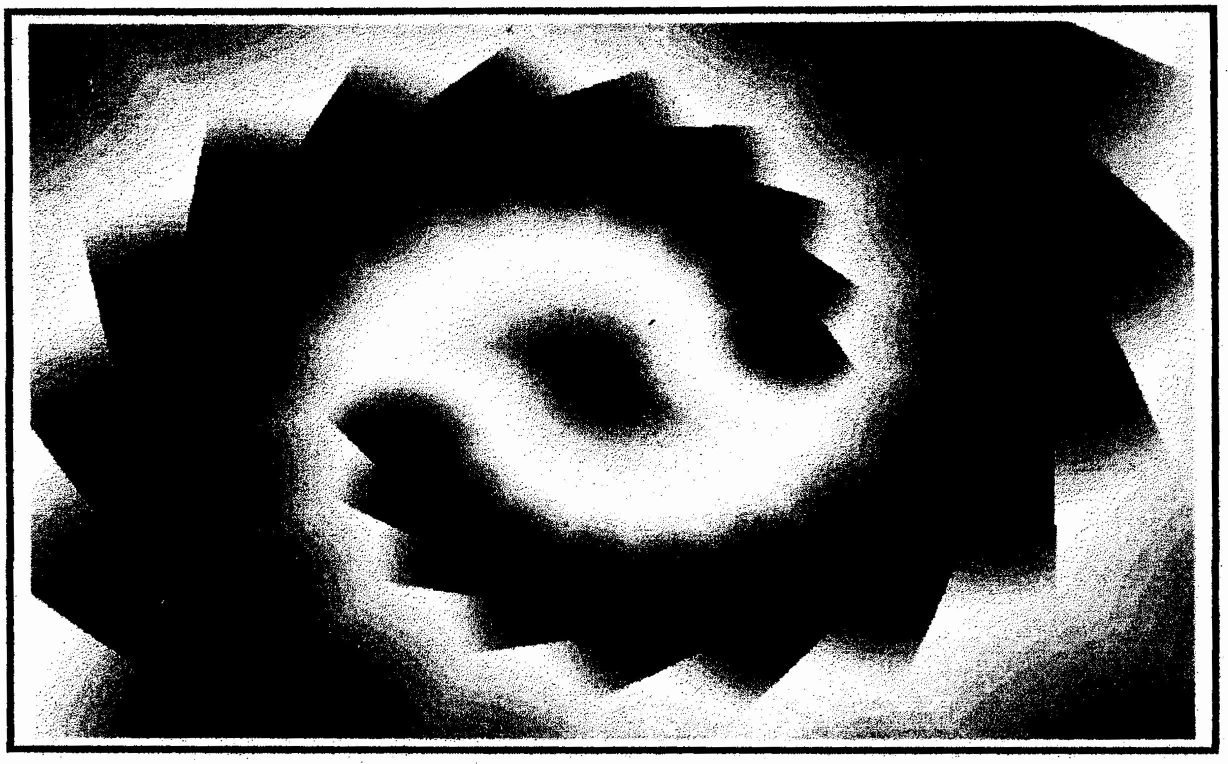
Apuntes

ÁLGEBRA LINEAL

IDALIA FLORES DE LA MOTA



Departamento de Sistemas

División de Estudios de Posgrado **Facultad de Ingeniería, U.N.A.M.**

Universidad Nacional Autónoma de México
Facultad de Ingeniería
División de Estudios de Posgrado
México D.F. C.P. 04510

1a. Edición 1991 con el título de Matemáticas aplicadas I
2a. Edición 2001

Tiraje: 200 ejemplares
Supervisión editorial. Lic. Ma. Guadalupe Castro Díaz

Como parte de las actividades del Departamento de Ingeniería de Sistemas de la División de Estudios de Posgrado, Facultad de Ingeniería UNAM, nos hemos propuesto el desarrollo de una serie de apuntes que sirvan de apoyo a los distintos cursos que se imparten y, desde luego, como material de referencia o de lectura para quien así lo desee.

En dichos apuntes se busca mantener un alto nivel, de manera que contribuyan a una sólida formación teórica del alumno, pero al mismo tiempo se intenta mantenerlo cerca de las aplicaciones de tales conocimientos, como es en síntesis la aspiración general del Posgrado de Ingeniería.

La presentación de los mismos es en siete capítulos cada uno consta de una introducción, el desarrollo del tema en cuestión, un resumen y unas notas históricas.

Agradecemos a la Jefatura de la División de Estudios de Posgrado el apoyo proporcionado para la publicación de los mismos.

PRÓLOGO

Los presentes apuntes son una reedición corregida de los Apuntes de Matemáticas Aplicadas, publicados por la División de Estudios de Posgrado a inicios de los años noventa. Como los apuntes correspondían al curso de Matemáticas Aplicadas I de ahí que tuvieran ese nombre pero en realidad corresponden al Álgebra Lineal.

El Álgebra Lineal es una estructura matemática de suma importancia en nuestro medio; ya que no sólo trata de resolver problemas de Investigación de Operaciones, también es una herramienta importante en electrónica, las estructuras y la ingeniería petrolera entre otras. Muchas veces los textos que tratan sobre esta disciplina son muy teóricos y se limitan a resolver unos cuantos ejemplos abundando en la demostración de teoremas, lo cual no es nuestro propósito. Más bien, se pretende resolver una gran cantidad de ejemplos sobre todo de ingeniería y efectuar las demostraciones de los teoremas más relevantes. Es bajo ésta perspectiva, que se presentan estos apuntes donde lo que se busca es tener un apoyo didáctico para la clase y como un complemento de la bibliografía sugerida para el curso.

Los presentes apuntes van dirigidos a estudiantes de posgrado que ya teniendo conocimiento de la materia requieren de un cierto grado de conceptualización y una buena dosis de ejemplos y aplicaciones, se presentan también al final de cada capitulo unas notas históricas que buscan dar al lector una idea mas completa de las matemáticas, viéndolas desde una perspectiva histórica, inmersas en la realidad de la época. Parafaseando a Isaac Asimov: "La ciencia gana realidad cuando es visualizada no como una abstracción, sino como la suma concreta de los científicos, pasados y presentes, vivos y muertos. No hay estamentos en ciencia, ni una observación, ni un pensamiento que exista por sí mismo. Cada uno de ellos es producto de un duro esfuerzo de algún hombre, y a menos que conozcáis al hombre y el mundo con el cual trabajaba, los supuestos que él aceptaba como verdades, los conceptos que él consideraba insostenibles, no podréis comprender sus afirmaciones, u observaciones, o pensamiento".

El orden en que se abordan los temas es el mismo que el del curso, este orden lo considero adecuado ya que el problema central del Álgebra Lineal es resolver sistemas de ecuaciones lineales y éste es el hilo conductor del curso.

Quiero agradecer el apoyo en revisión y correcciones a: Ivonne Peña Galeana, Griselda Águila Ramírez, Jair Gabriel Morales Camarena, y Leonardo Niño Ojeda. También quiero destacar el apoyo e interés de la Lic. Guadalupe Castro para la publicación de los mismos.

Finalmente quisiera mencionar que parte de este trabajo se apoyó en las notas de cursos anteriores impartidos por el Dr. Sergio Fuentes Maya.

SOBRE LA AUTORA

Idalia Flores De La Mota

Estudió el Doctorado en Investigación de Operaciones en la Facultad de Ingeniería de la UNAM. Obtuvo mención honorífica por su tesis de maestría así como la medalla Gabino Barreda por el mejor promedio de su generación. Ha asistido a varios congresos nacionales e internacionales, y ha sido miembro del Sistema Nacional de Investigadores. Ha publicado en memorias de congresos así como apuntes para las materias de Matemáticas Aplicadas, Teoría de Redes y Programación Entera. Ha impartido cursos en universidades del extranjero, en la Facultad de Economía de la Universidad de la Habana, Cuba así como en la Universidad Nacional Autónoma de Nicaragua. Es profesora de tiempo completo de la Facultad de Ingeniería desde 1990 y actualmente es la Jefa de la Sección de Investigación de Operaciones de la División de Estudios de Posgrado de la Facultad de Ingeniería.

ÍNDICE

CAPÍTULO I

MODELOS MATEMÁTICOS

CAPÍTULO 2

MATRICES

CAPÍTULO 3

SOLUCIÓN DE SISTEMAS LINEALES

CAPÍTULO 4

DETERMINANTES

CAPÍTULO 5

ESPACIOS VECTORIALES

CAPÍTULO 6

LOS CUATRO ESPACIOS FUNDAMENTALES

CAPÍTULO 7

VALORES Y VECTORES PROPIOS

CAPÍTULO 1
MODELOS MATEMÁTICOS

"La formulación de un problema es a menudo más importante que su solución"
A. Einstein

1.1 INTRODUCCIÓN

Un modelo es una representación de un problema o situación de la realidad. Esta representación la hacemos mediante diversos objetos o símbolos a través de un proceso de abstracción, que consiste en tomar de la realidad los elementos más importantes que intervienen en el problema y desechar todos aquellos que consideramos que no juegan un papel determinante en el mismo, estableciendo con precisión cuáles son las distintas relaciones que guardan entre sí dichos elementos. Una vez establecidas estas relaciones, podemos manipular los elementos del modelo en la búsqueda de una posible solución; o demostrar que tal solución no existe.

Se considera que las funciones de un modelo son la predicción y la comparación para proporcionar una forma lógica de predecir los resultados que siguen las acciones alternativas, e indicar una preferencia entre ellas. Aunque este uso de los modelos es importante, no es, de ninguna manera, su único propósito; la construcción de modelos proporciona una manera sistemática, explícita y eficiente para que un grupo de expertos y aquellos que toman las decisiones centren su juicio e intuición.

Uno de los principales elementos para resolver un problema es la construcción y el uso de un modelo. Dicho modelo puede tomar muchas formas, pero una de las más útiles, y ciertamente la que más se usa, es la matemática; lamentablemente, no siempre es posible crear un modelo matemático en un sentido riguroso y estricto. Al estudiar la mayoría de los sistemas militares e industriales, podemos definir objetivos, especificar restricciones y discernir si el diseño sigue las leyes de ingeniería y/o economía. Las relaciones esenciales se pueden descubrir y representar matemáticamente, de una manera u otra.

Representar algún objeto, sistema o idea con un modelo es tan general que es difícil clasificar todas las funciones que cumplen los modelos, algunas de ellas son:

a. Una ayuda para el pensamiento.
b. Una ayuda para la comunicación.
c. Para entrenamiento e instrucción.
d. Una herramienta de predicción.
e. Una ayuda para la experimentación.

En ingeniería, los modelos sirven como ayuda para diseñar nuevos sistemas o mejorar los existentes, mientras que en ciencias sociales y en economía, explican los sistemas existentes.

1.2 CARACTERÍSTICAS DE LOS MODELOS

Empezaremos nuestro análisis con un problema hipotético:

El viejo y el lobo

"A orillas del río Balsas, un viejo llevaba un lobo, una oveja y una paca de alfalfa. El viejo quiere cruzar el río con todas sus pertenencias, pero lo único que tiene para cruzarlo es una barca donde sólo cabe él y una de sus pertenencias de tal manera que tiene que llevarlas a la otra orilla una por una. Pero si deja al lobo con la oveja, éste se la comería y si deja a la oveja con la alfalfa, ésta se la comería". ¿Cómo puede el viejo llegar a la otra orilla del río con todas sus pertenencias?

Antes de resolver el problema, debe estar claro que podemos considerarlo como un sistema. Empecemos por diferenciar las relaciones que guardan entre sí los elementos del sistema:

1. En la barca sólo cabe el viejo con una de sus pertenencias (lobo, oveja, alfalfa).
2. El lobo no se puede quedar solo con la oveja.
3. La oveja no se puede quedar sola con la alfalfa.

Una vez enumeradas las principales relaciones de interdependencia del sistema, podemos intentar diseñar una posible solución.

A continuación se enuncian los pasos a seguir para una solución satisfactoria del problema:

1. El viejo cruza el río con la oveja.
2. El viejo regresa solo.
3. El viejo cruza el río con el lobo.
4. El viejo regresa con la oveja.
5. El viejo cruza el río con la alfalfa.
6. El viejo regresa solo.
7. El viejo cruza con la oveja.

Ahora pasemos a verificar si esta solución cumple las relaciones de interdependencia del sistema; esto es, hay que verificar que en ningún momento se pueda comer el lobo a la oveja ni la oveja a la alfalfa y que al final estén todos al otro lado del río. Una vez realizado lo anterior podemos, ahora sí, estar seguros de que la solución propuesta es adecuada.

Hasta ahora, sin decirlo, hemos seguido un método para resolver el problema; sin embargo, es importante que conozcamos otros más y sepamos diferenciarlos.

El problema anterior lo pudimos resolver de manera sencilla en forma mental, para lo cual lo único que tuvimos que hacer fue imaginar el sistema con sus relaciones de interdependencia y, a través de estas representaciones mentales, encontramos la solución; en otras palabras, elaboramos un modelo "mental" del sistema. A este método de solución le llamaremos RESOLUCIÓN MENTAL.

Sin embargo, no todos los problemas se pueden resolver tan fácilmente, existen problemas mucho más complicados para cuya solución necesitamos de otros métodos.

Regresemos al problema inicial y, con la finalidad de dejar más claro el razonamiento que hemos seguido, resolvamos las siguientes preguntas:

1. ¿Cuántas veces tuvo que cruzar el viejo el río?
2. Si la oveja no se comiera la alfalfa, ¿cuántas veces cruzaría el río?
3. ¿Hay otras soluciones al problema?
4. ¿Existe una solución más corta?
5. Si ponemos alfalfa donde diga lobo y lobo donde diga alfalfa en nuestra solución, ¿obtenemos otra solución?
6. Si intercambiamos alfalfa y oveja, ¿Es solución?
7. ¿Únicamente podemos resolver mentalmente nuestro problema?
8. ¿De qué otra manera, diferente a la mental, podemos resolver el problema?

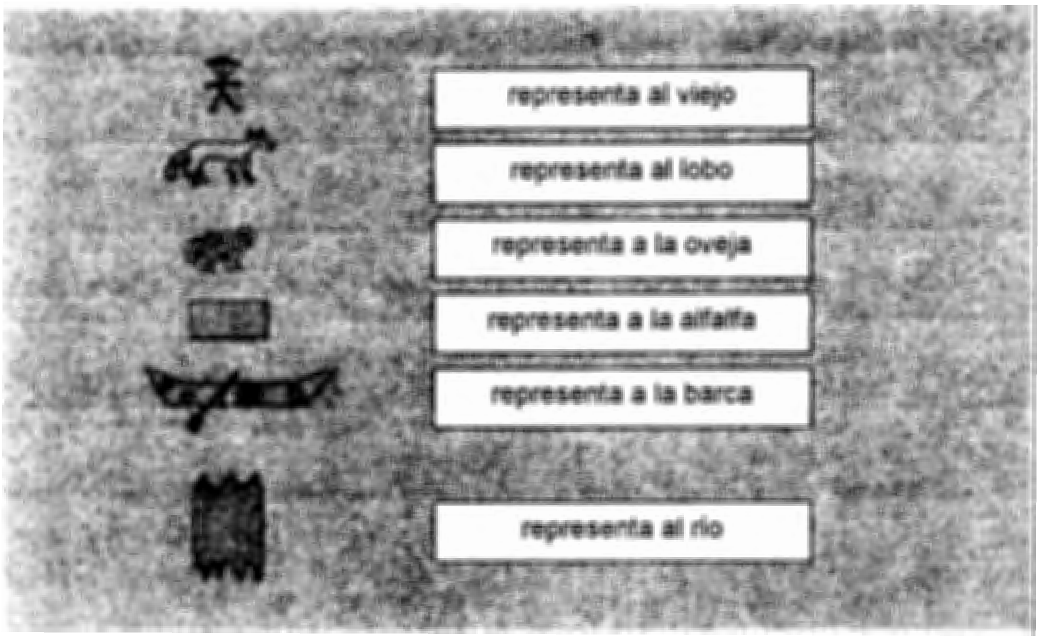
Preguntas como éstas, corresponden a lo que en investigación de operaciones se conoce como análisis de sensibilidad, las últimas nos sugieren que existen más métodos de solución y, en caso de existir otras soluciones, la siguiente pregunta obligada es ¿cuál es la mejor solución o la solución óptima?

Otro método de solución es trasladarnos al lugar de los hechos y ahí encontrar la solución del problema. Esto es: trasladarnos al Balsas, conocer al viejo y demás entidades del sistema, decirle al viejo cómo pasar el río, cuidando de que no haya ningún error y, una vez que se encontrase en la otra orilla sano y salvo con todas sus pertenencias, regresarnos. Este método de resolver el problema lo llamaremos RESOLUCIÓN DIRECTA. Obviamente, este último método presentaría dificultades y, en el supuesto de nuestro problema, nuestra simple presencia modificaría el problema original del viejo.

Existe otro método que es muy utilizado en la resolución de problemas y al cual le daremos el nombre de SIMULACIÓN. Como su nombre lo indica, este método consiste en hacer una representación del sistema en cuestión, con personas, objetos, etc., e interactuar con el modelo hasta encontrar la solución, que suponemos válida para el problema real.

Finalmente, un tipo de método muy importante y muy común es aquél que utiliza símbolos escritos en un papel. A éste se le ha dado el nombre de método SIMBÓLICO. Así, el método simbólico resuelve problemas a través de dibujos y diagramas (símbolos), mediante los cuales se representan las entidades del sistema, cuidando de utilizar símbolos diferentes para cada entidad distinta.

Por ejemplo, en nuestro problema del viejo podríamos representar los elementos del problema de la siguiente manera:



De esta manera podríamos representar cada uno de los pasos a seguir en la resolución del problema; sin embargo, esto entraña el problema de que algunos de nosotros no sabemos dibujar bien, por lo que el modelado sería muy complicado. Así que busquemos símbolos más sencillos de dibujar, por ejemplo sus iniciales (esto es: V, L, O, A, B, para designar al viejo,al lobo, la oveja, la alfalfa y la barca) y usemos dos líneas verticales para representar al río y una flecha para indicar la dirección en que se efectúa el cruce.

La solución de nuestro problema quedaría:

Todos los personajes en una orilla del rio	VLOA	B	
1. El viejo cruza el río con la oveja	LA	BVO	
	LA	B	VO
2. El viejo regresa solo	LA	VB	O
	LAV	B	O
3. El viejo cruza el río con el lobo	A	BVL	O
	A	B	VLO
4. El viejo regresa con la oveja	A	BVO	L
	AOV	B	L
5. El viejo cruza el río con la alfalfa	O	BVA	L
	O	B	VAL
6. El viejo regresa solo	O	BV	AL
7. El viejo cruza con la oveja		BVO	AL
		B	VOAL

De esta forma, el viejo y su pertenencias quedan a la otra orilla del río. Si observamos detenidamente nuestro modelo simbólico, notamos que se puede simplificar más. Se deja de ejercicio esa simplificación.

1.3 LA ABSTRACCIÓN

En todos los métodos de solución de problemas mencionados, se tiene en común la utilización de un modelo. En el método mental, nuestro modelo fueron imágenes en nuestro cerebro. En el método directo nuestro modelo fue la realidad misma (sin lugar a dudas el mejor modelo). En el de simulación, nuestro modelo consistió en representaciones del problema por medio de personas o cosas, suponiendo que cada una de ellas guardaba entre sí las mismas relaciones que las del problema original y, finalmente, en el método simbólico nuestro modelo consistió de símbolos en el papel.

Cabe resaltar que en los últimos dos métodos (de simulación y simbólico) utilizamos distintos modelos.

Fijaremos nuestra atención en los modelos simbólicos: primero, representamos la situación del problema y sus personajes por medio de dibujos, por ello nuestro modelo es de fácil comprensión para todos, pero tiene la desventaja de ser muy laborioso y, resulta casi imposible de construir a medida que el número de variables y entidades aumenta.

Como segunda opción, utilizamos letras para representar a nuestros personajes (letras relacionadas con ellos). Este modelo tiene la ventaja de que es fácil de elaborar y la desventaja de que, para entenderlo, necesitamos más información sobre el problema y requerimos un mayor esfuerzo mental, ésto debido a que nuestro modelo no se parece a la realidad.

1.4 RESUMEN

1. Los modelos simbólicos que utilizamos, serán cada vez menos reales y más abstractos.

2. En su elaboración efectuamos un proceso de abstracción.

3. Observe que, en los modelos que hemos construido, una vez que elegimos un objeto o símbolo para representar una entidad, éste permanece sin cambiar a lo largo de todo el proceso. Esta es una característica que todo modelo debe satisfacer.

De los métodos de solución que hemos visto, existe una gran diferencia entre el método directo y los demás. En el método directo nosotros no construimos el modelo, el modelo en este caso es la realidad misma. Además de las entidades que hemos considerado intervienen muchas más, como por ejemplo: hay rápidos o las aguas son tranquilas, el río es ancho o angosto, es de día o de noche, etc., factores que sin lugar a dudas intervienen y pueden facilitar o dificultar la solución del problema. En los demás métodos nosotros hemos construido el modelo, tomando de la realidad sólo las entidades que consideramos importantes y desechando las demás.

Otra consideración importante es que si en nuestro modelo nos equivocamos en la búsqueda de la solución, lo único que se habrá perdido es tiempo y un pedazo de papel, mientras que un error, en el método directo, implica la pérdida de una de las pertenencias del viejo. Un error puede corregirse sin mucha dificultad en los modelos, mientras que en la realidad difícilmente podremos hacerlo, ya que la experiencia nos ha enseñado que toda acción conlleva un riesgo.

Se preguntarán ahora, ¿De qué nos sirve resolver un problema en un modelo si, como vimos en nuestro ejemplo, hemos olvidado algunos factores que podrían imposibilitar la solución (tal como la posibilidad de que caiga una tormenta)? A lo que podríamos contestar: la solución encontrada en nuestro modelo nos sirve de guía para la acción, para intentar solucionar felizmente un problema. Además, en nuestro modelo proponemos la solución al problema de cualquier viejo o joven, que se encuentre en esas condiciones y no sólo en el río Balsas, sino en cualquier otro rio; proponemos la solución no de uno, sino de muchos problemas. Por esto la solución en un modelo es más general.

Podemos afirmar que entre más abstracto es un modelo, más generales son sus resultados; es decir, se aplican a un mayor número de casos, pero también hay que recordar que para elaborar e interpretar un modelo, entre más abstracto es más difícil hacerlo.

Finalmente, no hay que perder de vista que existen diferencias sustanciales entre resolver un problema en un modelo y resolverlo en la realidad. Lo que significa que no debemos ser tan mecánicos en la aplicación a la realidad de una solución obtenida a través de un modelo, ya que no podemos olvidar que en la realidad la situación es cambiante y, por consiguiente, elementos que en un principio desechamos al elaborar el modelo pudieron haberse convertido en determinantes, modificando con ello las condiciones reales del problema, lo cual imposibilita la aplicación de la solución encontrada.

CAPÍTULO 2
MATRICES

2.1 INTRODUCCIÓN

El problema central del Álgebra Lineal es la solución de ecuaciones lineales simultáneas; el caso más importante y el más simple, es aquél en que el número de incógnitas es igual al número de ecuaciones. Comenzaremos con dos ecuaciones lineales con dos incógnitas x_1 y x_2:

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

Donde a_{11}, a_{12}, a_{21}, a_{22}, b_1 y b_2 son números dados. Cada una de estas ecuaciones es la ecuación de una recta (en el plano x_1x_2). La pendiente de la primera recta es: $-a_{11}/a_{12}$; y la pendiente de la segunda recta es: $-a_{21}/a_{22}$ (si $a_{12} \neq 0$ y $a_{22} \neq 0$). Una solución de este sistema es una pareja de números (x_1, x_2) que satisfacen el sistema.

Otra interpretación consiste en observar el sistema en forma de columnas, se tiene entonces:

$$x_1\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Y se trata aquí de encontrar la combinación de los vectores en el lado izquierdo que dé como resultado el lado derecho. El ejemplo siguiente ilustra estas dos interpretaciones.

<u>EJEMPLO 2.1</u>

Resuelva el siguiente sistema de ecuaciones:

$$\begin{aligned} 2x_1 + \ \ 4x_2 &= 2 \qquad (1) \\ 4x_1 + 11x_2 &= 1 \qquad (2) \end{aligned}$$

multiplicando la primer ecuación por dos y restándola a la segunda se tiene:

$$\begin{aligned} 4x_1 + \ \ 8x_2 &= 4 \qquad (1) \\ 4x_1 + 11x_2 &= 1 \qquad (2) \end{aligned}$$

de donde x_2=-1 y x_1=3, entonces la solución es $(x_1, x_2) = (3,-1)$, y geométricamente se representa en la figura 2.1:

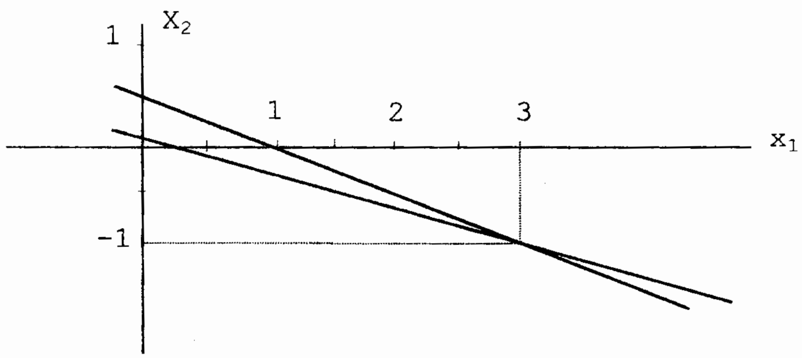


Figura 2.1

Si seguimos la segunda interpretación geométrica, consideramos entonces las columnas del sistema y no los renglones, y se tiene:

$$x_1\begin{bmatrix}2\\4\end{bmatrix}+x_2\begin{bmatrix}4\\11\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}$$

Como ya decíamos, se quiere encontrar la combinación de los vectores en el lado izquierdo que dé como resultado el lado derecho. Geométricamente se representa en la figura 2.2:

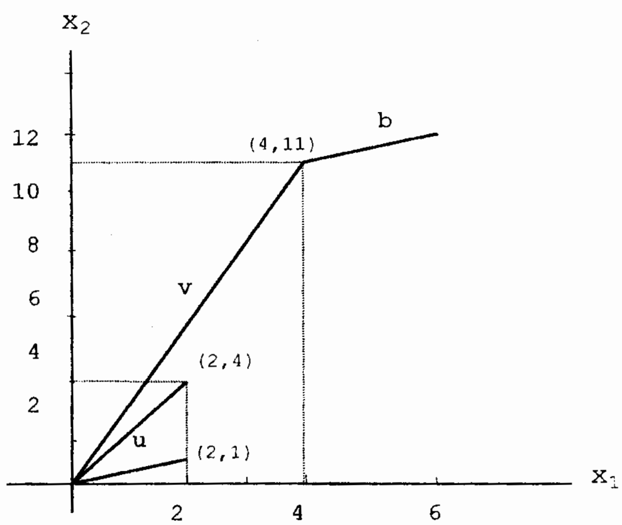


Figura 2.2

Multiplicamos el vector $[2,4]^t$ por 3 y su longitud se triplica:

$$3\begin{bmatrix}2\\4\end{bmatrix}=\begin{bmatrix}6\\12\end{bmatrix}$$

Cuando el vector se multiplica por -1 su dirección se invierte:

$$-\begin{bmatrix}4\\11\end{bmatrix}=\begin{bmatrix}-4\\-11\end{bmatrix}$$

Geométricamente, al sumar vectores se coloca un vector al inicio de donde el otro termina, o algebraicamente sumando sus componentes correspondientes:

$$3\begin{bmatrix}2\\4\end{bmatrix}-1\begin{bmatrix}4\\11\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}$$

Con ambos métodos se obtiene el mismo resultado $x_1 = 3$ y $x_2 = -1$

Podemos planteamos ahora las siguientes preguntas:

1. ¿ Cuándo tiene solución el sistema?

2. ¿ Cuántas soluciones tiene?

Veamos algunos ejemplos:

EJEMPLO 2.2

Considere el sistema:

$$\begin{aligned}x_1 - x_2 &= 7\\ x_1 + x_2 &= 5\end{aligned}$$

Si sumamos las dos ecuaciones obtenemos: $x_1 = 6$ y $x_2 = -1$. Esta es la solución del sistema, es decir, cumple con las dos ecuaciones y además es única, lo que se puede comprobar geométricamente:

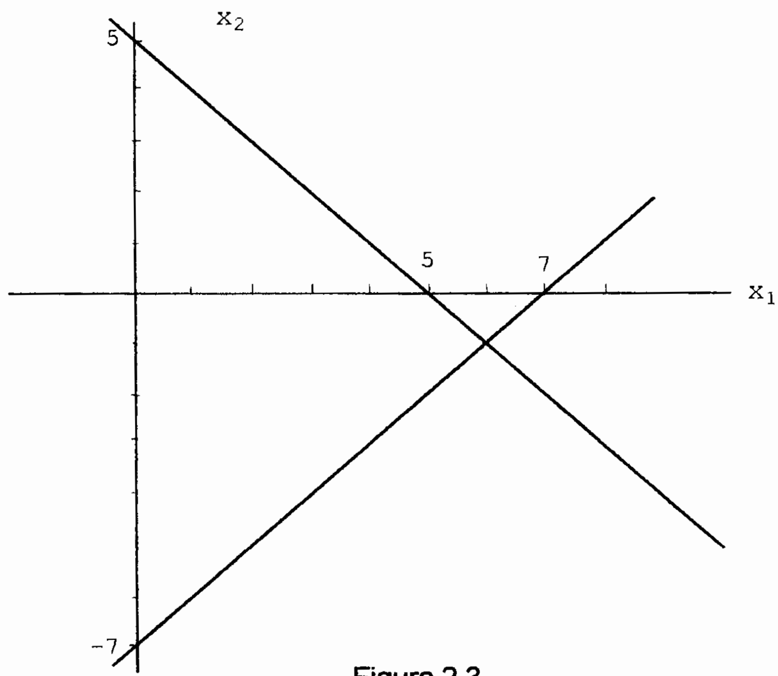


Figura 2.3

En la figura 2.3 se puede ver que la solución es el punto de intersección de las dos rectas.

EJEMPLO 2.3
Considere el sistema:

$$x_1 - \quad x_2 = 7 \qquad (1)$$
$$2x_1 - 2x_2 = 14 \qquad (2)$$

Se puede ver claramente que la ecuación (2) es resultado de multiplicar la primera por 2, de aquí tenemos que $x_2 = x_1 - 7$.

Así, el par $(x_2, x_1 - 7)$ es una solución al sistema para cualquier número real x_1 por lo tanto el sistema tiene un número infinito de soluciones, tantas como valores pueda tomar x_1. Por ejemplo, los siguientes pares son soluciones:

(7,0), (0,-7), (8,1), (1,-6), (3,-4), y (-2,-9)

Lo podemos verificar geométricamente en la figura 2.4:

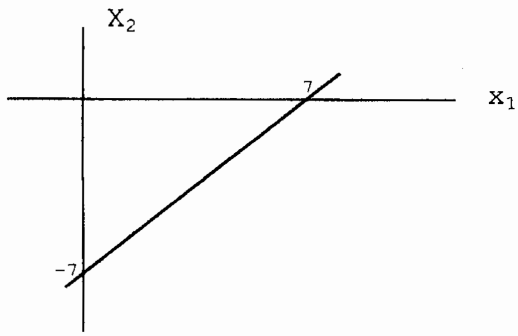


Figura 2.4

EJEMPLO 2.4
Considere el sistema:

$$x_1 - \ \ x_2 = 7 \qquad (1)$$
$$2x_1 - \ 2x_2 = 13 \qquad (2)$$

Si multiplicamos la primer ecuación por dos, tenemos:

$$2x_1 - \ 2x_2 = 14$$

Contrariamente a lo que se afirma en la segunda ecuación; por lo tanto, el sistema no tiene solución, como podemos ver en la figura 2.5:

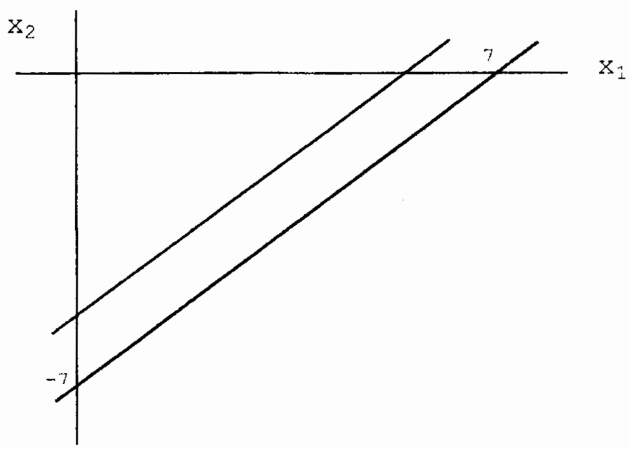


Figura 2.5

Considere nuevamente el sistema de dos ecuaciones con dos incógnitas:

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

Si multiplicamos la primer ecuación por a_{22} y la segunda por a_{12} tenemos:

$$\begin{aligned} a_{11}a_{22}x_1 + a_{12}a_{22}x_2 &= a_{22}b_1 \\ a_{12}a_{21}x_1 + a_{12}a_{22}x_2 &= a_{12}b_2 \end{aligned} \qquad (2)$$

El sistema (1) y el sistema (2) son equivalentes. Esto significa que cualquier solución del sistema (1) es una solución del sistema (2) y viceversa. Restando la segunda ecuación de la primera en el sistema (2) se tiene:

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2$$

Si $(a_{11}\, a_{22} - a_{12}\, a_{21}) \neq 0$, se puede dividir entre esta cantidad y se tiene:

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}$$

Entonces se sustituye este valor en el sistema (1) y se procede a resolver para x_2, encontrándose así la única solución del sistema.

Definición 2.1
Se define el *DETERMINANTE DEL SISTEMA* (1) como:

Determinante del sistema (1) = $(a_{11}\, a_{22} - a_{12}\, a_{21})$

Y se ha mostrado que si el determinante del sistema (1) es diferente de cero, entonces el sistema tiene una única solución. Si es igual a cero, no tiene solución o tiene un número infinito de soluciones.

2.2 MATRICES

Como veremos más adelante, la solución de un sistema de ecuaciones lineales se trata eficientemente ignorando las letras y los signos de igual. El resultado es un ordenamiento de números conocido como **matriz**. Las matrices son una herramienta fundamental para realizar cálculos eficientes en álgebra lineal.

Una **matriz** es un ordenamiento rectangular de números, por lo general encerrados en paréntesis o corchetes. Por ejemplo:

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

es la matriz de los coeficientes del sistema de ecuaciones lineales:

$$\begin{aligned} x_1 - x_2 &= 7 \\ x_1 + x_2 &= 5 \end{aligned}$$

En general, A es una matriz de mxn si:

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

Si n = m, la matriz es cuadrada, y si $n \neq m$, la matriz es rectangular.

OPERACIONES CON MATRICES

Definición 2.2

SUMA ALGEBRAICA DE MATRICES. sean A y B dos matrices de orden mxn, la suma o diferencia de ambas A±B es otra matriz C, de orden mxn, en la que cada elemento de C es igual a la suma o diferencia de los elementos correspondientes de A y B. Por ejemplo:

<u>EJEMPLO 2.5</u>

Sean A y B como sigue:

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

Entonces A + B está dada por:

$$A + B = \begin{bmatrix} 1+2 & 2+3 & 3+0 \\ 0-1 & 1+2 & 4+5 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 3 \\ -1 & 3 & 9 \end{bmatrix}$$

Dos matrices de distinto orden no se pueden sumar ni restar. La suma de k matrices A es otra matriz del mismo orden que A cuyos elementos son iguales a los correspondientes de A multiplicados por k.

EJEMPLO 2.6

Considere la matriz A como sigue:

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

Se tiene:

$$A + A + A = \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -6 \\ 6 & 9 \end{bmatrix} = 3A$$

En particular, -A se conoce como la **matriz opuesta** de A.

MULTIPLICACIÓN DE MATRICES

Considere el sistema lineal:

$$\begin{aligned} 2x_1 - 3x_2 &= 7 \\ 3x_1 - 4x_2 &= 2 \end{aligned} \qquad (3)$$

Los datos importantes para este sistema lineal están contenidos en la matriz de coeficientes:

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

y el vector columna:

$$b = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

de las constantes que aparecen en el lado derecho.

Sea también:

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

el vector columna de las incógnitas del sistema (3). Es costumbre abreviar el sistema (3) como:

$$Ax = b$$

o bien:

$$\begin{bmatrix} 2 & -3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

Esta multiplicación sugiere la multiplicación de la matriz A por un vector columna x. Se define entonces:

Definición 2.3
PRODUCTO PUNTO O ESCALAR. Dados dos vectores a y b con el mismo número de componentes, el producto punto o escalar de a y b está dado por:

$$a\,b = a_1\,b_1 + a_2\,b_2 + \ldots + a_n\,b_n$$

Es esencial diferenciar entre los conceptos de vector y número. El término **escalar** se suele utilizar en álgebra lineal en lugar de número. El producto punto también se llama **producto escalar**, pues $a \cdot b$ es una cantidad escalar.

EJEMPLO 2.7
Efectúe el producto escalar de los vectores a y b donde:

a) $a = (1, -3, 4)$ y $b = (2, 0, 8)^t$.

Se tiene entonces: $a.b = (1)(2) + (-3)(0) + (4)(8) = 2 + 0 + 32 = 34$

b) $a = (1, 5, 6)$ y $b = (2, 4)^t$

Se observa que a.b no está definido, ya que no tienen el mismo número de componentes.

El producto AB de dos matrices A y B se forma tomando todos los productos punto posibles de los vectores renglón de A por los vectores columna de B. Para que estos productos estén definidos, el número de componentes de cada renglón de A debe ser igual al número de componentes de cada columna de B. Si AB está definido y A es de mxn, de modo que sus renglones tengan n componentes, entonces B debe ser de nxs para que las columnas de B tengan también n componentes.

$$AB = (c_{ij}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & b_{1s} \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nj} & b_{ns} \end{bmatrix}$$

En términos de sumatoria tenemos:

$$c_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

AB está definida sólo cuando el número de columnas de A es igual al número de renglones de B. La matriz producto tiene tamaño:

(número de renglones de A)x(número de columnas de B)

EJEMPLO 2.8

Efectúe el producto de las siguientes matrices A y B:

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

$$AB = \begin{bmatrix} -3 & 8 & 9 & -1 \\ 38 & -10 & 4 & 20 \end{bmatrix}$$

2.3 PROPIEDADES DE LAS MATRICES

Consideremos ahora las siguientes propiedades de las matrices, donde se involucran la suma y el producto de las mismas.

1. A(B + C) = AB + AC (primer propiedad distributiva).
2. (A + B)C = AC + BC (segunda propiedad distributiva).
3. A(BC) = (AB)C (propiedad asociativa).

Sin embargo:

4. $AB \neq BA$ en general.
5. AB = 0 no implica necesariamente que A=0 ó B=0.
6. AB = AC no implica necesariamente que B = C.

Si A y B son dos matrices cuadradas y se verifica que AB = BA, dichas matrices se llaman **permutables o conmutativas**.

EJEMPLO 2.9

Sean las matrices A, B, C y D como sigue:

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

Se tiene entonces que AB = AC y B $\neq$C:

$$AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

Además: AD = 0, con A $\neq$ 0 y D $\neq$ 0.

CASOS PARTICULARES DE MATRICES CUADRADAS

Definición 2.4

Una matriz A de manera que A^{k+1} = A, siendo k un número entero positivo, se llama **periódica**. Si k es el menor número entero positivo para el cual A^{k+1} = A, la matriz A tiene periodo k. Si k = 1, esto es: A^2 = A, la matriz A se llama **idempotente**.

EJEMPLO 2.10

a) Sea la matriz A como sigue:

$$A = \begin{bmatrix} 1 & -2 & -6 \\ -3 & 2 & 9 \\ 2 & 0 & -3 \end{bmatrix}$$

es periódica de periodo 2, ya que A^3 = A.

b) La matriz A es idempotente, esto es: A^2 = A, con A definida como sigue:

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

Definición 2.5

Una matriz A tal que A^p = 0 siendo p un número entero positivo se llama **nilpotente**. Si p es el menor número entero positivo para el cual A^p = 0, la matriz A es nilpotente de índice p.

EJEMPLO 2.11

Sea la matriz A como sigue:

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

Es nilpotente de orden 3, $A^3 = 0$.

Definición 2.6

MATRIZ TRANSPUESTA. La matriz transpuesta de una matriz A de orden mxn es la matriz A^t de orden nxm que se obtiene permutando los renglones por las columnas.

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

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

EJEMPLO 2.12

Encuentre la transpuesta de la siguiente matriz A:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Observe que el elemento a_{ij} de A (renglón i, columna j) es el a_{ji} de A^t (renglón j, columna i).

O bien, si

$$A = [a_1 \quad \dots \quad a_n] \qquad A^t = \begin{bmatrix} a_1 \\ \dots \\ a_n \end{bmatrix}$$

el producto escalar de dos vectores columna <x, y> está dado por:

$$\langle x, y\rangle = x^t y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

O bien:

$$\langle x, y\rangle = y^t x = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

De donde:

$$x^t y = y^t x = \sum_{i=1}^{n} x_i y_i$$

PROPIEDADES DE LAS MATRICES TRANSPUESTAS

Sean A^t y B^t las transpuestas de las matrices A y B respectivamente, y k un escalar cualquiera; de aqui que:

a) $(A^t)^t = A$
b) $(kA)^t = kA^t$
Además, se verifica que si $A = A^t$, entonces la matriz A es simétrica.

2.4 RESUMEN

1. Una matriz de mxn es una disposición rectangular ordenada de números que contiene m filas y n columnas.
2. Una matriz de mx1 es un vector columna con m componentes y una matriz de 1xn es un vector fila con n componentes.
3. El producto punto del vector a, con componentes $a_1, a_2, \ldots, a_n$, por el vector b, con componentes $b_1, b_2, \ldots, b_n$, es el escalar:

$$a \cdot b = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

4. El producto AB de una matriz A de mxn y de una matriz B de nxs, es la matriz C de mxs cuyo elemento c_{ij} en la i-ésima fila y la j-ésima columna es el producto punto del i-ésimo vector fila de A por el j-ésimo vector columna de B. En general, $AB \neq BA$.
5. Si $A = (a_{ij})$ y $B = (b_{ij})$ son matrices del mismo tamaño, entonces A + B es la matriz de ese tamaño con elemento $a_{ij} + b_{ij}$ en la i-ésima fila y la j-ésima columna.
6. Para cualquier matriz A y escalar r, la matriz rA se halla multiplicando cada elemento de A por r.
7. La transpuesta de una matriz A de mxn es la matriz A^t de nxm que tiene como k-ésimo vector fila al k-ésimo vector columna de A.

2.5 NOTAS HISTÓRICAS

El término **matriz** se mencionó por primera vez en la literatura matemática en un artículo de 1850 de James Joseph Sylvester (1814-1897). El significado usual no técnico de ese término es *"lugar donde algo se crea, produce o desarrolla"*. Para Sylvester, entonces, una matriz (definida como "un ordenamiento oblongo de términos") era una entidad a partir de la cual uno podía formar varias porciones cuadradas para producir determinantes. Estas últimas cantidades, formadas a partir de matrices cuadradas, eran bastante bien conocidas en esa época.

La **multiplicación de matrices** se originó en la composición de sustituciones lineales. Aunque hay indicios de dichas ideas en el trabajo de Euler y Lagrange, fue Karl Friedrich Gauss (1777-1855) quien las estudió de modo exhaustivo en su trabajo de 1801, *Disquisitiones Arithmeticae*, en conexión con formas cuadráticas (funciones con dos variables de la forma $Ax^2 + 2Bxy + Cy^2$). En particular, una sustitución lineal de la forma:

$$x = ax' + by' \qquad y = cx' + dy' \qquad (1)$$

Convierte una de dichas formas F en x e y; en otra forma, F' en x', y'. Si una segunda sustitución:

$$x' = ex'' + fy'' \qquad y' = gx'' + hy'' \qquad (2)$$

Transforma F' en una forma F'' en x'', y'', entonces la composición de las sustituciones, hallada al reemplazar x', y' en (1) por sus valores en (2), da una sustitución que transforma F en F'':

$$x = (ae + bg)x'' + (af + bh)y'' \qquad y = (ce+dg)x'' + (cf+dh)y'' \qquad (3)$$

La matriz de coeficientes de la sustitución (3) es el producto de las matrices de coeficientes de las sustituciones (1) y (2). Gauss realizó una composición análoga de sustituciones para formas con 3 variables, que da la multiplicación de matrices de 3 x 3.

A Gauss se le suele llamar "príncipe de las matemáticas", pues durante una larga carrera científica hizo importantes contribuciones a campos tan variados como teoría de los números, álgebra, geometría, análisis complejo, astronomía, geodesia y mecánica.

CAPÍTULO 3
SOLUCIÓN DE SISTEMAS LINEALES

3.1 INTRODUCCIÓN

A menudo se compara la tarea de los matemáticos con la de los criminalistas que, forzosamente, tenían que hacer prender una incógnita. Hay, sin embargo, un aspecto en el que esta comparación nos deja abandonados en el problema: cualquier tribunal reputará de mucho más difícil la captura de un malhechor solitario, que sólo se confía a sí mismo prescindiendo de cómplices peligrosos; una gran banda de ladrones se traiciona a sí misma mucho antes. En las matemáticas y en particular en el álgebra lineal por el contrario, cada nueva incógnita nos hace la vida considerablemente más difícil.

Imaginemos para empezar, con un ejemplo numérico sencillo:

$$x + y = 62$$

x e y son dos números desconocidos, que precisamente tratamos de encontrar. Pero, evidentemente, esta exigencia está completamente indeterminada, puesto que hay toda una serie de números que satisfacen a la ecuación, como por ejemplo:
$x = 1$, $y = 61$; $x = 6$, $y = 56$; $x = 70$, $y = -8$, por no hablar de la avalancha de fracciones. Hay pues una infinidad de pares de números, x, y que representan soluciones a la ecuación; la cosa tiene así poca apariencia de problema matemático.

El caso recuerda aquella otra pregunta: Entre mis amistades hay un matrimonio que entre los dos suman 62 años. ¿Qué edad tiene cada uno de ellos? La cuestión cambia enseguida cuando se agrega una segunda condición, por ejemplo que la diferencia de sus edades es de dos años. Con ello el problema queda determinado y se tiene un sistema de ecuaciones lineales. Este problema o ejemplo tan simple da lugar a un sistema, mismo que podemos encontrar en problemas con más incógnitas y más ecuaciones referidas por ejemplo a restricciones industriales.

En este capítulo analizaremos más de cerca lo referente a sistemas de ecuaciones lineales y sus soluciones.

3.2 EL MÉTODO DE REDUCCIÓN DE GAUSS

Recordemos nuevamente que se pretende resolver sistemas de ecuaciones lineales cuadrados, es decir: con el mismo número de ecuaciones que de incógnitas.

Sea el sistema:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned} \qquad (3.1)$$

Este sistema está completamente determinado por su matriz de coeficientes $A=(a_{ij})$ y el vector columna b con i-ésimo elemento b_i. La matriz aumentada se denota $(A|b)$ y se escribe:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{2n} & b_2 \\ a_{n1} & a_{n2} & a_{nn} & b_n \end{array}\right] \qquad (3.2)$$

El vector columna:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

cuyas componentes satisfacen el sistema (3.1) es una *solución* de dicho sistema.

Es fácil ver que las soluciones del sistema (3.1) son las mismas que las de cualquier sistema obtenido de (3.1) por los siguientes procedimientos.

1. Intercambio de dos ecuaciones.
2. Multiplicación de una ecuación del sistema (3.1) por una constante $\neq 0$.
3. Sustitución de una ecuación por la suma de sí misma con un múltiplo de una ecuación diferente del sistema.

Aplicados al sistema (3.1), estos procedimientos corresponden a las **operaciones elementales en las filas** que se aplican a toda la matriz aumentada (3.2)

Para resolver el sistema (3.1) mediante el método de Gauss, tratamos de usar operaciones elementales en las filas para reducir la matriz (3.2) a una matriz de la forma:

$$\left[\begin{array}{ccc|c} u_{11} & u_{12} & u_{1n} & c_1 \\ & u_{22} & u_{2n} & c_2 \\ & \ddots & \vdots & \vdots \\ & & u_{nn} & c_n \end{array}\right]$$

Donde la parte cuadrada U a la izquierda de la partición tiene entradas igual a cero debajo de la diagonal principal. La matriz de coeficientes A original del sistema (3.1) se transforma en una **matriz triangular superior U** con ceros debajo de la diagonal principal. El sistema se convierte en:

$$Ux = c$$

Si una matriz B se puede obtener de una matriz A a través de operaciones elementales en las filas, entonces B es equivalente por filas a A. Por lo cual las matrices A y U descritas antes son equivalentes por filas.

EJEMPLO 3.1

Si se tiene la matriz triangular superior aumentada:

$$\left[\begin{array}{ccc|c} -5 & -1 & 3 & 3 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 2 & -4 \end{array}\right]$$

Halle la solución del sistema.

Solución:

Las ecuaciones correspondientes serían:

$$\begin{aligned} -5x_1 - x_2 + 3x_3 &= 3 \\ 3x_2 + 5x_3 &= 8 \\ 2x_3 &= -4 \end{aligned}$$

de la última ecuación se tiene que $x_3 = -2$; sustituyendo en la segunda ecuación se tiene:

$$3x_2 + 5(-2) = 8$$

y de aquí $\quad x_2 = 6$

Sustituyendo x_2 y x_3 en la primer ecuación se obtiene:

$$-5x_1 - 6 + 3(-2) = 3$$

y de aquí: $\quad x_1 = -3$

Así, la solución es $x^T = [-3, 6, -2]$. Este procedimiento recibe el nombre de **sustitución regresiva**.

EJEMPLO 3.2

Resuelva el sistema lineal a través del método de Gauss con sustitución regresiva:

$$\begin{aligned} x_2 - 3x_3 &= -5 \\ 2x_1 + 3x_2 - x_3 &= 7 \\ 4x_1 + 5x_2 - 2x_3 &= 10 \end{aligned}$$

Solución:

Se reduce la matriz aumentada por medio de operaciones elementales en las filas. Los pivotes se encierran en un círculo.

Se intercambian las filas 1 y 2:

$$\left[\begin{array}{ccc|c} 0 & 1 & -3 & -5 \\ 2 & 3 & -1 & 7 \\ 4 & 5 & -2 & 10 \end{array}\right] \approx \left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 4 & 5 & -2 & 10 \end{array}\right]$$

Se suma la fila 1 multiplicada por -2 a la fila 3:

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 4 & 5 & -2 & 10 \end{array}\right] \approx \left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & 0 & -4 \end{array}\right]$$

Se suma la fila 2 a la 3:

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & 0 & -4 \end{array}\right] \approx \left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & -3 & -9 \end{array}\right]$$

De la última matriz obtenemos, por sustitución regresiva:

$x_3=3$ $\qquad$ $x_2=4$ $\qquad$ $x_1=-1$

REDUCCIÓN DE GAUSS A LA FORMA TRIANGULAR SUPERIOR: CASO DE SOLUCIÓN ÚNICA

PASO 1 Si la entrada superior de la columna 1 es cero, entonces efectuar una operación de intercambio de filas para obtener un elemento distinto de cero en la parte superior de la columna. Siempre es posible esto en el caso de solución única. Llamamos pivote al elemento elegido distinto de cero.

PASO 2 Efectuar operaciones elementales en las filas de tal forma que las filas inferiores resultantes tengan cero como primer entrada.

PASO 3 Después del paso 2, la matriz tiene la forma:

$$
\left[\begin{array}{c|cccc}
u_{11} & u_{12} & \cdots & u_{1n} & c_1 \\
\hline
0 & x & \cdots & x & x \\
0 & \vdots & \ddots & \vdots & \vdots \\
0 & x & \cdots & x & x
\end{array}\right] \tag{3.3}
$$

Y la primer columna está en la forma que queremos. Se elimina (mentalmente) la fila superior y la primera columna de la matriz (3.3), dejando la parte enmarcada de ésta. Se vuelve al paso 1 con esta matriz más pequeña y se repite el procedimiento para componer la siguiente columna. Se continúa hasta obtener la forma triangular superior.

3.3 EL MÉTODO DE GAUSS-JORDAN

Una matriz cuadrada está en la forma diagonal si todas las entradas fuera de la diagonal principal son cero. El método de Gauss-Jordan es un avance del método de Gauss y usa operaciones elementales en las filas para reducir la parte izquierda de una matriz aumentada a la forma diagonal con todos los pivotes iguales a uno, es decir, a la matriz identidad $\mathtt{I}$.

Suponga, por ejemplo, que una matriz aumentada se reduce a la forma:

$$
\left[\begin{array}{ccc|c}
1 & 0 & 0 & 9 \\
0 & 1 & 0 & 7 \\
0 & 0 & 1 & -1
\end{array}\right]
$$

Se ve de inmediato que la solución al sistema original está dada por:

$$x_1=9 \qquad x_2=7 \qquad x_3=-1$$

<u>EJEMPLO 3.3</u>
Encuentre, mediante la reducción de Gauss-Jordan, la solución del sistema:

$$
\begin{aligned}
x_1 - 2x_2 + x_3 &= 7 \\
2x_1 - 5x_2 + 2x_3 &= 6 \\
3x_1 + 2x_2 - x_3 &= 1
\end{aligned}
$$

Solución:

1. - Se suma la fila 1, multiplicada por 2, a la fila 2 y se suma la fila 1, multiplicada por -3, a la fila 3:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 2 & -5 & 2 & 6 \\ 3 & 2 & -1 & 1 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & -1 & 0 & -8 \\ 0 & 8 & -4 & -20 \end{array}\right]$$

2. - Se multiplica la fila 2 por -1:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & -1 & -0 & -8 \\ 0 & 8 & -4 & -20 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & 1 & 0 & 8 \\ 0 & 8 & -4 & -20 \end{array}\right]$$

3. - Se suma la fila 2, multiplicada por 2 a la fila 1 y se suma la fila 2, multiplicada por -8 a la fila 3:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 7 \\ 0 & 1 & 0 & 8 \\ 0 & 8 & -4 & -20 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & -4 & -84 \end{array}\right]$$

4. - Se multiplica la fila 3 por $-\frac{1}{4}$

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & -4 & -84 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{array}\right]$$

5. - Se suma la fila 3, multiplicada por -1, a la fila 1:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 23 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{array}\right] \approx \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{array}\right]$$

De aquí se observa que la solución es: $x_1=2$ $x_2=8$ $x_3=21$.

3.4 SISTEMAS CON LA MISMA MATRIZ DE COEFICIENTES

Dados dos sistemas lineales $Ax = b$ con la misma matriz de coeficientes A, pero distintos vectores columna b, suponga que queremos resolver los dos sistemas simultáneamente:

$$Ax = b \qquad Ay = b'$$

En lugar de resolver uno después de otro, es más eficiente tomar una sola matriz aumentada con la matriz de coeficientes A a la izquierda de la partición y aumentarla con dos vectores columna b y b', a la derecha de la partición. Entonces se reduce esta matriz partida de modo que la parte izquierda tenga forma triangular superior o diagonal, resolviendo así ambos sistemas simultáneamente.

EJEMPLO 3.4
Resuelva los sistemas lineales utilizando el método de Gauss-Jordan.

$$\begin{aligned}
2x_1 - 4x_2 \qquad\qquad &= -10 \\
x_1 - 3x_2 \qquad + x_4 &= -4 \\
x_1 \qquad - x_3 + 2x_4 &= 4 \\
3x_1 - 4x_2 + 3x_3 - x_4 &= -11
\end{aligned}$$

$$\begin{aligned}
2y_1 - 4y_2 \qquad\qquad &= -8 \\
y_1 - 3y_2 \qquad + y_4 &= -2 \\
y_1 \qquad - y_3 + 2y_4 &= 9 \\
3y_1 - 4y_2 + 3y_3 - y_4 &= -15
\end{aligned}$$

Solución:

$$\left[\begin{array}{cccc|cc} 2 & -4 & 0 & 0 & -10 & -8 \\ 1 & -3 & 0 & 1 & -4 & -2 \\ 1 & 0 & -1 & 2 & 4 & 9 \\ 3 & -4 & 3 & -1 & -11 & -15 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & -2 & 0 & 0 & -5 & -4 \\ 1 & -3 & 0 & 1 & -4 & -2 \\ 1 & 0 & -1 & 2 & 4 & 9 \\ 3 & -4 & 3 & -1 & -11 & -15 \end{array}\right]$$

$$\left[\begin{array}{cccc|cc} 1 & -2 & 0 & 0 & -5 & -4 \\ 0 & -1 & 0 & 1 & 1 & 2 \\ 0 & 2 & -1 & 2 & 9 & 13 \\ 0 & 2 & 3 & -1 & 4 & -3 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & -2 & 0 & 0 & -5 & -4 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 2 & -1 & 2 & 9 & 13 \\ 0 & 2 & 3 & -1 & 4 & -3 \end{array}\right]$$

$$\left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 0 & -1 & 4 & 11 & 17 \\ 0 & 0 & 3 & 1 & 6 & 1 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 0 & 1 & -4 & -11 & -17 \\ 0 & 0 & 3 & 1 & 6 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 0 & 1 & -4 & -11 & -17 \\ 0 & 0 & 0 & 13 & 39 & 52 \end{array}\right] \approx \left[\begin{array}{cccc|cc} 1 & 0 & 0 & -2 & -7 & -8 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 1 & 0 & 1 & -4 & -11 & -17 \\ 0 & 0 & 0 & 1 & 3 & 4 \end{array}\right]$$

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

A partir de esta última matriz vemos que las soluciones son:

$x_1=-1$ $\quad$ $x_2=2$ $\quad$ $x_3=1$ $\quad$ $x_4=3$
$y_1=0$ $\quad$ $y_2=2$ $\quad$ $y_3=-1$ $\quad$ $y_4=4$

De aquí se puede concluir un resultado que se resume en el siguiente teorema:

TEOREMA 3.1
Sea Ax = b un sistema de n ecuaciones lineales en n incógnitas. El sistema tiene una solución única si y sólo si la matriz A de los coeficientes es equivalente en filas a la matriz identidad I de nxn.

3.5 MATRICES ELEMENTALES

Recordemos que las operaciones fila en una matriz son:

a) Intercambio de filas.
b) Multiplicación de una fila por una constante diferente de cero.
c) Sumar el múltiplo de una fila a otra.

Asimismo, las operaciones columna de una matriz son las mismas que las operaciones fila. La colección de operaciones fila y columna en una matriz se dicen **operaciones matriciales elementales**.

Definición 3.1
Una matriz E se dice elemental si resulta de efectuar una operación matricial en la matriz identidad $\mathcal{I}$ de orden nxn.

Las matrices elementales son:

E_{pq} = intercambio de filas (columnas) p y q.
$E_p^{(\alpha)}$ = multiplicar una fila (columna) p por $\alpha \neq 0$.
$E_{pq}^{(\beta)}$ = multiplicar la fila p por β y sumarla a la fila q.

EJEMPLO 3.5
Sea:

$$\mathcal{I}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

Considere las siguientes matrices elementales de 3x3:

a) Multiplicar la segunda fila de $\mathcal{I}_3$ por -3:

$$E_2^{(-3)}=\begin{bmatrix}1 & 0 & 0\\ 0 & -3 & 0\\ 0 & 0 & 1\end{bmatrix}$$

b) Intercambiar las filas 2 y 3:

$$E_{32}=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}$$

* **Nota**: $E_{32} = E_{23}$

c) Multiplicar la fila 1 por 4 y sumarla a la fila 3:

$$E_{13}^{(4)}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 4 & 0 & 1\end{bmatrix}$$

Observe que para:

$$E_{32}A=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{31}&a_{32}&a_{33}\\a_{21}&a_{22}&a_{23}\end{bmatrix}$$

se intercambian las filas dos y tres de la matriz A.

Si se hace A E_{32} se tiene:

$$AE_{32}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}=\begin{bmatrix}a_{11}&a_{13}&a_{12}\\a_{21}&a_{23}&a_{22}\\a_{31}&a_{33}&a_{32}\end{bmatrix}$$

y se intercambian las columnas dos y tres de la matriz A.

De aquí podemos concluir que: Para efectuar una transformación elemental de fila sobre una matriz A de nxn se aplica la transformación a la matriz $\mathcal{I}$ de nxn obteniéndose la matriz elemental E correspondiente y a continuación se multiplica por la izquierda A por E; equivalentemente, para efectuar una transformación elemental de columna sobre una matriz A, se multiplica A, por la derecha, por la matriz E correspondiente.

USO DE MATRICES ELEMENTALES

La reducción de una matriz cuadrada a la forma triangular superior U o diagonal se puede lograr a través de multiplicaciones sucesivas por la izquierda por matrices elementales en las filas, entonces existen matrices elementales E_1, E_2, ..., E_t tales que:

$$U = (E_t \ldots E_2\, E_1)\, A$$

<u>EJEMPLO 3.6</u>

Considere las matrices:

a) $A=\begin{bmatrix}2&4&1\\3&2&2\\1&1&3\end{bmatrix}$ $\quad B=\begin{bmatrix}4&4&1\\6&2&2\\2&1&3\end{bmatrix}$

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

Especifique si existe una matriz elemental E tal que AE=B ó EA=B.

Solución:
a) $AE_1^{(2)}=B$ ó $BE_1^{(½)}=A$
b) $E_{13}^{(-1)}A=B$ ó $E_{13}^{(1)}B=A$

EJEMPLO 3.7
Sea:

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}=\begin{bmatrix} 0 & 1 & -3 \\ 2 & 3 & -1 \\ 4 & 5 & -2 \end{bmatrix}$$

Encuentre una matriz C tal que CA sea una matriz triangular superior equivalente en filas a A.

Solución:
Sean: $E_1=E_{12}$ $E_2=E_{13}^{(-2)}$ $E_3=E_{23}^{(1)}$
Entonces: $C=E_3E_2E_1$

3.6 MATRICES INVERSAS

Definición 3.2
Sea A una matriz de nxn. Una matriz C de nxn es una inversa de A si CA = AC = I, donde I es la matriz identidad de nxn.

TEOREMA 3.2 UNICIDAD DE LA INVERSA
Sea A una matriz de nxn con inversa C de modo que CA = AC = I. Si D es una matriz de nxn tal que AD = I, entonces C = D.

Demostración:
Como el producto de matrices es asociativo, se tiene:

$$C(AD) = (CA)D$$

pero AD = I y CA = I; entonces:

$$C(AD) = CI = C \qquad y \qquad (CA)D = ID = D$$

de donde: C = D.

La inversa de una matriz A se denota por A^{-1}.

Definición 3.3
Una matriz cuadrada que tiene inversa se llama **invertible**. Si tal matriz cuadrada no tiene inversa se llama **singular**.

EJEMPLO 3.8
Determine la inversa de la matriz A.

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

Solución:

$$AC = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Haciendo el producto de AC e igualando a I se tiene:

$$\begin{array}{ll} c_{11}=\tfrac{1}{2} & c_{12}=c_{13}=0 \\ c_{22}=\tfrac{1}{4} & c_{21}=c_{23}=0 \\ c_{33}=1 & c_{31}=c_{32}=0 \end{array}$$

de donde:

$$A^{-1} = C = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

EJEMPLO 3.9
Determine la inversa de la matriz B, donde B_i son matrices cuadradas:

$$B = \begin{bmatrix} B_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & B_3 \end{bmatrix}$$

Solución:
Sea C la inversa de B, entonces:

$$BC = \begin{bmatrix} B_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & B_3 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} In_1 & 0 & 0 \\ 0 & In_2 & 0 \\ 0 & 0 & In_3 \end{bmatrix}$$

de aquí se tiene:

$$B_1c_{11}=I_{n1}$$
$$B_2c_{22}=I_{n2}$$
$$B_3c_{33}=I_{n3}$$

Se despejan las c_i's y se obtiene:

$$B^{-1} = C = \begin{bmatrix} B_1^{-1} & 0 & 0 \\ 0 & B_2^{-1} & 0 \\ 0 & 0 & B_3^{-1} \end{bmatrix}$$

EJEMPLO 3.10
Sea la matriz A como sigue:

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

donde $a \neq 0$, $b \neq 0$ y $c \neq 0$. Calcule la inversa de A.

Solución:
Se puede expresar A como producto de matrices A_1, A_2 y A_3 tales que:

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} = \underbrace{\begin{bmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{A_1} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{A_2} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}}_{A_3}$$

entonces A^{-1} se puede ver como:

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

$$A^{-1} \qquad A_3^{-1} \qquad A_2^{-1} \qquad A_1^{-1}$$

PROPOSICIÓN 3.1

Sean A y B matrices de nxn cuyas inversas son conocidas, entonces:

a) $(A^{-1})^{-1} = A$ (significa que A^{-1} tiene inversa)

b) $(AB)^{-1} = B^{-1} A^{-1}$ (significa que AB tiene inversa)

Demostración:

Se demostrará el inciso b); el inciso a) se deja al estudiante como ejercicio.

b) Se da por supuesto que existen matrices A^{-1} y B^{-1} tales que $AA^{-1} = I$ y $BB^{-1} = I$. Utilizando la ley asociativa para producto de matrices, tenemos que:

$$(AB)(B^{-1} A^{-1}) = [A(BB^{-1})] A^{-1} = (AI)A^{-1} = AA^{-1} = I$$

En forma similar:

$$(B^{-1} A^{-1})(AB) = [B^{-1} (A^{-1} A)]B = [B^{-1} (I)]B = BB^{-1} = I$$

Por lo tanto: $(AB)^{-1} = B^{-1} A^{-1}$

INVERSAS DE MATRICES ELEMENTALES

Sea E_{ik} una matriz elemental de intercambio de renglones obtenida al intercambiar los renglones ik. Recuerde que $E_{ik}A$ afecta intercambiando los renglones ik de A; donde A es cualquier matriz del mismo tamaño que E_{ik}. En particular al tomar $A = E_{ik}$ vemos que $E_{ik}E_{ik}$ intercambia las filas ik de E, y por lo tanto, devuelve a E_{ik} a I, así:

$$E_{ik} E_{ik} = I$$

Si multiplicamos el renglón i por un escalar r distinto de cero, entonces:

$$E_i^{(r)} E_i^{(1/r)} = I$$

Si se tiene $E_{pq}^{(ß)}$ entonces su inversa está dada por:

$$E_{pq}^{(ß)} E_{pq}^{(-ß)} = I$$

Por lo tanto: ***toda matriz elemental es invertible***.

EJEMPLO 3.11
Encuentre las inversas de las siguientes matrices elementales:

$$E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_1^{(3)} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_{31}^{(4)} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solución:

$$[E_{12}]^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\left[E_1^{(3)}\right]^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\left[E_{31}^{(4)}\right]^{-1} = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

CÁLCULO DE INVERSAS

Sea $A = (a_{ij})$ una matriz de nxn. Para hallar A^{-1}, si es que existe, se debe encontrar una matriz $X = (x_{ij})$ de nxn tal que $AX = I$, esto es, tal que:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \qquad (3.4)$$

La ecuación matricial (3.4) corresponde a n^2 ecuaciones lineales con las n^2 incógnitas x_{ij}, hay una ecuación lineal para cada una de las n^2 entradas en una matriz de nxn. Por ejemplo, si se igualan las entradas de la posición de la segunda fila y primera columna a cada lado de la ecuación (3.4) se obtiene la ecuación lineal:

$$a_{21}\, x_{11} + a_{22}\, x_{21} + \ldots + a_{2n}\, x_{n1} = 0$$

De estas n^2 ecuaciones lineales, n de ellas comprenden las n incógnitas x_{i1} para $i=1,2,\ldots,n$ y estas ecuaciones están dadas por la ecuación con vectores columna:

$$A\begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.5}$$

Hay también n ecuaciones que incluyen las n incógnitas x_{i2} para $i = 1,2,\ldots,n$, y así sucesivamente. Además de la ecuación (3.5) hay que resolver:

$$A\begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad ,\ldots, \quad A\begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Cada sistema tiene la matriz de coeficientes A, por lo que podemos formar la matriz aumentada:

$$[A|I] = \left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}\right]$$

Mediante operaciones elementales y usando Gauss-Jordan se obtiene:

$$[I|D] = \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & d_{11} & d_{12} & \cdots & d_{1n} \\ 0 & 1 & \cdots & 0 & d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & d_{n1} & d_{n2} & \cdots & d_{nn} \end{array}\right]$$

de donde $D = A^{-1} = (d_{ij})$.

EJEMPLO 3.12

Encuentre la inversa de la matriz A.

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

Solución:

Escribimos $[A \mid I]$ y efectuamos operaciones elementales para obtener $[I \mid A^{-1}]$

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

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & \frac{2}{5} & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & \frac{7}{5} & \frac{2}{5} & -\frac{1}{5} & 1 \end{array}\right] \approx \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{1}{7} & \frac{3}{7} & -\frac{15}{7} \\ 0 & 1 & 0 & \frac{2}{7} & -\frac{1}{7} & -\frac{2}{7} \\ 0 & 0 & 1 & \frac{2}{7} & -\frac{1}{7} & \frac{5}{7} \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{3}{7} & \frac{5}{7} & -\frac{11}{7} \\ 0 & 1 & 0 & \frac{2}{7} & -\frac{1}{7} & -\frac{2}{7} \\ 0 & 0 & 1 & \frac{2}{7} & -\frac{1}{7} & \frac{5}{7} \end{array}\right] =$$

de aquí se tiene:

$$\begin{bmatrix} -\frac{3}{7} & \frac{5}{7} & -\frac{11}{7} \\ \frac{2}{7} & -\frac{1}{7} & -\frac{2}{7} \\ \frac{2}{7} & -\frac{1}{7} & \frac{5}{7} \end{bmatrix} = A^{-1}$$

Dada una matriz, a menudo no se sabe de antemano si es invertible o no. Si se trata de aplicar este procedimiento a una matriz que no es invertible, en algún momento durante el proceso, aparecerá un renglón de ceros en el lado izquierdo, como lo ilustra el siguiente ejemplo:

EJEMPLO 3.13

Encuentre la inversa de la matriz:

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

Solución:

Escribimos $[A \mid I]$ y efectuamos operaciones elementales para obtener $[I \mid A^{-1}]$

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

de aquí se puede observar que ya no es posible seguir efectuando operaciones elementales, por lo cual la matriz A no tiene inversa.

Podemos resumir el procedimiento para encontrar la inversa de una matriz de la siguiente manera:

Cálculo de A^{-1}:

PASO 1 Formar la matriz aumentada $[A \mid I]$.

PASO 2 Aplicar el método de Gauss-Jordan para reducir $[A \mid I]$ a $[I \mid D]$ Si se puede hacer la reducción entonces $A^{-1}=D$. De no ser así, A^{-1} no existe.

CONDICIONES PARA QUE EXISTA A^{-1}

TEOREMA 3.3

Si A es una matriz de nxn, entonces las siguientes afirmaciones son equivalentes, es decir, todas son verdaderas o todas son falsas:

a) A es invertible.

b) $AX = 0$ únicamente tiene la solución trivial.

c) A es equivalente por renglones a I_n

Demostración:

a) $\Rightarrow$ b)

Suponga que A es invertible siendo x_0 cualquier solución de $AX = 0$, por lo tanto $Ax_0=0$. Multiplicando los dos lados de la ecuación por A^{-1} resulta $A^{-1}(Ax_0)=A^{-1}0$, ó bien: $(A^{-1}A)x_0=0$, es decir: $Ix_0=0$ de donde $x_0=0$, por lo cual $AX = 0$ únicamente tiene la solución trivial.

b) $\Rightarrow$ c)
Sea AX=0 la forma matricial asociada al sistema:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= 0 \end{aligned} \tag{3.6}$$

Suponga también que el sistema solamente tiene la solución trivial. Si se resuelve empleando la eliminación de Gauss-Jordan, el sistema de ecuaciones correspondiente a la forma escalonada de la matriz aumentada será:

$$\begin{aligned} x_1 \qquad\qquad &= 0 \\ x_2 \qquad &= 0 \\ \ldots\ldots\ldots\ldots \\ x_n &= 0 \end{aligned}$$

Así, la matriz aumentada asociada a (3.6) es igual a:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right]$$

y se puede llevar a la matriz aumentada

$$\left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right]$$

mediante una serie de operaciones elementales en los renglones. Si se elimina la última columna (de ceros) en ambas matrices, se puede concluir que A se redujo a I_n mediante una serie de operaciones elementales en los renglones, es decir, A es equivalente por renglones a I_n.

c) $\Rightarrow$ a)
Suponga que A es equivalente por renglones a I_n o, lo que es lo mismo, A se puede reducir a I_n mediante una serie finita de operaciones elementales en los renglones, esto se puede hacer multiplicando por la izquierda por una matriz elemental adecuada. De esta manera, es posible hallar matrices elementales E_1, E_2, ..., E_k tales que:

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

Como E_1, E_2,..., E_k son invertibles, premultiplicando en forma sucesiva los dos lados de la ecuación por $E_k^{-1}, \ldots, E_2^{-1}, E_1^{-1}$ se obtiene:

$$A = E_1^{-1} E_2^{-1} \ldots E_k^{-1} I_n = E_1^{-1} E_2^{-1} \ldots E_k^{-1}$$

Como A se puede expresar como un producto de matrices invertibles, entonces A es invertible.

<u>EJEMPLO 3.14</u>
Calcule la inversa de la matriz:

$$A = \begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix}$$

<u>Solución:</u>

$$\left[\begin{array}{cc|cc} 2 & 9 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right] \approx \underset{E_{12}}{\left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 2 & 9 & 1 & 0 \end{array}\right]} \approx$$

$$\approx \underset{E_{12}^{(-2)}}{\left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 0 & 1 & 1 & -2 \end{array}\right]} \approx \underset{E_{21}^{(-4)}}{\left[\begin{array}{cc|cc} 1 & 0 & -4 & 9 \\ 0 & 1 & 1 & -2 \end{array}\right]}$$

por lo tanto:

$$A^{-1} = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix}$$

Ahora expresamos A como producto de matrices elementales:

si $\quad E_{21}^{(-4)} E_{12}^{(-2)} E_{12} A = I$

entonces $\quad A = E_{12} E_{12}^{(2)} E_{21}^{(4)}$

donde:

$$E_{12}=\begin{bmatrix}0&1\\1&0\end{bmatrix}\qquad E_{12}^{(2)}=\begin{bmatrix}1&0\\2&1\end{bmatrix}\qquad E_{21}^{(4)}=\begin{bmatrix}1&4\\0&1\end{bmatrix}$$

Recuerde que queremos resolver un sistema Ax = b, esto lo podemos hacer a través del método de eliminación de Gauss y obtener un sistema equivalente Ux = c, donde U es una matriz triangular superior.

EJEMPLO 3.15

Sea el sistema:

$$\begin{aligned}2x_1 + x_2 + x_3 &= 1\\ 4x_1 + x_2 &= -2\\ -2x_1 + 2x_2 + x_3 &= 7\end{aligned}$$

En términos matriciales se tiene:

$$\begin{bmatrix}2&1&1\\4&1&0\\-2&2&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1\\-2\\7\end{bmatrix}$$

Efectuando operaciones elementales se tiene:

$$\left[\begin{array}{ccc|c}2&1&1&1\\4&1&0&-2\\-2&2&1&7\end{array}\right]\approx\underset{E_{12}^{(-2)}}{\left[\begin{array}{ccc|c}2&1&1&1\\0&-1&-2&-4\\-2&2&1&7\end{array}\right]}\approx$$

$$\approx\underset{E_{13}^{(1)}}{\left[\begin{array}{ccc|c}2&1&1&1\\0&-1&-2&-4\\0&3&2&8\end{array}\right]}\approx\underset{E_{23}^{(3)}}{\left[\begin{array}{ccc|c}2&1&1&1\\0&-1&-2&-4\\0&0&-4&-4\end{array}\right]}$$

y se obtiene:

$$\begin{bmatrix}2&1&1\\0&-1&-2\\0&0&-4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1\\-4\\-4\end{bmatrix}$$

Las matrices elementales usadas son:

$$E_{12}^{(-2)} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_{13}^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad E_{23}^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$$

Todas estas matrices elementales son triangulares inferiores ya que:

a) Si se tiene $E_p^{(a)}$ las matrices son diagonales; es decir, son triangulares inferiores y superiores.
b) Si se tiene $E_{pq}^{(\beta)}$ las matrices son triangulares inferiores si q>p.

Por lo tanto, las operaciones matriciales que convierten a A en U son:

$$E_{23}^{(3)}\, E_{13}^{(1)}\, E_{12}^{(-2)}\, A = U$$

De manera semejante:

$$E_{23}^{(3)}\, E_{13}^{(1)}\, E_{12}^{(-2)}\, b = c$$

3.7 FACTORIZACIÓN LU

¿Qué hacemos para regresar de U a A? ¿Cómo podemos deshacer los pasos de la eliminación gaussiana?

No es difícil deshacer un solo paso por ejemplo; en el paso a), en lugar de restar, sumamos al segundo renglón el doble del primero, de esta manera se invierte la matriz elemental $E_{12}^{(-2)}$ y se tiene:

$$E_{12}^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Análogamente, para b) y c) tenemos las matrices inversas respectivas:

$$E_{13}^{(-1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad E_{23}^{(-3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix}$$

Dichas matrices deshacen por separado cada uno de los pasos a), b) y c). Para deshacer todo el proceso y ver qué matriz regresa a la matriz U en A, nos fijamos en que: c) fue el último paso en realizarse al ir de A a U, por lo que debe ser el primero en invertirse cuando se va en dirección contraria, de aquí que:

$$A = E_{12}^{(2)}\, E_{13}^{(-1)}\, E_{23}^{(-3)}\, U$$

la matriz L que lleva a U de regreso a A debe ser el producto de las 3 matrices:

$$L = E_{12}^{(2)}\, E_{13}^{(-1)}\, E_{23}^{(-3)}$$

De aquí que A = LU, es la factorización de la matriz A como producto de dos matrices L triangular inferior (L = lower) y U triangular superior (U = upper).

La matriz L es la clave de la eliminación gaussiana, es el vínculo entre la matriz A y la matriz U a la que llegamos.

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

La matriz L así construida es triangular inferior con unos en la diagonal principal. Los elementos debajo de la diagonal son precisamente los multiplicadores 2, -1, -3, usados en los 3 pasos de la eliminación. También podemos observar que:

$$L^{-1} = E_{23}^{(3)}\, E_{13}^{(1)}\, E_{12}^{(-2)}$$

Por lo tanto:

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

<u>EJEMPLO 3.16</u>
Reduzca la matriz A, a la forma triangular superior U y cree la matriz L.

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

Solución:

Para hacer más explícita la forma en que se construye la matriz L, haremos el proceso en dos columnas: en una de ellas se va a iniciar con la matriz A y en la otra con la matriz identidad; se van a ir efectuando las operaciones elementales simultáneamente en ambas matrices para así, al finalizar el proceso, obtener las matrices U y L.

Reducción de A a U | **Creación de L**

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

$$E_{12}^{(-2)} \qquad E_{12}^{(2)}$$

$$\approx \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ -1 & 3 & -15 \end{bmatrix} \qquad \approx \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_{13}^{(1)} \qquad E_{13}^{(-1)}$$

$$\approx \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ 0 & 6 & 3 \end{bmatrix} \qquad \approx \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$E_{23}^{(-3)} \qquad E_{23}^{(3)}$$

$$\approx \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ 0 & 0 & -15 \end{bmatrix} \qquad \approx \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 3 & 1 \end{bmatrix}$$

Ahora veremos cómo se puede utilizar la información conservada en L de este ejemplo para resolver un sistema lineal Ax = b con esta matriz A.

EJEMPLO 3.17

Utilice la matriz L para hallar el vector columna c que se presentaría si redujéramos (A | b) a (U | c) aplicando esas mismas operaciones en los renglones.

$$b = \begin{bmatrix} -4 \\ 2 \\ 4 \end{bmatrix}$$

$I_{21}=2$ Significa multiplicar el renglón 1 por -2 y sumarlo al renglón 2 $\approx \begin{bmatrix} -4 \\ 10 \\ 4 \end{bmatrix}$

$I_{31}=-1$ Significa multiplicar el renglón 1 por 1 y sumarlo al renglón 3 $\approx \begin{bmatrix} -4 \\ 10 \\ 0 \end{bmatrix}$

$I_{32}=3$ Significa multiplicar el renglón 2 por -3 y sumarlo al renglón 3 $\approx \begin{bmatrix} -4 \\ 10 \\ 30 \end{bmatrix} = c$

PROPIEDADES

1. - La matriz A puede escribirse, mientras que ninguno de los pivotes sea cero, como un producto LU de una matriz triangular inferior L y una matriz triangular superior U.

2. - Los elementos de la diagonal principal de L son unos, y debajo de la diagonal están los multiplicadores I_{ij}.

3. - U es la matriz de coeficientes que obtenemos después de la eliminación y antes de la sustitución regresiva, los elementos de su diagonal son los pivotes.

<u>EJEMPLO 3.18</u>
Considere la matriz:

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

entonces una factorización de A = LU sería:

$$L = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \qquad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}$$

Los 16 elementos de A se pueden usar para determinar parcialmente los 10 elementos desconocidos de L y el mismo número que los de U. Sin embargo, si el procedimiento nos debe llevar a una solución única, se necesitan cuatro condiciones adicionales para los elementos de L y U. El método que hemos usado consiste en que $l_{11} = l_{22} = l_{33} = l_{44} = 1$; este método se conoce como **Método de Doolittle**. Para este ejemplo se tiene:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{6} & \frac{1}{5} & 1 & 0 \\ -\frac{1}{6} & \frac{1}{10} & -\frac{9}{37} & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 6 & 2 & 1 & -1 \\ 0 & \frac{10}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & \frac{37}{10} & -\frac{9}{10} \\ 0 & 0 & 0 & \frac{191}{74} \end{bmatrix}$$

Otro método comúnmente usado para factorizar es el **Método de Crout**, el cual requiere que los elementos de la diagonal principal de U sean iguales a uno, y que la matriz A sea positiva definida o estrictamente dominante diagonalmente.

TEOREMA 3.4

Si el procedimiento de eliminación gaussiana puede aplicarse al sistema Ax = b sin intercambio de renglones, entonces la matriz A puede factorizarse como producto de una matriz triangular inferior L con una matriz triangular superior U.

$$A = LU$$

VENTAJAS DE LA FORMA LU

1. - Se puede ver que la forma LU sirve para ahorrar trabajo si se tiene una matriz de coeficientes A y se modifica el vector b por un nuevo vector b', entonces ya no se tiene que volver a resolver todo el sistema.

2. - Mayor rapidez en tiempo de computadora. Supongamos que a cada división y cada multiplicación-sustracción le asignamos una sola operación. Al principio, cuando la primer ecuación tiene longitud n, lleva n operaciones por cada cero que logramos en la primera columna: una para encontrar el múltiplo l y las otras para encontrar las nuevas entradas a lo largo del renglón. Hay n-1 renglones bajo la primera, de modo que el primer paso de la eliminación necesita $n(n-1) = n^2 - n$ operaciones. (Otra forma de obtener $n^2 - n$ es la siguiente: es necesario cambiar las n^2 entradas exceptuando la primera fila). Obsérvese ahora que las etapas posteriores son más rápidas, porque las ecuaciones se vuelven progresivamente más cortas: cuando la eliminación ha llegado a k ecuaciones sólo se necesitan k^2-k operaciones para desalojar la columna debajo del pivote, empleando el mismo razonamiento aplicado a la primera etapa, cuando k era igual a n, esto es:

$$P=(1^2+\ldots+n^2)-(1+\ldots+n)=\frac{n(n+1)(2n+1)}{6}=\frac{-n(n+1)}{2}=\frac{n^3-n}{3}$$

para n suficientemente grande, $P \approx \frac{n^3}{3}$

La sustitución regresiva es más rápida. La última incógnita se encuentra en una operación (una división entre el último pivote), la penúltima requiere dos, y asi sucesivamente. La sustitución regresiva requiere un total de operaciones de:

$$Q=1+2+\ldots+n=\frac{n(n+1)}{2}\approx\frac{n^2}{2}\quad \text{para n grande}.$$

Una vez que conocemos LU podemos encontrar la solución x' correspondiente a cualquier nuevo lado derecho b', en sólo n^2 operaciones. (Usar L^{-1} ó U^{-1} sólo nos haría perder tiempo, igual si se tiene A^{-1}, aunque el producto $A^{-1}b$ también requiere n^2 operaciones). Utilizar la técnica LU tiene al menos dos ventajas: En primer lugar, para hallar U se requieren $\frac{n^3}{3}$ operaciones para n grande, y para hallar A^{-1} se necesitan n^3 operaciones; si n = 1000, la diferencia en tiempo de computadora es grande. Además lo que queremos es la solución, no todas las entradas de la inversa.

OBSERVACIONES A LA FORMA LU

1. - La forma LU es "asimétrica" en un sentido: a lo largo de su diagonal principal U contiene los pivotes, mientras que L siempre tiene unos. Esto es fácil de corregir: basta factorizar de U una matriz diagonal D constituida por los pivotes $d_1, d_2, \ldots, d_n$:

$$U=\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}\begin{bmatrix} 1 & \frac{u_{12}}{d_1} & \frac{u_{13}}{d_1} & \cdots & \frac{u_{1n}}{d_1} \\ & 1 & \frac{u_{23}}{d_2} & \cdots & \frac{u_{2n}}{d_2} \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix}$$

y se escribe A como A = LDU. L es triangular inferior con unos en la diagonal, U es triangular superior con unos en la diagonal y D es la matriz diagonal de los pivotes.

2. - No hay mucha libertad al efectuar el proceso de eliminación, ya que las operaciones en los renglones de manera aleatoria podrían fácilmente destruir en un paso los ceros generados en un paso anterior.

EJEMPLO 3.19

Se expresa A en la factorización LDU como sigue:

$$A = \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{bmatrix}} = \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix}} \underset{D}{\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{bmatrix}} \underset{U}{\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}$$

3.8 LA FORMA PA = LU

¿Qué pasa cuando el número que se va a usar como pivote es igual a cero? Por ejemplo:

$$\begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

No podemos usar ningún múltiplo de la primer ecuación para aniquilar el coeficiente 3. Entonces intercambiamos las dos ecuaciones y en este caso sencillo la matriz será triangular superior y el sistema:

$$\begin{aligned} 3x_1 + 4x_2 &= b_2 \\ 2x_2 &= b_1 \end{aligned}$$

Se puede resolver fácilmente por sustitución regresiva. Expresando esto en términos matriciales, necesitamos encontrar la matriz de permutación que produce el intercambio de renglones, esto es:

Entonces:

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

lo mismo sucede con b por lo que $PAx = Pb$.

Sí A es una matriz más grande; por ejemplo, de 4x4 y con $a_{22} = 0$

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & d & 6 \\ 0 & c & 7 & 8 \end{bmatrix}$$

Nos fijamos en la columna con pivote igual a cero; si hay algún elemento diferente de cero, se intercambia el renglón, si no, entonces la matriz es singular. En este caso se intercambian las filas 2 y 4; la matriz de permutación P_{24} que produce el intercambio es:

$$P_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Note que P_{24} es la matriz I_4 con las filas 2 y 4 intercambiadas. En general, la matriz de permutación P_{kl} es la idéntica con las filas k y l intercambiadas y $P_{kl}A$ produce el mismo intercambio de filas en A, esto es:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & d & 6 \\ 0 & c & 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & c & 7 & 8 \\ 0 & 0 & d & 6 \\ 0 & 0 & 5 & 6 \end{bmatrix}$$

Continuamos con la siguiente columna: si d = 0 hacemos P_{34} y observamos que el sistema es no singular, entonces tenemos $P_{34}P_{24}$ A y terminamos. Otros dos ejemplos serían:

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 6 & 9 \\ 2 & 8 & 8 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 2 & 4 \end{bmatrix} \Rightarrow \overset{PA}{\begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 5 \end{bmatrix}} \text{ (no singular)}$$

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 6 & 9 \\ 3 & 9 & 8 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 0 & 2 \end{bmatrix} \text{ (singular)}$$

¿Qué sucede con la factorización LU (o LDU) cuando hay intercambio de renglones? Se obtiene U triangular superior, pero ahora no sólo tenemos matrices elementales E_{ij} sino también matrices de permutación P_{kl}, cuyo producto no es triangular inferior. Así que A no puede expresarse como A = LU. Sin embargo, podemos reemplazar A por la matriz PA. Donde P es una matriz de permutación que es el producto de cada una de las permutaciones, resumiendo:

OBSERVACIÓN:

En el caso no singular existe una matriz de permutación P que reordena los renglones de A de modo que PA admite una factorización con pivotes diferentes de cero, y PA = LU (o LDU). En éste caso hay solución única a Ax = b, la cual se encuentra a través de la eliminación. En el caso singular ningún reordenamiento puede producir pivotes distintos de cero.

EJEMPLO 3.20

Dada la matriz A, factorícela a la forma LU

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

Solución:

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

Se observa que se necesitan intercambiar las filas 2 y 3; así podemos usar:

$$P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

de tal forma que:

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

de donde:

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

y se verifica que PA = LU

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

La factorización LU también se puede usar para resolver sistemas donde la matriz A está elevada a una potencia, como se muestra en el siguiente ejemplo:

EJEMPLO 3.21
Resuelva el sistema lineal $A^3x = b$.

$$A = \begin{bmatrix} 1 & 2 & -1 \\ -2 & -5 & 3 \\ -1 & -3 & 0 \end{bmatrix} \qquad b = \begin{bmatrix} 9 \\ -17 \\ -44 \end{bmatrix}$$

Solución:

$A^3x = b$	lo podemos escribir como:
$A(A^2x) = b$	sea $y = A^2x$ entonces nos queda
$Ay = b$	como $A^2x = y$ entonces $A(Ax) = y$ entonces sea $z = Ax$ y obtenemos
$Az = y$	se resuelve $Az = y$ para z y finalmente resolvemos
$Ax = z$	para la x deseada.

Como en todo el proceso siempre usamos la misma matriz de coeficientes A, es rentable hallar L y U.

$$U = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

Resolvemos para $Ay = b$ con la información conservada en L y se tiene

$$b = \begin{bmatrix} 9 \\ -17 \\ -44 \end{bmatrix} \cong \begin{bmatrix} 9 \\ 1 \\ -35 \end{bmatrix} \cong \begin{bmatrix} 9 \\ 1 \\ -36 \end{bmatrix}$$

entonces, usando U tenemos

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 1 \\ -36 \end{bmatrix}$$

y de aquí:

$$y = \begin{bmatrix} -7 \\ 17 \\ 18 \end{bmatrix}$$

Para resolver Az = y se aplica la información conservada en L a y

$$y = \begin{bmatrix} -7 \\ 17 \\ 18 \end{bmatrix} \cong \begin{bmatrix} -7 \\ 3 \\ 11 \end{bmatrix} \cong \begin{bmatrix} -7 \\ 3 \\ 8 \end{bmatrix}$$

y nuevamente usamos U

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} -7 \\ 3 \\ 8 \end{bmatrix}$$

y de aquí

$$z = \begin{bmatrix} 3 \\ -7 \\ -4 \end{bmatrix}$$

para resolver Ax = z usamos la información conservada en L para aplicarla a z y obtenemos

$$z = \begin{bmatrix} 3 \\ -7 \\ -4 \end{bmatrix} \approx \begin{bmatrix} 3 \\ -1 \\ -4 \end{bmatrix} \approx \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix} \approx \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$$

y nuevamente usamos U

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$$

y se tiene: $x = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

3.9 MATRICES ESPECIALES

Definición 3.4
Se dice que la matriz A de nxn es estrictamente dominante diagonalmente si satisface:

$$|a_{ii}| > \sum_{\substack{j=i \\ j \neq i}}^{n} |a_{ij}| \qquad \text{para toda } i = 1, \ldots\ldots, n$$

y sólo dominante diagonalmente sí

$$|a_{ii}| \geq \sum_{\substack{j=i \\ j \neq i}}^{n} |a_{ij}| \qquad \text{para toda } i = 1, \ldots\ldots, n$$

<u>EJEMPLO 3.22</u>
Sea la matriz A como sigue:

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

se dice que A es estrictamente dominante diagonalmente ya que

$$|7| > |2| + |0| \qquad |5| > |3| + |-1| \qquad |-6| > |0| + |5|$$

TEOREMA 3.5
Si A es una matriz de nxn estrictamente dominante diagonalmente, entonces A es no-singular. Además, se puede efectuar eliminación gaussiana en cualquier sistema lineal de la forma Ax = b para obtener su solución única sin intercambio de renglones o columnas.

Demostración:

Para probar que A es no singular, sea el sistema lineal $Ax = 0$, y supongamos que existe una solución $x = (x_i) \neq 0$ a este sistema, en este caso para alguna k tal que:

$$0 < |x_k| = \max_{1 \leq j \leq n} |x_j|$$

como $\sum_{j=1}^{n} a_{ij} \; x_j = 0$ para toda $i = 1,2,\ldots,n$ es decir

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n = 0$$
$$a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n = 0$$

$$a_{n1} x_1 + a_{n2} x_2 + \ldots + a_{nn} x_n = 0$$

para $i = k$

$$a_{k1} x_1 + a_{k2} x_2 + \ldots + a_{kk} x_k + \ldots + a_{kn} x_n = 0$$

despejando se tiene

$$a_{kk} \; x_k = -\sum_{\substack{j=i \\ j \neq k}}^{n} a_{kj} \; x_j$$

esto implica que

$$|a_{kk}| \; |x_k| \leq - \sum_{\substack{j=i \\ j \neq k}}^{n} |a_{kj}| \; |x_j|$$

ó

$$|a_{kk}| \leq \sum_{\substack{j=i \\ j \neq k}}^{n} |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{j=1}^{n} a_{kj}$$

Lo que contradice el hecho de que A sea estrictamente dominante diagonalmente, y esto implica que la única solución de $Ax = 0$ es $x=0$ que como ya habíamos visto en el teorema 3.3, es equivalente a la no singularidad de A.

Definición 3.5
Sea A una matriz simétrica, A es positiva definida si $x^tAx > 0$ para toda $x \neq 0$

EJEMPLO 3.23

Verifique si las siguientes matrices son positivas definidas

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

Solución:

$$x^tAx = [x_1,x_2]\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = x_1^2 + 2x_2^2 > 0$$

para toda x, por lo tanto A es positiva definida.

De igual forma se efectúa para B y se tiene:

$$x^tBx = 4x_1^2 - x_2^2$$

Si $x_1 = 1$ y $x_2 = 3$ entonces $x^tBx < 0$ por lo tanto B no es positiva definida.

Para C se tiene:

$$x^tCx = x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 > 0$$

a menos que $x_1 = x_2 = x_3 = 0$.

TEOREMA 3.6
Sí A es una matriz de nxn positiva definida, entonces A es no singular. Además, la eliminación gaussiana se puede aplicar a cualquier sistema lineal de la forma Ax = b para obtener su solución única sin intercambios de renglones o de columnas.

Demostración:

Si $x \neq 0$ es un vector que satisface Ax = 0, entonces $x^tAx = 0$ es imposible ya que A es positiva definida, lo que implica que Ax = 0 sólo tiene la solución trivial y, por lo tanto, es no singular. La otra parte de la demostración se omite.

3.10 RESUMEN

1. Un sistema lineal tiene asociada una matriz partida (o aumentada), que tiene la matriz de coeficientes del sistema a la izquierda de la partición y el vector columna de constantes a la derecha de la partición.

2. Las operaciones elementales en las filas son:
   a) Intercambio de filas.
   b) Multiplicación de una fila por una constante distinta de cero.
   c) Multiplicación de una fila por una constante distinta de cero y sumarla a otra fila.

3. En el método de Gauss, se resuelve un sistema lineal cuadrado usando operaciones elementales en filas y reduciendo la matriz partida de modo que la parte izquierda de la partición se convierta en triangular superior. Luego se halla la solución mediante sustitución regresiva.

4. El método de Gauss-Jordan es análogo al método de Gauss, excepto que la matriz de coeficientes del sistema se reduce a la matriz identidad I. Entonces, la solución aparece como el vector columna a la derecha de la partición en la matriz aumentada reducida.

5. Se pueden resolver al mismo tiempo varios sistema lineales con la misma matriz de coeficientes aumentando la matriz de coeficientes con los vectores columna de las constantes de todos los sistemas.

6. Sea A una matriz cuadrada. Una matriz cuadrada C tal que CA = AC = I es la inversa de A y se expresa $C = A^{-1}$. Si tal inversa A^{-1} de A existe, entonces se dice que A es invertible. La inversa de una matriz invertible A es única. Una matriz cuadrada que no tiene inversa se llama singular.

7. La inversa de una matriz cuadrada A existe si, y sólo si, A puede reducirse a la matriz identidad I mediante operaciones elementales en las filas, o también, si y sólo si, A es igual a un producto de matrices elementales. En este caso, A es igual al producto, tomado de izquierda a derecha, de las inversas de las matrices elementales sucesivas que corresponden a la sucesión de operaciones en las filas que reduce A a I.

8. Para hallar A^{-1}, si existe, se forma la matriz partida (A³I) y se aplica el método de Gauss-Jordan para reducir esta matriz a (I³D). Si se puede hacer esto, entonces $A^{-1} = D$. De lo contrario, A no es invertible.

9. La inversa de un producto de matrices invertibles es el producto de las inversas en orden contrario.

10. Si A es una matriz invertible de nxn que puede reducirse por filas a una matriz triangular superior U sin intercambio de filas, entonces existe una matriz triangular inferior L de nxn tal que A = LU

11. La matriz L se puede hallar como sigue: se comienza con la matriz identidad I de nxn. Si durante la reducción de A a U se multiplica la fila i por r y se suma el resultado a la fila k, reemplazar el cero de la fila k y columna i de la matriz identidad por -r. El resultado final obtenido de la matriz identidad es la matriz L.

12. Una vez que A se ha reducido a U y se ha hallado L se puede encontrar la solución de Ax = b para cualquier nuevo vector columna b.

13. Si A es como en el punto (10), entonces A tiene una factorización única de la forma A = LDU, donde:

L es triangular inferior con todas las entradas diagonales 1,
U es triangular superior con todas las entradas diagonales 1,
D es una matriz diagonal con todas las entradas en la diagonal igual a los pivotes.

14. Para cualquier matriz invertible A, existe una matriz de permutación P tal que PA se puede reducir por filas a una matriz triangular superior U y tiene las propiedades descritas para A en los puntos 10, 11, 12 y 13.

3.11 NOTAS HISTÓRICAS

El método de solución de Gauss se llama así debido a que Gauss lo describió en un artículo detallando los cálculos que hizo para deteminar la órbita del asteriode Pallas. Los parámetros de la órbita tenían que determinarse mediante observaciones del asteroide durante el periodo de seis años comprendido entre 1803 y 1809. Esto dio lugar a seis ecuaciones con seis incógnitas con coeficientes bastante complicados. Gauss mostró cómo resolver estas ecuaciones reemplazándolas sistemáticamente por un nuevo sistema en el que sólo la primer ecuación tenía seis incógnitas, la segunda tenía cinco, la tercera sólo cuatro y así sucesivamente, hasta que la sexta ecuación tenía sólo una incógnita. Evidentemente, esta última ecuación se podía resolver con facilidad, las incógnitas restantes se hallaron entonces por sustitución regresiva.

La parte Jordan del método de Gauss-Jordan es, en esencia, una técnica sistemática de "sustitución regresiva". En esta forma la describió Wilhelm Jordan (1842-1899), un profesor alemán de geodesia, en su “Manual de geodesia”.

Este trabajo se publicó por primera vez en alemán en 1873 y desde entonces ha tenido diez ediciones, además de traducciones a otras lenguas. Por otro lado, el propio Jordan atribuye el método a Friedrich Robert Helmert y Peter Andreas Hansen, unos años antes.

Wilhelm Jordan fue importante en su campo a finales del siglo XIX; participó en varias exploraciones geodésicas en Alemania, así como en la primera exploración importante en el desierto de Libia. También fue editor fundador de la revista alemana de geodesia. Su interés por hallar un método sistemático para resolver grandes sistemas de ecuaciones surge de la frecuente aparición de éstos en problemas de triangulación.

Un método de reducción matricial para resolver un sistema de ecuaciones lineales se encuentra en el antiguo trabajo chino "Nueve capítulos del arte matemático". El autor plantea la siguiente solución al sistema.

$$\begin{aligned} 3x + 2y + z &= 39 \\ 2x + 3y + z &= 34 \\ x + 2y + 3z &= 26 \end{aligned}$$

El diagrama de los coeficientes se dispone en un ábaco como sigue:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 26 & 34 & 39 \end{bmatrix}$$

Lo que nos induce a multiplicar la columna del centro por 3 y después restar a la columna de la derecha "el máximo de veces posible"; lo mismo se hace con la columna de la izquierda. Los nuevos diagramas son entonces:

$$\begin{bmatrix} 1 & 0 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 1 \\ 26 & 24 & 39 \end{bmatrix} \approx \begin{bmatrix} 0 & 0 & 3 \\ 4 & 5 & 2 \\ 8 & 1 & 1 \\ 39 & 24 & 39 \end{bmatrix}$$

La siguiente instrucción es multiplicar la columna de la izquierda por 5 y restar de nuevo la columna del centro el máximo de veces posible, lo que da:

$$\begin{bmatrix} 0 & 0 & 3 \\ 0 & 5 & 2 \\ 36 & 1 & 1 \\ 99 & 24 & 39 \end{bmatrix}$$

Así, el sistema se redujo al sistema $3x + 2y + z = 39$, $5y + z = 24$, $36z = 99$, del cual se halla fácilmente la solución completa.

La noción de la inversa de una matriz apareció por primera vez en 1855 en un escrito de Arthur Cayley (1821-1895) y se amplió en un artículo publicado tres años más tarde titulado "Una memoria sobre la teoría de matrices". En este trabajo, Cayley describe las propiedades básicas de las matrices, puntualizando que la mayoría se deriva del trabajo con conjuntos de ecuaciones lineales. En particular, la inversa procede de la idea de resolver un sistema

$$\begin{aligned} X &= ax + by + cz \\ Y &= a'x + b'y + c'z \\ Z &= a''x + b''y + c''z \end{aligned}$$

Para x, y, z en función de X, Y, Z. Cayley da una construcción explícita para la inversa en función de los determinantes de la matriz original y de los menores.

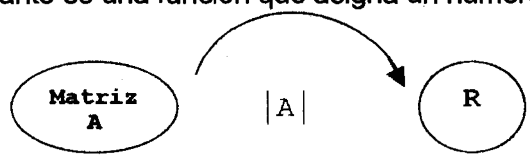
En 1842, Arthur Cayley se graduó en el Trinity College, Cambridge, pero no pudo hallar un puesto adecuado de profesor por lo que, al igual que Sylvester, estudió leyes y se hizo abogado en 1849. Durante sus catorce años de ejercicio escribió alrededor de 300 artículos matemáticos; finalmente, en 1863 se hizo profesor de Cambridge, donde permaneció hasta su muerte. Fue durante su época de abogado cuando conoció a Sylvester; sus discusiones durante los 40 años siguientes fueron muy fructíferas para el progreso del álgebra. Durante su vida, Cayley escribió alrededor de 1000 artículos sobre matemática pura, dinámica teórica y astronomía matemática.

CAPÍTULO 4
DETERMINANTES

4.1 INTRODUCCIÓN

En el capítulo dos de estos apuntes hemos visto cómo podemos representar en forma muy sencilla sistemas de ecuaciones lineales a través de matrices. Asociadas a dichas matrices existe otro concepto que es el que nos ocupa en este capítulo y es el de determinantes. Sus usos y aplicaciones son variados, como podrá verse después. Podemos adelantar sin embargo que el objetivo de los determinantes ha sido el de simplificar cálculos numéricos asociados a los sistemas de ecuaciones lineales.

Definición 4.1
Un determinante es una función que asigna un número a una matriz cuadrada.



Comenzaremos este tema con una introducción a determinantes de 2x2 y de 3x3 motivada por cálculos de área y volumen.

4.2 ÁREA DE UN PARALELOGRAMO

En la figura siguiente se muestra el paralelogramo determinado por dos vectores distintos de cero, no paralelos, $a=(a_1, a_2)$ y $b=(b_1, b_2)$ de R^2. Este paralelogramo tiene un vértice en el origen; consideramos las flechas que representan a y b como los dos lados del paralelogramo, con el origen como vértice común.

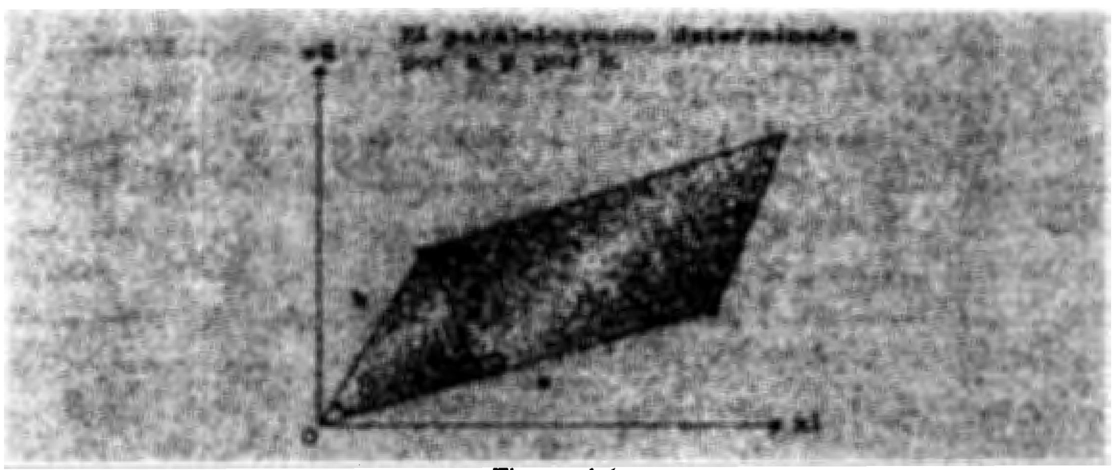


Figura 4.1

Podemos hallar el área de este paralelogramo multiplicando la longitud $\|a\|$ de su base por su altura h obteniendo:

$$\text{Área}\ (\Diamond) = \|a\| h = \|a\|\ \|b\|\ (\text{sen}\ \Theta) = \|a\|\ \|b\|\ \sqrt{[1 - \cos^2\Theta]} \qquad (4.1)$$

OBSERVACIÓN:

Recuerde que la magnitud de un vector a en R^n se define:

$$\|a\| = \sqrt{[a_1^2 + a_2^2 + \ldots + a_n^2]}$$

y el ángulo entre dos vectores a y b como:

$$\text{arcos} = \left[\frac{a\ \ b}{\|a\|\ \|b\|}\right]$$

donde se aplicó la ley de los cosenos:

$$\|a\|\ \|b\|\ \cos\Theta = a_1b_1 + \ldots + a_nb_n$$

<u>EJEMPLO 4.1</u>
Encuentre el ángulo Θ entre los vectores (1, 2, 0, 2) y (-3, 1, 1, 5) en R^4.

<u>Solución:</u>
Tenemos:

$$\cos\theta = \frac{(1,2,0,2)\cdot(-3,1,1,5)}{\sqrt{1^2+2^2+0^2+2^2}\sqrt{(-3)^2+1^2+5^2}} = \frac{9}{18} = \frac{1}{2}$$

así $\Theta = 60°$

Volviendo a la ecuación (4.1) tenemos que si usamos el resultado de que $a\, b = \|a\|\ \|b\|\ \cos\Theta$ y elevamos al cuadrado la ecuación:

$$(\text{Area})^2 = \|a\|^2\ \|b\|^2 - \|a\|^2\ \|b\|^2 \cos^2\Theta = \|a\|^2\ \|b\|^2 - (a\, b)^2$$

$$= (a_1^2 + a_2^2)(b_1^2 + b_2^2) - (a_1b_1 + a_2b_2)^2$$

$$= (a_1 b_2 - a_2 b_1)^2$$

Tomando raíces cuadradas tenemos:

$$\text{Area}\ (\Diamond) = |a_1b_2 - a_2 b_1| = \det A$$

donde A es la matriz:

$$A = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$$

EJEMPLO 4.2

Encuentre el área del paralelogramo en R^2 con vértices (1, 1), (2, 3), (2, 1), (3, 3)

Solución:

El paralelogramo se ilustra en la figura siguiente. Los lados que tienen (1,1) como vértice común se pueden considerar como los vectores a y b:

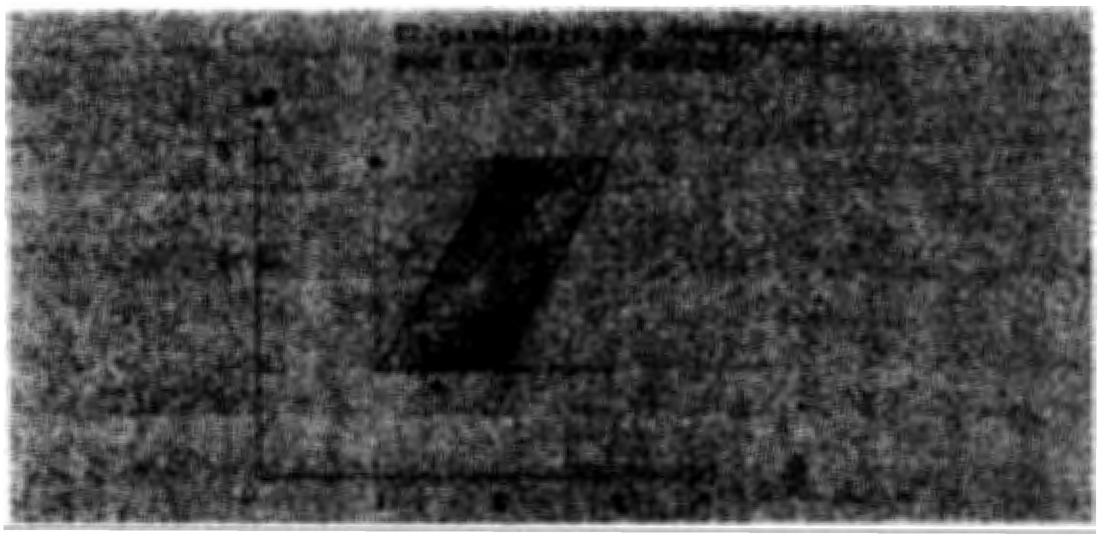


Figura 4.2 El paralelogramo determinado por a y b.

De aquí podemos encontrar que a y b están dados por:

a = (2, 1) - (1, 1) = (1, 0)
b = (2, 3) - (1, 1) = (1, 2)

Por lo tanto, el área del paralelogramo está dada por el determinante:

$$\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = (1)(2) - (0)(1) = 2$$

4.3 PRODUCTO CRUZ

El determinante asociado con una matriz cuadrada de 2x2 se denomina **determinante de segundo orden**, otra aplicación de estos determinantes aparece cuando buscamos un vector de R^3 que sea perpendicular a cada uno de los dos vectores independientes dados:

$$b = (b_1, b_2, b_3) \text{ y } c = (c_1, c_2, c_3).$$

Los vectores coordenados unitarios de R^3 son:

$$i = (1, 0, 0)\; j = (0, 1, 0) \text{ y } k = (0, 0, 1),$$

Entonces se verifica que:

$$p = \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} i - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} j + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} k \qquad (4.2)$$

es un vector ortogonal a b y c. Esto se puede observar calculando p b = p c = 0.
El vector p se conoce como producto cruz de b y c, y se denota por p = bxc. Hay una forma sencilla de recordar la fórmula (4.2) para el producto cruz bxc. Se forma la siguiente matriz de 3x3:

$$\begin{bmatrix} i & j & k \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

cuyo determinante es igual a:

$$\begin{bmatrix} i & j & k \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} = i \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - j \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + k \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

que no es otra que la fórmula (4.2).

EJEMPLO 4.3
Encuentre un vector perpendicular a (2, 1, 1) y (1, 2, 3) de R^3.

Solución:
Se forma la matriz:

$$\begin{bmatrix} i & j & k \\ 2 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$$

y se encuentra que:

$$(2,1,1) \times (1,2,3) = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} i - \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} j + \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} k$$

$= i - 5j + 3k = (1, 0, 0) - 5\,(0, 1, 0) + 3\,(0, 0, 1)$
$= (1, -5, 3)$

El producto cruz $p = bxc$ no sólo es perpendicular a b y c, sino que apunta en la dirección determinada por la conocida **regla de la mano derecha:** cuando los dedos de la mano derecha se curvan en la dirección de b a c, entonces el pulgar apunta en la dirección de bxc como en la siguiente figura:

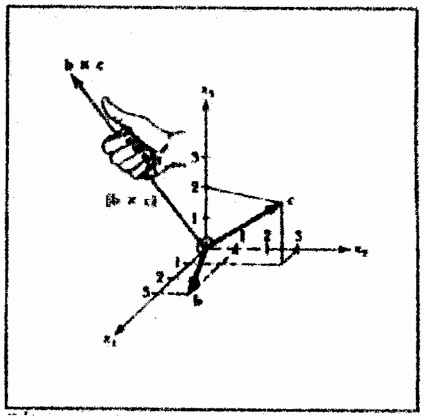


Figura 4.3

También tiene interés la magnitud del vector $p = bxc$ de (4.2): es el área del paralelogramo con un vértice en el origen de R^3 y aristas en ese vértice dadas por los vectores b y c. Para ver esto se considera un diagrama como el de la figura 4.1, pero con a reemplazado por c y se calcula nuevamente el área como sigue:

$(Area)^2 = \|c\|^2 \|b\|^2 - (c\,b)^2$

$$= (c_1^2 + c_2^2 + c_3^2)(b_1^2 + b_2^2 + b_3^2) - (c_1b_1 + c_2b_2 + c_3b_3)^2$$

$$= \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}^2 + \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}^2 + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}^2$$

Tomando raíces cuadradas obtenemos:

El Área del paralelogramo en R^3 dado por b y c = $\|b \times c\|$.

<u>EJEMPLO 4.4</u>
Encuentre el área del paralelogramo en R^3 determinado por los vectores b = (3, 1, 0) y c = (1, 3, 2).

<u>Solución:</u>
Se forma la matriz:

$$\begin{bmatrix} i & j & k \\ 3 & 1 & 0 \\ 1 & 3 & 2 \end{bmatrix}$$

De aquí:

$$b \times c = \begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix} i - \begin{vmatrix} 3 & 0 \\ 1 & 2 \end{vmatrix} j + \begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} k = (2,-6,8)$$

Por lo tanto: $\|bxc\| = 2\sqrt{[(1)^2 + (-3)^2 + (4)^2]} = 2\sqrt{26}$

<u>EJEMPLO 4.5</u>
Encuentre el área del triángulo en R^3 con vértices (-1, 2, 0), (2, 1, 3) y (1, 1, -1)

<u>Solución:</u>

Si consideramos (-1, 2, 0) como un origen local y tomamos los vectores correspondientes a flechas que comiencen ahí y lleguen a (2, 1, 3) y (1, 1, -1), a saber:

a = (2, 1, 3) - (-1, 2, 0) = (3, -1, 3)
b = (1, 1,-1) - (-1, 2, 0) = (2, -1,-1)

ahora $\|axb\|$ es el área del paralelogramo determinado por estos vectores y el área del triángulo es la mitad del área del paralelogramo:

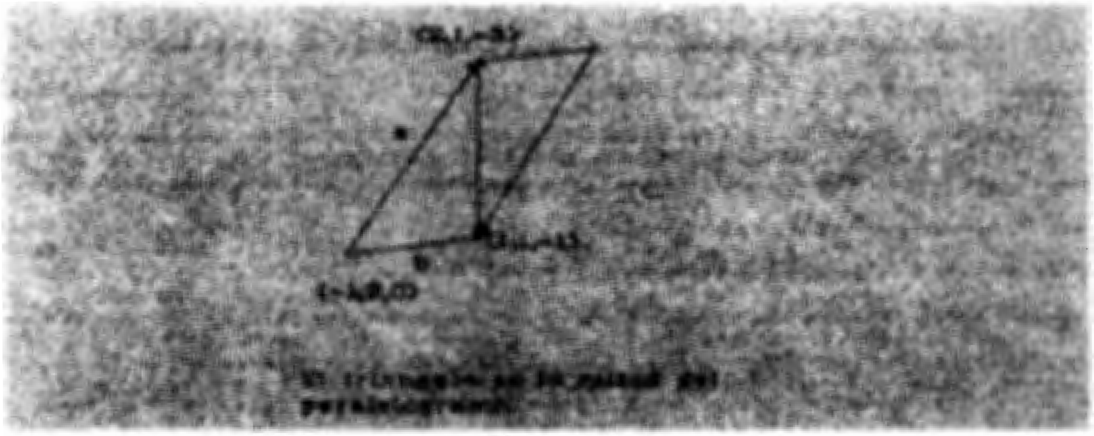


Figura 4.4

Se forma la matriz:

$$\begin{bmatrix} i & j & k \\ 3 & -1 & 3 \\ 2 & -1 & -1 \end{bmatrix}$$

De aquí:

$$a \times b = \begin{vmatrix} -1 & 3 \\ -1 & -1 \end{vmatrix} i - \begin{vmatrix} 3 & 3 \\ 2 & 1 \end{vmatrix} j + \begin{vmatrix} 3 & -1 \\ 2 & -1 \end{vmatrix} k = (4,9,-1)$$

Por lo tanto: $\|axb\| = \sqrt{[(4)^2 + (9)^2 + (-1)^2]} = \sqrt{98} = 7\sqrt{2}$ de aquí que el área del triángulo es $7\sqrt{2}/2$.

4.4 VOLUMEN DE UNA CAJA.

El producto cruz es útil para hallar el volumen de la caja o paralelepípedo determinado por tres vectores distintos de cero $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$ y $c = (c_1, c_2, c_3)$ en R^3 como se muestra en la figura siguiente:

El volumen de la caja se puede calcular multiplicando el área de la base por la altura h: el área de la base de la caja es igual a $\|bxc\|$ y la altura se puede encontrar calculando

$$h = \| a \| \, |\cos\theta| = \frac{\| bxc \| \, \| a \| \, |\cos\theta|}{\| bxc \|} = \frac{|(bxc)a|}{\| bxc \|}$$

El valor absoluto se usa en el caso en que cos Θ es negativo. Este sería el caso si bxc tuviera dirección opuesta a la que se muestra en la figura 4.5. Así:

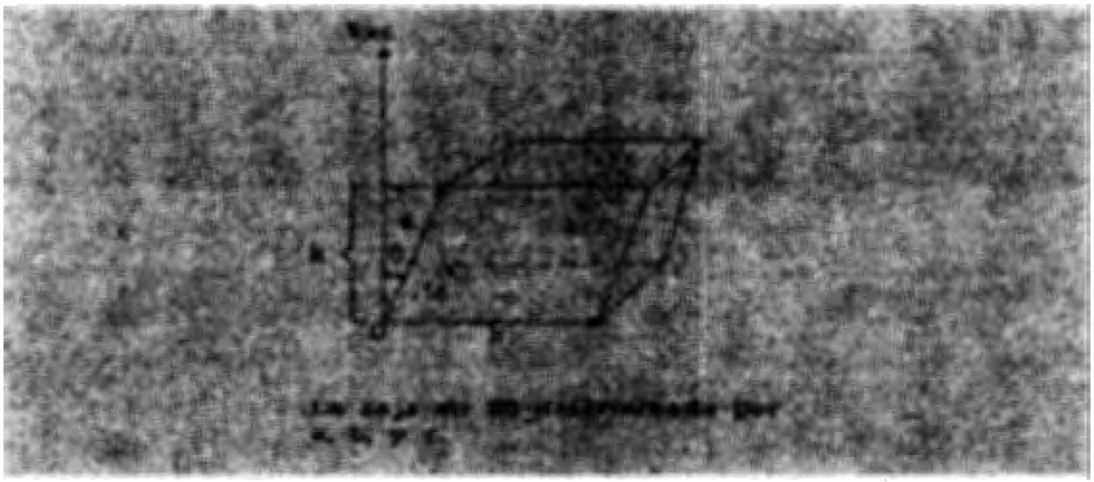


Figura 4.5

Volumen (caja) = área de la base x altura

$$= \frac{\| bxc \| \, |(bxc)a|}{bxc}$$

$$= |(bxc)\, a|$$

es decir; el volumen queda definido como:

volumen $= | a_1 (b_2 c_3 - b_3 c_2) - a_2(b_1 c_3 - b_3 c_1) + a_3(b_1 c_2 - b_2 c_1) |$

$= \det A$

Donde:

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

EJEMPLO 4.6

Encuentre el volumen de la caja con vértice en el origen, determinada por los vectores a = (4, 1, 1), b = (2, 1, 0) y c = (0, 2, 3)

Solución:

La caja se muestra en la figura 4.6, y se forma la matriz, cuyos renglones son los vectores a, b y c. Su volumen está dado por el valor absoluto del determinante:

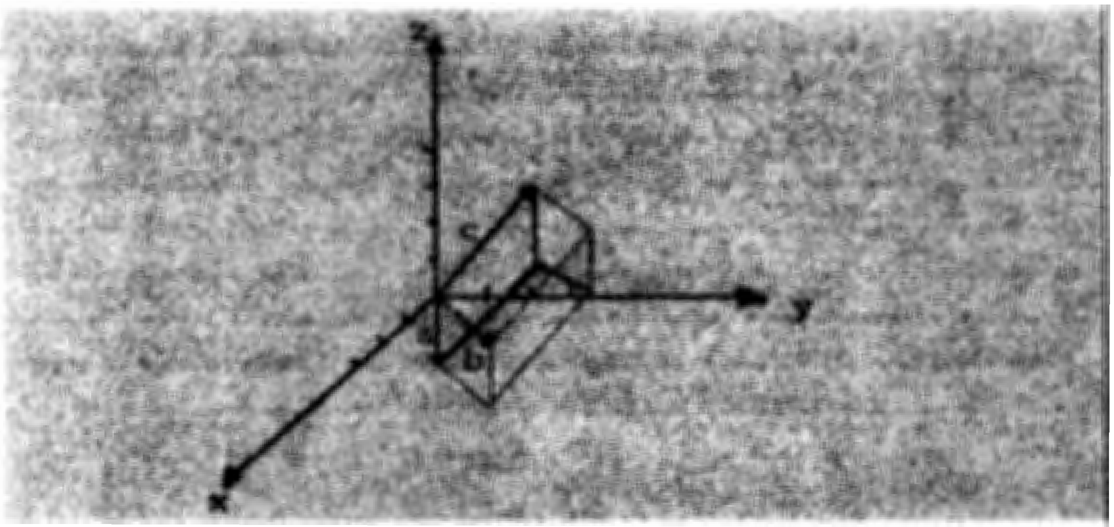


Figura 4.6

$$\begin{vmatrix} 4 & 1 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 3 \end{vmatrix} = 4\begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} - \begin{vmatrix} 2 & 0 \\ 0 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix} = 10$$

4.5 EL DETERMINANTE DE UNA MATRIZ CUADRADA.

El determinante de una matriz de 1x1 es su único registro. Así, un determinante de segundo orden se puede definir como ya se dijo:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,a_{22} - a_{21}a_{a12}$$

Y un determinante de tercer orden en términos de uno de segundo orden, como:

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

o bien:

$$\begin{array}{ccc|cc}
a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\
a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\
a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\
- & - & - & + \quad + & +
\end{array} = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Definimos un determinante de n-ésimo orden en términos de determinantes de orden n-1. Para lo cual se introduce la matriz menor A_{ij} de una matriz A de nxn, que es la matriz de (n-1)x(n-1) obtenida al eliminar la i-ésima fila y la j-ésima columna de A.

$$A_{ij} = \begin{array}{cl}
\begin{bmatrix} a_{11}\cdots & a_{1j}\cdots & a_{1n} \\ a_{i1}\cdots & a_{ij}\cdots & a_{in} \\ a_{n1}\cdots & a_{nj}\cdots & a_{nn} \end{bmatrix} & \begin{array}{l} \longrightarrow \\ \text{i– ésima fila} \\ \\ \end{array} \\
\downarrow \ \text{j-ésima columna} &
\end{array} \tag{4.3}$$

Usando $|A_{ij}|$ como notación para el determinante de la matriz menor A_{ij} podemos expresar el determinante de una matriz A de 3x3 como:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}|$$

los números $a_{11}' = |A_{11}|$, $a_{12}' = -|A_{12}|$ y $a_{13}' = |A_{13}|$ son apropiadamente llamados **cofactores** de a_{11}, a_{12} y a_{13}.

Definición 4.2

El determinante de una matriz de 1x1 es su único registro. Sea n > 1 y supongamos definidos los determinantes de orden menor que n. Sea A = (A_{ij}) una matriz de orden nxn. El cofactor de a_{ij} en A es:

$$a_{ij}' = (-1)^{i+j} \det(A_{ij})$$

Donde A_{ij} es la matriz menor de A dada en (4.3). El determinante de A es:

$$\det A = \begin{vmatrix} a_{11}\cdots\cdots\cdots & a_{1n} \\ a_{i1}\cdots\cdots\cdots & a_{in} \\ a_{n1}\cdots\cdots\cdots & a_{nn} \end{vmatrix} = a_{11}a_{11}' + a_{12}a_{12}' + \cdots\cdots + a_{1n}a_{1n}'$$

EJEMPLO 4.7

Encuentre el cofactor del registro 3 ó entrada a_{21} en la matriz:

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

Solución:

$$a_{21}' = (-1)^{2+1}\begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 2 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 4 \\ 2 & 1 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 0 & 2 \end{vmatrix} = 7$$

EJEMPLO 4.8

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

Usando la definición 4.3, encuentre el determinante de la matriz:

Solución:

$$\text{Det A}= \begin{vmatrix} 5 & -2 & 4 & -1 \\ 0 & 1 & 5 & 2 \\ 1 & 2 & 0 & 1 \\ -3 & 1 & -1 & 1 \end{vmatrix}$$

$$= 5(-1)^2\begin{vmatrix} 1 & 5 & 2 \\ 2 & 0 & 1 \\ 1 & -1 & 1 \end{vmatrix} + (-2)(-1)^3\begin{vmatrix} 0 & 5 & 2 \\ 1 & 0 & 1 \\ -3 & -1 & 1 \end{vmatrix} + 4(-1)^4\begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ -3 & 1 & 1 \end{vmatrix} + (-1)(-1)^5\begin{vmatrix} 0 & 1 & 5 \\ 1 & 2 & 0 \\ -3 & 1 & -1 \end{vmatrix}$$

Calculando los determinantes de tercer orden se tiene:

$$\begin{vmatrix} 1 & 5 & 2 \\ 2 & 0 & 1 \\ 1 & -1 & 1 \end{vmatrix} = 1\begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} - 5\begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} + 2\begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} = -8$$

$$\begin{vmatrix} 0 & 5 & 2 \\ 1 & 0 & 1 \\ -3 & -1 & 1 \end{vmatrix} = 0\begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} - 5\begin{vmatrix} 2 & 1 \\ -3 & 1 \end{vmatrix} + 2\begin{vmatrix} 1 & 0 \\ -3 & -1 \end{vmatrix} = -22$$

$$\begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ -3 & 1 & 1 \end{vmatrix} = 0\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ -3 & 1 \end{vmatrix} + 2\begin{vmatrix} 1 & 2 \\ -3 & -1 \end{vmatrix} = 10$$

$$\begin{vmatrix} 0 & 1 & 5 \\ 1 & 2 & 0 \\ -3 & 1 & -1 \end{vmatrix} = 0\begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} - 1\begin{vmatrix} 1 & 0 \\ -3 & -1 \end{vmatrix} + 5\begin{vmatrix} 1 & 2 \\ -3 & 1 \end{vmatrix} = 36$$

de donde det (A) = 5(-8) + 2(-22) + 4(10) + 1(36) = -8

TEOREMA 4.1: DESARROLLO GENERAL POR MENORES

Sean A una matriz de 2x2 y r y s cualesquiera selecciones de la lista de números 1, 2,...,n entonces

$$\det(A) = a_{r1}\,a'_{r1} + a_{r2}\,a'_{r2} + \ldots + a_{rn}\,a'_{rn} \qquad (4.4)$$

$$= a_{1s}\,a'_{1s} + a_{2s}\,a'_{2s} + \ldots + a_{ns}\,a'_{ns} \qquad (4.5)$$

La ecuación (4.4) es el desarrollo del determinante de A por menores en la r-ésima fila de A y la ecuación (4.5) es el desarrollo del determinante de A por menores en la s-ésima columna de A. Este teorema establece que se puede hallar el determinante de A desarrollando por menores en cualquier fila o columna de A.

EJEMPLO 4.9

Encuentre el determinante de la matriz:

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

Solución:

Se puede agilizar el cálculo desarrollando por menores en cada paso, en la fila o columna que contenga la mayoría de ceros:

$$\det(A) = 2(1)^{5+5}\begin{vmatrix} 3 & 2 & 0 & 1 \\ -2 & 4 & 1 & 2 \\ 0 & -1 & 0 & 1 \\ -1 & 2 & 0 & -1 \end{vmatrix}$$ desarrollando en la 5ª fila

$$= 2(1)(-1)^{2+3}\begin{vmatrix} 3 & 2 & 1 \\ 0 & -1 & 1 \\ -1 & 2 & -1 \end{vmatrix}$$ desarrollando en la 3ª columna

$$=-2\left[3(-1)^{1+1}\begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} - 1(-1)^{3+1}\begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix}\right]$$ desarrollando en la 1ª columna

$= 12$

El uso de esta definición para determinantes de 4x4 ya es un trabajo tremendo pues se involucran 4! productos, como se podrá observar en la sección 4.7, por lo cual se usan propiedades de los determinantes para calcularlos más fácilmente, ésto se verá en la siguiente sección.

PROPIEDADES DE LOS DETERMINANTES

1. Si cualquier renglón o columna de A, tiene sólo componentes cero entonces:

   det (A) = 0.

   Por ejemplo:

   $$A = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$$ entonces, $\det A = a(0) - b(0) = 0$

2. Si A tiene dos renglones (o columnas) iguales o proporcionales, entonces:

   det (A) = 0

   Por ejemplo:

   $$A = \begin{bmatrix} a & b \\ 2a & 2b \end{bmatrix}$$ entonces, $\det A = 2ab - 2ab = 0$

3. Sea B una matriz que resulta de intercambiar dos renglones (columnas) de la matriz cuadrada A, entonces:

   det B = -det A.

Por ejemplo:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

$\Rightarrow$ det A = ad – bc   det B = cb - ad = -(ad - bc)

4. Si A es una matriz cuadrada entonces:

   det A = det A^t

5. Si B es la matriz que resulta de multiplicar por α un renglón i de A entonces:

   det B = α det A

   Por ejemplo:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} \alpha a & \alpha b \\ c & d \end{bmatrix}$$

$\Rightarrow$ det A = ad - bc   det B $= \alpha ad - \alpha bc$
$= \alpha(ad - bc)$

6. Sea B = αA, donde A es una matriz de nxn entonces det B = α^n det A.

   Por ejemplo:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} \alpha a & \alpha d \\ \alpha c & \alpha d \end{bmatrix}$$

$\Rightarrow$ det A = ad - bc   det B $= (\alpha a)(\alpha d) - (\alpha b)(\alpha c)$
$= \alpha^2(ad) - \alpha^2(bc)$
$= \alpha^2(ad - bc)$

   Como A es de 2x2 entonces $\alpha^n = \alpha^2$.

7. El determinante de la matriz identidad es igual a 1.

8. Si B se obtiene de A por medio de la operación elemental de sumar a un renglón un múltiplo de otro: $E_{pq}^{(\alpha)}$, entonces:

detB = det A

Por ejemplo:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} a+\alpha c & b+\alpha d \\ c & d \end{bmatrix}$$

det A = ad – bc

det B = (a + αc)d - (b + αd) c
= ad + αcd - bc - αdc
=ad - bc

9. Para A y B de orden nxn el determinante del producto es igual al producto de los determinantes, entonces:

det (AB) = (det A)(det B)

Por ejemplo:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

det A = ad – bc

det B = eh - gf

det A det B = (ad-bc)(eh-gf) = adeh - adgf - bceh + bcgf

$$AB = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix}$$

det AB = (ae+bg)(cf+dh) - (ce+dg)(af+bh) = aedh - afdg + bgcf - bhce

En particular, si A es invertible, entonces:

det A det A^{-1} = det AA^{-1} = det I = 1 luego det A^{-1} = 1/ det A

10. Si A es singular, entonces:

det A = 0. Si A es no singular det A $\neq$ 0.

EJEMPLO 4.10

Sea la matriz:

$$A = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}$$

cuyo determinante es igual a 10. Usando propiedades del determinante calcule:

a) det (αA)
b) det (αA^{-1})
c) [det $(\alpha A)]^{-1}$
d) det $(4AA^{t})$

Solución:

a) α^4 (det A) = α^4 (10)
b) α^4 (det A^{-1}) = α^4 (1/det A) = $\alpha^4/10$
c) $(\alpha^4\ 10)^{-1}$
d) 4^4 (10)(10) = 4^4(100)

4.6 CÁLCULO DE DETERMINANTES MEDIANTE LA REDUCCIÓN A UNA MATRIZ TRIANGULAR SUPERIOR O INFERIOR

Hemos visto que el cálculo de determinantes de orden superior es una tarea irracional si se realiza directamente a partir de la definición 4.3, usando solamente desarrollos repetidos por menores; sabemos también que una matriz se puede reducir a una matriz triangular superior al efectuar operaciones elementales por filas. El siguiente teorema nos permite usar ésto para calcular determinantes de matrices de orden mayor que 3.

TEOREMA 4.2

Sea A = (A_{ij}) una matriz de orden nxn triangular superior o inferior, entonces:

$$\det A = a_{11}\, a_{22} \ldots a_{nn}$$

lo que significa que el determinante de una matriz triangular es igual al producto de sus componentes diagonales.

Demostración:

Sea U una matriz triangular superior tal que:

$$U=\begin{bmatrix} u_{11} & u_{12}\cdots & u_{1n} \\ 0 & u_{22}\cdots & u_{2n} \\ 0 & 0\ \cdots & u_{nn} \end{bmatrix}$$

Si desarrollamos cada vez en las primeras columnas, tenemos:

$$\det U = u_{11}\begin{vmatrix} u_{22} & \cdots & u_{2n} \\ 0 & u_{33}\cdots & u_{3n} \\ 0 & 0\ \cdots & u_{nn} \end{vmatrix} = u_{11}\,u_{22}\begin{vmatrix} u_{33} & \cdots & u_{3n} \\ 0 & u_{44}\cdots & u_{4n} \\ 0 & 0\ \cdots & u_{nn} \end{vmatrix} = u_{11}\,u_{22},\ldots\ldots\ldots,u_{nn}$$

Además, el efecto de las matrices elementales en los determinantes lo podemos encontrar en la siguiente proposición:

PROPOSICIÓN 4.1

Sea A una matriz de nxn y E una matriz elemental. Entonces:

det EA = det E det A, o bien det AE = det A det E.

Demostración:

Existen tres tipos de matrices elementales:

a) $E = E_{pq}$ sea $B = E_{pq}A$ entonces:

$\det E_{pq}A = \det B = -\det A = (-1)\det A = \det E_{pq}\det A$

b) $E = E_p^{(\alpha)}$ Sea $B = E_p^{(\alpha)}A$ entonces:

$\det E_p^{(\alpha)}A = \det B = \alpha \det A = \det E_p^{(\alpha)}\det A$

c) $E = E_{pq}^{(\beta)}$ se deja de ejercicio.

EJEMPLO 4.11

El determinante de la matriz A:

$$A=\begin{bmatrix} 2 & 1 & 7 \\ 0 & 2 & -5 \\ 0 & 0 & 1 \end{bmatrix} \text{ es igual a } |A| = (2)(2)(1) = 4$$

EJEMPLO 4.12

Encuentre el determinante de la matriz:

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

Solución:

$$\begin{vmatrix} 2 & 2 & 0 & 4 \\ 3 & 3 & 2 & 2 \\ 0 & 1 & 3 & 2 \\ 2 & 0 & 2 & 1 \end{vmatrix} = 2\begin{vmatrix} 1 & 1 & 0 & 2 \\ 3 & 3 & 2 & 2 \\ 0 & 1 & 3 & 2 \\ 2 & 0 & 2 & 1 \end{vmatrix}$$

Propiedad de multiplicación por un escalar.

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

El 1er renglón es multiplicado por –2.y se suma al 4º renglón.

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

Se multiplica el 1er renglón por –3 y se suma al 2º.

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

Se intercambian renglones

$$= 2\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 2 & -4 \\ 0 & 0 & 8 & 1 \end{vmatrix}$$

Se multiplica el 2º renglón y se suma al 4º.

$$= (-2)(2)\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 8 & 1 \end{vmatrix}$$

Propiedad de multiplicación por un escalar.

$$= (-2)(2)\begin{vmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 17 \end{vmatrix}$$

Se multiplica el 3er renglón por –8 y se suma al 4º.

De aquí det A = (-2)(2)(17) = -68

El ejemplo anterior ilustra que si una matriz A de orden nxn se puede reducir a una matriz triangular superior U con n pivotes (distintos de cero), entonces det (A) es igual a ± el producto de los pivotes. Si en algún punto de la reducción no se puede hallar un pivote distinto de cero, entonces el determinante de esta matriz y el determinante de A son igual a cero.

Por ejemplo:

Sea A matriz de 5x5 que se ha reducido y se tiene:

$$\begin{bmatrix} p_1 & X & X & X & X \\ 0 & p_2 & X & X & X \\ 0 & 0 & 0 & X & X \\ 0 & 0 & 0 & X & X \\ 0 & 0 & 0 & X & X \end{bmatrix}$$

Desarrollando repetidamente por menores en las primeras columnas vemos fácilmente que el determinante es cero.

TEOREMA 4.3: CRITERIO DE INVERTIBILIDAD PARA UNA MATRIZ CUADRADA

Sea A una matriz de nxn, entonces A tiene inversa si y sólo si det A $\neq$ 0.

Demostración:

$\Rightarrow$)Sí A es invertible entonces I = AA^{-1}, por lo que 1 = det (I) = det (A) det (A^{-1}); por lo tanto, det A $\neq$ 0.

$\Leftarrow$)Se tiene det A $\neq$ 0, entonces se pueden aplicar operaciones elementales en A para expresarla como una matriz equivalente U, esto es:

$$E_k \ldots\ldots E_2 E_1 A = U \quad ó \quad A = E_1^{-1} E_2^{-1} \ldots\ldots\ldots E_k^{-1} U$$

de aquí que:

$$\det A = \det (E_1^{-1}) \det (E_2^{-1}) \ldots\ldots\ldots \det (E_k^{-1}) \det U$$

como det A $\neq$ 0, entonces det (U) $\neq$0, de donde U no tiene renglones iguales a cero y por lo tanto U = I.

TEOREMA 4.4

Para una matriz A de nxn las siguientes afirmaciones son equivalentes:

a) La ecuación $Ax = 0$ tiene la única solución $x = 0$.

b) El sistema lineal $Ax = b$ tiene una solución única para cualquier vector columna b n-dimensional.

c) La matriz A es no-singular, es decir A^{-1} existe.

d) El determinante de A es distinto de cero.

e) El algoritmo de eliminación gaussiana se puede aplicar al sistema lineal $Ax = b$.

Ya hemos visto el método de desarrollo por cofactores para encontrar un determinante, ahora encontraremos una fórmula para calcular la inversa de una matriz.

Definición 4.3

Si A es una matriz de nxn y a'_{ij} es el cofactor de a_{ij} entonces la matriz:

$$A' = \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n1} & a'_{n2} & \cdots & a'_{nn} \end{bmatrix}$$

se llama la **matriz de cofactores de A**. La transpuesta de esta matriz se denomina la **adjunta de A** y se denota por adj(A).

<u>EJEMPLO 4.13</u>

Sea la matriz:

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

Los cofactores de A son:

$a'_{11} = 12$ $\quad a'_{12} = 6$ $\quad a'_{13} = -16$

$a'_{21} = -4$ $\quad a'_{22} = 2$ $\quad a'_{23} = 16$

$a'_{31} = 12$ $\quad a'_{32} = -10$ $\quad a'_{33} = 16$

Por lo que la matriz de cofactores es:

$$A' = \begin{bmatrix} 12 & 6 & -16 \\ -4 & 2 & 16 \\ 12 & -10 & 16 \end{bmatrix} \text{ y la adjunta de A es: } \operatorname{adj} A = \begin{bmatrix} 12 & -4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix}$$

Ahora es posible establecer una fórmula para calcular la inversa de una matriz invertible.

TEOREMA 4.5

Si A es una matriz invertible, entonces:

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

Demostración:

Primero se demostrará que:

$$A\ \mathrm{adj}(A) = \det(A)\ I$$

Considere el producto:

$$A\ \ \mathrm{adj}(A) = \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{11} & a'_{12} & \cdots & a'_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{11} & a'_{12} & \cdots & a'_{1n} \end{bmatrix} \begin{bmatrix} a'_{11} & a'_{21} & a'_{j1} \ldots & a'_{n1} \\ a'_{11} & a'_{21} & a'_{j1} \ldots & a'_{n1} \\ \vdots & \vdots & \vdots & \vdots \\ a'_{11} & a'_{21} & a'_{j1} \ldots & a'_{n}1 \end{bmatrix}$$

El elemento que se encuentra en el i-ésimo renglón y en la j-ésima columna de A adj(A) es:

$$a_{i1}\, a'_{j1} + a_{i2}\, a'_{j2} + \ldots + a_{in}\, a'_{jn} \qquad (4.6)$$

Si i=j entonces (4.6) es el desarrollo por cofactores de det(A) a lo largo del i-ésimo renglón de A.

Si $i \neq j$ entonces los elementos y los cofactores provienen de diferentes renglones de A, por lo que el valor de (4.6) es cero.

Note que:

$$a_{21}\, a'_{11} + a_{22}\, a'_{12} + \ldots\ldots + a_{2n}\, a'_{2n} = \begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$$

Por lo tanto:

$$A\ \mathrm{adj}\,(A) = \begin{bmatrix} \det A & & \\ & \det A & \\ & & \det A \end{bmatrix} = (\det A)\ I \qquad (4.7)$$

Como A es invertible det (A) $\neq 0$, entonces la ecuación (4.7) se puede reescribir como:

$$\frac{1}{\det A}\left[A\ \mathrm{adj}(A)\right] = I$$

entonces:

$$A = \left[\frac{1}{\det A}\right]\mathrm{adj}(A) = I$$

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

<u>EJEMPLO 4.14</u>

Encuentre la inversa de la matriz

$$A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 2 & 0 \\ 3 & 1 & 1 \end{bmatrix}$$ usando la fórmula para la inversa.

<u>Solución:</u>

Primero se verifica que el determinante de A sea diferente de cero: como el det (A) = 4, entonces A es invertible.
Para calcular la adjunta de A tenemos que los cofactores de A son:

$a'_{11} = 2$ $\quad a'_{12} = -2$ $\quad a'_{13} = -4$
$a'_{21} = 1$ $\quad a'_{22} = 1$ $\quad a'_{23} = -4$
$a'_{31} = -2$ $\quad a'_{32} = 2$ $\quad a'_{33} = 8$

Por lo que la matriz de cofactores es:

$$A' = \begin{bmatrix} 2 & -2 & -4 \\ 1 & 1 & -4 \\ -2 & 2 & 8 \end{bmatrix}$$ y la adjunta de A es: $$\mathrm{adj}\,A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 1 & -4 \\ -4 & -4 & 8 \end{bmatrix}$$

de aquí que:

$$A' = \frac{\text{adj A}}{\det A} = \begin{bmatrix} 1/2 & 1/4 & -1/2 \\ -1/2 & 1/4 & 1/4 \\ -1 & -1 & 2 \end{bmatrix}$$

Para matrices de orden mayor que 3x3 este método para invertir matrices es operacionalmente inferior a la técnica de Gauss-Jordan.

De la misma manera, con frecuencia es útil tener una fórmula en términos de determinantes para la solución de un sistema de ecuaciones $Ax = b$, donde A es una matriz invertible, por lo cual enunciamos el siguiente teorema.

TEOREMA 4.6: REGLA DE CRAMER

Si $Ax = b$ es un sistema de n ecuaciones lineales en n incógnitas tal que $\det(A) \neq 0$, entonces el sistema tiene solución única y está dada por:

$$x_k = \frac{\det B_k}{\det A} \qquad \text{para } k = 1,2,\ldots..,n \qquad (1)$$

donde B_k es la matriz que se obtiene al reemplazar los elementos de la j-ésima columna de A por los elementos del vector columna b.

<u>Demostración:</u>

Como A es invertible, la solución única del sistema $Ax = b$ se puede expresar como:

$$x_k = A^{-1}b = \frac{\det A}{\det A} b \qquad \text{comparando con (1):}$$

vemos que sólo se necesita demostrar que la k-ésima componente del vector columna adjA b está dado por el determinante de la matriz:

$$B_k = \begin{bmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{bmatrix} = 0$$

obtenida de A mediante el reemplazo de la k-ésima columna por b; esto es, el reemplazo de a_{jk} por b_j. Al desarrollar $\det B_k$ por menores en la k-ésima columna se tiene:

$$\det B_k = \sum_{j=1}^{n} a'_{jk}\, b_j$$

pero ésta es la k-ésima componente de:

$$(\operatorname{adj} A)\ b = \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{n1} \\ a'_{1k} & a'_{2k} & \cdots & a'_{nk} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{1n} & a'_{2n} & \cdots & a'_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= a'_{1k}\, b_1 + \ldots + a'_{nk}\, b_n$$

lo que completa la demostración.

EJEMPLO 4.15

Resuelva el siguiente sistema lineal usando la regla de Cramer.

$$\begin{aligned} x_1 + + 2x_3 &= 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solución:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix} \qquad \det A = 44$$

$$B_1 = \begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{bmatrix} \qquad B_2 = \begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{bmatrix} \qquad B_3 = \begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix}$$

$\det B_1 = -40$ $\qquad$ $\det B_2 = 72$ $\qquad$ $\det B_1 = 152$

Por lo tanto:

$$x_1 = \frac{\det(B_1)}{\det(A)} = \frac{-40}{44} = \frac{-10}{11}$$

$$x_2 = \frac{\det(B_2)}{\det(A)} = \frac{72}{44} = \frac{18}{11}$$

$$x_3 = \frac{\det(B_3)}{\det(A)} = \frac{152}{44} = \frac{38}{11}$$

4.7 RESUMEN

1. Un determinante de segundo orden está definido por

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

2. El área del paralelogramo con vértice en el origen determinado por los vectores distintos de cero a y b de R^2, es el valor absoluto del determinante de la matriz que tiene vectores fila a y b.

3. El producto cruz de los vectores b y c de R^3 se puede calcular usando el determinante simbólico:

$$\begin{bmatrix} i & j & k \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_{31} \end{bmatrix}$$

Este producto cruz bxc es perpendicular a: b y c.

4. El área del paralelogramo determinado por los vectores distintos de cero b y c de R^3 es $\| b \times c \|$.

5. El volumen de la caja determinada por los vectores distintos de cero a, b y c de R^3 es el valor absoluto del determinante de la matriz que tiene vectores fila a, b y c. Este determinante también es igual a: a(bxc).

6. El cofactor de un elemento a_{ij} en una matriz cuadrada A es $(-1)^{i+j} \, |A_{ij}|$, donde A_{ij} es la matriz obtenida de A al eliminar la i-ésima fila y la j-ésima columna.

7. El determinante de una matriz de nxn se puede definir inductivamente desarrollando por menores en la primera fila. El determinante se puede calcular desarrollando por menores usando cualquier fila o columna; es la suma de los productos de los registros en esa fila o columna con los cofactores de los registros. En general, dicho cálculo es muy largo.

8. Las operaciones elementales en filas tienen el efecto siguiente en el determinante de una matriz cuadrada A:

a) Si se intercambian dos filas diferentes de A; cambia el signo del determinante.
b) Si se multiplica una sola fila de A por un escalar, el determinante se multiplica por el escalar.
c) Si un múltiplo de una fila se suma a una fila diferente, el determinante no cambia.

9. La inversa de una matriz invertible A está dada por la fórmula:

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

10. Si A es invertible, entonces un sistema de ecuaciones Ax = b tiene la solución única x, cuya k-ésima componente está dada explícitamente por la fórmula.

$$x_k = \frac{\det B_k}{\det(A)}$$

donde la matriz B_k se obtiene de A reemplazando la k-ésima columna de A por b.

4.8 NOTAS HISTÓRICAS

La noción de producto cruz surgió de los intentos de Sir William Rowan Hamilton para desarrollar una multiplicación para "temas"; esto es, vectores de R^3. La simbología actual para el producto cruz apareció por primera vez en el breve texto "Elements of Vector Análisis" (1881) del norteamericano Josiah Willard Gibbs (1839- 1903), profesor de física matemática en Yale, que lo escribió para sus cursos de electricidad y magnetismo de Yale, y de mecánica en Johns Hopkins.

La interpretación como volumen de un determinante apareció por primera vez en 1773, en un artículo sobre mecánica de Joseph Louis Lagrange (1736- 1813). Lagrange observó que si los puntos M, M', M" tienen coordenadas (x, y, z), (x', y', z') y (x", y", z") respectivamente, entonces el tetraédro con vértices en el origen y esos tres puntos tendrá volumen:

$$(1/6)\,[z(x'y'' - y'x'') + z'(yx'' - xy'') + z''(xy' - yx')]$$

Esto es:

$$(1/6)\ \det\begin{bmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{bmatrix}$$

Lagrange nació en Turín y pasó la mayor parte de su carrera matemática en Berlín y Paris; aportó resultados importantes en temas tan variados como cálculo de variaciones, mecánica celeste, teoría de números y teoría de ecuaciones. Entre sus trabajos más famosos están el Tratado de Mecánica Analítica (1788), en el cual presentó los distintos principios de la mecánica desde un solo punto de vista, y la Teoría de Funciones Analíticas (1797), donde intentó basar el cálculo diferencial en la teoría de series de potencias.

La teoría de los determinantes se desarrolló a partir de los esfuerzos de muchos matemáticos de finales del siglo XVIII y principios del XIX. Además de Gabriel Cramer (1704-1752), Etienne Bezout (1739-1783), en 1764, y Alexandre- Theophile Vandermonde (1735-1796), en 1771, dieron varios métodos para calcular determinantes. En un trabajo sobre cálculo integral, Pierre Simon Laplace (1749-1827) trató sistemas de ecuaciones lineales repitiendo el trabajo de Cramer, pero también enunció y probó la regla de que el intercambio de dos columnas adyacentes del determinante cambia el signo y mostró que un determinante con dos columnas iguales es cero.

El más completo de los primeros trabajos sobre determinantes es el de Augustin - Louis Cauchy (1789-1857) en 1812. En éste trabajo, Cauchy introducjo el nombre de **"determinante"** para reemplazar varios términos antiguos, usó la notación hoy en boga del doble subíndice para un arreglo cuadrado de números, definió el arreglo de adjuntos (o menores) a un arreglo dado, y mostró que podemos calcular el determinante desarrollando en cualquier fila o columna. Además, Cauchy probó de nuevo muchos teoremas estándar sobre determinantes más o menos conocidos durante los 50 años anteriores.

Cauchy fue el matemático mas prolífico del siglo XIX, hizo aportaciones en áreas como análisis complejo, cálculo, ecuaciones diferenciales y mecánica. En particular, escribió el primer texto de cálculo usando el enfoque moderno de ε-δ para la continuidad. Políticamente fue conservador; cuando la Revolución de julio de 1830 reemplazó al rey Borbón Carlos X por el rey Orleáns Luis Felipe Cauchy se negó a prestar el juramento de fidelidad, perdiendo así sus puestos en la Ecole Polytechnique y en el College de France y se exilió en Turín y Praga.

La primera aparición del determinante de una matriz cuadrada en Europa occidental ocurrió en 1683, en una carta de Gottfried von Leibniz al marqués de L'Hopital. Leibniz escribe un sistema de tres ecuaciones con dos incógnitas con coeficientes "numéricos" abstractos como sigue:

$$\begin{gathered} 10 + 11x + 12y = 0 \\ 20 + 21x + 22y = 0 \\ 30 + 31x + 32y = 0 \end{gathered}$$

En donde se observa que cada número coeficiente tenía "dos caracteres: el primero que marca en qué ecuación se presenta y el segundo marca a qué letra pertenece". Después, procedió a eliminar y luego x, para mostrar que el criterio para que el sistema de ecuaciones tenga solución es:

$$(10)(21)(32) + (11)(22)(30) + (12)(20)(31) = (10)(22)(31) + (11)(20)(32) + (12)(21)(30).$$

Es evidente que esto equivale a nuestra condición de que el determinante de la matriz de coeficientes se debe anular. Lamentablemente, la carta no se publicó hasta 1850, por lo que no influyó en trabajos posteriores.

Los determinantes también aparecieron en el trabajo contemporáneo del matemático japonés Takakazu Seki (1642-1708).

La regla de Cramer apareció por primera vez en toda su generalidad en una obra de Gabriel Cramer (1704-1752) titulada "Introducción al análisis de curvas algebraicas" (1750). El problema en el que Cramer estaba interesado era el de determinar la ecuación de una curva plana de grado dado que pasara por cierto número de puntos dados. Formuló el teorema de que una curva de grado n-ésimo está determinada cuando se conocen (½)n (n+3) puntos de la curva. Por ejemplo, una curva de segundo grado que él escribía como:

$$A + By + Cx + Dy^2 + Exy + x^2 = 0 \qquad (*)$$

Está determinada por cinco puntos. Entonces, la cuestión es cómo determinar A, B, C, D, E dados los cinco puntos. El método obvio es sustituir sucesivamente en la ecuación (*) las coordenadas de cada uno de los cinco puntos. Esto da cinco ecuaciones para los coeficientes desconocidos. Cramer cita el apéndice del folleto, en donde da su regla general: "se halla el valor de cada incógnita formando n fracciones cuyo común denominador tiene tantos términos como permutaciones de n objetos hay" y pasa a explicar exactamente cómo se calculan estos términos como productos de ciertos coeficientes de las n ecuaciones, cómo se determina el signo apropiado para cada término y cómo se determinan los n numeradores de las fracciones al reemplazar ciertos coeficientes de este cálculo por los términos constantes del sistema.

Cramer fue un matemático suizo que enseñó en la Académie de Calvin, en Ginebra, desde 1724 hasta su muerte. Fue un gran sabio pero su trabajo se vió oscurecido por contemporáneos tan brillantes como Euler y los Bernoulli.

CAPÍTULO 5
ESPACIOS VECTORIALES

5.1 INTRODUCCIÓN

Hemos visto cómo se puede efectuar la eliminación para simplificar un sistema lineal $Ax = b$; afortunadamente, no sólo se facilita el cálculo del valor de x, sino que también se responde a las cuestiones teóricas acerca de su existencia y su unicidad. El objetivo fundamental de esta parte es adquirir un conocimiento diferente y más profundo del problema.

Para ello necesitamos el concepto de espacio vectorial, introducimos la idea con los espacios más importantes R, R^2, R^3,... El espacio R^2 se representa por el ya conocido plano x-y, donde las dos componentes del vector son las coordenadas x, y del punto correspondiente. También es familiar R^3 con las tres componentes que dan un punto en el espacio tridimensional y R es una recta. Lo importante en álgebra lineal es que la extensión a n dimensiones es directa.

En estos espacios, como en todos los espacios vectoriales, son posibles dos operaciones: la suma de dos vectores cualesquiera y la multiplicación de escalares por vectores.

5.2 ESPACIOS VECTORIALES Y ALGEBRA EN R^N

Sean u y v vectores en R^n, donde $u = (u_1, u_2,..., u_n)$ y $v = (v_1, v_2,..., v_n)$

La suma u + v es el vector

$$u + v = (u_1 + v_1, u_2 + v_2,..., u_n + v_n)$$

si r es un escalar, el múltiplo escalar ru se define

$$ru = (ru_1, ru_2,..., ru_n)$$

si u es un vector en R^n se define el negativo (o inverso aditivo) de u mediante -u y se define como:

$$- u = (-u_1, -u_2,...,-u_n)$$

la sustracción de vectores en R^n se define como v - u = v + (-u) o en términos de las componentes:

$$v - u = (v_1, v_2,...,v_n) + (-u_1,-u_2,...,-u_n) = v_1 - u_1, v_2 - u_2,..., v_n - u_n$$

Nota: Para sumar dos vectores, ambos deben estar contenidos en el mismo espacio R^n. La suma de vectores con diferente número de componentes no está definida.

Las propiedades más importantes de la suma y el producto por un escalar en R^n se enumeran en el siguiente teorema

TEOREMA 5.1

Sean u, v y w elementos de R^n y r, s escalares, entonces

a. $u + v = v + u$
b. $u + (v + w) = (u + v) + w$
c. $u + 0 = 0 + u = u$
d. $u + (-u) = 0$ es decir $u - u = 0$
e. $r(su) = (rs)u$
f. $r(u + v) = ru + rv$
g. $(r + s)u = ru + su$
h. $1u = u$

Este teorema permite manejar los vectores en R^n sin tener que expresarlos en términos de sus componentes, exactamente como se trabaja con los números reales.

Para extender las nociones de distancia, norma y ángulo a R^n se comenzará con la siguiente generalización de producto punto en R^2 y en R^3.

Definición 5.1

Si u y v son vectores en R^n, entonces el producto interior euclideano se define como:

$$u \bullet v = u_1 v_1 + u_2 v_2 + ... + u_n v_n$$

los vectores u y v se dicen ortogonales (o perpendiculares) si su producto interior es igual a cero, es decir si $u \bullet v = 0$

<u>EJEMPLO 5.1</u>

Encuentre el producto interior euclideano de los vectores u y v en R^4 donde u y v son

$$u = (-1, 3, 5, 7) \qquad v = (5, -4, 7, 0)$$

Solución:

$$u \bullet v = (-1)(5) + (3)(-4) + (5)(7) + (7)(0) = 18$$

TEOREMA 5.2
Si u, v y w son vectores en R^n y r es un escalar, entonces:

a. $u \bullet v = v \bullet u$
b. $(u + v) \bullet w = u \bullet w + v \bullet w$
c. $r(u \bullet v) = (ru) \bullet v = u \bullet (rv)$
d. $u \bullet u \geq 0$ y $u \bullet u = 0$ si y sólo si $u = 0$

Demostración del inciso d.

$u \bullet u = u_1^2 + u_2^2 + \ldots + u_n^2 \geq 0$ además la igualdad se cumple si y sólo si:
$u_1 = u_2 = \ldots = u_n$, es decir sólo si $u = 0$

Definición 5.2 **Norma y distancia en R^n**
Dados los vectores u y v en R^n se define la distancia *d(u,v)* entre los puntos u y v como:

$$d(u, v) = \sqrt{(u_1 - v_1)^2 + (u_2 + v_2)^2 + \ldots\ldots + (u_n + v_n)^2}$$

la norma o longitud del vector u denotada $|u|$ es definida como

$$|u| = u \bullet u = \sqrt{u_1^2 + u_2^2 + \ldots\ldots + u_n^2}$$

donde $u \bullet u \geq 0$ y entonces la raíz cuadrada existe. Observe que

$$d(u, v) = | u - v |$$

Es importante tener una buena apreciación del concepto de norma, dada su relación con otras ideas del álgebra lineal. Una primera relación de interés es la que se observa entre norma y distancia. Considere los puntos $P_1 = (a, b)$ y $P_2 = (c, d)$ en R^2

$$| P_1 | = \sqrt{a^2 + b^2} \quad y \quad d(P_1, P_2) = \sqrt{(a - c)^2 + (b - d)^2}$$

así $| P_1 |$ corresponde a la distancia del origen al punto P_1 y $d(P_1, P_2)$ corresponde a la distancia entre los puntos P_1 y P_2

EJEMPLO 5.2

Dados los vectores u = (1, 2, 0, 3) y v = (0, 3, 5, 2) encuentre la norma de u, la norma de v y la distancia entre ellos.

Solución:

$$|u| = \sqrt{(1)^2 + (2)^2 + (0)^2 + (3)^2} = \sqrt{14}$$

$$|v| = \sqrt{(0)^2 + (3)^2 + (5)^2 + (2)^2} = \sqrt{38}$$

$$d(u, v) = \sqrt{(1-0)^2 + (2-3)^2 + (0-5)_2 + (3-2)^2} = \sqrt{28}$$

Propiedades de la Magnitud o Norma

a) $|v|^2 = v\,v$

b) $|rv| = |r|\;|v|$

TEOREMA 5.3 DESIGUALDAD DE CAUCHY-SCHWARZ

Sean v y w vectores en R^n, entonces

$$|v \bullet w| \leq |v|\;|w|$$

Demostración

Para escalares r y s cualesquiera se tiene

$$|rv + sw|^2 \geq 0 \qquad (5.1)$$

$$|rv + sw|^2 = (rv + sw)\,(rv + sw) = r^2(vv) + 2rs(vw) + s^2(ww) \qquad (5.2)$$

haciendo r = ww y s = -(vw) se obtiene de las ecuaciones (5.1) y (5.2)

$$(ww)^2\,(vv) - 2(ww)(vw)(vw) + (vw)^2\,(ww) = (ww)^2\,(vv) - (ww)(vw)^2 \geq 0$$

ó $$(ww)\,[(vv)(ww) - (vw)^2] \geq 0$$

si (ww) = 0 entonces w = 0 y se tiene la igualdad con cero. Si (ww) $\neq$ 0 ,entonces: ww > 0 y se tiene

$$(vv)(ww) - (vw)^2 \geq 0$$
$$(vv)(ww) \geq (vw)^2$$

ó $$|v|^2\;|w|^2 \geq (vw)^2$$

tomando raíces cuadradas se tiene

$$|v|\ |w| \geq |v\,w|$$

Otra desigualdad importante que es derivada de la desigualdad de Cauchy- Schwarz es la desigualdad del triángulo: dados v y w w en R^n se tiene:

$$|v+w| \leq |v| + |w|$$

la figura 5.1 muestra el origen del nombre de desigualdad triangular

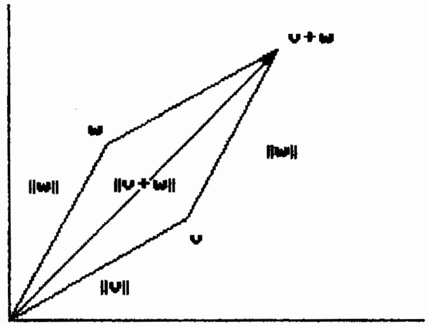


Figura 5.1

Demostración

Usando las propiedades del producto punto así como la desigualdad de Cauchy-Schwárz tenemos:

$$|v+w|^2 = (v+w)(v+w) = v \cdot v + 2v \cdot w + w \cdot w$$
$$\leq v \cdot v + 2\,|v|\,|w| + w w = |v|^2 + 2|v|\,|w| + |w|^2$$
$$= (|v| + |w|)^2$$

y tomando raíces cuadradas se tiene:

$$|v+w| \leq |v| + |w|$$

Ya se han visto algunas propiedades de los vectores en R^n, ahora procederemos a generalizarlas y definir un espacio vectorial.

Definición 5.3

Un espacio vectorial real es un conjunto no vacío V de objetos llamados vectores junto con una regla para sumar dos vectores cualesquiera v y w para producir un vector v + w $\in V$ y una regla para multiplicar cualquier vector $v \in V$ por cualquier escalar $r \in \mathbb{R}$ a fin de producir un vector $r\,v \in V$ llamado múltiplo escalar. Además, para todos los vectores u, v y w en V y para todos los escalares r y s, se deben cumplir los siguientes axiomas:

Propiedades de la suma

1. Si u y v son vectores en V entonces $u + v \in V$ (cerrado bajo la suma)
2. $u + v = v + u$ (ley conmutativa)
3. $u + (v + w) = (u + v) + w$ (ley asociativa)
1. Existe un vector $0 \in V$ llamado vector cero
   $0 + u = u + 0 = u$ para toda $u \in V$ (naturaleza del vector cero).
2. Para toda $u \in V$ existe $-u \in V$ tal que
   $u + (-u) = (-u) + u = 0$ (-u como inverso aditivo de u)

Propiedades que incluyen el producto por un escalar

6. Sí $r \in R$ y $u \in V$, entonces $r\,u \in V$ (cerrado bajo el producto por un escalar)
7. $r(u + v) = r\,u + r\,v$ (ley distributiva)
8. $(r + s)u = r\,u + su$ (ley distributiva)
9. $r(su) = (r\,s)u$ (ley asociativa)
10. $1v = v$ (preservación de escala)

EJEMPLO 5.3
El conjunto $V = R^n$ con las operaciones ordinarias de adición y multiplicación por escalares es un espacio vectorial.

EJEMPLO 5.4
Sea V cualquier plano que pasa por el origen en R^3. Los puntos de V forman un espacio vectorial bajo las operaciones ordinarias para vectores en R^3. Por el ejemplo anterior se sabe que R^3 es un espacio vectorial bajo esas operaciones. Los axiomas (2) (3) (7) (8) ((9) y (10) se cumplen para todos los puntos de R^3 y por lo tanto para todos los puntos del plano V. Entonces solo es necesario demostrar que se cumplen los axiomas (1) (4) (5) y (6).

Solución:
Dado que el plano V pasa por el origen tiene una ecuación de la forma

$$ax + by + cz = 0 \qquad (5.3)$$

por lo tanto, si $v = (v_1, v_2, v_3)$ y $w(w_1, w_2, w_3)$ son puntos en V entonces

$$av_1 + bv_2 + cv_3 = 0$$
$$aw_1 + bw_2 + cw_3 = 0$$

sumando estas ecuaciones se tiene:

$$a(v_1 + w_1) + b(v_2 + w_2) + c(v_3 + w_3) = 0$$

Esta igualdad indica que las coordenadas del punto $v + w = (v_1 + w_1, v_2 + w_2, v_3 + w_3)$ satisfacen (5.3) y por lo tanto $v + w \in V$, lo que muestra que el axioma 1 se cumple.

Si multiplicamos la ecuación $av_1 + bv_2 + cv_3 = 0$ por -1 resulta

$$a(-v_1) + b(-v_2) + c(-v_3) = 0$$

por lo tanto $-v \in V$ y se cumple el axioma 5. La demostración de los axiomas 4 y 6 se dejan como ejercicio.

EJEMPLO 5.5

Muestre que el conjunto P de todos los polinomios con coeficientes en R es un espacio vectorial, usando como suma de vectores y producto por un escalar la suma usual de polinomios y el producto usual de un polinomio por un escalar.

Solución:

Sean p y q polinomios

$$p = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$
$$q = b_0 + b_1x + b_2x^2 + \ldots + b_mx^m$$

si $m \geq n$, recordemos que la suma de p y q está dada por:

$$p + q = (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_n + b_n)x^n + b_{n+1}x_{n+1} + \ldots + b_mx^m$$

se hace una definición similar si $m < n$. El producto de p por un escalar r está dado por:

$$rp = ra_0 + ra_1x + ra_2x^2 + \ldots + ra_nx^n$$

se tiene el polinomio cero y -p y se cumplen los 8 axiomas restantes.

EJEMPLO 5.6

Demuestre que el conjunto M de todas las matrices de mxn es un espacio vectorial, usando como suma de vectores y multiplicación por un escalar la suma y producto por un escalar usual para las matrices.

Solución:

La suma de matrices de mxn y el producto de una matriz de mxn por un escalar produce de nuevo una matriz de mxn. Así M es cerrado bajo la suma de vectores y el producto por un escalar. Tomamos como vector cero en M la matriz cero usual, cuyas entradas son todas iguales a cero. Para cualquier matriz A en M consideremos -A como la matriz (-1)A.

EJEMPLO 5.7

Este ejemplo muestra que no necesariamente la suma vectorial ha de estar relacionada con la suma ordinaria, y que el vector cero no necesariamente es el número real cero.

Sea $V = \{ x \mid x \in R, x>0 \}$ la suma y producto escalar se definen como:

$$\underbrace{x \oplus y}_{\text{suma vectorial}} = \underbrace{x\,y}_{\text{producto ordinario}}$$

y para $r \in R$, $x \in V$, se define

$$r \otimes x = x^r \text{ (potencia ordinaria)}$$

Muestre que V con estas operaciones es un espacio vectorial.

Solución:

Como los elementos de V son $x \in R$ entonces se tiene

$$2 \oplus 3 = 2\,3 = 6$$
$$2 \otimes 3 = 3^2 = 9$$
$$-2 \otimes 3 = 3^{-2} = 1/9$$

Entonces tenemos que demostrar que se satisfacen los axiomas:

1. Cerradura para la suma
   Sean $x \in V$, $y \in V$ entonces $x \oplus y = x\,y$ como $x, y \in R^+$ y como el producto de dos reales positivos es otro real positivo $x\,y \in V$ entonces $x \oplus y \in V$

2. Cerradura para la multiplicación por un escalar. Sean $r \in R$ y $x \in V$ $r \otimes x = x^r$ como $x \in R^+$ al elevarlo a cualquier potencia real se obtiene un número real positivo. En consecuencia $r \otimes x \in V$.

3. Conmutatividad de la suma
   $$x \oplus y = x\,y = y\,x = y \oplus x$$
   por lo tanto $x \oplus y = y \oplus x$

4. Asociatividad
   $$(x \oplus y) \oplus z = (x\,y) \oplus z = (x\,y)z = x(y\,z) = x \oplus (y\,z) = x \oplus (y \oplus z)$$

5. Idéntico aditivo
   $x \oplus \theta = x$ por definición $x \oplus \theta = x\,\theta$ de donde $\theta = 1$, es decir el real positivo 1 es el idéntico aditivo o cero para este espacio
   $$x \oplus 1 = x$$

6. Inverso aditivo. Para toda $x \in V$, $-x$ debe satisfacer $x \oplus (-x) = \theta$ pero $-x$ no es "menos x" es el símbolo para el inverso aditivo y θ no es el cero, sino 1 entonces:

$$x \oplus (-x) = \theta$$

es equivalente a $x(-x) = 1$

de donde $-x = 1/x$ como $x > 0$ entonces $1/x > 0$ por lo tanto $-x \in V$

7. Ley asociativa para la multiplicación escalar. Sea $r \in R$, $s \in R$, $x \in V$ entonces:

$$(r\,s) \otimes x = x^{rs} \quad y \; r \otimes (s \otimes x) = r \otimes (x^s) = (x^s)^r$$

que por las leyes de los exponentes

$(x^s)^r = x^{sr} = x^{rs}$ en consecuencia

$(rs) \otimes x = r \otimes (s \otimes x)$

8 y 9. Leyes distributivas. Sean $r \in R$, $s \in R$, $x \in V$, $y \in V$ entonces

$$(r + s) \otimes x = x^{r+s} = x^r x^s = x^r \oplus x^s = (r \otimes x) \oplus (s \otimes x)$$
$$r \otimes (x \; y) = r \otimes (x\,y) = (x\,y)^r = x^r y^r = x^r + y^r = (r \otimes x) \oplus (r \otimes y)$$

10. Sea $x \in V$, entonces $\quad 1 \otimes x = x^1 = x$

5.3 SUBESPACIOS

Hemos visto que R^n es un espacio vectorial en el que la suma de vectores consiste en sumar las componentes correspondientes de tales vectores y el producto por un escalar se obtiene multiplicando cada componente por el escalar. Cualquier subconjunto de R^n que con estas mismas operaciones sea también un espacio vectorial, se llama subespacio de R^n

Por ejemplo, el plano x_1x_2 de R^3 formado por todos los vectores que tienen cero como tercera componente (como se puede ver en la figura 5.2), es un subespacio de R^3.

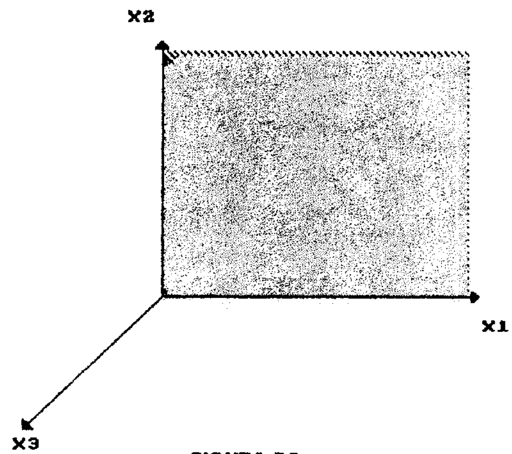


FIGURA 5.2

Por su parte, el subconjunto $\{(m, n, p) \mid m, n, p \in Z\}$ de R^3 que consta de todos los vectores de R^3 con componentes enteras no es un subespacio, pues aunque este subconjunto es cerrado bajo la suma de vectores, no es cerrado bajo el producto por un escalar, por ejemplo:

$$0.5\,(1, 2, 5) = (0.5, 1, 2.5) \text{ no está en el subconjunto}$$

Definición 5.4
Un subconjunto W de un espacio vectorial V es un subespacio de V si W es por si sólo un espacio vectorial bajo las operaciones de suma de vectores y de producto por un escalar definidas en V.

Esto significa que el conjunto W debe ser cerrado bajo la suma y el producto por un escalar. Las propiedades requeridas por un espacio vectorial se cumplen para W pues son válidas para todo V.

TEOREMA 5.4 CRITERIO PARA UN SUBESPACIO
Un subconjunto no vacío W de un espacio vectorial V es un subespacio de V si y sólo si satisface las siguientes condiciones

a) Si u y v son vectores de W entonces u + v está en W
b) Si r es cualquier escalar de R y v está en W, entonces r v está en W

Las condiciones a) y b) indican que W es cerrado bajo la suma y bajo el producto por un escalar.

Demostración

$\Rightarrow$)

Si W es un subespacio de V, entonces satisface todos los axiomas de espacio vectorial, en particular, cumple con los axiomas 1 y 6 pero estos axiomas son las condiciones a y b.

$\Leftarrow$)

Suponemos que se cumplen a y b. Dado que estas condiciones son los axiomas 1 y 6 de espacio vectorial, solamente resta demostrar que W satisface los otros 8 axiomas. Los axiomas 2, 3, 7, 8, 9, 10 se satisfacen automáticamente ya que todos los vectores en V los satisfacen. Para completar la demostración sólo falta verificar que W satisface los axiomas 4 y 5.

4) Sea $u \in W$, por la condición b, $ru \in W$ para toda $r \in R$. Haciendo $r = -1$ se tiene:

$(-1)\,u \in W$ de donde

$(-1)u + u \in W$ por a)

y esto implica que

$(-1)u + u = 0$ por lo que $0 \in W$

5) Usando 4 consideramos $r = -1$ y $(-1)u = -u \in W$

Todo espacio vectorial V tiene al menos 2 subespacios. V es un subespacio de sí mismo y el conjunto {0} que consiste únicamente del vector cero en V, este subespacio recibe el nombre de subespacio trivial. Los subespacios que no son ni {0} ni V se conocen como **subespacios propios**.

El teorema 5.3 dice que para probar que W es un subespacio de V basta verificar que se cumplen las condiciones a y b.

Algunos ejemplos de subespacios son:

EJEMPLO 5.8

Para todo espacio vectorial V, el subconjunto {0} que contiene solamente el vector cero, es un subespacio, puesto que:

$$0 + 0 = 0$$
$$r0 = 0 \text{ para toda } r \in R$$

y se le conoce como subespacio trivial.

EJEMPLO 5.9

El conjunto $M_{22}{}^{0}$ de todas las matrices de 2x2 que tiene ceros en la diagonal principal es un subespacio del espacio vectorial M_{22} de todas las matrices de orden 2x2.

Solución:
Sean A y B dos matrices en M_{22}^0 como sigue:

$$A=\begin{bmatrix}0 & a_{12}\\ a_{21} & 0\end{bmatrix} \quad B=\begin{bmatrix}0 & b_{12}\\ b_{21} & 0\end{bmatrix}$$

entonces r A y A + B son iguales a:

$$rA=\begin{bmatrix}0 & ra_{12}\\ ra_{21} & 0\end{bmatrix} \quad A+B=\begin{bmatrix}0 & a_{12}+b_{12}\\ a_{21}+b_{21} & 0\end{bmatrix}$$

Que como se ve tienen ceros en la diagonal principal, entonces pertenecen a M_{22}^0. Por lo tanto M_{22}^0 es un subespacio de M_{22a}.

EJEMPLO 5.10
Sea Ax = 0 un sistema homogéneo de m ecuaciones lineales con n incógnitas y sea W el conjunto de todos los vectores solución del sistema y sean u y v dos vectores en W. Se demostrará que W es un subespacio de R^n

Solución:
Se dice que un vector $v \in R^n$ es un vector solución del sistema si:

$x_1 = v_1, x_2 = v_2, \ldots, x_n = v_n$ es una solución del sistema.

Para mostrar que es cerrado bajo la suma y bajo el producto por un escalar, es necesario demostrar que u +v y r u son vectores solución.

Dado que u y v son vectores solución se tiene que:

$$Au = 0 \quad y \quad Av = 0$$

Y de aquí

$$A(u + v) = Au + Av = 0 + 0 = 0$$

Entonces: u + v satisface Ax = 0 lo que implica $u + v \in W$
además

$$A(ru) = r(Au) = r0 = 0$$

Se puede concluir que r u satisface Ax = 0 entonces $r\,u \in W$, por lo tanto W es un subespacio de R^n

EJEMPLO 5.11

Sea H = { (x, y) | y = 2x x,y∈R } esto es, H consiste de los vectores en R^2 que están sobre una recta que pasa por el origen. Verificaremos que H es un subespacio de R^2

Solución:

Sean

$$a = (x_1, y_1) \in H \text{ entonces } y_1 = 2x_1$$
$$b = (x_2, y_2) \in H \text{ entonces } y_2 = 2x_2$$

$$a + b = (x_1 + x_2, y_1 + y_2) = (x_1 + x_2, 2x_1 + 2x_2) = (x_1 + x_2, 2(x_1 + x_2))$$

de aquí se concluye que $a+b \in H$

$$ra = r(x_1, y_1) = (rx_1, ry_1) = (rx_1, r2x_1) = (rx_1, 2(rx_1))$$

de aquí se concluye que $ra \in H$ y por lo tanto H es un subespacio de R^2

EJEMPLO 5.12

Considere todos los vectores en R^2 cuyas componentes sean positivas o cero. Si el espacio vectorial V es R^2 con las operaciones usuales de suma y producto para vectores entonces este subconjunto es el primer cuadrante, las coordenadas satisfacen $x \geq 0$, $y \geq 0$.

Este subconjunto sin embargo no es un subespacio ya que aunque contenga al cero y sea cerrado bajo la suma, no es cerrado bajo el producto ya que si r = -1 se tiene r(x, y) = -1(x, y) = (-x, -y) que se encuentra en el tercer cuadrante. Si se incluye el tercer cuadrante además del primero, no es cerrado bajo la suma ya que:
(1, 2) + (-2, -1) = (-1, 1) que está en el segundo cuadrante.

Por lo tanto, el menor subespacio que contiene al primer cuadrante es todo el espacio R^2
Nota : No todo espacio vectorial tiene subespacios propios.

EJEMPLO 5.13

Sea W un subespacio de R. Si $W \neq \{0\}$ entonces W contiene un número real $\alpha \neq 0$ por el axioma 6 existe $1/\alpha$ tal que

$$(1/\alpha)\,\alpha = 1 \in W$$

y $\quad 1\,\alpha = \alpha \in W$ para toda $\alpha \in R$

Así si W no es el subespacio trivial entonces W = R, esto es, R no tiene subespacios propios.

Se puede pensar que cualquier recta o plano de R^3 es un subespacio, sin embargo, para que esto sea cierto, este subespacio de R^3 debe contener al cero. Geométricamente significa que esta recta o plano debe pasar por el origen.

5.4 COMBINACIONES LINEALES

Como ya se decía, un plano o una recta en R^3 es un subespacio sólo si pasa por el origen. Un plano que pasa por el origen está totalmente determinado por dos vectores cualesquiera v_1 y v_2 distintos de cero y no paralelos que están en él, como se ve en la figura 5.3

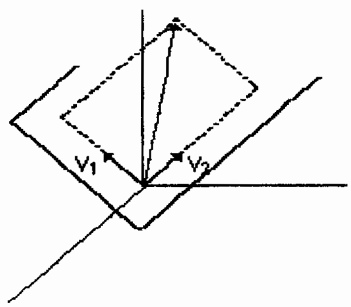


Figura 5.3

La figura muestra que cualquier vector de este plano tiene la forma

$$r_1 v_1 + r_2 v_2$$

esta expresión es una combinación lineal de los vectores v_1 y v_2

Definición 5.5

Se dice que un vector w es un combinación lineal de los vectores $v_1, v_2, \ldots, v_n$ si es posible expresarlo en la forma:

$$w = r_1 v_1 + r_2 v_2 + \ldots + r_n v_n$$

donde $r_1, r_2, \ldots, r_n$ son escalares.

EJEMPLO 5.14
Sean los vectores u = (1, 2, -1) y v = (6, 4, 2) en R^3. Demuestre que w = (9, 2, 7) es una combinación lineal de u y v.

Solución:
Para que w sea una combinación lineal de u y v, deben existir escalares r_1 y r_2 tales que

$$w = r_1\, u + r_2\, v$$

es decir

$$(9, 2, 7) = r_1\,(1, 2, -1) + r_2\,(6, 4, 2)$$

que equivale a

$$(9, 2, 7) = (r_1 + 6r_2,\ 2r_1 + 4r_2,\ -r_1 + 2r_2)$$

igualando las componentes correspondientes resulta:

$$\begin{aligned} r_1 + 6r_2 &= 9 \\ 2r_1 + 4r_2 &= 2 \\ -r_1 + 2r_2 &= 7 \end{aligned}$$

resolviendo este sistema se obtiene $r_1 = -3$ y $r_2 = 2$ de donde

$$w = -3u + 2v$$

en general definimos:

Definición 5.6
Sea { $v_1, v_2,\ldots,v_n$ } un conjunto de vectores en el espacio vectorial V. El espacio generado por { $v_1, v_2,\ldots,v_n$ } es el conjunto de todas las combinaciones posibles de $v_1, v_2,\ldots,v_n$ y se denota:

$$gen(v_1, v_2,\ldots,v_n) = \{r_1\, v_1 + r_2\, v_2 + \ldots + r_n\, v_n \mid r_i \in R \ i=1,2,..,n\}$$

EJEMPLO 5.15
Para el ejemplo del plano en R^3 tendríamos que el conjunto de todas las combinaciones lineales de v_1 y v_2 sería

$$gen(v_1, v_2) = \{ r_1\, v_1 + r_2\, v_2 \mid r_1, r_2 \in R\}$$

y geométricamente se puede ver en la figura 5.4. Además, $gen(v_1, v_2,\ldots,v_n)$ C V es un subespacio de V.

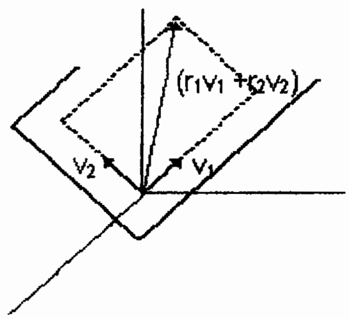


Figura 5.3

TEOREMA 5.5
Si $S = \{ v_1, v_2,...,v_n \} \subset V$ siendo V un espacio vectorial, entonces gen S es un subespacio de V

Demostración:

Sean x, y $\in$gen S, entonces x, y son combinaciones lineales

$$x = r_1 v_1 +...+ r_n v_n$$
$$y = s_1 v_1 +...+ s_n v_n$$

entonces

$$x + y = (r_1 + s_1)v_1 +...+ (r_n + s_n)v_n$$

y

$$kx = (kr_1)v_1 +...+ (kr_n)v_n$$

que también se encuentra en gen S. Por lo tanto gen S es un subespacio de V. Se le llama subespacio de V generado por S.

EJEMPLO 5.16
Determine si $v_1 = (1, 1, 2)$, $v_2 = (1, 0, 1)$ y $v_3 = (2, 1, 3)$ generan a R^3

Solución:
Es necesario determinar si un vector arbitrario $b = (b_1, b_2, b_3)$ en R^3 se puede expresar como una combinación lineal de v_1, v_2 y v_3 esto es:

$$b = r_1 v_1 + r_2 v_2 + r_3 v_3$$

en términos de sus componentes resulta:

$$(b_1, b_2, b_3) = (r_1 + r_2 + 2r_3, r_1 + r_3, 2r_1 + r_2 + 3r_3)$$

ó bien

$$\begin{aligned} r_1 + r_2 + 2r_3 &= b_1 \\ r_1 + \quad\ + r_3 &= b_2 \\ 2r_1 + r_2 + 3r_3 &= b_3 \end{aligned}$$

Entonces el problema se reduce a determinar si este sistema es consistente para todos los valores de b_1, b_2 y b_3. Esto es equivalente a que la matriz de los coeficientes sea invertible.

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

o lo que es lo mismo det A $\neq$ 0, como det A = 0, A no es invertible y por lo tanto v_1, v_2 y v_3 no generan a R^3.

<u>EJEMPLO 5.17</u>

Describa geométricamente el subespacio de R^4 generado por el vector v = (2, 3,-1, 4)

<u>Solución:</u>

El subespacio gen(v) es una recta de R^4 que contiene al origen, cada punto de esta recta es un múltiplo escalar rv de v, esto es, cada punto tiene la forma (x_1, x_2, x_3, x_4) donde

$$x_1 = 2r,\ x_2 = 3r,\ x_3 = -r \text{ y } x_4 = 4r \text{ para algún escalar } r.$$

Nota: Que un determinado vector u se pueda expresar como una combinación lineal de otro u otros vectores v_i se puede ver también, como el hecho de que ese vector u pertenezca al espacio generado por los vectores v_i.

<u>EJEMPLO 5.18</u>

¿Es posible escribir (3, -1, 4) como combinación lineal de (0, 1, 1), (1, -1, 0) y (3, -5, -2)? O lo que es lo mismo ¿está (3, -1, 4) en gen {(1, -1, 0), (0, 1, 1), (3, -5, -2)}?

<u>Solución:</u>

En cualquier caso se tiene:

$$(3, -1, 4) = r_1 (1, -1, 0) + r_2 (0, 1, 1) + r_3 (3, -5, -2)$$

o equivalentemente:

$$
\begin{aligned}
3 &= r_1 + \quad + 3r_3 \\
-1 &= -r_1 + r_2 - 5r_3 \\
4 &= \quad\quad r_2 - 2r_3
\end{aligned}
$$

escribiendo en forma matricial y efectuando eliminación gaussiana se tiene:

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

Entonces no existe solución y (3, -1, 4) no se puede escribir como combinación lineal de los vectores dados.

5.5 INDEPENDENCIA LINEAL

Hemos visto que un conjunto de vectores gen(S) genera un espacio vectorial V si S C V y cada vector en V es una combinación lineal de v_1, v_2,...,v_r. Los conjuntos de generadores son muy útiles, ya que en muchos casos es posible estudiar un espacio vectorial V si primero se trabaja con los vectores que pertenecen a algún conjunto de generadores S y después se aplican los resultados a la totalidad de V. El problema de encontrar los conjuntos más pequeños de generadores para un espacio vectorial depende del concepto de independencia lineal.

EJEMPLO 5.19

Determine si gen {(1, -1, 0), (0, 1, 1), (3, -5, -2)} genera todo R^3

Solución:

Sea (x_1, x_2, x_3) un vector cualquiera en R^3, entonces queremos saber si es posible escribir

$$(x_1, x_2, x_3) = r_1 (1, -1, 0) + r_2 (0, 1, 1) + r_3 (3, -5, -2)$$

que es equivalente a

$$A = \left[\begin{array}{ccc|c} 1 & 0 & 3 & x_1 \\ -1 & 1 & -5 & x_2 \\ 0 & 1 & -2 & x_3 \end{array}\right]$$

La cual se reduce a:

$$A = \left[\begin{array}{ccc|lll} 1 & 0 & 3 & x_1 & & \\ 0 & 1 & -2 & x_1 & +x_2 & \\ 0 & 0 & 0 & x_3 & -x_2 & -x_1 \end{array}\right]$$

Hay una solución sólo si $x_3 - x_2 - x_1 = 0$, pero esto impone una restricción a (x_1, x_2, x_3) y por lo tanto la primera de las ecuaciones no se puede resolver para un vector arbitrario (x_1, x_2, x_3), esto significa que este conjunto de vectores no genera R^3 lo que no implica que no se pueda generar otro espacio con este conjunto, entonces se puede describir gen(S) donde

$$S = \{(1, -1, 0), (0, 1, 1), (3, -5, -2)\}$$

Suponga que (x_1, x_2, x_3) está en gen(S) entonces la ecuación

$$(x_1, x_2, x_3) = r_1 (1, -1, 0) + r_2 (0, 1, 1) + r_3 (3, -5, -2)$$

debe tener solución, pero se tiene que $x_3 - x_2 - x_1 = 0$ por lo cual:
gen(S) = $\{ (x_1, x_2, x_3) \mid x_3 = x_1 + x_2\}$ Es decir, el espacio generado por S consta de todos los vectores cuya tercera componente es la suma de las dos primeras, así por ejemplo (1, 3, 5) no está en gen(S) y $(1, 3, 4) \in$ gen(S).

Por ejemplo R^2 está generado por S = { (1, 0), (0, 1)} pero también puede generarse por el conjunto mayor S' = { (1, 0), (0, 1), (1, 1)}. R^2 y funciones sobre R^2 se pueden analizar a través de los conjuntos generadores, así entonces por economía, habría que encontrar los conjuntos generadores más pequeños. Para lograr esto se requiere del concepto de **independencia lineal**.

Definición 5.7
Un conjunto $\{v_1, v_2, ..., v_n\}$ de vectores en un espacio vectorial V se llama linealmente independiente sí

$$r_1 v_1 + r_2 v_2 + ... + r_n v_n = 0$$

tiene como única solución $r_1 = r_2 = ... = r_n = 0$, es decir la solución trivial.

Definición 5.8
Un conjunto $\{v_1, v_2, ..., v_n\}$ de vectores en un espacio vectorial V se llama linealmente dependiente si

$$r_1 v_1 + r_2 v_2 + ... + r_n v_n = 0$$

tiene una solución no trivial, es decir, existen escalares $r_1, r_2, ..., r_n$ no todos cero tales que el sistema descrito tiene solución

<u>EJEMPLO 5.20</u>
En cada uno de los espacios vectoriales especificados determine si los vectores listados son linealmente independientes o dependientes

a) $V = R^2$ con $S = \{ x_1 = (2, 1), x_2 = (6, 3), x_3 = (0, 1)\}$

Geométricamente se tiene:

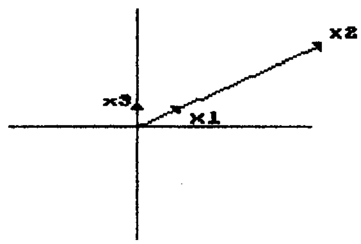


FIGURA 5.5

Note que

$$-3x_1 + x_2 + 0x_3 = 0$$

de donde x_1, x_2, x_3 son linealmente dependientes

Otra forma sería

$$r_1 x_1 + r_2 x_2 + r_3 x_3 = 0$$

ó

$$r_1\begin{bmatrix}2\\1\end{bmatrix}+r_2\begin{bmatrix}6\\3\end{bmatrix}+r_3\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Y en forma matricial

$$\begin{bmatrix}2&6&0\\1&3&1\end{bmatrix}\begin{bmatrix}r_1\\r_2\\r_3\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

de donde:

$$\begin{bmatrix}2&6&0\\1&3&1\end{bmatrix}\approx\begin{bmatrix}2&6&0\\0&0&1\end{bmatrix}$$

lo que implica $r_3 = 0$, $r_2 = t$, $r_1 = -3t$ $\quad t\in R$

Por lo tanto: x_1, x_2, x_3 son linealmente dependientes.

b) Si $V = R^3$ con

$$x_1=\begin{bmatrix}1\\1\\1\end{bmatrix}\quad x_2=\begin{bmatrix}0\\2\\1\end{bmatrix}\quad x_3=\begin{bmatrix}0\\0\\4\end{bmatrix}$$

Entonces:

$$r_1\begin{bmatrix}1\\1\\1\end{bmatrix}+r_2\begin{bmatrix}0\\2\\1\end{bmatrix}+r_3\begin{bmatrix}0\\0\\4\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

ó bien

$$A=\left[\begin{array}{ccc|c}1 & 0 & 0 & 0\\1 & 2 & 0 & 0\\1 & 1 & 4 & 0\end{array}\right]$$

Como se tiene una matriz triangular inferior es invertible y además la solución es la trivial.

$$r_1 = r_2 = r_3 = 0$$

por lo cual, x_1, x_2, y x_3 son linealmente independientes.

c) Si $V = M_{2x2}$ con

$$M_1=\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\quad M_2=\begin{bmatrix}4 & 0\\0 & 0\end{bmatrix}\quad M_3=\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}$$

entonces escribimos

$$r_1 M_1 + r_2 M_2 + r_3 M_3 = 0$$

Esto es:

$$r_1\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}+r_2\begin{bmatrix}4 & 0\\0 & 0\end{bmatrix}+r_3\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

ó bien

$$\begin{bmatrix}r_1+4r_2 & r_3\\r_3 & r_3\end{bmatrix}=\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

de aquí

$$r_1 + 4r_2 = 0$$
$$r_3 = 0$$

lo que implica que $r_1 = -4r_2$ si $r_2 = t$, entonces: $r_1 = -4t \quad t\in R$

Por lo tanto M_1, M_2, M_3 son linealmente dependientes.

EJEMPLO 5.21
Determine si el conjunto formado por los vectores $v_1 = (1, -2, 3)$, $v_2 = (5, 6, -1)$ y $v_3 = (3, 2, 1)$ es linealmente dependiente o independiente.

Solución:
En términos de los componentes, la ecuación vectorial

$$r_1 v_1 + r_2 v_2 + r_3 v_3 = 0$$

se transforma en

$$r_1 (1, -2, 3) + r_2 (5, 6, -1) + r_3 (3, 2, 1) = (0, 0, 0)$$

o el sistema:

$$\begin{aligned} r_1 + 5r_2 + 3r_3 &= 0 \\ -2r_1 + 6r_2 + 2r_3 &= 0 \\ 3r_1 - r_2 + r_3 &= 0 \end{aligned}$$

resolviendo este sistema se tiene

$$r_1 = -\tfrac{1}{2} t \qquad r_2 = -\tfrac{1}{2} t \qquad r_3 = t$$

por lo que el sistema tiene soluciones no triviales y en consecuencia v_1, v_2 y v_3 son linealmente dependientes.

Nota: El término *linealmente dependiente* sugiere que los vectores *dependen* unos de otros de alguna manera. Para ver esta dependencia, suponga que $v = \{v_1, v_2, \ldots, v_r\}$ es un conjunto linealmente dependiente por lo tanto, la ecuación

$$r_1 v_1 + r_2 v_2 + \ldots + r_n v_n = 0$$

tiene otra solución además de $r_1 = r_2 = \ldots = r_r = 0$. Supongamos que $r_1 \neq 0$ entonces multiplicando esta ecuación por $1/r_1$ y resolviendo para v_1 se obtiene:

$$v_1 = \left(-\frac{r_2}{r_1}\right) v_2 + \ldots\ldots + \left(-\frac{r_r}{r_2}\right) v_r$$

por lo que v_1 se puede expresar como una combinación lineal de los vectores restantes $v_2, \ldots, v_r$. Por lo que podemos concluir que: un conjunto de vectores es linealmente dependiente si y sólo si al menos uno de los vectores es una combinación lineal de los restantes.

EJEMPLO 5.22

Dos vectores v_1 y v_2 forman un conjunto l.d. si y sólo si uno de los vectores es un múltiplo escalar del otro. Para ver esto, suponga que S = {v_1, v_2} es linealmente dependiente. Dado que la ecuación vectorial

$$r_1 v_1 + r_2 v_2 = 0$$

tiene una solución además de la trivial $r_1 = r_2 = 0$, esta ecuación se puede expresar como:

$$v_1 = -\left(\frac{r_2}{r_1}\right)v_2 \quad \text{o} \quad v_2 = -\left(\frac{r_1}{r_1}\right)v_1$$

Estas ecuaciones indican que v_1 es múltiplo escalar de v_2 y viceversa. Por consecuencia, dos vectores en R^2 o en R^3 son l.d. si y sólo si pertenecen a la misma recta que pasa por el origen.

Para R^3 decimos que si v_1, v_2 y v_3 son tres vectores en R^3. Entonces, el conjunto S = {v_1, v_2, v_3} es l.d. si y sólo si los tres vectores pertenecen al mismo plano que pasa por el origen.

TEOREMA 5.6

Sea V = { v_1, v_2,...,v_r } un conjunto de vectores en R^n. Si $r > n$, entonces V es linealmente dependiente.

Demostración:

Suponga que

$$v_1 = (v_{11}, v_{12},\dots,v_{1n})$$
$$v_2 = (v_{21}, v_{22},\dots,v_{2n})$$
$$\dots\dots\dots\dots$$
$$v_r = (v_{r1}, v_{r2},\dots,v_{rn})$$

considere la ecuación

$$r_1 v_1 + r_2 v_2 + \dots + r_r v_r = 0$$

expresando ambos lados de la ecuación en términos de sus componentes se tiene que:

$$v_{11} r_1 + v_{21} r_2 + \dots + v_{r1} r_r = 0$$
$$v_{12} r_1 + v_{22} r_2 + \dots + v_{r2} r_r = 0$$
$$\dots\dots\dots\dots$$
$$v_{1n} r_1 + v_{2n} r_2 + \dots + v_{rn} r_r = 0$$

Este es un sistema homogéneo de n ecuaciones con r incógnitas. Dado que $r > n$ entonces por el resultado que dice que si un sistema homogéneo de ecuaciones lineales tiene más incógnitas que ecuaciones, siempre tiene un número infinito de soluciones, lo que implica que el sistema tiene soluciones no triviales y por lo tanto S es un conjunto l.d.

5.6 BASES Y DIMENSIÓN

El conjunto de vectores $S = \{(1, 1), (1, -1)\}$ genera a R^2. Es decir, cualquier vector de R^2 es una combinación lineal de (1, 1), (1,-1). El conjunto de vectores

$$T = \{ (1, 1), (1, -1), (1, 0)\}$$

también genera a R^2. Los conjuntos S y T difieren en que S es linealmente independiente, mientras que T es linealmente dependiente. Esto hace que haya una diferencia al expresar un vector como una combinación lineal de los vectores de cada conjunto. Por ejemplo, al expresar (2, 4) en términos de los vectores en S, se tiene como única posibilidad

$$(2, 4) = 3(1, 1) - 1(1, -1)$$

sin embargo, en términos de los vectores de T, hay varias posibilidades:

$$(2, 4) = 3(1, 1) - 1(1, -1) + 0(1, 0)$$
$$(2, 4) = 0(1, 1) - 4(1, -1) + 6(1, 0)$$
$$(2, 4) = 4(1, 1) + 0(1, -1) - 2(1, 0)$$

o en general

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

La conclusión es: si un conjunto S de vectores genera V y S es linealmente dependiente, entonces la representación de un vector x en términos de los vectores en S no es única. Si se desea unicidad, el conjunto generador también deberá ser linealmente independiente. Un conjunto con esas características recibe el nombre de **base** para V. Las bases se emplean en teoría de códigos entre otras áreas.

Definición 5.9

Si V es un espacio vectorial y $S = \{v_1, v_2,...,v_r\}$ es un conjunto finito de vectores en V, entonces se dice que S es una base de V si

a) S es linealmente independiente
b) S genera a V

EJEMPLO 5.23

Sean $v_1 = (1, 2)$ y $v_2 = (3, -1)$. Demuestre que el conjunto $S = \{ v_1, v_2 \}$ es una base de R^2

Solución:

Se tiene que demostrar que S es linealmente independiente y que genera a R^2

Para demostrar que S genera a R^2, se debe mostrar que un vector arbitrario $b = (b_1, b_2)$ se puede expresar como una combinación lineal de v_1 y v_2, esto es:

$$r_1 (1, 2) + r_2 (3, -1) = (b_1, b_2) \qquad (5.4)$$

tiene una solución única para cualquier $b = (b_1, b_2)$

Y para demostrar que v_1 y v_2 son linealmente independientes se debe demostrar que

$$r_1 (1, 2) + r_2 (3, -1) = (0, 0) \qquad (5.5)$$

tiene solamente la solución $r_1 = r_2 = 0$. Las ecuaciones (5.4) y (5.5) escritas en términos de sus componentes son iguales a:

$$\begin{aligned} r_1 + 3r_2 &= b_1 \\ 2r_1 - r_2 &= b_2 \end{aligned}$$

y

$$\begin{aligned} r_1 + 3r_2 &= 0 \\ 2r_1 - r_2 &= 0 \end{aligned}$$

en forma matricial se tiene:

$$\left[\begin{array}{cc|c} 1 & 3 & b_1 \\ 0 & -1 & b_2 \end{array}\right] \qquad \left[\begin{array}{cc|c} 1 & 3 & 0 \\ 0 & -1 & 0 \end{array}\right]$$

Lo que se puede resolver escribiendo

$$\left[\begin{array}{cc|c|c} 1 & 3 & b_1 & 0 \\ 0 & -1 & b_2 & 0 \end{array}\right]$$

Usando eliminación gaussiana se tiene:

La independencia lineal se comprueba con esta parte

$$\left[\begin{array}{cc|c|c} 1 & 3 & b_1 & 0 \\ 2 & -1 & \dfrac{-b_2+2b_1}{7} & 0 \end{array}\right]$$

El espacio generado se comprueba con esta parte

Entonces para demostrar que S genera a R^2 tenemos que:

$$r_2 = \frac{2b_1 - b_2}{7} \qquad r_2 = \frac{b_1 + 3b_2}{7}$$

Por lo cual dando valores arbitrarios a b_1 y b_2 se obtienen los valores para r_1 y r_2 que multiplicados por v_1 y v_2 pueden expresar como combinación lineal a b_1 y b_2.

Para la independencia lineal, se tiene que:

$$\begin{aligned} r_1 + 3r_2 &= 0 \\ r_2 &= 0 \end{aligned}$$

lo que implica $r_1 = r_2 = 0$ por lo tanto S es una base de R^2

TEOREMA 5.7

Sea $S = \{ v_1, v_2,\ldots,v_n \}$ una base para el espacio vectorial V, y sea $v \in V$. Los coeficientes en la representación

$$v = r_1 v_1 + \ldots + r_n v_n$$

son únicos.

<u>Demostración:</u>

Supóngase que se tienen las dos representaciones

$$v = a_1 v_1 + \ldots + a_n v_n$$
$$v = b_1 v_1 + \ldots + b_n v_n$$

para v; se demostrará que los coeficientes son iguales. Para ello, formamos v + (-v) que es igual a cero y se combinan los términos para obtener

$$0 = (a_1 - b_1)v_1 + \ldots + (a_n - b_n)v_n$$

Como S es una base, es un conjunto l.i. por lo tanto

$$a_1 - b_1 = a_2 - b_2 = ... = a_n - b_n = 0$$

lo que implica

$$a_1 = b_1,\ \ a_2 = b_2,...,a_n = b_n$$

<u>EJEMPLO 5.24</u>

Sean $v_1 = (1, 2, 1)$, $v_2 = (2, 9, 0)$ y $v_3 = (3, 3, 4)$. Demuestre que el conjunto $S = \{v_1, v_2, v_3\}$ es una base de R^3

<u>Solución:</u>

Entonces se tiene que demostrar que para cualquier vector $b = (b_1, b_2, b_3)$ la ecuación

$$r_1 v_1 + r_2 v_2 +...+r_3 v_3 = b$$

siempre tiene solución, y que la única solución de

$$r_1 v_1 + r_2 v_2 +...+r_3 v_3 = 0$$

es la trivial, esto es $r_1 = r_2 = r_3 = 0$

De estos dos sistemas podemos concluir, que la matriz de coeficientes asociada

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix}$$

es invertible, esto es que existe solución única para $Ar = 0$ y para $Ar = b$. Que A es invertible se puede verificar con su determinante, si det $A \neq 0$ entonces A es invertible.

$$\det A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix} = -1$$

por lo tanto, S es una base de R^3

TEOREMA 5.8
Sean $v_1, v_2, \ldots, v_n$ vectores en R^n las siguientes condiciones son equivalentes

1. Los vectores son linealmente independientes
2. Los vectores generan todo R^n
3. La matriz A que tiene estos vectores como vectores columna es invertible.

EJEMPLO 5.25
Sea el conjunto S = { M_1, M_2, M_3} tal que

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

demuestre que S es una base del espacio vectorial M_{22} de las matrices de 2x2

Solución:
Para comprobar que S genera a M_{22}, observe que, dada una matriz arbitraria

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

se puede expresar como:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = aM_1 + bM_2 + cM_3 + dM_4$$

es decir, una combinación lineal de M_1, M_2, M_3, y M_4

Para verificar que S es l.i. hacemos

$$a\,M_1 + b\,M_2 + c\,M_3 + d\,M_4 = 0$$

o bien

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

entonces

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

por lo tanto a = b = c = d = 0 y S es linealmente independiente.

EJEMPLO 5.26
Muestre que el conjunto $S = \{(1, 2), (3, -1), (1, 0)\}$ no es una base de R^2.

Solución:
El conjunto S es l.d. porque:

$$(1, 2) + 2(3, -1) - 7(1, 0) = (0, 0)$$

por lo cual S no es una base de R^2

TEOREMA 5.9 REDUCCIÓN DE UN CONJUNTO GENERADOR A UNA BASE
Si $S = \{v_1,..., v_m\}$ es un conjunto de vectores no nulos que genera a un subespacio W de un espacio vectorial V, entonces algún subconjunto de S es una base para W. Dicha base se puede encontrar excluyendo de S aquellos vectores que son combinaciones lineales de predecesores.

Demostración:

Si S es un conjunto linealmente independiente, entonces por definición S es una base para W.

Si S es linealmente dependiente, uno de los vectores puede escribirse como una combinación lineal de los otros. Supóngase que v_m es dicho vector. Se afirma que $S'= \{v_1,..., v_{m-1}\}$ aún genera a W. Para comprobar esto, sea x un elemento de W con

$$x = c_1 v_1 + ...+ c_{m-1} v_{m-1} + c_m v_m$$

pero $v_m = d_1 v_1 + ...+ d_{m-1} v_{m-1}$ por lo que esta expresión se puede sustituir en la combinación lineal anterior para obtener

$$\begin{aligned} x &= c_1 v_1 + ...+ c_{m-1} v_{m-1} + c_m (d_1 v_1 + ...+ d_{m-1} v_{m-1}) = \\ &= (c_1 + c_m d_1) v_1 +...+(c_{m-1} + c_m d_{m-1}) v_{m-1} \end{aligned}$$

por lo tanto S' genera a W. Si S' es linealmente independiente S' es una base para W. Si S' es l.d. uno de los vectores en S' es una combinación lineal de los demás y se procede a efectuar el mismo procedimiento. En esta forma debe llegarse al fin a un conjunto linealmente independiente que genere a W y que, por lo tanto, sea una base de W.

Se llega así al siguiente resultado fundamental:

Cualquier espacio vectorial finito, generado por un conjunto de vectores no nulos, tiene por lo menos una base.

EJEMPLO 5.27

Considere el subespacio W de R^3 generado por el conjunto de vectores $S = \{(2, 1, 3), (-1, 2, 0), (1, 8, 6)\}$ encuentre una base de W.

Solución:

Comenzamos por examinar el segundo vector, como (-1, 2, 0) no es múltiplo de (2, 1, 3) entonces (-1, 2, 0) se queda en la lista. Procedemos con el vector (1, 8, 6) y verificamos si puede expresarse en la forma

$$(1, 8, 6) = r_1(2, 1, 3) + r_2(-1, 2, 0)$$

para escalares r_1, r_2, este sistema de ecuaciones se puede escribir en forma matricial como

$$\left[\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 2 & 8 \\ 3 & 0 & 6 \end{array}\right]$$

para encontrar la solución en r_1 y r_2 procedemos a efectuar eliminación gaussiana:

$$\left[\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 2 & 8 \\ 3 & 0 & 6 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 2 & -1 & 1 \\ 3 & 0 & 6 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 0 & -5 & -15 \\ 0 & -6 & -18 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right]$$

Obtenemos la solución $r_2 = 3$ y $r_1 = 2$ y concluimos que (1, 8, 6) está en el conjunto $S' = \{(2, 1, 3), (-1, 2, 0)\}$ donde S' es una base de W.

Tal como se ha presentado hasta ahora, la técnica de exclusión parece potencialmente laboriosa. Por ejemplo, para aplicarla a una lista $v_1, v_2, \ldots, v_6$ parece que debemos probar si tienen soluciones para cinco ecuaciones vectoriales

$$v_2 = r_1 v_1, \qquad v_3 = r_1 v_1 + r_2 v_2$$

y así sucesivamente.

En el espacio vectorial R^n, cada una de éstas ecuaciones conducirá a un sistema lineal, según se ilustró en el ejemplo anterior. En realidad se pueden resolver cinco sistemas lineales al mismo tiempo, reduciendo una sola matriz, según se enuncia en el recuadro.

Técnicas de exclusión en R^n
La eliminación de los vectores distintos de cero que son combinación lineal de los predecesores en una lista $v_1, v_2, \ldots, v_k$ se puede lograr como sigue:
1. Formar la matriz A con vj como el j-ésimo vector columna.   2. Reducir A a la forma escalonada por filas como en el método de Gauss.   3. Los vectores de las columnas de A que dan lugar a pivotes (distintos de cero) deben permanecer y generarán todo el espacio columna de A. Esto es, se excluyen de la lista aquellos vectores de las columnas de A que no proporcionen pivotes distintos de cero.

Explicaremos la validez de esta técnica en el siguiente ejemplo.

EJEMPLO 5.28

Sea W el subespacio de R^5 generado por:

$$v_1 = (1, -1, 0, 2, 1) \quad v_2 = (2, 1, -2, 0, 0)$$
$$v_3 = (0, -3, 2, 4, 2) \quad v_4 = (3, 3, -4, -2, -1)$$
$$v_5 = (2, 4, 1, 0, 1) \quad v_6 = (5, 7, -3, -2, 0)$$

use la técnica de exclusión para intentar acortar la lista v_i con $i = 1, \ldots, 6$ de generadores de W.

Solución:

Se reduce la matriz que tiene v_j como j-ésimo vector columna y se tiene una forma escalonada.

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ -1 & 1 & -3 & 3 & 4 & 7 \\ 0 & -2 & 2 & -4 & 1 & -3 \\ 2 & 0 & 4 & -2 & 0 & -2 \\ 1 & 0 & 2 & -1 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ 0 & 3 & -3 & 6 & 6 & 12 \\ 0 & -2 & 2 & -4 & 1 & -3 \\ 0 & -4 & 4 & -8 & -4 & -12 \\ 0 & -2 & 2 & -4 & -1 & -5 \end{bmatrix}$$

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

De la última matriz vemos que el subespacio W está generado por las columnas que contienen pivotes distintos de cero y corresponden a v_1, v_2 y v_5 en la matriz original. Por lo tanto se concluye que:

$$W = \text{gen}\{v_1, v_2, v_5\}$$

Definición 5.10

Un espacio vectorial V tiene dimensión finita si hay un subconjunto finito de vectores de V, $\{v_1, v_2,...,v_n\}$ que formen una base de V (la dimensión es igual al número de vectores en una base de V). Si no existe tal subconjunto, V tiene *dimensión infinita* (además se convendrá en que el espacio trivial tiene dimensión finita, aún cuando no tenga conjuntos l.i. y por consecuencia, no tenga ninguna base)

La dimensión de un espacio vectorial V se denota por **dim V**

TEOREMA 5.10

1. Si $S = \{v_1,..., v_m\}$ es un conjunto de n vectores linealmente independientes en un espacio V de n dimensiones, entonces S es una base de V.

1. Si $S = \{v_1,..., v_m\}$ es un conjunto de n vectores que generan un espacio V de n dimensiones, entonces S es una base de V.

2. Si $S = \{v_1,..., v_m\}$ es un conjunto linealmente independiente en un espacio V de n dimensiones y $r < n$, entonces a S se le pueden agregar vectores hasta formar una base de V, es decir, existen vectores $v_{r+1},...,v_n$ tales que: $\{v_1, v_2,...,v_r, v_{r+1},...,v_n\}$ es una base de V.

Este teorema dice que si se conoce la dimensión n de un espacio vectorial V de dimensión finita, entonces los incisos 1 y 2 simplifican el problema de determinar si un conjunto S de n vectores es una base. Para mostrar que S es una base, basta con mostrar que S genera a V, o que S es linealmente independiente. Es claro que si no se conoce la dimensión de V se deben probar ambas cosas. La utilidad del inciso 3 se verá más adelante.

TEOREMA 5.11

Si $S = \{v_1, \ldots, v_n\}$ es una base de V, entonces

a) Cualquier conjunto de m vectores con m > n es linealmente dependiente, y por lo tanto no es una base para V.

1. Cualquier conjunto de m vectores con m < n no puede generar a V y por lo tanto, no es una base de V.

Este teorema simplemente dice que el número de vectores es único en una base.

Demostración:

a) Sea $T = \{w_1, w_2, \ldots, w_n, w_{n+1}\}$. Es decir, T contiene exactamente n+1 vectores. Se demostrará que T no puede ser una base probando que T es linealmente dependiente. Para ello considérese

$$c_1 w_1 + c_2 w_2 + \ldots + c_n w_n + c_{n+1} w_{n+1} = 0 \qquad (5.6)$$

pero cada w_k puede escribirse como:

$$w_k = a_{1k} v_1 + a_{2k} v_2 + \ldots + a_{nk} v_n \qquad k = 1, 2, \ldots, n+1$$

ya que S genera a V. Sustituyendo esto en la ecuación (5.6) se tiene:

$$\sum_{k=1}^{n+1} c_k w_k = \sum_{k=1}^{n+1} c_k \sum_{j=1}^{n} a_{jk} v_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{n+1} a_{jk} c_k\right) v_j = 0$$

desarrollando se tiene:

$$\begin{array}{c}
a_{11} c_1 v_1 + a_{12} c_2 v_2 + \ldots + a_{1,n+1} c_{n+1} v_n = 0 \\
a_{21} c_1 v_1 + a_{22} c_2 v_2 + \ldots + a_{2,n+1} c_{n+1} v_n = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1} c_1 v_1 + a_{n2} c_2 v_2 + \ldots + a_{n,n+1} c_{n+1} v_n = 0
\end{array}$$

Observe que este sistema homogéneo tiene menos ecuaciones que incógnitas, por lo que existe entonces solución no trivial para $c_1, c_2, \ldots, c_{n+1}$. Esto quiere decir que T es l.d. Ahora suponga que T contiene más de n+1 vectores. Sea T' un subconjunto de n+1 vectores, tal como se acaba de demostrar, debe ser linealmente dependiente. Como $T' \subset T$, el conjunto T contiene un subconjunto linealmente dependiente y es entonces un conjunto linealmente dependiente.

b) Suponga que T contiene n-1 vectores y que genera a V. Entonces por el teorema 5.8, T debe contener una base S^* para V. Si S^* contiene r vectores, entonces deberá ser $r \leq n-1$. Como S^* es una base y S tiene r+1 o más vectores, se infiere de la parte a) que S es linealmente dependiente. Lo que contradice que S sea una base.

TEOREMA 5.12

Cualesquiera dos bases de un espacio vectorial de dimensión finita tienen el mismo número de vectores

Demostración:

Sean $S = \{v_1,\ldots,v_n\}$ y $S' = \{w_1, w_2,\ldots,w_m\}$ dos bases de un espacio vectorial V de dimensión finita. Dado que S es una base y S' es un conjunto l.i. por el teorema 5.9 se afirma que $m \leq n$. De igual manera puesto que S' es una base y S es un conjunto l.i. se tiene $n \leq m$ lo que implica que $m = n$.

EJEMPLO 5.29 (BASE CANÓNICA)

Muestre que R^n tiene la base canónica $E = \{e_1, e_2,\ldots,e_n\}$ donde $e_j = (0, 0,\ldots, \underset{\text{j-ésima componente}}{1}, 0,\ldots,0)$

Solución:

Para el espacio generado considere cualquier $x = (x_1, x_2,\ldots, x_n) \in R^n$ y observe que:

$$x = x_1 (1, 0,\ldots,0) + x_2 (0, 1,\ldots,0) + \ldots + x_n (0,\ldots,0, 1) =$$
$$= x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

para el caso de la independencia lineal considere

$$c_1 e_1 + c_2 e_2 + \ldots + c_n e_n = (0, 0,\ldots,0)$$

de donde se ve que $c_1 = c_2 = \ldots = 0$. Por lo tanto, E es una base de R^n y $\dim R^n = n$

esto es razonable ya que se tiene:

R – recta	objeto unidimensional
R^2 – plano	objeto bidimensional
R^3 – espacio	objeto tridimensional

EJEMPLO 5.29
Demuestre que M_{23} tiene dimensión 6

Solución:
Una base que es la canónica de M_{23} es la siguiente

$$S = \left\{ \begin{matrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{matrix} \right.$$

por lo que dim M_{23} = número de vectores en S = 6

Ahora que ya sabemos qué es una base de un espacio vectorial, podemos plantear el segundo problema fundamental del álgebra lineal.

EL PROBLEMA DE LA BASE

Sea V un espacio vectorial. El problema de la base puede presentarse en una de las siguientes formas:

Problema 1. Construir una base para V. Tomando los vectores de V.

Problema 2. Dado un conjunto S de vectores en V, construir una base para V añadiendo o bien eliminando algunos (pero no todos) los vectores de S, o ambas cosas.

Se puede resolver esto planteando lo siguiente:

Problema 1. Si puede escogerse un conjunto de vectores de V que genere V entonces eliminando vectores dependientes se obtendrá una base para V.

Problema 2. Si el conjunto dado S genera V, se procede como en el problema 1. Si no, se aumenta S añadiendo más vectores hasta que se obtenga un conjunto generador y l.i.

EJEMPLO 5.30 (2o. PROBLEMA DE LA BASE)
Sea S = {(1, 0, 3), (2, 1, 4), encuentre una base T para R^3 que contenga a S.

Como R^3 tiene dimensión 3, entonces T debe contener 3 vectores. El conjunto S, como está, ya es l.i. por lo que sólo hace falta añadir un vector a S, de tal forma que este nuevo vector sea l.i. con los dos ya existentes.

<u>*Solución 1:* (prueba y error)</u>
Como {(1, 0, 0), (0, 1, 0), (0, 0, 1)} es una base para R^3 entonces verificamos la independencia lineal de cada uno de estos vectores con los de S.

Esto es, si existen r_1, r_2 y r_3 tales que

$$r_1\,(1, 0, 3) + r_2\,(2, 1, 4) + r_3\,(1, 0, 0) = (0, 0, 0)$$

lo que implica que

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

así los vectores son linealmente independientes y forman una base

$$T = \{(1, 0, 3), (2, 1, 4), (1, 0, 0)\}$$

<u>*Solución 2:*</u>
Otra forma de resolverlo es la siguiente:

$$\begin{aligned} \text{gen}\,\{(1, 0, 3), (2, 1, 4)\} &= \{x \mid x = a\,(1, 0, 3) + b\,(2, 1, 4)\} \\ &= \{x \mid x = (a + 2b, b, 3a + 4b)\} \end{aligned}$$

por lo tanto, hay que procurar que el nuevo vector no sea de la forma (a + 2b, b, 3a + 4b). Para ello se supone que el nuevo vector es (x_1, x_2, x_3) y se hace que la ecuación

$$(a + 2b, b, 3a + 4b) = (x_1, x_2, x_3)$$

no tenga solución para a y b. Igualando las componentes se obtienen las ecuaciones cuya representación matricial es

$$\left[\begin{array}{cc|c} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 3 & 4 & x_3 \end{array}\right]$$

lo que se reduce a

$$\left[\begin{array}{cc|llll} 1 & 2 & x_1 & & & \\ 0 & 1 & x_2 & & & \\ 3 & 4 & x_3 & -3x_1 & +2x_2 & \end{array}\right]$$

por lo tanto, si x_1, x_2 y x_3 se eligen de tal manera que $x_3 - 3x_1 + 2x_2 \neq 0$ se tendrá un vector que no está en gen{(1, 0, 3), (2, 1, 4)}, en tal caso x = (0, 1, 0) es una elección aceptable y

$$T = \{ (1, 0, 3), (2, 1, 4), (0, 1, 0)\}$$

es una base de R^3.

Solución 3:
Si el tercer vector (x_1, x_2, x_3) hiciera que $\{(1, 0, 3), (2, 1, 4), (x_1, x_2, x_3)\}$ fuera linealmente dependiente, entonces:

$$\det\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ x_1 & x_2 & x_3 \end{bmatrix} \neq 0$$

Calculando el determinante se tiene

$$x_3 + 2x_2 - 3x_1 + 0$$

que es la misma condición obtenida mediante la solución 2. Algunos terceros vectores posibles son (1, 0, 0), (0, 1, 0), (0, 1, 1) ó (1, 1, 0) en realidad hay un número infinito de posibilidades.

5.7 RESUMEN

1. Un espacio vectorial es un conjunto no vacío V de objetos llamados vectores junto con reglas para sumar dos vectores cualesquiera v y w de V y para multiplicar cualquier vector v de V y cualquier escalar r de R. Se requiere que V sea cerrado bajo esta suma de vectores y multiplicación por un escalar, de modo que v+w y rv estén en V. Es más, los axiomas siguientes se deben satisfacer para todos los vectores u, v y w de V y todos los escalares r y s de R.

2. Para todos los vectores v y w de R^n, tenemos

Desigualdad de Schwarz: $|v \cdot w| \leq \|v\| \cdot \|w\|$

Desigualdad triangular: $|v + w| \leq \|v\| + \|w\|$

3.Un subconjunto S de un espacio vectorial V es un subespacio de V si y sólo si, es no vacio y satisface las dos propiedades de cierre:

$v + w \in S$ para todos los vectores v y $w \in S$
$rv \in S$ para todos los vectores $v \in S$ y escalares $r \in R$

4.El espacio generado por $\{ v_1, v_2,...,v_n \}$ es el conjunto de todas las combinaciones posibles de $v_1, v_2,...,v_n$ y se denota

$$gen(v_1, v_2,...,v_n) = \{r_1 v_1 + r_2 v_2 +...+ r_n v_n \mid r_i \in R \ i=1,2,..,n \}$$

5.Sea V un espacio vectorial, un conjunto de vectores $\{v_1, v_2,...,v_k\}$ de V es linealmente dependiente si existe una relación de dependencia

$$r_1 v_1 + r_2 v_2 +...+ r_n v_n = 0, \text{ para algún } r_i \neq 0$$

El conjunto es linealmente independiente si no existe dicha relación de dependencia, de modo que una combinación lineal de los v_i es el vector cero sólo si todos los coeficientes son cero.

6.Una lista finita de vectores distintos de cero en V forma un conjunto linealmente independiente si, y sólo si, ningún vector de la lista se puede expresar como combinación lineal de sus predecesores.

7.Un conjunto de vectores $\{v_1, v_2,...,v_k\}$ es una base para un subespacio S de un espacio vectorial V si es linealmente independiente y constituye un conjunto generador para S.

8.Cualquier subconjunto generador finito de un subespacio S, distinto de cero, de un espacio vectorial V se puede reducir, de ser necesario, a una base para S quitando los vectores cero y excluyendo después aquellos que son combinaciones lineales de predecesores.

9.Las siguientes afirmaciones son equivalentes para n vectores de R^n

a) Los vectores forman una base para R^n
b) Los vectores son linealmente independientes
c) Los vectores generan R^n
d) Una matriz que tenga a los vectores como columnas es invertible.

10.Sea S un subespacio de dimensión finita de un espacio vectorial V. Todas las bases de S tienen el mismo número de elementos. Este número, dim(S), se llama dimensión de S.

5.8 NOTAS HISTÓRICAS

La idea de un espacio n-dimensional para n>3 se aceptó gradualmente durante el siglo XIX; por tanto, es difícil señalar una primera "invención" de este concepto. Entre los diversos usos tempranos de este concepto está su presencia en un trabajo del matemático ruso Mikhail Ostrogradskii (1801-1862) sobre el teorema de la divergencia, escrito en 1836; en los tratados geométricos de Hermann Grassmann (1809-1877) al principio de la década de 1840, y en un breve artículo de Arthur Cayley, en 1846. Lamentablemente, los dos primeros autores fueron virtualmente ignorados en vida. En particular,el trabajo de Grassman era bastante filosófico y muy difícil de leer. El artículo de Cayley simplemente decía que se podían generalizar ciertos resultados a dimensiones mayores que tres "sin recurrir a ninguna noción metafísica respecto a la posibilidad de un espacio de cuatro dimensiones". Sir William Rowan Hamilton (1805-1865), en una carta de 1841, también observó que "debe de ser posible, de una u otra manera, introducir no sólo ternas sino poliadas (polyplets), de modo que en algún sentido se satisfaga la ecuación simbólica
$a = (a_1, a_2, \ldots, a_n)$ donde a es un símbolo indicativo de un pensamiento (complejo) y $a_1, a_2, \ldots, a_n$ denotan n números reales, positivos o negativos".

Aunque los vectores de r^n fueron tratados por los matemáticos y físicos durante la segunda mitad del siglo XIX (como sucedió con otros objetos que hoy consideramos vectores), hasta la aparición del tratado de Hermann Weyl (1885-1955) *Space-Time Matter* en 1918, no se imprimió una definición axiomática y abstracta de un espacio vectorial. Weyl escribió este libro como introducción a la teoría general de la relatividad de Einstein; en el capítulo 1 analizó la naturaleza del espacio euclideano y, como parte de ese análisis, formuló lo que ahora son los axiomas comunes de un espacio vectorial. Como trató sólo los espacios V de dimensión finita, incluyó el axioma de dimensionalidad: para algún número entero h hay h vectores linealmente independientes en V, pero todo conjunto de h+1 vectores es linealmente dependiente.

Los axiomas de espacio vectorial se conocían desde hacía años, pero en general se demostraban como consecuencia de otras definiciones de vectores. Por ejemplo, aparecen en el breve texto en italiano de Giuseppe Peano (1858-1932) *Cálculo geométrico* (1888), donde se explica el trabajo de Hermann Grassmann.

La desigualdad de Schwarz se debe independientemente al trabajo de Augustin-Louis Cauchy (1789-1857) y el de Hermann Amandus Schwarz (1843-1921) y Viktor Yakovlevich Bunyakovsky (1804-1889). Primero se formuló como un teorema sobre coordenadas en un apéndice al texto de Cauchy de 1821 para su curso de análisis en la *Ecole Polytechnique* como sigue:

$$|a\alpha + a'\alpha' + a''\alpha'' + \ldots| \leq \sqrt{a^2 + a'^2 + a''^2 + \ldots}\sqrt{\alpha^2 + \alpha'^2 + \alpha''^2 + \ldots}$$

Bunyakovsky demostró en 1859 la desigualdad para funciones, esto es, formuló el resultado

$$\left[\int_a^b f(x)g(x)dx\right]^2 \leq \int_a^b f^2(x)dx\int_a^b g^2(x)dx$$

donde la integral del producto de las funciones f y g se puede considerar como el producto interior de las funciones continuas en [a,b]. Bunyakovsky fue vicepresidente de la Academia de Ciencias de San Petesburgo desde 1864 hasta su muerte. En 1875 la Academia estableció un premio matemático en su nombre, como reconocimiento a sus 50 años de enseñanza e investigación.

Schwarz formuló la desigualdad en 1884. En su caso, los vectores eran funciones Φ, Ω, de dos variables en una región T del plano, y el producto interior de estas funciones estaba dado por la doble integral de su producto, donde se supone que esta integral existe. Entonces la desigualdad establece que

$$\left|\iint_T \Phi\Omega dxdy\right| \leq \sqrt{\iint_T \Phi^2 dxdy} * \sqrt{\iint_T \Omega^2 dxdy}$$

Schwarz era el principal matemático de Berlín alrededor del cambio de siglo, el trabajo donde aparece la desigualdad está dedicado a una cuestión sobre superficies minimales.

CAPÍTULO 6
LOS CUATRO ESPACIOS FUNDAMENTALES

6.1 INTRODUCCIÓN

Cuando tenemos un sistema de n ecuaciones con n incógnitas, podemos resolverlo escribiéndolo en forma matricial y llevando su matriz de coeficientes a una matriz triangular mediante una eliminación gaussiana para finalmente proceder a resolver el sistema por sustitución regresiva.

Para el caso de sistemas con m ecuaciones y n incógnitas (o viceversa), se procede a efectuar una eliminación gaussiana, pero la matriz resultante, no es una matriz triangular superior estrictamente hablando, sino una matriz escalonada.

La forma final U es **"triangular superior"**, donde los pivotes no están en la diagonal principal y las entradas distintas de cero quedan en forma escalonada.

$$U = \begin{bmatrix} a & g & h & i \\ 0 & 0 & k & b \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Antes de estudiar todas las posibilidades de encontrar soluciones a este tipo de sistemas, conviene que estudiemos ciertos espacios asociados a las matrices; los resultados nos proporcionarán un procedimiento sencillo para elaborar bases reduciendo una matriz determinada a la forma escalonada y, en consecuencia tendremos la posibilidad de analizar soluciones a tales sistemas.

6.2 ESPACIOS ASOCIADOS A UNA MATRIZ

Definición 6.1 Considere la matriz

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

y los vectores

$$r_1 = (a_{11} \ \ a_{12} \ldots \ a_{1n})$$

$$r_2 = (a_{21} \ \ a_{22} \ldots \ a_{2n})$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$r_m = (a_{m1} \ \ a_{m2} \ldots \ a_{mn})$$

que se forman con los renglones de A, se denominan los **vectores renglón de A**. El subespacio de R^n generado por los vectores renglón se llama **espacio de los renglones de A**.

EJEMPLO 6.1

Los vectores renglón de la matriz:

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

son $r_1 = (2, 1, 0)$ y $r_2 = (3, -1, 4)$

El siguiente teorema es de gran utilidad para los cálculos que se efectuarán posteriormente.

TEOREMA 6.1

Las operaciones elementales en los renglones no alteran el espacio de los renglones de una matriz.

Supongamos que se quiere elaborar una base para el subespacio de R^n generado por un determinado conjunto de vectores $\{v_{11}\ v_2,\ldots,v_m\}$; entonces, si se forma la matriz B cuyos vectores renglón son v_1, $v_2,\ldots,v_m$, el problema se reduce a encontrar una base para el espacio de los renglones de B; por ejemplo, se puede obtener una base para el espacio generado por los vectores:

$$v_1 = (1, -2, 0, 0, 3)$$

$$v_2 = (2, -5, -3, -2, 6)$$

$$v_3 = (0, 5, 15, 10, 0)$$

$$v_4 = (2, 6, 18, 8, 6)$$

Encontrando una base para el espacio de los renglones de la matriz, entonces:

$$B=\begin{bmatrix}1 & -2 & 0 & 0 & 3\\ 2 & -5 & -3 & -2 & 6\\ 0 & 5 & 15 & 10 & 0\\ 2 & 6 & 18 & 8 & 6\end{bmatrix}$$

Por el teorema 6.1, se deduce que el espacio de los renglones de una matriz no se altera si la matriz se lleva a la forma escalonada.

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

U es la forma escalonada de B. Dado que el espacio de los renglones de B coincide con los de U, una base para el espacio de los renglones de U será también una base para el espacio de los renglones de B. Es fácil verificar que los vectores renglón de U, diferentes de cero, son linealmente independientes, por lo tanto estos vectores forman una base para el espacio de los renglones de U y en consecuencia para el espacio de los renglones de B lo que nos conduce al teorema 6.2.

$r_1 = (1, -2, 0, 0, 3)$

$r_2 = (0, 1, 3, 2, 0)$

$r_3 = (0, 0, 1, 1, 0)$

TEOREMA 6.2

Los vectores renglón diferentes de cero, en la forma escalonada de una matriz A, constituyen una base para el espacio de los renglones de A.

Nota 1: Las columnas que contienen los pivotes son linealmente independientes.

Nota 2: La dimensión del espacio de los renglones es igual al rango de la matriz.

Definición 6.2

El **espacio columna** de una matriz de orden mxn es el espacio generado por las columnas. Es un subespacio del espacio m-dimensional R^m. Si m=n entonces tanto el espacio renglón como el espacio columna son subespacios de R^n, incluso pueden ser el mismo subespacio. Al espacio columna a menudo se le llama el recorrido de A: R(A); similar a la idea de recorrido de una función, como se muestra en la Figura 6.0.

Figura 6.0

En este caso la función es f(x) = Ax, cuyo dominio consta de todas las x en R^n y el recorrido todos los posibles vectores Ax. Esto es, todas las b para las cuales puede resolverse Ax = b es igual a todas las posibles combinaciones de las columnas(el recorrido es el espacio columna).

EJEMPLO 6.2

Sea la matriz:

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

Las columnas de esta matriz son linealmente dependientes, ya que la segunda columna es tres veces la primera. Los renglones son también linealmente dependientes, el tercero es dos veces el segundo menos cinco veces el primero.

Si escalonamos la matriz A, los renglones distintos de cero deben ser linealmente independientes.

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

Además si elegimos las columnas que contienen los pivotes, también son linealmente independientes (en este caso las columnas 1 y 3 son linealmente independientes).

Las cuatro columnas generan el espacio columna por definición, sin embargo no son l.i.; por lo tanto, proponemos que las columnas que contengan los pivotes distintos de cero, sean una base del espacio columna.

Los dos primeros renglones (por el teorema 6.2) son una base para el espacio renglón. De todo esto, podemos concluir que los r renglones distintos de cero de una matriz escalonada U son l.i. y también lo son las r columnas que contienen los pivotes distintos de cero.

La forma escalonada entonces se puede describir de la siguiente manera:

1. Primero vienen los renglones distintos de cero (de haber alguno se habría intercambiado).

2. Debajo de cada pivote, hay una columna de ceros que se obtiene por eliminación.

3. Cada pivote está a la derecha del pivote del renglón anterior, esto produce la figura escalonada.

Nos interesa localizar las soluciones. Si existe $Ax = b$, donde A es una matriz de mxn, $x \in R^n$ y $b \in R^m$, las columnas de A las podemos ver como vectores de R^m, esto es:

$$A = (a_1, a_2, \ldots, a_n) \text{ con } a_i \in R^m$$

donde, Ax podemos verla como una combinación lineal de los vectores columna de A, esto es:

$$Ax = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$$

tal que $Ax = b$ tiene solución, si el vector $be R^m$ puede expresarse como combinación lineal de los vectores columna de A, de donde:

1. Si $m > n$, entonces los vectores columna son vectores de R^m y el número de vectores columna de A es menor que m, de donde el espacio generado por las columnas de A es un subespacio de R^m (de dimensión menor o igual a n)

Por ejemplo:

$$3x_1 + 2x_2 = 3$$

$$5x_1 + 6x_2 = 1$$

$$3x_1 - 4x_2 = 5$$

Tenemos entonces que (3, 5, 3) y (2, 6, 4) son vectores en R^3, donde no pueden generar a R^3 pero si un subespacio de R^3. Si son linealmente independientes pueden generar un plano en R^3 y si no, una recta.

Tenemos que considerar además que:

a) El sistema no tenga solución; lo que sucede si b no se encuentra en el subespacio generado por las columnas de A.

b) Que tenga solución única; si los vectores columna de A son l.i. y formen por lo tanto una base del espacio generado por las columnas de A.

c) Que tenga una infinidad de soluciones, esto ocurre si la dimensión del espacio generado por las columnas es menor que n, por lo cual, los vectores columna son linealmente dependientes y conforman un conjunto generador (pero no base, ya que generan un subespacio de dimensión menor que n) de donde, si b está en el espacio generado por las columnas de A, puede representarse en una infinidad de formas de combinaciones lineales de las columnas de A.

2. Si m<n los vectores columna de A no pueden ser una base para R^m, ya que son más de los que se necesitan y para R^n solo pueden ser base para un subespacio.

EJEMPLO 6.3

Comenzamos con el caso homogéneo Ax = 0, donde A es la matriz:

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ 1 & -3 & 3 & 0 \end{bmatrix}$$

que se puede reducir a Ux = 0 de la siguiente manera:

$$Ux = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Las incógnitas x, y, z, w van en dos grupos. Un grupo consta de las variables básicas, aquellas que corresponden a las columnas con pivotes distintos de cero (en este caso la 1a. y la 3a.) así que x, z son **variables básicas**. El otro grupo consta de las **variables libres**, que corresponden a las columnas sin pivotes, estas son la 2a y la 4a. columna, entonces, y, w son las variables libres.

Para encontrar la solución general a Ux = 0 (ó Ax = 0) asignamos valores arbitrarios a las variables libres; esto es, y, w, entonces:

$$3z + w = 0 \rightarrow z = -1/3\, w$$
$$x + 3y + 3z + 2w = 0 \rightarrow x = -3y - w$$

hay una infinidad de soluciones al sistema con dos parámetros libres w, y. La solución general es la combinación de:

$$x = \begin{bmatrix} -3y - w \\ y \\ -1/3w \\ w \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 1 \end{bmatrix}$$

Si damos valores y = 1, w = 0, tenemos la solución (-3, 1, 0, 0), si y = 0, w = 1 tenemos (-1, 0, -1/3, 1). Todas las soluciones son combinación lineal de estas dos.

Entonces las soluciones de Ax = 0 forman un subespacio bidimensional, **el espacio nulo de A**, del espacio de 4 dimensiones de todos los vectores x. Este espacio nulo lo podemos describir como un "plano" generado por (-3, 1, 0, 0) y (-1, 0, -1/3, 1). Las combinaciones de estos vectores forman un conjunto cerrado bajo la suma y el producto por un escalar y todas estas combinaciones constituyen el espacio nulo.

En el caso de n > m, no debe haber mas de m pivotes distintos de cero, entonces debe haber al menos n-m variables libres por lo cual: Cada sistema homogéneo Ax = 0 tiene una solución no trivial, si tiene más incógnitas que ecuaciones (n > m): existe alguna solución x diferente de la trivial x = 0.

Para el caso no homogéneo $b \neq 0$ tenemos Ax = b , aplicando operaciones elementales tenemos Ux = c:

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b \end{bmatrix}$$

No está claro que el sistema tenga solución. Ya que el tercer renglón de la matriz es igual a cero entonces:

$$b_3 - 2b_2 + 5b_1 = 0$$

El conjunto de vectores obtenibles b no es todo el espacio tridimensional. Es posible resolver Ax = b ,si b está en el espacio columna de A generado por las columnas:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix} \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix}$$

Aunque son 4 vectores, solo obtenemos un plano en el espacio tridimensional (ya que la 2a columna es 3 veces la primera, y la cuarta es la primera más una fracción de la tercera). Son estas columnas las que no tienen pivotes.

Si suponemos que b está en el plano generado por el espacio columna, por ejemplo b = (1, 5, 5) efectuando eliminación y fijando las variables libres obtenemos:

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$

$$\begin{aligned} 3z + w = 3 \quad &\rightarrow \quad z = 1 - 1/3\, w \\ x + 3y + 3z + 2w = 1 \quad &\rightarrow \quad x = -2 - 3y - w \end{aligned}$$

de donde:

$$x = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 1 \end{bmatrix}$$

Si comparamos la solución con la del caso homogéneo Ax = 0, la única diferencia es la inclusión del vector (-2, 0, 1, 0) que es una solución de las ecuaciones dadas, es una solución particular de Ax = b, y la solución general x es una suma de esta solución particular y la solución general de Ax = 0. Geométricamente las soluciones se encuentran de nuevo en un plano en el espacio de 4 dimensiones, pero ahora no constituyen un subespacio, ya que el plano no pasa por el origen, el origen x = 0 no es una solución cuando b $\neq$ 0. El plano es paralelo al espacio nulo que se tenía antes, pero desplazado por el vector que da la solución particular.

Resumiendo tenemos:

El propósito original de la eliminación, era simplificar un sistema de ecuaciones lineales sin alterar ninguna de las soluciones. Se reduce el sistema Ax = 0 a Ux = 0 y éste proceso es reversible. Por lo tanto, el espacio nulo de A es el mismo que el espacio nulo de U. De las m aparentes restricciones impuestas por la m ecuaciones Ax = 0, sólo r son independientes, están especificadas por cualesquiera r renglones de U distintos de cero.

Definición 6.3
El espacio nulo de A: N(A) tiene dimensión n-r y consta de todos aquellos vectores x para los cuales el sistema homogéneo Ax = 0 (o su equivalente Ux = 0) tiene solución. Se obtiene una base para el espacio nulo reduciendo el sistema original al sistema Ux = 0, que tiene n-r variables libres, correspondientes a las columnas de U que no contienen pivotes. Al espacio nulo también se le llama el núcleo de A y su dimensión n-r nulidad de A.

De todo lo expuesto hasta aquí, podemos concluir que A no tiene el mismo espacio columna que U. La eliminación no altera el espacio de los renglones ni el espacio nulo, pero las columnas son completamente diferentes. Sin embargo, si el conjunto de columnas de A es independiente, entonces lo mismo es cierto para las correspondientes columnas de U y viceversa; entonces U nos puede ser útil en términos de indicarnos cuáles son las columnas de A que forman una base.

Para encontrar una base de R(A) consideremos el siguiente resultado:

TEOREMA 6.3
La dimensión del espacio columna R(A) es igual al rango r, que también es igual a la dimensión del espacio fila: el número de columnas independientes es igual al número de renglones independientes. Una base de R(A) está formada por aquellas r columnas de A correspondientes en U, a las columnas que contienen los pivotes distintos de cero.
Este teorema también afirma para el caso de matrices cuadradas, que si los renglones de una matriz cuadrada son l.i. también lo son las columnas.

Definición 6.4
El espacio nulo de A^t ó espacio nulo izquierdo $N(A^t)$,es un subespacio de R^m que consta de todos aquellos vectores; y tales que $A^t y = 0$ ó
$y^t A = 0$ y ,se le llama a y el vector nulo izquierdo de A.

Es fácil encontrar la dimensión de $N(A^t)$. Sabiendo que:

rango + nulidad = dimensión del espacio columna + dimensión del espacio nulo
= número de columnas

Esta regla también se aplica para A^t la cual tiene m columnas.

Pero el rango de los renglones = rango de las columnas = r, entonces:

$$r + \dim N(A^t) = m$$
$$\dim N(A^t) = m - r$$

Esquemáticamente se puede ver en la figura 6.1

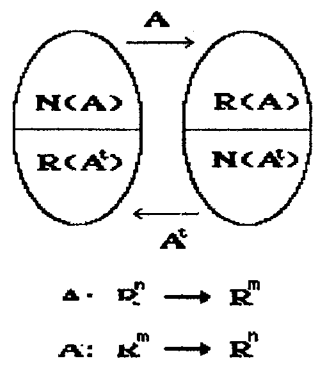


FIGURA 6.1

Sabemos ahora cual es la dimensión de los cuatro espacios asociados a una matriz y resumimos este resultado en el siguiente teorema

***TEOREMA 6.4*: *TEOREMA FUNDAMENTAL DEL ALGEBRA LINEAL* (1a. Parte)**

1.	$R(A^t)$ = espacio renglón de A:	dimensión = r
2.	$N(A)$ = espacio nulo de A:	dimensión = n-r
3.	$R(A)$ = espacio columna de A:	dimensión = r
4.	$N(A^t)$ = espacio nulo izquierdo de A:	dimensión = m-r

tal que:

$$n = \dim R(A) + \dim N(A)$$
$$m = \dim R(A^t) + \dim N(A^t)$$

Este teorema se ilustra con el siguiente ejemplo:

<u>EJEMPLO 6.3</u>

Encuentre la dimensión y construya una base de los cuatros subespacios asociados a la matriz:

$$A = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix}$$

Solución:

1.- El espacio de los renglones de A: $R(A^t)$

Escalonando la matriz A tenemos:

$$\begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Entonces el vector (0,1,4,0) forma una base de los renglones de A y:

$$\text{Dim } [R(A^t)]=1.$$

2.- El espacio de las columnas. Buscamos las columnas cuyos pivotes sean distintos de cero y vemos que la única es la segunda columna $(1,2)^t$,de donde:

$$\text{Dim } (R(A))= 1 \text{ y una de sus bases } (1,2)^t$$

3.- Espacio nulo de A. N(A) con dimensión n-r $\Rightarrow$ 4-1=3 y para determinar una base, procedemos a resolver el sistema Ax = 0, (también puede hacerse con x=0)

$$Ax = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

de donde:

$$x_2 + 4x_3 = 0 \Rightarrow x_2 = -4x_3$$

Sea x_3 =t, entonces x_2 = -4t y haciendo x_1 = u y x_4 = v, se obtiene:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} u \\ -4t \\ t \\ v \end{bmatrix} = t\begin{bmatrix} 0 \\ -4 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Por lo cual:

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

forman una base para N(A), y dim (N(A)) = n - r = 4 - 1 = 3 , que son el número de vectores en la base.

4.- Espacio nulo izquierdo de A :N(A^t) con dimensión m-r $\Rightarrow$ 2-1= 1. Para encontrar una base procedemos a resolver A^t y = 0

$$A^t y = \begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 4 & 8 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

de donde

$y_1 = -2y_2$ haciendo $y_2 = t$
$y_1 = -2t$

Por lo cual una base de N(A^t) está dada por:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

y dim (N(A^t)) = m - r = 2 - 1 = 1

EJEMPLO 6.4

Encuentre la dimensión y construya una base de los cuatro subespacios asociados a la matriz:

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

Solución:

1.- El espacio de los renglones de A: R (A^t)

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

Entonces los vectores {(1, 3, 3, 2), (0, 0, 3, 1)} forman una base de los renglones de A y Dim R(A^t) = 2

2.- El espacio de las columnas. Buscamos las columnas cuyos pivotes sean distintos de cero:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix}$$

$$\text{Dim } R(A) = 2$$

3.- Espacio nulo de A. N(A) con dimensión n-r $\Rightarrow$ 4-2=2 y para determinar una base, procedemos a resolver el sistema Ax=0, (también puede hacerse con Ux=0)

$$Ax = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

de donde:

$$3x_3 + x_4 = 0 \quad \Rightarrow \quad x_3 = -1/3\, x_4$$

Sea $x_4 = t$, entonces $x_3 = -1/3t$ y haciendo $x_1 = -3x_2 - 3x_3 - 2x_4$ y $x_2 = u$ se obtiene:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3u-t \\ u \\ -t/3 \\ t \end{bmatrix} = u\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 1 \end{bmatrix}$$

Por lo cual:

$$\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 0 \end{bmatrix}$$

forman una base para N(A).

4.- Espacio nulo izquierdo de A :N(A^t) con dimensión m-r $\Rightarrow$ 3 - 2 = 1. Para encontrar una base procedemos a resolver $A^t\, y = 0$

$$A^t y = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \\ 3 & 9 & 3 \\ 2 & 5 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Por lo cual, una base de $N(A^t)$ está dada por:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = t \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$$

Hemos visto hasta aquí que para encontrar una base del espacio renglón y una base para el espacio columna se procede a escalonar la matriz A en una matriz U, es entonces factible preguntar si estas matrices están relacionadas como antes, mediante una matriz triangular inferior L, tal que A = LU, la respuesta es si existe una matriz L tal que A = LU, por ejemplo:

Si se tienen las siguientes matrices A y su forma escalonada U

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

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

entonces la matriz L está dada por:

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

y se verifica que A = LU, note además que L no es rectangular, sino cuadrada. Es una matriz del mismo orden m = 3 que el número de renglones en A y en U. En este caso como se requirió de un intercambio de renglones entonces se introduce una matriz de permutación P con lo que enunciamos el siguiente teorema:

TEOREMA 6.5

A cada matriz A de orden mxn corresponde una matriz de permutación P, una matriz triangular inferior L con diagonal unitaria y una matriz escalonada U de orden mxn tales que PA = LU

MATRICES DE RANGO UNO

Un ejemplo de una matriz de rango uno, r = 1 está dado por la siguiente matriz:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix}$$

Cada renglón es un múltiplo del primero de modo que el espacio renglón es unidimensional. Se puede escribir toda la matriz de la siguiente manera:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} (2, 1, 1)$$

como el producto de un vector columna con un vector fila, además, todas las columnas son múltiplos del mismo vector el espacio columna tiene dimensión r = 1.

Cualquier matriz de rango uno se puede factorizar en la forma simple $A = uv^t$. Todos los renglones son múltiplos del mismo vector v^t y todas las columnas son múltiplos de un mismo vector u.

6.3 SOLUCIÓN DE M ECUACIONES EN N INCÓGNITAS

Ya conocemos el proceso de eliminación para matrices cuadradas. La eliminación misma se realiza directamente sin mayor cambio, pero hay algunas diferencias al obtener la solución mediante la sustitución regresiva.

Consideremos la ecuación escalar ax = b (un sistema de una ecuación con una incógnita), podemos considerar las siguientes tres alternativas:

1. Si $a \neq 0$ entonces existe la solución x = b/a para cualquier b y ésta solución es única. Este es el caso **no singular**, de una matriz de 1x1 invertible.

2. Si a=0 y b=0 hay una infinidad de soluciones, cualquier x satisface 0x = 0. Este es el caso **indeterminado,** existe solución pero no es única.

3. Si a=0 y $b \neq 0$ no hay solución a 0x = b este es el caso **inconsistente.**

Estas alternativas pueden ocurrir para matrices cuadradas. Con una matriz rectangular no podemos tener existencia y también unicidad para cada b. Esto se puede ver a través de la existencia de la inversa.

Sabemos que una matriz A tiene inversa si existen BA = I y AC = I donde $B = C = A^{-1}$. Ahora, del rango de una matriz es fácil decir qué matrices tienen realmente estas inversas, se puede decir en general que existe una inversa sólo cuando el rango es lo más grande posible.

El rango siempre satisface r□m y r□n ya que una matriz de mxn no puede tener más de m filas independientes o de n columnas independientes. Entonces si r = m existe una inversa derecha y si r = n existe una inversa izquierda.

Definición 6.5

Sea A matriz de orden mxn. Una matriz G de nxm (si existe) tal que $GA = I_n$ se dice inversa izquierda de A. Asimismo una matriz H de orden nxm (si existe), tal que:
$AH = I_m$ se dice inversa derecha.

<u>EJEMPLO 6.5</u>

Sea la matriz de 2x1 siguiente

$$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

calcule (si existe) la inversa izquierda y derecha de A

<u>Solución:</u>

Sea $G = [g_1, g_2]$ Si GA = 1

$$[g_1, g_2]\begin{bmatrix} 1 \\ 1 \end{bmatrix} = g_1 + g_2 = 1$$

Si $g_1 = \alpha$ entonces $g_2 = 1-\alpha$ de donde $G = [\alpha, 1-\alpha]$ para $\alpha \in R$

Sea $H = [h_1, h_2]$ tal que AH = I

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}(h1,\ h2) = \begin{bmatrix} h_1 & h_2 \\ h_1 & h_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

pero $h_1 = 1$ y $h_1 = 0$ por lo que se afirma que no existe matriz derecha además el rango de A = 1 < 2 = m

TEOREMA 6.6

EXISTENCIA: El sistema $Ax = b$ tiene al menos una solución x para toda b si y sólo si las columnas de A generan R^m, entonces $r = m$. En este caso existe una inversa derecha H de nxm tal que $AH = I_m$, la matriz identidad de orden m. Esto es posible sólo sí $m \leq n$.

UNICIDAD: El sistema $Ax = b$ tiene a lo sumo una solución x para cada b si y sólo si las columnas son linealmente independientes, entonces $r=n$. En éste caso existe una inversa izquierda G de nxm tal que $GA = I_n$, la matriz identidad de orden n. Esto es posible sólo si $m \geq n$.

En el primer caso, una posible solución es $x = Hb$, ya que entonces $Ax = AHb = b$ pero habrá otras soluciones si hay otras inversas derechas.

En el segundo caso, si existe una solución a $Ax = b$, tiene que ser $x = GAx = Gb$, pero no puede haber otra solución.

<u>EJEMPLO 6.6</u>

Considere la siguiente matriz A de 2x3 y de rango 2

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}$$

verifique si existen inversa izquierda o derecha de A.

<u>Solución:</u>

como $r = m = 2$ el teorema garantiza una inversa derecha H:

$$AH = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \\ c_{31} & c_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

de hecho hay muchas inversas derechas. Ya que el último renglón de H es totalmente arbitrario, entonces si escribimos un sistema $Ax = b$ con esta matriz A se puede afirmar que existe solución.

Para ver si existe inversa izquierda solo hay que verificar que $r = n$, en este caso $r = 2 < 3$ por lo que no existe inversa izquierda.

TEOREMA 6.7
Una matriz A de nxn es invertible si y sólo si rango (A) = n

EJEMPLO 6.7
Determine si la matriz siguiente es invertible hallando su rango.

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

Solución:
Escalonando A, se tiene:

$$\begin{bmatrix} 4 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 3 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 1 & -12 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & -14 \end{bmatrix}$$

Por lo que rango A = 3 y A es invertible

6.4 SUBESPACIOS ORTOGONALES

Recordemos que la longitud de un vector $\|x\|$ en R^n está dada por:

$$\|x\| = \sqrt{x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + \ldots + x_n^{\,2}}$$

Geométricamente esto se obtiene aplicando la fórmula de Pitágoras n-1 veces, añadiendo en cada paso una dimensión más.

Si tenemos dos vectores x, y ¿Cómo podemos saber si son perpendiculares?

En el caso de R^2 la respuesta se puede obtener usando trigonometría, en este caso x es ortogonal a y si forman un triángulo rectángulo, usando Pitágoras se tiene:

$$\|x\|^2 + \|y\|^2 = \|x - y\|^2$$

Para R^n se tiene:

$$(x_1^2 + x_2^2 + \ldots + x_n^2) + (y_1^2 + y_2^2 + \ldots + y_n^2) = (x_1 - y_1)^2 + \ldots + (x_n - y_n)^2$$

Desarrollando el lado derecho se tiene:

$$(x_1^2 + x_2^2 + \ldots + x_n^2) - 2\,(x_1y_1 + \ldots + x_ny_n) + (y_1^2 + y_2^2 + \ldots + y_n^2)$$

Así, la igualdad se cumple si los "términos cruzados" son iguales a cero:

$$x_1\,y_1 + \ldots + x_n\,y_n = 0$$

Esta cantidad no es otra, que el producto interior de $x^t y$ en R^n; por lo que se puede afirmar que es igual a cero ,sólo si "x " e " y "son ortogonales.

Existe una relación sencilla entre independencia y ortogonalidad: si los vectores son mutuamente ortogonales, entonces son linealmente independientes.

En el espacio tridimensional ordinario se representan los subespacios mediante rectas o planos que pasan por el origen y en casos extremos por el origen mismo o por todo el espacio.

Definición 6.6

Dos subespacios V y W del mismo espacio R^n son ortogonales si cada vector $v \in V$ es ortogonal a cada vector $w \in W$: $v^T w = 0$ para toda v y w.

EJEMPLO 6.8

Suponga que V es el plano generado por $v_1 = (1,0,0,0)$ y $v_2 = (1,1,0,0)$ y W es una recta generada por $w_1 = (0,0,4,5)$, como $w_1^{\,T} v_1 = 0$ y $w_2^{\,T} v_2 = 0$ entonces la recta W será ortogonal a todo el plano V.

En el caso de los cuatro subespacios fundamentales asociados a una matriz dos son subespacios de R^n : el espacio nulo y el espacio renglón o fila y los otros dos están en R^m, como se muestra en la figura 6.2

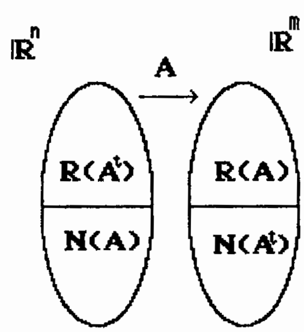


FIGURA 6.2

TEOREMA 6.8
Para cualquier matriz A de orden mxn, el espacio nulo y el espacio fila son subespacios ortogonales de R^n. Análogamente, el espacio nulo izquierdo y el espacio columna son subespacios ortogonales de R^m.

Demostración:

Supongamos que $w \in N(A)$ y $v \in R(A^t)$, entonces $Aw = 0$ y v es de la forma $v = A^t x$ para algún vector x (esto es, v es una combinación de las columnas de A^t). Por lo tanto

$$w^t v = w^t(A^t x) = (w^t A^t)x = (Aw)^t x = 0^t x = 0$$

EJEMPLO 6.9
Sea la matriz A y su correspondiente U como sigue:

$$A = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \quad U = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

la segunda columna es básica y las otras tres son variables libres, por lo tanto resolviendo para $Ux = 0$, se obtiene que una base para N(A) está dada por:

$$\{[1,0,0,0]^t, [0,-4,1,0]^t, [0,0,0,1]^t\}$$

la cual debe ser ortogonal a los renglones de A.

El espacio columna de A es unidimensional y está generado por la única columna básica $[1,2]^t$. Por otro lado hallamos una base del espacio nulo izquierdo, dada por el vector $y^t = [-2,1]$, el cual es ortogonal al espacio columna:

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

Sin embargo cabe aclarar que el espacio nulo N(A) no sólo contiene algunos de los vectores ortogonales al espacio fila, contiene todos esos vectores. El espacio nulo se formó con todas las soluciones de $Ax = 0$.

Definición 6.7
Dado un subespacio V de R^n, el espacio de todos los vectores ortogonales a V es el complemento ortogonal de V y se denota por $V^{\perp}$.

Así se tiene entonces que el espacio nulo N(A) es el complemento ortogonal de $R(A^T)$: $N(A) = (R(A^T))^{\perp}$. Al mismo tiempo es válida la relación opuesta: el espacio fila $R(A^T)$ contiene todos los vectores que son ortogonales al espacio nulo.

Al aplicar el mismo razonamiento a A^T, se produce el resultado dual: el espacio nulo izquierdo $N(A^T)$ y el espacio columna R(A) son complementos ortogonales entre si en R^m. Esto completa la segunda parte del teorema fundamental del álgebra lineal.

TEOREMA 6.9: TEOREMA FUNDAMENTAL DEL ALGEBRA LINEAL (2ª parte)

$$N(A) = (R(A^T))^{\perp}, \qquad R(A^T) = N(A)^{\perp}$$
$$N(A^T) = (R(A))^{\perp}, \qquad R(A) = (N(A^T))^{\perp}.$$

La última igualdad significa que Ax = b tiene una solución si y sólo si b es ortogonal a $N(A^T)$; b está en el espacio columna si y sólo si es ortogonal a cada solución y de la ecuación homogénea transpuesta $A^T y = 0$.

Sin embargo dos subespacios V y W pueden ser ortogonales sin ser complementos ortogonales entre sí. En R^3, la recta V generada por (1,0,0) es ortogonal a la recta W generada por (0,0,1), pero V no es igual $W^{\perp}$. El complemento ortogonal de W es un subespacio bidimensional (un plano) que contiene todos los vectores de la forma (x_1, x_2, 0), esto es todos los generados por los vectores (1,0,0) y (0,1,0). La recta V sólo puede ser parte de $W^{\perp}$ porque su dimensión es muy pequeña. Sin embargo, si las dimensiones son correctas, los dos subespacios ortogonales son necesariamente complementos ortogonales, así fue en el caso del espacio fila y el espacio nulo. Además si $W = V^{\perp}$, esto asegura que las dimensiones son las correctas y automáticamente $V = W^{\perp}$. Simplemente se descompone el espacio en dos partes perpendiculares V y W

TEOREMA 6.10

Si V y W son subespacios de R^n, entonces cualquiera de las siguientes condiciones los obliga a ser complementos ortogonales entre sí:

1. $W = V^{\perp}$ (W consta de todos los vectores ortogonales a V)
2. $V = W^{\perp}$ (V consta de todos los vectores ortogonales a W)
3. V y W son ortogonales y dim V + dim W = n.

Suponiendo cualquiera de estas tres condiciones equivalentes, cada vector x puede descomponerse de una sola manera en una suma x = v + w con $v \in V$ y $w \in W$. Estas componentes, las proyecciones de x en V y W son ortogonales $v^T w = 0$.

Una x arbitraria se descompone en $x_r + x_n$, y A transforma la componente x_r del espacio fila en un vector $Ax_r = Ax$ en el espacio columna, mientras que transforma la componente x_n del espacio nulo en cero.

EJEMPLO 6.10

Sea la matriz A como sigue:

$$A = \begin{bmatrix} 0 & 1 & 4 \\ 0 & 2 & 8 \end{bmatrix}$$

su correspondiente forma escalonada queda como sigue:

$$U = \begin{bmatrix} 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

de aquí que $\qquad \dim R(A^t) = 1, \dim N(A) = 2$

una base de $\qquad R(A^t) = \{ (0, 1, 4)\}$

una base de N(A) está dada por $\{(1, 0, 0),(0, 4, 1)\}$ entonces cualquier $x \in R^3$ se puede escribir como:

$$x = \underbrace{\alpha_1 \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}}_{\text{basedeR}(A^T)} \underbrace{+\alpha_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_{32} \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}}_{\text{basedeN}(A)}$$

6.5 ESPACIOS CON PRODUCTO INTERIOR Y EL PROCESO DE GRAM-SCHMIDT

Definición 6.8

Un producto interior en un espacio vectorial V es una función que a cada par de vectores x e y en V asocia un número real $\langle x, y\rangle$ de tal manera que se satisfacen los siguientes axiomas, para todos los vectores x, y, z en V y cualquier escalar k.

1. $\langle x, y\rangle = \langle y, x\rangle$ — axioma de la simetría
2. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ — axioma de la aditividad
3. $\langle kx, y\rangle = k \langle x, y\rangle$ — axioma de la homogeneidad
4. $\langle x, x\rangle \geq 0$ y $\langle x, x\rangle = 0 \iff x = 0$ — axioma de positividad.

Un espacio vectorial que tiene definido un producto interior se denomina espacio vectorial con producto interior.

Las siguientes propiedades adicionales son una consecuencia inmediata de los cuatro axiomas del producto interior:

a) $<0, x> = <x, 0> = 0$
b) $<x, y+z> = <x, y> + <x, z>$
c) $<x, ky> = k<x, y>$

EJEMPLO 6.11

Considere el espacio R^n con producto interior dado como

$$< x, y > \; = x^t y = \sum_{i=1}^{n} x_i \, y_i$$

donde x, y son vectores columna en R^n.

Note que $<.,\ ,>$ ó $<\,|\,>$ es un producto interior ya que:

1. $<x, y> = x^t y = y^t x = <y, x>$
2. $<x+y, z> = (x+y)^t z = (x^t + y^t) z = x^t z + y^t z = <x, z> + <y, z>$
3. $<kx, y> = (kx)^t y = kx^t y = k <x, y>$
4. $<x, x> = x^t x = \Sigma x_i^2 \square 0$ si $\Sigma x_i^2 = 0$ entonces $x = 0 \in R^n$

EJEMPLO 6.12

Sean $x = <x_1, x_2>$, $y = <y_1, y_2>$ vectores en R^2, entonces

$$<x, y> = 3x_1 y_1 + 2x_2 y_2$$

define un producto interior.

Solución:

Para verificar esta afirmación observe que:

1. $<x, y> = 3x_1 y_1 + 2x_2 y_2 = 3y_1 x_1 + 2y_2 x_2 = <y, x>$
2. $<x+y, z> = 3(x_1 + y_1)z_1 + 2(x_2 + y_2)z_2 = (3x_1 z_1 + 2x_2 z_2) + (3x_1 z_1 + 2x_2 z_2) =$
   $<x, z> + <y, z>$
3. $<kx, y> = 3(kx_1)y_1 + 2(kx_2)y_2 = k(3x_1 y_1 + 2x_2 y_2) = k<x, y>$
4. $<x, x> = 3x_1 x_1 + 2x_2 x_2 = 3x_1^2 + 2x_2^2 \geq 0$ y $3x_1^2 + 2x_2^2 = 0$ si y sólo si $x_1 = x_2 = 0$.

El producto interior de este ejemplo es distinto al producto interior de R^2 lo que muestra que un espacio vectorial puede tener más de un producto interior definido en él.

EJEMPLO 6.13

Sean $x = <x_1, x_2>$, $y = <y_1, y_2>$ vectores en R^2, entonces:

$$<x, y> = x_1 + y_1$$

verifique si $<x, y>$ define un producto interior.

Solución:

Para verificar esta afirmación observe que:

1. $<x, y> = x_1 + y_1 = y_1 + x_1 = <y, x>$
2. $<x+y, z> = (x_1 + y_1) + z_1 = (x_1 + z_1) + y_1 = <x+z, y> \neq <x, z> + <y, z>$

por lo que se puede concluir que $<x, y>$ así definido no es un producto interior.

EJEMPLO 6.14

(Caso de polinomios) Sea P_n el espacio vectorial de los polinomios de grado $\leq$ n. Un producto interior en este espacio está dado por:

$$\langle x, y \rangle = \int_a^b x(t)y(t)dt$$

1. $\langle x, y \rangle = \int_a^b x(t)y(t)dt = \int_a^b y(t)x(t)dt = \langle y, x \rangle$
2. $\langle x, y + z \rangle = \int_a^b x(t)[y(t) + z(t)]dt = \int_a^b [x(t)y(t) + x(t)z(t)]dt$
   $= \int_a^b x(t)y(t)dt + \int_a^b x(t)z(t)dt = \langle x, y \rangle + \langle x, z \rangle$
3. $\langle kx, y \rangle = \int_a^b kx(t)y(t)dt = k\int_a^b x(t)y(t)dt = k\langle x, y \rangle$
4. $\langle x, x \rangle = \int_a^b x^2(t)dt = 0$ si y solo si $x = 0$

Definición 6.9

Si V es un espacio con producto interior, entonces se llama un espacio normado si existe una función $\|.\|$: $V \to R$ tal que:

$$\|x+y\| \leq \|x\| + \|y\|$$
$$\|\alpha x\| = |\alpha| \|x\|$$
$$\|x\| \geq 0 \quad y \quad \|x\| = 0 \leftrightarrow x = 0$$

EJEMPLO 6.15

Muestre que el espacio $V = R^n$ es normado con cualquiera de las siguientes normas:

$$a) \quad \|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$$b) \quad \|x\|_1 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

$$c) \quad \|x\|_\infty = \max \{x_i | i = 1,2,\dots n\}$$

Solución:

Vamos a demostrar que efectivamente a) b) y c) son normas, esto es:

$$a) \quad \|x+y\|_1 = \sum |x_i + y_i| \leq \sum |x_i| + \sum |y_i| = \|x\|_1 + \|y\|_1$$
$$\|\alpha x\| = \sum |\alpha x_i| = |\alpha| \sum x_i = |\alpha| \|x\|_1$$
$$\|x\| = \sum |x_i| \geq 0 \quad y \quad \|x\| = 0 \leftrightarrow x = 0$$

$$b) \quad \|x+y\|_2 = \sum (x_i + y_i)^2 = \sum \left(x_i^2 + 2x_i y_i + y_i^2\right) =$$
$$\sum x_i^2 + 2\sum x_i y_i + \sum y_i^2 = \|x\|_1 + \|y\|_1$$

$$c) \quad \|x+y\|_\infty = \max\{x_i + y_i|, i = 1,2,..n\} \leq \max\{x_i| + |y_i| i = 1,2..n\}$$
$$\max\{x_i|i = 1,2..,n\} + \max\{y_i|i = 1,2..n\} = \|x\|_\infty + \|y\|_\infty$$
$$\|\alpha x\|_\infty = \max\{\alpha x_i | i = 1,2..,n\} = \max\{\alpha \; |x_i| i = 1,2..,n\} =$$
$$|\alpha| \max\{x_i | i = 1,2..,n\} = \alpha \|x\|_\infty$$

6.6 ÁNGULO EN LOS ESPACIOS VECTORIALES, PROYECCIONES ORTOGONALES Y PROCESO DE GRAM-SCHMIDT

Ya vimos en la sección anterior que el hecho de que dos vectores x e y sean ortogonales se puede expresar como $x^t y = 0$ donde x y son vectores columna.

Considere el espacio vectorial R^2 y los vectores a y b donde sus longitudes están dadas por $\|a\|$ y $\|b\|$ respectivamente. Sea α el ángulo que forma el vector a con el eje de las x y ß el ángulo del vector b con el mismo eje como se ilustra en la figura 6.3

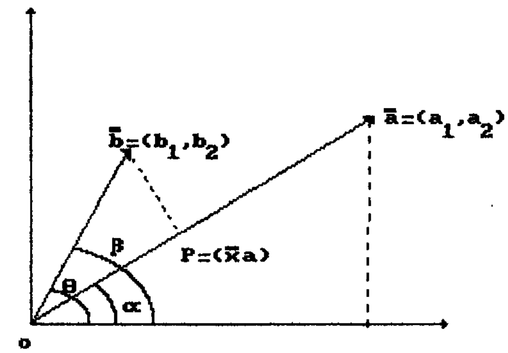


FIGURA 6.3

entonces se definen el seno y coseno de a como

$$\operatorname{sen}\alpha = \frac{a_2}{\|a\|} \qquad \cos\alpha = \frac{a_1}{\|a\|}$$

y para ß se tiene

$$\operatorname{sen}\beta = \frac{b_2}{\|a\|} \qquad \cos\beta = \frac{b_1}{\|a\|}$$

si Θ = ß - α entonces usando una identidad trigonométrica se tiene:

$$\cos\Theta = \cos\beta\cos\alpha + \operatorname{sen}\beta\operatorname{sen}\alpha = a_1 b_1$$

por lo tanto el coseno del ángulo entre dos vectores cualesquiera a y b está dado por

$$\cos\Theta = \frac{\langle a,b \rangle}{\|a\| \; \|b\|}$$

Si queremos encontrar la proyección de un punto del vector b, sobre el vector a, este punto debe ser algún múltiplo p = xa del vector a y el problema se reduce a calcular el coeficiente x. Todo lo que necesitamos es el hecho de que la recta desde b hasta el punto más cercano p = xa es ortogonal al vector a, como se puede ver en la figura 6.3, entonces:

$$(b - xa) \perp a \quad \text{ó} \quad <a, b\text{-}xa> = 0$$

de donde

$$x = \frac{\langle a, b\rangle}{\langle a, a\rangle} = \frac{\langle a, b\rangle}{\|a\|^2}$$

entonces, la proyección p del punto b sobre la recta generada por el vector a está dada por:

$$p = \frac{\langle a, b\rangle}{\|a\|^2} a$$

De esta relación se puede encontrar nuevamente la desigualdad de Cauchy- Schwarz ya que la distancia al cuadrado del punto a la recta es:

$$\left\| b - \frac{\langle a,b\rangle}{\|a\|^2} a \right\|^2 = \frac{\|b\|^2 \|a\|^2 - \langle a,b\rangle}{\|a\|^2}$$

como la distancia de un punto a una recta es mayor o igual a cero lo mismo sucede con el cuadrado de la distancia entonces

$$\|b\|^2 \|a\|^2 - \langle a,b\rangle \geq 0$$

$$\|b\|^2 \|a\|^2 \geq \langle a,b\rangle \text{ que es la desigualdad de Schwarz}$$

Al seleccionar una base se busca que ésta simplifique la solución del problema que se trate. En la mayoría de los casos, la mejor selección será la base en la cual todos los vectores son ortogonales entre sí, es así como definimos:

Definición 6.10

Se dice que un conjunto de vectores en un espacio con producto interior es ortonormal; si cada vector del conjunto tiene norma 1 y si además dos vectores cualesquiera distintos en el conjunto son ortogonales.

EJEMPLO 6.16
Sean

$v_1 = (0, 1, 0)$, $v_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$, $v_3 = (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$, el conjunto $S = \{v_1, v_2, v_3\}$ es ortonormal si en R^3 se tiene que con el producto interior euclideano:

$$\langle v_1, v_2\rangle = \langle v_2, v_3\rangle = \langle v_1, v_3\rangle = 0$$

y además

$$\|V_1\| = \|V_2\| = \|V_3\| = 1$$

EJEMPLO 6.17
Si v es un vector en un espacio con producto interior, y v es diferente de cero, entonces el vector

$$\frac{1}{\|V\|}V = 1 \text{ tiene norma } 1$$

ya que aplicando $\|ku\| = |k|\|u\|$ se tiene

$$\left\|\frac{1v}{\|v\|}\right\| = \frac{1}{\|v\|}\|v\| = 1$$

al proceso de multiplicar un vector v diferente de cero por el recíproco de su longitud para obtener un vector de norma 1 se denomina **normalización de v.**

TEOREMA 6.11: TEOREMA DE LA BASE ORTONORMAL (GRAM-SCHMIDT)

Todo subespacio distinto de cero en R^n tiene una base ortonormal.

Demostración:

Queremos producir de dos vectores independientes ay b dos vectores perpendiculares v_1 y v_2. Claramente, el primero puede ir en la dirección de a $v_1 = a$. El problema es entonces encontrar un segundo vector que sea perpendicular.

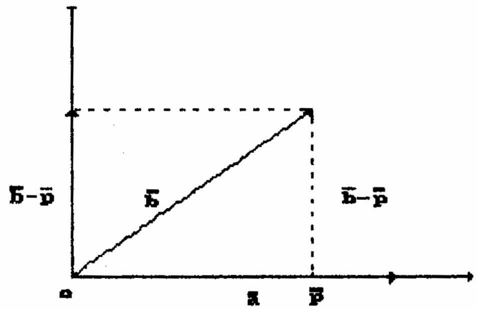


FIGURA 6.4

En la figura 6.4 se puede ver que ese vector es b-p ya que se le sustrajo la proyección de b en la dirección de a, entonces el segundo vector queda:

$$v_2 = b - p = b - \frac{\langle a, b\rangle}{\|a\|^2} a = b - \frac{\langle v_1, b\rangle}{\|v_1\|^2} v_1$$

si hay un tercer vector independiente entonces la idea es la misma, sustraemos las proyecciones de c en la dos direcciones v_1 y v_2

$$v_3 = c - \frac{\langle v_1, c\rangle}{\|v_1\|^2} v_1 - \frac{\langle v_2, c\rangle}{\|v_2\|^2} v_2$$

es inmediato que v_3 es perpendicular a v_1 y v_2.

Para que estos vectores sean ortonormales deben transformarse en vectores unitarios entonces

$$q_1 = \frac{v_1}{\|v_1\|}, \quad q_2 = \frac{v_2}{\|v_2\|}, \quad q_3 = \frac{v_3}{\|v_3\|}$$

Este proceso se puede resumir de la siguiente manera:

Es posible convertir cualquier conjunto de vectores independientes a_1, a_2,...,a_n en un conjunto de vectores ortogonales mediante el proceso de Gram-Schmidt: primero se fija $a_1 = v_1$ después cada v_i es ortogonal a las v_1,..., v_{i-1} precedentes:

$$v_i = a_i - \frac{\langle v_1, a_i\rangle}{\|v_1\|^2} v_1 - \dots - \frac{\langle v_{i-1}, a_i\rangle}{\|v_{i-1}\|^2} v_{i-1}$$

EJEMPLO 6.18

Aplique el proceso de Gram-Schmidt a los siguientes vectores independientes:
$a_1 = [1, 1, 0]$, $a_2 = [1, 0, 1]$, $a_3 = [0, 1, 1]$

Solución:

Sea $a_1 = v_1$, v_2 y v_3 se calculan:

$$v_2 = a_2 - \tfrac{1}{2} v_1 = [\tfrac{1}{2}, -\tfrac{1}{2}, 1]$$

$$v_3 = a_3 - \tfrac{1}{2} v_1 - 1/3\, v_2 = [-2/3, 2/3, 2/3]$$

los vectores ortonormales finales son:

$$q_1 = \sqrt{\tfrac{1}{2}}\,[1, 1, 0]$$

$$q_2 = \sqrt{\tfrac{2}{3}}\,[\tfrac{1}{2}, -\tfrac{1}{2}, 1]$$

$$q_3 = \sqrt{\tfrac{3}{4}}\,[-2/3, 2/3, 2/3]$$

6.7 MATRICES DE PROYECCIÓN Y MÍNIMOS CUADRADOS

Considere ahora la proyección p de un vector sobre un subespacio general $S = \{a_1, a_2, \ldots, a_k\}$ de R^n donde las a_i son vectores independientes de R^n. Los detalles se elaboran para el caso k = 2, pero los cálculos son igualmente válidos para el caso general. Recordemos que $\{a_1, a_2\}$ corresponde a un plano en R^n que contiene al origen, y sus miembros son todas las combinaciones lineales de a_1 y a_2.

La figura 6.5 muestra al vector proyección p de b en $\{a_1, a_2\}$

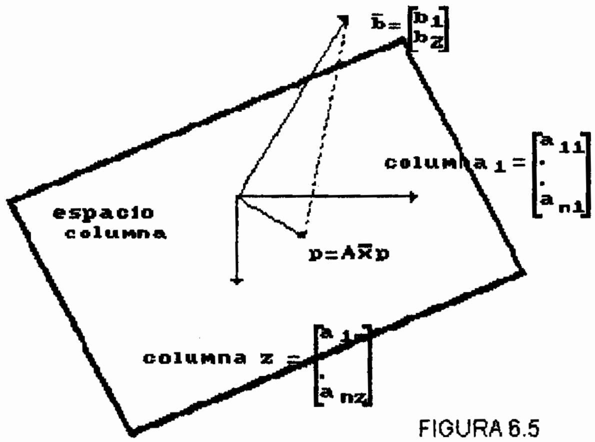


FIGURA 6.5

Observe que p satisface:

1. p debe estar en el subespacio $S = \{a_1, a_2\}$ y
2. b-p debe ser perpendicular a cada vector de S.

Si escribimos los vectores de R^n como vectores columna, el subespacio S es el espacio columna de la matriz A de nx2 cuyas columnas son a_1 y a_2 entonces como $Ax = b$ tiene solución si y sólo si b está en el espacio columna de A, todos los vectores de S tienen la forma Ax donde $x = [x_1, x_2]^t$. Como p está en el espacio S entonces $p = Ax_p$ donde $x_p = [r_1, r_2]^t$, r_1, r_2 escalares, como $b - Ax_p$ debe ser perpendicular a cada vector de S entonces el producto punto de $b- Ax_p$ y Ax debe ser igual a cero, para toda x.

$$<b-Ax_p, Ax> = 0$$

o bien

$$(Ax)^t (b - Ax_p) = x^t(A^tb - A^tAx_p) = 0$$

es decir, el producto punto de los vectores x y $A^tb - A^tAx_p$ debe ser cero para todos los vectores x, pero esto solo sucede sí

$$A^t b - A^tA x_p = 0$$

La matriz A^tA de 2x2 es invertible* pues tiene el mismo rango que A, entonces despejando x_p se tiene

$$x_p = (A^tA)^{-1} A^t b$$

Denotando el vector proyección $p = Ax_p$ de b sobre S por medio de $p = b_s$ y escribiendo b como un vector columna obtenemos la forma general:

Sea $S = \{a_1,\dots,a_k\}$ un subespacio de R^n y sea la matriz A cuyas columnas son los vectores $a_1, a_2,\dots, a_k$. El vector proyección de b en R^n sobre S está dado por :

$$b_s = A(A^tA)^{-1} A^tb$$

este resultado se puede usar para una proyección de un vector en otro, por ejemplo:

<u>EJEMPLO 6.19</u>

Sea el vector a = (2, 4, 3), encuentre el vector proyección sobre a del vector b = (1, 2, 3)

Solución:

$$p = \frac{\langle a, b\rangle}{\|a\|^2} a = \frac{19}{29}(2,\ 4,\ 3)$$

Considerando la fórmula $b_s = A(A_tA)^{-1} A^tb$ se tiene que la matriz A consiste de una sola columna que es el vector a, haciendo cálculos se tiene:

$$A^tA = 29$$

$$b_s = A(A^tA)^{-1} A^tb = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} \frac{1}{29}(2,4,3)\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{29}\begin{bmatrix} 4 & 8 & 6 & 1 \\ 8 & 16 & 12 & 2 \\ 6 & 12 & 9 & 3 \end{bmatrix} = \frac{1}{29}\begin{bmatrix} 38 \\ 76 \\ 57 \end{bmatrix}$$

A la matriz $P_s = A(A^tA)^{-1}A^t$ se le conoce como la matriz de proyección para el subespacio S.

EJEMPLO 6.20

Encuentre la matriz de proyección para el plano x_2x_3 de R^3.

Solución:

El plano x_2x_3 es el subespacio S generado por los vectores e_2e_3 que son las columnas de la siguiente matriz A.

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

donde $A^tA = I$ y por lo tanto

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

Así P_s proyecta cada vector b en R^3 en el plano x_2x_3 de la siguiente manera:

$$Ps\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ b_2 \\ b_3 \end{bmatrix}$$

EJEMPLO 6.21

Sea la matriz A y el vector b como sigue:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 4 \\ 3 \\ 9 \end{bmatrix}$$

encuentre la proyección del vector b sobre el plano generado por los vectores columna de la matriz A.

Solución:

El espacio columna de A es el plano x-y en el espacio tridimensional ya que ambas columnas terminan con un cero. La proyección de b sobre este plano no altera las componentes x, y que son 4 y 3 pero desaparecerá la componente z y se puede confirmar.

$$A^tA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 5 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 7 \\ 7 & 29 \end{bmatrix}$$

$$\left(A^tA\right)^{-1} = \left(\tfrac{1}{9}\right)\begin{bmatrix} 29 & -7 \\ -7 & 2 \end{bmatrix}$$

$$b_2 = \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$$

TEOREMA 6.12

Para cualquier matriz A de mxn, de rango r, la matriz simétrica A^tA de nxn también tiene rango r.

Demostración:

Trabajando con los espacios nulos se tiene que si v es cualquier vector solución del sistema Ax = 0, de modo que Av = 0 entonces, al multiplicar a la izquierda ambos miembros de esta última ecuación por A^t, vemos que v también es una solución al sistema $(A^t)Ax = 0$. Recíprocamente, supongamos que $(A^t)\,Aw = 0$ para un vector w de nx1. Entonces:

$$0 = w^t\,[(A^t)Aw] = (Aw)^t(Aw)$$

Que se puede escribir como $[Aw]^2$. Esto es, Aw es un vector de magnitud cero y, en consecuencia, Aw = 0. Resulta así que A y $(A^t)A$ tienen el mismo espacio nulo. Como tienen el mismo número de columnas, se concluye que el rango de A y el rango de $(A^t)A$ son iguales.

De este teorema se desprende el siguiente resultado: Si A tiene columnas linealmente independientes de modo que r = n, entonces A^tA es una matriz cuadrada, simétrica e invertible.

MATRICES ORTOGONALES

Una matriz ortogonal es simplemente una matriz cuadrada con columnas ortonormales. Se usará la letra Q para denotar una matriz ortogonal y a $q_1,\ldots,q_n$ para denotar sus columnas.

Como Q es cuadrada de nxn entonces si sus columnas son independientes, esto es su rango r es tal que r = n y es invertible, entonces: si Q^t es inversa izquierda entonces es la inversa y $QQ^t = I$

Una matriz ortogonal tiene las siguientes propiedades:

$$Q^tQ = I = QQ^t$$

$$Q^t = Q^{-1}$$

Q no sólo tiene las columnas ortonormales, sino que también sus filas son ortonormales, en otras palabras si Q es ortogonal también lo es Q^t.

Además cumple: La multiplicación por una matriz ortogonal Q preserva la longitud

$$\|Qx\| = \|x\| \quad \text{para toda vector x}$$

y también preserva los productos internos:

$$(Qx)^t\,(Qy) = x^t\,y \text{ para todos los vectores x,y.}$$

EJEMPLO 6.22

Sea la matriz Q como sigue:

$$Q = \begin{bmatrix} \cos\Theta & -\operatorname{sen}\Theta \\ \operatorname{sen}\Theta & \cos\Theta \end{bmatrix}$$

su transpuesta es igual a su inversa

$$Q^t = Q^{-1} = \begin{bmatrix} \cos\Theta & -\operatorname{sen}\Theta \\ \operatorname{sen}\Theta & \cos\Theta \end{bmatrix}$$

Q rota cada vector en un ángulo Θ y Q^t lo rota de regreso en $-\Theta$

EJEMPLO 6.23

Verifique que la matriz siguiente es una matriz ortogonal y encuentre su inversa.

$$A = \left(\tfrac{1}{7}\right) \begin{bmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \\ 6 & 2 & -3 \end{bmatrix}$$

Solución:

$$A^t A = I$$

$$A^{-1} = A^t = \left(\tfrac{1}{7}\right) \begin{bmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \\ 6 & 2 & -3 \end{bmatrix}$$

MÉTODO DE LOS MÍNIMOS CUADRADOS

Supongamos que se obtienen datos de mediciones de la forma (a_i, b_i) mediante observación o experimentación y se ubican como puntos de datos en el plano x-y. Es conveniente hallar una relación matemática $Y = f(x)$ que represente razonablemente bien los datos, de manera que podamos efectuar predicciones de valores no medidos. Dependiendo de la naturaleza del experimento y de la configuración de los puntos de datos localizados, podemos decidir acerca de un tipo apropiado de función $y = f(x)$ como una función lineal, una función cuadrática o exponencial, entre otras. Se pueden presentar problemas como el siguiente:

De acuerdo con la ley de Hooke, la distancia que se estira un resorte es proporcional a la fuerza aplicada. Supongamos que colocamos 4 pesos distintos, a_1, a_2, a_3 y a_4 sucesivamente en la parte inferior de un resorte. Medimos las cuatro longitudes b_1, b_2, b_3 y b_4 del resorte estirado y suponga que se obtienen los datos de la siguiente tabla:

ai = peso en gramos	2.0	4.0	5.0	6.0
bi = longitud en cms.	2.5	4.5	7.0	8.5

Debido a la ley de Hooke, esperamos que los puntos de datos (a_i, b_i) estén cerca de alguna recta con ecuación

$$y = f(x) = r_0 + r_1x$$

donde r_0 es la longitud del resorte y r_1 es la constante del resorte. Esto es, si nuestras mediciones fueran exactas y el resorte ideal, tendríamos $b_i = r_0 + r_1\, a_i$ para valores específicos r_0 y r_1.

Como solamente tenemos las dos incógnitas r_0 y r_1, bastarán dos mediciones para hallarlas, sin embargo, en la práctica esperamos tener algún error en las mediciones físicas, por lo tanto hacemos más mediciones de las técnicamente necesarias con la esperanza de que en general, los errores se cancelen unos con otros. La sustitución de cada punto de datos (a_i, b_i) en la ecuación da una ecuación lineal con dos incógnitas r_0 y r_1, así los 4 puntos de datos del problema, dan lugar a un sistema lineal de 4 ecuaciones con 2 incógnitas. Dicho sistema lineal con más ecuaciones que incógnitas se llama <u>sobredeterminado</u>, y se espera que dicho sistema sea inconsistente, entonces debemos encontrar valores para las incógnitas r_0 y r_1 que estén los más cerca posible, de satisfacer las 4 ecuaciones. Geométricamente esto equivale a encontrar la recta en el plano que esté más cerca de pasar por los 4 puntos de datos.

En general, consideramos el problema de hallar una recta o función lineal $f(x) = r_0 + r_1x$ que ajuste "mejor" los datos (a_i, b_i) para i = 1, 2,...,m donde m>2. Si no hubiera error en las mediciones y los datos fueran realmente lineales, entonces para algún r_0 y r_1 se tendría:

$$b_i = r_0 + r_1\, a_i \quad \text{para } i = 1,2,\dots,m$$

Estas m ecuaciones en las 2 incógnitas r_0 y r_1 forman un sistema sobredeterminado de ecuaciones que probablemente no tenga solución. Pero nuestros puntos satisfacen realmente un sistema de aproximaciones lineales que se pueden expresar como:

$$\begin{bmatrix} b1 \\ b2 \\ . \\ . \\ bn \end{bmatrix} \approx \begin{bmatrix} 1 & a1 \\ 1 & a2 \\ \multicolumn{2}{c}{\cdots\cdots} \\ \multicolumn{2}{c}{\cdots\cdots} \\ 1 & an \end{bmatrix} \begin{bmatrix} r0 \\ r1 \end{bmatrix} \qquad (6.1)$$

O simplemente $b \approx Ar$.

Tratamos de encontrar un vector solución óptimo r para el sistema (6.1). Para cada vector r el vector error Ar - b mide a qué distancia está el sistema (6.1) de un sistema con solución r. Las normas de las componentes del vector Ar - b representan las distancias $d_i = r_0 + r_1 a_i - b_i$ como se muestra en la figura 6.6

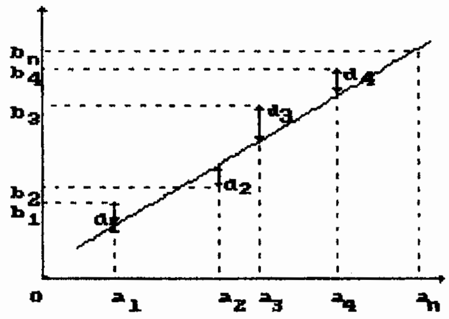


FIGURA 6.6

queremos entonces minimizar nuestro vector de error Ar - b. Minimizaremos la longitud $\|Ar - b\|$ del vector error, pero esto equivale a minimizar $\|Ar - b\|^2$ o lo que es lo mismo, minimizar

$$d_1^2 + d_2^2 + \ldots + d_m^2$$

de aquí el nombre de mínimos cuadrados.
Si a_1 y a_2 denotan las columnas de A en el sistema (6.1) entonces el vector $Ar = r_0 a_1 + r_1 a_2$ está en el espacio columna de A.

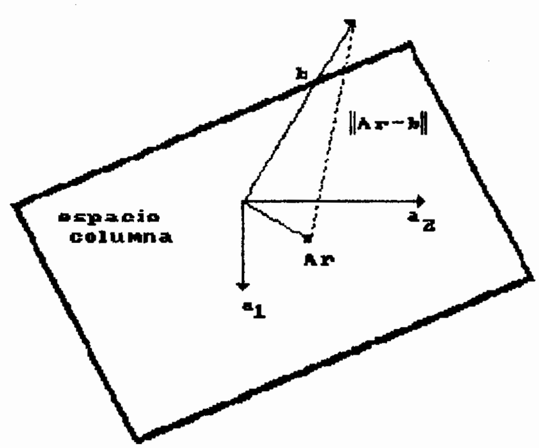


FIGURA 6.7

Entonces de la figura 6.7, resulta claro que de todos los vectores Ar en el espacio columna el que minimiza $\|Ar - b\|$ es la proyección b_s = Ar de b en el espacio columna S. Entonces $Ar = A(A^tA)^{-1}A^t\,b$; y de aqui se tiene que el vector solución r que es óptimo está dado por:

$$r = (A^tA)^{-1}\,A^t\,b.$$

Ahora si, volviendo a los datos del problema, hacemos el ajuste de mínimos cuadrados por medio de una recta, esto es:

$$y = r_0 + r_1\,x$$

Formamos el sistema $y \approx Ar$

$$\begin{bmatrix} 2.5 \\ 4.5 \\ 7 \\ 8.5 \end{bmatrix} \approx \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 15 \\ 16 \end{bmatrix} = \begin{bmatrix} 4 & 17 \\ 17 & 81 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 1/35 \end{bmatrix} \begin{bmatrix} 81 & -17 \\ -17 & 4 \end{bmatrix}$$

entonces $r = (A^tA)^{-1}\,A^t\,b = [-0.9,\ 1.5]^t$

de donde la recta que mejor ajusta los datos está dada por la ecuación
$y = -0.9 + 1.5\,x$

EJEMPLO 6.24

Si se varía la carga que se aplica a una estructura y se mide la deformación que produce donde x es la carga, y es la lectura del medidor de la deformación. A menos que la carga sea tan grande que el material se haga plástico lo normal en la teoría de la elasticidad es la relación lineal $y = r_0 + r_1\,x$. Dadas las siguientes mediciones encuentre la recta que mejor se ajusta.

x	0	1	3	4
y	0	1	2	5

Solución:

Formamos el sistema y = Ar

$$\begin{bmatrix} 0 \\ 1 \\ 2 \\ 5 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 4 & -17 \\ 17 & 81 \end{bmatrix}$$

$$\left(A^tA\right)^{-1} = \left(\tfrac{1}{20}\right)\begin{bmatrix} 13 & -4 \\ -4 & 2 \end{bmatrix}$$

Entonces : $r = (A^tA)^{-1} A^t b = [-0.2, 1.1]^t$

de donde, la recta que mejor ajusta los datos está dada por la ecuación
$y = -0.2 + 1.1\, x$

<u>EJEMPLO 6.25</u>

Una población de conejos de una gran isla se estimó todos los años desde 1981 hasta 1984 y se obtuvieron los datos que se listan en la siguiente tabla:

ai = año observado	1	2	3	4
bi = # de conejos en unidades de 1000	3	4.5	8	17
zi = Ln bi	1.1	1.5	2.08	2.83

Sabiendo que el crecimiento de la población es exponencial en ausencia de enfermedades, depredadores, hambre, etc. se espera que una función exponencial

$$b_i = re^{sa_i} \quad \text{o} \quad b = re^{sa}$$

sea la mejor representación de los datos. Encuentre la función exponencial que mejor ajusta los datos y con ella haga una proyección de la población de la isla para 1991.

Solución:

Observe que usando logaritmos es posible convertir esta función exponencial a la forma lineal

$$\text{Ln } b = \text{Ln } r + s(a)$$

de donde en la tabla se agrega el renglón de z = ln (b_i)

Entonces como es lineal la función así escrita se usa

$$x = (A^tA)^{-1}A^t b$$

donde

$$\begin{bmatrix} 1.1 \\ 1.5 \\ 2.08 \\ 2.83 \end{bmatrix} \approx \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}$$

$$[A^tA]^{-1} = \begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 1/5 \end{bmatrix}$$

entonces $(A^tA)^{-1}\, A^t\, b = [0.435, 0.577]^t = [\ln r, s]^t$

por lo tanto ln b = ln r + s(a_i) = 0.435 + 0.577 a de donde

$$b = re^{sa} = e^{0.435}\, e^{0.577a} = 1.54\, e^{0.577a}$$

en términos de funciones se tiene $f(x) = 1.54\, e^{0.577x}$

A partir de esta función se puede proyectar la población de conejos para 1991 haciendo f(11) 1000 = 570 778 conejos.

APROXIMACIÓN CUADRÁTICA

Ahora, pretendemos ajustar una curva cuadrática a los n puntos de datos. Recuerde que una cuadrática en x es cualquier expresión de la forma:

$$y = a + bx + cx^2$$

que representa a una parábola en el plano. Si los n puntos dados estuvieran en la parábola se tendría:

$$\begin{aligned} y_1 &= a + bx_1 + cx_1^2 \\ y_2 &= a + bx_2 + cx_2^2 \\ &\ldots\ldots\ldots\ldots \\ &\ldots\ldots\ldots\ldots \\ y_n &= a + bx_n + cx_n^2 \end{aligned} \tag{6.2}$$

para

$$y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix} \qquad A = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 \end{bmatrix} \qquad r = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

entonces (6.2) se puede escribir como:

$$y = Ar$$

igual que antes, si los puntos dados no están todos en la parábola entonces $y - Ar \neq 0$ y se tiene que para cualquier r nuestro problema vuelve a ser:

Encontrar un vector r en R^3 tal que $[y - Ar]$ sea mínimo y usando un razonamiento similar al anterior se tiene:

$$r = (A^tA)^{-1} A^t y$$

EJEMPLO 6.26

En una exhibición reciente de yates se hicieron las observaciones listadas en la tabla siguiente, donde se relacionan los precios b_i de las embarcaciones y sus pesos a_i

ai = peso en tons.	2	4	5	8
bi = precio en unidadesde 10 000	1	3	5	12

Al localizar los puntos de datos (a_i, b_i) como se muestra en la figura 6.8 podemos esperar que una función cuadrática de la forma

$$y = f(x) = r_0 + r_1 x + r_2 x^2$$

ajuste bien los datos

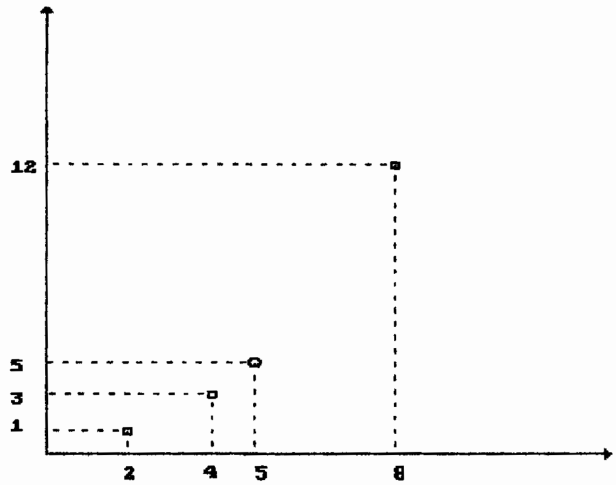


FIGURA 6.8

Solución:

Formamos el sistema y = Ar

$$\begin{bmatrix} 1 \\ 3 \\ 5 \\ 12 \end{bmatrix} \approx \begin{bmatrix} 1 & 2 & 14 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 8 & 64 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 4 & 19 & 109 \\ 19 & 109 & 709 \\ 109 & 709 & 4993 \end{bmatrix}$$

$$\left(A^tA\right)^{-1} = \left(1/_{5400}\right) \begin{bmatrix} 12744 & -3348 & -6624 & 2628 \\ -4464 & 2538 & 3744 & -1818 \\ 360 & -270 & -360 & 270 \end{bmatrix}$$

entonces $y = (A^tA)^{-1} A^t b = [0.207, 0.01, 0.183]^t$

Así la función cuadrática que mejor aproxima los datos en el sentido de los mínimos cuadrados es:

$$y = 0.207 + 0.01x + 0.183x^2$$

para graficar esta función se hace:

a_i	b_i	$f(a_i)$
2	1	0.959
4	3	3.175
5	5	4.832
8	12	11.999

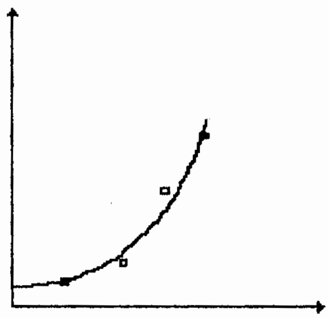


FIGURA 6.9

6.8 LA PSEUDOINVERSA Y LA DESCOMPOSICIÓN EN VALOR SINGULAR

En este punto cabe hacernos la siguiente pregunta: ¿Cuál es la solución óptima x para el sistema inconsistente Ax = b? deberíamos tener una regla que especifique x dada cualquier matriz A y cualquier lado derecho b.

Para cada b, Ax debe estar en el espacio de las columnas de A ya que es una combinación de las columnas ponderadas por las componentes de x. Por lo tanto, la elección óptima Ax es el punto p en el espacio columna que está más cerca a la b dada. En otras palabras, tenemos que proyectar a b sobre el espacio columna:

$$Ax = b_s = p$$

Claramente x está determinada cuando hay sólo una combinación de las columnas de A que producen p, los pesos de esta combinación serán las componentes de x. Si se tiene solución única existen varias condiciones equivalentes:

a) las columnas de A son linealmente independientes
b) el espacio nulo de A solo contiene al cero.
c) el rango de A es n
d) la matriz cuadrada $A^T A$ es invertible.

Si estas condiciones no son válidas, entonces x no está determinada en forma única y

$$x = (A^TA)^{-1} A^T b$$

es la expresión de dicha solución. Si A es una matriz invertible entonces x coincide con la única solución al sistema original

$$x = A^{-1} (A^T)^{-1} A^T b = A^{-1} b \qquad (6.3)$$

pero si A no es invertible definimos la pseudoinversa A+ (o inversa de Moore-Penrose), por lo tanto, si A es invertible $A+ = A^{-1}$.

Cuando la matriz A cumple con las cuatro condiciones anteriores, la pseudoinversa es la inversa izquierda que aparece en (6.3):

$$A+ = (A^T A)^{-1} A^T$$

pero la pseudoinversa queda por definirse cuando no son válidas las condiciones a)-d) y x no está determinada en forma única por Ax = p. Tenemos que escoger uno de los muchos vectores que satisfacen Ax = p y la elección será la solución óptima x = a + b al sistema inconsistente Ax = b.

Esta elección se efectúa de acuerdo a la regla siguiente: la solución óptima de entre todas las soluciones de Ax = p es aquella con longitud mínima. La clave para encontrarla es recordar que el espacio fila y el espacio nulo de A son complementos ortogonales en R^n. Esto significa que cualquier vector puede descomponerse en dos piezas perpendiculares, su proyección sobre el espacio fila y su proyección sobre el espacio nulo. Suponiendo que aplicamos esta descomposición a una de las soluciones x_0 de la ecuación Ax = p. Entonces $x_0 = x_r + x_n$ donde x_r está en el espacio fila y x_n está en el espacio nulo. Hay ahora tres puntos importantes:

1.- La componente x_r es una solución de Ax = p ya que $Ax_n = 0$,

$$Ax_0 = A(x_r + x_n) = Ax_r = p$$

2.- Todas las soluciones de Ax = p comparten ésta misma componente x_r en el espacio fila y difieren solamente en la componente x_n en el espacio nulo. La solución general es la suma de una solución particular x_r y una solución arbitraria x_n de la ecuación homogénea.

3.- La longitud de dicha solución $x_r + x_n$ obedece la ley de Pitágoras, ya que las dos componentes son ortogonales:

$$\|x_r + x_n\|^2 = \|x_r\|^2 + \|x_n\|^2$$

De lo anterior se desprenden las siguientes conclusiones:

La solución que tiene longitud mínima es x_r. Deberíamos elegir como cero a la componente en el espacio nulo, dejando una solución situada completamente en el espacio fila.

La solución óptima en mínimos cuadrados para cualquier sistema Ax = b es el vector x_r determinado por las dos condiciones:

1. Ax es igual a la proyección de b sobre el espacio columna de A.
2. x está en el espacio fila de A.

La matriz que resuelve Ax = b es la pseudoinversa A+ definida por $x = A^{+}b$.

Una mejor manera de entender la pseudoinversa es viéndola geométricamente, como se muestra en la figura 6.10.

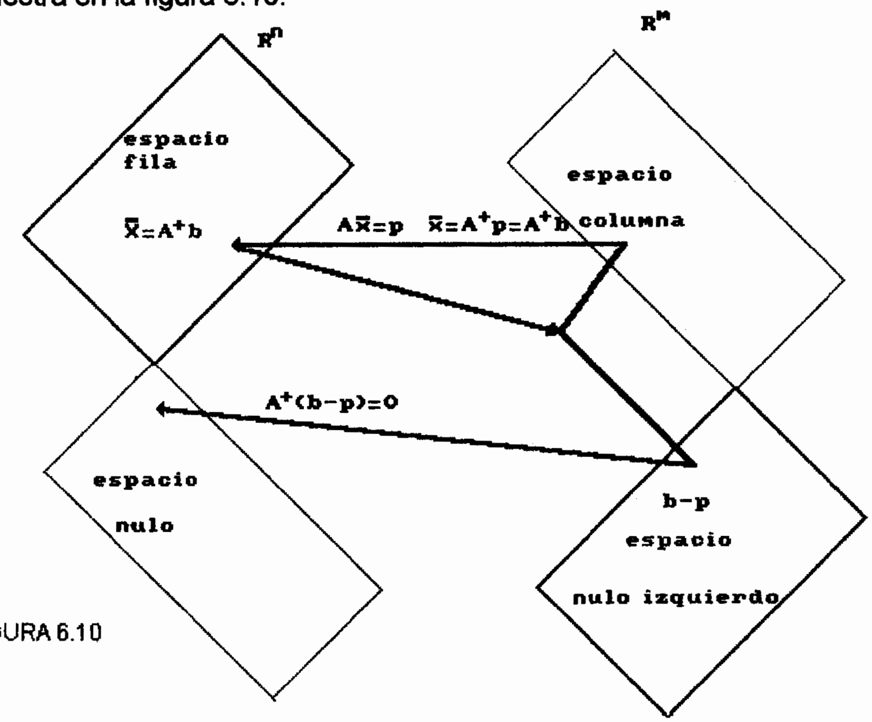


FIGURA 6.10

La matriz A+ combina el efecto de dos pasos separados: proyecta a b sobre el punto p y después encuentra el único vector x en el espacio fila que resuelve Ax = p. Un caso extremo es cuando b es perpendicular al espacio columna, en otras palabras, cuando b está en el espacio nulo izquierdo. Entonces p = 0 y x = 0, A+ envía todo b-p a cero. En el otro extremo tenemos a b dentro del espacio columna entonces p = b y encontramos x al "invertir" A (ya se había dicho que A es invertible si la consideramos como una aplicación de su espacio fila en su espacio columna con A+ como la inversa).

Una b arbitraria está entre estos dos extremos: la componente p se invierte para dar x y la otra componente b-p se aniquila. De esta descripción y de la figura 6.10 podemos listar algunas propiedades básicas de la pseudoinversa:

1. A+ es una matriz de nxm comienza con el vector $b \in R^m$ y produce el vector $x \in R^n$.
2. El espacio columna de A+ es el espacio fila de A y el espacio fila de A+ es el espacio columna de A, de aqui que rango A = rango A+.
3. La pseudoinversa de A+ es A.
4. En general AA+ $\neq$ I, ya que es posible que A no tenga inversa derecha, pero AA+ siempre es igual a la proyección P sobre el espacio columna:

$$AA^{+}b = Ax = p = Pb$$
$$AA^{+} = P$$

EJEMPLO 6.27

Sea la matriz A una matriz no invertible:

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

Observe que ésta no es otra que la matriz de proyección que manda cualquier punto de R^3 al plano x-y. De esta forma el espacio columna y el espacio fila de A coinciden con el plano x-y en R^3 que contiene todos los vectores (x, y, 0). El espacio nulo es el eje z, que es ortogonal al espacio fila. Para encontrar x se proyecta b sobre el espacio columna:

Sí

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \text{ entonces } \quad p = pb = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

y resolviendo Ax = p

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

se tiene:

$x_1 = b_1$ $\qquad$ $x_2 = b_2$ $\qquad$ y x_3 es arbitraria

Se elige entre esta familia infinita de soluciones aquella que tenga longitud mínima, donde la tercer componente debe ser cero, esto deja a $x = (b_1, b_2, 0)^T$ que está en el espacio fila (el plano x-y) y la pseudoinversa A+ está dada por A misma. Esto sucede porque A actúa como la matriz identidad en la manera como aplica su espacio fila en su espacio columna y la pseudoinversa ignora todo lo demás. Resumiendo, se tiene que la solución óptima de Ax = b que es un conjunto inconsistente

$$\begin{aligned} x_1 + 0x_2 + 0x_3 &= b_1 \\ 0x_1 + x_2 + 0x_3 &= b_2 \\ 0x_1 + 0x_2 + 0x_3 &= b_3 \end{aligned}$$

consiste en satisfacer las dos primeras ecuaciones y hacer $x_3 = 0$.

6.9 RESUMEN

1. La dimensión de los cuatro espacios asociados a una matriz es

1.	$R(A^t)$ = espacio renglón de A:	dimensión = r
2.	N(A) = espacio nulo de A:	dimensión = n-r
3.	R(A) = espacio columna de A:	dimensión = r
4.	$N(A^t)$ = espacio nulo izquierdo de A:	dimensión = m-r

tal que:

$$n = \dim R(A) + \dim N(A)$$
$$m = \dim R(A^t) + \dim N(A^t)$$

2. El sistema Ax = b tiene al menos una solución x para toda b si y sólo si las columnas de A generan R^m, entonces r = m. En este caso existe una inversa derecha H de nxm tal que $AH = I_m$, la matriz identidad de orden m. Esto es posible sólo si m£n.

El sistema Ax = b tiene a lo sumo una solución x para cada b si y sólo si las columnas son linealmente independientes, entonces r=n. En este caso existe una inversa izquierda G de nxm tal que $GA = I_n$, la matriz identidad de orden n. Esto es posible sólo si m³n.

3. Dos subespacios V y W del mismo espacio R^n son ortogonales si cada vector veV es ortogonal a cada vector weW: $v^Tw = 0$ para toda v y w.

4 .Dado un subespacio V de R^n, el espacio de todos los vectores ortogonales a V es el complemento ortogonal de V y se denota por V^L . Asi se tiene entonces

$$N(A) = (R(A^T))^L, \quad R(A^T) = N(A)^L$$
$$N(A^T) = (R(A))^L, \quad R(A) = (N(A^T))^L$$

5. Un producto interior en un espacio vectorial V es una función que a cada par de vectores x y y en V asocia un número real <x, y> de tal manera que se satisfacen los siguientes axiomas, para todos los vectores x, y, z en V y todos los escalares k.

1.	<x, y> = <y, x>	axioma de la simetría
2.	<x+y, z> = <x, z> + <y, z>	axioma de la aditividad
3.	<kx, y> = k <x, y>	axioma de la homogeneidad
4.	<x, x> ³ 0 y <x, x> = 0 $\Leftrightarrow$ x = 0	axioma de positividad.

6. Es posible convertir cualquier conjunto de vectores independientes a_1, a_2,...,a_n en un conjunto de vectores ortogonales mediante el proceso de Gram-Schmidt: primero se fija $a_1 = v_1$ después cada v_i es ortogonal a las v_1,..., v_{i-1} precedentes:

$$v_i = a_i - \frac{\langle v_1, a_i \rangle}{\|v_1\|^2} v_1 - \dots - \frac{\langle v_{i-1}, a_i \rangle}{\|v_{i-1}\|^2} v_{i-1}$$

7. Sea Ax = b un sistema lineal de m ecuaciones con n incógnitas, donde m>n (un sistema sobredeterminado) y el rango de A es n. La solución de mínimos cuadrados del sistema correspondiente Ax » b de aproximaciones lineales es el vector x = r que minimiza la magnitud del vector error.

Es decir, de todos los vectores Ar en el espacio columna el que minimiza °Ar - b° es la proyección b_s = Ar de b en el espacio columna S. Entonces $Ar = A(A^tA)^{-1}A^t\, b$; y de aquí se tiene que el vector solución r que es óptimo está dado por:

$$r = (A^tA)^{-1} A^t\, b.$$

6.10 NOTAS HISTÓRICAS

El rango de una matriz lo definió en 1879 Georg Frobenius (1849-1917) como sigue: si se anulan todos los determinantes de grado (r+1), pero no todos los de grado r, entonces r es el rango de la matriz. Frobenius usó este concepto para tratar las cuestiones de formas canónicas para ciertas matrices de enteros y las soluciones de ciertos sistemas de congruencias lineales.

Por otro lado, James Sylvester definió en 1884 la nulidad para matrices cuadradas, como sigue: la nulidad de una matriz de nxn es i si todo menor (determinante) de orden n-i+1 (y, por tanto, de todo orden superior) es igual a 0 e i es el mayor de los números para los cuales esto es cierto. Aqui, Sylvester estaba interesado, como en buena parte de su carrera matemática, en descubrir invariantes, esto es, propiedades de objetos matemáticos particulares que no cambian bajo tipos específicos de transformación. El procedió a probar lo que llamó una de las leyes cardinales en la teoría de matrices: que la nulidad del producto de dos matrices no es menor que la nulidad de cualquier factor ni mayor que la suma de las nulidades de los factores.

El proceso de Gram-Schmidt debe su nombre al matemático danés Jorge P. Gram (1850-1916), y al alemán Erhard Schmidt (1876-1959). Lo publicó primero Gram en 1883 en un artícluo titulado "Desarrollo de series usando el método de los mínimos cuadrados", y Schmidt lo publicó de nuevo en 1907, con una cuidadosa demostración, en un trabajo sobre ecuaciones integrales. De hecho Schmidt incluso se refirió al trabajo de Gram. Para Schmidt igual que para Gram, los vectores eran funciones continuas definidas en un intervalo [a,b] con el producto interno de dos funciones s y F dado por la integral de su producto en ese intervalo. Sin embargo Schmidt fue más explícito que Gram al escribir el proceso con gran detalle y probar que el conjunto de funciones F_i derivado de su conjunto original s_i, era de hecho, un conjunto ortonormal.

Schmidt, que estuvo en la Universidad de Berlín desde 1917 hasta su muerte, es más conocido por su trabajo decisivo sobre los espacios de Hilbert, espacios de sucesiones de números complejos de cuadrado sumable. De hecho aplicó el proceso de Gram-Schmidt a conjuntos de vectores en esos espacios para ayudar a desarrollar condiciones necesarias y suficientes para que dichos conjuntos sean linealmente independientes.

Una técnica muy cercana a la de los mínimos cuadrados es la que desarrolló Roger Cotes (1682-1716), el genial matemático que editó la segunda edición de los **Principia de Isaac Newton**, en una obra que trataba errores en las observaciones astronómicas, escrita alrededor de 1715.

Sin embargo, fue Carl Gauss, a los 16 años el primero en formular el principio completo, mientras ajustaba aproximaciones relacionadas con la distribución de los números primos. Más tarde Gauss, afirmó que durante años había usado a menudo el método, por ejemplo en sus cálculos sobre las órbitas de asteroides. Gauss publicó el método en 1809 e hizo una exposición definitiva 14 años después.

Por otro lado, debemos a Adrien-Marie Legendre (1752-1833), fundador de la teoría de las funciones elípticas, la primer publicación del método de los mínimos cuadrados en un trabajo de 1806, sobre la determinación de las órbitas de cometas. Después de la publicación de Gauss en 1809, Legendre le escribió censurándolo por reclamar el método como propio. Todavía en 1827, Legendre seguía acusando a Gauss de apropiarse de los descubrimientos de otros. De hecho el problema radicaba en el fallo de Gauss de no publicar a tiempo sus descubrimientos; solamente los mencionaba cuando ya habían sido publicados por otros.

CAPÍTULO 7
VALORES Y VECTORES PROPIOS

7.1 INTRODUCCIÓN

En este capítulo veremos lo que se puede considerar la segunda parte de teoría de matrices, para ello es necesario entender lo que son los valores propios y cómo pueden ser útiles. Una de sus aplicaciones es la solución de un sistema de ecuaciones diferenciales ordinarias, donde al parecer la clave se encuentra en sus valores y vectores propios. Quizá el ejemplo más sencillo es el de los soldados que cruzan un puente. Cada vez que lo hacen dejan de marchar y lo cruzan sólo caminando. La razón es que podrían marchar a una frecuencia igual a uno de los valores propios del puente y entonces comenzaría a oscilar. Un ingeniero trata de que las frecuencias naturales de su puente o de su cohete estén alejadas de las del viento o el chapoteo de la gasolina. En el otro extremo, un corredor de bolsa pasa su vida tratando de estar en correspondencia con las frecuencias naturales del mercado. Los valores y los vectores propios son los rasgos más importantes de prácticamente cualquier sistema dinámico.

Otro ejemplo interesante es la rotación de la Tierra. Siempre habrá alguna dirección que permanezca sin cambio, a saber el eje de rotación. No necesariamente es aquel alrededor del cual la Tierra gira en la realidad, pero deberá haber algún polo norte y polo sur que permanezcan fijos. Estos polos son los vectores propios, con valores propios iguales a 1. En general todos los otros puntos se mueven y no vemos más vectores propios. Las únicas excepciones suceden cuando la rotación es de 360° (con vectores propios por donde sea) o de 180°. En el caso de 180°, el plano del ecuador está lleno de vectores propios; cada dirección del plano se ha invertido y el valor propio es −1. Este plano ecuatorial es un caso de un "espacio propio bidimensional"; el único valor propio $\lambda = 1$ tiene dos vectores propios independientes y, por lo tanto, un plano de vectores propios. En general los valores propios no son ± 1 y los vectores usualmente se estiran o comprimen. Pero lo más importante es que tomando una rotación en el espacio tridimensional $\mathbb{R}^3$, y exceptuando 180° y 360°; sólo hubo una recta de vectores propios cuando esperábamos tres. Como se puede ver la gama de aplicaciones del tema que aborda este capítulo es muy amplia y es conveniente hacer algunas definiciones.

7.2 VALORES Y VECTORES PROPIOS

Sea $T: V \rightarrow V$ una transformación lineal. En una gran variedad de aplicaciones resulta útil encontrar un vector $v \in V$ tal que Tv y v sean paralelos. Esto es, buscamos un vector v y un escalar α tal que:

$$Tv = \alpha v \qquad (7.1)$$

Si v $\neq$ 0 y α satisface (7.1), entonces α se conoce como un valor propio ó característico de T y v es un vector propio o característico de T correspondiente al valor propio α.

Definición 7.1

Si A es una matriz de nxn, entonces un escalar α es un valor propio de A sí hay un vector v distinto de cero tal que:

$$Av = \alpha v$$

El vector v $\neq$ 0 se llama un vector propio de A correspondiente a un valor propio α.

<u>EJEMPLO 7.1</u>

Sea A tal que:

$$A = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}$$

entonces:

$$A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

De esta forma α_1 = 1 es un valor propio (característico) de A, correspondiente al vector propio (característico) $v_1 = [2, 1]^t$, análogamente:

$$A\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -6 \\ -4 \end{bmatrix} = -2\begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Por lo que α_2 = -2 es un valor propio de A con su correspondiente vector propio $v_2 = [3, 2].^t$

Sea α un valor propio de A, entonces existe un vector no nulo $v = (x_1,...,x_n)^t \neq 0$, tal que,

$$Av = \alpha v = \alpha Iv \quad o$$
$$(A - \alpha I)v = 0 \qquad (7.2)$$

Si A es de orden nxn, entonces la ecuación (7.2) es un sistema homogéneo de n ecuaciones con n incógnitas, como por hipótesis (ya que v es un vector propio distinto de cero), el sistema tiene soluciones no triviales, entonces: det $(A - \alpha I) = 0$

A esta ecuación se le denomina la **ecuación característica de A.**

Si A = (a_{ij}) entonces la ecuación anterior se puede escribir

$$\begin{bmatrix} a_{11}-\alpha & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\alpha & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn}-\alpha \end{bmatrix} = 0$$

Si desarrollamos el determinante obtenemos una expresión polinomial p(α) de grado n con coeficientes que incluyen a_{ij}, es decir:

$$\det(A - \alpha I) = p(\alpha)$$

El polinomio p(α) es el polinomio característico de la matriz A. Los valores propios de A son precisamente las soluciones de la ecuación característica p(α) = 0.

EJEMPLO 7.2

Sea A = I, entonces para todo vector $v \in R^n$, Av = Iv = v. De esta manera I es el único valor característico de A y cada vector v es un vector característico de I.

EJEMPLO 7.3

Encuentre los vectores propios de la matriz A:

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

Solución:

El polinomio característico de A es:

$$\det(A - \alpha I) = \begin{vmatrix} 3 - \alpha & 2 \\ 2 & -\alpha \end{vmatrix} = \alpha^2 - 3\alpha - 4 = 0$$

$\alpha^2 - 3\alpha - 4 = (\alpha - 4)(\alpha + 1) = 0$, de donde, $\alpha_1 = -1$ y $\alpha_2 = 4$ son valores propios de A.

EJEMPLO 7.4

Encuentre los valores propios de la matriz A:

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

Solución:

$$\det(A-\alpha I)=\begin{vmatrix} 2-\alpha & 1 & 0 \\ -1 & -\alpha & 1 \\ 1 & 3 & 1-\alpha \end{vmatrix} = -(\alpha-2)(\alpha^2-\alpha-2) = -(\alpha-2)(\alpha-2)(\alpha+1)$$

$-(\alpha-2)(\alpha-2)(\alpha+1) = 0$, de donde, $\alpha_1 = -1$ y $\alpha_2 = \alpha_3 = 2$ son valores propios de A.

En general si se tiene una matriz A de orden nxn

$$A=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = 0$$

entonces:

$$p(\alpha) = \det(A-\alpha I) = \begin{vmatrix} a_{11}-\alpha & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\alpha & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn}-\alpha \end{vmatrix} = 0$$

y $p(\alpha)$ se puede escribir como:

$$p(\alpha) = \alpha^n + b_{n-1}\,\alpha^{n-1} + \ldots + b_1\,\alpha + b_0 = 0 \qquad (7.3)$$

La ecuación (7.3) tiene n raíces, varias de las cuales pueden repetirse. Si $\alpha_1,\ldots,\alpha_m$ son las diferentes raíces de (7.3) con multiplicidades $r_1, r_2,\ldots,r_m$, entonces (7.3) se puede factorizar para obtener:

$$p(\alpha) = (\alpha-\alpha_1)^{r1}\,(\alpha-\alpha_2)r^2\ldots(\alpha-\alpha_m)^{rm} = 0$$

Los números $r_1, r_2,\ldots,r_m$ se llaman **multiplicidades algebraicas** de los valores propios de
$\alpha_1,\ldots,\alpha_m$.

EJEMPLO 7.5

Todo polinomio de grado n con coeficientes reales o complejos tiene n raíces exactamente (contando multiplicidades), por ejemplo; el polinomio $(\alpha - I)^5$ tiene cinco raíces todas iguales al número uno. Puesto que todo valor característico de A es una raiz de la ecuación característica de A se tiene:

Definición 7.2
Si contamos multiplicidades, cada matriz de nxn tiene exactamente n valores característicos.

TEOREMA 7.1
Si A es una matriz de orden nxn, entonces las siguientes afirmaciones son equivalentes:

a) α es un valor característico de A.
b) El sistema (A - αI)v = 0 tiene soluciones no triviales.
c) Existe un vector v e n R^n v $\neq$ 0, tal que Av = αv

Si α es un valor característico de A, entonces el espacio solución del sistema de ecuaciones (A - αI)v = 0 se denomina el espacio característico de A correspondiente a α, y los vectores diferentes de cero en el espacio característico, se denominan los vectores característicos de A correspondientes a α.

INTERPRETACIÓN GEOMÉTRICA

Si v es un vector característico de A correspondiente a α, entonces Av = αv. Por lo tanto, la multiplicación por A mapea a v en un múltiplo escalar de si mismo; por consiguiente dependiendo del valor de α, la multiplicación por A dilata a v, lo contrae o invierte su dirección, ver figura 7.1.

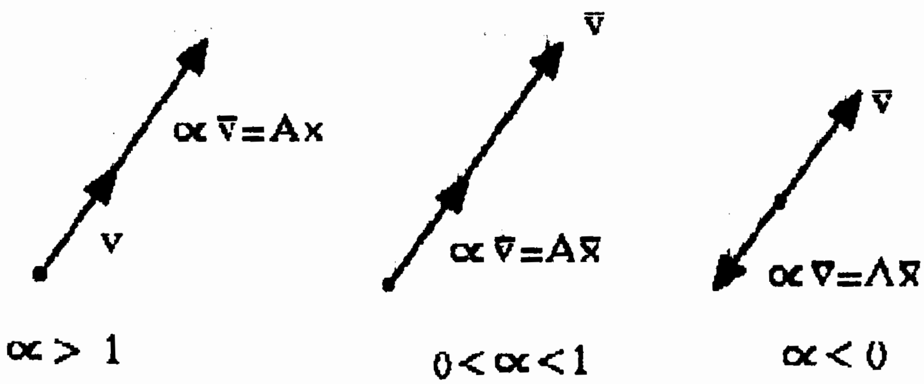


Figura 7.1

EJEMPLO 7.6
Sea la matriz A como sigue:

$$A = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$$

Encuentre los valores y los vectores propios.

Solución
Sea:

$$\det(A - \alpha I) = \begin{vmatrix} 4-\alpha & 1 \\ 0 & 4-\alpha \end{vmatrix} = (\alpha - 4)^2 = 0$$

$\alpha = 4$ es el valor característico de multiplicidad 2, de donde:

$$(A - 4I)V = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ 0 \end{bmatrix}$$

Lo que implica que $v_1 = [1, 0]^t$ es un vector propio.

EJEMPLO 7.7
Sea la matriz A como sigue:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

Encuentre los valores y vectores característicos de A.

Solución:
Sea

$$\det(A - \alpha I) = \begin{vmatrix} 4-\alpha & 0 \\ 0 & 4-\alpha \end{vmatrix} = (\alpha - 4)^2 = 0$$

$\alpha = 4$ es el valor característico de multiplicidad 2, de donde: $Av = 4v$
Lo que implica que $v_1 = [1, 0]$ y $v_2 = [0, 1]$ generan el espacio característico de A.

Definición 7.3
Sea α un valor característico de A. El subespacio E_α, se denomina el espacio característico de A correspondiente al valor característico α.

TEOREMA 7.2

Sea α un valor característico de la matriz A de orden nxn y sea $E^{\alpha} = \{v : Av = \alpha v\}$, entonces E^{α} es un subespacio de R^n (en general de E^n).

Demostración

Si $Av = \alpha v$,entonces $(A - \alpha I)\, v = 0$; de donde, E_{α} es el espacio nulo de la matriz $A - \alpha I$, la cual es un subespacio de R^n.

TEOREMA 7.3

Sea A una matriz de nxn y sea $\alpha_1, \alpha_2, ..., \alpha_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $v_1, v_2, ..., v_m$. Entonces v_i son linealmente independientes o bien, los vectores característicos correspondientes a valores característicos diferentes son linealmente independientes. (la demostración se hace por inducción).

EJEMPLO 7.8

Sea la matriz A, como sigue:

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

Encuentre los valores y vectores propios, así como el espacio característico correspondiente.

Solución:

Sea $\det(A - \alpha I) = -\alpha^3 + 6\alpha^2 + 15\,\alpha + 8 = -(\alpha + 1)^2 (\alpha - 8) = 0$

Lo que implica $\alpha_1 = 8$ y $\alpha_2 = -1$ con multiplicidad algebraica 2, para $\alpha_1 = 8$ se obtiene: $v_1 = [2, 1, 2]^t$ y $E_8 = \{[2, 1, 2]\}$, para $\alpha_2 = -1$ se obtiene: $v_2 = [1, -2, 0]^t$ y $v_3 = [0, -2, 1]^t$ con

$$E_{-1} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \right\}$$

En los ejemplos que hemos visto, encontramos un valor característico con una multiplicidad algebraica de dos o más. Sin embargo, el número de vectores característicos linealmente independientes no necesariamente es igual a la multiplicidad algebraica del valor característico (como en el ejemplo 7.6). Esto se puede corroborar con la siguiente definición.

Definición 7.4
Sea α un valor característico de A, entonces la multiplicidad geométrica de α es la dimensión del espacio característico correspondiente a α (que es la nulidad de la matriz A - αI) esto es:

$$\text{multiplicidad geométrica de } \alpha = \dim E_\alpha$$

Si A es una matriz de 2x2 y α un valor característico con multiplicidad algebraica 2, entonces la multiplicidad geométrica de α es menor o igual a 2; puesto que puede haber al menos dos vectores l.i. en un espacio de dos dimensiones.

Si A es de 3x3 con dos valores característicos α_1 y α_2 de multiplicidad algebraica 1 y 2 respectivamente, entonces la multiplicidad geométrica de α_2 es menor o igual a 2.

TEOREMA 7.4
Sea α un valor característico de A, entonces:

$$\text{multiplicidad geométrica de } \alpha \leq \text{multiplicidad algebraica de } \alpha$$

7.3 DIAGONALIZACIÓN

TEOREMA 7.5
Sea A una matriz de orden nxn, entonces A tiene n vectores característicos linealmente independientes si y solo si; la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, A tiene n vectores característicos l.i. si todos sus valores característicos son distintos (pues, si no, podrían ser menores que n).

Como ya se ha visto en varios ejemplos, algebraicamente y desde el punto de vista de los cálculos, el problema de los valores propios es mucho más difícil que Ax = b; sin embargo se pueden usar algunos resultados para que sea más fácil trabajar con ellos.

Definición 7.5
La suma de los n valores propios, es igual a la suma de las n entradas de la diagonal de A:

$$\alpha_1 + \ldots + \alpha_n = a_{11} + \ldots + a_{nn}$$

Esta suma es la **traza** de A. Además, el producto de los n valores propios es igual al determinante de A.

EJEMPLO 7.9
Sea la matriz simétrica:

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

con valores propios $\alpha_1 = 0$, $\alpha_2 = 1$ y $\alpha_3 = 3$ entonces:

$$0 + 1 + 3 = 1 + 2 + 1 = 4 \text{ y } \det A = 0$$

Nota: No deben confundirse los valores propios de una matriz y sus entradas diagonales. Normalmente son completamente diferentes; sin embargo, se tiene la siguiente definición:

Definición 7.6
Si la matriz A es triangular (puede ser superior y en particular diagonal), entonces los valores propios $\alpha_1,..,\alpha_n$ son exactamente los mismos que las entradas de la diagonal $a_{11},...,a_{nn}$.

EJEMPLO 7.10
Sea la matriz A como sigue:

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

Su polinomio característico es:

$$\det \begin{vmatrix} 1-\alpha & 1/4 & 0 \\ 0 & 3/4-\alpha & 0 \\ 0 & 0 & 1/2-\alpha \end{vmatrix} = (1-\alpha)(3/4-\alpha)(1/2-\alpha)$$

El determinante es el producto de las entradas diagonales. Obviamente, las raíces son:

$\alpha_1 = 1$, $\alpha_2 = 3/4$ y $\alpha_3 = ½$; los valores propios ya estaban colocados a lo largo de la diagonal principal.

Este ejemplo, en el cual, se pueden encontrar los valores propios por inspección, señala lo más importante de este tema: **transformar a una matriz A en una matriz diagonal o triangular sin cambiar sus valores propios.** (La factorización LU no sirve para este caso, ya que los valores propios de U no son los valores propios de A).

Otra situación en que los cálculos son fáciles es: si ya tenemos los valores y vectores propios de una matriz A, entonces los valores propios de A^2 son exactamente $\alpha_1^2, ..., \alpha_n^2$ y cada vector propio de A es también un vector propio de A^2.

TEOREMA 7.6

Supongamos que la matriz A de nxn tiene n vectores propios linealmente independientes; entonces, si se eligen estos vectores como las columnas de una matriz S, se sigue que:

$S^{-1}AS$ es una matriz diagonal Λ con los valores propios de A en su diagonal.

$$S^{-1}AS = \Lambda = \begin{bmatrix} \alpha_1 & & \\ & \alpha_2 & \cdots \\ & \vdots & \alpha_n \end{bmatrix}$$

Demostración.

Colocamos los vectores propios x_i en las columnas de S y se calcula el producto AS, una columna a la vez:

$$AS = A\begin{bmatrix} \vdots & \vdots & \vdots \\ x_1 & x_2 \cdots & x_n \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \alpha_1 x_1 & \alpha_2 x_2 \cdots & \alpha_3 x_n \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Entonces:

$$\begin{bmatrix} \vdots & \vdots & \vdots \\ \alpha_1 x_1 & \alpha_2 x_2 \cdots & \alpha_3 x_n \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ x_1 & x_2 \cdots & x_n \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \alpha_1 & & \\ & \alpha_2 & \cdots \\ & \vdots & \alpha_n \end{bmatrix} = S\Lambda$$

$$AS = S\Lambda \qquad ó \qquad S^{-1}AS = \Lambda \qquad ó \qquad A = S\Lambda S^{-1}$$

La matriz S es invertible, ya que supusimos que sus columnas (los vectores propios son l.i.).

Observación 1
Si la matriz A no tiene valores propios repetidos (los números α_1, ..., α_n son distintos), entonces los n vectores propios son independientes y; por lo tanto, cualquier matriz con valores propios distintos puede diagonalizarse.

Observación 2
La matriz diagonalizada S no es única, ya que un vector propio x puede multiplicarse por una constante y seguir siendo vector propio; por lo tanto, se pueden multiplicar las columnas de S por cualquier constante distinta de cero y producir una nueva diagonalización S.

Observación 3
La ecuación AS = SΛ es válida si las columnas de S son los vectores propios de A y no de otra manera. Supongamos que la primer columna de S es algún vector *y*, entonces la primer columna de SΛ es $\alpha_1 y$. Si esto va a corresponder con la primer columna de AS, que debido a la multiplicación de matrices es A*y*, entonces *y* debe ser un vector propio $Ay = \alpha_1 y$.

Observación 4
No todas las matrices poseen n vectores propios l.i. y por lo tanto, no todas las matrices son diagonalizables.

<u>EJEMPLO 7.11</u>
Sea la matriz A como sigue:

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

sus valores propios son: $\alpha_1 = \alpha_2 = 0$, ya que es triangular

$$\det(A - \alpha I) = \begin{vmatrix} -\alpha & 1 \\ 0 & -\alpha \end{vmatrix} = \alpha^2$$

Si x es un vector propio entonces:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{ó} \quad x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

Aunque $\alpha = 0$ es un valor propio de multiplicidad algebraica 2, solo tiene un espacio unidimensional de vectores propios. La multiplicidad geométrica de éste valor propio es 1 y no podemos construir S.

Como $\alpha_1 = \alpha_2 = 0$, *A* tendría que ser la matriz de ceros; pero si $S^{-1}AS = 0$, entonces:

$SS^{-1}AS = 0$; lo que implica que $ASS^{-1} = 0$ y, de aquí se deduce que $A = 0$; lo que resulta imposible. Por lo tanto, no existe S tal que $S^{-1}AS = A$

Algunas matrices con valores propios repetidos pueden diagonalizarse. Si se repite un valore propio α, m veces, entonces para que se pueda diagonalizar A debe haber m vectores propios correspondientes.

EJEMPLO 7.12

Diagonalice la siguiente matriz A:

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

Solución:

Los valores propios de A son: $\alpha_1 = 1$ y $\alpha_2 = 5$.

Para $\alpha_2 = 5$ se tienen los vectores propios:

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

y para $\alpha_1 = 1$ se tiene:

$$v_{31} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

de donde; $\{v_1, v_2, v_3\}$ es un conjunto de vectores l.i. entonces:

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

diagonaliza a A.

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

El orden de las columnas de S no altera los valores propios sólo su orden.

TEOREMA 7.7
Una matriz A de nxn es diagonalizable, si y sólo si, la multiplicidad algebraica de cada valor propio es igual a su multiplicidad geométrica.

7.4 PROPIEDADES DE VALORES Y VECTORES PROPIOS

Sea la matriz A de orden nxn

1. Si α es un valor propio de A con v como vector propio correspondiente, entonces Λ^k es un valor propio de A^k, de nuevo con v como vector propio correspondiente, para cualquier entero positivo k.

2. Si α es un valor propio de una matriz invertible A con v como vector propio correspondiente, entonces $1/\alpha$ es un valor propio de A^{-1}, de nuevo con v como vector propio correspondiente.

3. Si α es un valor propio de A, entonces el conjunto E es un subespacio de R^n.

COROLARIO 7.1
Sean A y S matrices de orden nxn, entonces $A^k = A\,\Lambda^k S^{-1}$ para cada entero positivo k.

Demostración
De $S^{-1}AS = \Lambda$ obtenemos $A = S\,\Lambda\,S^{-1}$ así:

$$A^k = \underbrace{(S\,\Lambda\,S^{-1})(S\,\Lambda\,S^{-1})\ \ldots\ (S\,\Lambda S^{-1})}_{\text{k veces}}$$

Como los teoremas adyacentes $S^{-1}S$ se cancelan queda:

$$A^k = S\,\Lambda^k S^{-1}$$

MATRICES SIMÉTRICAS Y DIAGONALIZACIÓN ORTOGONAL

TEOREMA 7.8
Sea A una matriz simétrica real de nxn, entonces los valores propios de A son reales. Para cualquier matriz simétrica, la multiplicidad algebraica de cada valor propio es

igual a su multiplicidad geométrica, de modo que toda matriz simétrica es diagonalizable.

Ya vimos que vectores característicos correspondientes a valores característicos diferentes son l.i. Para matrices simétricas el resultado es más fuerte; los vectores característicos de una matriz simétrica correspondientes a valores diferentes son ortogonales.

TEOREMA 7.9

Sea A una matriz simétrica real de nxn. Si α_1 y α_2 son valores característicos distintos con correspondientes vectores característicos reales v_1 y v_2, entonces v_1 y v_2 son ortogonales.

Demostración.

Calculamos

$$Av_1\, v_2 = \alpha_1 v_1\, v_2 = \alpha_1(v_1\, v_2)$$

y $$Av_1\, v_2 = v_1\, A^t v_2 = v_1\, Av_2 = v_1\,(\alpha_2 v_2) = \alpha_2(v_1 v_2)$$

entonces: $\alpha_1(v_1\, v_2) = \alpha_2(v_1\, v_2)$ como $\alpha_1 + \alpha_2$

y de aquí: $(v_1\, v_2) = 0$ i.q.q.d.

TEOREMA 7.10

Sea A una matriz simétrica real de nxn, entonces A tiene n vectores característicos reales ortonormales.

Este teorema nos dice, que si A es simétrica, entonces R^n tiene una base $B = \{u_1, u_2, \ldots, u_n\}$ de vectores característicos ortonormales de A. Sea Q la matriz cuyas columnas son $u_1, u_2, \ldots, u_n$, entonces Q es una matriz ortogonal.

Definición 7.7

Una matriz A de nxn se dice diagonalizable ortogonalmente, si existe una matriz ortogonal Q tal que:

$$Q^tAQ = \Lambda$$

donde: Λ = diag. $(\alpha_1, \ldots, \alpha_n)$ los valores característicos de A.

Como Q es ortogonal $\quad Q^t = Q^{-1}$, lo que implica $Q^{-1}AQ = \Lambda$

TEOREMA 7.11

Sea A una matriz real de nxn; entonces A es diagonalizable ortogonalmente si y solo si, A es simétrica.

Antes de ver ejemplos, enunciaremos los 3 pasos que sirven para encontrar la matriz ortogonal Q que diagonaliza a la matriz simétrica A.

1. Encuentre una base para cada espacio característico de A.
2. Encuentre una base ortonormal para cada espacio característico de A, usando el proceso de Gram-Schmidt.
3. Escriba Q como la matriz cuyas columnas son los vectores característicos ortonormales obtenidos en el paso 2.

EJEMPLO 7.13
Encuentre una matriz ortogonal Q que diagonalice a la matriz A :

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Solución:

$$\det(A - I) = \begin{vmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{vmatrix} = (\alpha-2)^2(\alpha-8) = 0$$

Por lo que: $\alpha_1 = 2$ y $\alpha_2 = 8$

Para $\alpha_1 = 2$, se tiene:

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Forman una base para E_2. Aplicando Gram-Schmidt se tiene:

$$v_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

El espacio característico correspondiente a $\alpha_2 = 8$, tiene como base:

$$v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Aplicando Gram-Schimdt a v_3 se tiene:

$$v'_3 \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

Diagonaliza ortogonalmente a A

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \qquad S = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

7.5 ECUACIONES EN DIFERENCIAS

Las ecuaciones en diferencias se dirigen a un número finito de pasos finitos, mientras que una ecuación diferencial lleva un numero infinito de pasos infinitesimales (sin embargo ambas teorías son paralelas).

EJEMPLO 7.14

Supongamos que se invierten \$100 por 5 años a un interés de 6%. Si se compone una vez al año, entonces el capital se multiplica por 1.06 y $P_{k+1} = 1.06P_k$ donde ésta, es una ecuación en diferencias con un lapso de tiempo de un año. Relacione el capital después de k+1 años con el capital del año anterior; después de 5 años el capital original B=1000 se ha multiplicado 5 veces y:

$$P_5=(1.06)^5 \ P_0 = (1.06)_5 \ 1000 = \$ \ 1338.$$

Si el lapso de tiempo se reduce a un mes:

$$P_{k+1} = (1 + 0.6/12)P_k.$$

y después de 5 años o 60 meses:

$$P_{60} = (1 + 0.06/12)^{60}\, p_0 = (1.005)^{60}\, 1000 = 1349$$

Si se compone diariamente el interés, se tiene:

$$(1+0.06/365)^{5.365}\, 1000 = \$\ 1349.83.$$

Si se pretende componer continuamente el interés, se añade a cada instante, entonces se tiene un proceso de limite:

$$(1+0.06/n)^{5n}\, 1000 \rightarrow \ e^{0.30}\, 1000 = \$\ 1349.87$$

ó se puede cambiar a una ecuación diferencial, que será el límite de la ecuación en diferencias $P_{k+1} = (1+0.06\ \Delta t)P_k$, lo que implica:

$$P_{k+1} = P_k + 0.06\ \Delta t\ P_k$$

$$P_{k+1} - P_k = 0.06\ \Delta t\ P_k$$

de donde:

$$\frac{P_{k+1} - P_k}{\Delta} = 0.06P_k \qquad \Rightarrow \qquad \frac{dp}{dt} = 0.06p$$

cuya solución es $p(t) = e^{0.06t}\, p_0$,para 5 anos se tiene \$ 1349.07.

Este ejemplo incluyó tanto ecuaciones en diferencias como ecuaciones diferenciales, con una tendiendo a la otra a medida que desapareció el lapso de tiempo. Pero hay muchas ecuaciones en diferencias por derecho propio y aquí nuestro siguiente ejemplo:

<u>EJEMPLO 7.15</u> (Los conejos de Fibonacci)

Suponga que las parejas de conejos recién nacidos no tienen descendencia durante su primer mes de vida, pero que a partir de ahí cada pareja produce otra nueva pareja cada mes. Comenzando con $F_1 = 1$ que es la pareja recién nacida en el primer mes, hallar el número F_k de parejas en el k - ésimo mes, suponiendo que no muere ningún conejo.

En el k-ésimo mes el número de parejas de conejos es

F_k = (número de parejas más el mes anterior) + (número de parejas recién nacidas en el k-ésimo mes)

Como nuestros conejos no tienen descendencia durante el primer mes de vida, vemos que el número de parejas recién nacidas en el k-ésimo mes es el número F_{k-2} de parejas vivas dos meses antes. Así, podemos escribir la ecuación anterior como:

$$F_k = F_{k-1} + F_{k-2}$$

Esta es una ecuación en diferencias, conocida como **Relación de Fibonacci**.

Es conveniente hacer $F_0 = 0$ para denotar 0 parejas en el mes 0, antes de la llegada de la primera pareja recién nacida. Así la sucesión:

$$F_0, F_1, F_2, \dots, F_k.$$

para el número de parejas de conejos se convierte en la sucesión de Fibonacci.

$$0,1,1,2,3,5,8,13,21,34,\dots$$

Donde cada término, comenzando con $F_2 = 0+1=1$ es la suma de los dos términos anteriores. Para cualquier k particular, podemos calcular F_k prolongando lo suficiente la sucesión. Sin embargo esto puede ser una tarea tediosa; aunque solo queramos calcular F_{30}.

Una forma de hacerlo, consiste en resolver la ecuación en diferencias $F_{k+2} = F_{k+1} + F_k$; y como primer paso podemos reducirla a una ecuación $u_{k+1}=Au_k$ precisamente como el interés compuesto $P_{k+1} = 1.06\ P_k$; excepto que ahora la incógnita es un vector y el multiplicador A tiene que ser una matriz:

$$\text{Si} \qquad u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$$

entonces:

$$F_{k+2} = F_{k+1} + F_k$$
$$F_{k+1} = F_{k+1}$$

se transforma en:

$$\text{Si} \qquad u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} u_k$$

Formalmente la ecuación en diferencias $u_{k+1} = Au_k$ es fácil de resolver. Como cada paso, conlleva una multiplicación por A, la solución u_k esta relacionada con el valor inicial u_0 mediante $u_k = A^k u_0$.; entonces el problema se reduce a encontrar A^k de un modo rápido y esto se puede resolver a través de los valores y vectores característicos de A.

TEOREMA 7.12

Si puede diagonalizarse A, $A=S\Lambda S^{-1}$,entonces:

$$u_k = A^k u_0 = S\,{}^k S^{-1} u_0$$

ó lo que es lo mismo:

$$u_k = \begin{bmatrix} \vdots & \vdots & \vdots \\ x_1 & \cdots & x_n \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \Lambda^k{}_1 & & \\ & \ddots & \\ & & \Lambda^k{}_n \end{bmatrix} S^{-1} u_0$$

donde $S^{-1}u_0$ es un vector columna de R^n, ya que S^{-1} tiene rango n. De modo que sus columnas forman una base de R^n, de aquí que:

$$S^{-1}U_0 = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

entonces:

$$u_k = C_1\,\Lambda^k{}_1\,x_1 + \ldots + C_n\,\Lambda^k{}_1\,x_n$$

La solución general es una combinación de las soluciones especiales $\alpha^k{}_i\,x_i$ y los coeficientes c_i correspondientes con la condición inicial u_0 son :

$$C_1\,\alpha^0{}_1\,x_1 + \ldots + C_n\,\alpha^0{}_n\,x_n = u_0 \quad \text{o} \quad Sc = u_0$$

En el caso específico de la ecuación de Fibonacci, el primer paso es diagonalizar la matriz A:

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

$$\det(A - \alpha I) = \alpha^2 - \alpha - 1$$

Lo que implica: $\alpha_1 = \dfrac{1+\sqrt{5}}{2}$ $\qquad$ $\alpha_{21} = \dfrac{1-\sqrt{5}}{2}$

$$A = S\,\Lambda\,S^{-1} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 & \\ & \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & -\alpha_2 \\ -1 & \alpha_1 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

Como $F_0 = 0$ y $F_1 = 1$, se tiene $u_0 = [1,0]$, entonces:

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = u_k = A^k u_0 = S^k S^{-1} u_0 = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^k{}_1 & \\ & \alpha^k{}_2 \end{bmatrix} \begin{bmatrix} 1 & -\alpha_2 \\ -1 & \alpha_1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^k{}_1 & \\ & \alpha^k{}_2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^k{}_1 \\ -\alpha^k{}_2 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha^{k+1}{}_1 & -\alpha^{k+1}{}_2 \\ \alpha^k{}_1 & \alpha^k{}_2 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

Entonces el número de Fibonacci F_k, es la segunda componente de este producto:

$$F_k = \frac{\alpha^k{}_1}{\alpha_1 - \alpha_2} - \frac{\alpha^k{}_2}{\alpha_1 - \alpha_2} = 1/\sqrt{5}[(1+\sqrt{5}/2)^k - (1-\sqrt{5}/2)^k]$$

Como el segundo término $[1-\sqrt{5}/2]^k/\sqrt{5}$ siempre es menor que ½, solo debe moverse el primer término al entero más cercano, de modo que:

$$F_k \approx 1/\sqrt{5}\,(1 + \sqrt{5}/2)^k$$

ó

F_k = entero más cercano a $1/\sqrt{5}\,(1+\sqrt{5}/2)^k$ para toda k

Por ejemplo:

Si k=1000 entonces:

F_{1000} = entero más cercano a $1/\sqrt{5}\,(1+\sqrt{5}/2)^{1000}$

Pero este número como se puede ver es muy grande y aún mas grande es F_{1001}, la razón F_{1000}/F_{1001} debe estar muy cerca de la cantidad $1+\sqrt{5}/2 \approx 1.1618$ que los griegos llamaron la **"sección dorada"**. En otras palabras $\alpha^k{}_2$ se vuelve insignificante comparado con $\alpha^k{}_1$ y la razón F_{k+1}/F_k se acerca a $\alpha_1^{k+1}/\alpha_1^k = \alpha_1$.

7.6 CADENAS DE MARKOV

Las Cadenas de Markov se utilizaron inicialmente para analizar procesos en física y meteorología; sin embargo, las aplicaciones más recientes incluyen el análisis de los movimientos de precios de bienes, el mantenimiento de maquinaria, el comportamiento de los animales en el laboratorio, la selección de productos por el consumidor, la longitud de las colas en un supermercado o un aeropuerto, en el manejo de inventarios en cuanto a nivel y en la administración de plantas.

Para definir una cadena de Markov, considere un experimento con un espacio muestra finito $S = \{E_1, E_2, ..., E_n\}$ y una secuencia (ó cadena) de experimentos realizados. Se dice que el experimento está en el estado E_i en el intento m-ésimo, si E_i es el resultado del m-ésimo ensayo del experimento.

Definición 7.8
Una secuencia de intentos de m experimentos es una cadena de Markov, si:

a) El resultado del intento m-ésimo depende solo del resultado del intento (m-1)-ésimo y no de los resultados en los intentos anteriores, y

b) La probabilidad de pasar del estado E_i al estado E_j en dos intentos sucesivos del experimento permanece constante.

Por ejemplo
Si el clima de hoy depende solamente del clima de ayer, entonces la observación y predicción del clima es un problema de cadenas de Markov. Si la probabilidad de escoger una marca particular de coche la próxima vez que piense comprar uno, depende únicamente del auto que ya se tiene, entonces el problema del patrón de ventas de automóviles es un problema que se puede resolver por cadenas de Markov.

Por otra parte, si el clima de hoy es determinado por el clima de varios días anteriores, entonces no se trata de una cadena de Markov.

Una cadena de Markov se caracteriza por las probabilidades de que el sistema pase de un estado a otro en intentos sucesivos.

MATRIZ DE TRANSICIÓN DE UNA CADENA DE MARKOV

La matriz de transición de una cadena de Markov, es la matriz de nxn de probabilidades $T=(p_{ij})$ cuya componente ij-esima p_{ij} es la probabilidad de que el sistema pase del estado E_i al estado E_j en intentos sucesivos del experimento; por lo cual, todos los componentes son no negativos y la suma de ellos por columna es 1.

EJEMPLO 7.16

Cada año 1/10 de la gente que vive fuera de California se muda dentro y 2/10 de la gente que vive dentro se muda fuera.

Esto sugiere una ecuación en diferencias. Comenzamos con y_0 gente que vive fuera y z_0 que vive dentro y al final del primer año hay

$$y_1 = 0.9\, y_0 + 0.2\, z_0$$
$$z_1 = 0.1\, y_0 + 0.8\, z_0$$

ó bien:

$$\begin{bmatrix} y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

La matriz T es una matriz de transición, ya que:

$$0.9+0.1=1 \qquad\qquad 0.2+0.8=1$$

y son todos no negativos. Por lo tanto, las potencias T^k son no negativas.
Resolvemos esta ecuación en diferencias usando $S\ ^kS^{-1}u_0$; posteriormente vemos si la población tiende a un estado estacionario. Esto es, que después de un tiempo las probabilidades de que el sistema se encuentre en cada uno de sus posibles estados son invariables con el tiempo. Para comenzar los cálculos, T tiene que diagonalizarse.

$$T = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$$

$$\det(T-\alpha I) = \alpha_2 - 1.7\,\alpha + 0.7$$

$$\alpha_1 = 1 \ \text{ y } \ \alpha_2 = 0.7$$

$$A = S\Lambda S^{-1} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & \\ & 0.7 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$$

Podemos encontrar A^k y la distribución después de k años:

$$\begin{bmatrix} y_k \\ z_k \end{bmatrix} = A^k \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & \\ & 0.7 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

$$= y_0 z_0 \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} + (y_0 - 2z_0)(0.7)^k \begin{bmatrix} 1/3 \\ -1/3 \end{bmatrix}$$

Esta es la solución que queríamos; a largo plazo el factor $(0.7)^k$ se vuelve muy pequeño y la solución tiende a un estado límite.

$$\begin{bmatrix} y_0 \\ z_0 \end{bmatrix} = (y_0 + z_0)\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

La población total es todavía y_0+z_0, la misma que al principio pero en el límite 2/3 de esta población está fuera de California y 1/3 está dentro. Esto es cierto independientemente de las distribuciones iniciales. Si se comienza en el año con 2/3 fuera y 1/3 dentro, entonces termina de la misma manera:

$$\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

ó $$Au_\infty = u_\infty$$

El estado estacionario es el vector propio de A correspondiente a $\alpha = 1$. La multiplicación por A que nos lleva de un lapso de tiempo a otro, no altera u_∞.

En el ejemplo de California, si el individuo está fuera, se mudará adentro con una probabilidad de 1/10, si está adentro, entonces se mudará afuera con probabilidad 2/10. Su movimiento se vuelve un proceso aleatorio y la matriz A que lo gobierna es una matriz de transición. No sabemos donde está, pero cada año las componentes de $u_k = A^k u_0$ especifican la probabilidad de que esté fuera del estado y la probabilidad de que esté dentro. Estas probabilidades suman 1; el individuo tiene que estar en algún lado y nunca son negativas. Por que $\alpha = 1$ es siempre un valor propio y porque su vector propio es el estado estacionario.

Como se vio en el caso de los conejos de Fibonacci, se tiene que el vector propio o de información después del k-ésimo estado, es igual a $A^k x_1$ para un vector inicial x_1 y una matriz diagonalizable A.. Como es evidente, un aspecto importante es si alguno de los valores propios de A tiene magnitud mayor que 1. Cuando éste es el caso, la magnitud de las componentes del vector de información pueden crecer exponencialmente como es el caso de la sucesión de Fibonacci donde $|\alpha_1|>1$ y, se dice que es inestable.

Por otro lado, si todos los valores propios tienen magnitud menor que 1, las componentes del vector de información deben tender a cero a medida que k crece. Si las probabilidades de Markov decrecieran a cero, entonces la ecuación sería estable, cosa que no sucede porque en cada estado deben sumar 1.

Dada cualquier ecuación en diferencias $U_{k+1} = A_{uk}$, donde queremos estudiar su comportamiento cuando $k \to \infty$. Suponiendo que A puede diagonalizarse, la solución U_k será una combinación de soluciones para:

$$U_k = S\,\Lambda^k S^{-1}\, U_0 = C_1\ {}^{k}_{1}\, x_1 + \dots + C_n\ {}^{k}_{n}\, x_n$$

Definición 7.9

La ecuación en diferencias $U_{k+1} = A_{uk}$ es estable y $U_k \rightarrow 0$. Si todos los valores propios satisfacen $|\alpha_i| < 1$, es neutralmente estable y u_k está acotado si todo $|\alpha_i| \leq 1$ y es inestable (u_k no está acotado) cuando al menos un valor propio de A tiene $|\alpha_i| > 1$.

7.7 FORMAS CUADRÁTICAS

Las formas cuadráticas son una aplicación inmediata de los valores propios, el teorema del eje principal asegura que toda forma cuadrática se puede diagonalizar, este resultado tiene importantes aplicaciones a vibración de cuerpos elásticos, mecánica cuántica y circuitos eléctricos.

Definición 7.10

Se dice que una ecuación de la forma

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0 \qquad (7.4)$$

Donde a,b,c,d,e,f son números reales y al menos uno de ellos es diferente de cero (es una ecuación cuadrática en x,y). La expresión $ax^2 + 2by + cy^2$, se denomina la forma cuadrática asociada.

EJEMPLO 7.17

En la ecuación cuadrática

$$3x^2 + 5xy - 7y^2 + 2x + 7 = 0$$

las constantes son:

$a = 3 \qquad b = 5/2 \qquad c = -7 \qquad d = 2 \qquad e = 0 \qquad f = 7.$

EJEMPLO 7.18

Ecuación cuadrática	**Forma cuadrática asociada**
$3x^2 + 5xy - 7y^2 + 2x + 7 = d$	$3x^2 + 5xy - 7y^2$
$4x^2 - 5y^2 + 8y + 9 = d$	$4x^2 - 5y^2$
$xy + y = d$	xy

Las gráficas de las ecuaciones cuadráticas en x,y se llaman **cónicas** ó **secciones cónicas**.

Las cónicas mas importantes son la elipse, la circunferencia, la hipérbola y la parábola.

Una forma cuadrática con una variable x es un polinomio $f(x)= ax^2$ con $a \neq 0$. Una forma cuadrática en 2 variables x,y es un polinomio $f(x) = ax^2 + bxy + y^2$, donde a, b ó c es $\neq 0$. El término cuadrático significa de grado 2, el término forma significa homogénea, es decir, cada sumando tiene un producto del mismo número de variables, a saber 2. Se dice que una cónica tiene una posición normal relativa a los ejes de coordenadas, si no se encuentra trasladada ni rotada.

EJEMPLO 7.19

La ecuación

$$\frac{x^2}{4}+\frac{y^2}{9}=1$$

es de la forma:

$$\frac{x^2}{k^2}+\frac{y^2}{1^2}=1 \qquad \text{con}\, k = 2, \qquad 1 = 3$$

Por lo tanto, su gráfica es una elipse en posición normal que intersecta al eje x en: (-2,0), (2,0) y al eje *y* en (0,-3) y (0,3).

La ecuación $x^2 - 8y^2 = -16$ se puede expresar como $y^2/2 - x^2/16 = 1$ que es de la forma

$y^2/k^2 - x^2/l^2 = 1$ con $k = \sqrt{2}$ $l=4$

Por lo tanto, su gráfica es una hipérbola en posición normal que corta al eje *y* en $(0, -\sqrt{2})$ y $(0, \sqrt{2})$.

La ecuación $5x^2 + 2y = 0$ se puede reescribir como: $x^2 = (-2/5)$ y ,que es de la forma $x^2 = ky$ con $k=-2/5$ como $k < 0$. Su gráfica es una parábola en posición normal que se abre hacia abajo.

Ninguna cónica en posición normal tiene términos cruzados (xy) en su ecuación. Si éstos aparecen, significa que la cónica se giró con respecto a su posición normal. Tampoco se tienen cónicas en posición normal que tengan simultáneamente un término en x^2 y en x o en y^2 y en y. Si tales términos aparecen, significa que la cónica se trasladó de su posición normal. Ver figura 7.2.

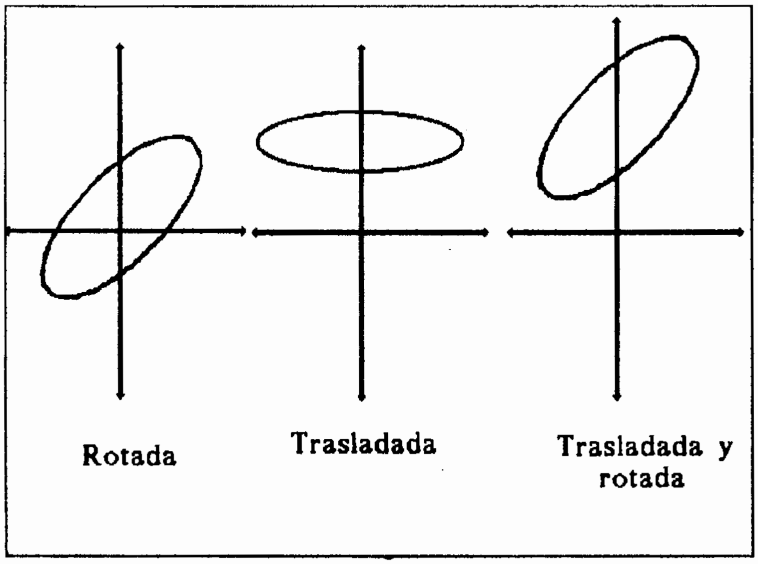


Figura 7.2

EJEMPLO 7.20

Dada la ecuación cuadrática.

$$2x^2 + y^2 - 12x - 4y + 18 = 0$$

Tiene términos en x^2, x, y^2, y ; pero no en xy ; entonces, su gráfica es una cónica que se trasladó pero no se giró: Agrupamos términos para obtener:

$$(2x^2 - 12x) + (y^2 - 4y) + 18 = 0$$
$$2(x^2 - 6x) + (y^2 - 4y) = -18$$

Completando cuadrados ;donde para completar el cuadrado de una expresión de la forma x^2+px se suma y se resta $(P/2)^2$.

$$x^2 + px = x^2 + px + (p/2)^2 - (p/2)^2 = (x+p/2)^2 - (p/2)^2$$

Se tiene entonces:

$$2(x^2 - 6x + 9) + (y^2 - 4y + 4) = -18 + 18 + 4$$

$$2(x - 3)^2 + (y - 2)^2 = 4$$

Si se trasladan los ejes de coordenadas mediante

$$x' = x-3 \qquad y' = y-2$$

se tiene: $$2x'^2 + y'^2 = 4$$

Entonces:

$$\frac{x'^2}{2} + \frac{y'^2}{4} = 1$$

es una elipse en posición normal con respecto a x'y' (Ver figura 7.3)

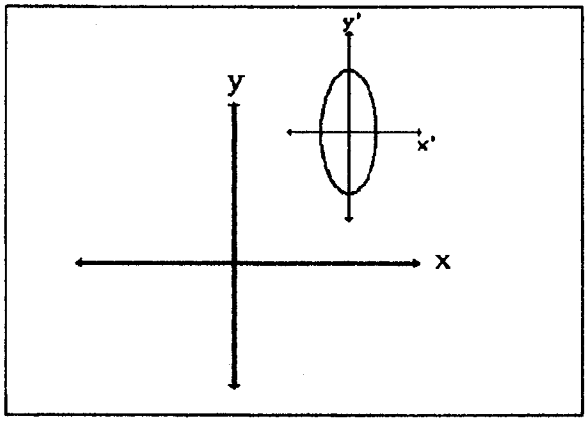


Figura 7.3

La ecuación (7.4)

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

se puede expresar en forma matricial como:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + f = 0$$

es decir:

$$x^t A x + kx + f = 0 \qquad (7.5)$$

donde:

$$x = \begin{bmatrix} x \\ y \end{bmatrix} \qquad A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \qquad k = \begin{bmatrix} d & e \end{bmatrix}$$

Con esta notación, la forma cuadrática asociada a (7.5) es:

$$x^t A x$$

La matriz simétrica A se llama la matriz de la forma cuadrática $x^t A x$.

EJEMPLO 7.21

La matriz de la forma cuadrática

$$3x^2 + 5xy - 7y^2$$

es:

$$\begin{bmatrix} 3 & 5/2 \\ 5/2 & -7 \end{bmatrix}$$

Considere que la ecuación de una cónica C, es:

$$x^t Ax + kx + f = 0 \qquad (7.6)$$

A continuación se muestra que siempre es posible rotar los ejes de coordenadas x-y de tal manera, que la ecuación de la cónica con respecto al sistema de coordenadas x'-y' no contenga un término en xy.

Etapa 1

Encuentre la matriz

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix}$$

que diagonalice ortogonalmente a A..

Etapa 2

Si es necesario, intercambiar las columnas de P de tal suerte que det(P) = 1. Esto garantiza que la transformación ortogonal de coordenadas (7.7) es una rotación:

$$x=Px' \qquad (7.7)$$

Etapa 3

Para obtener la ecuación de C con respecto al sistema x'- y' ,se sustituye (7.6) en (7.7)

$$(Px')^t \ A(Px') + k(Px') + f = 0$$

ó

$$x^t \ (P^tAP) \ x' + (kP)x' + f = 0 \qquad (7.8)$$

Como P diagonaliza ortogonalmente a A..

$$P'AP = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}$$

donde α_1 y α_2 son valores propios de A, por lo cual (7.8) se puede expresar como:

$$[x' \quad y'] \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + [d \quad C] \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + f = 0$$

Entonces:

$$\alpha_1 x'^2 + \alpha_2 y'^2 + d'x' + e'y' + f = 0.$$

donde: $d' = dP_{11} + eP_{21} \quad e' = dP_{12} + eP_{22}$.

El siguiente teorema resume todo el proceso.

TEOREMA 7.13 DE LOS EJES PRINCIPALES PARA R^2

Sea

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0$$

la ecuación de una cónica C y sea

$$x^t Ax = ax^2 + 2bxy + cy^2$$

la forma cuadrática asociada. Entonces los ejes de coordenadas se pueden rotar de tal manera que la ecuación C con respecto al nuevo sistema de coordenadas x' - y' tiene la forma:

$$\alpha_1 x^{12} + \alpha_2 y^{12} + d'x' + e'y' + f = 0$$

donde α_1 y α_2 son valores propios de A. La rotación se puede efectuar mediante la sustitución $x = Px'$,donde P diagonaliza ortogonalmente a A.

EJEMPLO 7.22

Describa la cónica C cuya ecuación está dada por:

$$5x^2 - 4xy + 8y^2 + (20/\sqrt{5})x - (80/\sqrt{5})y + 4 = 0$$

La forma matricial de esta ecuación es:

$$x^t Ax + kx + 4 = 0$$

donde:

$$A=\begin{bmatrix}5 & -2\\ -2 & 8\end{bmatrix} \qquad k=\left[20/\sqrt{5}, -80/\sqrt{5}\right]$$

$$\det(A - \alpha I) = (\alpha - 9)(\alpha - 4)$$

y v_1 v_2 son:

$$P=\begin{bmatrix}2/5 & -1/\sqrt{5}\\ 1/5 & -2/\sqrt{5}\end{bmatrix}$$

Diagonaliza ortogonalmente a A. Sustituyendo x = Px'

$$(Px')^t A(Px') + k(Px') + y = 0$$

es decir:

$$(x'^t)(P^tAP)x' + (kP)x' + y = 0$$

Dado que

$$P^tAP=\begin{bmatrix}4 & \\ & 9\end{bmatrix} \qquad kP=\left[20/\sqrt{5}, -80/\sqrt{5}\right]\begin{bmatrix}2/5 & -1/5\\ 1/5 & 2/5\end{bmatrix}=[-8, -36]$$

entonces:

$$4x'^2 + 9y'^2 - 8x' - 36y' + 4 = 0$$

trasladando: $$4(x'^2 - 2x') + 9(y'^2 - 4y') = -4$$

Completando cuadrados:

$$4(x'^2 - 2x' + 1) + 9(y'^2 - 4y' + 4) = -4+4+36$$

ó bien: $$4(x'-1)^2 + 9(y'-2)^2 = 36$$

Sean

$$x'' = x' - 1$$
$$y'' = y' - 2$$

las ecuaciones de traslación, entonces:

$$4x''^2 + 9y''^{12} = 36$$

ó $$\frac{x''^2}{9} + \frac{y''^2}{4} = 1$$ es la ecuación de una elipse.

EJEMPLO 7.23

Describa la cónica C cuya ecuación está dada por:

$$5x^2 - 4xy + 8y^2 - 36 = 0$$

ó bien

$$x^tAx - 36 = 0$$

donde:

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}$$

$$\det(A - \alpha I) = (\alpha - 9)(\alpha - 4) = 0$$

Para $\alpha = 4$, se tiene:

$$v_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

Para $\alpha = 9$, se tiene:

$$v_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

Por lo tanto:

$$P = \begin{bmatrix} 2/5 & -1\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

digonaliza ortogonalmente a A; además el det P = 1; por lo cual, se puede concluir que
x = Px' es una rotación ó bien:

$$(x')_t (P_tAP)\, x' - 36 = 0$$

donde:

$$P'PA = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$$

$$\begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} - 36 = 0$$

$$4x'^2 + 9y'^2 - 36 = 0$$

ó $$\frac{x'^2}{9} + \frac{y'^2}{4} = 1$$ es una elipse.

7.8 RESUMEN

1. Si A es una matriz de nxn, entonces un escalar α es un valor propio de A si hay un vector v distinto de cero; tal que:

$$Av = \alpha v$$

El vector $v \neq 0$ se llama un vector propio de A correspondiente a un valor propio α

2. Si A es de orden nxn, entonces la ecuación

$$(A - \alpha I)v = 0$$

es un sistema homogéneo de n ecuaciones con n incógnitas, como por hipótesis (ya que v es un vector propio distinto de cero) el sistema tiene soluciones no triviales, entonces:

$$\det (A - \alpha I) = 0$$

A esta ecuación se le denomina la ecuación característica de A.

3. Si contamos multiplicidades, cada matriz de nxn tiene exactamente n valores característicos.

4. Si A es una matriz de orden nxn, entonces las siguientes afirmaciones son equivalentes:

a) α es un valor característico de A.
b) El sistema $(A - \alpha I)v = 0$ tiene soluciones no triviales.
c) Existe un vector v e n R^n $v \neq 0$ tal que $Av = \alpha v$

5. Si α es un valor característico de A, entonces el espacio solución del sistema de ecuaciones $(A - \alpha I)v = 0$ se denomina el espacio característico de A correspondiente a α, y los vectores diferentes de cero en el espacio característico, se denominan los vectores característicos de A correspondientes a α.

6. Sea A una matriz de nxn y sea $\alpha_1, \alpha_2, \ldots, \alpha_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $v_1, v_2, \ldots, v_m$. Entonces v_i son linealmente independientes o bien, los vectores característicos correspondientes a valores característicos diferentes son linealmente independientes. (la demostración se hace por inducción).

7. Sea α un valor característico de A, entonces la multiplicidad geométrica de α es la dimensión del espacio característico correspondiente a α (que es la nulidad de la matriz $A - \alpha I$) ,esto es:

$$\text{multiplicidad geométrica de } \alpha = \dim E_\alpha$$

8. Sea A una matriz de orden nxn, entonces A tiene n vectores característicos linealmente independientes si y solo si la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, A tiene n vectores característicos l.i. si todos sus valores característicos son distintos (pues si no, podrían ser menores que n)

9. Sea la matriz A de orden nxn.

9.1 Si α es un valor propio de A con v como vector propio correspondiente, entonces ^{k} es un valor propio de A^k, de nuevo con v como vector propio correspondiente, para cualquier entero positivo k.

9.2 Si α es un valor propio de una matriz invertible A con v como vector propio correspondiente, entonces $1/\alpha$ es un valor propio de A^{-1}, de nuevo con v como vector propio correspondiente.

9.3 Si α es un valor propio de A, entonces el conjunto E es un subespacio de R^n.

10. Una secuencia de intentos de m experimentos es una cadena de Markov si:

a) El resultado del intento m-ésimo, depende solo del resultado del intento (m-1)-ésimo y no de los resultados en los intentos anteriores, y

b) La probabilidad de pasar del estado E_i al estado E_j en dos intentos sucesivos del experimento permanece constante.

11. Se dice que una ecuación de la forma

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0$$

Donde a,b,c,d,e,f son números reales y al menos uno de ellos es diferente de cero, es una ecuación cuadrática en x,y. La expresión $ax^2 + 2by + cy^2$, se denomina la forma cuadrática asociada.

12. Sea

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0$$

la ecuación de una cónica C y sea

$$x^tAx = ax^2 + 2bxy + cy^2$$

la forma cuadrática asociada. Entonces los ejes de coordenadas se pueden rotar de tal manera que la ecuación C con respecto al nuevo sistema de coordenadas x' - y' tiene la forma:

$$\alpha_1 x^{12} + \alpha_2 y^{12} + d'x' + e'y' + f = 0$$

donde α_1 y α_2 son valores propios de A. La rotación se puede efectuar mediante la sustitución $x = Px'$,donde P diagonaliza ortogonalmente a A.

7.9 NOTAS HISTÓRICAS.8

La primera aparición de los **valores propios**, fue en relación con su uso para resolver ecuaciones diferenciales. En 1743, Leonhard Euler introdujo por primera vez el método estándar de resolución de una ecuación diferencial de n - ésimo orden con coeficientes constantes

$$y^{(n)} + a_{n-1}\, y^{(n-1)} + \ldots + a_1\, y^1 + a_0 y = 0$$

usando funciones de la forma $y = e^{at}$,donde a es una raíz de la ecuación característica:

$$z^n + a_{n-1}\, z^{n-1} + \ldots + a_1\, z + a_0 = 0$$

Esta es la misma ecuación que se obtiene al hacer las sustituciones

$$y_1 = y,\ y_2 = y',\ y_3 = y'', \ldots,\ y_n = y^{(n-1)},$$

reemplazando la única ecuación de orden n por un sistema de n ecuaciones de primer orden

$$\left|\begin{array}{lcl} y'_1 & = & y_2 \\ y'_2 & = & y_3 \\ y'_n & = & -a_0 y_1 - a_1 y_2 - \ldots - a_n - 1 y_n \end{array}\right.$$

y calculando la ecuación característica de la matriz de coeficientes de este sistema.

Unos 20 años más tarde, Lagrange dio una versión más explícita de esta misma idea, al hallar la solución de un sistema de ecuaciones diferenciales, encontrando las raíces de lo equivalente a la ecuación característica de la matriz de coeficientes. El sistema particular de ecuaciones diferenciales surgió del examen de los "movimientos infinitesimales" de un sistema mecánico en la vecindad de su posición de equilibrio. En 1774, Lagrange resolvió un problema similar de mecánica celeste usando la misma técnica.

El nombre de **Cadenas de Markov** se debe al matemático ruso Andrei Andreevich Markov (1856-1922), quien las definió por primera vez en un artículo de 1906 que trataba de la ley de los grandes números y posteriormente demostró muchos resultados estándar sobre ellas. Su interés en estas sucesiones se originó en las necesidades de la teoría de probabilidad. Markov nunca trató sus aplicaciones a las ciencias. Los únicos ejemplos reales que utilizó eran de textos literarios, donde los dos estados posibles eran vocales y consonantes. Para ilustrar sus resultados hizo un estudio estadístico de la alternancia de vocales y consonantes en el libro de Pushkin: Eugen Onegin.
Andrei Markov dio clases en la Universidad de San Petersburgo de 1880 a 1905, y se retiró para dar paso a matemáticos más jóvenes. Además de su trabajo en probabilidad, hizo contribuciones a campos como teoría de números, fracciones continuas y teoría de la aproximación. Fue un participante activo en el movimiento liberal ruso en la época anterior a la Primer Guerra Mundial; en varias ocasiones criticó públicamente la actuación de las autoridades estatales. En 1913, cuando como miembro de la Academia de Ciencias se le pidió participar en las pomposas ceremonias de celebración del 300 aniversario de la dinastía Romanov, prefirió organizar una conmemoración del 200 aniversario de la publicación de la ley de los grandes números de Jacobo Bernoulli.

FIBONACCI (1175 - 1230) El notable matemático italiano Leonardo de Pisa mejor conocido por su sobrenombre de Fibonacci, una abreviación de "filus Bonacci" que significa el hijo de Bonaccio de Pisa. Su padre, notario público de Bugía, en Argel, confió su hijo a las enseñanzas de un maestro de cálculo musulmán: dedicó "algunos días" al estudio del ábaco, aprendió los signos numéricos de los indios, y penetró en seguida en los secretos de la ciencia arábiga. Ulteriores viajes lo llevaron posteriormente a través de Oriente y de Europa, hacia Egipto, Siria; Grecia, Sicilia y Provenza. Por doquier planteó nuevas cuestiones a los más afamados maestros con los que medía sus fuerzas para perfeccionar su sabiduría. "Pero todo esto, y el algoritmo y los arcos de Pitágoras, sólo me parecieron otras tantas equivocaciones, comparadas con el método de los indios".

En el año 1202, a los 27 años de edad, fue cuando escribió Leonardo, resumiendo su sabiduría singular, un gran libro sobre el arte de calcular, el Liber abaci, en el que por primera vez un matemático cristiano ofrecía a sus contemporáneos una imagen completa, acabada y ordenada con la mayor claridad, del arte del cálculo tal y como éste se había desarrollado más allá del mundo europeo. Nos informa en ella que, habiendo aprendido en Berbería, el sistema numérico de los árabes lo transmitía "a la raza latina para que ésta no permaneciese más tiempo privada de él".

En el Liber abaci comezaba Leonardo explicando los signos numéricos, enseñando a continuación a calcular con los dedos, así como las operaciones más sencillas con los números enteros y las fracciones. Una gran parte de sus problemas se referían a casos prácticos. Así, por ejemplo, en la sección 8 trataba " de los precios de las mercancías a largo plazo", en la 9, "del intercambio de mercancías y cosas semejantes"; en la 10, "de la aligación y compañía"; en la 11, "de la aleación y mezcla de las monedas"; en la 15, "de las reglas pertenecientes a la Geometría y de los problemas del Algebra y Almucabala". Es en este mismo libro donde expone su ejemplo de los conejos y como se construye la sucesión que posteriormente llevará su nombre.

En 1826, Cauchy analizó las formas cuadráticas con tres variables, es decir, formas del tipo $Ax^2 + By_2 + Cz^2 + 2Dxy + 2Exz + 2Fyz$. Demostró que la ecuación característica formada a partir del determinante:

$$\begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix}$$

que permanece igual bajo cualquier cambio de ejes rectangulares, lo que nosotros llamaríamos un cambio coordenado ortogonal.

BIBLIOGRAFÍA

1. ANTON, Howard, Introducción al Álgebra Lineal, *Ed. Limusa*, México D.F.,1984.

2. FUENTES MAYA, Sergio, Apuntes de Matemáticas Aplicadas I, *DEPFI*, México D.F.,1989.

3. FRALEIGH, J.B. y BEAUREGARD, R.A. Álgebra Lineal, *Ed. Addison-Wesley Iberoamericana*, México D.F.,1989.

4. GROSSMAN, Stanley I. Aplicaciones de Álgebra Lineal, *Ed. Grupo Editorial Iberoamérica*, México D.F.,1988.

5. LOPEZ DE MEDRANO, Santiago, Modelos Matemáticos, *Ed. ANUIES*, México D.F., 1973.

6. LOPEZ DE MEDRANO, Santiago, Lenguajes Simbólicos, *Ed. ANUIES*, México D.F.,1973.

7. PERRY, William L. Álgebra Lineal con Aplicaciones, *Ed. Mc. Graw Hill*, México D.F.,1990.

8. PROSKURIAKOV, I. Problemas de Álgebra Lineal, *Ed. MIR*, URSS Hayca,1986.

9. RORRES, C. y ANTON, H. Aplicaciones de Álgebra Lineal, *Ed. Limusa*, México D.F.,1979.

10. STRANG, Gilbert, Álgebra Lineal y sus Aplicaciones, *Ed. Fondo Educativo Interamericano*, México D.F.,1982.

11. STRANG, Gilbert, Introduction To Applied Mathematics, *Ed. Wellesley-Cambridge Press*, México D.F.,1986.

Esta obra se terminó de imprimir
en abril de 2001
en el taller de imprenta del
Departamento de Publicaciones
de la Facultad de Ingeniería,
Ciudad Universitaria, México, D.F.
C.P. 04510

Secretaría de Servicios Académicos

El tiraje consta de 200 ejemplares
más sobrantes de reposición.