporciona la relación que existe entre cada material y cada fabricante. Además, podemos descomponer este arreglo de dos formas distintas: por columnas, que fue como se tomó la información inicialmente (fabricantes), y por renglones (que daría lugar a vectores material). Con esto, este arreglo se dice que es de doble entrada.

Ejemplo 2.- Cinco animales de laboratorio se nutren con tres alimentos diferentes. Si definimos c_{ij} como el consumo diario del i-ésimo alimento por el j-ésimo animal, entonces el arreglo

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

registra el consumo diario de cada alimento por cada animal.

Otra vez, como columnas tenemos los consumos de cada animal y como renglones la cantidad de cada alimento consumida diariamente, es decir, es un arreglo de doble entrada.

Estos arreglos reciben el nombre de matriz y lo definimos como sigue:

Definición 3.4.1.- Una matriz A (*) de mxn es un arreglo rectangular de mn números de la forma:

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

que tiene m renglones y n columnas.

* El término matriz se utilizó por vez primera en 1850 por el matemático británico James Joseph Sylvester (1814-1897) para distinguir a las matrices de los determinantes (que se estudiaran en el siguiente capítulo). De hecho, el término ma---

Notación.- A las matrices las representaremos con letras mayúsculas y a sus elementos con la misma letra pero minúscula. Por ser un arreglo de doble entrada utilizamos [ ] para representar a los elementos de la matriz en su interior. El arreglo anterior lo podemos escribir en forma abreviada como $A = [\, a_{ij} \,]_{mxn}$, $i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$, donde a_{ij} representa al elemento que se encuentra situado en la intersección del i-ésimo renglón y la j-ésima columna, y además, $a_{ij} \in R$. Otra forma de representar al ij-ésimo elemento de la matriz A es: $[\, A \,]_{ij} = a_{ij}$; esto es, el ij-ésimo elemento de la matriz A es el término a_{ij}.

El i-ésimo renglón de la matriz está dado por $[a_{i1}, a_{i2}, \ldots, a_{in}]$, y la j-ésima columna es $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$.

Observaciones:

- El número de renglones (m) y de columnas (n) de la matriz, son los que determinan el tamaño mXn de la misma. Así, la matriz A del ejemplo 1 es de tamaño 4X3 y la matriz C del ejemplo 2 es de tamaño 3x5.
- Una matriz de 1xn que consiste únicamente de un renglón se llama simplemente matriz renglón; de igual manera, una matriz de mx1 que consiste sólo de una columna es llamada una matriz columna. Una matriz $A = [\, a \,]$ que consiste de un número solamente es identificada con el número. Esto significa que los vectores se pueden interpretar como casos particulares de las matrices; así, el vector renglón (-1, 0, 8, 1)) es una matriz de 1x4 y, de manera análoga, el vector columna $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ es una matriz de 3x1.

Cabe señalar que más adelante, cuando se defina el producto de matrices, éste se relacionará con el producto punto de vectores y se verá como una generalización de él.

Resumiendo un poco, podemos afirmar que, al igual que los vectores, las matrices surgen de situaciones prácticas y se utilizan como almacén de datos.

Definición 3.4.2.- Si m = n entonces la matriz se llama cuadrada y diremos que tiene orden n.

Definición 3.4.3.- Si todos los elementos de una matriz de mxn son iguales a cero, la matriz se llama matriz nula o matriz cero y se denota por 0_{mxn}.

Definición 3.4.4.- Dos matrices $A = [\, a_{ij} \,]_{mxn}$ y $B = [\, b_{ij} \,]_{mxn}$ son iguales si y sólo si tienen el mismo tamaño y sus correspondientes componentes son iguales, esto es, -

Ejemplo 3.- Sean las matrices $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 5 & 0 \end{bmatrix}$ y ----

$B = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$. $A \neq B$, pues no son del mismo tamaño.

A continuación definiremos algunas operaciones algebraicas que muestren la verdadera utilidad de las matrices.

OPERACIONES ENTRE MATRICES

1.- Adición de matrices.- Si las matrices son una generalización de los vectores, deben entonces satisfacer una regla de adición que sea equivalente a la definida para vectores. -- Puesto que los vectores se suman componente a componente, entonces la adición de matrices deberá definirse de manera semejante.

Definición 3.4.5.- Sean A y B matrices de m x n ambas. La suma de A y B se define como la matriz A + B de m x n dada por

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

Esto es, A + B es la matriz de m x n obtenida al sumar las -- componentes correspondientes de A y B.

Mediante notación abreviada, la definición anterior se - simplifica y entonces :

$[A + B]_{ij} = a_{ij} + b_{ij} = [A]_{ij} + [B]_{ij}$.

Observación.- La suma de matrices está definida únicamente para matrices del mismo tamaño.

Ejemplo 4.- Supongamos que en el ejemplo 1 los materiales pueden adquirirse en dos localidades distintas. Supongamos que la matriz dada

$$A = \begin{bmatrix} 4 & 2 & 3 \\ 8 & 10 & 9 \\ 13 & 16 & 15 \\ 17 & 13 & 12 \end{bmatrix}$$

representa las órdenes atendidas por una de estas localidades surtidoras, y que la matriz

$$B = \begin{bmatrix} 2 & 2 & 4 \\ 3 & 5 & 1 \\ 6 & 5 & 6 \\ 8 & 10 & 9 \end{bmatrix}$$

representa el suministro en la segunda localidad.

La matriz que representa los suministros totales de los 4 materiales para los tres fabricantes está dada por:

$$A + B = \begin{bmatrix} 6 & 4 & 7 \\ 11 & 15 & 10 \\ 19 & 21 & 21 \\ 25 & 23 & 21 \end{bmatrix}.$$

Al igual que sucede con los vectores, toda matriz $A = [\, a_{ij} \,]$ posee un inverso aditivo, éste es, una matriz que se denota $-A$ definida por $-A = [\, -a_{ij} \,]_{mxn}$ tal que $A + (-A) = 0_{mxn}$. En base a éste podemos definir la diferencia de matrices como sigue:

Definición 3.4.6.- Sean $A = [\, a_{ij} \,]_{mxn}$, $B = [\, b_{ij} \,]_{mxn}$. La resta o diferencia de A y B está definida por :

$$A - B = A + (-B).$$

En otras palabras, $A - B$ es A más el inverso aditivo de B, y $[\, A - B \,]_{ij} = [\, A \,]_{ij} + [\, -B \,]_{ij}$.

Ejemplo 5.- En el problema anterior, si $A + B = C$ entonces

$$C = \begin{bmatrix} 6 & 4 & 7 \\ 11 & 15 & 10 \\ 19 & 21 & 21 \\ 25 & 23 & 21 \end{bmatrix}.$$

$$C - B = C + (-B) = \begin{bmatrix} 6 & 4 & 7 \\ 11 & 15 & 10 \\ 19 & 21 & 21 \\ 25 & 23 & 21 \end{bmatrix} + \begin{bmatrix} -2 & -2 & -4 \\ -3 & -5 & -1 \\ -6 & -5 & -6 \\ -8 & -10 & -9 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 2 & 3 \\ 8 & 10 & 9 \\ 13 & 16 & 15 \\ 17 & 13 & 12 \end{bmatrix} = A,$$

lo cual era de esperarse, ya que si $C = A + B$, entonces se tiene que $C - B = (A + B) - B = A + B - B = A + 0 = A$. De aquí concluimos que una propiedad válida para la adición de matrices es la asociatividad. Además, si alteramos el orden de suma de las matrices resulta que:

$$B + A = \begin{bmatrix} 2 & 2 & 4 \\ 3 & 5 & 1 \\ 6 & 5 & 6 \\ 8 & 10 & 9 \end{bmatrix} + \begin{bmatrix} 4 & 2 & 3 \\ 8 & 10 & 9 \\ 13 & 16 & 15 \\ 17 & 13 & 12 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 4 & 7 \\ 11 & 15 & 10 \\ 19 & 21 & 21 \\ 25 & 23 & 21 \end{bmatrix} = A + B ,$$

lo cual ilustra la conmutatividad de la adición de matrices.

Como ya sabemos, existe una matriz (la matriz nula), cuyas componentes son todas cero; por este hecho y con la definición dada para la adición de matrices, la matriz nula se comporta como el cero real bajo esta operación, éste es:

$$A + 0 = \begin{bmatrix} 4 & 2 & 3 \\ 8 & 10 & 9 \\ 13 & 16 & 15 \\ 17 & 13 & 12 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 2 & 3 \\ 8 & 10 & 9 \\ 13 & 16 & 15 \\ 17 & 13 & 12 \end{bmatrix} = A .$$

Observación.- Nótese que las matrices que se utilizan a lo largo del capítulo son todas de componentes reales; ésto se debe a que al definir el concepto de matriz se pidió que $a_{ij} \in R$, es decir, que cada componente de la matriz A sea un número real. Como algunas operaciones entre matrices se definen en términos de la misma operación entre los elementos de las matrices (como se verá también en la siguiente operación) y estos son reales, tendremos que, bajo estas operaciones, se cumplen las mismas propiedades. Adviértase que este es el argumento principal para el establecimiento de varias propiedades de operaciones entre matrices.

Definición 3.4.7.- Al conjunto de matrices de mxn de componentes reales, llamadas matrices reales, se les denota R_{mxn}

2.- Multiplicación por un escalar.- Introduciremos ahora una nueva operación con matrices llamada multiplicación por un escalar.

Análogamente a lo que sucede con vectores en los cuales la multiplicación por un escalar da lugar a otro vector en el mismo espacio que el original (es decir, con el mismo número de componentes) y cuyas componentes son las originales cada una de ellas multiplicada por el escalar, definiremos la multiplicación de una matriz por un escalar.

Definición 3.4.8.- Sea $A = [a_{ij}]_{mxn}$ y $k \in R$, entonces kA se define como la matriz $[ka_{ij}]_{mxn}$ obtenida al multiplicar cada elemento de A por k. Así, si:

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

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

De manera abreviada: $[\, kA \,]_{ij} = ka_{ij} = k[A]_{ij}$.

Ejemplo 6.- Si queremos saber a cuánto equivale las 2/3 partes del suministro total de los fabricantes del ejemplo 4, tomamos $k = 2/3$ y

$$C = \begin{bmatrix} 6 & 4 & 7 \\ 11 & 15 & 10 \\ 19 & 21 & 21 \\ 25 & 23 & 21 \end{bmatrix} \text{ de donde:}$$

$$kC = \begin{bmatrix} 4 & 8/3 & 14/3 \\ 22/3 & 10 & 20/3 \\ 38/3 & 14 & 14 \\ 50/3 & 46/3 & 14 \end{bmatrix}.$$

Con las definiciones dadas de adición de matrices y multiplicación de una matriz por un escalar es bastante sencillo comprobar que el conjunto $R_{m\times n}$ satisface las siguientes propiedades básicas:

a) Para $A, B \in R_{m\times n}$, $A + B \in R_{m\times n}$. (Cerradura para la suma).

b) Para $A, B \in R_{m\times n}$. $A + B = B + A$. (Conmutatividad).

c) Para $A, B, C \in R_{m\times n}$, $A + (B + C) = (A + B) + C$. (Asociatividad).

d) Existe $0 \in R_{m\times n}$, tal que $A + 0 = 0 + A = A$, para toda $A \in R_{m\times n}$. (Existencia del neutro aditivo).

e) Para toda $A \in R_{m\times n}$ existe $-A \in R_{m\times n}$ tal que $A + (-A) = 0$. (Existencia del inverso aditivo).

f) Para toda $A \in R_{m\times n}$ y $k \in R$, $kA \in R_{m\times n}$. (Cerradura para el producto por un escalar).

g) Si $k = 1$ entonces $1 \cdot A = A$; si $k = 0$ entonces $0 \cdot A = 0$. (Existencia de la identidad para la multiplicación por un escalar).

h) Si $k \in R$ y $A, B \in R_{m\times n}$, entonces $k(A + B) = kA + kB$. (Distributividad).

i) Si $k_1, k_2 \in R$ y $A \in R_{m\times n}$, entonces $(k_1 + k_2)A = k_1A + k_2A$. (Distributividad).

j) Si $k_1, k_2 \in R$ y $A \in R_{m\times n}$, entonces $(k_1k_2)A = k_1(k_2A) = k_2(k_1A)$. (Asociatividad).

Nótese que cada propiedad involucra una igualdad de matrices. El demostrar cada una de las propiedades consiste en demostrar la igualdad en los tamaños y la igualdad componente a componente.

Observación.- Las anteriores operaciones con matrices pueden considerarse como una generalización de las operaciones correspondientes definidas en el Capítulo I, por lo que también el conjunto R_{mxn} bajo estas dos operaciones se clasificará como un espacio vectorial.

3.- Producto Matricial.- Esta es una operación en la cual intervienen únicamente matrices. Si recordamos el producto entre vectores (producto punto), resulta que de acuerdo a su definición se obtiene al final un número (que puede asociarse a una matriz de 1x1) cuyo significado depende del tipo de problema en que se aplica. Ahora veremos si se conserva la analogía para esta operación bajo matrices y en qué forma nos conviene definir la operación para que preste aplicaciones más útiles.

A manera de motivación, recordemos la definición de producto de vectores:

Sean $A = (a_1, a_2, \ldots, a_n)$, $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ dos vectores en R^n.

Entonces:

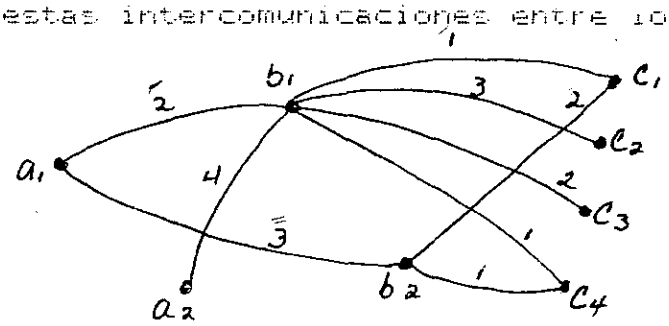
$$A \cdot B = (a_1, a_2, \ldots, a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n .$$

Pero al mismo tiempo A es una matriz de 1xn y B es otra matriz de nx1; A·B se asocia con una matriz de 1x1. ¿Cómo relacionaríamos esto de manera general para dos matrices de tamaños arbitrarios? Obsérvese que el número de columnas de la primera matriz y el número de renglones de la segunda son los que indican el número de componentes de cada una vista como vector y que, además, la matriz resultante tiene el mismo número de renglones que la primera matriz y de columnas la segunda.

Veamos el siguiente ejemplo:

Ejemplo 7.- Tres países A,B,C, mantienen sus relaciones comerciales a través de diversas líneas aéreas que salen de los diferentes aeropuertos situados en estos países pero que pueden seguir la misma ruta de vuelo. El siguiente esquema

lustra estas intercomunicaciones entre los tres países:



Cada a_i, b_j, c_k denota los aeropuertos de cada país respectivamente, donde $i, j = 1,2$, $k = 1,2,3,4$. Los números sobre las líneas de unión entre los aeropuertos indican el número de posibles elecciones de líneas aéreas para cada ruta; por ejemplo, el número 4 que conecta a_2 con b_1 indica que hay 4 compañías aéreas que siguen esa ruta. ¿Cuál es el número de posibles rutas a seguir que conecten al país A con el C, (por cada aeropuerto respectivamente)?

Podemos verlo así:
Una ruta de a_1 a c_1 puede pasar por b_1 o b_2. En el primer caso existen (2)(1) = 2 posibilidades, y en el segundo caso (3)(2) = 6 posibilidades, es decir, 8 posibilidades en total. Análogamente se determinan los demás resultados:
de a_1 a c_2 se tienen (2)(3) = 6 posibilidades,
de a_1 a c_3 se tienen (2)(2) = 4 posibilidades,
de a_1 a c_4 se tienen (2)(1) + (3)(1) = 5 posibilidades,
de a_2 a c_1 se tienen (4)(1) = 4 posibilidades,
de a_2 a c_2 se tienen (4)(3) = 12 posibilidades,
de a_2 a c_3 se tienen (4)(2) = 8 posibilidades,
de a_2 a c_4 se tienen (4)(1) = 4 posibilidades.

Pero como estos datos nos relacionan a cada aeropuerto del país A con cada aeropuerto del país C y nos proporciona el número de posibles rutas a seguir en esta relación, toda esta información obtenida la podemos expresar como sigue:

$$\begin{array}{c} \\ a_1 \\ a_2 \end{array}\begin{array}{c} \begin{array}{cccc} c_1 & c_2 & c_3 & c_4 \end{array} \\ \begin{bmatrix} 8 & 6 & 4 & 5 \\ 4 & 12 & 8 & 4 \end{bmatrix} \end{array} = P .$$

Si hacemos lo mismo con la información contenida en el esquema podemos formar dos cuadros (arreglos) donde se contemplen las rutas de las líneas aéreas que conectan al país A con el B (matriz M) y el país B con el C (matriz N):

$$\begin{array}{c} \\ a_1 \\ a_2 \end{array}\begin{array}{c} \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix} \end{array} = M, \qquad \begin{array}{c} \\ b_1 \\ b_2 \end{array}\begin{array}{c} \begin{array}{cccc} c_1 & c_2 & c_3 & c_4 \end{array} \\ \begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 0 & 0 & 1 \end{bmatrix} \end{array} = N.$$

Como en el desarrollo de los cálculos lo que se hizo fue multiplicar el número de rutas de A a B por el número de rutas de B a C, tal vez estos mismos resultados se obtengan en base a las matrices M y N, pues en ellas está contenida esta información. Pero entonces tendríamos que estudiar la multi-

plicación matricial de tal manera que se obtenga como resul-- tado a P. Pero para esto:

$$MN = \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix}_{mn} \begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 0 & 0 & 1 \end{bmatrix}_{nk} \quad (n = m)$$

$$= \begin{matrix} (2)(1)+(3)(2) & (2)(3)+(3)(0) & (2)(2)+(3)(0) & (2)(1)+(3)(1) \\ (4)(1)+(0)(2) & (4)(3)+(0)(0) & (4)(2)+(0)(0) & (4)(1)+(0)(1) \end{matrix}$$

$$= \begin{bmatrix} 8 & 6 & 4 & 5 \\ 4 & 12 & 8 & 4 \end{bmatrix} = P.$$

Pero ésto es multiplicar cada vector renglón de la ma--- triz M por cada vector columna de N. Sólo así se satisface - que MN = P.

Observe otra vez que M es de tamaño 2x2, N de 2x4 y la - matriz producto P es de 2x4, es decir, su tamaño quedó determinado por el número de renglones de M y de columnas de N.

Con ésto llegamos a la siguiente
Definición 3.4.9.- Sean A una matriz de mxn y B una matriz de nxk. El producto AB es una matriz de mxk denotada por C, cuyos elementos están definidos de la siguiente forma: el elemento del i-ésimo renglón y j-ésima columna de la matriz C es igual a la suma de los productos de los elementos del i-ésimo renglón de la matriz A con los correspondientes elementos de la j-ésima columna de la matriz B. Así:

$$A\,B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mk} \end{bmatrix}$$

donde $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$, con $i = 1,2,\ldots,m$, $j = 1,2,\ldots,k$.

Esta operación puede ilustrarse mediante el siguiente -- diagrama:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{nk} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & c_{ij} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mk} \end{bmatrix}$$

Observaciones:
- Esta definición se aplica solamente cuando el número de -- columnas de la primera matriz coincide con el número de renglones de la segunda. Si este producto está definido se dice que las matrices son conformables, es decir, se ajustan. Si - no es así AB será definida ...

B es una matriz de nxk, para algún número entero k:

$A_{mxn} \times B_{nxk} = C_{mxk}$, con C = AB.

- Si introducimos una nueva notación e identificamos como A_i al i-ésimo renglón de la matriz A, y si $A^{(j)}$ denota la j-ésima columna de A, entonces esta definición puede también representarse como:

Si A es una matriz de mxn y B es una matriz de nxk tales que

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{bmatrix} \quad \text{entonces}$$

$$AB = \begin{bmatrix} A_1 \cdot B^{(1)} & A_1 \cdot B^{(2)} & \cdots & A_1 \cdot B^{(k)} \\ A_2 \cdot B^{(1)} & A_2 \cdot B^{(2)} & \cdots & A_2 \cdot B^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ A_m \cdot B^{(1)} & A_m \cdot B^{(2)} & & A_m \cdot B^{(k)} \end{bmatrix},$$

que es una matriz de mxk.

Por consiguiente, la multiplicación de matrices es una generalización del producto interior de vectores.

Además, de acuerdo a la definición de esta operación, resulta que podemos representar matricialmente a los sistemas de ecuaciones (es decir, podemos expresarlos mediante notación matricial). Veamos ésto.

Sea el sistema

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

Tomemos

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

donde A es una matriz cuadrada de orden n y X y B son matrices columna de nx1. Tendremos que AX = B, siempre y cuando -

$$AX = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + a_{2n}x_n \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{nn}x_n \end{bmatrix}$$

$$= \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = B \ .$$

Esto podrá ser si y sólo si X es una matriz columna tal que sus componentes satisfacen el SEL dado. Así la ecuación matricial AX = B es equivalente al sistema dado. Por lo tanto, el resolver un SEL se reduce a resolver la ecuación matricial AX =B, para la matriz incógnita X: para ésto necesitamos determinar la matriz columna X que satisfaga el SEL correspondiente. Esto lo veremos más adelante.

Observación.- El ejemplo anterior ilustra un hecho muy importante: el producto de matrices, en general, no es conmutativo. Esto es, en general AB ≠ BA, aunque en ciertas ocasiones haya excepciones y sí se dé que AB = BA. En el ejemplo anterior vimos que MN es una matriz de 2x4, mientras que NM no está definido puesto que en este orden las matrices no se ajustan:

$$\underset{2\times4}{N} \quad \underset{}{X} \quad \underset{2\times2}{M} \qquad\qquad \underset{2\times2}{M} \quad \underset{}{X} \quad \underset{2\times4}{N} = \underset{2\times4}{P}$$

Con ésto, debe tenerse especial cuidado en el orden en que multipliquemos las matrices.

A continuación resolveremos algunos de los problemas incluídos en la sección anterior.

Problema 1.- Contactos de primero y segundo orden de una enfermedad contagiosa. Suponga que en un cierto poblado tres personas han contraído una enfermedad infecciosa mucho muy contagiosa. Existe el interés de estudiar los contactos entre personas contagiadas por otras anteriormente contagiadas, y por las inicialmente enfermas. Para ésto, las autoridades sanitarias están efectuando averiguaciones en las cuales seleccionaron un grupo de 6 personas que declararon haber tenido contacto con una o varias de las personas enfermas. Pero como estas personas, antes de enterarse de la enfermedad de las primeras tuvieron contacto con otras personas, se formó un segundo grupo de 7 personas que tuvieron contacto con alguna(s) de las 6 personas del primer grupo.

Con la información obtenida, las autoridades sanitarias definen una matriz A de 3x6 tomando $a_{ij} = 1$ si la j-ésima persona del primer grupo ha tenido algun contacto con la i-ésima persona enferma y $a_{ij} = 0$ si no ha sido así. De igual modo, define otra matriz B de 6x7 haciendo $b_{ij} = 1$ si la j-ésima persona del segundo grupo ha tenido contacto con la i-ésima persona del primer grupo y $b_{ij} = 0$ en caso contrario. Estas dos matrices describen los contactos de primer orden o directos entre grupos.

Así, obtuvieron las matrices:

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

Según ésto, $a_{33} = 1$ significa que la tercera persona del primer grupo tuvo contacto con la tercera persona infectada; en cambio $a_{22} = 0$ significa que la segunda persona del primer grupo no tuvo contacto de ninguna especie con la segunda persona infectada. De manera análoga se explica la matriz B (--pero ahora entre los dos grupos de 6 y 7 personas, respecti--vamente).

Como las 7 personas del segundo grupo adquieren la en---fermedad por haber tenido contacto con las del primer grupo, y éstas a su vez la adquirieron de los enfermos, se dice que las personas del segundo grupo adquieren la enfermedad por --contacto indirecto o de segundo orden con las personas enfermas. Este número de contactos indirectos está descrito por - la matriz producto AB, y la ij-ésima componente representa el número de contactos de segundo orden entre la j-ésima persona del segundo grupo y la i-ésima persona infectada. Así, tenemos:

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

En donde $c_{31} = 2$ indica que la primera persona del se---gundo grupo tuvo dos contactos indirectos con la tercera persona infectada, y, en total, dicha persona tuvo 2 + 1 = 3 --- contactos indirectos con las personas enfermas. Sólo la ---- quinta persona no tuvo ningún contacto.

Observación.- Este problema ilustra el comportamiento --epidémico de una enfermedad, ya que por contagio dicha enfermedad se transmite de una persona a otra.

Debemos hacer notar que mediante matrices podemos simplificar y escribir en forma compacta la información obtenida en el problema y además que, utilizando dichas matrices, podemos obtener el número de contactos entre personas contagiadas por otras y las originadores de la epidemia.

Propiedades del producto de matrices.- Aunque muchas de las propiedades algebraicas de los números relaes son heredadas por las matrices reales (por el hecho de que sus compo---nentes son números reales), existen varias excepciones. Una de ellas ya se mencionó anteriormente: el producto de matri--ces, en general, no es conmutativo, ya que podría suceder que

el producto esté definido pero sin embargo no exista igualdad entre ellas.

Pero ésto no significa que no haya semejanza en algunas otras propiedades algebraicas con los números reales. Algu-- nas de las más importantes son:

a) $A(BC) = (AB)C$. (Asociatividad para la multiplicación).
b) $(A + B)C = AC + BC$. (Distributividad por la derecha).
c) $A(B + C) = AB + AC$. (Distributividad por la izquierda).
d) $k(AB) = (kA)B = A(kB)$.
e) Existe I, matriz cuadrada de la forma:

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

tal que $AI = A$; I recibe el nombre de matriz identidad por la derecha. Si sucede que $IA = A$, I se llama matriz identidad - por la izquierda.

A, B, C, I son matrices que se ajustan a la operación -- indicada y $k \in R$.

Observación.- Sobre el cómo demostrar estas propiedades, es válido el comentario que se hizo en las operaciones ante-- riores. Demostraremos sólo algunas de ellas.

Demostración.- a) Sean A_{mxn}, B_{nxr}, C_{rxk}. Con respecto a los tamaños, AB será una matriz de mxr y [(AB)C] será de mxk. Por otro lado BC es una matriz de nxk y [A(BC)] será una ma-- triz de mxk. Por lo tanto, A(BC) y (AB)C tienen el mismo ta- maño. En lo que se refiere a la igualdad componente a compo- nente puede verse mediante ejemplos particulares.

Nota: La demostración formal se incluye posteriormente - en el apéndice ya que requiere de dobles sumatorias que no -- todos los lectores conocen.

b) Lo demostraremos basándonos en propiedades del producto.

Sean A,B matrices de mxn y C de nxr. Con respecto al -- tamaño, A + B también es de mxn y (A + B)C = F es de mxr. Por otro lado, si D = AC es de mxr y E = BC es de mxr, en--- tonces D + E también es de mxr. Por lo tanto, (A + b)C y --- AC + BC son del mismo tamaño.

Además, por definición el ij-ésimo elemento de (A +B)C - es $[(A + B)C]_{ij}$

$$\begin{aligned} &= (A_i + B_i) \cdot C^{(j)} = (A_i + B_i) \cdot C^{(j)} \\ &= A_i \cdot C^{(j)} + B_i \; C^{(j)} \\ &= (AC)_{ij} + (BC)_{ij} \\ &= (AC + BC)_{ij} . \end{aligned}$$

Es decir, el ij-ésimo elemento de (A + B)C es igual al - ij-ésimo elemento de AC + BC para toda i, j. Por lo tanto -- (A + B)C = AC + BC

d) Sean A_{mxn} y B_{nxk} y $k \in R$. Por definición $(AB)_{mxk}$ y $[k(AB)]_{mxk}$. Por otro lado kA es de mxn y $(kA)B$ es de mxk, entonces $k(AB)$ y $(kA)B$ tienen igual tamaño.

En componentes: $[k(AB)]_{ij} = k(AB)_{ij} = k(A_i \cdot B^{(j)})$
$= (kA_i)_i B_j = (kA)_i \cdot B^{(j)})$
$= [(kA)B]_{ij}$.

Por lo tanto $k(AB) = (kA)B$.

Al trabajar con matrices debemos tener cuidado en no utilizar reglas aritméticas no válidas para matrices, como: $(A + B)^2 = A^2 + 2AB + B^2$, con A,B matrices cuadradas de orden n, pues en general $AB \neq BA$.

Para el cálculo de potencias de matrices tenemos que $A^n = (A^{n-1})A$, $n>2$, donde $A^n = A \cdot A \cdot A \ldots A$ (n veces).

Este tipo de definiciones es muy útil en la programación de computadoras. La computadora se dirige a un ciclo de cálculo repetitivo hasta que se obtiene la potencia deseada de A. Esto es lo que se llama una definición recurrente.

Antes del siguiente ejemplo cabe hacer notar que si A es una matriz cuadrada y r y s son enteros, entonces $A^r A^s = A^{r+s}$ $(A^r)^s = A^{rs}$, $A^0 = I$ (por definición). Las demostraciones son semejantes a las que se han hecho hasta el momento y por lo tanto no las haremos aquí.

Nota: Un espacio vectorial puede tener otra operación más (que llamaremos multiplicación) y si esta nueva operación satisface ciertas propiedades, el sistema se llama Algebra. La múltiplicación de matrices tal como se definió hace que el espacio R_{mxn} de matrices cuadradas sea un álgebra.

Problema 2.-Cadenas de Markov.- En un proceso de Cadena de Markov hay n estados: $s_1, s_2, \ldots, s_n$. En un tiempo dado, el proceso está en uno y solamente uno de los estados dados.

La probabilidad de que el proceso se encuentre en un determinado estado depende únicamente del estado inmediatamente anterior en el que estuvo el proceso. Si P_{ij} denota la probabilidad de transición, o sea, la probabilidad de que el proceso pase del estado j al estado i, entonces $0 \leq P_{ij} \leq 1$ y $P_{1j} + P_{2j} + \ldots + P_{nj} = 1$. (Hay seguridad de que el proceso pase del estado j a cualquier otro estado). La matriz:

$$\begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix}$$

se llama Matriz de Transición. Una cadena de Markov queda determinada al especificar la matriz de transición y el estado inicial. Como ilustración consideremos el

ma:

Suponga que en una cierta actividad laboral cada persona está clasificada en uno solo de los tres estados siguientes:
- Estado 1: Trabajador Profesional (TP)
- Estado 2: Trabajador No Calificado (N)
- Estado 3: Trabajador Calificado (C).

Suponga además que cada persona tiene un hijo y que la matriz de transición (de probabilidad)

$$\begin{matrix} & TP & N & C \\ \begin{matrix} TP \\ N \\ C \end{matrix} & \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.3 & 0.2 & 0.5 \end{bmatrix} \end{matrix}$$

dá la probabilidad de que la próxima generación pase de un estado a otro. Si el 20% de los trabajadores son trabajadores profesionales, el 50% de los trabajadores son trabajadores no calificados, y el 30% restante son trabajadores calificados, encuentre las probabilidades respectivas de que el nieto de un trabajador sea profesional, trabajador no calificado o trabajador calificado.

Solución.- La matriz de transición

$$P = \begin{matrix} & TP & N & C \\ \begin{matrix} \\ \\ \end{matrix} & \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.3 & 0.2 & 0.5 \end{bmatrix} \end{matrix} \begin{matrix} TP \\ N \\ C \end{matrix}$$

puede explicarse como sigue: $P_{11} = 0.6$ significa que existe una probabilidad de 0.6 (60%) de que el hijo de un trabajador profesional sea a su vez trabajador profesional; $P_{32} = 0.2$ significa que hay un 20% de probabilidades de que el hijo de un trabajador no calificado sea un trabajador calificado.

Como lo que se quiere es encontrar las probabilidades respectivas de que el nieto de un trabajador pertenezca al estado 1, 2, ó 3, definimos

$$X_k = \begin{matrix} x_{1,k} \\ x_{2,k} \\ x_{3,k} \end{matrix}$$

como el vector que dá estas probabilidades en la k-ésima generación.

Si el 20% de los trabajadores son profesionales, el 50% son no calificados y el 30% son calificados, entonces

$$X_0 = \begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix}$$

es el vector que dá las probabilidades iniciales. Queremos encontrar X_2, pues el nieto corresponde a la segunda genera-

Consideremos el producto

$$PX_0 = X_1 = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.3 & 0.2 & 0.5 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.28 \\ 0.41 \\ 0.31 \end{bmatrix}$$

En este caso,

$$X_1 = \begin{bmatrix} x_{1,1} \\ x_{2,1} \\ x_{3,1} \end{bmatrix} = \begin{bmatrix} 0.28 \\ 0.41 \\ 0.31 \end{bmatrix}$$

se explica en la siguiente forma, por ejemplo, en la compo--- nente $x_{1,1} = 0.12 + 0.1 + 0.06$, $x_{1,1}$ representa la probabilidad de que el hijo de un trabajador sea profesional; 0.12 da la - probabilidad de que el hijo de un trabajador profesional sea también profesional; 0.1 da la probabilidad de que el hijo de un trabajador no calificado sea profesional; 0.06 da la pro-- babilidad de que el hijo de un trabajador calificado sea profesional.

Esto es así porque cada trabajador, independientemente del estado a que pertenezca, puede tener un hijo que sea profesional, es decir, cada uno tiene una cierta probabilidad de que su hijo sea del estado 1 y la suma de ellas dará la pro-- babilidad total buscada.

Tenemos también que, como las probabilidades 0.2, 0.5 y 0.3 permanecen constantes en cada generación, entonces para - X_2, $X_2 = PX_1 = P(PX_0) = (PP)X_0 = P^2X_0$.

Esto es:

$$X_2 = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.3 & 0.2 & 0.5 \end{bmatrix} \begin{bmatrix} 0.28 \\ 0.41 \\ 0.31 \end{bmatrix} = \begin{bmatrix} 0.312 \\ 0.367 \\ 0.321 \end{bmatrix}$$

que son las probabilidades buscadas. En general, $X_n = P^nX_0$, , con lo que calculando P^n podemos obtener las componentes de - X_n que nos representan las probabilidades de que un trabaja-- dor de la n-ésima generación esté en el estado 1, 2 ó 3.

Nota: mediante ciertas características de una matriz --- llamadas Valores y Vectores Propios se tiene otra forma de -- calcular P^n para problemas de este tipo.

Observación.- En adelante, podemos escribir el producto de varias matrices sin utilizar paréntesis, ya que obtenemos el mismo resultado sin importar cómo se efectúe la multipli-- cación (siempre que no conmutemos ninguna de las matrices).

Nota: Una Matriz de Probabilidad es una matriz cuadrada que tiene dos características:

i) Cada componente es no negativa (es decir ≥ 0).

ii) La suma de los elementos de cada renglón es 1. La ma--- triz de transición del problema anterior es, por lo tanto una matriz de probabilidad.

3.5.- T I P O S E S P E C I A L E S D E M A T R I C E S .

Consideremos ahora ciertas clases de matrices que desempeñan un importante papel en el Algebra Lineal.

Dentro de las matrices cuadradas es importante referirnos a la Diagonal Principal de la matriz $A = (a_{ij})_{n\times n}$, que es el conjunto ordenado de las componentes $(a_{11}, a_{22}, \ldots, a_{nn})$,- ésto es, la Diagonal Principal es la línea de las componentes que comienza en la esquina superior izquierda de la matriz y termina en la esquina inferior derecha.

1.-Matrices Diagonales.- Entre las matrices cuadradas -- juegan un importante papel las llamadas Matrices Diagonales, que son matrices que se caracterizan por la propiedad de que todos los elementos por encima y por debajo de la diagonal -- principal son nulos, ($a_{ij} = 0$ si $i \neq j$), ésto es, matrices -- cuyos elementos diferentes de cero están colocados a lo largo de la diagonal principal. Las matrices diagonales se denotan por $(k_1, k_2, \ldots, k_n)$ tal que

$$(k_1, k_2, \ldots, k_n) = \begin{bmatrix} k_1 & 0 & \ldots & 0 \\ 0 & k_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & k_n \end{bmatrix}.$$

Es claro que la suma de matrices diagonales es otra vez una matriz diagonal y lo mismo sucede con el producto, es decir, si A y B son matrices diagonales entonces AB = C también lo es.

Nota: Una propiedad muy importante de las matrices diagonales es que permiten el cálculo de potencias de ellas sin mayor dificultad.

En el cálculo de estas potencias todo el problema se reduce a calcular las potencias de los elementos diagonales, y ésto puede verse aplicando la operación producto.

2.- Matriz Identidad.- Otra manera de definir a la matriz identidad es describirla como aquella en la cual la diagonal principal está constituída por elementos todos iguales a 1, siendo los demás elementos todos nulos y se denota por I_n. Obsérvese que la matriz identidad es un caso particular de las matrices diagonales con $k_1 = k_2 = \ldots k_n = 1$.

3.- Matrices Escalares.- De modo más general que las anteriores, si todos los números k_i en las matrices diagonales son iguales, la matriz se denomina Escalar y se denota

$$[k] = \begin{bmatrix} k & 0 & \ldots & 0 \\ 0 & k & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & k \end{bmatrix} = k \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix} = kI$$

Esto es, una matriz escalar es simplemente un múltiplo escalar de la matriz identidad.

En casos particulares la multiplicación de matrices puede ser conmutativa. Así, por ejemplo, es fácil ver que las matrices escalares conmutan con todas las matrices cuadradas del mismo orden, pues,

$$\begin{bmatrix} k & 0 & \dots & 0 \\ 0 & k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & k \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{n1} & ka_{n2} & \dots & ka_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} k & 0 & \dots & 0 \\ 0 & k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

Podemos abreviar y decir: si [k] es una matriz escalar tal que [k] = kI, entonces:
[k]A = (kI)A = I(kA) = kA = (Ak)I = A(kI) = A[k].

De todo ésto podemos concluir que el papel especial que juega la matriz identidad en la multiplicación matricial es exactamente el mismo papel que juega el número 1 entre los números. De hecho, IA = AI = A, es decir, es la unidad multiplicativa.

Observación.- Con ésto podemos ver que existe una importante relación entre la multiplicación por un escalar y el producto de matrices.

4.- Matrices Triangulares.- Otra clase particular de matrices cuadradas es la de aquellas matrices tales que todos los elementos por debajo de la diagonal principal son nulos, como en el caso de la matriz

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

Esto, es $a_{ij} = 0$ siempre que $i>j$. Tales matrices se llaman matrices triangulares superiores. Un caso parecido es cuando los ceros están por encima de la diagonal principal; son las llamadas Matrices Triangulares Inferiores:

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

En este caso $a_{ij} = 0$, $i<j$.

5.- Matrices Idempotentes.- Una matriz A se dice Idempotente si $A^2 = A$. Observe que si A es idempotente, entonces $A^n = A$, cualquiera que sea el entero positivo n. Un ejemplo particular de matriz idempotente lo es la matriz identidad.

6.- Matrices Nulipotentes.- Una matriz A se llama Nulipotente de grado p, si existe un entero positivo p tal que $A^p = 0$ y $A^{p-1} \neq 0$, (por supuesto, con A diferente de la matriz nula).

7.- Matrices Simétricas.- Antes de proceder a aclarar lo que es una matriz simétrica, introduciremos una nueva operación singular definida entre matrices llamada Transposición. Cada matriz tiene una matriz correspondiente que, como se verá después, tiene propiedades similares a las de la matriz original.

Sea $A = [\, a_{ij} \,]$ una matriz de mxn. Definimos la Transpuesta de A, denotada A^t, como la matriz de nxm obtenida al intercambiar los renglones y las columnas de A. De forma abreviada denotaremos $A^t = [\, a_{ji} \,]_{nxm}$. En otras palabras, si

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \text{ entonces } A^t = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdot & \cdot & \cdots & \cdot \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

Dicho simplemente, el i-ésimo renglón de A es la i-ésima columna de A^t y la j-ésima columna de A es el j-ésimo renglón de A^t.

Observación.- Es obvio que la transpuesta de una matriz renglón será una matriz columna que contenga los mismos elementos.

Es posible demostrar que:

i) $(A^t)^t = A$.
ii) Si A y B son matrices del mismo tamaño, entonces, $(A + B)^t = A^t + B^t$.
iii) Si A y B son matrices conformables, entonces, $(AB)^t = B^t A^t$.
iv) $(kA)^t = kA^t$, donde k es un escalar.

Estas demostraciones son fáciles de efectuar mediante las definiciones de las operaciones.

Una matriz cuadrada A se llama Simétrica si $A^t = A$. La denominación "simétrica" viene del hecho de que los elementos de A están colocados simétricamente con relación a la diagonal principal.

Nota: Algunas propiedades de las matrices simétricas, -

3.6.- INVERSA DE UNA MATRIZ CUADRADA.

Un concepto importante dentro de la teoría de matrices es el de Inversa de una matriz. En esta sección estudiaremos este concepto y posteriormente lo utilizaremos en la solución de algunos problemas.

3.6.1.- Definición.- Sean A y B matrices de orden n. Si $AB = BA = I$, entonces B se llama la Inversa de A y se denota A^{-1}. Si A tiene una inversa decimos que A es invertible o no singular (o también no degenerada).

Observaciones:
- De la definición concluimos inmediatamente que $(A^{-1})^{-1} = A$. (Si A es invertible).
- Esto no significa de ningún modo que todas las matrices cuadradas forzosamente sean invertibles; podemos toparnos con matrices cuadradas que no lo sean.

Es conveniente preguntarnos si esta inversa, si es que existe, es única o existen muchas. Para esto veremos la siguiente:

Proposición 3.6.2.- Si A es una matriz cuadrada invertible, entonces su inversa es única.

Demostración.- En otras palabras, si B es una inversa de A y C es otra matriz (todas de orden n), tal que $AC = CA = I$, entonces $B = C$. Esto es así, puesto que $C = CI = C(AB) = (CA)B = IB = B$.

Por lo tanto, concluimos que cuando una matriz admite una inversa, esta es única. Otro hecho importante referente a matrices inversas es el siguiente:

Teorema 3.6.3.- Si A y B son matrices de orden n no singulares, entonces AB es no singular y además $(AB)^{-1} = B^{-1}A^{-1}$.

Demostración.- Por hipótesis A^{-1} y B^{-1} existen. Es claro que $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = A(IA^{-1}) = AA^{-1} = I$.
Por otro lado, $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}(IB) = B^{-1}B = I$,
de donde $B^{-1}A^{-1}$ es inversa de AB. Pero como ésta sí existe es única, entonces $(AB)^{-1} = B^{-1}A^{-1}$.

Este resultado es extensivo al producto de n matrices no singulares y puede probarse repitiendo el proceso anterior.

Enseguida veremos una proposición que nos ayudará a reconocer cuando una matriz tiene o no inversa.

Proposición 3.6.4.- Sea A una matriz de orden n. Si existe una matriz no nula X tal que $AX = 0$, entonces A no admite inversa.

Demostración.- Lo haremos demostrando lo contrario, es - decir que si AX = 0 y A es invertible, entonces X = 0 necesariamente. Así, sea AX = 0 y A^{-1} existe. Entonces, $A^{-1}(AX) = A^{-1}0 = 0$, pero $A^{-1}(AX) = (A^{-1}A)X = IX = X = 0$. De - donde, si A es invertible y AX = 0, entonces X = 0.

Observación.- Este resultado se puede relacionar con --- sistemas de ecuaciones pues al considerar la ecuación matri-- cial AX = 0 en realidad nos referimos al sistema homogéneo -- AX = 0. Si ésto se da con $X \neq 0$ significa que dicho sistema homogéneo tendrá varias soluciones. En este caso A^{-1} no existe. Hemos visto que para comprobar que $B = A^{-1}$ debemos de--- mostrar que AB = BA = I. Una forma de simplificar el trabajo de comprobar si una matriz es la inversa de otra esta dada -- por el siguiente teorema:

Teorema 3.5.5.- Sean A y B matrices de nxn. Entonces A es - invertible y $B = A^{-1}$ si:

a) BA = I
b) AB = I.

Demostración.- Para demostrar a) supongamos que BA = I y consideremos la ecuación matricial AX = 0. Entonces B(AX) = B0 = 0 y B(AX) = (BA)X = IX = X = 0, por lo que, por la pro-- posición anterior, X = 0 es la única solución a AX = 0 y A es invertible. Falta demostrar que $B = A^{-1}$. Para ésto, como -- A^{-1} existe, post-multiplicamos BA = I por A^{-1}, y obtenemos que $A^{-1} = (BA)A^{-1} = B(AA^{-1}) = BI = B$, de donde $B = A^{-1}$.

b) Sea AB = I. De a) tenemos que $A = B^{-1}$ y por definición - AB = BA = I, lo que prueba que A es invertible y $B = A^{-1}$.

3.7 I N V E R S I O N D E M A T R I C E S.

A continuación veremos un método para encontrar la in--- versa de una matriz dada.

3.7.1.- Algoritmo para el cálculo de la Inversa de una - matriz mediante el Método de Gauss-Jordan. Una técnica para encontrar la inversa de una matriz, cuando existe, emplea el método de Gauss-Jordan. Esta técnica hace dos cosas a la --- vez: nos dice si la matriz tiene inversa y, al mismo tiempo, nos da esta inversa. Esta técnica o algoritmo consiste en lo siguiente:

Dada una matriz A de orden n formemos la matriz aumentada -- [A | I] escribiendo junto a ella la matriz identidad de or- den n obteniéndose con ello una matriz de nx2n. Apliquemos -- operaciones elementales sobre los renglones de esta matriz -- aumentada para convertir el bloque A en la matriz identidad. Supongamos que de ésto resulte la matriz [I | B]. Entonces $B = A^{-1}$. Si durante el proceso llegamos a una matriz ----- [A' | B] donde A' tiene una línea de ceros, entonces A no es invertible.

Nota 1.- Este método es una especie de Esquema de Divi--

sión Simple para inversión de matrices. (Recuérdese este esquema para la solución de SEL).

Nota 2.- Posteriormente veremos otro método para encontrar la inversa de una matriz mediante determinantes.

La justificación para este algoritmo es la siguiente: Sea A una matriz cuadrada de orden n, invertible y sea X de n×n, la inversa de A: entonces X es una matriz tal que satisface que AX = I, es decir:

$$AX = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n$$

Desarrollando tenemos que:

$$\begin{aligned} a_{11}x_{11} + a_{12}x_{21} + \cdots + a_{1n}x_{n1} &= 1 \\ a_{21}x_{11} + a_{22}x_{21} + \cdots + a_{2n}x_{n1} &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{n1}x_{11} + a_{n2}x_{21} + \cdots + a_{nn}x_{n1} &= 0 \end{aligned}$$

$$\begin{aligned} a_{11}x_{12} + a_{12}x_{22} + \cdots + a_{1n}x_{n2} &= 0 \\ a_{21}x_{12} + a_{22}x_{22} + \cdots + a_{2n}x_{n2} &= 1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{n1}x_{12} + a_{n2}x_{22} + \cdots + a_{nn}x_{n2} &= 0 \end{aligned}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\begin{aligned} a_{11}x_{1n} + a_{12}x_{2n} + \cdots + a_{1n}x_{nn} &= 0 \\ a_{21}x_{1n} + a_{22}x_{2n} + \cdots + a_{2n}x_{nn} &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{n1}x_{1n} + a_{n2}x_{2n} + \cdots + a_{nn}x_{nn} &= 1 \end{aligned}$$

Si cada sistema resultante a resolver lo representamos mediante notación matricial tenemos n sistemas de la forma:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ \cdots \\ x_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} ;$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ \cdots \\ x_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} ;$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{1n} \\ x_{2n} \\ \cdots \\ x_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Obsérvese que todos estos sistemas tienen la misma matriz de coeficientes, por lo tanto podemos resolverlos simul-

táneamente mediante el método de solución para varios SEL tomando a la matriz de coeficientes común a todos los SEL y a-- gregando enseguida los vectores columna de las constantes de cada uno de los n SEL, y ésto da lugar a una matriz de nx2n - elementos,

$$\left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right] = [\, A \mid I \,].$$

Al aplicar operaciones elementales en el bloque A para - escalonarlo y reducirlo (Gauss-Jordan), en el lado derecho -- (bloqueI) van apareciendo las soluciones para cada SEL, es -- decir, el vector solución X tal que satisfaga a su correspondiente SEL. Con ésto en I van apareciendo X y ésto es A^{-1}.

$$\left[\begin{array}{c|ccc} I & X_1 & X_2 \dots & X_n \end{array}\right] = X = A^{-1}$$

Observación.- Para controlar los cálculos y verificar el resultado, es conveniente comprobar que $AA^{-1} = I$.

Veamos un ejemplo:

Ejemplo 8.- Sea

$$\begin{bmatrix} 15 & 10 & 18 \\ 15 & 0 & 12 \\ 10 & 20 & 16 \end{bmatrix}$$

la matriz asociada a un SEL. Queremos determinar la inversa - de esta matriz mediante el método de Gauss-Jordan.

Solución.- Siguiendo el algoritmo representado anteriormente tenemos:

$$\left[\begin{array}{ccc|ccc} 15 & 10 & 18 & 1 & 0 & 0 \\ 15 & 0 & 12 & 1 & 0 & 1 \\ 10 & 20 & 16 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_{13}} \left[\begin{array}{ccc|ccc} 10 & 20 & 16 & 0 & 0 & 1 \\ 15 & 0 & 12 & 0 & 1 & 0 \\ 15 & 10 & 18 & 1 & 0 & 0 \end{array}\right]$$

$$\begin{array}{l} 1/10\ R_1 \\ -15\ R_1 + R_2 \\ -15\ R_1 + R_3 \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & 8/5 & 0 & 0 & 1/10 \\ 0 & -30 & -12 & 0 & 1 & -3/2 \\ 0 & -20 & -6 & 1 & 0 & -3/2 \end{array}\right]$$

$$\begin{array}{l} -1/30\ R_2 \\ -2\ R_2 + R_1 \\ -2R_2 + R_3 \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 4/5 & 0 & 1/15 & 0 \\ 0 & 1 & 2/5 & 0 & -1/30 & 1/20 \\ 0 & 0 & 2 & 1 & -2/3 & -1/2 \end{array}\right]$$

$$\begin{array}{l} 1/2\ R_3 \\ -4/5\ R_3 + R_1 \\ -2/5\ R_3 + R_2 \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2/5 & 1/3 & 1/5 \\ 0 & 1 & 0 & -1/5 & 1/10 & 3/20 \\ 0 & 0 & 1 & 1/2 & -1/3 & -1/4 \end{array}\right]$$

De aquí que

$$A^{-1} = \begin{bmatrix} -2/5 & 1/3 & 1/5 \\ -1/5 & 1/10 & 3/20 \\ 1/2 & -1/3 & -1/4 \end{bmatrix}$$

3.8 A L G O M A S A C E R C A D E S E L .

En esta sección estableceremos algunos resultados también importantes a sistemas de ecuaciones (SEL) con los que daremos otro método de solución para sistemas de n ecuaciones con n incógnitas (i.e. en cierto tipo de problemas el más eficiente) y finalmente daremos un criterio de

Primero, indicamos que si un SEL no tiene solución.

Teorema 3.8 : Si A es una matriz invertible de tamaño nxn, entonces para cada matriz de nx1 el SEL AX = B tiene exactamente una solución a saber $A^{-1}B$.

Demostración: Ya que $AA^{-1} = I$ para cualquier matriz cuadrada A es invertible, tenemos en particular que $A(A^{-1}B) = AA^{-1}B = IB = B$, de modo que $A^{-1}B$ es una solución de AX = B. Falta mostrar que es la única solución. Supongamos que X_0 es cualquier solución. Vamos a demostrar que X_0 tiene que ser la solución $A^{-1}B$. Si X_0 es una solución, entonces $AX_0 = B$. Premultiplicando por A^{-1} en ambos lados tenemos que $AA^{-1}X_0 = A^{-1}B$ o también $X_0 = A^{-1}B$.

Podemos ahora resumir mediante alguna relación todas las propiedades presentadas anteriormente :

a) A es invertible.
b) AX = 0 tiene únicamente la solución trivial.
c) A es equivalente por renglones a I.
d) AX = B es consistente para toda matriz de nx1.

De aquí podemos concluir que si A es invertible entonces el sistema asociado tiene solución única; en caso contrario, el sistema tendrá varias soluciones o no existirá solución.

Ejemplo 3.- A partir de la información dada en el ejemplo anterior obtenemos el siguiente SEL a resolver :

$$\begin{aligned} 15x_1 - 11x_2 - 12x_3 &= 1{,}01 \\ 2x_1 - 2x_2 - 1x_3 &= 0{,}10 \\ 1x_1 - 1x_2 + 1x_3 &= 0{,}01 \end{aligned}$$

Para resolver este sistema utilizando la matriz

De aquí que

$$A^{-1} = \begin{bmatrix} -2/5 & 1/3 & 1/5 \\ -1/5 & 1/10 & 3/20 \\ 1/2 & -1/3 & -1/4 \end{bmatrix}$$

3.8 A L G O M A S A C E R C A D E S E L .

En esta sección estableceremos algunos resultados también concernientes a sistemas de ecuaciones (SEL) con lo que daremos otro método de solución para sistemas de n ecuaciones con n incógnitas que, en cierta clase de problemas, son más eficientes que la eliminación de Gauss.

Además, analizaremos qué sucede con los SEL que no tienen solución.

Teorema 3.8.1.- Si A es una matriz invertible de nxn, entonces, para cada matriz de nx1, el SEL AX = B tiene exactamente una solución, a saber $X = A^{-1} B$.

Demostración.- Ya que $A(A^{-1} C) = C$ para cualquier matriz C cuando A es invertible, tenemos entonces en particular que $A(A^{-1}B)=B$, de donde $X = A^{-1} B$ es una solución de AX = B. Para mostrar que es la única solución supongamos que X_0 es cualquier solución. Vamos a demostrar que X_0 tiene que ser la solución $A^{-1} B$. Si X_0 es una solución, entonces $AX_0 = B$. Premultiplicando por A^{-1} en ambos lados tenemos que $AA^{-1} X_0 = A^{-1} B$, entonces $X_0 = A^{-1} B$.

Podemos ahora resumir mediante alguna relación todas las propiedades presentadas anteriormente :

a) A es invertible.
b) AX = 0 tiene únicamente la solución trivial.
c) A es equivalente por renglones a I.
d) AX = B es consistente para toda matriz de nx1.

De aquí podemos concluir que si A es invertible, entonces su sistema asociado tiene solución única; en caso contrario, el sistema tendrá varias soluciones o no existirá solución.

Problema 3.- A partir de la información dada en el problema obtenemos el siguiente SEL a resolver :

$$\begin{aligned} 15x + 10y + 18z &= 1800 \\ 10x + \quad\; + 12z &= 1800 \\ 10x + 20y + 16z &= 1200 \end{aligned}$$

Este sistema mediante notación matricial lo representa-

mos como :

$$\begin{bmatrix} 15 & 10 & 18 \\ 15 & 0 & 12 \\ 10 & 20 & 16 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1800 \\ 1800 \\ 1200 \end{bmatrix} .$$

Observe que la matriz de coeficientes del sistema ya es conocida por nosotros ; su inversa existe y es

$$A^{-1} = \begin{bmatrix} -2/5 & 1/3 & 1/5 \\ -1/5 & 1/10 & 3/20 \\ 1/2 & -1/3 & -1/4 \end{bmatrix} .$$

De donde la solución estará dada por :

$$X = A^{-1} B = \begin{bmatrix} -2/5 & 1/3 & 1/5 \\ -1/5 & 1/10 & 3/20 \\ 1/2 & -1/3 & -1/4 \end{bmatrix} \begin{bmatrix} 1800 \\ 1800 \\ 1200 \end{bmatrix} = \begin{bmatrix} 120 \\ 0 \\ 0 \end{bmatrix} .$$

Esto es, se pueden producir 120 unidades del fertilizante de tipo I solamente.

Problema 4.- Alimentación y dietas.- Si x,y,z nos representan a cada comida (menú) respectivamente, interpretando el problema obtenemos el sistema

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

donde $x,y,z > 0$. En notación matricial :

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

Mediante A^{-1} :

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 4 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{-4R_1 + R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & 0 & -3 & -4 & 0 & 1 \end{array}\right] \xrightarrow[-1/3\ R_3]{1/2\ R_2}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 4/3 & 0 & -1/3 \end{array}\right] \xrightarrow[-3/2\,R_3 + R_2]{-R_3 + R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/3 & 0 & 1/3 \\ 0 & 1 & 0 & -2 & 1/2 & 1/2 \\ 0 & 0 & 1 & 4/3 & 0 & -1/3 \end{array}\right]$$

Entonces

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

y

$$X = A^{-1} B = \begin{bmatrix} -1/3 & 0 & 1/3 \\ -2 & 1/2 & 1/2 \\ 4/3 & 0 & -1/3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/2 \\ 1/3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

La respuesta es : 2/3 partes del menú 1, la mitad del 2 y una tercera parte del 3.

Problema 5.- En este problema los sectores o compañías se interrelacionan de una forma elaborada. Se hace una distinción entre las compañías o sectores (que convierten un producto en otros productos de tal manera que los productos suministrados no son realmente consumidos, sino reemplazados en el mercado por los productos de la producción)y el público considerado como los "consumidores finales", (que en este caso corresponde a ahorro).

Leontief trabaja con dos tipos de modelos: cerrado y abierto. En un modelo cerrado, el público se trata como una compañía o sector más cuyo ingreso es la canasta de mercado, mientras que la producción es el ingrediente trabajo en el suministro de más sectores tradicionales. La distinción entre "buenos consumos finales" en nuestro modelo abierto y los ingresos que el sector construcción "procesa" en un producto de salida llamado mano de obra en un modelo cerrado es de interés económico pero no presenta alguna diferencia matemática.

En el capítulo de Sistemas de Ecuaciones Lineales se resolvió un problema representado por un modelo cerrado de Leontief en el que se enfatizó que en este tipo de modelo económico se trabajaría suponiendo igualdad entre la producción y el consumo. Aquí supondríamos lo mismo.

El arreglo dado se interpreta como sigue: el sector de la agricultura produce un total de 20 unidades, de las cuales se autosuministra con 4, la manufactura consume 6 y el sector

de la construcción consume 3; las 7 restantes son los productos de inversión (que pasan a formar los fondos invertidos o ahorros y que son consumidos por el sector de la construcción). El mismo sector, agricultura, consume 4 unidades de su propia construcción, 8 de la manufactura y 8 del sector de la construcción, 20 unidades en total.

Como en este caso el sector ahorros es sector abierto, el problema se interpreta como un modelo de sectores: agricultura, manufactura y construcción, pero todavía no corresponde a un modelo de Leontief. Lo convertiremos en uno en la siguiente forma: los insumos (consumos) de la agricultura es 4/20 de ella misma, 8/20 de la manufactura y 8/20 de la construcción, etc.

Con ésto resulta que, en este problema, necesita satisfacerse cierta demanda, que en este caso es la inversión, de manera que se formen los ahorros y sean utilizados por el sector de la construcción. En otras palabras, el vector demanda d tiene como elementos $\begin{bmatrix}7\\3\\2\end{bmatrix}$, y la matriz a trabajar será:

$$A = \begin{bmatrix} 4/20 & 6/30 & 3/20 \\ 8/20 & 18/30 & 1/20 \\ 8/20 & 6/30 & 4/20 \end{bmatrix} = \begin{bmatrix} 1/5 & 1/5 & 3/20 \\ 2/5 & 3/5 & 1/20 \\ 2/5 & 1/5 & 1/5 \end{bmatrix}.$$

Lo resolveremos mediante $(I-A)^{-1}d = X$, donde X = vector producción. Así:

$$I-A = \begin{bmatrix} 4/5 & -1/5 & -3/20 \\ -2/5 & 2/5 & -1/20 \\ -2/5 & -1/5 & 4/5 \end{bmatrix}.$$

$$(I-A)^{-1} = \left[\begin{array}{ccc|ccc} 4/5 & -1/5 & -3/20 & 1 & 0 & 0 \\ -2/5 & 2/5 & -1/20 & 0 & 1 & 0 \\ -2/5 & -1/5 & 4/5 & 0 & 0 & 1 \end{array}\right]$$

$$\begin{array}{l} 5/4\ R_1 \\ \underline{2/5\ R_1 + R_2} \\ 2/5\ R_1 + R_3 \end{array} \left[\begin{array}{ccc|ccc} 1 & -1/4 & -3/16 & 5/4 & 0 & 0 \\ 0 & 3/10 & -1/8 & 1/2 & 1 & 0 \\ 0 & -3/10 & 29/40 & 1/2 & 0 & 1 \end{array}\right]$$

$$\begin{array}{l} 10/3\ R_2 \\ \underline{1/4\ R_2 + R_1} \\ 3/10\ R_2 + R_3 \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & -7/24 & 5/3 & 5/6 & 0 \\ 0 & 1 & -5/12 & 5/3 & 10/3 & 0 \\ 0 & 0 & 3/5 & 1 & 1 & 1 \end{array}\right]$$

$$\begin{matrix} 5/3\ R_3 \\ 7/24\ R_3 + R_1 \\ 5/12\ R_3 + R_2 \end{matrix} \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 155/72 & 95/72 & 35/72 \\ 0 & 1 & 0 & 85/36 & 145/36 & 25/36 \\ 0 & 0 & 1 & 5/3 & 5/3 & 5/3 \end{array}\right]$$

De donde

$$(I - A)^{-1} = \left[\begin{array}{ccc} 155/72 & 95/72 & 35/72 \\ 85/36 & 145/36 & 25/36 \\ 5/3 & 5/3 & 5/3 \end{array}\right]$$

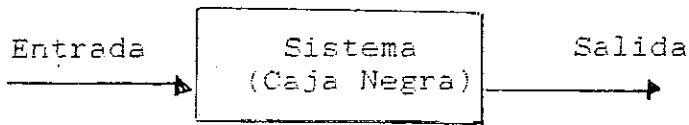
$$= 1/72 \left[\begin{array}{ccc} 155 & 95 & 35 \\ 170 & 290 & 50 \\ 120 & 120 & 120 \end{array}\right]$$

Por lo tanto,

$$X = (I - A)^{-1} d = 1/72 \left[\begin{array}{ccc} 155 & 95 & 35 \\ 170 & 290 & 50 \\ 120 & 120 & 120 \end{array}\right] \left[\begin{array}{c} 7 \\ 3 \\ 2 \end{array}\right] = \left[\begin{array}{c} 20 \\ 30 \\ 20 \end{array}\right]$$

Este resultado era de esperarse debido a la suposición hecha y a los datos dados en la tabla.

Finalmente, sólo comentaremos dos últimos problemas. Podemos considerar un tipo de problema muy importante en Física: en ciertas aplicaciones se trabaja con sistemas físicos llamados "cajas negras", por haber sido reducidos a sus características esenciales. En ellas consideramos que si al sistema le aplicamos una cierta entrada, el sistema producirá una cierta salida; sin embargo, el funcionamiento interno del sistema se desconoce o no importa, de ahí el término "caja negra". En casos como éste, tanto la entrada como la salida podemos describirlas como matrices de una columna. Por ejemplo, si la caja negra está formada por un cierto circuito electrónico, entonces la entrada puede ser una matriz de nx1 cuyos elementos son los n voltajes medidos entre ciertas terminales de entrada, y la salida puede ser una matriz de mx1 cuyos elementos son las corrientes resultantes en m cables de salida. En términos matemáticos lo que hace el sistema es transformar una matriz de entrada de nx1 en una matriz de salida de mx1.



Para un gran número de sistemas de este tipo la relación es como sigue: al sistema le asociamos una cierta matriz ---- constante A de mxn, cuyos elementos son parámetros físicos --

determinados por el sistema. La matriz de salida B de mx1, - que resulta de una matriz de entrada C de nx1, se puede obtener mediante B = AC.

Un sistema de este tipo es un ejemplo de sistema físico lineal. En las aplicaciones es importante con frecuencia determinar qué entrada C se le debe aplicar al sistema para --- conseguir cierta salida B, es decir, resolver la ecuación --- AX = B para la entrada incógnita X dada la salida B requerida.

Un último problema fundamental consiste en lo siguiente: Sea A una matriz constante de mxn. Encuentre todas las matrices B de mx1, tales que el sistema AX = B sea consistente.

Si A es una matriz invertible, el teorema 3.8.1 resuelve en forma completa el problema al afirmar que para toda matriz B de mx1, AX = B tiene la solución única $X = A^{-1}B$. Si A no es cuadrada o no es invertible sería muy conveniente determinar en qué condiciones el sistema AX =B es consistente y ésto --- puede hacerse utilizando alguno de los métodos de reducción - desarrollados en el capítulo anterior (Gauss, Gauss-Jordan).

C A P I T U L O 4
ESPACIOS VECTORIALES

4.1 I N T R O D U C C I O N .

Como se ha observado en los capítulos anteriores, existen ciertos conjuntos de objetos (físicos o matemáticos) que, bajo dos operaciones definidas en determinada forma, tienen una serie de propiedades en común : asociatividad, distributividad, conmutatividad, etc. Casos particulares de estos conjuntos se han estudiado ya con anterioridad : vectores, matrices e incluso el conjunto solución de sistemas de ecuaciones lineales homogéneos. Como se vé son conjuntos de elementos bastante diferentes que lo único que tienen en común son algunos conceptos : una operación suma, una operación multiplicación por un escalar y las propiedades mencionadas.

Esto nos lleva a una clasificación especial para este tipo de conjuntos en la cual generalizamos varios conceptos (entre ellos el de vector) a espacios más abstractos donde lo que nos interesa son las propiedades analíticas que se satisfacen y no las geométricas, y en esta forma podemos visualizar el comportamiento de los elementos de estos conjuntos, sin graficarlos, en R^n. La teoría completa se desarrolla sin hacer referencia exclusivamente a n-adas ; esto nos permite que polinomios o funciones, por ejemplo, puedan manejarse de manera semejante a ellas.

En otros textos este material se desarrolla antes de sistemas de ecuaciones, pero nosotros alteramos el orden y lo cambiamos de lugar puesto que, como se verá más adelante, varias definiciones incluyen material correspondiente a solución de sistemas de ecuaciones por lo que preferimos desarrollar primero toda la teoría concerniente a este punto y después basarnos en ella para comprender y aplicar sin dificultad estos nuevos conceptos. Aprovecharemos para completar lo que faltó desarrollar en sistemas de ecuaciones, que es lo que corresponde a sistemas de ecuaciones lineales sin solución.

Debemos aclarar que el material contenido en este capítulo es más complejo que el de los anteriores debido a su naturaleza propiamente teórica, pero, aún así, ilustraremos con algunos ejemplos en base a temas ya conocidos para tratar de aclarar las ideas vertidas en él.

Sin embargo, se tiene la gran ventaja de que toda la teoría que se desarrolla para este tipo especial de conjuntos también será válida para cada ejemplo particular de ellos y no tendrá que comprobarse cada vez que se necesite.

En este capítulo se formalizan algunos conceptos utilizados anteriormente y se refuerzan con otros más para confor-

de Algebra Lineal I. Aquí se contempla toda la justificación teórica que respalda a lo desarrollado en temas anteriores.

4.2 R E S E Ñ A H I S T O R I C A

Es al llamado "Príncipe de las Matemáticas", Karl Friedrich Gauss (1777-1855), al que se le atribuyen las primeras ideas acerca del concepto de n dimensiones, así como el método para sumar vectores de dos componentes : esto último debido a que aparecen cuando dá la demostración del Teorema Fundamental del Algebra, al identificar puntos del plano con números complejos.

Posteriores a él, encontramos los trabajos de Arthur Cayley (1821-1895) y de Hermann Gunther Grossman (1809-1877), en los cuales ya se maneja el concepto de espacio vectorial.

En las publicaciones de Grossman encontramos además nociones como subespacio, independencia lineal, conjunto de generadores, dimensión, suma e intersección de subespacios.

La vida de Grossman es algo curiosa, porque a pesar de ser un gran matemático, tardíamente se dió valor a su obra, quizá debido a la dificultad de la lectura de su trabajo. Prácticamente la totalidad de su vida la pasó como maestro de secundaria en Stettin, pequeña ciudad de Pomerania situada a poca distancia del Mar Báltico.

Giuseppe Peano (1858-1932), lógico y matemático italiano, en 1888 da la definición axiomática de un espacio vectorial con escalares reales empleando una notación moderna.

Continuaremos con algunos resultados importantes para el caso R^n y después generalizaremos al caso abstracto.

4.3 C O M B I N A C I O N L I N E A L D E V E C T O R E S, D E P E N D E N C I A E I N D E P E N D E N C I A L I N E A L D E V E C T O R E S Y G E N E R A D O R E S : Caso R^n.

En esta sección estudiaremos algunos conceptos que nos servirán para después obtener a partir de ellos una definición muy importante : la de base. Observaremos cómo estos conceptos están muy relacionados unos con otros, existe una seriación entre ellos de tal forma que trataremos de que el último concepto que se presente quede lo bastante firme como para comprenderse sin mucha dificultad. Empecemos.

Como se recordará, las operaciones multiplicación por un escalar y suma de vectores, para vectores en R^2 o en R^3, tienen un significado geométrico : si r, es un escalar y V, es un vector en R^2 o R^3, ...

mos llegar a cualquier punto que esté colocado sobre la misma dirección de V_1. Este punto determina otro vector con la misma dirección que V_1 pero diferente magnitud.

Si queremos alcanzar otro punto fuera de esta dirección podemos tomar otro vector V_2 (por supuesto no colineal a V_1) de manera que V_1 y V_2 forman un paralelogramo (o plano), cuya resultante está definida por $V_1 + V_2$, y contiene a ambos vectores.

Tomando múltiplos convenientes de V_1 y V_2 podemos alcanzar cualquier punto en este plano mediante $r_1V_1 + r_2V_2$, con r_1, r_2 escalares, pero no podemos movernos fuera de este plano. Para llegar a un punto fuera de este plano podemos introducir un tercer vector V_3 que no esté contenido en el plano definido por V_1 y V_2 (es decir, V_3 no coplanar a V_1 y V_2). Análogamente, tomando sumas de múltiplos de la forma $r_1V_1 + r_2V_2 + r_3V_3$, con r_1, r_2, r_3 escalares, podemos llegar a cualquier punto en R^3 que determina a un cierto vector.

Generalizando, si tenemos n vectores fijos $V_1, V_2, \ldots, V_n \in R^n$ y queremos llegar a un cierto punto que determina a un vector $V \in R^n$, podemos hacerlo mediante la suma de múltiplos escalares convenientes de $V_1, V_2, \ldots, V_n$, es decir :
$V = r_1V_1 + r_2V_2 + \ldots + r_nV_n$, con $r_1, r_2, \ldots, r_n$ escalares.
Por el hecho de poder representar a V de esta forma, decimos que V es una Combinación Lineal de $V_1, V_2, \ldots, V_n$. Así, enunciamos la siguiente

Definición 4.3.1.- Sean $V_1, V_2, \ldots, V_n$ vectores fijos en R^n. Un vector V en R^n es una COMBINACION LINEAL de $V_1, V_2, \ldots, V_n$ si podemos expresarlo como

$$V = r_1V_1 + r_2V_2 + \ldots + r_nV_n$$

para algunos escalares $r_1, r_2, \ldots, r_n$.

Observación.- Nótese que las combinaciones lineales de vectores también son vectores. Asímismo, el verificar si un vector V es una combinación lineal de los vectores $V_1, V_2, \ldots V_n$ nos lleva a resolver un sistema de ecuaciones.

Si variamos indistintamente los valores de $r_1, r_2, \ldots, r_n$ podemos formar vectores distintos ; si reunimos todos estos vectores que se obtienen mediante combinaciones lineales de $V_1, V_2, \ldots, V_n$ y lo denotamos por $\mathcal{L}(V_1, V_2, \ldots, V_n)$, es decir si

$$\mathcal{L}(V_1, V_2, \ldots, V_n) = \{ V \in R^n / V = r_1V_1 + r_2V_2 + \ldots + r_nV_n, \text{ con } r_1, r_2, \ldots, r_n \text{ escalares} \},$$

resulta que este conjunto contiene, además de todas las combinaciones lineales de $V_1, V_2, \ldots, V_n$ a los mismos $V_1, V_2, \ldots V_n$, para valores adecuados de los escalares $r_1, r_2, \ldots, r_n$.

Este conjunto lo utilizaremos con frecuencia más adelante y nos llevará a otros conceptos muy importantes. En el

en el concepto de combinacion lineal de vectores es la de dependencia e independencia lineal de vectores ; más adelante veremos cómo este concepto es central también en la teoría general de conjuntos abstractos. Es de esperarse que estos conceptos estén mucho muy relacionados con sistemas de ecuaciones y la existencia de los escalares $r_1, r_2, \ldots, r_m$.

Recordemos la forma de un sistema de mxn :

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

Este sistema puede reescribirse en la siguiente forma :

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

es decir,

$$x_1V_1 + x_2V_2 + \ldots + x_nV_n = b \text{ , donde } V_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{mi} \end{pmatrix}, \quad i = 1, 2, \ldots n.$$

son vectores columna en R^m. Lo que hemos hecho entonces es escribir al sistema como combinación lineal de los vectores $V_1, V_2, \ldots, V_n$. Cuando el sistema es consistente es porque existen las constantes $x_1, x_2, \ldots, x_n$ que permiten hacer lo anterior, es decir $b \in \mathcal{L}(V_1, V_2, \ldots, V_n)$; si no es posible encontrar esas constantes es porque el sistema es inconsistente y $b \notin \mathcal{L}(V_1, V_2, \ldots, V_n)$.

Pero además existe una relación muy estrecha entre las soluciones de un sistema de ecuaciones lineales no homogéneo y su homogéneo asociado (teorema 2.6.4) ; resulta que en la solución del sistema $x_1V_1 + x_2V_2 + \ldots + x_nV_n = b$, el comportamiento que tenga el sistema homogéneo $x_1V_1 + x_2V_2 + \ldots + x_nV_n = 0$ es muy importante, pues :

Proposición 4.3.2.- Si el sistema

$$x_1V_1 + x_2V_2 + \ldots + x_nV_n = b$$

tiene solución, entonces tiene infinitas soluciones si y sólo si el sistema $x_1V_1 + x_2V_2 + \ldots + x_nV_n = 0$ tiene solución no trivial.

Demostración.- Primero verificaremos que si el sistema $x_1V_1 + x_2V_2 + \ldots + x_nV_n = b$ tiene varias soluciones, entonces el sistema homogéneo $x_1V_1 + x_2V_2 + \ldots + x_nV_n = 0$ tiene solución no trivial, es decir, existen constantes $r_1, r_2, \ldots, r_n$ no todas cero tales que

$$r_1V_1 + r_2V_2 + \ldots + r_nV_n = 0$$

Si el sistema $x_1V_1 + x_2V_2 + \ldots + x_nV_n = b$ tiene varias

rias formas distintas en términos de $V_1, V_2, \ldots, V_n$. Sean las siguientes dos de estas expresiones :

$$\begin{array}{l} x_1V_1 + x_2V_2 + \ldots + x_nV_n = b \\ y_1V_1 + y_2V_2 + \ldots + y_nV_n = b \end{array} \Rightarrow \begin{array}{r} x_1V_1 + x_2V_2 + \ldots + x_nV_n = b \\ -y_1V_1 - y_2V_2 - \ldots - y_nV_n = -b \end{array}$$

$$\underbrace{(x_1-y_1)}_{r_1}V + \underbrace{(x_2-y_2)}_{r_2}V + \ldots + \underbrace{(x_n-y_n)}_{r_n}V = 0$$

donde $r_i \neq 0$, pues $x_i \neq y_i$ para alguna i. Así, $r_1V_1 + r_2V_2 + \ldots + r_nV_n = 0$, con alguna $r_i \neq 0$.

En el otro sentido, si el sistema

$$x_1V_1 + x_2V_2 + \ldots + x_nV_n = 0$$

tiene solución no trivial forzosamente tiene varias soluciones. Si además, el sistema $x_1V_1 + x_2V_2 + \ldots + x_nV_n = b$ tiene solución, el que dicho sistema tiene varias soluciones puede concluirse a partir del teorema 2.6.4.

Observación.- En otro caso, si $r_1 = r_2 = \ldots = r_n = 0$ significa que la forma de expresar el vector b en términos de $V_1, V_2, \ldots, V_n$ es única y que la solución del sistema homogéneo $x_1V_1 + x_2V_2 + \ldots + x_nV_n = 0$ deberá ser trivial, para que sea válido el teorema 2.6.4.

Por otro lado, supongamos que existen escalares no todos cero tales que $r_1V_1 + r_2V_2 + \ldots + r_nV_n = 0$ y tomemos $r_1 \neq 0$. Reescribiendo el sistema tenemos que

$$V_1 = (-r_2/r_1)\,V_2 + (-r_3/r_1)\,V_3 + \ldots + (-r_n/r_1)\,V_n,$$

es decir, podemos expresar un vector del conjunto como combinación lineal de los restantes. Sin embargo, si $r_1 = r_2 = \ldots = r_n = 0$, no podemos hacer lo anterior. En el primer caso podemos decir que V_1 depende de $V_2, V_3, \ldots, V_n$, en otras palabras $V_1 \in \mathcal{L}(V_2, V_3, \ldots, V_n)$ y por lo tanto $V_1, V_2, \ldots, V_n$ son coplanares, esto es $\mathcal{L}(V_1, V_2, \ldots, V_n) = \mathcal{L}(V_2, V_3, \ldots, V_n)$ es definición de coplanares ; en cambio, en el segundo caso no lo son. Esto nos permite enunciar la siguiente

Definición 4.3.3.- Diremos que $V_1, V_2, \ldots, V_n$ son LINEALMENTE DEPENDIENTES (L.D.) si en el conjunto existe algún vector que se escribe como combinación lineal de los restantes. En caso contrario, diremos que son LINEALMENTE INDEPENDIENTES (L.I).

En lo que respecta al conjunto $\mathcal{L}(V_1, V_2, \ldots, V_n)$ lo anterior corresponde a verificar si el vector 0 pertenece a él o, dicho de otro modo, si podemos expresar el vector 0 como combinación lineal de $V_1, V_2, \ldots, V_n$ y resultó que de ser así esta forma de expresarlo puede ser única o no, dependiendo de los valores que tomen los escalares $r_1, r_2, \ldots, r_n$. Desde este punto de vista, podemos reescribir la definición anterior como sigue :

Definición 4.3.3'.- El conjunto de vectores $V_1, V_2, \ldots, V_n$ es L.D. si existen escalares $r_1, r_2, \ldots, r_n$ no todos cero tales que $r_1V_1 + r_2V_2 + \ldots + r_nV_n = 0$

Si $V_1, V_2, \ldots, V_m$ no son L.D., entonces son L.I.

Existen varias técnicas para determinar si un conjunto de vectores son L.I. o L.D. Una de ellas se relaciona con la solución de sistemas de ecuaciones ; esto lo podemos ver como sigue :

Sean los vectores $V_1 = (x_1, x_2, \ldots, x_m)$, $V_2 = (y_1, y_2, \ldots y_m), \ldots, V_m = (z_1, z_2, \ldots, z_m) \in R^m$. Para saber si dichos vectores son L.I. o L.D. necesitamos determinar los valores de las constantes $r_1, r_2, \ldots, r_m$ que permitan que

$$r_1 V_1 + r_2 V_2 + \ldots + r_m V_m = 0 ,$$

esto es :

$$r_1(x_1, x_2, \ldots, x_m) + r_2(y_1, y_2, \ldots, y_m) + \ldots + r_m(z_1, z_2, \ldots, z_m) = (0 \ldots 0)$$

Esto se reduce a resolver el sistema homogéneo

$$\begin{aligned} r_1 x_1 + r_2 y_1 + \ldots + r_m z_1 &= 0 \\ r_1 x_2 + r_2 y_2 + \ldots + r_m z_2 &= 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ r_1 x_m + r_2 y_m + \ldots + r_m z_m &= 0 \end{aligned}$$

Note que el sistema resultante será el mismo si $V_1, V_2, \ldots, V_m$ son vectores columna.

El que un conjunto de vectores resulte ser L.I. indica que la solución del sistema homogéneo asociado a ellos es única y, por tanto, trivial ; el que dicho conjunto de vectores sea L.D. significa que el sistema homogéneo asociado tiene una infinidad de soluciones.

La definición anterior, aunada a la observación posterior a la proposición 4.3.2, nos sugiere la siguiente

Definición 4.3.4.- Si el sistema $x_1V_1 + x_2V_2 + \ldots + x_mV_m = b$ tiene solución, entonces tiene una única solución si y sólo si los vectores son L.I.

Otra forma de referirnos a esto es que si $V_1, V_2, \ldots, V_m$ son L.I., entonces existe un único conjunto de constantes $x_1, x_2, \ldots, x_m$ tales que $x_1V_1 + x_2V_2 + \ldots + x_mV_m = b$, es decir, se tiene una única forma de expresar el vector b en términos de $V_1, V_2, \ldots, V_m$. Si son L.D. hay infinitos conjuntos de constantes que expresan a b como combinación lineal de $V_1, V_2, \ldots, V_m$, en caso de existir alguna.

En base a la definición de dependencia lineal de vectores se obtiene el siguiente resultado :

Proposición 4.3.5.- $V_1, V_2, \ldots, V_m$ son L.D. si y sólo si existen constantes no todas cero que dan la combinación lineal cero.

Demostración.- Veremos primeramente que si $V_1, V_2, \ldots,$ son L.D. entonces existen constantes no todas cero que

tir de la definición 4.3.3, pues si estos vectores son L.D. - entonces uno de ellos puede escribirse como combinación li--- neal de los restantes ; esto es, si $V_1 = a_2V_2 + a_3V_3 + \ldots + a_nV_n$, entonces $a_2V_2 + a_3V_3 + \ldots + a_nV_n - V_1 = 0$, con el coeficiente de - V_1 distinto de cero.

En el otro sentido, si existen constantes $r_1, r_2, \ldots, r_n$ no todas cero tales que $r_1V_1 + r_2V_2 + \ldots + r_nV_n = 0$, entonces $V_1, V_2, \ldots, V_n$ son L.D., es exactamente la definición 4.3.3'.

Observación.- Esta proposición valida el que la defini- ción 4.3.3 se puede reescribir como 4.3.3'.

Por otro lado, ¿qué sucede cuando un cierto vector V --- puede escribirse en términos de n vectores L.D.?. Veamos.

Sean $V_1, V_2, \ldots, V_n$ vectores L.D., entonces existen ---- constantes no todas cero $r_1, r_2, \ldots, r_n$ tales que $r_1V_1 + r_2V_2 + \ldots + r_nV_n = 0$ pero además, por la definición 4.3.3, uno de es- tos vectores puede expresarse como combinación lineal de los restantes, digamos

$$V_1 = \underbrace{a_2V_2 + a_3V_3 + \ldots + a_nV_n}_{(n-1) \text{ vectores}} \qquad (1)$$

Sea V un vector tal que

$$V = x_1V_1 + x_2V_2 + \ldots + x_nV_n. \qquad (2)$$

Sustituyendo (1) en (2) tenemos que

$$\begin{aligned} V &= x_1(a_2V_2 + a_3V_3 + \ldots + a_nV_n) + x_2V_2 + \ldots + x_nV_n \\ &= (x_1a_2 + x_2)V_2 + (x_1a_3 + x_3)V_3 + \ldots + (x_1a_n + x_n)V_n. \end{aligned}$$

Es decir, pudimos escribir el vector V en términos de -- (n-1) vectores. Esto lo afirma la siguiente

Proposición 4.3.6.- Si $V_1, V_2, \ldots, V_n$ son L.D. y un vector - se escribe como combinación lineal de estos n vectores, en--- tonces existen (n-1) vectores con los que se puede escribir - dicho vector.

Si tomamos a todos los vectores V que puedan expresarse de esta forma, es decir, si tomamos el conjunto $\mathcal{L}(V_1, V_2, \ldots, V_n)$ con $V_1, V_2, \ldots, V_n$ L.D. los resultados anteriores sugie-- ren como consecuencia de ellos los dos siguientes

Corolario 4.3.7.- Si $V_1, V_2, \ldots, V_n$ son L.D., entonces cual- quier vector en $\mathcal{L}(V_1, V_2, \ldots, V_n)$ se puede escribir con ---- (n-1) vectores. Es más, si $V_1, V_2, \ldots, V_n$ son L.D., entonces $\mathcal{L}(V_1, V_2, \ldots, V_n)$ se puede escribir con (n-1) vectores, es decir $\mathcal{L}(V_1, V_2, \ldots, V_n) = \mathcal{L}(V_2, V_3, \ldots, V_n)$.

Corolario 4.3.8.- Si $\mathcal{L}(V_1, V_2, \ldots, V_n)$ se escribe con menos vectores, entonces $V_1, V_2, \ldots, V_n$ son L.D.

Enseguida veremos en el caso de dos vectores cuándo son L.D. y la interpretación geométrica de este hecho. Lo hare-- mos así para facilitar su comprensión.

Teorema 4.3.9.- Dos vectores son L.D. si y sólo si uno de ellos es múltiplo escalar del otro.

Demostración.- Supongamos que $V_2 = rV_1$, con $r \neq 0$. En-- tonces $rV_1 - V_2 = 0$ y V_1, V_2 son L.D. Por otro lado, supon- gamos que V_1 y V_2 son L.D. ; entonces existen constantes r_1, r_2 al menos una diferente de cero tales que $r_1V_1 + r_2V_2 = 0$. Si $r_1 \neq 0$ entonces $V_1 + (r_2/r_1)V_2 = 0$, de donde $V_1 = -(r_2/r_1)V_2$, esto es V_1 es múltiplo escalar de V_2. Si $r_1 = 0$, entonces $r_2 \neq 0$ y $V_2 = 0V_1 = 0$.

Geométricamente, el que dos vectores sean L.D. significa que dichos vectores son paralelos (figura a)).

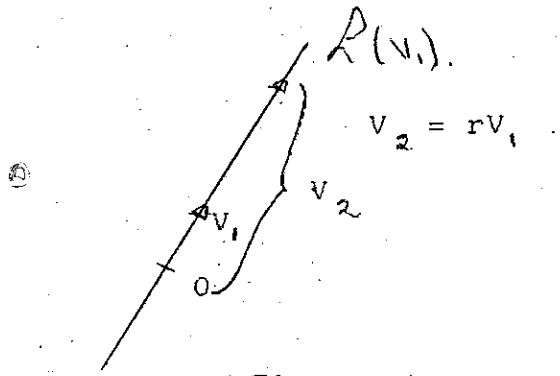


Figura a)

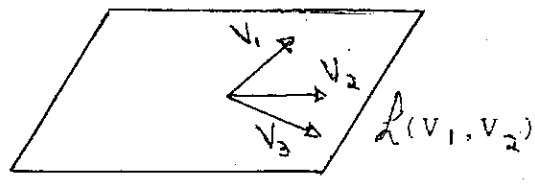
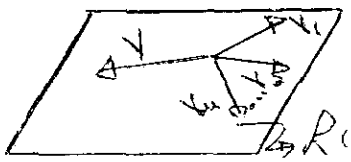


Figura b)

Si se tienen tres vectores L.D. geométricamente significa que estos tres vectores están sobre el mismo plano (figura b)), pues uno de ellos puede expresarse como combinación li-- neal de los otros dos (simplemente, la Ley del Paralelogramo); en otras palabras, son coplanares. En caso contrario, los -- vectores son no coplanares y L.I.

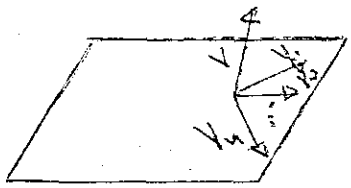
En general cuando se tienen n vectores L.D. en R^m y to- mamos al conjunto $\mathcal{L}(V_1, V_2, \ldots, V_n)$, dicho conjunto lo aso-- ciamos con un plano (por ser la idea geométrica más simple -- que tenemos) ; pero este conjunto no es exactamente un plano sino una generalización de él en R^m de forma tal que cual---- quier elemento $V \in \mathcal{L}(V_1, V_2, \ldots, V_n)$ pueda expresarse mediante una combinación lineal de $V_1, V_2, \ldots, V_n$. En caso contrario, el que $V \notin \mathcal{L}(V_1, V_2, \ldots, V_n)$ significa que el vector V no pue- de escribirse en términos de $V_1, V_2, \ldots, V_n$, es decir, V no - se encuentra en esta especie de plano.

La idea gráfica para esto es : si $V_1, V_2, \ldots, V_n$ son L.D entonces para todo $V \in \mathcal{L}(V_1, V_2, \ldots, V_n)$ se tiene que



$r_1V_1 + r_2V_2 + \ldots + r_mV_m = V$, entonces
$r_1V_1 + r_2V_2 + \ldots + r_mV_m - V = 0$
$\in \mathcal{R}(V_1, V_2, \ldots, V_m)$.

Si $V \notin \mathcal{R}(V_1, V_2, \ldots, V_m)$, indica que



$r_1V_1 + r_2V_2 + \ldots + r_mV_m = V$ para cualesquiera $r_1, r_2, \ldots, r_m \in R$, de modo que $r_1V_1 + r_2V_2 + \ldots + r_mV_m + rV = 0$ sólo con $r_1 = r_2 = \ldots = r_m = r = 0$.

Observaciones.-
- El hecho de que V esté o no en $\mathcal{R}(V_1, V_2, \ldots, V_m)$ únicamente depende de si existe o no una combinación lineal de $V_1, V_2, \ldots, V_m$ que nos exprese a V.
- Por su semejanza con un plano denominaremos al conjunto $\mathcal{R}(V_1, V_2, \ldots, V_m)$ como HIPERPLANO.

Basándonos de nuevo en las figuras a) y b), vemos que en la primera toda la recta $\mathcal{R}(V_1)$ que contiene a V_1 puede formarse a partir de este vector y mediante múltiplos escalares de él, ésto es, la recta $\mathcal{R}(V_1)$ puede "generarse" a partir de V_1 y decimos que V_1 genera a $\mathcal{R}(V_1)$; asímismo en b), si todo elemento en $\mathcal{R}(V_1, V_2)$ se puede "crear" o "generar" a partir de V_1 y V_2, decimos que V_1 y V_2 son generadores del plano $\mathcal{R}(V_1, V_2)$

Cabe señalar lo siguiente : el plano de la figura b) se obtuvo a partir de V_1, V_2 y V_3 y sin embargo líneas arriba se dijo que el plano puede generarse únicamente con V_1 y V_2. En otras palabras estamos afirmando que $\mathcal{R}(V_1, V_2, V_3) = \mathcal{R}(V_1, V_2)$. ¿Qué tan válido es ésto ?. Completamente, pues como estos tres vectores son L.D., podemos sustituir uno de ellos, digamos V_3, como combinación lineal de los otros dos de forta tal que en realidad estamos generando el plano en base a V_1 y V_2 solamente.

Este concepto de generadores también puede extenderse a un conjunto de n vectores $V_1, V_2, \ldots, V_m$. Cuando estos vectores son L.D. generamos al hiperplano $\mathcal{R}(V_1, V_2, \ldots, V_n)$ y además podemos sustituir uno de los vectores (igual que en el caso anterior) y cambiar el conjunto de generadores, ésto es, $\mathcal{R}(V_1, V_2, \ldots, V_m) = \mathcal{R}(V_2, V_3, \ldots, V_n)$. A ésto hace referencia el corolario 4.3.7.

Antes de estudiar algunos ejemplos sencillos de ésto, aprovecharemos para desarrollar lo que corresponde a sistemas de ecuaciones lineales sin solución.

4.3.b S I S T E M A S D E E C U A C I O N E S L I N E A - L E S S I N S O L U C I O N .

Hasta el momento hemos estudiado varios métodos diferentes para solución de sistemas de ecuaciones. Posteriormente y conforme se vaya avanzando en la teoría a lo largo de los siguientes capítulos, desarrollaremos algunos otros. Sin embargo, nos falta por analizar una situación muy importante : ¿Qué podemos decir acerca de los sistemas de ecuaciones lineales que no tienen solución ?. ¿Cuál es la relación existente entre lo que se ha visto en el transcurso de toda la teoría anterior y este nuevo problema ?.

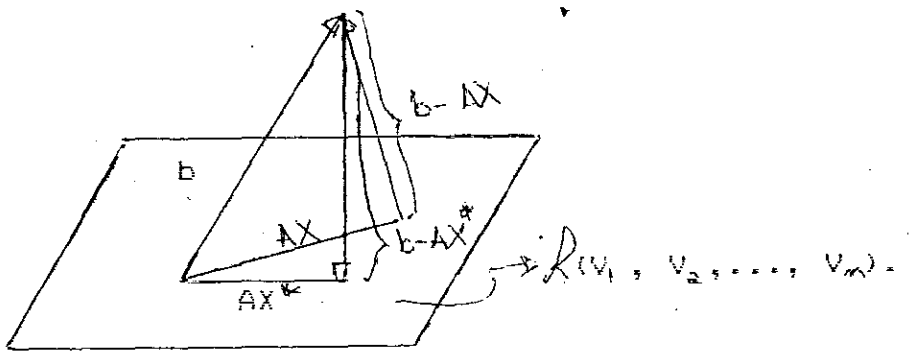
Ya vimos que si no podemos determinar el valor de las constantes $x_1, x_2, \ldots, x_m$ en el sistema $x_1V_1 + x_2V_2 + \ldots + x_mV_m = b$ es porque éste es inconsistente ; en este caso pudiera intentarse buscar lo que esté más cercano a ser una solución.

Si tratamos el problema desde su enfoque geométrico tenemos :
a) El sistema tiene solución cuando existe un vector X tal que la diferencia $b - AX = 0$.
b) Si el sistema no tiene solución es porque, obviamente, $b - AX \neq 0$.

Recordemos que

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$$

¿Cuál será entonces la "mejor solución posible" ?. Aquélla que nos permita afirmar que la diferencia anterior es lo más pequeña posible. Podemos concluir, auxiliándonos de la siguiente figura (que considera el caso en R^3) y del Teorema Generalizado de Pitágoras, que $b - AX$ es mínima cuando $b - AX$ es perpendicular a $V_1, V_2, \ldots, V_n$ simultáneamente (es decir, cuando $b - AX$ es perpendicular al hiperplano $\mathcal{L}(V_1, V_2, \ldots, V_n)$).



pendicular a $\mathcal{L}(V_1, V_2, \ldots V_m)$. De acuerdo con la defini--- ción de perpendicularidad entre vectores tenemos que debe --- cumplirse que :

$(AX) \cdot (b - AX^*) = 0$ para toda X, pues $AX \in \mathcal{L}(V_1, V_2, \ldots, V_m)$.

O lo que es lo mismo

$$(AX)^t \cdot (b - AX^*) = 0$$

Usando propiedades de la transpuesta :

$$(X^t A^t) \cdot (b - AX^*) = 0$$

Entonces

$$X^t (A^t b - A^t AX^*) = 0$$

Y ésto sucede si y sólo si

$$A^t b - A^t AX^* = 0.$$

Resolviendo esta ecuación para X^* tenemos :

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

Y claro que ésto podrá suceder siempre y cuando $A^t A$ sea invertible.

Con esta parte completamos todas las posibles opciones - de solución para un sistema de ecuaciones lineales.

Ahora sí, veamos algunos ejemplos que ilustren lo que -- hemos visto.

Ejemplo.- ¿Puede expresarse el vector (1,2,4) como com- binación lineal de los vectores $V_1 = (1,1,1)$, $V_2 = (0,1,1)$ y $V_3 = (0,0,1)$?

Para responder a ésto necesitamos encontrar elementos -- constantes r_1, r_2 y r_3 tales que

$$r_1(1,1,1) + r_2(0,1,1) + r_3(0,0,1) = (1,2,4).$$

$r_1 = 1$

$r_1 + r_2 = 2$, entonces $r_2 = 2 - r_1 = 2 - 1 = 1$

$r_1 + r_2 + r_3 = 4$, entonces $r_3 = 4 - r_2 - r_1 = 4 - 1 - 1 = 2$

Por lo tanto (1,2,4) sí es combinación lineal de V_1, V_2 y V_3, pues

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

Ejemplo.- Si cambiamos los vectores anteriores por (1,2,1), (1,1,1) y (0,1,0), ¿sigue siendo (1,2,4) una combi-- [illegible] de V_1, V_2 y V_3 ?

$$r_1(1,2,1) + r_2(1,1,1) + r_3(0,1,0) = (1,2,4).$$

$$\begin{aligned} r_1 + r_2 &= 1 \\ 2r_1 + r_2 + r_3 &= 2 \\ r_1 + r_2 &= 4. \end{aligned}$$

Pero no podemos encontrar valores de r_1 y r_2 tales que

$$\begin{aligned} r_1 + r_2 &= 1 \\ r_1 + r_2 &= 4 \end{aligned}$$

Por lo tanto, (1,2,4) NO es combinación lineal de estos vectores.

Ejemplo.- ¿Qué relación existe entre los vectores $V_1 = (1,2)$ y $V_2 = (2,4)$?
Rápidamente vemos que $V_2 = 2V_1$ o de otro modo $2V_1 - V_2 = 0$, por lo tanto V_1 y V_2 son L.D.

Ejemplo.- ¿Y entre los vectores $V_1 = (1,2,3)$, $V_2 = (-4,1,5)$ y $V_3 = (-5,8,19)$?. Resolviendo

$$r_1V_1 + r_2V_2 + r_3V_3 = 0$$

para saber si son L.I. o L.D. :

$$r_1(1,2,3) + r_2(-4,1,5) + r_3(-5,8,19) = (0,0,0)$$

$$\begin{aligned} r_1 - 4r_2 - 5r_3 &= 0 \\ 2r_1 + r_2 + 8r_3 &= 0 \\ 3r_1 + 5r_2 + 19r_3 &= 0 . \end{aligned}$$

Utilizando notación matricial y eliminación gaussiana :

$$\begin{bmatrix} 1 & -4 & -5 \\ 2 & 1 & 8 \\ 3 & 5 & 19 \end{bmatrix} \xrightarrow[-3R_1 + R_3]{-2R_1 + R_2} \begin{bmatrix} 1 & -4 & -5 \\ 0 & 9 & 18 \\ 0 & 17 & 34 \end{bmatrix} \xrightarrow[-17R_2 + R_3]{(1/9)R_2} \begin{bmatrix} 1 & -4 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} .$$

Entonces

$$\begin{aligned} r_1 - 4r_2 - 5r_3 &= 0 \\ r_2 + 2r_3 &= 0. \end{aligned}$$

Entonces $r_2 = -2r_3$, de donde $r_1 + 8r_3 - 5r_3 = 0$ y así $r_1 = -3r_3$.

Por lo tanto, existen muchas soluciones, son L.D. En particular, si $r_1 = 3$, $r_2 = 2$, $r_3 = -1$ entonces $3V_1 + 2V_2 - V_3 = 0$ o $V_3 = 3V_1 + 2V_2$.

Ejemplo.- Considere el sistema

$$\begin{aligned} x &= 0 \\ x + y &= 2 \\ x + 2y &= 7 . \end{aligned}$$

Fácilmente se observa que dicho sistema NO tiene solución. Determine la mejor solución posible.

Utilizando notación matricial tenemos $AX = b$, con

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} , \quad X = \begin{bmatrix} x \\ y \end{bmatrix} , \quad b = \begin{bmatrix} 0 \\ 2 \\ 7 \end{bmatrix}$$

Para determinar la mejor solución posible, necesitamos resolver la ecuación

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

Tenemos :

$$A^{t}A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} .$$

Determinando $(A^{t}A)^{-1}$ mediante Gauss-Jordan :

$$\left[\begin{array}{cc|cc} 3 & 3 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right] \xrightarrow[-3R_1+R_2]{1/3\ R_1} \left[\begin{array}{cc|cc} 1 & 1 & 1/3 & 0 \\ 0 & 2 & -1 & 1 \end{array}\right] \xrightarrow[-R_2+R_1]{1/2\ R_2} \left[\begin{array}{cc|cc} 1 & 0 & 5/6 & -1/2 \\ 0 & 1 & -1/2 & 1/2 \end{array}\right]$$

Entonces

$$(A^{t}A)^{-1} = \begin{bmatrix} 5/6 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} .$$

Con ésto :

$$(A^{t}A)^{-1} A^{t} = \begin{bmatrix} 5/6 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 5/6 & 1/3 & -1/6 \\ -1/2 & 0 & 1/2 \end{bmatrix}$$

De aquí :

$$X^{*} = (A^{t}A)^{-1} A^{t} b = \begin{bmatrix} 5/6 & 1/3 & -1/2 \\ -1/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 7/2 \end{bmatrix}$$

Nota : Esta técnica es la utilizada para adaptar, en forma óptima, una recta o una curva a un conjunto de puntos en el plano que se han determinado previamente en forma experimental, esto es, ajuste de datos por el método de mínimos cuadrados donde, en el caso general

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdots & \cdots \\ 1 & x_n \end{bmatrix}, \quad X = \begin{bmatrix} a \\ b \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{bmatrix}.$$

El ejemplo corresponde a la obtención de la recta de adaptación por mínimos cuadrados a los puntos (0,0), (1,2) y (2,7).

Pasaremos ahora a generalizar estos resultados a conjuntos más abstractos llamados Espacios Vectoriales.

4.4 E S P A C I O S V E C T O R I A L E S : Definición y propiedades básicas.

Definición 4.4.1.- Sea V un conjunto no vacío y K un campo(*). Definimos un Espacio Vectorial $\mathbf{V}$ sobre K a un conjunto de objetos con dos operaciones definidas : una llamada suma (+) y otra llamada multiplicación por un escalar elemento de K tales que se cumplen los siguientes axiomas :

1) Para $V,W \in \mathbf{V}$ se tiene que $(V+W) \in \mathbf{V}$ (cerradura para la suma).
2) $V+W = W+V$, para $V,W \in \mathbf{V}$ (conmutatividad para la suma)
3) Si $V,W,Z \in \mathbf{V}$ entonces $V+(W+Z) = (V+W)+Z$ (asociatividad para la suma).
4) Existe $\theta \in \mathbf{V}$ tal que $\theta+V = V+\theta = V$, para toda $V \in \mathbf{V}$ (θ = neutro aditivo).
5) Para toda $V \in \mathbf{V}$ existe $(-V) \in \mathbf{V}$ tal que $V+(-V) = \theta$ (inverso aditivo)
6) Si $\alpha \in K$ y $V \in \mathbf{V}$ entonces $(\alpha V) \in \mathbf{V}$ (cerradura)
7) $\alpha(V+W) = \alpha V + \alpha W$, para $\alpha \in K$, $V,W \in \mathbf{V}$ (distributividad)
8) $(\alpha+\beta)V = \alpha V + \beta V$, para $\alpha,\beta \in K$, $V \in \mathbf{V}$ (distributividad)
9) $\alpha(\beta V) = (\alpha\beta)V = \beta(\alpha V)$, para $\alpha,\beta \in K$, $V \in \mathbf{V}$.
10) $1 \cdot V = V$, con $1 \in K$, $V \in \mathbf{V}$ (neutro multiplicativo).

Notación : En el caso en que se tenga una estructura algebraica como la anterior diremos que "$\mathbf{V}$ es un espacio vectorial sobre el campo K" y se denotará $\mathbf{V}(K)$.

Antes de estudiar algunos ejemplos de espacios vectoriales, necesitamos hacer algunas aclaraciones :

1) Por comodidad, dado que ya estamos familiarizados con el término, llamaremos vectores a los elementos del espacio vectorial .

(*) Recuerde que un campo es un sistema (que puede ser de números) que contiene a la suma, diferencia, producto y cocien-

2) K no tiene que ser necesariamente el campo de los reales, ni siquiera el de los complejos, en general cualquier conjunto con estructura de campo nos sirve.

3) Es conveniente referirse al "espacio vectorial V" , ya que V es solamente un conjunto de vectores mientras que un espacio vectorial es un sistema matemático que consiste en una terna ordenada $(V,+,\cdot)$, donde $\cdot$ es la multiplicación por un escalar.

Ahora sí, analizaremos algunos ejemplos de espacios vectoriales.

Ejemplo.- Sea $V = P_n$ el conjunto de polinomios con coeficientes reales de grado menor o igual a n, es decir $P_n = \{ p(x) / p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n , a_i \in R \}$. Definimos la suma de dos elementos de P_n como:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n ,$$
$$q(x) = b_0 + b_1 x + b_2 x^2 + \ldots + b_n x^n ,$$

entonces

$$p(x) + q(x) = (a_0+b_0) + (a_1+b_1)x + (a_2+b_2)x^2 + \ldots + (a_n+b_n)x^n ,$$

y la multiplicacion por un escalar $\alpha \in R$ como

$$\alpha p(x) = \alpha a_0 + \alpha a_1 x + \alpha a_2 x^2 + \ldots + \alpha a_n x^n .$$

Es claro que la suma de dos polinomios de grado menor o igual a n es otro polinomio también de grado menor o igual a n, por lo que se satisface la propiedad 1).

Las propiedades 2 y 3 se siguen de la conmutatividad y asociatividad para la suma de reales. Por argumentos semejantes es fácil verificar que se satisfacen las propiedades 6,7,8,9, y 10.

Para verificar la propiedad 4 necesitamos definir el polinomio cero como $\Theta = 0 + 0x + 0x^2 + \ldots + 0x^n$. Así definido, $\Theta \in P_n$ y se satisface 4.

Finalmente, haciendo $-p(x) = -a_0 - a_1 x - \ldots - a_n x^n$, vemos que vale la propiedad 5, por lo que concluimos que P_n es un espacio vectorial con coeficientes reales.

Ejemplo.- Consideremos ahora el conjunto de funciones definidas y continuas en $I = [0,1]$ y con valores reales, es decir $f:[0,1] \longrightarrow R$, f continua en I.

Definimos $(f+g)(x) = f(x) + g(x)$,
$(\alpha f)(x) = \alpha [f(x)]$.

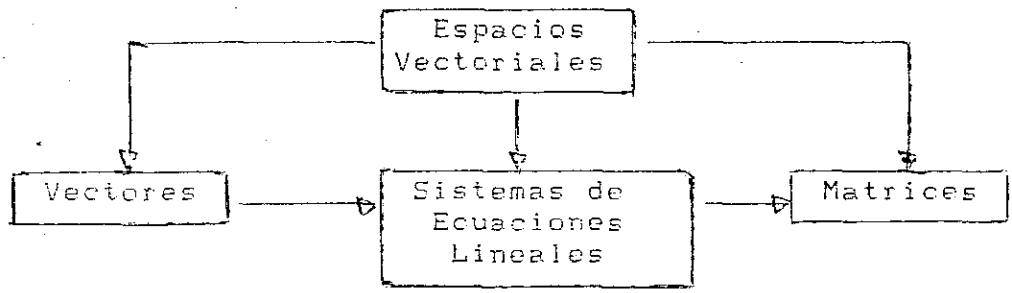
Dado que la suma de dos funciones continuas es continua (argumentos analizados en los cursos de Cálculo), la propiedad 1 se satisface ; las propiedades restantes se verifican fácilmente. Tomando O como la función cero, es decir $f(x)=O$ y $(-f)(x) = -f(x)$ demostramos las propiedades 4 y 5

Aunados a los anteriores tenemos como ejemplos de espa-- cios vectoriales a los tres primeros capítulos de esta tesis, es decir, vectores, matrices y solución de sistemas de ecua-- ciones lineales homogéneos. Los primeros dos están demostra- dos ya en el desarrollo de la teoría ; veremos el último.

Sea el conjunto solución de un sistema de ecuaciones li- neales homogéneo formado por elementos en R o C (es decir, -- K = R o C). El teorema 2.6.2 verifica las propiedades 1 y 6 y, en consecuencia, 5,6 y 10. Las propiedades 2,3,7 y 8 se - siguen de propiedades de reales (o complejos).

La propiedad 4 se satisface también, pues es la solución trivial.

Con todo ésto, podemos relacionar estos temas gráfica--- mente como sigue :



pues los espacios vectoriales son la generalización de los -- otros tres temas y, además, éstos están relacionados entre sí como ya se vió antes.

De los axiomas requeridos en la definición podemos fá--- cilmente obtener algunas propiedades elementales de los espa- cios vectoriales tales como :

a) Unicidad del cero, es decir, si θ y $\theta' \in \mathbb{V}$ son tales que -- para toda $V \in \mathbb{V}$ se tiene que $V + \theta = V$ y $V + \theta' = V$, enton-- ces $\theta = \theta'$.

b) Unicidad del inverso aditivo.

c) $\theta X = X\theta = \theta$.

d) $-X = (-1)X$.

e) Si $W \in \mathbb{V}$ tal que $W + V = V$, para toda $V \in \mathbb{V}$, entonces $W = \theta$

Por lo sencillo y lógico de estos enunciados no nos de-- tendremos a demostrarlos.

Generalizando las definiciones de la sección anterior -- tenemos :

Definición 4.4.2.- Sean $V_1, V_2, \ldots, V_n \in \mathbb{V}(K)$. Un vector V - - -- se dice que es una combinación lineal de los vectores $V_1, V_2, \ldots, V_n$ si y sólo si

Definición 4.4.3.- Sean $V_1, V_2, \ldots, V_n \in \mathcal{V}(K)$. Diremos que $V_1, V_2, \ldots, V_n$ forman un conjunto de vectores linealmente independientes, L.I., si existen $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ tales que

$$\lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n = \Theta$$

únicamente con $\lambda_1, \lambda_2, \ldots, \lambda_n = 0$, y Θ = elemento cero de $\mathcal{V}$. Si ésto se obtiene con algún $\lambda_i \neq 0$, entonces $V_1, V_2, \ldots, V_n$ son L.D.

Definición 4.4.4.- Si todo vector $V \in \mathcal{V}(K)$ puede escribirse como $\lambda V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n = V$, entonces decimos que $V_1, V_2, \ldots, V_n$ generan $\mathcal{V}(K)$.

4.5 S U B E S P A C I O S V E C T O R I A L E S : Definición y ejemplos.

Ya que algebraicamente un espacio vectorial $\mathcal{V}$ es un conjunto de objetos (con una serie de propiedades comunes), podemos determinar varios subconjuntos de él de manera que algunos "heredan" dichas propiedades y otros no ; con ésto, los subconjuntos en los que se preservan estas propiedades (axiomas) son a su vez espacios vectoriales, por lo que decimos que estos subconjuntos son subespacios vectoriales de $\mathcal{V}$. En esta sección estudiaremos estos conjuntos tan importantes.

Definición 4.5.1.- Sea $\mathcal{W}$ un subconjunto no vacío de $\mathcal{V}(K)$. Si $\mathcal{W}$ es él mismo un espacio vectorial, bajo las mismas operaciones definidas en $\mathcal{V}$, entonces se dice que $\mathcal{W}$ es un subespacio vectorial de $\mathcal{V}$.

De aquí resulta, aparentemente, que el demostrar si un subconjunto $\mathcal{W}$ de un espacio vectorial $\mathcal{V}$ también es espacio vectorial bajo las mismas operaciones implica la demostración de los 10 axiomas que definen a un espacio vectorial, lo cual sería muy engorroso. Pero podemos simplificar bastante este trabajo utilizando la "heredabilidad" de varias de las propiedades y pidiendo únicamente que sea válida la cerradura para ambas operaciones tal y como se enuncia en el siguiente

Teorema 4.5.2.- Sea $\mathcal{W}$ un sunconjunto no vacío del espacio vectorial $\mathcal{V}$. Si para todo $V, W \in \mathcal{W}$ se satisface que

i) $V + W \in \mathcal{W}$

ii) $\lambda V \in \mathcal{W}$, $\lambda \in K$

afirmamos que $\mathcal{W}$ mismo tiene estructura de espacio vectorial y se dice que $\mathcal{W}$ es un subespacio vectorial de $\mathcal{V}$.

Demostración.- Para demostrar este teorema debemos verificar que $\mathcal{W}$ satisface los 10 axiomas necesarios para ser un espacio vectorial.

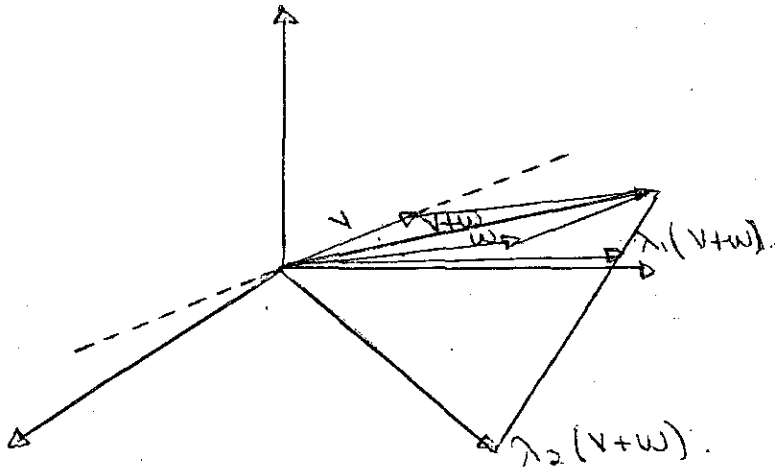
Las cerraduras para la suma (axioma 1) y para la multiplicación por un escalar (axioma 6) se satisfacen por hipótesis. Ahora, si $\lambda = 0$, entonces $0V \in \mathcal{W}$, por ii), pero $0V = \Theta \in \mathcal{V}$ y $\Theta \in \mathcal{W}$, por la unicidad del cero en $\mathcal{V}$, es válida la propiedad

4.

La asociatividad y conmutatividad para la suma (axiomas 2 y 3), asociatividad para el producto (axioma 9), las dos -- leyes distributivas (axiomas 7 y 8) y la existencia del ele-- mento neutro multiplicativo (axioma 10) se verifican también, dado que $V, W \in \mathcal{W} \subset \mathcal{V}$.

Faltaría por verificar únicamente el axioma 5, ésto es, la existencia de elementos inversos. Pero ésto es fácil, --- basta con tomar $\alpha = -1$, y por el axioma 6, $(-1)V = -V \in \mathcal{W}$, con lo cual terminamos la demostración.

La idea geométrica de un subespacio vectorial $\mathcal{W}$ es la -- siguiente : Supongamos que estamos en $\mathbb{R}^3$. El que $\lambda V \in \mathcal{W}$, con $\lambda \in K$ y $V \in \mathcal{V}$ significa que toda la recta generada por el vector V (y que por supuesto lo contiene a él) está contenida en el conjunto $\mathcal{W}$; igualmente, el que $\theta \in \mathcal{W}$ significa que dicha recta pasa por el origen del espacio R^3, y el que $(V+W) \in \mathcal{W}$ con $V, W \in \mathcal{W}$ indica que este nuevo vector (y las sumas de múltiplos de V y W) también forman parte de $\mathcal{W}$, es decir, $\mathcal{W}$ no admite do-- bladuras u ondulaciones.



Ahora veremos algunos ejemplos de subespacios. Trataremos con los mismos ejemplos de espacios vectoriales que ya se vieron.

Ejemplo.- Sea $\mathcal{V} = R^n$ y $\mathcal{W} = \{(x_1, x_2, \ldots, x_n) \ / \ x_1 = x_2\}$. Sean $V = (x_1, x_1, x_3, \ldots, x_n)$ y $W = (y_1, y_1, y_3, \ldots, y_n)$ y las operaciones de suma y multiplicación por un escalar ya cono-- cidas. Entonces $V + W = (x_1 + y_1, x_1 + y_1, x_3 + y_3, \ldots, x_n + y_n)$, de donde $(V+W) \in \mathcal{W}$. Por lo tanto, se satisface i).

$\lambda V = (\lambda x_1, \lambda x_1, \lambda x_3, \ldots, \lambda x_n)$, por lo que se satis-face ii).

Por lo tanto $\mathcal{W}$ es un subespacio vectorial de .

Ejemplo.- Sea $\mathcal{V} = P_n$, el conjunto de polinomios con ---

$$P_n = \{ a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \ / \ a_i \in R \} .$$

Sea
$W = \{ p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \ / \ a_1 = 0 \}$,
y las operaciones ya conocidas.

Sean $p(x) = a_0 + a_2 x^2 + \ldots + a_n x^n$
$q(x) = b_0 + b_2 x^2 + \ldots + b_n x^n$,
entonces
$p(x) + q(x) = (a_0+b_0) + (a_2+b_2)x^2 + \ldots + (a_n+b_n)x^n$.
Por lo tanto, se cumple i).

$\lambda p(x) = \lambda a_0 + \lambda a_2 x^2 + \ldots + \lambda a_n x^n \in W$. Se satisface --- ii).

Por lo tanto, W es un subespacio vectorial de V.

En base a lo visto anteriormente podemos enunciar la siguiente proposición :

Proposición 4.5.3.- Sean V un espacio vectorial arbitrario y $V_1, V_2, \ldots, V_n$ vectores de V . Si
$W = \{ \lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n$, con $\lambda_1, \lambda_2, \ldots, \lambda_n \in K \}$
entonces W es un subespacio vectorial de V .
En otras palabras, el conjunto de combinaciones lineales de $V_1, V_2, \ldots, V_n$ es un subespacio de V .

Demostración.-
i) Tenemos $V, W \in W$, entonces :
$V = \lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n$
$W = \alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n$.
De aquí que
$V + W = (\lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n) + (\alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n)$
$= (\lambda_1 + \alpha_1) V_1 + (\lambda_2 + \alpha_2) V_2 + \ldots + (\lambda_n + \alpha_n) V_n$.
Como esto también es una combinación lineal de $V_1, V_2, \ldots, V_n$, entonces $V + W \in W$.

ii) Si $V \in W$, entonces $\alpha V = \alpha (\lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n)$
$= \alpha \lambda_1 V_1 + \alpha \lambda_2 V_2 + \ldots + \alpha \lambda_n V_n$
$= (\alpha \lambda_1) V_1 + (\alpha \lambda_2) V_2 + \ldots + (\alpha \lambda_n) V_n$,
que, claramente vemos, también es una combinación lineal de $V_1, V_2, \ldots, V_n$. Por lo tanto $\alpha V \in W$.

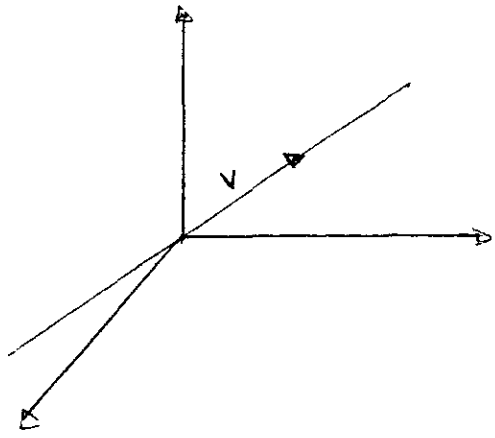
Esto prueba que W es un subespacio vectorial de V .

Observaciones :
- El conjunto anterior W podemos denotarlo por $W = \mathcal{L}(V_1, V_2, \ldots, V_n)$ y se le conoce como el subespacio generado por $V_1, V_2, \ldots, V_n$.
- Se le llama subespacio generado por $V_1, V_2, \ldots, V_n$ porque --- cualquier elemento de $\mathcal{L}(V_1, V_2, \ldots, V_n)$ se puede "crear" o --- "generar" a partir de $V_1, V_2, \ldots, V_n$.

cio generado por un vector no nulo $V \in \mathcal{V}$ está formado por

$$\mathcal{L}(V) = \{ \lambda V, \ \lambda \in K \} .$$

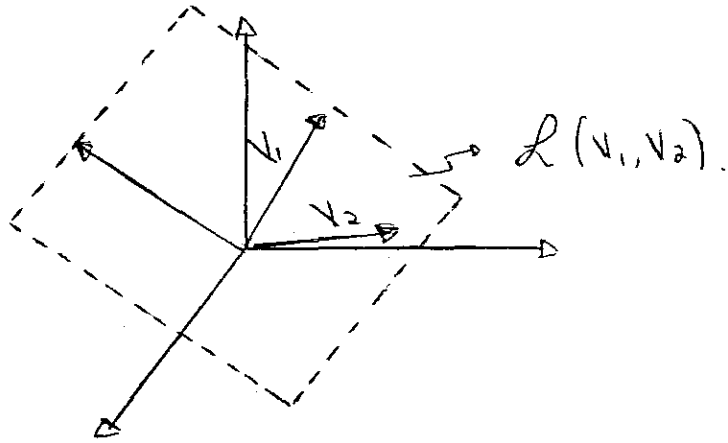
es decir, por todos los múltiplos escalares de V. Geométricamente, es la recta que pasa por el origen y contiene a V.



Ejemplo.- El subespacio generado por dos vectores cualquiera $V_1, V_2 \in R^3$ tales que no son múltiplos uno del otro lo representamos por

$$\mathcal{L}(V_1, V_2) = \{ \lambda_1 V_1 + \lambda_2 V_2 \ , \lambda_1, \lambda_2 \in K \}$$

Es decir, geométricamente nos forma un paralelogramo (plano) que contiene a V_1 y V_2 y que pasa por el origen.



Ejemplo.- Sean $V_1 = (1,1,0)$, $V_2 = (2,1,0)$ y

$$\mathcal{W} = \mathcal{L}(V_1, V_2) = \{(x,y,0), \ x,y \in K \}$$

¿Es $\mathcal{W}$ un subespacio generado por V_1 y V_2?. Si acaso lo es deberá cumplir con que

$$\mathcal{W} = \{ \lambda_1 V_1 + \lambda_2 V_2 \ , \ \lambda_1, \lambda_2 \in K \}$$

Esto es, $\lambda_1 V_1 + \lambda_2 V_2 = \lambda_1 (1,1,0) + \lambda_2 (2,1,0)$.

Pero, además, por como está definido :

$$\lambda_1 (1,1,0) + \lambda_2 (2,1,0) = (x,y,0).$$

De aquí, que todo se reduce a resolver el sistema

$$\begin{aligned} \lambda_1 + 2\lambda_2 &= x \\ \lambda_1 + \lambda_2 &= y \end{aligned} .$$

Mediante notación matricial y el método de Gauss :

$$\left[\begin{array}{cc|c} 1 & 2 & x \\ 1 & 1 & y \end{array}\right] \xrightarrow{R_{12}} \left[\begin{array}{cc|c} 1 & 1 & y \\ 1 & 2 & x \end{array}\right] \xrightarrow{-R_1+R_2} \left[\begin{array}{cc|c} 1 & 1 & y \\ 0 & 1 & x-y \end{array}\right] \xrightarrow{-R_2+R_1}$$

$$\left[\begin{array}{cc|c} 1 & 0 & -x+2y \\ 0 & 1 & x-y \end{array}\right] .$$

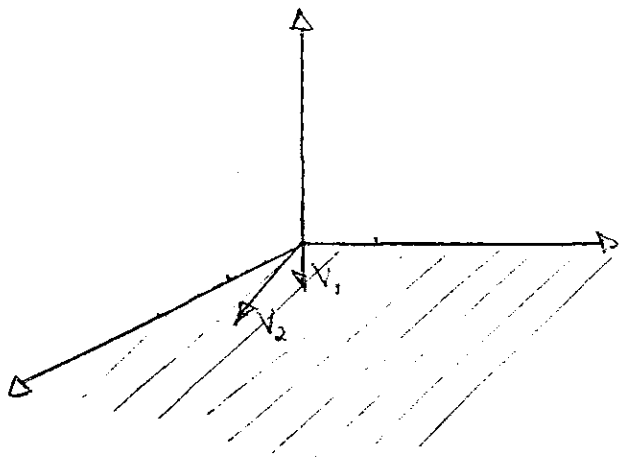
De donde el sistema asociado es :

$$\lambda_1 = -x + 2y$$
$$\lambda_2 = x - y \quad .$$

Comprobando :

$$\begin{aligned}(-x + 2y)(1,1,0) + (x - y)(2,1,0) &= (-x + 2y, -x + 2y, 0) + (2x - 2y, x - y, 0) \\ &= (-x+2y+2x-2y, -x+2y+x-y, 0) \\ &= (x,y,0).\end{aligned}$$

Por lo tanto, W sí es un subespacio vectorial de R^3 . Geométricamente, ésto se vé como sigue :



$\mathcal{L}(V_1, V_2) = (x,y,0)$

Observación.- Generalizando la idea geométrica del subespacio generado por $V_1, V_2, \ldots, V_n$, al conjunto $\mathcal{L}(V_1, V_2, \ldots V_n)$ se le conoce también como el Hiperplano Generado por $V_1, V_2, \ldots, V_n$.

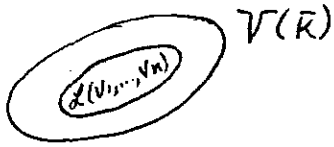
4.6 B A S E S Y D I M E N S I O N .

Hasta el momento tenemos definido formalmente el concepto de espacio vectorial, analizamos algunas de sus propiedades básicas e investigamos como clasificar a un conjunto de vectores como un espacio o subespacio vectorial ; además, definimos la dependencia e independencia lineal de un conjunto de vectores y comprobamos que algunos conjuntos generan un espacio vectorial

Todos estos conceptos, especialmente el de espacio vectorial, son importantísimos en Matemáticas por lo que queremos conocer más acerca de estos conjuntos en términos, de ser posible, de sus características más elementales. En esta sección estudiaremos algunas de estas características para las cuales es fundamental el concepto de base ; veremos esto para $\mathcal{L}(V_1, V_2, \ldots, V_n)$ y después generalizaremos a espacios vectoriales. Del mismo modo definiremos formalmente el concepto de dimensión de un espacio vectorial trabajando la idea intuitiva que se tiene de ello pero en forma más general.

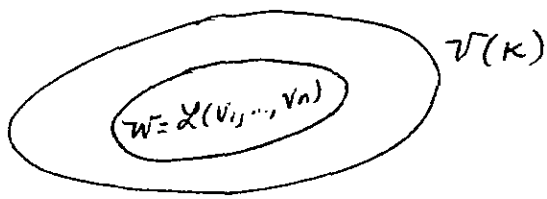
Supongamos que tenemos n vectores $V_1, V_2, \ldots, V_n$ y tenemos al conjunto $\mathcal{L}(V_1, V_2, \ldots, V_n)$. Dependiendo de si $V_1, V_2, \ldots V_n$ son L.I. o L.D. cada vector $V \in \mathcal{L}(V_1, V_2, \ldots, V_n)$ puede expresarse en forma única en términos de $V_1, V_2, \ldots, V_n$ o no. Esto es, si $V_1, V_2, \ldots, V_n$ son L.I. y además generan un hiperplano $\mathcal{L}(V_1, V_2, \ldots, V_n)$ significa que cualquier vector $V \in (V_1, V_2, \ldots, V_n)$ puede obtenerse de manera única a partir de $V_1, V_2, \ldots, V_n$ o' en base a ellos y precisamente así se denomina este conjunto de vectores. Así, podemos enunciar la siguiente

Definición 4.6.1.- Sea $\{V_1, V_2, \ldots, V_n\}$ un conjunto finito de vectores en un subespacio vectorial $\mathcal{L}(V_1, V_2, \ldots, V_n)$. Se dice que $\{V_1, V_2, \ldots, V_n\}$ es una BASE de $\mathcal{L}(V_1, V_2, \ldots, V_n)$ si $V_1, V_2, \ldots, V_n$ son L.I.



Todavía más : si $\mathcal{W}$ es un subespacio vectorial de $\mathcal{V}(K)$ y $V_1, V_2, \ldots, V_n \in \mathcal{W}$, entonces si $\mathcal{W} = \mathcal{L}(V_1, V_2, \ldots, V_n)$ y $V_1, V_2, \ldots V_n$ son L.I. resulta que ocurre exactamente lo mismo que en el caso anterior : cualquier $V \in \mathcal{V}$ puede expresarse de manera única en base a $V_1, V_2, \ldots, V_n$ con lo que, ampliando un poco la definición anterior, tenemos que :

Definición 4.6.1'.- Sea $\{V_1, V_2, \ldots, V_n\}$ un conjunto finito de vectores en $\mathcal{W} \subset \mathcal{V}(K)$. Decimos que $\{V_1, V_2, \ldots, V_n\}$ es una base para $\mathcal{W}$ si :
i) $\{V_1, V_2, \ldots, V_n\}$ es L.I.
ii) $\mathcal{W} = \mathcal{L}(V_1, V_2, \ldots, V_n)$.



Observación.- Los sistemas donde aparezcan vectores $V_1, V_2, \ldots, V_n$ que son base siempre tienen solución y es, además, única. Por otro lado, si $V_1, V_2, \ldots, V_n$ generan a $\mathcal{W}$ y son L.D. siempre se tienen infinitas soluciones.

Recordemos uno de los ejemplos de la sección anterior : vimos que el conjunto $\mathcal{A} = \{(x, y, 0) / x, y \in K\}$ es un

subespacio de R^3 generado por $V_1 = (1,1,0)$ y $V_2 = (2,1,0)$. Falta determinar si estos vectores son L.I. para clasificarlos como una base para $\mathcal{L}(V_1,V_2)$. Así :

$$\lambda_1(1,1,0) + \lambda_2(2,1,0) = (0,0,0)$$
$$\lambda_1 + 2\lambda_2 = 0$$
$$\lambda_1 + \lambda_2 = 0.$$

Mediante notación matricial y el método de Gauss :

$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \xrightarrow{-R_1 + R_2} \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$$

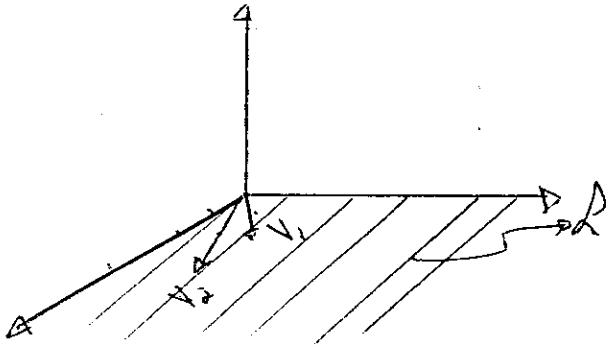
Entonces

$$\lambda_1 + 2\lambda_2 = 0$$
$$-\lambda_2 = 0.$$

De donde $\lambda_1 = \lambda_2 = 0$.

De aquí que V_1 y V_2 son L.I. y, por lo tanto, son una base para $\mathcal{L}(V_1,V_2)$

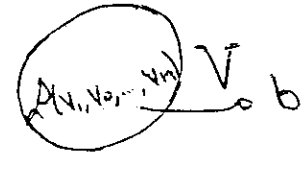
En esta misma sección vimos el significado de esto :



$\mathcal{L}(V_1,V_2) = \{(x,y,0) / x,y \in K\}$ es simplemente el plano xy.

Podemos generalizar este concepto a espacios vectoriales y enunciarla como :

Definición 4.6.2.- Sea $V(K)$. Decimos que $\{V_1,V_2,\ldots,V_m\}$ es una base para V si satisface que
i) $\{V_1,V_2,\ldots,V_m\}$ es L.I.
ii) $\mathcal{L}(V_1,V_2,\ldots,V_m) = V(K)$.



para $x_1V_1 + x_2V_2 + \ldots + x_mV_m = b$

Es claro que en el ejemplo mencionado líneas antes el espacio vectorial R^3 No tiene como base a $\{V_1,V_2\}$ pues, aunque estos dos vectores satisfacen la primera condición no cumplen la segunda, es decir, no generan a todo R^3 sino sólo a una parte de él : el plano xy. Necesitaríamos tener en

forme un conjunto L.I. y además genere a R^3 . Por ejemplo, tomemos $V_1 = (1,1,0)$, $V_2 = (2,1,0)$ y $V_3 = (0,0,1)$.

Necesitamos determinar $\lambda_1, \lambda_2, \lambda_3$ tales que

$$\lambda_1(1,1,0) + \lambda_2(2,1,0) + \lambda_3(0,0,1) = (0,0,0).$$

De aquí :

$$\begin{aligned} \lambda_1 + 2\lambda_2 &= 0 \\ \lambda_1 + \lambda_2 &= 0 \\ \lambda_3 &= 0. \end{aligned}$$

Entonces $\lambda_1 = \lambda_2 = \lambda_3 = 0$, por lo que V_1, V_2 y V_3 son L.I. Falta comprobar que generan a R^3 . Para esto tomamos un vector general en R^3 :

$$\lambda_1(1,1,0) + \lambda_2(2,1,0) + \lambda_3(0,0,1) = (x,y,z).$$

De donde

$$\begin{aligned} \lambda_1 + 2\lambda_2 &= x \\ \lambda_1 + \lambda_2 &= y \\ \lambda_3 &= z. \end{aligned}$$

Así :

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & x \\ 1 & 1 & 0 & y \\ 0 & 0 & 1 & z \end{array}\right] \xrightarrow{-R_1+R_2} \left[\begin{array}{ccc|c} 1 & 2 & 0 & x \\ 0 & -1 & 0 & -x+y \\ 0 & 0 & 1 & z \end{array}\right] \xrightarrow[-R_2]{2R_2+R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -x+2y \\ 0 & 1 & 0 & x-y \\ 0 & 0 & 1 & z \end{array}\right]$$

De aquí que $\lambda_1 = -x + 2y$
$\lambda_2 = x - y$
$\lambda_3 = z$.

Comprobando :

$$\begin{aligned} (-x+2y)(1,1,0) + (x-y)(2,1,0) + z(0,0,1) &= (-x+2y+2x-2y, \\ &\quad -x+2y+x-y, z) \\ &= (x,y,z). \end{aligned}$$

De donde concluimos que efectivamente V_1, V_2 y V_3 son una base para R^3 .

Pero, ¿por qué esta diferencia en el número de vectores de una base para $\mathcal{L}(V_1, V_2)$ y para R^3 ?. ¿Tendrá algo que ver en esta diferencia el hecho de que uno represente geométricamente un plano y el otro el espacio ?. Para aclarar esta interrogante es esencial el concepto de dimensión. Nos apoyaremos, como dijimos al inicio de esta sección, en la idea intuitiva que tenemos de dimensión.

Según cursos de Geometría pensamos en una recta (R) como un espacio unidimensional, es decir, algo con una sola dimensión ; así también identificamos al plano como un espacio bidimensional (R^2), y en un espacio como algo de tres dimensiones (R^3). Curiosamente, observe que, en los ejemplos anteriores, coinciden el número de vectores en una base y la dimensión de cada conjunto, respectivamente. Esto es: $\mathcal{L}(V_1, V_2)$ es un plano con los vectores V_1, V_2 formando una base para él y es de dimensión 2 ; R^3 es un espacio de dimensión 3 con $\{V_1, V_2, V_3\}$ una base de él.

En pocas palabras, el número de vectores que constituyen

una base de un conjunto es una cantidad muy especial. Esto nos sugiere la

Definición 4.6.3.- La dimensión de un espacio vectorial $\mathcal{V}(K)$ es el número de vectores en una base de $\mathcal{V}$. Dicho de otro modo, es el número máximo de vectores L.I. de $\mathcal{V}$, y se denota dim.$\mathcal{V}$. (En el caso $\mathcal{V}=\{0\}$, definimos su dimensión como cero).

Para aclarar un poco más la definición anterior, daremos la siguiente

Definición 4.6.4.- Sea $\mathcal{V}(K)$. Diremos que $V_1, V_2, \ldots, V_n \in \mathcal{V}$ es un conjunto máximo L.I. si
a) $V_1, V_2, \ldots, V_n$ son L.I.
b) Para cualquier $W \in \mathcal{V}$ se tiene que $V_1, V_2, \ldots, V_n, W$ son L.D.

Esto es, si a un conjunto de vectores L.I. en $\mathcal{V}$ le añadimos otro vector más $W \in \mathcal{V}$ y este nuevo conjunto de vectores es ahora L.D., decimos que el conjunto inicial es un conjunto máximo L.I.

Con esta definición tenemos un criterio más para afirmar cuándo un conjunto de vectores de un espacio vectorial $\mathcal{V}$ constituyen una base. Esto lo enunciamos en la siguiente

Proposición 4.6.5.- $V_1, V_2, \ldots, V_n \in \mathcal{V}(K)$ forman un conjunto máximo de vectores L.I. si y sólo si $V_1, V_2, \ldots, V_n$ son una base para $\mathcal{V}$.

Demostración.- Primero demostraremos que si $V_1, V_2, \ldots, V_n$ forman un conjunto máximo de vectores L.I., entonces son una base verificando que :
i) $V_1, V_2, \ldots, V$ son L.I.
ii) $\mathcal{V} = \mathcal{L}(V_1, V_2, \ldots, V_n)$.

La primera condición se sigue por hipótesis, no necesita demostración. Para demostrar ii) hay que ver que cualquier elemento $W \in \mathcal{V}$ puede expresarse en términos de $V_1, V_2, \ldots, V_n$, es decir, hay que demostrar que existen constantes $\lambda_1, \ldots, \lambda_n$ tales que $W = \lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n$.

Como por hipótesis $V_1, V_2, \ldots, V_n$ forman un conjunto máximo L.I., entonces $V_1, V_2, \ldots, V_n, W$ son L.D., es decir,

$$h_1 V_1 + h_2 V_2 + \ldots + h_n V_n + h_0 W = 0,$$

con alguna $h_0 \neq 0$. Es sencillo observar que $h_0 = 0$ y entonces

$$W = \underbrace{(-h_1/h_0)}_{\lambda_1} V_1 + \underbrace{(-h_n/h_0)}_{\lambda_2} V_2 + \ldots + \underbrace{(-h_n/h_0)}_{\lambda_n} V$$

En el otro sentido : Supongamos que $V_1, V_2, \ldots, V_n$ son una base para $\mathcal{V}$; hay que demostrar que estos vectores forman un conjunto máximo L.I.

Aplicando la definición anterior hay que verificar que :
a) $V_1, V_2, \ldots, V_n$ son L.I.

b) Para cualquier $W \in V$ se tiene que $V_1, V_2, \ldots, V_n, W$ son L.D.

El primer inciso se cumple por hipótesis, pues al ser $V_1, V_2, \ldots, V_n$ una base para V necesariamente estos vectores son L.I. En b) hay que demostrar que para $W \in V$ existen constantes $\lambda_1, \lambda_2, \ldots, \lambda_n, \lambda$ tales que $\lambda_1 V_1 + \lambda_n V_n + \ldots + \lambda V = 0$, con no todas las constantes nulas. Como $V_1, V_2, \ldots, V_n$ son una base, entonces se satisface que $h_1 V_1 + h_2 V_2 + \ldots + h_n V_n = W$, de donde podemos concluir el resultado.

Observación.- Sea $V(K) = R^n$. Aquí una base está formada por los vectores canónicos $e_1, e_2, \ldots, e_n$ (con $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_n = (0, 0, \ldots, 0, 1)$), pues puede comprobarse que dichos vectores son L.I. y además generan a R^n, ya que podemos expresar cualquier vector $(x_1, x_2, \ldots, x_n)$ como sigue :

$$(x_1, x_2, \ldots, x_n) = x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \ldots + x_n(0, 0, \ldots, 0, 1).$$

Pero en R^n, ¿siempre habrá bases con n elementos ?. O de manera aún mas general, ¿todas las bases para un mismo espacio vectorial tendrán el mismo número de elementos ?

Este es un resultado mucho muy importante que veremos en la siguiente sección.

4.7 M A S T E O R E M A S S O B R E D I M E N S I O N .

Debemos aclarar que el número de vectores de una base para un espacio vectorial $V(K)$ puede ser finito o infinito ; en el primer caso decimos que V es un espacio vectorial de dimensión finita y en el segundo caso hablaremos de un espacio de dimensión infinita. En este caso, una definicion para una base infinita es :

Definición 4.7.1.- Un conjunto infinito $W = \{W_1, W_2, \ldots, W_n\}$ es una base para $V(K)$ si :

a) Ningún elemento de W es combinación lineal de otros elementos de W,
b) W genera a V, es decir, $V = \mathcal{L}(W_1, W_2, \ldots, W_n)$.

A continuación veremos una serie de proposiciones y teoremas bastante relacionados con los conceptos ya vistos.

Proposición 4.7.2 (Reemplazo de Steinitz).- Sea $V(k)$ y $\{V_1, V_2, \ldots, V_n\}$ una base para V. Si $W = \{W_1, W_2, \ldots, W_m\}$ es un conjunto de vectores en V con $m > n$ entonces se tiene que $W_1, W_2, \ldots, W_m$ son L.D.

Demostración.- Como $\{V_1, V_2, \ldots, V_n\}$ es una base para V y todos los $W_i \in V$ (con $i = 1, 2, \ldots, m$), entonces existen constantes $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que $W_1 = \lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_n V_n$.

los anteriores es el siguiente :

Proposición 4.7.8.- Sea $\mathcal{V}$ (K) de dimensión n y $\mathcal{W}$ un subespacio vectorial de $\mathcal{V}$, entonces dim.$\mathcal{W}$ < dim.$\mathcal{V}$.

Si $\mathcal{W} = \{0\}$, no hay vectores L.I. y dim.$\mathcal{W} = 0$. Si consta de otros elementos además del cero, entonces contiene conjuntos de vectores L.I. de, a lo más, n elementos, por lo --- tanto los conjuntos máximos de vectores L.I. de $\mathcal{W}$ tienen a lo más n elementos ; es decir, una base para $\mathcal{W}$ tiene cuando mucho n elementos, de donde dim.$\mathcal{W} \leq n$ = dim.$\mathcal{V}$.

Por último, el siguiente resultado nos proporciona otra forma más de determinar si un conjunto de elementos de un espacio vectorial forma una base de él conociendo su dimensión.

Proposición 4.7.9.- Sea $\mathcal{V}$ (K) de dimensión n. Si $V_1, \ldots, V_n \in \mathcal{V}$ satisfacen que $\mathcal{L}(V_1, \ldots, V_n) = \mathcal{V}$, entonces $\{V_1, \ldots, V_n\}$ es una base para $\mathcal{V}$.

Demostración.- Como dim.$\mathcal{V}$ = n, si $V_1, \ldots, V_n$ son L.I. - entonces $\{V_1, \ldots, V_n\}$ es una base para $\mathcal{V}$. Si no es así, supongamos entonces que $V_1, \ldots, V_n$ son L.D., entonces

$$\lambda_1 V_1 + \ldots + \lambda_n V_n = 0,$$

con algún $\lambda_i \neq 0$. Sea $\lambda_1 \neq 0$, entonces

$$V_1 = (-\lambda_2/\lambda_1)V_2 + \ldots + (-\lambda_n/\lambda_1)V_n .$$

Siguiendo el mismo procedimiento para los otros V_i podemos - obtener un subconjunto de vectores $V_1', V_2', \ldots, V_r'$ que sean L.I. con r < n, y entonces todo lo que escribimos en términos de - $V_1, \ldots, V_n$ podemos reescribirlo ahora en términos de $V_1', \ldots, V_r'$ pero entonces $\mathcal{L}(V_1, \ldots, V_n) = \mathcal{V} = \mathcal{L}(V_1', \ldots, V_r')$, y $\{V_1', \ldots, V_r'\}$ sería una base para $\mathcal{V}$ con r < n elementos, lo cual no es posible.

4.8 ESPACIO DE LOS RENGLONES DE UNA MATRIZ, OBTENCION DE BASES, RANGO DE UNA MATRIZ.

Estudiaremos ahora ejemplos de espacios vectoriales relacionados con matrices ; veremos también cómo podemos construir bases para estos espacios mediante el proceso de reducción de una matriz dada a la forma escalonada a través de --- operaciones elementales.

Definición 4.8.1.- Sea A una matriz de mxn tal que

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

tos que cualquier base de $\mathcal{V}$ (teorema 4.7.2). De aquí concluimos que $\{V_1,\ldots,V_m\}$ es un conjunto máximo L.I. y por la proposición 4.6.5 es una base para $\mathcal{V}$.

Ahora pasaremos a relacionar estos conceptos entre dos conjuntos $\mathcal{V}$ y $\mathcal{W}$ que sean espacio y subespacio vectorial, respectivamente.

Proposición 4.7.5.- Sea $\mathcal{V}(K)$ de dimensión n y $\mathcal{W}$ un subespacio vectorial de dimensión m, entonces $n \geq m$.

Demostración.- Si $m > n$ aseguramos que en $\mathcal{W}$ hay bases con m elementos L.I. que también estarían en $\mathcal{V}$, lo cual nos llevaría a afirmar que en $\mathcal{V}$ hay conjuntos L.I. de más elementos que una base para $\mathcal{V}$, lo cual no es posible. Por lo tanto, $m < n$.

Proposición 4.7.6.- Sea $\mathcal{V}(K)$ de dimensión n y $\mathcal{W}$ subespacio vectorial con la misma dimensión, entonces $\mathcal{V} = \mathcal{W}$.

Demostración.- Si $\dim.\mathcal{W} = n$, podemos elegir una base arbitraria $\{W_1, W_2, \ldots, W_n\}$ cuyos elementos cumplen con que :
i) $W_1,\ldots,W_n$ son L.I.
ii) $\mathcal{L}(W_1,\ldots,W_n) = \mathcal{W}$.

De i) concluimos que $\{W_1,\ldots,W_n\}$ es una base para $\mathcal{V}$ por el teorema 4.7.4. y por esto mismo $\mathcal{L}(W_1,\ldots,W_n) = \mathcal{V}$. Pero entonces $\mathcal{W} = \mathcal{L}(W_1,\ldots,W_n) = \mathcal{V}$. De aquí que $\mathcal{V} = \mathcal{W}$.
Esto es, si la dimensión de un subespacio es igual a n, entonces coincide con el espacio completo pues de hecho una base de un subespacio consiste de un conjunto de vectores L.I. igual en número a la dimensión del espacio completo, por lo tanto también es una base para el espacio completo.

La siguiente proposición garantiza que cualquier conjunto de vectores L.I. de un espacio vectorial puede extenderse hasta llegar a formar una base de él.

Proposición 4.7.7.- Sea $\mathcal{V}(K)$ de dimensión n y r un entero positivo tal que $r < n$. Entonces para cualquier conjunto $\{V_1,\ldots,V_r\}$ que sea L.I. se pueden encontrar elementos $V_{r+1},\ldots,V_m$ tales que $\{V_1,\ldots,V_r,V_{r+1},\ldots,V_m\}$ es una base para $\mathcal{V}$.

Demostración.- Como $\{V_1,\ldots,V_r\}$ es L.I., si al agregar $W_0 \in \mathcal{V}$ el conjunto pasa a ser L.D. resultará que $\{V_1,\ldots,V_r\}$ es un conjunto máximo L.I. y, por lo tanto, una base para $\mathcal{V}$ con menos elementos que n, lo cual no es posible. Por lo tanto, existe $W_0 \in \mathcal{V}$ tal que $\{W_0,V_1,\ldots,V_r\}$ es L.I.
Sea $W_0 = V_{r+1}$, entonces $\{V_1,\ldots,V_r,V_{r+1}\}$ es L.I.
Si $r < n$ podemos proceder del mismo modo y llegar a obtener n elementos L.I. $V_1,\ldots,V_n$ que, por el teorema 4.7.4, debe ser una base para $\mathcal{V}$.

Los vectores renglón de A son $R_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, ------- $R_2 = (A_{21}, a_{22}, \ldots, a_{2n})$, ..., $R_m = (a_{m1}, a_{m2} \ldots, a_{mn})$, que son -- vectores en $\mathbb{R}^n$, generan un subespacio de R^n llamado Espacio de los Renglones de A. Esto es,

Espacio de los Renglones de A $= \mathcal{L}(R_1, R_2, \ldots, R_m)$

De la misma forma, las columnas de A (que son vectores - en R^m) forman un subespacio de R^m llamado Espacio de las -- Columnas de A, es decir

Espacio de las Columnas de A $= \mathcal{L}(A^1, A^2, \ldots, A^n)$.

Según la definición anterior tenemos que a una matriz A podemos asociarle dos espacios : el de los renglones y el de las columnas. Pero, ¿estos espacios se alterarán al cambiar los elementos de los vectores (renglón o columna) de A ? Al referirnos a "cambiar los elementos de los vectores" significa simplificar los elementos de A mediante operaciones ele--- mentales para transformarla en otra matriz B con elementos -- más sencillos y fáciles de operar con ellos.

Esta pregunta es interesante de responder puesto que da lugar a un concepto muy importante en matrices que nos pone - en condiciones de aclarar muchos detalles respecto a la teo-- ría de matrices y además facilita la construcción de bases de espacios vectoriales.

Recordemos los tres tipos de operaciones elementales entre renglones de A :

I) $R_i \longrightarrow R_j$

II) $R_i \longrightarrow kR_i$, $k \neq 0$.

III) $R_i \longrightarrow kR_j + R_i$.

Supongamos que mediante ellas obtenemos una matriz B. Bajo una operación elemental de tipo I, B y A tienen los mismos vectores renglón y, por tanto, el mismo espacio de ren--- glones. Mediante una operación elemental de tipo II o III -- cada renglón de B es claramente una combinación lineal de los renglones de A ; puesto que un espacio vectorial es cerrado - bajo la suma y la multiplicación por un escalar tenemos que - todo vector en el espacio de los renglones de B está también en el espacio de los renglones de A.

Por otro lado, si operamos en sentido contrario y a partir de B obtenemos a A, empleando los mismos argumentos, tendremos que el espacio de los renglones de A está contenido en el de B. Consecuentemente, A y B tienen el mismo espacio de renglones. Todo esto lo podemos resumir en el siguiente teorema :

Teorema 4.8.2.- Las operaciones elementales en los renglones de una matriz no alteran el espacio de renglones de -- ella.

Recuérdese que cuando una matriz A se transforma en otra

matriz B por medio de operaciones elementales entre renglones se dice que A y B son equivalentes por renglones. Con esto - podemos reescribir el teorema anterior como sigue :

Teorema 4.8.2'.- Matrices equivalentes por renglón tienen el mismo espacio de renglones.

De manera más particular, si tenemos dos matrices escalonadas A y B, éstas tendrán el mismo espacio de renglones si y sólo si sus renglones no nulos son iguales. Esto se formaliza en el siguiente teorema :

Teorema 4.8.3.- Sean A y B dos matrices escalonadas que se han reducido por renglones (*). A y B tienen el mismo espacio de renglones si y sólo si tienen iguales los renglones no nulos.

Demostración.- Si A y B tienen los renglones no nulos - iguales, es obvio que tienen el mismo espacio de renglones. Sólo falta probar que si tienen el mismo espacio de renglones entonces tienen los mismos renglones no nulos (iguales).

Supóngase que A y B tienen el mismo espacio de renglones y $R \neq 0$, con R = i-ésimo renglón de A. Entonces existen escalares $c_1, c_2, \ldots, c_s$ tales que

$$R = c_1 R_1 + c_2 R_2 + \ldots + c_s R_s \quad , \qquad (1)$$

donde R_i = renglones no nulos de la matriz B. Bastará con -- probar que $R = R_i$ o $c_i = 1$ y $c_k = 0$ con $i \neq k$.
Sea a_{ij_i} la primera componente no nula de R . De (1) tenemos que

$$a_{ij_i} = c_1 b_{1j_i} + c_2 b_{2j_i} + \ldots + c_s b_{sj_i} \quad (2)$$

Pero como b_{ij_i} es un elemento no nulo de B y ésta está reducida por renglones, es el único en la j_i-ésima columna de B, entonces $a_{ij_i} = c_i b_{ij_i}$. Como $a_{ij_i} = 1$ y $b_{ij_i} = 1$, entonces $c_i = 1$.

Supongamos ahora que $k \neq i$ y $b_{kj_k} \neq 0$ en R_k . Por (1) tenemos que

$$a_{kj_k} = c_1 b_{1j_k} + c_2 b_{2j_k} + \ldots + c_s b_{sj_k} .$$

Por los mismos argumentos anteriores $b_{kj_k} = 1$ y además $a_{kj_k} = c_k b_{kj_k}$. Como $a_{kj_k} = 0$, entonces $c_k = 0$. Por lo -- tanto $R = R_i$.

Observación.- Los anteriores resultados pueden aplicarse al espacio columna, trabajando con la transpuesta de la -- matriz original y aplicándolos a su espacio de renglones.

(*) Recuerde que al escalonar una matriz los elementos principales son 1's.

Pero dado que finalmente estamos hablando de espacios y

subespacios generados por los vectores renglón de determina-- das matrices, ¿podremos construir bases para estos subespa--- cios de vectores renglón ?.

Vimos por la definición 4.8.1. que los vectores renglón $R_1, R_2, \ldots, R_m$ generan un subespacio de R^n de modo tal que espacio de renglones = $\mathcal{L}(R_1, R_2, \ldots, R_m)$.

Si queremos encontrar una base para el espacio de ren--- glones de una matriz bastará con demostrar que $R_1, R_2, \ldots, R_m$ - son L.I. Pero este trabajo puede simplificarse aplicando el teorema 4.8.2. para manejar a la matriz escalonada equivalente a la matriz original y así reducir el número de vectores - con los que se va a operar (únicamente los vectores no nulos)

Si los vectores dados $R_1, R_2, \ldots, R_m$ pertenecen a un mismo espacio y queremos construir una base para el subespacio ge-- nerado por ellos, formamos una matriz A cuyos renglones estén determinados por $R_1, R_2, \ldots, R_m$ y aplicamos el razonamiento anterior. Como se observa, estas propiedades pertenecientes a teoría de matrices facilita en mucho la elaboración o cons--- trucción de bases de espacios y subespacios vectoriales .

Esto lo enunciamos en el siguiente teorema sin demostración.

Teorema 4.8.4.- Los vectores renglón no nulos de una -- matriz escalonada A forman una base para el espacio de los -- renglones de A.

El concepto tan importante al que hicimos referencia al inicio de esta sección al responder a la pregunta planteada - después de la definición 4.8.1. es el concepto de RANGO de -- una matriz. Recordemos que el espacio de renglones de A es - el subespacio de R^n generado por sus renglones y el espacio - columna de A es aquél generado por sus columnas. Las dimen-- siones del espacio de renglones de A y del espacio columna de A se llaman el Rango por Renglones y el Rango Columna de A, - respectivamente. Además, el valor común de ellos es el RANGO de A. Esto lo enunciamos en la siguiente definición.

Definición 4.8.5.- Las dimensiones del espacio renglón y el espacio columna de A se denominan el Rango por Renglones y el Rango Columna, respectivamente. El valor común de ellos es el RANGO de A y se denota rang(A).

Con ésto, el rango de una matriz es el máximo número de renglones (y columnas) independientes. Además, la obtención del rango de una matriz se simplifica con la reducción de --- ella a su forma escalonada.

Un resultado muy importante relacionado directamente con la existencia del rango de una matriz es el que afirma que

Teorema 4.8.6.- El rango por renglones y el rango columna de una matriz A son iguales.

Demostración.- Sea $A_{m\times n}$ la matriz

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

y sean $R_1, R_2, \ldots, R_m$ sus renglones : $R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \ldots, R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$. Supóngase que el rngo por renglones es r y que los vectores $S_1, S_2, \ldots, S_r$ forman una base del es-- pacio de los renglones, con $S_1 = (b_{11}, b_{12}, \ldots, b_{1n})$, $S_2 = (b_{21}, b_{22}, \ldots, b_{2n}), \ldots, S_r = (b_{r1}, b_{r2}, \ldots, b_{rn})$. Entonces cada R_i es combinación lineal de los S_i, es decir, existen constantes c_{ij} tales que

$$\begin{aligned} R_1 &= c_{11} S_1 + c_{12} S_2 + \ldots + c_{1r} S_r \\ R_2 &= c_{21} S_1 + c_{22} S_2 + \ldots + c_{2r} S_r \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ R_m &= c_{m1} S_1 + c_{m2} S_2 + \ldots + c_{mr} S_r \end{aligned}$$

Pero esto es :

$$\begin{aligned} (a_{11}, a_{12}, \ldots, a_{1n}) &= c_{11} S_1 + c_{12} S_2 + \ldots + c_{1r} S_r \\ &= c_{11} (b_{11}, \ldots, b_{1n}) + c_{12} (b_{21}, \ldots, b_{2n}) + \ldots \\ &\quad + c_{1r} (b_{r1}, \ldots, b_{rn}). \\ (a_{21}, a_{22}, \ldots, a_{2n}) &= c_{21} S_1 + c_{22} S_2 + \ldots + c_{2r} S_r \\ &= c_{21} (b_{11}, \ldots, b_{1n}) + c_{22} (b_{21}, \ldots, b_{2n}) + \ldots \\ &\quad + c_{2r} (b_{r1}, \ldots, b_{rn}). \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (a_{m1}, a_{m2}, \ldots, a_{mn}) &= c_{m1} S_1 + c_{m2} S_2 + \ldots + c_{mr} S_r \\ &= c_{m1} (b_{11}, \ldots, b_{1n}) + c_{m2} (b_{21}, \ldots, b_{2n}) + \ldots \\ &\quad + c_{mr} (b_{r1}, \ldots, b_{rn}). \end{aligned}$$

De aquí :

$$\begin{aligned} a_{1i} &= c_{11} b_{1i} + c_{12} b_{2i} + \ldots + c_{1r} b_{ri} \\ a_{2i} &= c_{21} b_{1i} + c_{22} b_{2i} + \ldots + c_{2r} b_{ri} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{mi} &= c_{m1} b_{1i} + c_{m2} b_{2i} + \ldots + c_{mr} b_{ri}, \end{aligned}$$

$i = 1, 2, \ldots, n$. Entonces

$$\begin{pmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{mi} \end{pmatrix} = \begin{pmatrix} c_{11} \\ c_{21} \\ \cdots \\ c_{m1} \end{pmatrix} b_{1i} + \begin{pmatrix} c_{12} \\ c_{22} \\ \cdots \\ c_{m2} \end{pmatrix} b_{2i} + \ldots + \begin{pmatrix} c_{1r} \\ c_{2r} \\ \cdots \\ c_{mr} \end{pmatrix} b_{ri} .$$

De lo que resulta que cada columna de A es una combinacion -- lineal de los vectores de constantes

$$\begin{pmatrix} c_{11} \\ c_{21} \\ \cdots \\ c_{m1} \end{pmatrix}, \begin{pmatrix} c_{12} \\ c_{22} \\ \cdots \\ c_{m2} \end{pmatrix}, \ldots, \begin{pmatrix} c_{1r} \\ c_{2r} \\ \cdots \\ c_{mr} \end{pmatrix}.$$

Entonces el espacio columna de A tiene a lo más dimensión r, es decir Rango Columna $\leq r$. (En otras palabras, rango co-

lumna ≤ rango de renglones).

Análogamente, considerando A^t concluimos que rango de renglones ≤ rango columna, con lo que rango renglón = rango columna.

4.9 APLICACIONES A LAS ECUACIONES LINEALES.

Sea el siguiente sistema de mxn sobre un campo K :

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

o lo que es lo mismo, la ecuación matricial AX=B. Recordemos que la matriz aumentada del sistema está dada por

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a & \ldots & a & b_m \end{array}\right] .$$

Obsérvese que al reducir esta matriz a su forma escalonada, el número de ecuaciones (o vectores renglón) independientes es igual al rango de la matriz aumentada. Además, reescribiendo el sistema como

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

vemos que existe solución siempre y cuando el vector columna B sea una combinación lineal de las columnas de la matriz A, es decir, si B pertenece al espacio columna de A. En otras palabras establecemos el siguiente teorema.

Teorema 4.9.1.- El sistema de ecuaciones lineales AX=B tiene una única solución si y sólo si la matriz de coeficientes y la matriz aumentada del sistema tienen el mismo rango.

Por último, si recordamos el teorema 2.6.4. se cumple que si u es una solución particular del sistema AX=B y W es la solución general del sistema AX = 0, entonces (u + W) es solución de AX = B, donde $W = (W_1, \ldots, W_n)$. Observe que $W \subset R^n$ y tiene, por tanto, una cierta dimensión. Para determinar esto, tenemos el siguiente teorema :

Teorema 4.9.2.- La dimensión del espacio solución del sistema de ecuaciones lineales homogeneo AX = 0 es n-r, con n el número de incógnitas y r el rango de la matriz de coeficientes A.

Demostración.- Si el sistema AX = 0 puede reducirse só-

lo a su forma escalonada, entonces por el teorema 2.5.3. b) - resulta que existen n-r variables libres, es decir, n-r ecuaciones independientes que forman, por lo tanto, una base para el espacio solución. Pero n = número de incógnitas en el --- sistema y r = número de ecuaciones no nulas (o renglones no - nulos) de la matriz A, ésto es, el rango de A. L.Q.Q.D.

C A P I T U L O 5

DETERMINANTES

5.1 I N T R O D U C C I O N .

En este capítulo vamos a desarrollar la teoría de los deter-- minantes y su importancia en la resolución de Sistemas de --- Ecuaciones Lineales. Hay que hacer notar, sin embargo, que - el uso de los determinantes no se restringe exclusivamente a éso. Si bien es cierto que nacieron a partir de esa idea, -- (recordemos el trabajo de Gabriel Cramer, que en 1750 dio una regla general para la resolución de Sistemas de Ecuaciones -- Lineales), en la actualidad este concepto se ha desarrollado tanto, que tiene valor por sí mismo. Esto es, se ha cons---- truido toda una teoría alrededor suyo y sus aplicaciones son- variadas: en Geometría, Física, en el estudio de valores y -- vectores propios de una matriz, en el procedimiento para lo-- grar la inversa de una matriz, etc.

Antes de iniciarnos en el estudio de los determinantes, debemos hacer otra aclaración: existen varias maneras de a--- bordar el tema. Una de las más usadas es aquélla que involu- cra el estudio de las permutaciones. Vamos a sacrificar un - tanto la formalidad que nos proporcionaría el enfocar el tema de esa forma, en aras de lograr una mayor facilidad en la --- comprensión del mismo. Trataremos de que la presentación da- da sea lo mas intuitiva posible, aprovechando para ello in--- terpretaciones geométricas.

Como una recomendación importante diremos que debe po--- nerse especial atención a las propiedades de los determinan-- tes, que son las que en un momento dado, nos permiten calcu-- larlo de una manera mas práctica.

Tal y como lo hemos venido haciendo hasta este momento, presentaremos algunos problemas que se resuelven mediante el empleo de los determinantes. Conforme se vaya avanzando en - la teoría se irán resolviendo, pero antes mostraremos un pe-- queño bosquejo histórico.

5.2 B O S Q U E J O H I S T O R I C O .

La teoría de los determinantes fue trabajada en el siglo XVIII por los matemáticos MacLaurin, Cramer, Bezout, Vander-- monde, Lagrange y Laplace, aunque algunos algoritmos utiliza- dos no fueron explicitados suficientemente. Su desarrollo -- más amplio tuvo lugar durante el transcurso del siglo XIX, -- donde los matemáticos Euler, Gauss, Cauchy, Jacobi y otros -- tuvieron gran influencia.

Otras contribuciones a la teoría de los determinantes -- la dieron matemáticos tales como Heinrich F. Scherk, quien -- formuló las reglas para la adición de dos determinantes que -

tienen en común una fila o una columna y para la multiplicación de un determinante por una constante.

Enunció también que el determinante de una matriz en la que un renglón es una combinación lineal de dos o más renglones es cero y que el valor del determinante de una matriz --- diagonal es igual al producto de los elementos de la diagonal principal. No se debe dejar de mencionar a Sylvester, que -- participó en el estudio de los determinantes durante más de - 50 años.

Cayley, fundador de la teoría de matrices, aportó tam--- bién resultados nuevos a la teoría de los determinantes, y al final del siglo XIX, Charles Dogson, mejor conocido como Le-- wis Carroll, y algunos otros, enriquecieron toda esta teoría con un caudal de nuevos conocimientos.

5.3 P R O B L E M A S D I V E R S O S .

Pasaremos, ahora sí, a enunciar algunos problemas de - - - - - aplicación de los determinantes.

Problema 1.- La Criptografía.- En muchas revistas hemos visto que con frecuencia aparecen problemas que consisten en descifrar algún mensaje misterioso. Son los llamados criptogramas. Para lograr descifrarlos, se proporciona el código - de traducción que permite hacerlo. Este pequeño ejemplo no - es más que una aplicación de lo más sencilla de una ciencia que en nuestros días, época de guerras militares y comercia-- les, ha cobrado gran auge. La llamada Criptografía: ciencia que hace y descifra códigos.

Es muy conocido por todos nosotros, que tanto las gran-- des potencias militares como los grandes emporios comerciales necesitan enviar y recibir una gran cantidad de información - secreta. Se han desarrollado técnicas enormemente sofisticadas.

Obviamente no pretendemos aquí desarrollarlas. Trataremos de ilustrar con un ejemplo sencillo, cómo las matrices, - sus inversas y sus determinantes, pueden utilizarse para re-- cibir y transmitir información en forma secreta.

Problema 2.- La recta que pasa por dos puntos.- Sabemos por nuestros conocimientos de Geometría Analítica, que si tenemos dos puntos diferentes del plano, de coordenadas (x_1, y_1) y (x_2, y_2) podemos hacer pasar una única línea recta que los - una. La ecuación de esta recta es:

$$c_1 x + c_2 y + c_3 = 0,$$

donde c_1, c_2, c_3 no pueden ser cero simultáneamente y son ú-- nicas para una recta dada.

¿Cómo podemos obtener la ecuación de esta línea recta -- empleando determinantes?

Problema 3.- Circunferencia que pasa por 3 puntos.- Si - tenemos como datos 3 puntos del plano diferentes y no alineados, de coordenadas (x_1, y_1), (x_2, y_2), (x_3, y_3), sabemos que por ellos pasa una única circunferencia, cuya ecuación es:

$$c_1(x^2 + y^2) + c_2x + c_3y + c_4 = 0.$$

¿Puede obtenerse la ecuación de esta circunferencia mediante el uso de los determinantes?

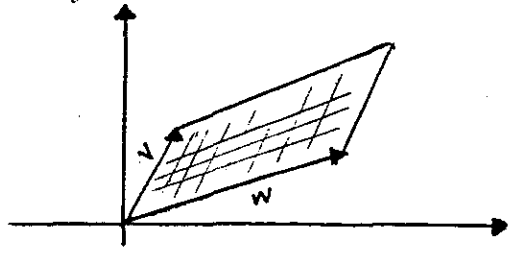
Problema 4.- ¿Pudiera seguirse un razonamiento similar a los anteriores para encontrar, digamos la ecuación del plano que pasa por 3 puntos o la ecuación de la esfera que pasa por 4 puntos?

Problema 5.- Trayectoria de una avión.- La trayectoria de un avión se puede considerar formada por segmentos de recta. Para que una construcción no interfiera con la trayectoria del avión, deberá estar situada bajo el segmento de la -- trayectoria. Las construcciones cercanas al aeropuerto tienen asignadas dos coordenadas, una de las cuales es la distancia al aeropuerto (x), y la otra su altura (y).

Se supone que la línea recta representada en la ecuación $\begin{vmatrix} x & y & 1 \\ 2 & 3 & 1 \\ 6 & 6 & 1 \end{vmatrix} = 0$, representa la trayectoria del avión cerca del aeropuerto. ¿Se podrá construir un edificio a 6 unidades del aeropuerto con 9 unidades de altura?

5.4 L A F U N C I O N D E T E R M I N A N T E .

Vamos a considerar el paralelogramo que descansa sobre - los vectores V y W, esto es, el paralelogramo cuyos lados son los vectores V y W.



Notación.- $S(V,W)$, cuando $V, W \in R^2$, denotará la superficie del paralelogramo construido en R^2 con dos lados que coinciden con los vectores V,W respectivamente.

Resulta claro que para cada pareja de vectores en R^2, -- siempre se puede construir el multicitado paralelogramo. Pudiéra pensarse entonces en la existencia de una función que a

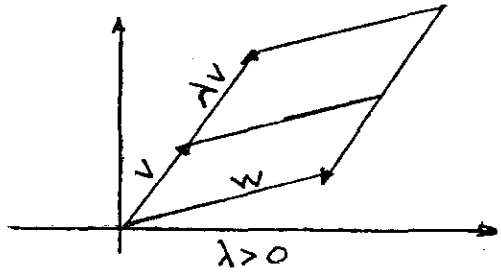
cada pareja $V, W \in R^2$, le asocia el número $S(V,W)$.

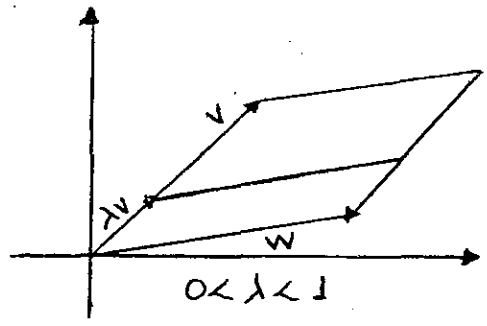
¿Qué propiedades interesantes se espera tenga esta función?

a) $S(kv, W) = k\, S(V, W)$,

a') $S(V, kW) = k\, S(V, W)$.

Es decir, al alargar uno de los lados del paralelogramo, en términos de operaciones vectoriales significaría que multiplicando a V o W por un escalar, el efecto que ocasiona es que el área queda multiplicada por ese mismo escalar. Por ejemplo:

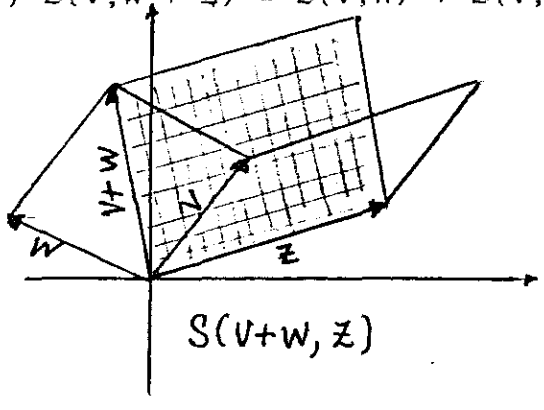


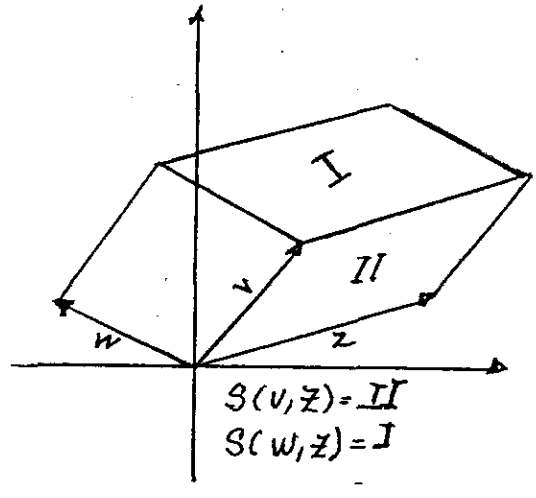


Observación.- La función $S(V, W)$ puede tomar valores positivos o negativos, o sea podemos hablar de área o superficie orientada.

b) $S(V + W, Z) = S(V, Z) + S(W, Z)$

b') $S(V, W + Z) = S(V, W) + S(V, Z)$.





Observación.- Esto es lo mismo que decir que S tiene un comportamiento lineal en cada componente.

c) $S(e_1, e_2) = 1$, donde $e_1 = (1,0)$, $e_2 = (0,1)$

d) $S(V,V) = 0$

Enunciaremos a continuación un teorema que, englobando las cuatro propiedades anteriores, nos define y nos proporciona además una manera de calcular el área orientada.

Teorema 5.4.1.- Si $S : R^2 \times R^2 \to R$ es una función que satisface las propiedades a), b), c), d) anteriores, entonces

si $V = (v_1, v_2)$, $W = (w_1, w_2)$,
$S(V, W) = v_1 w_2 - v_2 w_1$.

Esta función recibe el nombre de determinante y para -- calcularse pueden disponerse los vectores V y W como columnas de una matriz,

$$\det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} = v_1 w_2 - v_2 w_1 .$$

Para demostrar el teorema anterior es necesario demos--- trar primero el siguiente lema:

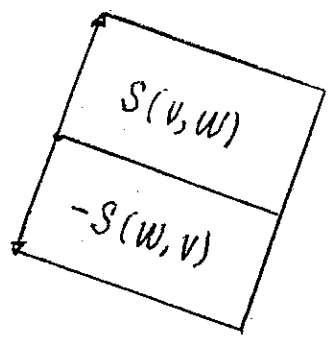
Lema 5.4.2.- $S(V,W) = -S(W,V)$.

Demostración.- Consideremos $S(V + W, V + W) = 0$, por d).

$S(V + W, V+W) = S(V,V+W) + S(W,V + W) = 0$, por b),
$= S(V,V) + S(V,W) + S(W,V) + S(W,W) = 0$,
$= S(V,V) + S(W,V) = 0$.

Por lo tanto, $S(V,W) = -S(W,V)$.

Geométricamente,



Volviendo entonces a la demostración del Teorema,

Demostración.- Verificaremos primero que efectivamente
$S(V,W) = v_1 w_2 - v_2 w_1$.

Sabemos que cualquier vector en R^2 se puede escribir como combinación lineal de los llamados vectores canónicos unitarios, $e_1 = (1,0)$ y $e_2 = (0,1)$. Esto vale pues, para V,W.

$V = v_1 e_1 + v_2 e_2 = v_1(1,0) + v_2(0,1)$
$= (v_1, 0) + (0, v_2) = (v_1, v_2)$
$W = w_1 e_1 + w_2 e_2 = w_1(1,0) + w_2(0,1)$
$= (w_1, 0) + (0, w_2) = (w_1, w_2)$
$S(V,W) = S(v_1 e_1 + v_2 e_2, w_1 e_1 + w_2 e_2)$
$= S(v_1 e_1, w_1 e_1 + w_2 e_2) + S(v_2 e_2, w_1 e_1 + w_2 e_2)$ (por b)),
$= S(v_1 e_1, w_1 e_1) + S(v_1 e_1, w_2 e_2) + S(v_2 e_2, w_1 e_1) +$
$S(v_2 e_2, w_2 e_2)$, (por b)),
$= v_1 S(e_1, w_1 e_1) + v_1 S(e_1, w_2 e_2) + v_2 S(e_2, w_1 e_1) + v_2 S(e_2, w_2 e_2)$
(por a)),

$= v_1 w_1 S(e_1, e_1) + v_1 w_2 S(e_1, e_2) + v_2 w_1 S(e_2, e_1) + v_2 w_2 S(e_2, e_2)$ (por a)),

$= v_1 w_1 (0) + v_1 w_2 (1) + v_2 w_1 (-1) + v_2 w_2 (0)$, (por c), d), y Lema 5.4.2)

$= v_1 w_2 - v_2 w_1$.

Para tener el teorema completamente demostrado hay que probar que S definida de tal manera, efectivamente satisface las propiedades a), b), c), d).

Empezando por a). P.D. $S(kV, W) = kS(V, W)$.

Para facilitar la demostración, utilizaremos el dispositivo nemotécnico que se menciona en párrafos anteriores,

$$\det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} = v_1 w_2 - v_2 w_1 ,$$

$k V = k(v_1, v_2) = (kv_1, k v_2)$, entonces

$$\det \begin{pmatrix} kv_1 & w_1 \\ kv_2 & w_2 \end{pmatrix} = kv_1 w_2 - kv_2 w_1 = k(v_1 w_2 - v_2 w_1) = kS(V, W).$$

b) P.D. $S(V + W, Z) = S(V, Z) + S(W, Z)$, donde $Z = (z_1, z_2)$ y $V + W = (v_1 + w_1, v_2 + w_2)$.

$$\det \begin{pmatrix} v_1+w_1 & z_1 \\ v_2+w_2 & z_2 \end{pmatrix} = (v_1 + w_1) z_2 - (v_2 + w_2) z_1$$

$$= v_1 z_2 + w_1 z_2 - v_2 z_1 - w_2 z_1$$

$$= (v_1 z_2 - v_2 z_1) + (w_1 z_2 - w_2 z_1)$$

$$S(V, Z) + S(W, Z).$$

c) P.D. $S(e_1, e_2) = 1$.

$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 - 0 = 1 .$$

d) P.D. $S(V, V) = 0$.

$$\det \begin{pmatrix} v_1 & v_1 \\ v_2 & v_2 \end{pmatrix} = v_1 v_2 - v_2 v_1 = 0.$$

Con lo que queda completamente demostrado el teorema.

Observación.- Como ya se mencionó, la función determinante puede considerarse como una función cuyo dominio es el conjunto de las matrices de 2x2 y cuyo contradominio es el conjunto de los números reales. De aquí en adelante así lo trabajaremos.

Algunas propiedades extras de los determinantes de 2x2:

1) Si B es la matriz que resulta de sumar un múltiplo de una columna de A a otra, entonces det A = det B, ésto es,

$$\det A = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} = \det B = \begin{pmatrix} v_1 + kw_1 & w_1 \\ v_2 + kw_2 & w_2 \end{pmatrix}$$

$$= (v_1 + kw_1)w_2 - (v_2 + kw_2)w_1$$
$$= v_1 w_2 + kw_1 w_2 - v_2 w_1 - kw_2 w_1$$
$$= v_1 w_2 - v_2 w_1 .$$

2) El determinante de la matriz A es igual al determinante de la transpuesta de A.

$$\det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} = \det \begin{pmatrix} v_1 & v_2 \\ w_1 & w_2 \end{pmatrix} = v_1 w_2 - v_2 w_1 .$$

Esta última propiedad es muy interesante, pues permite - asegurar que todas las aseveraciones que hemos hecho para las columnas de A son válidas también para los renglones.

Observación.- No debemos dejar pasar más tiempo sin re-- calcar que el determinante solamente está definido para ma--- trices cuadradas.

Sin gran dificultad y en base a que ya se ha definido -- claramente el determinante de 2x2, podemos pasar a definir el determinante de 3x3. Aclaramos que esta definición quedará - en términos del determinante de 2x2.

Definición 5.4.3.- Sea A una matriz de 3x3.

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

definimos el determinante de A, mediante su desarrollo por el primer renglón de la siguiente manera:

$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$, donde A_{1j}, $j = 1, 2, 3$, es la matriz que se obtiene de eliminar el primer ren--- glón y la j-ésima columna de A, esto es:

$$\det A = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} +$$
$$a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$

Haciendo el desarrollo:

$$\det A = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$
$$= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22}.$$

Notemos que al hacer el cálculo del determinante de A -- mediante el desarrollo por el primer renglón, lo que se ha - hecho es expresar al determinante de A como una suma de 3 --- términos, cada uno de los cuales es un producto de 2 facto--- res, uno de ellos es un elemento del primer renglón y el otro es el determinante de la matriz que se obtiene al eliminar el

primer renglón y la columna donde se encuentra el elemento -- que se mencionó anteriormente.

Los signos de cada uno de los 3 términos iniciales se asigna mediante la siguiente técnica:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

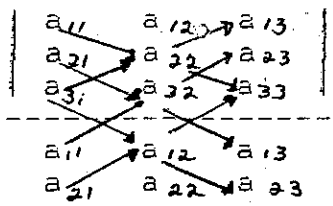
Notación.- El determinante se designará mediante la expresión $|A|$, "det A" o encerrando las componentes de A entre barras verticales.

Ejemplo: Si $A = \begin{pmatrix} 5 & 6 \\ 4 & 2 \end{pmatrix}$, $\det A = \begin{vmatrix} 5 & 6 \\ 4 & 2 \end{vmatrix} = |A|$.

Observaciones:
- El desarrollo anterior hubiera podido realizarse no únicamente por el primer renglón, sino por cualquiera de los renglones de A.
- También hubiera podido efectuarse, siguiendo el mismo esquema, mediante cualquiera de las columnas de A.

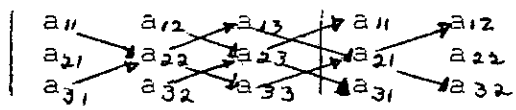
El cálculo de un determinante de 3x3 se efectúa frecuentemente mediante las siguientes técnicas, que procuran simplificar de alguna manera los cálculos, o más bien procuran que sea fácilmente memorizable la manera de obtenerlo.

Diagrama 1.- En este esquema se repiten los dos primeros renglones de A, sumándose los productos indicados con flechas. Los que tengan flechas hacia abajo se multiplican por 1 (se dejan igual, pues) y los que tengan flechas hacia arriba se multiplican por -1.



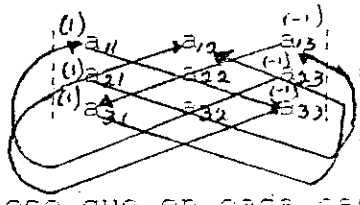
$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}.$$

Diagrama 2.- Aquí se repiten las dos primeras columnas, sumándose los productos indicados con flechas y con la misma observación para el signo que el caso anterior.



$$= a_{11}a_{22}a_{23} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$$

Diagrama 3.- Esta forma es conocida como "regla de la canasta". También se suman los productos indicados con flechas, multiplicándolos por el factor que aparece al inicio de la flecha.



$$= a_{11}\,a_{22}\,a_{33} + a_{21}\,a_{32}\,a_{13} + a_{31}\,a_{23}\,a_{12}$$
$$- a_{13}\,a_{22}\,a_{31} - a_{23}\,a_{32}\,a_{11} - a_{33}\,a_{21}\,a_{12}$$

Nótese que en cada caso se obtiene el mismo resultado que en la definición original.

Observación.- Todas las propiedades que se enunciaron anteriormente para los determinantes de 2x2 se pueden enunciar y demostrar para los determinantes de 3x3. Las demostraciones son un tanto sencillas, basta con efectuar el cálculo directo. Las omitiremos por ello, pero cuando se defina el determinante para el caso general, y se vea que también satisface esas mismas propiedades, se demostrarán algunas de ellas.

No está demás el advertir que las técnicas descritas por los diagramas 1,2,3, sólo son válidas para el determinante de 3x3.

5.5 E L P R O D U C T O V E C T O R I A L .

Antes de seguir adelante, haremos un alto para, aprovechando lo que hemos visto, definir una operación entre vectores que no se ha manejado anteriormente y que tiene características tan interesantes como útiles.

Adelantaremos que esta operación, llamada Producto Cruz o Producto Vectorial, solamente está definida para vectores en R^3. A continuación damos la definición formal:
Definición 5.5.1.- Sean $V,W \in R^3$ tales que $V = (v_1, v_2, v_3)$, $W = (w_1, w_2, w_3)$. Sabemos que cualquier vector se puede poner en términos de los vectores canónicos, en este caso de R^3.

$i = (1,0,0)$, $j = (0,1,0)$, $k = (0,0,1)$. (Se les ha llamado i,j,k, en lugar de la notación acostumbrada e_1, e_2, e_3 porque usualmente así se les denota en física, y el producto cruz es muy usado en esa rama del conocimiento humano).

$$V = v_1\,i + v_2\,j + v_3\,k \quad ; \quad W = w_1\,i + w_2\,j + w_3\,k.$$

Definimos entonces al Producto Cruz o Producto Vectorial de V con W, denotado VXW, al vector obtenido mediante el cálculo del determinante

$$\begin{vmatrix} i & j & k \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

La siguiente definición también es importante:

Definición 5.5.2.- Sean $v, W, Z \in R^3$. Definimos

$$V \cdot (WXZ) = \begin{vmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

Es fácil verificar que esta afirmación es válida.

Observación.- El producto vectorial satisface la importante propiedad de que es perpendicular tanto a V como a W y además satisface la Igualdad de Lagrange, misma que la relaciona con el producto punto.

Formalmente, establecemos:

Teorema 5.5.3.- Sean $V, W \in R^3$, para ellos se cumple:

a) $V \cdot (VXW) = 0$, es decir, VXW es ortogonal a V.
b) $W \cdot (VXW) = 0$, es decir, VXW es ortogonal a W.
c) $\| VXW \|^2 = \| V \|^2 \| W \|^2 - (V \cdot W)^2$. (Esta es la llamada Igualdad de Lagrange).

Demostración:

a) $V \cdot (VXW) = \begin{vmatrix} v_1 & v_2 & v_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = 0$, por ser un determinante que tiene dos renglones iguales.

b) Demostración similar a la anterior.

c) Basta desarrollar ambos lados de la igualdad y es fácilmente verificable que se satisface.

Además de las anteriores, VXW satisface:

Teorema 5.5.4.- Sean $V, W, Z \in R^3$, $k \in R$, tenemos:

a) $VXW = -(WXV)$
b) $VX(W + Z) = (VXW) + (VXZ)$
c) $(V + W)XZ = (VXZ) + (WXZ)$
d) $k(VXW) = (kV)XW = VX(kW)$
e) $VX\theta = \theta XV = \theta$
f) $VXV = \theta$

Todas estas definiciones son fáciles de comprobar, basta aplicar la definición.

Utilizando la definición de producto vectorial puede darse también la interpretación geométrica del determinante de 2x2.

Para ello se enuncia:

Proposición 5.5.5.- $\| VXW \| = \| V \| \| W \| \operatorname{Sen} \theta$, donde

θ es el ángulo formado por V y W.

Demostración.- Se probará la proposición equivalente:

$$|| VXW ||^2 = || V ||^2 || W ||^2 \operatorname{Sen}^2 \theta.$$

Por un lado:

$$\begin{aligned} || V ||^2 || W ||^2 \operatorname{Sen}^2 \theta &= || V ||^2 || W ||^2 (1 - \operatorname{Cos}^2 \theta) \\ &= || V ||^2 || W ||^2 - || V ||^2 || W ||^2 \operatorname{Cos}^2 \theta \\ &= || V ||^2 || W ||^2 - || V ||^2 || W ||^2 (V \cdot W / ||V|| \, ||W||)^2 \\ &= || V ||^2 || W ||^2 - (V \cdot W)^2 \quad \ldots\ldots\ldots\ldots (1) \end{aligned}$$

Por otro lado:

$$\begin{aligned} || VXW ||^2 &= (v_2 w_3 - v_3 w_2)^2 + (v_1 w_3 - v_3 w_1)^2 + (v_1 w_2 - v_2 w_1)^2 \\ &= v_2^2 w_3^2 - 2 v_2 w_3 v_3 w_2 + v_3^2 w_2^2 + v_1^2 w_3^2 - 2 v_1 w_3 v_3 w_1 \\ &\quad + v_3^2 w_1^2 + v_1^2 w_2^2 - 2 v_1 w_2 v_2 w_1 + v_2^2 w_1^2 \\ &= v_1^2 (w_3^2 + w_2^2) + v_2^2 (w_3^2 + w_1^2) + v_3^2 (w_1^2 + w_2^2) - \\ &\quad v_1 w_1 (w_3 v_3 + v_2 w_2) - v_2 w_2 (w_3 v_3 + v_1 w_1) - \\ &\quad v_3 w_3 (v_2 w_2 + v_1 w_1) - V W (V W + V W) \\ &= (v_1^2 + v_2^2 + v_3^2)(w_1^2 + w_2^2 + w_3^2) - \\ &\quad (v_1 w_1 + v_2 w_2 + v_3 w_3)^2 \\ &= || V ||^2 || W ||^2 - (V \cdot W)^2 \quad \ldots\ldots\ldots (2) \end{aligned}$$

De (1) y (2) se concluye que

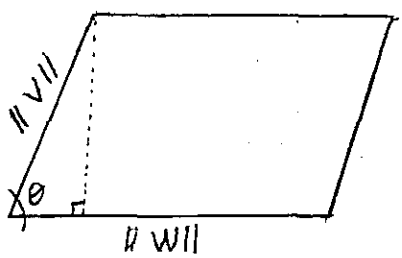
$$|| VXW ||^2 = || V ||^2 || W ||^2 \operatorname{Sen}^2 \theta, \text{ por lo tanto,}$$

$$|| VXW || = +- || V || \, || W || \operatorname{Sen} \theta.$$

Consideramos $0 < \theta < 180^\circ$, entonces

$$|| VXW || = || V || \, || W || \operatorname{Sen} \theta .$$

Si nuevamente observamos el paralelogramo generado por V y W,



$\operatorname{Sen} \theta = h / || V ||$
$h = || V || \operatorname{Sen} \theta$
$\text{Area} = (||V|| \operatorname{Sen} \theta) ||W||$
$\text{Area} = ||W|| \, ||V|| \operatorname{Sen} \theta$

Concluimos que Area = || VXW ||.

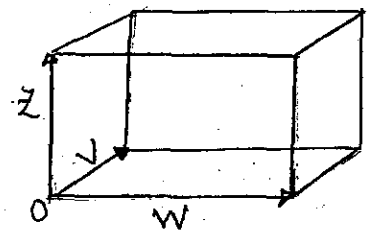
Las causas por las que se ha definido el producto vectorial así, son que resulta ser de gran utilidad en física. -- Con gran frecuencia se pueden encontrar cantidades físicas - que son vectores, cuyo producto, definido en la forma ante--- rior, es una cantidad vectorial que tiene un significado fí-- sico importante.

Algunos ejemplos que se pueden dar son el movimiento es-

tático, la cantidad de movimiento angular, la fuerza sobre -- una carga que se mueve en un campo magnético y el flujo de -- energía electromagnética. En análisis vectorial también es - importante, pues || VXW || aparece cuando se desea calcular - la integral de una superficie.

Pasaremos ahora a tratar la interpretación geométrica de los determinantes de 3x3.

Para ello fijémonos en el paralelepípedo determinado por los vectores $V = (v_1, v_2, v_3)$, $W = (w_1, w_2, w_3)$, $Z = (z_1, z_2, z_3)$.



El volumen de cualquier paralelepípedo es V^* = (área de la base)(altura). La altura es la longitud de la proyección del vector Z sobre la base del paralelepípedo, que como ya se vió, tiene por área || VXW ||.

Entonces, el volumen del paralelepípedo será:

$$V^* = (\| VXW \|) \left\| \frac{Z\cdot(VXW)}{(VXW)\cdot(VXW)} \cdot (VXW) \right\|$$

$$= \| VXW \| \frac{| Z\cdot(VXW) |}{\| VXW \|^2} \| VXW \| =$$

$$= \| VXW \| \left(\frac{Z\cdot(VXW)}{\| VXW \|} \right)$$

$$= | Z\cdot(VXW) |.$$

Desarrollando:

$$\begin{aligned} |Z\cdot(VXW)| &= | (z_1, z_2, z_3)\ (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_3 w_1) \\ &= | z_1 (v_2 w_3 - v_3 w_2) + z_2 (v_3 w_1 - v_1 w_3) + z_3 (v_1 w_2 - v_2 w_1) | \\ &= | z_1 (v_2 w_3 - v_3 w_2) - z_2 (v_1 w_3 - v_3 w_1) + z_3 (v_1 w_2 - v_2 w_1) | \\ &= | z_1 \ \text{Det}\ V^*_{11} - z_2 \ \text{Det}\ V^*_{21} + z_3 \ \text{Det}\ V^*_{31} |. \end{aligned}$$

Que no es otra cosa que el valor absoluto del determi--- nante de la matriz

$$V^* = \begin{pmatrix} z_1 & v_1 & w_1 \\ z_2 & v_2 & w_2 \\ z_3 & v_3 & w_3 \end{pmatrix}$$

cuando se ha desarrollado por la primera columna.

Las interpretaciones geométricas que se han hecho anteriormente se pueden generalizar. Pudiéramos hablar por ejemplo, de que el determinante de la matriz con columnas V_1, V_2, ..., V_n, $V_i \in R^n$, es igual al volumen del "cuerpo" generado por V_1, V_2, ..., V_n.

5.6 GENERALIZACION DEL CONCEPTO DE DETERMINANTE.

Hasta este momento sólo se ha trabajado con matrices de 2x2 y de 3x3. No hay nada que nos impida hablar del determinante de una matriz de nxn. Al igual que en el caso de 3x3, la definición se puede dar en términos del desarrollo por renglones o desarrollo por alguna columna.

Definicion 5.6.1.- Si A es una matriz de nxn, de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

entonces el determinante de A se puede calcular como:

$$\det A = |A| = a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}| + \dots +- a_{1n}|A_{1n}|,$$

donde $|A_{1j}|$ es el determinante de la matriz obtenida al eliminar el primer renglón y la j-ésima columna de A.

El signo que antecede a cada uno de los sumandos anteriores se obtiene del siguiente diagrama que es muy fácil de recordar si observamos que al elemento a_{11} le corresponde siempre el signo + y que, enseguida de él, por columna o renglón no puede haber un signo igual. La misma observación se hace a cada elemento.

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ . & . & . & \dots \\ . & . & . & \dots \end{pmatrix}$$

Este signo puede obtenerse anteponiéndole a cada sumando el factor $(-1)^{i+j}$.

Observaciones:

-Cuando el determinante de A se ha obtenido de la manera anterior, se dice que se ha expandido o desarrollado mediante el primer renglón.

- La expansion no necesariamene ha de realizarse por el primer renglón. Puede llevarse a cabo mediante cualquier renglón o mediante cualquier columna, aseveración que se demuestra al final de este capítulo, cuando se hayan demostrado algunas afirmaciones un poco más sencillas que emplean el mismo método de demostración: la Inducción Matemática.

Pudiéramos redefinir entonces el determinante de A, de la siguiente manera:

Definición 5.6.2.- Si A es una matriz de nxn, su determinante puede calcularse, si se desarrolla mediante el i-ésimo renglón,

$|A| = a_{i1}|A_{i1}| - a_{i2}|A_{i2}| + a_{i3}|A_{in}| - a_{i4}|A_{i4}| + \ldots$

$\ldots + (-1)^{i+n} a_{in}|A_{in}|$, donde $|A_{ij}|$ es el determinante de la matriz obtenida al eliminar el i-ésimo renglón y la j-ésima columna de A, j, i = 1,2,... , n, o mediante :

$|A| = a_{1j}|A_{1j}| - a_{2j}|A_{2j}| + a_{3j}|A_{3j}| - a_{4j}|A_{4j}| + \ldots$

$\ldots + (-1)^{n+j} a_{nj}|A_{nj}|$, si se ha expandido mediante la j-ésima columna.

Las propiedades que se enunciaron para determinantes de 2x2 son válidas también en el caso general. Las enunciaremos de nueva cuenta, aunque en su demostración intervienen argumentos engorrosos. Para no pecar de totalmente informales, se demostrarán aquellos resultados de razonamiento sencillo.

Casi todas, por no decir que la totalidad de las demostraciones se puede verificar mediante el empleo del Método de Inducción Matemática, que básicamente consiste en lo siguiente:

i) Se desea demostrar una proposición que se asegura vale para todos los naturales.

ii) Se comprueba que la proposición vale para ciertos naturales, digamos n = 1,2, ..., etc.

iii) Se elabora la hipótesis de inducción; ésto es, se supone válida la proposición para un natural arbitrario.

iv) A partir de aquí, se debe demostrar que la proposición vale para cualquier natural.

Observación.- Esta es una de las maneras de enunciar el Principio de Inducción Matemática.

Finalmente, antes de enunciar las propiedades, introduciremos una notación que es muy utilizada y que nos permitirá considerar al determinante de A como función de sus renglones o de sus columnas.

Notación.- El determinante de A puede denotarse:
$|A| = \det A = D(A_1, A_2, \ldots, A_n)$ cuando se considera como una función de los renglones de A, o $\det A = D(A^1, A^2, \ldots, A^n)$ cuando se considera como una función de las columnas de A.

Enunciaremos entonces:

Proposición.5.6.3.- Si A es una matriz de nxn, el determinante de A satisface:

Ahora se escoge una matriz de, en este caso 2x2, (el tamaño de la matriz se escoge de acuerdo al tamaño de las partes en que se ha fraccionado el mensaje), tal que tenga componentes enteros, que sea invertible y que su determinante sea +1 o -1. Todas estas condiciones aseguran que A^{-1} también tendrá componentes enteros.

Tomemos por ejemplo:

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

Ya que se ha escogido A, la multiplicamos por cada una de las matrices de 2x1 (o vectores columna de 2 componentes).

$$A\begin{pmatrix} 14 \\ 1 \end{pmatrix} = \begin{pmatrix} 57 \\ 43 \end{pmatrix}.$$

$$A\begin{pmatrix} 19 \\ 1 \end{pmatrix} = \begin{pmatrix} 77 \\ 58 \end{pmatrix}.$$

$$A\begin{pmatrix} 14 \\ 10 \end{pmatrix} = \begin{pmatrix} 66 \\ 52 \end{pmatrix}.$$

$$A\begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ 7 \end{pmatrix}.$$

$$A\begin{pmatrix} 22 \\ 10 \end{pmatrix} = \begin{pmatrix} 100 \\ 78 \end{pmatrix}.$$

$$A\begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 17 \\ 14 \end{pmatrix}.$$

Como se ve, se ha encontrado un nuevo conjunto de vectores, que acomodándolos, nos dan el arreglo:
57 43 77 58 66 52 8 7 100 78 17 14 ,
que sería el que transmitiríamos.

Para descifrarlo, necesitamos conocer la matriz A y A^{-1}. Fijémonos:

$$A\begin{pmatrix} 14 \\ 1 \end{pmatrix} = \begin{pmatrix} 57 \\ 43 \end{pmatrix}, \text{ entonces } \begin{pmatrix} 14 \\ 1 \end{pmatrix} = A^{-1}\begin{pmatrix} 57 \\ 43 \end{pmatrix}.$$

En este caso,

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

$$A^{-1}\begin{pmatrix}57\\43\end{pmatrix} = \begin{pmatrix}14\\1\end{pmatrix} = \text{NA}.$$

$$A^{-1}\begin{pmatrix}77\\58\end{pmatrix} = \begin{pmatrix}19\\1\end{pmatrix} = \text{RA}.$$

$$A^{-1}\begin{pmatrix}66\\52\end{pmatrix} = \begin{pmatrix}14\\10\end{pmatrix} = \text{NJ}.$$

$$A^{-1}\begin{pmatrix}8\\7\end{pmatrix} = \begin{pmatrix}1\\4\end{pmatrix} = \text{AD}.$$

$$A^{-1}\begin{pmatrix}100\\78\end{pmatrix} = \begin{pmatrix}22\\12\end{pmatrix} = \text{UL}.$$

$$A^{-1}\begin{pmatrix}17\\14\end{pmatrix} = \begin{pmatrix}3\\5\end{pmatrix} = \text{CE}.$$

La matriz A recibe el nombre de Matriz de Codificación y A^{-1} es llamada Matriz de Desciframiento o Descodificación.

Es claro que este proceso se puede volver tan complicado como se quiera, escogiendo particiones más grandes y por ende, matrices de mayor tamaño.

5.9.- Vamos a agregar un último detalle. Se ha estado mencionando constantemente, desde que se definió el determinante mediante su desarrollo por el primer renglón, que este número será el mismo independientemente de el renglón, e incluso la columna, ppor el cual se haga el desarrollo. Entonces, lo que pretendemos a continuación es demostrar esta aseveración, misma que enunciaremos como el:

Teorema 5.9.1.- El determinante de una matriz es el mismo, independientemente del renglón o la columna por el (la) cual se desarrolle. En símbolos:

$$\det A = \sum_{j=1}^{n} a_{pj}(-1)^{p+j} D(A_{pj}) = \sum_{i=1}^{n} (-1)^{i+q} a_{iq} D(A_{iq}).$$

Demostración.- Se hará también por inducción sobre n, es es decir sobre el tamaño de la matriz.

Vamos a verificar que la proposición vale para n = 2.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} D(A_{11}) + (-1)^{1+2} a_{12} A_{12}$$

$$= a_{11} a_{22} - a_{12} a_{21} \text{ (primer renglón)};$$

$$= (-1)^{2+1} a_{21} D(A_{21}) + a_{22} D(A_{22})$$
$$= -a_{21} a_{12} + a_{22} a_{11} \quad \text{(segundo renglón)};$$
$$= (-1)^{1+1} a_{11} D(A_{11}) + (-1)^{2+1} a_{21} D(A_{21})$$
$$= a_{11} a_{22} - a_{21} a_{12} \quad \text{(primera columna)};$$
$$= (-1)^{1+2} a_{12} D(A_{12}) + (-1)^{2+2} a_{22} D(A_{22})$$
$$= - a_{12} a_{21} + a_{22} a_{11} \quad \text{(segunda columna)}.$$

Como se ve, al comparar los resultados, nos damos cuenta de que el valor obtenido es el mismo. Elaboramos pues la hipótesis de inducción. Suponemos que la proposición vale para matrices de tamaño n-1.

P.D. que vale para matrices de tamaño n.

Haremos esto en tres etapas:

a) Se verificará que el desarrollo por el primer renglón es el mismo que el desarrollo por la primera columna.

b) Se demostrará que el desarrollo por el primer renglón es el mismo que el desarrollo por cualquier renglón.

c) Por último, veremos que el desarrollo por la primera columna es el mismo que por cualquier columna.

Empezamos:

a) Demostraremos a), en símbolos,

$$\sum_{i=1}^{n} (-1)^{i+1} a_{i1} D(A_{i1}) = \sum_{j=1}^{n} a_{1j} D(A_{1j}) (-1)^{1+j}.$$

Es claro que al desarrollar por la primera columna, tenemos una expresión de n sumandos, en cada uno de los cuales se ha de calcular el determinante de una matriz de tamaño n-1 que carece del i-ésimo renglón y de la primera columna. Esto es, el término a_{i1} no aparece en alguno de los A_{i1}.

Similarmente, al desarrollar por el primer renglón, los a_{1j} no aparecen en alguno de los A_{1j}.

El coeficiente de a_{11} es el mismo en ambos lados de nuestra igualdad. Analizaremos el coeficiente del término $a_{i1} a_{1j}$, para $i>1$, $j>1$. Por la izquierda, este coeficiente lo obtenemos al desarrollar A_{i1} y sería:

$$(-1)^{i+1} (-1)^{j-1+1} D(A_{i1,1j}) = (-1)^{i+j+1} D(A_{i1,1j}).$$

¿Por qué? Recordemos que a_{1j} se encuentra colocado en la columna j-1 de A_{i1} y además, que estamos suponiendo que el teorema vale para determinantes de orden n-1.

$A_{i1,1j}$ es la matriz de tamaño n-2 obtenida al eliminar los renglones i-ésimo y primero de las columnas primera y j-ésima..

Vamos a ser más explícitos:

Sea $\sum_{i=1}^{n} (-1)^{c+1} a_{c,1} D(A_{c,1})$, ¿$D(A_{c,1})$ =? Es el determinante de la matriz

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{nn} \end{pmatrix} \longleftarrow .$$

Si desarrollamos este determinante mediante el renglón señalado, (aplicando la hipótesis de inducción):

$D(A_{c,1}) = \sum_{j=2}^{n} a_{1j} (-1)^{1+j-1} D(A_{c,1,1j})$, sustituyendo:

$D(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \sum_{j=2}^{n} a_{1j} (-1)^{1+j-1} D(A_{i,1,1j})$.

El coeficiente como ya se dijo de $a_{i,1} a_{1j}$ es:

$(-1)^{i+1+j} D(A_{i,1,1j})$.

Un argumento muy parecido se utiliza al analizar el extremo derecho de la igualdad que se quiere probar.

Sea $\sum_{j=1}^{n} a_{1j} (-1)^{1+j} D(A_{1j})$, ¿$D(A_{1j})$ = ? Es el determinante de la matriz

$$\begin{matrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{matrix} .$$

Al desarrollar este determinante por la columna señalada (aplicando la hipótesis de inducción), tenemos:

$D(A_{1j}) = \sum_{i=2}^{n} (-1)^{i-1+1} a_{i,1} D(A_{1,i,j,1})$.

Sustituyendo:

$D(A) = \sum_{j=1}^{n} a_{1j} (-1)^{1+j} \left(\sum_{i=2}^{n} (-1)^{i} a_{i,1} D(A_{1,i,j,1}) \right)$.

El coeficiente del elemento $a_{1j}\, a_{i1}$ es $(-1)^{1+j+i} D(A_{1,i,j,1})$ que es el mismo que en el término de la izquierda de la igualdad deseada, que es lo que queríamos demostrar.

Pasaremos ahora a b), que en símbolos, significaría demostrar:

$\sum_{i=1}^{n} (-1)^{1+j} a_{ij} D(A_{ij}) = \sum_{k=1}^{n} a_{pk} D(A_{pk}) (-1)^{p+k}$, $p = 2,3,..,n$, --llamaremos (1) a la expresión anterior.

Retomaremos el lado izquierdo de (1),

$$\begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pj} & \cdots & a_{pk} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}$$

Cuando $j>k$, a_{pk} queda situado en la p-ésima columna y el renglón (k-1) de A_{ij}. Si $j<k$, entonces a_{pk} queda colocado en el renglón (k-1) y la columna (p-1) de A_{ij}.

Si se considera ahora el lado derecho de 1) y se examina (usando un razonamiento muy similar al que ya se hizo en la primera parte) que coeficiente tiene $a_{pk}a_{ij}$, se ve que éste es:

$$(-1)^{p+k}(-1)^{i+j-1}\ D(A_{pi},{}_{kj}),\quad (\text{si } j>k),$$

$$(-1)^{p+k}(-1)^{i+j}\ D(A_{pi},{}_{kj}),\quad (\text{con } j<k),$$

y que coincide con lo obtenido en (2).

c) Por último, si aplicamos el resultado b) a la transpuesta de A, vemos que el desarrollo del determinante de A por la q-ésima columna da el mismo resultado que por la primera columna.

De a), b), c) concluimos la veracidad del teorema.

C A P I T U L O 6

TRANSFORMACIONES LINEALES.

6.1 I N T R O D U C C I O N.

Con este capítulo terminamos lo que creemos es suficiente material para cubrirse dentro del curso de Algebra Lineal I y que deja sentadas las bases, tanto teóricas como prácticas para que el alumno continúe posteriormente sus estudios dentro del mismo campo.

Hablar de las transformaciones lineales es remontarse a una de las ideas pilares de las matemáticas: el concepto de función. Resulta que las transformaciones lineales no son más que funciones que tienen la particularidad de que tanto su dominio como su contradominio son espacios vectoriales. Es por ello que pensamos que el manejo de la teoría que se va a cubrir en este capítulo no se vá a dificultar, dado que a estas alturas existirá suficiente familiaridad con las ideas tanto de función como de espacios vectoriales.

Las transformaciones lineales tienen también una fuerte conexión con las matrices, dado que éstas son una forma abreviada de aquéllas. Incluso puede pensarse en las primeras como una correspondencia $X \longrightarrow AX$, donde A es una matriz.

Además de la gran cantidad de aplicaciones teóricas que presentan en distintas ramas de la física y la matemática, las transformaciones lineales se utilizan bastante en la ingeniería y en las ciencias sociales.

6.2 B R E V E B O S Q U E J O H I S T O R I C O.

Tal y como se advirtió en la introducción, si queremos hablar de la historia de las transformaciones lineales, debemos considerar la historia de las funciones.

Uno de los grandes matemáticos de todos los tiempos, Leonard Euler pensaba en una función como una fórmula o ecuación en la que aparecían variables y constantes.

Alrededor de 1860, Cayley, a quien ya mencionamos en la sección de determinantes como uno de los estudiosos de la teoría de matrices, fué quien aseguró que al fin de cuentas éstas no eran más que una forma abreviada de representar transformaciones lineales.

El concepto abstracto de función que se utiliza en la actualidad fue presentado por Peter Gustav Lejeune Dirichlet, y Giuseppe Peano fue el autor de la idea de transformación que en nuestros días usamos.

Más recientemente, un grupo de matemáticos formado por - Emmy Noether, Emil Artin y W. Krull, además de Hasse, dan la concepción abstracta que hoy conocemos.

6.3 P R O B L E M A S D I V E R S O S .

Siguiendo con la costumbre, se presentan algunos problemas de aplicación que se resuelven al emplear transformaciones li--- neales.

Problema 1.- El dueño de una fábrica produce cuatro productos distintos: vestidos, faldas, blusas y pantalones. Las materias primas que utiliza en la elaboración de estos pro--- ductos son: algodón, hilo y botones.

En la tabla que a continuación se muestra, está la in--- formación de cuántas unidades, de cada una de las materias -- primas utilizadas, se requieren para producir una unidad de - cada producto.

Una notación que pudiéramos utilizar es: vestidos, V; -- faldas, F; blusas, B; pantalones, P; algodón, A; hilo, H; botones, T.

		Utilizadas en la fabricación de un			
		V	F	B	P
Número de Unidades de Materias Primas	A	1	2/3	1/2	3/2
	H	3	2	2	2
	T	4	2	6	2

Pregunta: ¿Cuántas unidades de cada materia prima se necesitan si se produce un número determinado de cada uno de -- los cuatro productos?

Problema 2.- Ya es conocido por nosotros el llamado Teorema Fundamental del Cálculo. De cualquier forma vamos a recordarlo:

Llamemos $F(x) = \int_a^x f(t)\,dt$.

Teorema.- Si f es continua en [a,b] entonces la deriva-- da $F'(x)$ existe para toda $x \in [a,b]$ y $F'(x) = f(x)$.

Si tenemos una función P que satisface $P'(x) = f(x)$, (P es llamada una primitiva o antiderivada de f), entonces para toda k y x de [a,b] se tiene que

$$P(x) - P(k) = \int_a^x f(t)\,dt,$$

es decir, F(x) y cualquier otra antiderivada de f, P(x), di--

fieren en una constante:

$$P(x) - F(x) = P(k).$$

En efecto, ya que $F'(x) = f(x)$ y $P'(x) - F'(x) = f(x) - f(x) = 0$, debemos tener que $P(x) - F(x) =$ constante.

Si $x = k$, $P(k) = P(k) - F(k) =$ constante, de donde

$$P(x) - P(k) = F(x) = \int_a^x f(t)\,dt,$$

lo que nos proporciona un método para calcular $\int_a^b f(t)\,dt$:

$$\int_a^b f(t)\,dt = P(b) - P(a),$$

ésto es, bastará con encontrar una antiderivada P de f.

Así el problema de evaluar una integral se transfiere al problema de encontrar una antiderivada de f. Y toda fórmula de diferenciación nos dá una primitiva de alguna función f, - lo que resuelve el problema del cálculo de la integral de esa función f.

¿Puede abordarse este problema desde el punto de vista - de las transformaciones lineales?

Problema 3.- Cuando estudiamos Sistemas de Ecuaciones -- Lineales, vimos un teorema que establecía lo siguiente:

Teorema.- La solución general de un sistema de ecuaciones lineales es igual a la solución del homogéneo más una solución particular.

¿Es posible abordar este teorema desde el enfoque de las transformaciones lineales?

6.4 D E F I N I C I Ó N Y E J E M P L O S

Definición 6.4.1.- Una función T, que se define para todos -- los vectores V, que pertenecen a un espacio vectorial $\mathcal{V}$, y -- que nos dá como resultado vectores T(V), en un espacio vectorial $\mathcal{W}$, se llama transformación de $\mathcal{V}$ a $\mathcal{W}$.

Notación.- Como las transformaciones son funciones, se - usará la misma notación con la que usualmente se manejan es-- tas últimas, es decir una transformación T de $\mathcal{V}$ a $\mathcal{W}$ es una función $T: \mathcal{V} \to \mathcal{W}$.

Definición 6.4.2.- Si para $U, V \in \mathcal{V}$ y $\alpha \in K$ (un campo arbitrario), se satisface:

i) $T(U + V) = T(U) + T(V)$ y
ii) $T(\alpha V) = \alpha T(V)$, entonces la transformación se llama lineal

Observación.- Para transformaciones lineales, $\mathcal{V}$ y $\mathcal{W}$ deben ser espacios vectoriales sobre el mismo campo.

Vamos a ejemplificar.

Ejemplo 1.- Sean $\mathcal{V} = R^2$, $\mathcal{W} = R^2$. Diga si la transformación T, tal que T(U) = (-x,y), donde U = (x,y) es lineal.

Debemos verificar que se satisfagan las condiciones i), ii).

Sean U = (x,y), W = (x_1, y_1), U, V $\in \mathcal{V}$.

U + W = $(x + x_1, y + y_1)$.

T(U + W) = $(-(x + x_1), y + y_1) = (-x-x_1, y + y_1)$.

T(U) + T(W) = $(-x,y) + (-x_1, Y_1) = (-x-x_1, y + y_1)$,

ésto es T(U + W) = T(U) + T(W), lo cual permite concluir que i) se cumple.

Ahora verificamos ii)
$\alpha U = (\alpha x, \alpha y)$,
$T(\alpha U) = (-\alpha x, \alpha y) = \alpha(-x, y)$
$\alpha T(U) = \alpha(-x,y)$, ésto es, $T(\alpha U) = \alpha T(U)$, lo cual permite concluir que ii) se cumple.

Por lo que T(U) = (-x,y) es una transformación lineal.

Ejemplo 2.- Sean $\mathcal{V} = R^2$, $\mathcal{W} = R$, T: $R^2 \to R$ tal que T(x, y) = 2x - y.

Sean U = (x,y), W = (x_1, y_1),

U + W = $(x + x_1, y + y_1)$,

$$T(U + W) = T(x + x_1, y + y_1) = 2(x + x_1) - (y + y_1)$$
$$= 2x + 2x_1 - y - y_1.$$

Por otro lado,

$$T(U) + T(W) = (2x - y) + (2x_1 - y_1)$$
$$= 2x + 2x_1 - y - y_1, \text{ por lo que}$$
$$T(U + W) = T(U) + T(W).$$

Conclusión.- Se cumple i).

Verifiquemos ii)

$$T(\alpha U) = T(\alpha x, \alpha y) = 2(\alpha x) - \alpha y$$
$$= 2\alpha x - \alpha y.$$

$\alpha T(U) = \alpha(2x - y) = 2\alpha x - \alpha y$, esto es, $T(\alpha U) = \alpha T(U)$, por lo que se cumple ii), lo cual permite concluir que:

$T : R^2 \to R$ tal que $T(X, y) = 2x - y$, es una transformación - lineal.

En la proposición que a continuación enunciamos se resume la forma en la cual podemos verificar si una transforma--- ción es o no lineal. Veamos:

Proposición 6.4.3.- Sea $T: \mathcal{V} \to \mathcal{W}$, T es lineal si y sólo si - para toda $\alpha, \beta \in K$,

$$T(\alpha U + \beta V) = \alpha T(U) + \beta T(V), \text{ para toda } U, V \in \mathcal{V} .$$

Demostración.- Como es una proposición "si y sólo si", - hay que demostrar que se cumplen los dos sentidos, esto es:

" $\Longrightarrow$ ". Suponemos que T es lineal.

P.D. $T(\alpha U + \beta V) = \alpha T(U) + \beta T(V)$.

$$\begin{aligned} T(\alpha U + \beta V) &= T(\alpha U) + T(\beta V) \text{ (por ser T lineal)} \\ &= \alpha T(U) + \beta T(V) \text{ (por ser T lineal).} \end{aligned}$$

" $\Longleftarrow$ ". Suponemos que se tiene una transformación que:

$T(\alpha U + \beta V) = \alpha T(U) + \beta T(V)$, para toda $U, V \in \mathcal{V}, \alpha, \beta \in K$.

P.D. que T es lineal.

Tomamos $\alpha, \beta = 1$ (el neutro multiplicativo del campo).

$T(1\ U + 1\ V) = 1\ T(U) + 1\ T(V)$

$T(U + V) = T(U) + T(V)$.

Tomamos $\alpha \in K, \beta = 0$, (0, neutro aditivo del campo).

$T(\alpha U + 0\ V) = \alpha\ T(U) + 0\ T(V)$.

$T(\alpha U) = \alpha T(U)$.

Por lo tanto, T es lineal, con lo que se concluye la demostración.

Usualmente a $\mathcal{V}$ se le llama " espacio de salida" y a $\mathcal{W}$, - "espacio de llegada".

Prácticamente la definición de transformación lineal que acabamos de ver es lo único que necesitamos para resolver el problema 1.

Vamos a definir el vector de producción, $V_p = \begin{pmatrix} V \\ F \\ B \\ P \end{pmatrix}$, y a $V_m = \begin{pmatrix} A \\ H \\ T \end{pmatrix}$.

Sea $M = \begin{pmatrix} 1 & 2/3 & 1/2 & 3/2 \\ 3 & 2 & 2 & 2 \\ 4 & 2 & 6 & 2 \end{pmatrix}$.

Suponga que el vector de producción es $V_p = \begin{pmatrix} 10 \\ 30 \\ 20 \\ 50 \end{pmatrix}$.

¿Cuántas unidades de A, H, T, se necesitan para producir el número dado de cada uno de los cuatro productos?

De acuerdo con la tabla del principio,

A = V(1) + F(2/3) + B(1/2) + P(3/2),

H = V(3) + F(2) + B(2) + P(2),

T = V(4) + F(2) + B(6) + P(2),

ecuaciones que aplicando el caso particular del $V_p = \begin{pmatrix} 10 \\ 30 \\ 20 \\ 50 \end{pmatrix}$, nos daría:

A = 10(1) + 30(2/3) + 20(1/2) + 50(3/2)
= 10 + 20 + 10 + 75 = 115 unidades.

H = 10(3) + 30(2) + 20(2) + 50(2)
= 30 + 60 + 40 + 100 = 230 unidades.

T = 10(4) + 30(2) + 20(6) + 50(2)
= 40 + 60 + 120 + 100 = 320 unidades.

Para cualquier V_p, se cumple:

$M\ V_p = V_m$, ésto es,

$$\begin{pmatrix} 1 & 2/3 & 1/2 & 3/2 \\ 3 & 2 & 2 & 2 \\ 4 & 2 & 6 & 2 \end{pmatrix} \begin{pmatrix} V \\ F \\ B \\ P \end{pmatrix} = \begin{pmatrix} A \\ H \\ T \end{pmatrix}.$$

Tenemos entonces una función T que "transforma" al V_p en

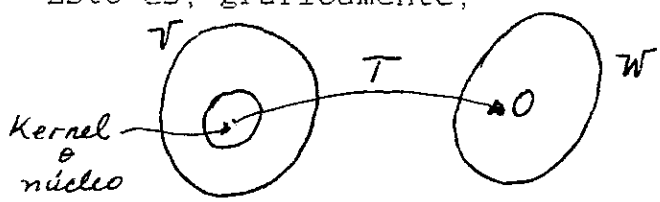
V_m, mediante una sencilla operación entre matrices, la multiplicación.

$T : V_p \rightarrow V_m$ tal que $T(V_p) = V_m = AV_p$, en donde A es la llamada matriz de transformación o transformación en sí.

6.5 EL NUCLEO O KERNEL DE UNA TRANSFORMACION.

Definición 6.5.1.- Sea $T : \mathcal{V} \rightarrow \mathcal{W}$ una transformación lineal. Definimos el núcleo o kernel de T como el conjunto de vectores $U \in \mathcal{V}$ tales que $T(U) = 0$.

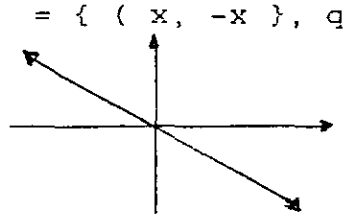
Esto es, gráficamente,



Veamos algunos ejemplos:

1) Sea $T : R^2 \rightarrow R$, $T(x, y) = x + y$.

Kernel $T = \{ (x, y) / T(x, y) = 0 \}$
$= \{ (x, y) / x + y = 0 \}$,
$= \{ (x, -x \}$, que geométricamente sería:



2) Sea $T: R^2 \rightarrow R^2$, $T(x, y) = (x + y, x - y)$.

Kernel $T = \{ (x, y) / T(x, y) = (0, 0) \}$
$= \{ (x, y) / (x + y, x - y) = (0,0) \}$
$= \{ (0,0) \}$.

$x + y = 0$ La única solución al sistema es (0,0).
$x - y = 0$

3) Sea $T: R^4 \rightarrow R$, tal que $T(x, y, z, w) = x + y + z + w$.

Kernel $T = \{ (x, y, z, w) / T(x, y, z, w) = 0 \}$
$= \{ (x, y, z, w) / x + y + z + w = 0 \}$.

4) Sea $T: R \rightarrow R^3$, tal que $T(x) = (x, 2x, -x)$.

Kernel $T = \{ x / (x, 2x, -x) = (0,0,0)$
$= \{ 0 \}$.

Hasta aquí ya podemos resolver los problemas 2 y 3 que se plantearon en la Sección 6.3.

Resolveremos el Problema 2:

Sean $\mathcal{V}$ = { f/f es derivable }, $\mathcal{W}$ = { h/h es continua }, K = R

Definamos T: $\mathcal{V} \to \mathcal{W}$ tal que T(f) = f'. ¿Es T una transformación lineal?

De acuerdo con la Proposición 6.4.3, debemos probar que T(mU + nV) = mT(U) + nT(V), para m,n $\in$ K, V,W $\in \mathcal{V}$.

Sean f, g, funciones derivables, es decir, f,g$\in \mathcal{V}$ y m,n $\in$ R.

T(mf + ng) = (mf + ng)'
= (mf)' + (ng)', (aplicando teoremas ya conocidos del Cálculo Dif.)
= mf' + ng'
= mT(f) + nT(g).

Concluimos entonces que T es lineal.

¿Qué sería el kernel de T?

Kernel T = { f/f T(f) = 0} = { f/f es constante }

Si tomamos h$\in \mathcal{W}$, ¿cuál sería la respectiva pre-imagen de h?

Pues sería un elemento f de $\mathcal{V}$ más un elemento del kernel de T. ¿Por qué? Porque estamos preguntando por la antiderivada de la función h, y de acuerdo con el Teorema Fundamental del Cálculo, las antiderivadas de una función sólo difieren en una constante.

A continuación trataremos el Problema 3:

Consideremos A una matriz de mxn, $\mathcal{V} = R^n$, $W = R^m$, K = R y $X = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ x_n \end{pmatrix}$.

Definamos T la transformación tal que T(X) = AX.

¿Es T lineal? Tomemos X, Y$\in R^n$, m,n $\in$ R.

T(mX + nY) = A(mX + nY)
= A(mX) + A(nY), (por las propiedades de suma y multiplicación de matrices que ya se vieron).
= m(AX) + n(AY)
= mT(X) + nT(Y)

De donde concluimos que T es lineal.

¿Cuál sería el kernel de T?

Kernel T = { X / T(X) = 0 } = { X / AX = 0 }. Con palabras, sería el conjunto de los vectores $X \in R^n$, tales que --- fueran solución de los sistemas homogéneos AX = 0 , (donde - 0 = (0, 0, ..., 0).

Si tomamos $W \in \mathcal{W}$, ¿cuál sería la respectiva pre-imagen de W? Pues W ha de ser tal que se satisfaga AX =W, entonces su pre-imagen será el vector X que cumpla la igualdad pedida, -- que no es otra cosa que la forma compacta de escribir un Sistema de Ecuaciones Lineales.

Si recordamos el teorema 2.6.4, que nos asegura que la - solución general de un SEL está dada por la solución del sistema homogéneo asociado más una solucion particular, éso significa que el vector pre-imagen de W quedará expresado como - un elemento de $\mathcal{V}$ más un elemento de su kernel.

Vamos a continuación a enunciar y demostrar algunas proposiciones que son necesarias para la demostración de uno de los resultados más importantes de transformaciones lineales.

Proposición 6.5.2.- Sea $T: \mathcal{V} \to \mathcal{W}$, una transformación li-- neal. T es uno a uno si y sólo si el kernel de T consta sólo del cero en $\mathcal{W}$.

Como ésta es una proposición que involucra una doble implicación, hay que demostrar dos cosas:

Hipótesis.- T es uno a uno.

Tesis.- El kernel de T sólo consta del vector cero.

Supongamos que el kernel de T consta al menos de los --- vectores V y U. Esto significa que T(U) = 0 = T(V), de donde T(U) = T(V) para U = V, lo que contradice la hipótesis de que T es uno a uno. Por tanto, el kernel sólo consta del vector cero.

Hipótesis.- El kernel de T sólo consta del vector cero.

Tesis.- T es uno a uno.

Supongamos que T no es uno a uno, éso significa que ---- existen un par de vectores U,V, $U \neq V$, tales que T(U) = T(V), entonces T(U) - T(V) = 0 = T(U - V), de donde U - V está en el kernel de T, lo que contradiría la hipótesis de que el --- kernel de T sólo consta del vector cero.

De donde T tiene que ser uno a uno, lo que completa la - demostración del teorema.

Proposición 6.5.3.- El kernel de T es un subespacio vec-

torial lineal de $\mathcal{V}$, ésto es, se cumplen las condiciones para ser un subespacio vectorial y además, si T es lineal T(0) = 0

Demostración:

i) Sean U,V en el kernel de T. P.D. U + V está en el kernel - de T.

$T(U + V) = T(U) + T(V) = 0 + 0 = 0.$

ii) Si U está en el kernel de T, entonces αU también estará.

$T(\alpha U) = \alpha T(U) = \alpha(0) = 0.$

iii) 0 está en el kernel de T.

$T(0) = T(0 \cdot U) = 0T(U) = 0.$

Proposición 6.5.4.- Sea T una transformación lineal uno a uno., Si $V_1, V_2, \ldots, V_n \in \mathcal{V}$ y son linealmente independien-- tes, entonces $T(V_1), T(V_2), \ldots, T(V_n) \in \mathcal{W}$ lo son también.

Demostración.- Sea $T: \mathcal{V} \to \mathcal{W}$. Hay que demostrar que si V_1, $V_2, \ldots, V_n$, son linealmente independientes, y T es uno a uno entonces $T(V_1), T(V_2), \ldots, T(V_n)$ también son linealmente independientes. Se hará por contradicción.

Supongamos que $T(V_1), T(V_2), \ldots, T(V_n)$, son linealmente dependientes, o sea que $\lambda_1 T(V_1) + \lambda_2 T(V_2) + \ldots + \lambda_n T(V_n) = 0$, con al menos un $\lambda_i \neq 0$. Eso significa que $\lambda_1 V_1 + \ldots + \lambda_n V_n$ está en el kernel de T.

Como por hipótesis T es uno a uno, entonces el kernel de T consta sólo del vector cero y ésto a su vez implica que -- $\lambda_1 V_1 + \ldots + \lambda_n V_n = 0$ ---› ‹---- Contradicción a lo que se había supuesto.

Por ser $V_1, V_2, \ldots, V_n$, linealmente independientes en-- tonces $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.

Definición 6.5.5.- Sea $T: \mathcal{V} \to \mathcal{W}$. Definimos el rango de T como el conjunto de elementos $W \in \mathcal{W}$ para los cuales existe V con T(V) = W.

Observación.- El rango de T también se llama Imagen de T

Proposición 6.5.6.- El rango de T es un subespacio vec-- torial de $\mathcal{W}$.

Demostración:

1) Sean $W_1, W_2 \in$ rango de T. P.D. $W_1 + W_2$ rango de T.

Como W_1, W_2 pertenecen al rango de T entonces existen -- $V_1, V_2 \in \mathcal{V}$ tales que $T(V_1) = W_1$, $T(V_2) = W_2$.

Tomando $V_1 + V_2$ se tendrá que:

$T(V_1 + V_2) = T(V_1) + T(V_2) = W_1 + W_2$.

2) Si W está en el rango de T, P.D. αW también está.

Como W pertenece al rango de T, entonces existe $V \in \mathcal{V}$ con $T(V) = W$. $T(\alpha V) = \alpha W$, entonces αW está en el rango de T.

3) ¿0 pertenece al rango de T? Sí, $T(0) = 0$.

Con todos estos preliminares llegamos a la proposición - que habíamos dicho es pilar en transformaciones lineales y -- cuyo enunciado se muestra renglones abajo.

Proposición 6.5.7.- Sea $\mathcal{V}$ un espacio vectorial y $T: \mathcal{V} \to \mathcal{W}$ una transformación lineal.

Si dimensión $\mathcal{V}$ = n, dimensión kernel T = m, dimensión --- imagen de T = s , entonces n = m + s, ésto es:

dimensión $\mathcal{V}$ = dimensión kernel de T + dimensión rango de T.

Observación.- Si kernel de T = { 0 }, no hay nada que -- demostrar. ¿Por qué? Sea $U \in \mathcal{V}$ arbitrario, dado por:

$U = \lambda_1 U_1 + \ldots + \lambda_n U_n$, $T(U) = \lambda_1 T(U_1) + \ldots + \lambda_n T(U_n)$. $T(U_1)$, $T(U_2), \ldots, T(U_n)$ generan al rango de T. Entonces $U_1, U_2, \ldots U_n$, son base de $\mathcal{V}$.

Por ser T uno a uno, $T(U_1), T(U_2), \ldots, T(U_n)$, son li--- nealmente independientes (por la Proposición 6.5.4) y por - lo tanto son base del rango de T.

dimensión $\mathcal{V}$ = dimensión kernel T + dimensión rango T.

$$n = 0 + s.$$

Supongamos que dimensión rango de T = s.

Existen $W_1, W_2, \ldots, W_s$ base del rango de T, si están en $\mathcal{W}$ entonces existen $V_1, \ldots, V_s \in \mathcal{V}$ tal que $T(V_1) = W_1, \ldots, T(V_s) = W_s$.

Como dimensión del kernel de T es m, entonces kernel de T tiene una base con m elementos, digamos $Z_1, Z_2, \ldots, Z_m$, - que están en el kernel de T.

Vamos a demostrar que $V_1, V_2, \ldots, V_s$ junto con $Z_1, Z_2, \ldots, Z_m$ son una base de $\mathcal{V}$.

Primero demostraremos que generan a $\mathcal{V}$.

Tomemos U en $\mathcal{V}$ arbitrario, T(U) en el rango de T, entonces $T(U) = \lambda_1 W_1 + \ldots + \lambda_S W_S$ (se puede escribir como combinación lineal de los elementos de la base de $\mathcal{W}$.

Consideremos al vector $\lambda_1 V_1 + \ldots + \lambda_S V_S$.

Afirmamos que $U - (\lambda_1 V_1 + \ldots + \lambda_S V_S) \in$ kernel de T. ¿Por qué?

$$T(U - (\lambda_1 V_1 + \ldots + \lambda_S V_S)) = T(U) - T(\lambda_1 V_1 + \ldots + \lambda_S V_S)$$

$$= (\lambda_1 W_1 + \ldots + \lambda_S W_S) - (\lambda_1 W_1 + \ldots + \lambda_S W_S)$$

$$= 0, \text{ de donde } U - (\lambda_1 V_1 + \ldots + \lambda_S V_S) \in \text{kernel T.}$$

Por ser $U - (\lambda_1 V_1 + \ldots + \lambda_S V_S)$ un elemento del kernel - y $Z_1, Z_2, \ldots, Z_m$ base del kernel, puedo escribir al primero como combinación lineal de los elementos de la base.

$U - (\lambda_1 V_1 + \ldots + \lambda_S V_S) = \alpha_1 Z_1 + \ldots + \alpha_m Z_m$, entonces se tiene que

$U = \lambda_1 V_1 + \ldots + \lambda_S V_S + \alpha_1 Z_1 + \ldots + \alpha_m Z_m$, por lo tanto - son generadores de $\mathcal{V}$.

Ahora hay que demostrar que $V_1, V_2, \ldots V_S, Z_1, \ldots, Z_m$ son linealmente independientes, ésto es:

$\lambda_1 V_1 + \ldots + \lambda_S V_S + \alpha_1 Z_1 + \ldots + \alpha_m Z_m = 0$, con $\lambda_1 = \lambda_2 = \ldots = \lambda_S = \alpha_1 = \ldots = \alpha_S = 0$.

Aplicamos la transformación a la expresión anterior:

$$T((\lambda_1 V_1 + \ldots + \lambda_S V_S) + (\alpha_1 Z_1 + \ldots + \alpha_m Z_m)) = 0.$$

$$\lambda_1 T(V_1) + \ldots + \lambda_S T(V_S) + \alpha_1 T(Z_1) + \ldots + \alpha_n T(Z_m) = 0.$$

$$\lambda_1 W_1 + \lambda_2 W_2 + \ldots + \lambda_S W_S + 0 .$$

Como $W_1, \ldots W_S$, son base, son vectores linealmente independiente, éso implica que $\lambda_1 = \ldots = \lambda_S = 0$.

De la misma manera se concluye que $\alpha_1, \ldots, \alpha_m = 0$, y ya se tiene el resultado,

$$n = m = s .$$

APENDICE

En esta sección completaremos la teoría referente a algunos capítulos desarrollados en este trabajo y aclaramos que el no incluir este material en un curso normal de la materia no le resta continuidad ni tampoco se pierde el sentido de lo que se está estudiando. Contiene los siguientes puntos :

a) Más problemas aplicados que involucran vectores.
b) Otro método interesante para la solución de SEL no homogéneos : Esquema Compacto.
c) Algo más de matrices y de solución de SEL.

En la primera parte analizamos algunos problemas adicionales relacionados con vectores y se incluyen en esta sección debido a que son problemas que requieren de conocimientos de otras áreas (Física, principalmente).

En la segunda parte presentamos un esquema de solución de SEL no homogéneos mediante el cual obtenemos la solucion de SEL a través de un proceso de acumulación que simplifica bastante los cálculos a realizar. De este modo ilustramos el hecho de que existen otros métodos de solucion de SEL que van más allá de los tradicionales y que son bastante útiles por su simplificación en los cálculos.

La última parte incluye lo siguiente :
i) La demostración de la propiedad asociativa del producto matricial.
ii) Otros tipos especiales de matrices.
iii) Esquema Compacto para la Inversión de Matrices y el Método de la Raíz Cuadrada para solución de SEL.

Esta última parte se incluye en esta sección porque es de interés básicamente para físicos y matemáticos, sobre todo ii) y iii). Además, en i) se introducen dobles sumatorias, que es una notación no muy utilizada por los estudiantes y puede presentar problemas al momento de operar con ellas.

Los métodos de inversión de matrices y de la raíz cuadrada para solución de SEL son bastante importantes e ilustramos una vez más el que existe una gran gama de métodos para la solución de dichos sistemas. Estos métodos a diferencia de los otros son aplicables en determinados casos y operan en base a operaciones entre matrices.

A manera de referencia, si alguien se interesa en profundizar más en este material, sugerimos consultar la obra "Computational Methods of Linear Algebra" de Faddeev-Faddeeva.

a) M A S P R O B L E M A S A P L I C A D O S Q U E I N -- V O L U C R A N V E C T O R E S .

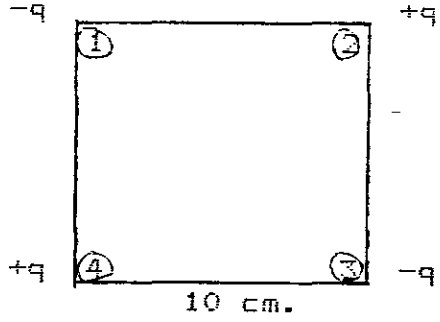
Carga eléctrica.- Ley de Coulomb.- Las cargas que consideró Coulomb fueron esferas, por lo que se supone que tienen simetría esférica. Esta ley relaciona fuerza entre cargas con éstas mediante la expresión

$$F = \frac{1}{4\pi \varepsilon_0} \frac{q_1 q_2}{r^2},$$

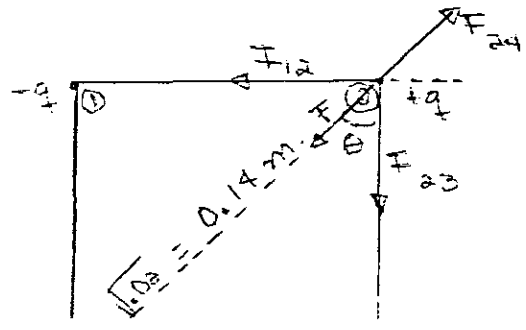
donde ε_0 = Permitividad = $8.89 \times 10^{}$ C^2 ./ $N.M.^2$, de ahí -- que $\frac{1}{4\pi \varepsilon_0} = 9 \times 10^9$ $N.M^2/C^2$, y r = distancia entre q_1 y q_2.

La ecuación de Coulomb se aplica por pares de cargas y - las fuerzas son colineales. La magnitud de las fuerzas (de - atracción o repulsión, dependiendo de la naturaleza de las -- cargas) es independiente de las mismas, pues es igual. La -- magnitud de cada fuerza se obtiene con la fórmula anterior, - pero la fuerza resultante se obtiene con la suma vectorial de todas las fuerzas.

Ejemplo : Se colocan dos partículas de carga +q en vértices diagonalmente opuestos de un cuadrado de lado a, y dos partículas de carga -q en los vértices restantes. Si $q = 5 \times 10^{-8}$ C. y a = 10 cm., determine :
a) La fuerza que actúa sobre la partícula colocada en el vértice superior derecho.
b) La fuerza que actúa sobre una de las partículas de carga - negativa.



Solución.- Mediante la fórmula encontramos el valor de $||F_{12}||$, $||F_{23}||$ y $||F_{24}||$ y por medio de la suma vectorial de ellas determinamos el valor de la fuerza F.



Así :

$$||F_{12}|| = \frac{1}{4\pi\varepsilon_o}\frac{q_1 q_2}{r_{12}^2} = (9 \times 10^9\ N.M.^2/C^2.)\frac{(5 \times 10^{-8}\ C.)^2}{(.10\ M.)^2}$$

$$= 2.25 \times 10^{-3}\ N.$$

$$||F_{23}|| = \frac{1}{4\pi\varepsilon_o}\frac{q_2 q_3}{r_{23}^2} = (9 \times 10^9)\frac{(5 \times 10^{-8})}{(0.1)^2} = 2.25 \times 10^{-3}\ N.$$

$$||F_{24}|| = \frac{1}{4\pi\varepsilon_o}\frac{q_2 q_4}{r_{24}^2} = (9 \times 10^9)\frac{(5 \times 10^{-8})}{(0.14)^2} = 1.148 \times 10^{-3}\ N$$

Para obtener las componentes de cada fuerza necesitamos el valor de θ, pero $\text{arcsen}\ \frac{0.1}{\sqrt{.02}} = 45$ grados. Con lo que :

$F_{12} = (-||F_{12}||, 0) = (-2.25 \times 10^{-3}, 0)$

$F_{23} = (0, -||F_{23}||) = (0, -2.25 \times 10^{-3})$

$F_{24} = (||F_{24}||\cos.45\ \text{grados}, ||F_{24}||\text{sen.}45\ \text{grados}) =$

$= (8.1176 \times 10^{-4}, 8.1176 \times 10^{-4})$

Entonces

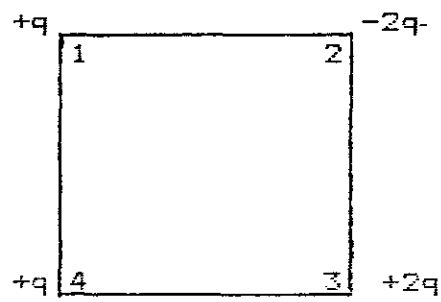
$F = F_{12} + F_{23} + F_{24} = (-1.4382 \times 10^{-3}, -1.4382 \times 10^{-3})$, cuya magnitud es $||F|| = 2.034 \times 10^{-3}$ N.

Análogamente se resuelve b).

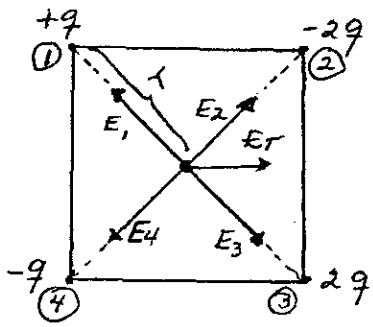
Campo eléctrico.- Es la fuerza eléctrica que actúa sobre la unidad de carga. Podemos determinar su magnitud mediante la expresión $E = \frac{F}{q_o}$ N./C., de donde $F = E\ q_o$.

Como $F = \frac{1}{4\pi\varepsilon_o}\frac{q q_o}{r^2}$, entonces $E = \frac{\frac{1}{4\pi\varepsilon_o}\frac{q q_o}{r^2}}{q_o} = \frac{1}{4\pi\varepsilon_o}\frac{q}{r^2}$.

Ejemplo : ¿Cuál es la magnitud y dirección del campo eléctrico en el centro del cuadrado de la figura ? Suponga que $q = 1 \times 10^{-8}$ C.



Solución.-



En este caso $r^2 = 2(.05/2)^2 = 12.5 \times 10^{-4}\ m.^2$

Al igual que sucede con las fuerzas $E_R = E_1 + E_2 + E_3 + E_4$. Con la fórmula dada podemos determinar sus normas como sigue:

$||E_1|| = ||(9 \times 10^9)(1 \times 10^{-8})/(12.5 \times 10^{-4})|| = 7.2 \times 10^4$ N.C

$||E_2|| = ||(9 \times 10^9)(-2 \times 10^{-8})/(12.5 \times 10^{-4})|| = 1.44 \times 10^5$ N.C

$||E_3|| = ||(9 \times 10^9)(2 \times 10^{-8})/(12.5 \times 10^{-4})|| = 1.44 \times 10^5$ N.C

$||E_4|| = ||(9 \times 10^9)(-1 \times 10^{-8})/(12.5 \times 10^{-4})|| = 7.2 \times 10^4$ N.C

Si fijamos el centro del cuadrado en el origen de nuestro sistema de referencia obsérvese que los ángulos de cada E_i con los ejes es de 45 grados. Con ésto podemos descomponer a los vectores y aplicar la definición de suma :

$E_1 = (-||E_1||\cos.45 \text{ grados}, ||E_1||\text{sen } 45 \text{ grados})$
$E_2 = (||E_2||\cos.45 \text{ grados}, ||E_2||\text{sen.45 grados})$
$E_3 = (||E_3||\cos.45 \text{ grados}, -||E_3||\text{sen.45 grados})$
$E_4 = (-||E_4||\cos.45 \text{ grados}, -||E_4||\text{sen.45 grados})$

Así :

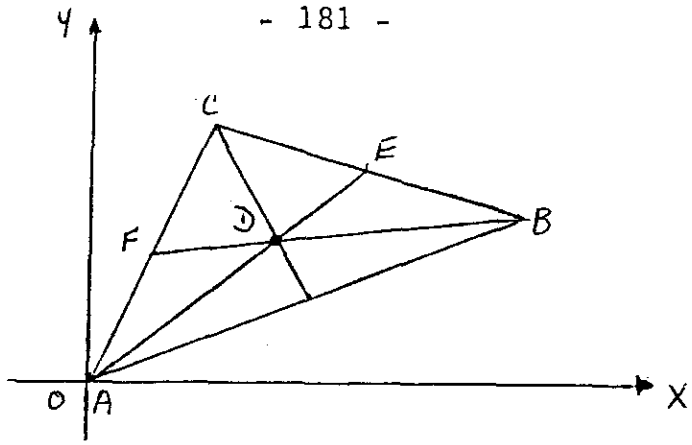
$E = ((-||E_1||+||E_2||+||E_3||-||E_4||)\cos.45 \text{ grados},$
$(||E_1||+||E_2||-||E_3||-||E_4||)\text{sen.45 grados}))$
$= (1.0182 \times 10^5, 0)$

y $||E|| = 1.0182 \times 10^5$ N.C.-

Geometría.- Verifique que las medianas de un triángulo se intersectan en un punto fijo que está localizado a 2/3 partes de la distancia de cada vértice a su lado opuesto.

Solución.- Queremos comprobar que las medianas de un triangulo con vertices A,B,C se intersectan en un punto fijo localizado a 2/3 partes de la distancia de cada vertice a su lado opuesto. Supondremos ésto válido y pasaremos a comprobarlo.

Situemos el triángulo con vértices en los puntos A,B,C de tal forma que A coincida con el origen del plano coordenado.



Sean $A = (0,0)$
$B = (x_1, y_1)$
$C = (x_2, y_2)$

Entonces $AB = (x_1, y_1)$ y $BE = 1/2\ BC$, pero como $BC = (x_2 - x_1, y_2 - y_1)$, entonces $BE = 1/2\ (x_2 - x_1, y_2 - y_1)$.

Claramente vemos que AB + BE = AE (primer mediana). Así :

$$AE = (x_1,y_1) + 1/2\ (x_2 - x_1, y_2 - y_1)$$
$$= (x_1 + (x_2 - x_1)/2, y_1 + (y_2 - y_1)/2)$$
$$= ((x_1 + x_2)/2, (y_1 + y_2)/2).$$

Por hipótesis :

$$AD = 2/3\ AE = 2/3\ ((x_1 + x_2)/2, (y_1 + y_2)/2) =$$
$$= ((x_1 + x_2)/3, (y_1 + y_2)/3).$$

Por otro lado vemos que :
(segunda mediana) AF - AB = BF , pero $AF = 1/2\ AC = 1/2\ (x_2, y_2)$ de donde

$$BF = 1/2\ (x_2, y_2) - (x_1, y_1) = ((x_2 - 2x_1)/2, (y_2 - 2y_1)/2).$$

Pero también

$$AD = AB + 2/3\ BF = (x_1, y_1) + 2/3\ ((x_2 - 2x_1)/2, (y_2 - 2y_1)/2)$$
$$= ((x_1 + x_2)/3, (y_1 + y_2)/3)$$

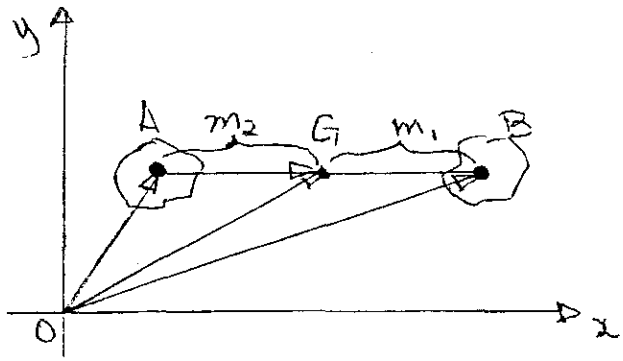
Por lo tanto, como no llegamos a una contradicción de las suposiciones planteadas, concluimos que son válidas ; es decir AD = 2/3 AE. Análogamente se trabajan las otras medianas.

Centros de masa.- El centro de masa de dos partículas con masas m_1 y m_2 localizadas en los puntos A y B respectivamente, es el punto G que divide al segmento AB en la proporción $m_2 : m_1$. Compruebe que (con relación a cierto origen fijo - O) :

$$OG = \frac{1}{m_1 + m_2}\ (m_1\, OA + m_2\, OB)$$

Solución.- Queremos comprobar la localización del centro de masa para estas dos partículas conociendo las masas de ambas y su colocación en el plano. Esto lo haremos en base a la suposición de que :

$$OG = \frac{1}{m_1 + m_2}\ (m_1\, OA + m_2\, OB)$$



Sean $A = (x_1, y_1)$, $B = (x_2, y_2)$. Como $AB = m_1 + m_2$, entonces $AG = \frac{m_2}{m_1 + m_2} AB$.

Ya que $AB = (x_2 - x_1, y_2 - y_1)$ entonces $AG = \frac{m_2}{m_1 + m_2} (x_2 - x_1, y_2 - y_1)$

Puesto que por hipótesis

$$OG = \frac{1}{m_1 + m_2} (m_1 OA + m_2 OB) = \frac{1}{m_1 + m_2} [m_1(x_1, y_1) + m_2(x_2, y_2)]$$

$$= \frac{1}{m_1 + m_2} (m_1 x_1 + m_2 x_2, m_1 y_1 + m_2 y_2),$$

y además, como se observa claramente en la figura, $AG = OG - OA$, entonces:

$$AG = \frac{1}{m_1 + m_2} (m_1 x_1 + m_2 x_2, m_1 y_1 + m_2 y_2) - (x_1, y_1)$$

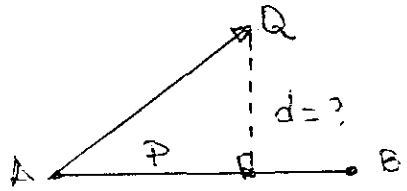
$$= \frac{1}{m_1 + m_2} (m_2 x_2 - m_2 x_1, m_2 y_2 - m_2 y_1) = \frac{m_2}{m_1 + m_2} (x_2 - x_1, y_2 - y_1),$$

que es lo que queríamos demostrar.

Geometría.- Hallar la distancia perpendicular del punto (5, -2,3) a la recta que pasa por los puntos A = (2,0,-3) y B = (8,3,3).

Solución.- Dados dos puntos A = (2,0,-3) y B = (8,3,3) por los cuales pasa una recta se nos pide determinar la distancia perpendicular del punto Q = (5,-2,3) a dicha recta.

En la siguiente figura se visualiza fácilmente el método de solución. Lo que se necesita es encontrar el vector V=AQ, su norma y la norma del vector proyección P. La distancia se encontrará mediante el Teorema de Pitágoras.



$V = AQ = OQ - OA = (5,-2,3) - (2,0,-3) = (3,-2,6)$.

$AB = OB - OA = (8,3,3) - (2,0,-3) = (6,3,6)$.

$$P = CB = \left(\frac{V \cdot AB}{\|AB\|^2} \right) AB$$

$$c = \frac{V \cdot (AB)}{||AB||^2} = \frac{(3,-2,6)\cdot(6,3,6)}{36 + 9 + 36} = \frac{48}{81}$$

$$P = \frac{48}{81}(6,3,6) = (288/81, 144/81, 288/81).$$

$$||P|| = \sqrt{(288/81)^2 + (144/81)^2 + (288/81)^2} = 5.333$$

De acuerdo con el Teorema de Pitágoras :

$$||V||^2 = ||P||^2 + d^2,$$

de donde

$$d = \sqrt{||V||^2 - ||P||^2} = \sqrt{49 - 28.4} = 4.534$$

Trabajo.- Como ya se mencionó en el problema 7, Capítulo 1, si la dirección en que se aplica la fuerza F y la dirección en la que desplaza la partícula no coinciden, el trabajo se calcula en base a la componente de la fuerza F en la dirección del movimiento y la distancia d recorrida durante dicho movimiento. Así, tenemos que

$$W = ||F|| \, ||d|| \cos \theta,$$

donde θ es el ángulo formado entre F y d. Ejemplo :

Una bestia jala un vagón con una fuerza de 215 N. que forma un ángulo de 30 grados con la horizontal y lo mueve a una velocidad de 10.5 Km./h. ¿Qué cantidad de trabajo hace la bestia en 10 minutos ?

Solución.- Los datos de este problema son :

$||F|| = 215$ N.

$||V|| = 10.5$ Km./h. $= 175$ m./min.

$\theta = 30$ grados.

Se pregunta acerca del trabajo W realizado por la bestia en un tiempo $t = 10$ minutos. Por definición $V = d/t$,entonces $d = Vt$. De aquí que $d = 9175)(10) = 1750$ m.

Con ésto, como $W = ||F|| \, ||d|| \cos \theta$, entonces

$W = (215)(1750)\cos.30$ grados $= 325\,842$ Joules.

b) E S Q U E M A S C O M P A C T O S P A R A L A S O ---
L U C I O N D E S I S T E M A S D E E C U A C I O N E S -
L I N E A L E S N O H O M O G E N E O S .

Como ya vimos anteriormente, el resolver un sistema de - ecuaciones lineales por el esquema de división simple es lo - mismo que determinar los coeficientes $a_{ij\cdot k}$ de las ecuaciones transformadas (incluyendo las constantes) y los coeficientes b_{ij} en las ecuaciones en el sistema triangular final. En realidad, para la obtención de una solución del sistema sólo necesitaríamos conocer los coeficientes b_{ij} , pues los $a_{ij\cdot k}$ son necesarios únicamente para determinarlos a ellos. Estos términos b_{ij} pueden obtenerse mediante un proceso de acumulación que simplifica en mucho los cálculos a hacer.

Escojamos elementos de la primera columna de cada matriz auxiliar $a_{ij\cdot j-1}$, $i \geq j$ y denotémoslos por c_{ij} , $i \geq j$.

Analizando los cálculos vemos que:

$$a_{ij\cdot k} = a_{ij\cdot k-1} - a_{ik\cdot k-1} b_{kj} = a_{ij\cdot k-1} - c_{ik} b_{kj}$$

$$= a_{ij\,k-2} - c_{i\,k-1} b_{k-1\,j} - c_{ik} b_{kj} \quad \ldots\ldots\ldots\ldots (1)$$

$$= a_{ij} - c_{i1} b_{1j} - c_{i2} b_{2j} - \ldots - c_{ik} b_{kj}$$

$$= a_{ij} - \sum_{l=1}^{k} c_{il} b_{lj} .$$

Con lo que cada $a_{ij\cdot k}$ se calcula mediante la suma (de -- ahí lo de acumulación) de los productos en términos c_{ij} y b_{ij} que necesitan determinarse.

En particular para los c_{ij} , $i \geq j$ y b_{ij} , $i < j$ son válidas las formulas de recurrencia:

$$c_{ij} = a_{ij\cdot j-1} = a_{ij} - \sum_{l=1}^{j-1} c_{il} b_{lj} \quad (i \geq j)$$

$$b_{ij} = \frac{a_{ij\cdot i-1}}{a_{ii\cdot i-1}} = a_{ij} - \sum_{l=1}^{i-1} c_{il} b_{lj} \quad (i<j) .$$

Obviamente las constantes transformadas también se calculan mediante estas fórmulas. El esquema para llevar el --- curso hacia adelante mediante este procedimiento se llama --- COMPACTO. El curso de regreso permanece inalterable.

Por conveniencia, para más simplificación en los cálculos, podemos arreglar los elementos del esquema compacto, y - calcular los elementos c_{ij} y b_{ij} en sucesión "por esquinas" - comenzando con los elementos c_{ij} como sigue:

c_{11}	b_{12}	b_{13}	b_{14}	Primer paso
c_{21}	c_{22}	b_{23}	b_{24}	Segundo paso
c_{31}	c_{32}	c_{33}	b_{34}	Tercer paso

Por ejemplo, en la siguiente tabla:

$$c_{42} = a_{42} - c_{41} b_{12} = 0 - (-6)(-1) = -6,$$

$$b_{23} = \frac{a_{23} - c_{21} b_{13}}{c_{22}} = \frac{0 - (6)(1)}{11} = -6/11 \quad .$$

El cálculo para el esquema compacto requiere la fijación de los elementos principales de antemano, por lo que es imposible dar una forma compacta al esquema de elementos principales.

TABLA 1.- ESQUEMA COMPACTO DEL METODO DE DIVISION SIMPLE.

a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	1	-1	1	0	1
a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	6	5	0	20	31
a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	0	5	4	30	39
a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	6	0	4	10	8
1 c_{11}	b_{12}	b_{13}	b_{14}	b_{15}	1	-1	1	0	1
c_{21}	1 c_{22}	b_{23}	b_{24}	b_{25}	6	11	-6/11	20/11	25/11
c_{31}	c_{32}	1 c_{33}	b_{34}	b_{35}	0	5	74/11	230/74	304/74
c_{41}	c_{42}	c_{43}	1 c_{44}	b_{45}	-6	-6	74/11	0	0 (*)
			0					0	0
		1					1	230/74	304/74
	1					1		260/74	334/74
1					1			30/74	104/74

(*) Recuérdese que la última ecuación se anula, de ahí el por que sucede ésto.

Obsérvese que este ejemplo, ya resuelto anteriormente, efectivamente llega a la solucion esperada.

Expliquemos un poco la construcción de la tabla. Las primeras tres columnas representan los coeficientes del sistema, la cuarta las constantes, y la quinta corresponde a la de control; en esta última columna ejecutamos las mismas operaciones que en todas las demas. Cuando hacemos ésto, cada número de la columna transformada deberá coincidir (si no se cometieron errores de cálculo), con la suma de los elementos de los correspondientes renglones de la matriz B, aumentada por incluir la columna de las constantes. De hecho la matriz B es la matriz de los coeficientes para el sistema obtenido

del sistema original despues de terminar el curso hacia adelante en el esquema de division simple.

Los elementos c_{ii} corresponden a los elementos principales (unos, 1's) de que se habló anteriormente. Así, la matriz B a que nos referimos en este problema es:

$$B = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -6/11 & 20/11 \\ 0 & 0 & 1 & 230/74 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad y \quad \begin{pmatrix} 1 \\ 25/11 \\ 304/74 \\ 0 \end{pmatrix}$$ corresponde a la columna de control.

Otro método para solución de Sistemas de Ecuaciones Lineales, el Método de la Raíz Cuadrada, se basa fundamentalmente en la multiplicación de matrices y es aplicable esencialmente a matrices simétricas. Por incluirse propiedades de matrices (o mejor dicho, operaciones entre ellas), este método se estudiará mas adelante.

c) M A T R I C E S

i) Propiedad Asociativa del Producto Matricial.

Dentro de lo que corresponde a propiedades para el producto matricial demostraremos aquí la que corresponde a la asociatividad: Para tres matrices $A_{m\times n}$, $B_{n\times r}$, $C_{r\times k}$, se tiene que $(AB)C = A(BC)$.

Primeramente veamos los tamaños. Sean $A_{m\times n}B_{n\times r} = D_{m\times r}$ y $(AB)C = D_{m\times r}C_{r\times k} = E_{m\times k}$. Por otro lado, si $B_{n\times r}C_{r\times k} = F_{n\times k}$ y $A(BC) = A_{m\times n}F_{n\times k} = G_{m\times k}$, vemos que tanto E como G tienen el mismo tamaño.

Además debemos comprobar si el ij-ésimo elemento de la matriz E es igual al ij-ésimo elemento de la matriz G para toda i y toda j. De tal modo que, en conjunto, se tenga la igualdad de matrices.

Por definición el lj-ésimo elemento de F es:

$$[F]_{lj} = [BC]_{lj} = B_l C^{(j)} = b_{l1}c_{1j} + b_{l2}c_{2j} + \dots + b_{lr}c_{rj},$$

pero esta suma la podemos abreviar utilizando notación $\sum$ y escribir:

$$[F]_{lj} = \sum_{p=1}^{r} b_{lp}c_{pj} = f_{lj}.$$

También por definición, la ij-ésima componente de G es:

$$[G]_{ij} = [AF]_{ij} = A_i F^{(j)} = \sum_{l=1}^{n} a_{il} f_{lj} = \sum_{l=1}^{n} a_{il}\left(\sum_{p=1}^{r} b_{lp}c_{pj}\right)$$
$$= \sum_{l=1}^{n}\sum_{p=1}^{r} a_{il}b_{lp}c_{pj} = [A(BC)]_{ij}.$$

Por otro lado, el ip-ésimo elemento de D es:

$$[D]_{ip} = [AB]_{ip} = A_i B^{(p)} = \sum_{l=1}^{n} a_{il}b_{lp} = d_{ip}.$$

Y además, el ij-ésimo elemento de E es:

$$[E]_{ij} = [DC]_{ij} = D_i C^{(j)}$$
$$= \sum_{p=1}^{r} d_{ip}c_{pj} = \sum_{p=1}^{r}\left(\sum_{l=1}^{n} a_{il}b_{lp}\right)c_{pj}$$
$$= \sum_{p=1}^{r}\sum_{l=1}^{n} a_{il}b_{lp}c_{pj} = [(AB)C]_{ij}$$, pero ésta es la ij ésima componente de A(BC).

ii) Más tipos especiales de matrices.

En lo que se refiere a la clasificación de las matrices, ésta se puede extender mucho más allá de lo que se hizo en el capítulo correspondiente de forma tal que nos muestra muchísimas propiedades adicionales interesantes. Algunos otros tipos de matrices son los siguientes:

1.- Matriz Conjugada Compleja.- Si reemplazamos los elementos de una matriz con números complejos conjugados obtenemos la llamada Matriz Conjugada Compleja $\bar{A}$. Si los elementos de A son reales, entonces $\bar{A} = A$.

2.- Matriz Conjugada.- La matriz $A^* = \bar{A}^t$, la compleja conjugada de la matriz transpuesta, se llama Matriz Conjugada para A o la Conjugada de A. Obsérvese que $(A^*)^* = A$. Si la matriz A es real, entonces su conjugada coincide con la matriz transpuesta.

3.- Matrices Antisimétricas.- Una matriz cuadrada A se llama Antisimétrica si $A = -A^t$.

Para que A sea una matriz antisimétrica es necesario que sea cuadrada y además que cumpla que:

$$a_{ij} = -a_{ji},\ i \neq j, \qquad a_{ij} = 0,\ i = j,$$

ya que por definición $a_{ij} = -a_{ii}$, y esto es únicamente con $a_{ii} = 0$.

A continuación veremos algunas propiedades de las matrices simétricas y antisimétricas (no demostraremos todas).

Proposición 1.- Si A es una matriz simétrica de elementos reales o complejos, se tiene que:
a) αA es simétrica, con $\alpha \in R$ o $\alpha \in C$
b) $AA^t = A^tA$
c) A^2 es simétrica.

En la demostración de estas propiedades no debemos olvidar que por hipótesis A es simétrica, es decir $A = A^t$.

Proposición 2.- Si A es una matriz antisimétrica con elementos reales o complejos, tenemos que:

a) A es antisimétrica, con $\alpha \in R$ o $\alpha \in C$
b) $AA^t = A^tA$
c) A^2 es simétrica.

Proposición 3.- Cualquier matriz cuadrada A puede descomponerse como la suma de una matriz simétrica con otra antisimétrica y esta descomposición es única.

Esta propiedad sí la demostraremos. Definamos una matriz simétrica S y otra antisimétrica U tales que:

$$S = \frac{A + A^t}{2} \quad , \quad U = \frac{A - A^t}{2} ,$$

donde A es cualquier matriz cuadrada. Así, de acuerdo con -- ésto,

$$s_{ij} = 1/2\ (a_{ij} + a_{ji}) \text{ y } u_{ij} = 1/2\ (a_{ij} - a_{ji}).$$

Con lo que:

$$\begin{aligned} s_{ij} + u_{ij} &= 1/2(a_{ij} + a_{ji}) + 1/2(a_{ij} - a_{ji}) \\ &= 1/2(a_{ij} + a_{ji} + a_{ij} - a_{ji}) \\ &= a_{ij}. \end{aligned}$$

Para demostrar la unicidad de tal descomposición basta - con ver que si existe otra descomposición tal que $A = \bar{S} + \bar{U}$, entonces $\bar{S} = S$ y $\bar{U} = U$.

Puede comprobarse también que si A es simétrica, entonces A^n, con $n \in \mathbb{Z}^+$, también es simétrica.

Además, son dignos de mención otros hechos importantes:

a) El producto de dos matrices simétricas no es, en general, una matriz simétrica, ya que si $A^t = A$ y $B = B^t$, $(AB)^t = B^t A^t = BA$, pero AB no necesariamente es igual a BA.

b) El producto de una matriz por su transpuesta es una matriz simétrica. Así:

$(AA^t)^t = (A^t)^t\ A^t = AA^t$ y $(A^tA)^t = A^t\ (A^t)^t = A^tA$, aunque AA^t no necesariamente es igual a A^tA.

c) Una matriz renglón multiplicada por una matriz columna -- (por supuesto, con el mismo número de elementos), da como resultado una matriz con un único elemento, que siempre es simétrica.

4.- Matrices Particionadas.- Con frecuencia es de bastante -- utilidad el reducir las operaciones entre matrices de orden - superior a operaciones entre matrices de orden menor. Esto - puede hacerse partiendo a las matrices dadas en "celdas", es decir, considerando a cada matriz como formada por varias matrices de orden menor (o celdas). Esto puede hacerse en varias formas, pero sólo consideraremos particiones de matrices cuadradas para las cuales las celdas diagonales son cuadradas.

Las operaciones básicas para matrices particionadas de - esta forma están muy conectadas con las operaciones en las -- celdas mismas. Esto es, si

$$A = \begin{bmatrix} A_{11} & A_{12} \dots A_{1K} \\ A_{21} & A_{22} \dots A_{2K} \\ \dots\dots\dots\dots \\ A_{K1} & A_{K2} \dots A_{KK} \end{bmatrix} \quad y \quad B = \begin{bmatrix} B_{11} & B_{12} \dots B_{1K} \\ B_{21} & B_{22} \dots B_{2K} \\ \dots\dots\dots\dots \\ B_{K1} & B_{K2} \dots B_{KK} \end{bmatrix}$$

donde A_{ii} y B_{ii} son matrices cuadradas conformables, entonces

$$A + B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \dots A_{1K} + B_{1K} \\ A_{21} + B_{21} & A_{22} + B_{22} \dots A_{2K} + B_{2K} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ A_{K1} + B_{K1} & A_{K2} + B_{K2} \dots A_{KK} + B_{KK} \end{bmatrix} \quad y$$

$$AB = \begin{bmatrix} C_{11} & C_{12} \dots C_{1K} \\ C_{21} & C_{22} \dots C_{2K} \\ \dots\dots\dots\dots \\ C_{K1} & C_{K2} \dots C_{KK} \end{bmatrix} \text{ donde } C_{ij} = A_{i1} B_{1j} + \dots + A_{iK} B_{Kj}, \quad i, j = 1, 2, \dots, k.$$

Además, si $c_{\alpha\beta}$ es algún elemento de la celda C_{ij} entonces:

$$c_{\alpha\beta} = (a_{\alpha 1} b_{1\beta} + \dots + a_{\alpha s_1} b_{s_1\beta}) + \dots + (a_{\alpha s_{K-1}+1} \; b_{s_{K-1}\beta} + \dots + a_{\alpha s_K} b_{s_K\beta}).$$

Aquí s_1, $s_2 - s_1$, $\dots$, $s_K - s_{K-1}$ denota los órdenes de las matrices A_{11}, A_{22}, $\dots$, A_{KK}. Los elementos entre paréntesis son elementos de las matrices $A_{i1}B_{1K}$, $\dots$, $A_{iK}B_{Kj}$ y ocupan en estas matrices la misma posición que tiene $c_{\alpha\beta}$ en la matriz C_{ij}.

Con ésto, $C_{ij} = A_{i1} B_{1j} + \dots + A_{iK} B_{Kj}$.

Observación.- Las definiciones así dadas nos muestran que las operaciones entre matrices particionadas en celdas se realizan exactamente en la misma forma que con matrices que tienen números en lugar de celdas.

5.- Matrices Limitadas.- Son un caso especial y muy importante de las matrices particionadas. Sea A una matriz cuadrada de orden n-1 de la forma

$$A = \begin{bmatrix} a_{11} & a_{12} \dots a_{1n-1} \\ a_{21} & a_{22} \dots a_{2n-1} \\ \dots\dots\dots\dots\dots \\ a_{n-1\,1} & a_{n-1\,2} \dots a_{n-1\,n-1} \end{bmatrix}$$

Podemos formar a partir de ella una matriz de n-ésimo orden agregando un renglón $V_{n-1} = (a_{n1}, a_{n2}, \dots a_{nn-1})$, una columna $U_{n-1} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{n-1\,n} \end{pmatrix}$ y un número a_{nn}. Así:

$$A_n = \begin{pmatrix} & & a_{1n} \\ & A_{n-1} & a_{2n} \\ & & \cdots \\ & & a_{n-1\,n} \\ a_{n1} \quad a_{n2} & \cdots \; a_{nn-1} & a_{nn} \end{pmatrix} = \begin{pmatrix} A_{n-1} & U_{n-1} \\ V_{n-1} & a_{nn} \end{pmatrix}$$

De aquí diremos que A_n se obtuvo limitando a la matriz A_{n-1}. Por supuesto que A_n es divisible en celdas, por lo tanto las operaciones entre ellas siguen las mismas reglas que para matrices particionadas; de acuerdo con ésto, si:

$$A = \begin{pmatrix} M & U \\ V & a \end{pmatrix} \quad y \quad B = \begin{pmatrix} N & Y \\ X & b \end{pmatrix}$$

son dos matrices limitadas de orden n, entonces

$$\alpha A = \begin{pmatrix} \alpha M & \alpha U \\ \alpha V & \alpha a \end{pmatrix}, \quad A + B = \begin{pmatrix} M + N & U + Y \\ V + X & a + b \end{pmatrix} y$$

$$AB = \begin{pmatrix} MN + UX & MY + Ub \\ VN + aX & VY + ab \end{pmatrix},$$

donde MN y Ux son matrices de orden n-1; My y Ub son columnas que contienen al (n-1)ésimo elemento; VN y aX son los correspondientes renglones y VY + ab es un número.

6.- Matrices cuasi-diagonales.- Son también un caso especial de las matrices particionadas. Una matriz se dice que es cuasi-diagonal si es una matriz cuadrada que consta de celdas cuadradas (no necesariamente del mismo orden) a lo largo de su diagonal principal y con los elementos restantes iguales a cero. Por ejemplo, la matriz

$$\left[\begin{array}{cc:ccc:cc} a_{11} & a_{12} & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & b_{11} & b_{12} & b_{13} & 0 & 0 \\ 0 & 0 & b_{21} & b_{22} & b_{23} & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & c_{11} & c_{12} \\ 0 & 0 & 0 & 0 & 0 & c_{21} & c_{22} \end{array}\right]$$

es una matriz cuasi-diagonal de séptimo orden, cuyas celdas son las matrices

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}, \quad \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

y seis matrices nulas.

Si la estructura de dos matrices cuasi-diagonales es la misma (es decir, tienen elementos del mismo tipo y están par-

ticionadas en la misma forma), el producto de ellas será otra vez una matriz cuasi-diagonal de la misma estructura, cuyas - celdas diagonales son iguales a los productos de las corres-- pondientes celdas de las matrices que se multiplican.

7.- Matrices Ortogonales.- Una matriz real se llama ortogo-- nal si la suma de los cuadrados de los elementos de cada co-- lumna es igual a 1, y si la suma de los productos de los ele- mentos correspondientes de dos columnas diferentes es igual a cero. La ortogonalidad de una matriz puede caracterizarse -- por una simple ecuación matricial : $A^t A = I$.

Los elementos diagonales de la matriz $A^t A$ son la suma de los cuadrados de los elementos de las columnas de la ma--- triz A, y los elementos no diagonales son iguales a las sumas de los productos de los elementos de las columnas correspon-- dientes. La matrices ortogonales tienen las siguientes pro-- piedades :

a) La matriz identidad es ortogonal.
b) Si A es ortogonal, entonces A^t también lo es.
c) El producto de dos matrices ortogonales es una matriz or-- togonal.

8.- Matrices Elementales de Rotación.- Pertenecen a la clase de matrices ortogonales y son de la forma

$$T_{ij} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & \cdots & -s & & \\ & & \vdots & & \vdots & & \\ & & s & \cdots & c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

donde $c^2 + s^2 = 1$. Esta última relación indica que existe un ángulo φ tal que $c = \cos\varphi$, $s = \operatorname{sen}\varphi$.

Nota : Las matrices elementales de rotación difieren de la matriz identidad sólo en cuatro elementos, situados en la intersección de dos renglones y dos columnas i y j, con $i<j$. Estos cuatro elementos forman la matriz

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$$

que coincide con la matriz de transformación cartesiana con - coordenadas en el plano por rotación del eje a través del án- gulo φ .

Estas matrices se utilizan en la transformación de una -

cierta matriz dada A mediante una secuencia de mutiplicaciones por ellas.

9.- Matrices Hermitianas.- Una matriz con elementos complejos se llama hermitiana si $a_{ij} = \overline{a_{ji}}$, es decir, si $A = A^*$. De aquí que los elementos diagonales de una matriz hermitiana son reales. Un caso especial de las matrices hermitianas son las matrices simétricas reales, y de ellas conservan muchas de sus propiedades ; una de ellas es que el producto de dos matrices hermitianas será hermitiana si y sólo si las matrices conmutan. Además, para cualquier matriz con elementos complejos, la matriz A^*A será hermitiana.

10.- Matrices Unitarias.- Una matriz con elementos complejos se llama unitaria si la suma de los cuadrados de los módulos de los elementos de cada columna es igual a 1, y también la suma de los productos de los elementos de una columna por números conjugados a los correspondientes elementos de otra columna es cero. Las matrices unitarias pueden caracterizarse por la ecuación matricial $A^*A = I$.

Un caso especial de las matrices unitarias lo son claramente las matrices ortogonales. Ambas clases de matrices pertenecen a la clase mas general de matrices complejas normales, que tienen como característica que cada una conmuta con su matriz conjugada.

Dentro de este tipo de matrices están también las matrices unitarias elementales, que son de la forma

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & ce^{i\theta_1} & \cdots & -se^{i\theta_2} \\ & & \vdots & & \vdots \\ & & se^{i\theta_3} & \cdots & ce^{i\theta_4} \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} ,$$

con $c, s > 0$, $c^2 + s^2 = 1$ y $\theta_1 - \theta_2 = \theta_3 - \theta_4$.

11.- Matrices Tridiagonales.- Una matriz tridiagonal es una matriz de la forma

$$\begin{bmatrix} a & b & 0 & \cdots & 0 & 0 \\ c & a & b & \cdots & 0 & 0 \\ 0 & c & a & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a & b \\ 0 & 0 & 0 & \cdots & c & a \end{bmatrix}$$

Una matriz real tridiagonal se llama Jacobiana si $b_i c_i > 0$ para $i = 1,2,\ldots,(n-1)$. Toda matriz simétrica tridiagonal con elementos no diagonales distintos de cero será automáticamente Jacobiana.

12.- Matrices Casi Triangulares.- Una matriz se dice casi triangular superior si es de la forma

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n-1} & a_{2n} \\ 0 & a_{32} & a_{33} & \ldots & a_{3n-1} & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{nn-1} & a_{nn} \end{bmatrix}$$

Nota : Haciendo uso del producto de matrices cuadradas podemos considerar polinomios de matrices. Si

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n$$

es un polinomio en x, dada una matriz A de orden n podemos definir el polinomio en A de grado n, que denotaremos p(A), como

$$p(A) = a_0 I + a_1 A + a_2 A^2 + \ldots + a_n A^n .$$

iii) Esquema Compacto para la Inversión de Matrices.

Cierto hecho importante referente a la expresión de una matriz como el producto de dos matrices triangulares nos permite construir un esquema compacto para el cálculo de la inversa de una matriz. Este esquema requiere $2n^2$ entradas, donde n^2 de ellas dan los elementos de la matriz inversa. Pero veamos primero cómo justificar esta descomposición de una determinada matriz como producto de otras dos.

Como ya vimos en la sección correspondiente, las matrices triangulares son un tipo especial de matriz con ciertas características igualmente especiales : la suma de matrices triangulares es a su vez una matriz triangular, etc. ; una de estas características asegura que cualquier matriz cuadrada A de orden n puede representarse como el producto de una matriz triangular superior y una inferior bajo ciertas condiciones. esto lo podemos explicar así :

Sea

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

Podemos particionarla como sigue :

$$A = \begin{bmatrix} & & a_{1n} \\ & A_{n-1} & \vdots \\ & & a_{n-1\,n} \\ a_{n1} & a_{n2} \cdots a_{nn-1} & a_{nn} \end{bmatrix} = \begin{bmatrix} A_{n-1} & U \\ V & a_{nn} \end{bmatrix} .$$

Necesitamos encontrar una descompisición de A tal que $A = CB$, con C y B matrices triangulares inferior y superior respectivamente ; si las particionamos obtenemos :

$$C = \begin{bmatrix} C_{n-1} & 0 \\ x & c_{nn} \end{bmatrix} , \quad B = \begin{bmatrix} B_{n-1} & y \\ 0 & b_{nn} \end{bmatrix} .$$

De aquí :

$$CB = \begin{bmatrix} C_{n-1} & 0 \\ x & c_{nn} \end{bmatrix} \begin{bmatrix} B_{n-1} & y \\ 0 & b_{nn} \end{bmatrix} = \begin{bmatrix} C_{n-1} B_{n-1} & yC_{n-1} \\ x B_{n-1} & xy + c_{nn} b_{nn} \end{bmatrix} = A.$$

Pero ésto será sólo si :

$$\begin{aligned} C_{n-1}\ B_{n-1} &= A_{n-1} \\ y\ C_{n-1} &= U \\ x\ B_{n-1} &= V \\ xy + c_{nn}\ b_{nn} &= a_{nn}. \end{aligned}$$

Esto se puede afirmar si requerimos que $a_{11} \neq 0$, que $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$, ..., $|A| \neq 0$. Como $|A_{n-1}| \neq 0$, entonces $|C_{n-1}| \neq 0$ y $|B_{n-1}| \neq 0$; además, si C y B son invertibles, entonces C_{n-1}^{-1} , B_{n-1}^{-1} existen y además $y = UC_{n-1}^{-1}$, $n = VB_{n-1}^{-1}$; y si a c_{nn} o b_{nn} se le determina un valor, de la última ecuación se calcula el valor del otro.

Con todo esto resulta que una matriz efectivamente puede expresarse como el producto de dos matrices triangulares.

Ahora pasaremos a construir el esquema compacto de inversión de matrices. Sea $A = CB$, con C y B como se requirieron anteriormente y cuyos elementos están definidos por:

$$c_{ij} = a_{ij} - \sum_{l=1}^{j-1} c_{il}\, b_{lj} \quad (i \geq j);$$

$$b_{ij} = \frac{a_{ij} - \sum_{l=1}^{i-1} c_{il}\, b_{lj}}{c_{ii}} \quad (i < j\ ,\ b_{ii} = 1).$$

(Ver esquema compacto para la solución de SEL no homogéneo, Capítulo 2).

Sea también $D = A^{-1} = B^{-1} C^{-1}$. Queremos ver si los elementos d_{ij} pueden determinarse sin la inversión de B y C.

De $D = B^{-1} C^{-1}$ se tiene que $DC = B^{-1}$, que también es una matriz triangular con unos a lo largo de la diagonal principal, con lo que conocemos $n(n+1)/2$ de sus elementos (de los cuales $n(n-1)/2$ son cero y los n restantes son unos).

Algo semejante sucede con $BD = C^{-1}$. Nótese que si conjuntamos las $n(n+1)/2$ ecuaciones de $DC = B^{-1}$ y las $n(n-1)/2$ ecuaciones de $BD = C^{-1}$ se obtiene un sistema recurrente que permite determinar los n^2 elementos de la matriz inversa. Tomemos n = 3; C será una matriz de la forma:

$$C = \begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \quad y \quad D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix}.$$

De aquí:

$$\begin{array}{lllll} & & i = 1 & 2 & 3 \\ c_{11} d_{i1} + c_{21} d_{i2} + c_{31} d_{i3} & = & 1 & 0 & 0 \\ c_{22} d_{i2} + c_{32} d_{i3} & = & & 1 & 0 \\ c_{33} d_{i3} & = & & & 1 \end{array}$$

$$\begin{array}{llll} & j = 2 & 3 \\ d_{1j} + b_{12} d_{2j} + b_{13} d_{3j} = & 0 & 0 \\ d_{2j} + b_{23} d_{3j} = & & 0 \end{array}.$$

La matriz B es de la forma:

$$B = \begin{bmatrix} 1 & b_{12} & b \\ 0 & 1 & b_{23} \\ 0 & 0 & 1 \end{bmatrix}.$$

De las ecuaciones del primer grupo para i= 3 podemos determinar sucesivamente d_{33}, d_{32} y d_{31} y obtener así los elementos del tercer renglón de D. Después, con las ecuaciones del segundo grupo y tomando j = 3 podemos determinar d_{23} y d_{13} respectivamente y obtener así los elementos de la tercer columna de D. Siguiendo estos mismos pasos podemos determinar todos los elementos de D. Vea el siguiente ejemplo.

TABLA A.2 .- ESQUEMA COMPACTO DE INVERSION DE UNA MATRIZ

10	20	50			
30	30	0			
60	50	50			
10 (1)	2	5	-1/20	-1/20	1/20
30	-30 (1)	5	1/20	5/60	-1/20
60	-70	100 (1)	1/100	-7/300	1/100

Sustituyendo en las ecuaciones de recurrencia :

$i = 3$: $\quad 100\, d_{33} = 1$ entonces $d_{33} = 1/100$

$-30\, d_{32} - 70(1/100) = 0$ entonces $d_{32} = (70/100)(-1/3) = -7/300$ y así sucesivamente.

Con ésto tenemos otro método más para invertir matrices y, por lo tanto, un método más para aplicar en la solución de Sistemas de Ecuaciones Lineales aplicando inversión.

Esta solución para Sistemas de Ecuaciones Lineales se simplifica bastante si la matriz asociada a él resulta ser simétrica ya que, como vimos en capítulos anteriores, una matriz simétrica puede expresarse como el producto de una matriz por su transpuesta; y esta simplificación se acentúa cuando dicha representación es en base a matrices triangulares. Esto se incluye en el llamado Método de la Raíz Cuadrada para solución de SEL.

iv) Método de la Raíz Cuadrada.

Sea $A = S^t S$, donde

$$S = \begin{pmatrix} s_{11} & s_{12} \cdots & s_{1n} \\ 0 & s_{22} \cdots & s_{2n} \\ \cdots & \cdots & \cdots \\ 0 & 0 \cdots & s_{nn} \end{pmatrix} \quad \text{y por tanto} \quad S^t = \begin{pmatrix} s_{11} & 0 \cdots & 0 \\ s_{12} & s_{22} \cdots & 0 \\ \cdots & \cdots & \cdots \\ s_{1n} & s_{2n} \cdots & s_{nn} \end{pmatrix}$$

Entonces

$$A = \begin{pmatrix} s_{11} & 0 \cdots & 0 \\ s_{12} & s_{22} \cdots & 0 \\ \cdots & \cdots & \cdots \\ s_{1n} & s_{2n} \cdots & s_{nn} \end{pmatrix} \begin{pmatrix} s_{11} & s_{12} \cdots & s_{1n} \\ 0 & s_{22} \cdots & s_{2n} \\ \cdots & \cdots & \cdots \\ 0 & 0 \cdots & s_{nn} \end{pmatrix}.$$

Así:

$$a_{ij} = (s_{1i}, s_{2i} \cdots, s_{ii}, 0, \cdots,) \begin{pmatrix} s_{1j} \\ s_{2j} \\ \cdots \\ s_{ij} \\ s_{jj} \\ \cdots \\ 0 \end{pmatrix}$$

(1) $a_{ij} = s_{1i} s_{1j} + s_{2i} s_{j2} + \cdots + s_{ii} s_{ij} + 0 + \cdots$ $\quad (i<j)$ y

(2) $a_{ii} = s_{1i}^2 + s_{2i}^2 + \cdots + s_{ii}^2$ $\quad (i=j)$.

De aquí podemos obtener expresiones con las cuales podamos determinar los elementos s_{ij} de la matriz S, como sigue:

$s_{11} = \sqrt{a_{11}}$, de (2); $s_{1j} = a_{1j}/s_{11}$, de (1), con $i = 1$;

$$s_{ii} = \sqrt{a_{ii} - \sum_{l=1}^{i-1} s_{li}^2} \quad (i>1), \text{ de (2)};$$

(3) $$s_{ij} = \frac{a_{ij} - \sum_{l=1}^{i-1} s_{li}\, s_{lj}}{s_{ii}} \quad (i<j)1, \text{ de (1) y por definición}$$

$s_{ij} = 0$, con $i>j$.

Por otra parte, el resolver el sistema $AX = B$ equivale a resolver dos sistemas triangulares : $S^tK = B$ y $SX = K$ (pues como $A = S^tS$ entonces $AX = S^tSX = B$; tomando $SX = K$ resulta $S^tK = B$). Los elementos del vector K los podemos determinar de manera semejante a los s_{ij}:

$$S^tK = B, \text{ entonces } \begin{bmatrix} s_{11} & 0 & \dots & 0 \\ s_{12} & s_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ s_{1n} & s_{2n} & \dots & s_{nn} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}.$$

De aquí:

$b_1 = k_1 s_{11}$ entonces $k_1 = b_1 / s_{11}$

$$b_i = (s_{1i}, s_{2i}, \dots, s_{ii}, 0, \dots 0) \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_i \\ \dots \\ k_n \end{bmatrix}$$

$= s_{1i} k_1 + s_{2i} k_2 + \dots + s_{ii} k_i + 0$ $(i>1)$, entonces

(4) $$k_i = \frac{b_i - \sum_{l=1}^{i-1} s_{li} k_l}{s_{ii}} \qquad (i>1).$$

Análogamente, la solución final la encontramos de la siguiente forma:

$$SX = K \text{ entonces } \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ 0 & s_{22} & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & s_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix}.$$

De aquí:

$$k_n = s_{nn} x_n \text{ entonces } x_n = k_n / s_{nn}.$$

$$k_i = (0, 0, \dots, s_{ii}, s_{i(i+1)}, \dots s_{in}) \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_i \\ \dots \\ x_n \end{pmatrix}$$

$$= 0 + s_{ii} x_i + s_{i(i+1)} x_{i+1} + \dots + s_{in} x_n.$$

De donde,

$$(5)\quad x_i = \frac{k_i - \sum_{l=i+1}^{n} s_{il}\, x_l}{s_{ii}} \quad (i<n).$$

Aquí podemos usar también la columna de control cuya ecuación, al igual que en el esquema compacto, es $\bar{k}_i = \sum_{k=1}^{n} s_{ik} + k_i$.

En este método basta con escribir $n(n+1)/2$ elementos de la matriz S y $2n$ componentes de los vectores K y X. Enseguida daremos un ejemplo:

TABLA A.3 METODO DE LA RAIZ CUADRADA

a_{11}	a_{12}	a_{13}	b_1	$\bar{b}_1$	1	1	3	10	15
	a_{22}	a_{23}	b_2	$\bar{b}_2$		2	4	15	22
		a_{33}	b_3	$\bar{b}_3$			14	35	56
s_{11}	s_{12}	s_{13}	k_1	$\bar{k}_1$	1	1	3	10	15
	s_{22}	s_{23}	k_2	$\bar{k}_2$		1	1	5	7
		s_{33}	k_3	$\bar{k}_3$			2	0	2
x_1	x_2	x_3			5	5	0		
$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$			6	6	1		

El cálculo de los elementos s_{ij} , k_i y $\bar{k}_i$ se aplica sucesivamente por renglones. Cualquier elemento diagonal se obtiene mediante la raíz cuadrada de la diferencia entre el elemento a correspondiente y la suma de los cuadrados de todos los elementos calculados s situados en la misma columna. Los elementos no diagonales s_{ij} se obtienen mediante el cociente que se obtiene al dividir la diferencia entre el elemento a_{ij} correspondiente y la suma de los productos de los elementos s situados en la columna i y j por el elemento diagonal del renglón. El curso de regreso está definido por (5)

En la actualidad se utiliza frecuentemente este método para la solución de sistemas simétricos y se recomienda como uno de los más efectivos.

Observación.- En el ejemplo referente a cajas negras visto en el Capítulo 3 una situación interesante es la siguiente: suponga que A es una matriz cuadrada y que se tiene una sucesión de matrices de salida diferentes $B_1, B_2, \ldots, B_k$. Si queremos determinar qué matrices de entrada producirían estas salidas deberíamos resolver sucesivamente cada sistema $AX = B_j$, $j = 1,2,\ldots,k$, donde A es la misma matriz de coeficientes para todos los sistemas. Aunque también puede aplicarse el método de solución para varios SEL.

C O N C L U S I O N E S Y R E C O M E N D A C I O N E S

Con el afán de situarnos, repetiremos las ideas básicas que motivaron este trabajo de tesis y que ya se establecieron en el Capítulo 0.

Habíamos mencionado que la idea original fue diseñar una propuesta de modificación curricular para el curso de Algebra Lineal I, en la cual se incluirían tanto aspectos teóricos propios de la materia, como el aspecto computacional. Esto es, en cada capítulo se analizaría el aspecto numérico del tema y se anexaría un diagrama de flujo con la correspondiente codificación y programa corrido.

Desafortunadamente se abandonó este último enfoque debido a dos causas fundamentales: la primera de ellas fué la cantidad de trabajo tan enorme que significaba el implementar las notas de esa manera, y la segunda fué que la materia de Programación de Computadoras se cursa al mismo tiempo que la de Algebra Lineal I, lo que no daría a los alumnos el dominio del material necesario para entenderlo.

La teoría se trató de abordar desarrollando las ideas básicas y fundamentales del Algebra Lineal I, ilustradas con una gran cantidad de ejemplos variados e interpretaciones geométricas.

Estas notas ya han tenido su implementación práctica en el aula, si bien ésta ha sido mínima. En el semestre 87-1 fueron los capítulos correspondientes a Vectores en R^n, Sistemas de Ecuaciones Lineales y Matrices, los que ya fueron expuestos ante el alumnado, en la manera tal y como se presentan actualmente, con modificaciones más de redacción que de contenido.

¿Qué comentarios podemos hacer al respecto? Creemos que el principal fue el interés que mostraron los alumnos ante las aplicaciones. Como ya se apuntaba en el Capítulo 0, pensamos que una de las mayores deficiencias de los cursos de Algebra Lineal había sido la ausencia de aplicaciones en áreas concernientes a las carreras que se manejan dentro del tronco común de Ciencias e Ingeniería.

Hubo también detalles que resultaron atractivos, entre ellos pudiéramos mencionar la justificación que se da al final del Capítulo I acerca de la definición del producto interior entre vectores en R^n, o la manera en que se motiva la multiplicación de matrices, definiéndolas al fin como una especie de generalización del producto interior.

Recomendaríamos al maestro interesado en utilizar esta propuesta, que pusiera especial énfasis en motivar a los alumnos mediante las interpretaciones geométricas y, aunque

suene repetitivo, mediante las aplicaciones.

Sugeriríamos la aplicación de cuatro exámenes parciales distribuídos así:

I Parcial.- Vectores en R^n, (20% de la calificación total)
II Parcial.- Sistemas de Ecuaciones Lineales y Matrices, (40% de la calificación).
III Parcial.- Espacios Vectoriales, (20% de la calificación).
IV Parcial.- Determinantes y Transformaciones Lineales, (20% de la calificación).

Pudiera pensarse en la completación de la calificación - mediante exposiciones, que son importantes.

¿Qué es lo que se pretende básicamente de un alumno que haya aprobado el curso?

a) Que tenga una panorámica de lo que es el Algebra Lineal y su utilización.
b) Que domine las técnicas de resolución de Sistemas de E--- cuaciones Lineales, pero no sólo la cuestión operativa, sino también lo teórico relacionado.
c) Que sin mayores problemas sea capaz de continuar estudios conectados.

Como todo trabajo que es presentado en una primera fase, éste puede ser completado de muchas maneras. La serie de --- problemas que se enuncia al principio de cada capítulo puede ser incrementada o incluso variada, dependiendo del tipo de - alumnos que se tenga, pudiera agregársele también un desglose por objetivos, otras sugerencias sobre evaluación, etc. Todo ésto ya se retomará con la colaboración de las personas interesadas.

Algunos de los proyectos que se tienen como continua--- ción a la propuesta son:

- Su exposición mediante un seminario ante los maestros del - Departamento de Matemáticas, con la intención de convencerlos de que lo sigan como texto de la materia.
- Abordar el enfoque computacional que se mencionaba al ini-- cio de esta sección.
- Elaboración de un trabajo similar para el curso de Algebra Lineal II.

Para todo ello, obviamente es indispensable la coopera-- ción del resto de los compañeros maestros e incluso de los -- mismos alumnos. Nuevamente hacemos un llamado a toda clase - de observaciones y sugerencias.

B I B L I O G R A F I A

1.- Applied Linear Algebra
Ben Noble
Ed. Prentice Hall.

2.- Algebra Lineal
Mina S. de Carakushansky, Guillerme de la Penha
Ed. Mc. Graw Hill.

3.- Fundamentos de Algebra Lineal y sus Aplicaciones
Francis G. Florey
Ed. Prentice Hall.

4.- Introducción al Algebra Lineal
Howard Anton
Ed. Limusa.

5.- Algebra Lineal
Serge Lang
Fondo Educativo Interamericano, S.A.

6.- Algebra Lineal
Stanley I. Grossman
Grupo Editorial Iberoamérica.

7.- Algebra Lineal
Seymour Lipschutz
Serie Schaum
Ed. Mc. Graw Hill.

8.- General Equilibrium: A Leontief Economic Model
Philip M. Tuchinsky
UMAP, unit 209.

9.- Applications of Matriz Methods: Food Service and Dietory Requirements
Sister Mary K. Killer
UMAP, unit 109.

10.- Linear Algebra
G. Hadley
Fondo Educativo Interamericano, S.A.

11.- Computational Methods of Linear Algebra
Fadeev Fadeeva
W.H. Freeman and Company.

12.- A History of Vector Analysis
Michael J. Crowe
University of Notre Dame Press
Notre Dame London.

13.- Historia de las Matemáticas, Tomo II
Jean Paul Collette
Ed. Siglo XXI.

14.- Algebra Lineal
Jorge Antonio Ludlow Wiechers
Ed. Limusa.

Reg 104
E. Morales Pons Cuevas