Apuntes de Álgebra Lineal

PROFESOR: Ricardo Visiers Bañón

Rev04 - Septiembre 2019

1 SISTEMAS DE ECUACIONES LINEALES

1.1 Ecuaciones lineales

Una **ecuación lineal en n variables** $x_1, x_2, x_3, \ldots.., x_n$ tiene la forma:

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b$$

Los coeficientes $a_1, a_2, a_3, \ldots., a_n$ y el **término constante** b son números reales.

Nota: Las primeras letras del abecedario se utilizan para representar constantes; las últimas para representar variables.

Las ecuaciones lineales no contienen productos o raíces de las variables, ni funciones trigonométricas, exponenciales o logarítmicas. Las variables aparecen sólo elevadas a la primera potencia.

Una **solución** de una ecuación lineal en n variables es una colección de n números reales $s_1, s_2, s_3, \ldots., s_n$ tales que la ecuación se satisface cuando se sustituyen en ella los valores:

$x_1 = s_1, \quad x_2 = s_2, \quad x_3 = s_3, \ldots\ldots. \quad x_n = s_n$

El conjunto de todas las soluciones de una ecuación lineal se llama **conjunto solución.** Para describir el conjunto solución de una ecuación suele emplearse una representación paramétrica, tal como se describe en el siguiente ejemplo.

Ejemplo

Resolver la ecuación lineal $\quad x_1 + 2x_2 = 4$

Para hallar el conjunto solución de una ecuación en dos variables, despejamos una de éstas en términos de la otra, por ejemplo:

$$x_1 = 4 - 2x_2$$

Con ello la variable x_2 queda **libre**, con lo que podemos asignarle cualquier valor real. Por el contrario la variable x_1, no es libre, ya que su valor depende del asignado a x_2. Para representar las infinitas soluciones de ésta ecuación conviene introducir una tercera variable λ, llamada **parámetro**. Así pues, haciendo $x_2 = \lambda$ podemos representar el conjunto solución como:

$x_1 = 4 - 2x_2$, $x_2 = \lambda$, λ es cualquier número real.

Soluciones particulares se obtienen sin más que asignar un valor al parámetro λ

1.2 Sistemas de ecuaciones lineales

Un **sistema de *m* ecuaciones lineales con *n* incógnitas** $\boldsymbol{x_1, x_2, \ldots, x_n}$—al que podemos llamar simplemente **sistema lineal**—, es un conjunto de *m* ecuaciones lineales, cada una con *n* incógnitas. Un sistema lineal puede denotarse sin problema mediante:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_n \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Nota: La notación de doble índice indica que a_{ij} es el coeficiente de x_j en la ecuación i-ésima.

Una solución de un sistema de ecuaciones lineales es una colección de números s_1, s_2, s_3, …., s_n que es solución de todas y cada una de las ecuaciones de sistema.

Número de soluciones de un sistema de ecuaciones lineales

Para un sistema de ecuaciones lineales, exactamente una de las siguientes posibilidades ocurre:

1. El sistema tiene un única solución (sistema compatible – determinado)
2. El sistema tiene infinitas soluciones (sistema compatible – indeterminado)
3. El sistema no tiene solución (sistema incompatible)

1.2.1 Resolución de sistemas de ecuaciones lineales

Sistemas equivalentes. Se dice que dos sistemas de ecuaciones son equivalentes si tienen el mismo conjunto solución. Para resolver un sistema cualquiera, lo pasamos a una forma equivalente escalonada. Para ello emplearemos las siguientes propiedades:

Cada una de estas operaciones pasa de un sistema de ecuaciones lineales a otro equivalente:

1. Intercambiar dos ecuaciones
2. Multiplicar una ecuación por una constante distinta de cero.
3. Sumar a una ecuación un múltiplo de otra.

Reescribir un sistema dado en forma escalonada suele requerir una cadena de sistemas equivalentes, cada uno de los cuales se obtiene del anterior por alguna de las tres operaciones básicas citadas. Este proceso se conoce como **eliminación gaussiana**.

Ejemplo. Sistema compatible

Resolver el sistema:

$$\begin{cases} x - 2y + 3z = 9 \\ -x + 3y \quad\quad = -4 \\ 2x - 5y + 5z = 18 \end{cases}$$

Solución

$$\begin{cases} x - 2y + 3z = 9 \\ y + 3z \quad = 5 \\ 2x - 5y + 5z = 18 \end{cases}$$ Sumando las ecuaciones 1ª y 2ª obtenemos una nueva 2ª ecuación

$$\begin{cases} x - 2y + 3z = 9 \\ y + 3z = 5 \\ -y - z = -1 \end{cases}$$ Sumando -2 veces la 1ª a la tercera obtenemos una nueva 3ª ecuación

Una vez eliminadas de la primera columna todas las x excepto la de arriba, vamos con la segunda columna.

$$\begin{cases} x-2y+3z=9 \\ \quad y+3z = 5 \\ \quad 2z=4 \end{cases}$$ Sumando las ecuaciones 2ª y 3ª obtenemos una nueva 3ª ecuación

$$\begin{cases} x-2y+3z=9 \\ \quad y+3z = 5 \\ \quad z=2 \end{cases}$$ Multiplicando por ½ la 3ª ecuación obtenemos una nueva 3ª ecuación

Basta sustituir el valor de $z = 2$ en la segunda ecuación para obtener $y = -1$, y sustituyendo z e y en la primera, obtenemos $x = 1$. Se trata por tanto, de un sistema compatible determinado (una única solución formada por (1, -1, 2), coordenadas del punto donde coinciden los tres planos.

Ejemplo. Sistema incompatible. Sin soluciones

Resolver el sistema:

$$\begin{cases} x_1 - 3x_2 + x_3 = 1 \\ 2x_1 - x_2 - 2x_3 = 2 \\ x_1 + 2x_2 - 3x_3 = -1 \end{cases}$$

Solución

Sumando -2 veces la 1ª ecuación a la segunda obtenemos una nueva 2ª ecuación:

$$\begin{cases} x_1 - 3x_2 + x_3 = 1 \\ \quad 5x_2 - 4x_3 = 0 \\ x_1 + 2x_2 - 3x_3 = -1 \end{cases}$$

Sumando -1 veces la 1ª ecuación a la tercera obtenemos una nueva 3ª ecuación:

$$\begin{cases} x_1 - 3x_2 + x_3 = 1 \\ \quad 5x_2 - 4x_3 = 0 \\ \quad 5x_2 - 4x_3 = -2 \end{cases}$$

Restamos la 2ª ecuación de la tercera para obtener una nueva 3ª ecuación:

$$\begin{cases} x_1 - 3x_2 + x_3 = 1 \\ \quad 5x_2 - 4x_3 = 0 \\ \quad 0 = -2 \end{cases}$$

Puesto que la tercera "ecuación" es absurda, este sistema carece de solución, como es equivalente al inicial, concluimos que el sistema propuesto no tiene solución, es decir, es un sistema incompatible. En este caso los tres planos no tienen ningún punto en común.

Ejemplo. Sistema compatible indeterminado. Infinitas soluciones

Resolver el sistema:

$$\begin{cases} x_2 - x_3 = 0 \\ x_1 - 3x_3 = -1 \\ -x_1 + 3x_2 = 1 \end{cases}$$

Solución

Intercambiamos las dos primeras ecuaciones:

$$\begin{cases} x_1 \quad\quad - 3x_3 = -1 \\ \quad\quad x_2 - x_3 = 0 \\ -x_1 + 3x_2 \quad\quad = 1 \end{cases}$$

Sumando la 1ª ecuación a la 3ª obtenemos una nueva 3ª ecuación:

$$\begin{cases} x_1 \quad\quad - 3x_3 = -1 \\ \quad\quad x_2 - x_3 = 0 \\ 3x_2 - 3x_3 = 0 \end{cases}$$

Sumando -3 veces la segunda ecuación a la 3ª obtenemos una 3ª ecuación trivial, que podemos eliminar:

$$\begin{cases} x_1 \quad\quad - 3x_3 = -1 \\ \quad\quad x_2 - x_3 = 0 \\ 0 = 0 \end{cases}$$

Obtenemos así un sistema equivalente:

$$\begin{cases} x_1 \quad\quad - 3x_3 = -1 \\ \quad\quad x_2 - x_3 = 0 \end{cases}$$

Para representar las soluciones, tomamos x_3 como variable libre y la denotamos con la letra λ. Dado que $x_2 = x_3$ y $x_1 = 3x_3 - 1$, podemos describir el conjunto solución como:

$x_1 = 3\lambda - 1, \quad x_2 = \lambda, \quad x_3 = \lambda,$ donde λ representa cualquier número real.

1.3 Eliminación gaussiana y eliminación de Gauss-Jordan[1]

Una matriz A de *m* x *n* es una distribución rectangular de $m \cdot n$ números reales (o complejos) ordenados en *m* filas (renglones) horizontales y *n* columnas verticales:

$$A = \begin{array}{cl} \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} & \leftarrow \text{fila (renglón) } i \\ \uparrow \text{ columna } j & \end{array}$$

La ***i*-ésima fila** de A es:

$$[a_{i1}, a_{i2}, a_{i3}, \cdots, a_{in}] \ (1 \leq i \leq m)$$

La ***j*-ésima columna** de A es:

[1] Este apartado se volverá a ver más adelante al tratar de matrices en el apartado 2.1

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad (1 \leq j \leq n)$$

Diremos que A es ***m* por *n*** (que se escribe $\boldsymbol{m} \times \boldsymbol{n}$). Si $m = n$, decimos que A es una **matriz cuadrada de orden *n***, y que los números $a_{11}, a_{22}, \ldots, a_{nn}$ forman la **diagonal principal** de A. Nos referimos al número a_{ij}, que está en la *i*-ésima fila (renglón) y la *j*-ésima columna de A, como el ***i*, *j*-ésimo elemento** de A, o la **entrada (*i*, *j*)** de A, y solemos escribir como:

$$A = [a_{ij}]$$

Restringiremos nuestra atención al análisis de las matrices cuyas entradas son números reales. Sin embargo, también se estudian las matrices con entradas complejas, que tienen gran importancia en muchas aplicaciones.

Matriz ampliada. Una de las aplicaciones más frecuentes de las matrices consiste en la representación de sistemas lineales. La matriz que consta de los coeficientes y los términos constantes del sistema se denomina matriz ampliada del sistema. La matriz formada sólo por los coeficientes se denomina matriz de coeficientes del sistema.

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

sistema *matriz ampliada* *matriz de coeficientes*

Operaciones elementales por filas.

* Intercambiar dos filas
* Multiplicar una fila por una constante distinta de cero
* Sumar a una fila un múltiplo de otra

Aplicada sobre la matriz ampliada de un sistema lineal, una de estas operaciones produce la matriz ampliada de un nuevo sistema equivalente. Se dice que dos matrices son equivalentes por filas si una de ellas se puede obtener de la otra mediante aplicación de una *secuencia finita* de operaciones elementales por filas.

Es conveniente anotar las operaciones que se realicen de modo que podamos detectar un error. Así, si realizamos las operaciones elementales en el ejemplo del sistema compatible determinado del punto anterior sobre la matriz ampliada tendremos:

$$\begin{cases} x - 2y + 3z = 9 \\ -x + 3y = -4 \\ 2x - 5y + 5z = 17 \end{cases} \qquad \begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix} \qquad \text{Sistema y matriz original}$$

$$\begin{cases} x - 2y + 3z = 9 \\ y + 3z = 5 \\ 2x - 5y + 5z = 17 \end{cases} \qquad \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 2 & -5 & 5 & 17 \end{bmatrix} \qquad \text{F1+F2} \rightarrow \text{F2}$$

$$\begin{cases} x-2y+3z=9 \\ \quad y+3z\ =\ 5 \\ \quad -y-z=-1 \end{cases} \qquad \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & -1 & -1 & -1 \end{bmatrix}$$

-2F1+F3 → F3

$$\begin{cases} x-2y+3z=9 \\ \quad y+3z\ =\ 5 \\ \quad 2z=4 \end{cases} \qquad \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

F2+F3 → F3

$$\begin{cases} x-2y+3z=9 \\ \quad y+3z\ =\ 5 \\ \quad z=2 \end{cases} \qquad \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$^1/_2$F3 → F3

Esta última matriz se dice que está de forma escalonada.

Definición de matriz de forma escalonada. Una matriz está en forma escalonada cuando:

* Las filas nulas, si las hay, ocupan las posiciones más bajas.
* En cada fila no nula, el primer elemento no nulo es 1 (se llama **1 dominante o pivote**)
* Dadas dos filas sucesivas (no nulas), el 1 dominante de la fila superior está más a la izquierda que el 1 dominante de la inferior.

Una matriz escalonada se dice que está en forma escalonada reducida si , además de cumplir las tres condiciones anteriores, en las columnas que contienen un 1 dominante o pivote, todos los demás elementos encima y debajo del 1, son cero.

Teorema. Cada matriz es equivalente por filas a una, y solo a una, matriz escalonada reducida.

Se puede demostrar que toda matriz es equivalente por filas a una matriz escalonada. Por lo general, la forma escalonada por filas de una matriz no es única. Es decir, una matriz puede ser equivalente, en sus filas, a más de una matriz en forma escalonada por filas.

1.3.1 Eliminación gaussiana con sustitución hacia atrás.

1. Escribir la matriz ampliada del sistema lineal.
2. Llevar la matriz ampliada a forma escalonada aplicando operaciones elementales por filas.
3. Escribir el sistema asociado a la matriz escalonada obtenida y resolverlo mediante sustitución hacia atrás.

Este sistema proporciona un algoritmo cómodo de implementar en un ordenador. En este algoritmo, el orden en que se realizan las operaciones elementales por filas es importante, hay que desplazarse de izquierda a derecha, haciendo cero en cada columna los elementos situados debajo del 1 dominante (primer elemento no nulo).

Ejemplo
Resolver el sistema:

$$\begin{cases} \quad +x_2+\ \ x_3-\ 2x_4\ =\ -3 \\ x_1+2x_2-\ \ x_3 \qquad\ \ =\ \ 2 \\ 2x_1+4x_2+\ x_3-3\,x_4=\ -2 \\ x_1-4x_2-7x_3-\ x_4=\ -19 \end{cases}$$

Solución

La matriz ampliada de este sistema es:

$$\begin{bmatrix} 0 & 1 & 1 & -2 & -3 \\ 1 & 2 & -1 & 0 & 2 \\ 2 & 4 & 1 & -3 & -2 \\ 1 & -4 & -7 & -1 & -19 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 2 & 4 & 1 & -3 & -2 \\ 1 & -4 & -7 & -1 & -19 \end{bmatrix} \quad F1 \leftrightarrow F2$$

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 1 & -4 & -7 & -1 & -19 \end{bmatrix} \quad F3 + (-2)F1 \longrightarrow F3$$

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 0 & -6 & -6 & -1 & -21 \end{bmatrix} \quad F4+(-1)F1 \longrightarrow F4$$

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 0 & 0 & 0 & -13 & -39 \end{bmatrix} \quad F4+6F2 \longrightarrow F4$$

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & -13 & -39 \end{bmatrix} \quad (1/3)F3 \longrightarrow F3$$

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix} \quad (-1/13)F4 \longrightarrow F4$$

Hemos obtenido una matriz escalonada equivalente cuyo sistema de ecuaciones asociado nos permite resolverlo (de abajo hacia arriba):
$x_4 = 3$, $x_3 = 1$, $x_2 = 2$, $x_1 = -1$

Al resolver un sistema lineal recuerda que es posible que no tenga solución (sistema incompatible) si durante el proceso de eliminación se obtiene una fila con todo ceros menos el último elemento.

1.3.2 Eliminación de Gauss-Jordan

El método de eliminación de Gauss-Jordan continúa el proceso de la eliminación gaussiana hasta obtener una forma escalonada reducida.

Ejemplo

Resolver el sistema:

$\begin{cases} x - 2y + 3z = 9 \\ -x + 3y \quad\quad = -4 \\ 2x - 5y + 5z = 17 \end{cases}$ la matriz ampliada del sistema es $\begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$

En el ejemplo anterior, llegamos a la matriz escalonada:

$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Ahora, en lugar de utilizar la sustitución hacia atrás, aplicamos operaciones elementales hasta conseguir una forma escalonada reducida. Para ello, hemos de convertir en ceros todos los elementos situados encima de cada 1 dominante:

$\begin{bmatrix} 1 & 0 & 9 & 19 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ F1 + 2 F2 ⟶ F1

$\begin{bmatrix} 1 & 0 & 9 & 19 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ F2 +(-3)F3 ⟶ F2

$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ F1 + (-9)F3 ⟶ F1

El sistema línea asociado es:

$$\begin{cases} x = 1 \\ y = -1 \\ z = 2 \end{cases}$$

Tanto el método de eliminación gaussiana como el de Gauss-Jordan se adaptan bien en un ordenador. Sin embargo ninguno de los dos evita el uso de coeficientes racionales.

1.4 Sistemas homogéneos de ecuaciones lineales

Se denominan sistemas homogéneos de ecuaciones lineales a aquellos en que el término constante es igual a cero en todas las ecuaciones. Un sistema lineal homogéneo de m ecuaciones con n variables tiene la forma:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0 \end{array}$$

Todo sistema homogéneo tiene la menos una solución (aquella en que todas las variables toman el valor cero). A dicha solución se denomina **solución trivial**.

Teorema. Todo sistema homogéneo de ecuaciones lineales es compatible. Además, si el sistema tiene menos ecuaciones que variables, tiene infinitas soluciones.

Ejemplo. Sistema homogéneo con infinitas soluciones

Resolver el sistema:

$\begin{cases} x_1 - x_2 + 3x_3 = 0 \\ 2x_1 + x_2 + 3x_3 = 0 \end{cases}$ su matriz ampliada asociada es: $\begin{bmatrix} 1 & -1 & 3 & 0 \\ 2 & 1 & 3 & 0 \end{bmatrix}$

Aplicando la eliminación de Gauss-Jordan a la matriz ampliada:

$\begin{bmatrix} 1 & -1 & 3 & 0 \\ 0 & 3 & -3 & 0 \end{bmatrix}$ F2 + (-2)F1 ⟶ F2

$\begin{bmatrix} 1 & -1 & 3 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$ (1/3)F2 ⟶ F2

$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$ F1+F2 ⟶ F1

El sistema asociado a esta matriz es:

$$\begin{cases} x_1 \quad + 2x_3 = 0 \\ \quad x_2 - x_3 = 0 \end{cases}$$

Usando como parámetro $t = x_3$, el conjunto solución viene dado por:

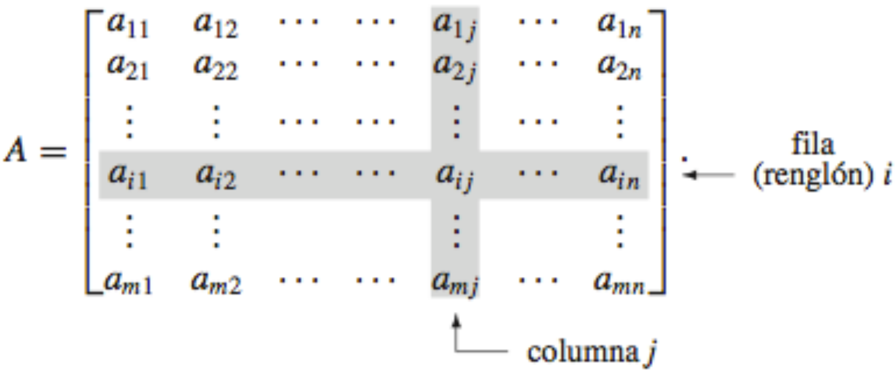
$x_1 = -2t \quad x_2 = t \quad x_3 = t$ Por tanto el sistema propuesto tiene infinitas soluciones, una de las cuales es la trivial (dada por $t = 0$).

---

2 MATRICES Y DETERMINANTES

2.1 Matrices. Definiciones

Una matriz A de $m \times n$ es una distribución rectangular de $m \cdot n$ números reales (o complejos) ordenados en m filas (renglones) horizontales y n columnas verticales:



La ***i*-ésima fila** de A es:

$$[a_{i1}, a_{i2}, a_{i3}, \cdots, a_{in}] \ (1 \leq i \leq m)$$

La ***j*-ésima columna** de A es:

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad (1 \leq j \leq n)$$

Diremos que A es ***m* por *n*** (que se escribe $\boldsymbol{m} \times \boldsymbol{n}$). Si $m = n$, decimos que A es una **matriz cuadrada de orden *n***, y que los números $a_{11}, a_{22}, \dots, a_{nn}$ forman la **diagonal principal** de A. Nos referimos al número a_{ij}, que está en la *i*-ésima fila (renglón) y la *j*-ésima columna de A, como el ***i*, *j*-ésimo elemento** de A, o la **entrada (*i*, *j*)** de A y solemos escribir como:

$$A = [a_{ij}]$$

Restringiremos nuestra atención al análisis de las matrices cuyas entradas son números reales. Sin embargo, también se estudian las matrices con entradas complejas, que tienen gran importancia en muchas aplicaciones.

Las matrices de $1 \times n$ ó $n \times 1$ también se denominan un ***n*-vectores**, y lo denotaremos mediante letras minúsculas en negritas. Cuando se sobreentienda el valor de n, nos referiremos a los *n*-vectores sólo como **vectores.**

$$\boldsymbol{u} = [\,1 \quad -1 \quad 0 \quad 4\,] \text{ es un 4-vector} \qquad \boldsymbol{v} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \text{ es un 3-vector}$$

Si todas las entradas de un *n*-vector son iguales a cero, se denota con **0**.

Observe que si A es una matriz de $n \times n$, las filas de A son matrices de $1 \times n$. El conjunto de todos los *n*-vectores con entradas reales se denota con $\mathbb{R}^n$.

Una matriz cuadrada $A = [a_{ij}]$ en donde cada término fuera de la diagonal principal es igual a cero, es decir, $a_{ij} = 0$ para $i \neq i$, es una **matriz diagonal**.

$$G = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix} \quad \text{y} \quad H = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

son matrices diagonales.

Una matriz diagonal A , en donde todos los términos de la diagonal principal son iguales, es decir, $a_{ij} = c$ para $i = j$ y $a_{ij} = 0$ para $i \neq j$, es una **matriz escalar**.

Las siguientes son matrices escalares:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad J = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

Dos matrices de $m \times n$, $A = [a_{ij}]$ y $B = [b_{ij}]$, son **iguales** si $a_{ij} = b_{ij}$, $1 \leq i \leq m, 1 \leq j \leq n$, es decir, si los elementos correspondientes son iguales.

Traspuesta de una matriz

Si $A = [a_{ij}]$ es una matriz de $m \times n$, la matriz traspuesta de A se denomina como $A^T = [a_{ij}^T]$ de $n \times m$, donde

$$a_{ij}^T = a_{ji}^T \ (1 \leq i \leq n, 1 \leq j \leq m)$$

En consecuencia, las entradas en cada fila de A^T son las entradas correspondientes en la columna de A.

Ejemplo:

Sea:

$$A =$$

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

entonces:

Matriz simétrica

Una matriz $A = [a_{ij}]$ cuyas entradas son números reales es **simétrica** si $A^T = A$. Es decir, A es simétrica si es una matriz cuadrada para la cual $a_{ij} = a_{ji}$

Si la matriz A es simétrica, los elementos de A son simétricos respecto de la diagonal principal de A.

2.2 Operaciones con matrices

2.2.1 Suma de matrices

Si $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices de $m \times n$, la **suma** de A y B da por resultado la matriz $C = [c_{ij}]$ de $m \times n$, definida por:

$c_{ij} = a_{ij} + b_{ij}$ $(i \leq i \leq m, 1 \leq j \leq n)$. Es decir, C se obtiene sumando los elementos correspondientes de A y B.

Ejemplo

Sean:

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

Entonces:

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

Observe que la suma de las matrices A y B sólo se define cuando A y B tienen el mismo número de filas (renglones) y el mismo número de columnas; es decir, sólo cuando A y B son del mismo tamaño.

Establecemos la convención, al escribir $A + B$ entendemos que A y B tienen el mismo tamaño.

2.2.2 Multiplicación por un escalar

Si $A = [a_{ij}]$ es una matriz de $m \times n$ y r es un número real, el **producto escalar** de A por r, rA, es la matriz $B = [b_{ij}]$ de $m \times n$, donde

$b_{ij} = r \cdot a_{ij}$ $(1 \leq i \leq m, 1 \leq j \leq n)$. Es decir, B se obtiene multiplicando cada elemento de A por r.

Si A y B son matrices de $m \times n$, escribimos $A + (-1)B$ como $A - B$, y denominamos a esto **diferencia** de A y B.

Ejemplo:

Sean:

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

Entonces:

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

Si $A_1, A_2, \ldots, A_k$ son matrices de $m \times n$ y $c_1, c_2, \ldots, c_k$ son números reales, entonces una expresión de la forma

$$c_1 A_1 + c_2 A_2 + \cdots + c_k A_k$$

se de nomina **combinación lineal** de $A_1, A_2, \ldots, A_k$, y $c_1, c_2, \ldots, c_k$ se llaman **coeficientes**.

Ejemplo:

Si:

$$A_1 = \begin{bmatrix} 0 & -3 & 5 \\ 2 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix} \quad \text{y} \quad A_2 = \begin{bmatrix} 5 & 2 & 3 \\ 6 & 2 & 3 \\ -1 & -2 & 3 \end{bmatrix}$$

entonces $C = 3 \cdot A_1 - ½\, A_2$ es una combinación lineal de A_1 y A_2. Por medio de la multiplicación por un escalar y la suma de matrices, podemos calcular C:

$$C = 3 \begin{bmatrix} 0 & -3 & 5 \\ 2 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 5 & 2 & 3 \\ 6 & 2 & 3 \\ -1 & -2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{5}{2} & -10 & \frac{27}{2} \\ 3 & 8 & \frac{21}{2} \\ \frac{7}{2} & -5 & -\frac{21}{2} \end{bmatrix}.$$

2.2.3 Producto de matrices

Si A es una matriz de $m \times p$, y B es una matriz de $p \times n$, el **producto** de A y B, que se denota mediante AB, es la matriz C de $m \times n$, definida como:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj} \ (1 \leq i \leq m, 1 \leq j \leq n)$$

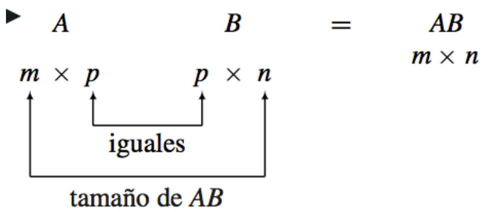
La ecuación dice que el elemento del producto AB situado en la fila i y en la columna j se obtiene multiplicando los elementos de la fila i de A y la columna j de B y sumando después los resultados; esto se muestra en la figura siguiente:

$$
\mathrm{fil}_i(A)\;\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix}
\begin{array}{c} \mathrm{col}_j(B) \\ \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pj} & \cdots & b_{pn} \end{bmatrix} \end{array}
$$

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

$$
\mathrm{fil}_i(A) \cdot \mathrm{col}_j(B) = \sum_{k=1}^{p} a_{ik} b_{kj}
$$

Observe que el producto de A y B sólo está definido cuando el número de filas de B es exactamente igual al número de columnas de A, como se indica en la figura siguiente:



Ejemplo:

Sean

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

Entonces:

$$
\begin{aligned}
AB &= \begin{bmatrix} (1)(-2) + (2)(4) + (-1)(2) & (1)(5) + (2)(-3) + (-1)(1) \\ (3)(-2) + (1)(4) + (4)(2) & (3)(5) + (1)(-3) + (4)(1) \end{bmatrix} \\
&= \begin{bmatrix} 4 & -2 \\ 6 & 16 \end{bmatrix}.
\end{aligned}
$$

2.2.4 Propiedades de las operaciones con matrices

2.2.4.1 Propiedades de la suma de matrices

Sean A, B, C y D matrices de $m \times n$.

a) $A + B = B + A$. (propiedad conmutativa)
b) $A + (B + C) = (A + B) + C$. (propiedad distributiva)
c) Existe una única matriz O de $m \times n$ tal que:

$$A + O = A$$

para cualquier matriz A de $m \times n$. La matriz O se denomina **neutro aditivo** de $m \times n$, **matriz nula** o **matriz cero**.

d) Para cada matriz A de $m \times n$, existe una única matriz D de $m \times n$ tal que:

$$A + D = O$$

Escribiremos D como $(-A)$, de modo que la ecuación anterior puede escribirse como

$$A + (-A) = O.$$

La matriz $(-A)$ se llama **inverso aditivo** o **negativo** de A.

2.2.4.2 Propiedades de la multiplicación de un escalar por una matriz

Sean A y B dos matrices del mismo tamaño y c, d dos números reales, entonces se verifica:

- $(c \cdot d) \cdot A = c \cdot (d \cdot A)$
- $1 \cdot A = A$
- $c \cdot (A + B) = c \cdot A + c \cdot B$
- $(c + d) \cdot A = c \cdot A + d \cdot A$

2.2.4.3 Propiedades de la multiplicación de matrices

(a) Si A, B y C son matrices de los tamaños apropiados, $A(BC) = (AB)C$.

(b) Si A, B y C son matrices de los tamaños apropiados, entonces $A(B + C) = AB + AC$.

(c) Si A, B y C son matrices de los tamaños apropiados, entonces $(A + B)C = AC + BC$.

Ejercicio: Demuestre que el producto de matrices no es conmutativo tomando dos matrices cualesquiera.

2.2.5 Matriz identidad, potencias de una matriz

La matriz escalar de $n \times n$:

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

cuyas entradas en la diagonal son todas iguales a 1, es la **matriz identidad de orden *n***.

Si A es una matriz de $m \times n$, es fácil verificar que: $I_m A = A I_n = A$.

También resulta sencillo ver que toda matriz escalar de $n \times n$ puede escribirse como rI_n para algún r.

Supongamos que *A* es una matriz cuadrada. Si *p* es un entero positivo, definimos las **potencias de una matriz** como sigue:

$$A^p = \underbrace{A \cdot A \cdots A}_{p\ factores}$$

Si A es de $n \times n$, también definimos:

$$A^0 = I_n.$$

En el caso de enteros no negativos *p* y *q*, algunas de las leyes conocidas de los exponentes de los números reales también pueden demostrarse para la multiplicación de una matriz cuadrada *A*

$$A^p A^q = A^{p+q}$$

$$(A^p)^q = A^{pq}$$

Obsérvese que:

$$(AB)^p \neq A^p B^p$$

para las matrices cuadradas en general. Sin embargo, si $AB = BA$, esta regla es válida.

A continuación llamamos la atención respecto de otras dos peculiaridades de la multiplicación de matrices.

Si *a* y *b* son números reales, *ab* = 0 se cumple sólo si *a* o *b* son cero. Sin embargo, esto no es válido para las matrices.
Si *a, b* y *c* son números reales para los cuales *ab* = *ac* y *a* ≠ 0, se tiene que *b* = *c*. Es decir, podemos cancelar *a*. Sin embargo, la ley de cancelación no se cumple para las matrices

2.2.6 Propiedades de la traspuesta

Si *r* es un escalar y A y B son matrices, entonces:

- $(A^T)^T = A$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$
- $(rA)^T = rA^T$

2.3 Inversa de una matriz

Consideremos la ecuación $ax = b$ entre números reales. Para resolverla, multiplicamos ambos lados por a^{-1}(supuesto que $a \neq 0$):

$$ax = b$$

$$a^{-1}ax = a^{-1}b$$

$$1x = a^{-1}b$$

$$x = a^{-1}b$$

El número a^{-1} es el inverso de a porque $a \cdot a^{-1} = 1$ (elemento identidad para la multiplicación). La definición de la inversa para las matrices es similar.

Definición. Una matriz A de $n \times n$ es **no singular** (o **invertible**) si existe una matriz B de $n \times n$ tal que $AB = BA = I_n$. La matriz B se denomina **inversa** de A. Si no existe tal matriz B, entonces B es **singular** (o **no invertible**).

Con base en la definición anterior, se deduce que $AB = BA = I_n$, por lo tanto, también A es una inversa de B.

> Teorema. Unicidad de la matriz inversa. *Si una matriz tiene inversa, la inversa es única y se denota por A^{-1}.*

Demostración: Sean B y C inversos de A. Entonces $BA = AC = I_n$. Por lo tanto,

$$B = BI_n = B \cdot (A \cdot C) = (B \cdot A) \cdot C = I_n \cdot C = C$$

Con lo cual concluye la demostración.

Escribiremos la inversa de A, si existe, como A^{-1}. Así,

$$AA^{-1} = A^{-1}A = I_n.$$

Ejemplo. Método para calcular la inversa de un matriz

Sea

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

Para determinar A^{-1} hacemos:

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Entonces, se debe cumplir que:

$$A \cdot A^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ de modo que } \begin{bmatrix} a + 2c & b + 2d \\ 3a + 4c & 3b + 4d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Al igualar las entradas correspondientes de estas dos matrices, obtenemos los sistemas lineales:

$$a + 2c = 1 \qquad b + 2d = 0$$

$$3a + 4c = 0 \qquad 3b + 4d = 1$$

cuyas soluciones son: $a = -2$, $c = 3/2$, $b = 1$ y $d = -1/2$. Además como la matriz:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

también se satisface la propiedad de que:

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

concluimos que A es no singular y que:

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

No todas las matrices tienen inversa.

2.3.1 Cálculo de la matriz inversa por eliminación de Gauss-Jordan

Sea A una matriz cuadrada de orden n.

1. Escribir la matriz $n\ x\ 2n$ $[A\ \vdots\ I]$ formada por la matriz A a la izquierda y la identidad I_n de orden n a su derecha separadas por una línea de puntos. A este proceso se le denomina unión de las matrices A e I.
2. Si es posible, reducir A a I mediante operaciones elementales por filas, que se deben aplicar a la matriz $[A\ \vdots\ I]$ completa. El resultado será la matriz $[I\ \vdots\ A^{-1}]$. Si no resulta posible esa reducción, A no es invertible.
3. Compruebe el resultado verificando que $AA^{-1} = A^{-1}A = I_n$.

Ejemplo.

Hallar la inversa de la matriz:

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

Empezamos añadiendo la matriz identidad:

$$[A\ \vdots\ I] = \begin{bmatrix} 1 & -1 & 0 & \vdots & 1 & 0 & 0 \\ 1 & 0 & -1 & \vdots & 0 & 1 & 0 \\ -6 & 2 & 3 & \vdots & 0 & 0 & 1 \end{bmatrix}$$

Ahora mediante operaciones elementales, reescribimos esta matriz en la forma $[I\ \vdots\ A^{-1}]$ como sigue:

$$\begin{bmatrix} 1 & -1 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & -1 & \vdots & -1 & 1 & 0 \\ -6 & 2 & 3 & \vdots & 0 & 0 & 1 \end{bmatrix} \qquad F2 + (-1)F1 \longrightarrow F2$$

$$\begin{bmatrix} 1 & -1 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & -1 & \vdots & -1 & 1 & 0 \\ 0 & -4 & 3 & \vdots & 6 & 0 & 1 \end{bmatrix} \qquad F3 + (6)F1 \longrightarrow F3$$

$$\begin{bmatrix} 1 & -1 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & -1 & \vdots & -1 & 1 & 0 \\ 0 & 0 & -1 & \vdots & 2 & 4 & 1 \end{bmatrix} \qquad F3 + (4)F2 \longrightarrow F3$$

$$\begin{bmatrix} 1 & -1 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & -1 & \vdots & -1 & 1 & 0 \\ 0 & 0 & 1 & \vdots & -2 & -4 & -1 \end{bmatrix} \qquad (-1)\ F3 \longrightarrow F3$$

$$\begin{bmatrix} 1 & -1 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 0 & \vdots & -3 & -3 & -1 \\ 0 & 0 & 1 & \vdots & -2 & -4 & -1 \end{bmatrix} \qquad F2 + F3 \longrightarrow F2$$

$$\begin{bmatrix} 1 & 0 & 0 & \vdots & -2 & -3 & -1 \\ 0 & 1 & 0 & \vdots & -3 & -3 & -1 \\ 0 & 0 & 1 & \vdots & -2 & -4 & -1 \end{bmatrix} \qquad F1 + F2 \longrightarrow F1$$

Por tanto A es invertible y su inversa es:

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

Si A no tiene inversa, el proceso lo detectará al obtener una fila de ceros en la matriz equivalente a A, como se ve a continuación:

Ejemplo

Probar que la siguiente matriz no tiene inversa:

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

Añadiendo la matriz identidad a A obtenemos:

$$[A \ \vdots\ I] = \begin{bmatrix} 1 & 2 & 0 & \vdots & 1 & 0 & 0 \\ 3 & -1 & 2 & \vdots & 0 & 1 & 0 \\ -2 & 3 & -2 & \vdots & 0 & 0 & 1 \end{bmatrix}$$

a la que aplicamos la eliminación de Gauss-Jordan:

$$\begin{bmatrix} 1 & 2 & 0 & \vdots & 1 & 0 & 0 \\ 0 & -7 & 2 & \vdots & -3 & 1 & 0 \\ -2 & 3 & -2 & \vdots & 0 & 0 & 1 \end{bmatrix} \qquad F2 + (-3)\, F1 \longrightarrow F2$$

$$\begin{bmatrix} 1 & 2 & 0 & \vdots & 1 & 0 & 0 \\ 0 & -7 & 2 & \vdots & -3 & 1 & 0 \\ 0 & 7 & -2 & \vdots & 2 & 0 & 1 \end{bmatrix} \qquad F3 + (2)\, F1 \longrightarrow F3$$

Ahora vemos que sumando la 2ª fila a la 3ª obtenemos una fila de ceros en el lado izquierdo de la matriz:

$$\begin{bmatrix} 1 & 2 & 0 & \vdots & 1 & 0 & 0 \\ 0 & -7 & 2 & \vdots & -3 & 1 & 0 \\ 0 & 0 & 0 & \vdots & -1 & 1 & 1 \end{bmatrix} \qquad F3 + F2 \longrightarrow F3$$

Como la "parte de A" de esa matriz tiene una fila nula, concluimos que no es posible reducir la matriz $[A \ \vdots\ I]$ a la forma $[I \ \vdots\ A^{-1}]$, lo que quiere decir que A no es invertible.

2.3.2 Propiedades de la matriz inversa

(a) Si A es una matriz no singular, entonces A^{-1} es no singular y

$$(A^{-1})^{-1} = A$$

(b) Si A y B son matrices no singulares, entonces AB es no singular y

$$(AB)^{-1} = B^{-1}A^{-1}.$$

(c) $(A^T)^{-1} = (A^{-1})^T$

Si k es un entero positivo y c un escalar, entonces A^k y cA son invertibles, con:

(d) $(A^k)^{-1} = A^{-1} \cdot A^{-1} \cdot \ldots \cdot A^{-1} = (A^{-1})^k$

(e) $(cA)^{-1} = \frac{1}{c} \cdot A^{-1}$

Demostración

(a) A^{-1}es no singular si podemos encontrar una matriz B tal que

$$A^{-1}B = BA^{-1} = I_n$$

Como A es no singular,

$$A^{-1}A = AA^{-1} = I_n$$

En consecuencia, $B = A$ es una inversa de A^{-1}, y como las inversas son únicas, concluimos que:

$$(A^{-1})^{-1} = A$$

Por tanto la inversa de la inversa de una matriz A no singular, es la misma matriz A.

(b) Tenemos

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

y

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$$

Por lo tanto, AB es no singular. Como la inversa de una matriz es única, concluimos que

$$(AB)^{-1} = B^{-1}A^{-1}$$

En consecuencia, la inversa de un producto de dos matrices no singulares es el producto de sus inversas en orden inverso.

(c) Tenemos

$$AA^{-1} = I_n \text{ y } \quad A^{-1}A = I_n$$

Al calcular las transpuestas, obtenemos

$$(AA^{-1})^T = {I_n}^T = I_n \text{ y } \; (A^{-1}A)^T = {I_n}^T = I_n.$$

Entonces

$$(A^{-1})^T A^T = I_n \quad \text{y} \quad A^T(A^{-1})^T = I_n$$

Estas ecuaciones implican que:

$$(A^T)^{-1} = (A^{-1})^T$$

En consecuencia, la inversa de la transpuesta de una matriz no singular, es la traspuesta de su inversa.

Corolario. Si $A_1, A_2, \dots, A_r$ son matrices no singulares de $n \times n$, entonces $A_1 A_2 \dots A_r$ es no singular y:

$$(A_1 A_2 \dots A_r)^{-1} = A_r^{-1} \cdot A_{r-1}^{-1} \cdot \dots \cdot A_1^{-1}$$

Propiedades de la cancelación

Si C es una matriz invertible, son válidas las siguientes propiedades:

Si $AC = BC$, entonces $A = B$

Si $CA = CB$, entonces $A = B$

Demostración. Para demostrar la propiedad 1 usamos que C es invertible de tal modo que:

$$AC = BC$$

$$(AC)C^{-1} = (BC)C^{-1}$$

$$A(CC^{-1}) = B(CC^{-1})$$

$$AI = BI$$

$$A = B$$

La segunda propiedad se demuestra de forma análoga.

2.3.3 Matrices y sistemas de ecuaciones lineales

Un sistema de m ecuaciones lineales con n incógnitas:

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

Puede escribirse como producto de matrices de la siguiente forma: $A\boldsymbol{x} = \boldsymbol{b}$

Donde:

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

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Como A es una matriz (matriz de coeficientes) de $m\ x\ n$ y $\boldsymbol{x}$ es una matriz de $n\ x\ 1$, su producto $A\boldsymbol{x}$ es una matriz de $m\ x\ 1.$

Teorema. Si A es una matriz invertible, el sistema de ecuaciones lineales $A\boldsymbol{x} = \boldsymbol{b}$ tiene solución única dada por:

$$\boldsymbol{x} = A^{-1}\,\boldsymbol{b}$$

Demostración.

Como A no es singular, los pasos siguientes son válidos:

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$A^{-1}A\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

$$I\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

Esta solución es única porque si $\boldsymbol{x_1}$ y $\boldsymbol{x_2}$ son soluciones, la propiedad de cancelación en $A\boldsymbol{x_1} = \boldsymbol{b} = A\boldsymbol{x_2}$ permite concluir que $\boldsymbol{x_1} = \boldsymbol{x_2}$

El teorema anterior tiene importancia teórica, pero no resulta práctico para resolver sistemas lineales. Exige más trabajo calcular A^{-1} y multiplicar después por b que resolver el sistema por eliminación gaussiana con posterior sustitución hacia atrás.

Sistema homogéneo asociado

Dado el sistema de ecuaciones lineales $A\boldsymbol{x} = \boldsymbol{b}$, $(\boldsymbol{b} \neq \boldsymbol{0})$ se denomina sistema homogéneo asociado al formado por el sistema: $A\boldsymbol{x} = \boldsymbol{0}$.

Teorema. Sean $\boldsymbol{x_1}$ y $\boldsymbol{x_2}$ soluciones a un sistema no homogéneo $A\boldsymbol{x} = \boldsymbol{b}$. Entonces su diferencia $\boldsymbol{x_1} - \boldsymbol{x_2}$ es solución del sistema homogéneo asociado $A\boldsymbol{x} = \boldsymbol{0}$.

Demostración

$$A(\boldsymbol{x_1} - \boldsymbol{x_2}) = A\boldsymbol{x_1} - A\boldsymbol{x_2} = b - b = 0$$

Corolario. Sea $\boldsymbol{x}$ e $\boldsymbol{y}$ soluciones particulares del sistema no homogéneo $A\boldsymbol{x} = \boldsymbol{b}$. Entonces existe una solución $\boldsymbol{h}$ al sistema homogéneo asociado, tal que

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{h}$$

Demostración.

Si $\boldsymbol{h}$ está definida como $\boldsymbol{h} = \boldsymbol{y} - \boldsymbol{x}$, entonces $\boldsymbol{h}$ es una solución del sistema homogéneo asociado a $A\boldsymbol{x} = \boldsymbol{b}$ por el teorema anterior, por lo que $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{h}$

Con objeto de encontrar todas las soluciones a un sistema no homogéneo $A\boldsymbol{x} = \boldsymbol{b}$, basta con encontrar una solución, que llamaremos solución particular (x_p), y todas las soluciones al sistema homogéneo asociado $A\boldsymbol{x} = \boldsymbol{0}$, que llamaremos solución homogénea (x_h).

Ejemplo.

Encuentre todas las soluciones al sistema de ecuaciones lineales:

$$\begin{cases} x_1 + 2x_2 - x_3 = 2 \\ 2x_1 + 3x_2 + 5x_3 = 5 \\ -x_1 - 3x_2 + 8x_3 = -1 \end{cases}$$

Primero encontramos una solución mediante la reducción por filas:

La matriz ampliada del sistema es: $\begin{bmatrix} 1 & 2 & -1 & 2 \\ 2 & 3 & 5 & 5 \\ -1 & -3 & 8 & -1 \end{bmatrix}$

Haciendo F2 ⟶ F2 – 2F1 y F3 ⟶ F3 + F1 obtenemos:

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

Haciendo F1 ⟶ F1+2F2 y F3 ⟶ F3 – F2 obtenemos:

$$\begin{bmatrix} 1 & 0 & 13 & 4 \\ 0 & -1 & 7 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las ecuaciones correspondientes a los primeros dos renglones del último sistema son:

$$x_1 = 4 - 13\, x_3 \quad \text{y} \quad x_2 = -1 + 7\, x_3$$

con lo que las soluciones son:

$$\boldsymbol{x} = (x_1, x_2, x_3) = (4 - 13\, x_3, -1 + 7\, x_3\ , x_3) = \boldsymbol{x}_p + \boldsymbol{x}_h$$

donde $x_p = (4, -1, 0)$ es una solución particular y $x_h = x_3\,(-13, 7, 1)$ donde x_3 es un número real, es una solución al sistema homogéneo asociado. Así $x_3 = 0$ lleva a la solución $(-4, 1, 0)$ y $x_3 = 1$ da la solución $(-9, 15, 2)$.

2.3.4 Matrices elementales

Se consideran operaciones elementales sobre las filas de una matriz:

- Intercambiar dos filas.
- Multiplicar una fila por un número no nulo.
- Sumar un múltiplo de una fila a otra fila.

Definición. Un matriz $n\,x\,n$ se llama **matriz elemental** si se puede obtener de la matriz identidad I_n mediante *una única operación elemental por filas*.

La matriz identidad I_n es elemental, según esta definición, ya que puede obtenerse de sí misma multiplicando una de sus filas por 1.

Ejemplos.

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

a) Elemental. Se obtiene multiplicando la segunda fila de I_3 por 3.
b) No elemental, porque no es cuadrada.
c) No elemental. Se ha obtenido multiplicando la tercera fila de I_3 por 0 (que no es una operación elemental.
d) Elemental. Se obtiene intercambiando las filas 2ª y 3ª de I_3.
e) Elemental. Obtenida de I_2 sumando a la segunda fila la primera multiplicada por 2.
f) No elemental, porque son necesarias dos operaciones elementales por filas para obtenerla a partir de I_3.

Notación. Una matriz elemental se denota por E, o por $cR_i, R_j + cR_i$, o por P_{ij} de acuerdo con la forma en que se obtuvo de I. En este caso, P_{ij} (la matriz de permutación) es la matriz obtenida a partir del intercambio de las filas de i y j de I.

Las matrices elementales son útiles porque permiten expresar las operaciones elementales por filas como producto de matrices, como ilustra el siguiente ejemplo.

2.3.4.1 Matrices elementales y operaciones elementales por filas.

En el siguiente producto, E es la matriz elemental obtenida de I_3 intercambiando sus dos primeras filas:

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

$$E \qquad A \qquad = \qquad E \cdot A$$

Nótese que las dos primeras filas de A han resultado intercambiadas al multiplicarla por E a la izquierda.

> Teorema: Sea E la matriz elemental obtenida al efectuar sobre las filas de I_n una operación elemental. El resultado de efectuar esa misma operación elemental sobre las filas de A viene dado por el producto $E \cdot A$

Nótese que A debe multiplicarse por E por la izquierda.

2.3.4.2 Definición de equivalencia por filas

Sean A y B matrices $m\,x\,n$. Se dice que B es **equivalente por filas** a la matriz A si existe una secuencia finita de matrices elementales $E_1, E_2, \ldots, E_k$, tal que:

$$B = E_k\, E_{k-1} \ldots E_2 E_1 A$$

2.3.4.3 Inversa de una matriz elemental

> Teorema. Todas las matrices elementales son invertibles, además la inversa de una matriz elemental es otra matriz elemental.

Para hallar la inversa de una matriz elemental E basta invertir la operación elemental utilizada para obtener E.

$$(cF_i)^{-1} = \frac{1}{c}F_i$$

$$(F_j + cF_i)^{-1} = F_j - cF_i$$

$$(P_{ij})^{-1} = P_{ij}$$

Toda matriz de permutación elemental es su propia inversa

Matriz elemental tipo E	Efecto de multiplicar A por la izquierda por E	Representación simbólica de las operaciones elementales	Al multiplicar por la izquierda, E^{-1} hace lo siguiente	Representación simbólica de la operación inversa
Multiplicación	Multiplica la fila i de A por $c \neq 0$	cF_i	Multiplica la fila i de A por $\frac{1}{c}$	$\frac{1}{c}F_i$
Suma	Multiplica la fila i de A por c y lo suma a la fila j	$F_j + cF_i$	Multiplica la fila i de A por $-c$ y lo suma a la fila j	$F_j - cF_i$
Permutación	Permuta las filas i y j de A	P_{ij}	Permuta las filas i y j de A	P_{ij}

Tabla 1. Matrices elementales y sus inversas

Teorema. Una matriz cuadrada A es invertible, si y solo si, se puede escribir como producto de matrices elementales.

Demostración. Para demostrar la segunda parte, supongamos que A es producto de matrices elementales. En tal caso, como toda matriz elemental es invertible y el producto de invertibles es invertible, concluimos que A es invertible.

Para probar la primera parte, supongamos A invertible. Por el teorema visto en 2.3.3 sabemos que el sistema lineal representado por $Ax = O$ tiene solo la solución trivial. Eso implica que la matriz ampliada $[A \ \vdots \ O]$ se puede llevar a la forma $[I \ \vdots \ O]$ (usando operaciones elementales asociadas a$E_1, E_2, \dots, E_k$. Por tanto:

$$E_k\, E_{k-1}\ \dots\ E_2E_1A = I$$

Y, por consiguiente:

$$A = E_1^{-1} \cdot E_2^{-1} \cdot \dots \cdot E_k^{-1}$$

Así pues, A se puede expresar como producto de matrices elementales con lo que la demostración está completa.

Ejemplo

Hallar la secuencia de matrices elementales cuyo producto sea la matriz:

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

Solución

Comenzamos buscando una secuencia de operaciones elementales por filas que lleven A a forma escalonada reducida:

$$\begin{bmatrix}1 & 2\\3 & 8\end{bmatrix} \qquad (-1)\cdot F1 \longrightarrow F1 \qquad E_1 = \begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 2\\0 & 2\end{bmatrix} \qquad F2 + (-3)\cdot F1 \longrightarrow F2 \qquad E_2 = \begin{bmatrix}1 & 0\\-3 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 2\\0 & 1\end{bmatrix} \qquad \left(\frac{1}{2}\right)\cdot F2 \longrightarrow F2 \qquad E_3 = \begin{bmatrix}1 & 0\\0 & \frac{1}{2}\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \qquad F1 + (-2)\cdot F2 \longrightarrow F1 \qquad E_4 = \begin{bmatrix}1 & -2\\0 & 1\end{bmatrix}$$

Ahora el producto de matrices:

$E_4 \cdot E_3 \cdot E_2 \cdot E_1 A = I$ despejamos $A = E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot E_4^{-1} \cdot I = E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot E_4^{-1} \cdot$

Esto implica que A es un producto de matrices elementales:

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

Teorema. Condiciones equivalentes

Si A es una matriz cuadrada de orden n, las siguientes afirmaciones son equivalentes:

1. A es invertible
2. $A\boldsymbol{x} = \boldsymbol{b}$ tiene solución única para toda matriz columna $\boldsymbol{b}$ de tamaño $n \, x \, 1$.
3. $A\boldsymbol{x} = \boldsymbol{O}$ solo tiene la solución trivial.
4. A es equivalente por filas a I_n.
5. A se puede escribir como producto de matrices elementales.

2.3.5 LU-factorización

Una matriz cuadrada se denomina **triangular superior** (y se denota con U de *upper*) si todas sus componentes debajo de la diagonal son cero. Es una matriz **triangular inferior** (y se denota con L de *lower*) si todas sus componentes encima de la diagonal son cero. Una matriz se denomina **diagonal** si todos los elementos que no se encuentran sobre la diagonal son cero; es decir, $A = [a_{ij}]$ es triangular superior si $a_{ij} = 0$ para $i > j$, triangular inferior si $a_{ij} = 0$ para $i < j$ y diagonal si $a_{ij} = 0$ para $i \neq j$. Observe que una matriz diagonal es tanto triangular superior como triangular inferior.

$$[\mathbf{A}]_L = \begin{bmatrix} a_{11} & 0 & 0 & 0 & \cdots & 0\\ a_{21} & a_{22} & 0 & 0 & \cdots & 0\\ a_{31} & a_{32} & a_{33} & 0 & \cdots & 0\\ a_{41} & a_{42} & a_{43} & a_{44} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & a_{m3} & a_{m4} & \cdots & a_{mn} \end{bmatrix} \qquad [\mathbf{A}]_U = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n}\\ 0 & a_{22} & a_{23} & a_{24} & \cdots & a_{2n}\\ 0 & 0 & a_{33} & a_{34} & \cdots & a_{3n}\\ 0 & 0 & 0 & a_{44} & \cdots & a_{4n}\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & a_{mn} \end{bmatrix}$$

Figura 1. Matrices triangular inferior (izquierda) y triangular superior (derecha)

Definición de LU-factorización. Si la matriz cuadrada de orden n, A se puede escribir como producto de una matriz triangular inferior L con todos los elementos de su diagonal principal valiendo 1 y una matriz triangular superior U, se dice que $A = LU$ es una LU-factorizacion de A.

Si una matriz cuadrada A se puede reducir por filas a una matriz triangular superior U *usando sólo operaciones elementales consistentes en sumar un múltiplo de una fila a otra fila*, es muy fácil encontrar una LU-factorización de A, basta con recordar las sucesivas operaciones por filas realizadas.

Este método nos facilita la resolución de ecuaciones lineales por ordenador al reducir el número de operaciones, como se verá más adelante.

Nótese que para que ello sea posible se debe cumplir que ninguno de los elementos de la diagonal principal de A sea nulo. En caso de que no fuera así, deberíamos primeramente intercambiar sus filas de forma que el elemento nulo dejara de estar situado en la diagonal principal, obteninedo una nueva matriz A' equivalente a la anterior, sobre la que buscaríamos su factorización $A' = LU$.

Ejemplo. LU-factorización de una matriz

Hallar la LU-factorización de la matriz $A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$

Solución.

En primer lugar, reducimos A a forma triangular superior, guardando constancia de la matriz elemental asociada a cada operación elemental realizada.

Matriz obtenida	Operación elemental	Matriz elemental E	Matriz elemental inversa E^{-1}
$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ -2 & 5 & 5 \end{bmatrix}$	$F2 \longrightarrow F2 + (-2) \cdot F1$	$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix}$	$F3 \longrightarrow F3 + F1$	$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$	$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U$	$F3 \longrightarrow F3 + 2 \cdot F2$	$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$	$E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$

La matriz U de la izquierda es triangular superior y se sigue que $E_3 E_2 E_1 A = U$, o sea $A = E_1{}^{-1} E_2{}^{-1} E_3{}^{-1} U$. Como el producto de matrices triangulares inferiores es también triangular inferior:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} = L \qquad {E_1}^{-1}{E_2}^{-1}{E_3}^{-1} = L$$

La factorización $A = LU$ está conseguida.

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

En general, si A se puede reducir por filas a una matriz triangular superior U usando sólo la operación de sumar un múltiplo de una fila a otra, entonces A admite LU-factorización:

$$E_k\, E_{k-1}\, \dots\, E_2 E_1 A = U$$

$$A = E_1^{-1} \cdot E_2^{-1} \cdot \dots \cdot E_k^{-1} U$$

$$A = LU$$

Donde L es el producto de las inversas de las matrices elementales usadas en la reducción.

Nótese que la matriz L puede obtenerse sin necesidad de encontrar las matrices elementales, construyendo la matriz L de la siguiente forma:

Partiendo de la matriz $L = \begin{bmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$ vamos añadiendo las operaciones elementales cambiando el signo al escalar que multiplica a la fila en cada caso:

Matiz U	Operación elemental	Matriz L
$\begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$		$\begin{bmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$
$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ -2 & 5 & 5 \end{bmatrix}$	$F2 \longrightarrow F2 + (-2) \cdot F1$	$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ * & * & 1 \end{bmatrix}$
$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix}$	$F3 \longrightarrow F3 + F1$	$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & * & 1 \end{bmatrix}$
$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U$	$F3 \longrightarrow F3 + 2 \cdot F2$	$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}$

El procedimiento utilizado en el ejemplo se puede llevar a cabo mientras no se requieran permutaciones para poder reducir A a la forma triangular, lo que no siempre es factible.

Procedimiento rápido para obtener la descomposición LU:

1. Reducir la matriz a su forma escalonada equivalente mediante el método de Gauss sin emplear permutaciones de filas y tomando nota de los multiplicadores empleados para obtener los pivotes y los empleados para obtener ceros debajo de los pivotes.
2. En cada posición de la diagonal principal de L poner los inversos de los multiplicadores empleados para obtener los pivotes.
3. En cada posición por debajo de la diagonal principal de L poner los multiplicadores empleados para obtener cero en cada posición cambiándole el signo.

Nótese que cada una las matrices elementales empleadas en la LU-factorización es una matriz triangular inferior. Esto se deduce del hecho de que E es de la forma $F_j + cF_i$ (no hay permutaciones). Más aún, los números que se hacen cero en la reducción por filas están siempre *debajo* de la diagonal de manera que en $F_j + cF_i$ siempre se cumple que $j > i$. De este modo, las c aparecen debajo de la diagonal.

Teorema. Propiedades de la multiplicación de matrices triangulares.

El producto de las matrices triangulares inferiores es una matriz triangular inferior. Más aún, el producto de matrices triangulares superiores es una matriz triangular superior.

Teorema de la LU-factorización

Sea A una matriz cuadrada $(n\ x\ n)$ y suponga que A se puede reducir por filas a una matriz triangular U sin hacer alguna permutación entre sus renglones ni multiplicando sus filas por una constante. Entonces existe una matriz triangular inferior L invertible tal que $A = LU$. Si, además, U tiene n pivotes (es decir, A es invertible), entonces esta factorización es única.

Demostración

U y L se obtienen como en el ejemplo. Sólo es necesario probar la unicidad en el caso de que A sea invertible. Como U tiene n pivotes, su forma escalonada por renglones también tiene n pivotes (para verificar esto divida cada renglón de U por el pivote en esa fila). Entonces, U es invertible.

Para demostrar que L es invertible, considere la ecuación $L\boldsymbol{x} = \boldsymbol{0}$.

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Se deduce que $x_1 = 0$, $a_{21}x_1 + x_2 = 0$, etc., lo que demuestra que $x_1 = x_2 = \cdots = x_n = 0$ y L es invertible. Para demostrar la unicidad, suponga que $A = L_1U_1 = L_2U_2$. Entonces

$$U_1U_2^{-1} = (L_1^{-1}L_1)(U_1U_2^{-1}) = L_1^{-1}(L_1U_1)U_2^{-1} = L_1^{-1}(L_2U_2)U_2^{-1} = (L_1^{-1}L_2)(U_2U_2^{-1}) = L_1^{-1}L_2$$

$L_1^{-1}L_2$ es una matriz triangular inferior con unos en la diagonal mientras que $U_1U_2^{-1}$es triangular superior. La única forma en que una matriz triangular superior y una inferior pueden ser iguales es si ambas son diagonales. Como $L_1^{-1}L_2$tiene unos en la diagonal se ve que:

$$U_1U_2^{-1} = L_1^{-1}L_2 = I$$

De lo que se deduce que:

$$U_1 = U_2 \quad y\, L_1 = L_2$$

2.3.5.1 Resolución de sistemas lineales mediante LU-factorización

Suponga que se quiere resolver el sistema $A\boldsymbol{x} = \boldsymbol{b}$, donde A es invertible. Si A satisface la hipótesis del teorema de la LU-factorización, se puede escribir:

$$LU\boldsymbol{x} = \boldsymbol{b}$$

Como L es invertible, existe un vector único $\boldsymbol{y}$ tal que $L\boldsymbol{y} = \boldsymbol{b}$. Como U también es invertible, existe un vector único $\boldsymbol{x}$ tal que $U\boldsymbol{x} = \boldsymbol{y}$. Entonces $A\boldsymbol{x} = L(U\boldsymbol{x}) = L\boldsymbol{y} = \boldsymbol{b}$ y nuestro sistema está resuelto. Observa que $L\boldsymbol{y} = \boldsymbol{b}$ se puede resolver directamente mediante la sustitución hacia adelante, mientras que el sistema $U\boldsymbol{x} = \boldsymbol{y}$ se puede resolver por sustitución hacia atrás. Esto se ilustra en el siguiente ejemplo.

Ejemplo

Resolver el sistema $A\boldsymbol{x} = \boldsymbol{b}$ donde:

$$A = \begin{bmatrix} 2 & 3 & 2 & 4 \\ 4 & 10 & -4 & 0 \\ -3 & -2 & -5 & -2 \\ -2 & 4 & 4 & -7 \end{bmatrix} \quad y \quad b = \begin{bmatrix} 4 \\ -8 \\ -4 \\ -1 \end{bmatrix}$$

Solución

Comenzamos con la LU-factorización de la matriz A.

$$\begin{bmatrix} 2 & 3 & 2 & 4 \\ 4 & 10 & -4 & 0 \\ -3 & -2 & -5 & -2 \\ -2 & 4 & 4 & -7 \end{bmatrix} \begin{matrix} F2 \longrightarrow F2 - 2F1 \\ F3 \longrightarrow F3 + \frac{3}{2} \cdot F1 \\ F4 \longrightarrow F4 + F1 \end{matrix} \begin{bmatrix} 2 & 3 & 2 & 4 \\ 0 & 4 & -8 & -8 \\ 0 & 5/2 & -2 & 4 \\ 0 & 7 & 6 & -3 \end{bmatrix} \begin{matrix} F3 \longrightarrow F3 - \frac{5}{8}F2 \\ F4 \longrightarrow F4 - \frac{7}{4}F2 \end{matrix} \begin{bmatrix} 2 & 3 & 2 & 4 \\ 0 & 4 & -8 & -8 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 20 & 11 \end{bmatrix}$$

$$F4 \longrightarrow F4 - \frac{20}{3}F3 \quad \begin{bmatrix} 2 & 3 & 2 & 4 \\ 0 & 4 & -8 & -8 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 0 & -49 \end{bmatrix} = U$$

Usando las matrices elementales se puede escribir:

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{20}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -\frac{7}{4} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{5}{8} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{3}{2} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} A$$

Despejando A y calculando las inversas de las seis matrices elementales:

$$A=\begin{bmatrix}1&0&0&0\\2&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\-\frac{3}{2}&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\-1&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&\frac{5}{8}&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&\frac{7}{4}&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&\frac{20}{3}&1\end{bmatrix}U$$

Se ha escrito A como un producto de seis matrices elementales y una matriz triangular superior. Sea L el producto de las matrices elementales. Debe verificar que:

$L=\begin{bmatrix}1&0&0&0\\2&1&0&0\\-\frac{3}{2}&\frac{5}{8}&1&0\\-1&\frac{7}{4}&\frac{20}{3}&1\end{bmatrix}$, que se trata de una matriz triangular inferior con unos en la diagonal principal.

Después se puede escribir $A = LU$, donde L es triangular inferior y U es triangular superior. Los elementos de la diagonal de L son todos iguales a 1 y los elementos de la diagonal de U son los pivotes.

El sistema $L\boldsymbol{y} = \boldsymbol{b}$ conduce a las ecuaciones:

$$\left\{\begin{array}{ccccccccc} y_1 & & & & & & & = & 4\\ 2y_1 & + & y_2 & & & & & = & -8\\ -\frac{3}{2}y_1 & + & \frac{5}{8}y_2 & + & y_3 & & & = & -4\\ -y_1 & + & \frac{7}{4}y_2 & + & \frac{20}{3}y_3 & + & y_4 & = & -1 \end{array}\right.$$

de donde realizando la sustitución hacia atrás obtenemos:

$$\begin{cases} y_1 = 4\\ y_2 = -8 - 2y_1 = -16\\ y_3 = -4 + \frac{3}{2}y_1 - \frac{5}{8}y_2 = 12\\ y_4 = -1 + y_1 - \frac{7}{4}y_2 - \frac{20}{3}y_3 = -49 \end{cases}$$ es decir $y=\begin{bmatrix}4\\-16\\12\\-49\end{bmatrix}$

De $U\boldsymbol{x} = \boldsymbol{y}$ se obtiene:

$$\left\{\begin{array}{ccccccccc} 2x_1 & + & 3x_2 & + & 2x_3 & + & 4x_4 & = & 4\\ & & 4x_2 & - & 8x_3 & - & 8x_4 & = & -16\\ & & & & 3x_3 & + & 9x_4 & = & 12\\ & & & & & & 49x_4 & = & -49 \end{array}\right.$$

de donde sustituyendo hacia delante se obtiene:

$$\begin{cases} x_4 = 1 \\ 3x_3 = 12 - 9x_4 = 3, de\ donde\ x_3 = 1 \\ 4x_2 = -16 + 8x_3 + 8x_4 = 0, de\ donde\ x_2 = 0 \\ 2x_1 = 4 - 3x_2 - 2x_3 - 4x_4 = -2, de\ donde\ x_1 = -1 \end{cases}$$

Por tanto la solución al sistema es:

$$X = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Esta descomposición conduce a un algoritmo para resolver un sistema lineal $A\boldsymbol{x} = \boldsymbol{b}$, que es el más utilizado en los ordenadores para resolver sistemas lineales. La razón principal por la que este método es tan utilizado, radica en que proporciona la manera más económica de resolver un sistema lineal en el que tenemos que cambiar de manera repetida el lado derecho (columna de términos independientes). Este tipo de situación suele presentarse en problemas de aplicación. Por ejemplo, una compañía de servicio eléctrico debe determinar las entradas (las incógnitas) que necesita para producir algún resultado requerido (los términos independientes). Las entradas y los resultados pueden estar relacionados por un sistema lineal, cuya matriz de coeficientes es fija, mientras que el lado derecho cambia día con día, o incluso cada hora.

Si *A* es de $n\ x\ n$, entonces el número total de operaciones (multiplicaciones y divisiones) requeridas para resolver un sistema lineal $A\boldsymbol{x} = \boldsymbol{b}$ usando una factorización *LU* de *A*, es $T(n) \approx {n^3}/{3}$, el mismo que se requiere para eliminación gaussiana. Esto no es extraño pues la fase de eliminación hacia delante produce la factorización LU en $\approx {n^3}/{3}$ pasos, mientras que las sustituciones hacia delante y hacia atrás requieren $\approx {n^2}/{2}$ pasos. Por tanto, para valores grandes de n, el término ${n^3}/{3}$ es dominante. Desde este punto de vista la eliminación gaussiana y la factorización LU son equivalentes.

Sin embargo, la factorización LU tiene otras ventajas:

1. Desde el punto de vista de almacenamiento, la factorización LU es muy compacta, porque puede sobrescribir las entradas de A con entradas de L y U conforme se calculan.

Por ejemplo si:

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

Esto puede almacenarse como:

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

con las entradas colocadas en el orden $(1,1), (1,2), (1,3), (2,1), (3,1), (2,2), (2,3), (3,2), (3,3)$. En otras palabras, las entradas subdiagonales de *A* se sustituyen con los multiplicadores correspondientes.

2. Una vez calculada la factorización *A* de *LU*, puede usarse para resolver tantos sistemas lineales de la forma $A\boldsymbol{x} = \boldsymbol{b}$ como se quiera.

3. Para matrices con ciertas formas especiales, sobre todo aquellas con un gran número de ceros (las llamadas matrices "dispersas") concentrados fuera de la diagonal, existen métodos que simplificarán el cálculo de una factorización *LU*. En estos casos, este método es más rápido que la eliminación gaussiana para resolver $A\boldsymbol{x} = \boldsymbol{b}$.

4. Para una matriz invertible *A*, puede usarse una factorización *LU* de *A* para encontrar A^{-1}, si es necesario. Más aún, esto puede hacerse en tal forma que simultáneamente produzca una factorización de A^{-1} .

2.4 Determinantes

2.4.1 Definición de determinante de una matriz de 2 x 2

Sea $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ una matriz de 2 x 2

Se define determinante de $A, det\ (A)$ = $a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$

A menudo se denota como $|A|$ ó $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$

El determinante de una matriz puede ser positivo, cero o negativo.

El determinante de una matriz de orden 1 se define simplemente como el valor de su único elemento.

Para definir el determinante de una matriz de orden mayor que 2, conviene usar las nociones de *menores* y *cofactores*.

2.4.2 Definición de los menores y cofactores de una matriz

Si A es una matriz cuadrada, el **menor** M_{ij} del elemento a_{ij} es el determinante[2] de la matriz obtenida al eliminar en A la fila i y la columna j. El **cofactor** C_{ij} de ese mismo elemento viene dado por:

$$C_{ij} = (-1)^{i+j} M_{ij}$$

Por ejemplo, si A es una matriz de 3 x 3, algunos de sus menores y cofactores son los que se muestran a continuación:

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

[2] Algunos autores denominan menor a la matriz obtenida al eliminar la fila i y la columna j, no al determinante. Entonces el cofactor se define como $(-1)^{i+j} \cdot M_{ij}$

$M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$ Cofactor de $a_{21} = C_{21} = (-1)^{2+1} M_{21} = -M_{21}$

$M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$ Cofactor de $a_{22} = C_{22} = (-1)^{2+2} M_{22} = M_{22}$

Los menores y cofactores de una matriz difieren a lo sumo en el signo. Para hallar los cofactores de una matriz, calculamos primero los menores y, a continuación, aplicamos la siguiente regla: si la suman de la *fila + columna* del elemento es par, el signo del cofactor es positivo, mientras que será negativo si la suma es impar.

2.4.3 Determinante de una matriz de orden 3

El determinante de una matriz de orden 3 se define mediante:

$$|A| = a_{11} \cdot C_{11} + a_{12} \cdot C_{12} + a_{13} \cdot C_{13} = a_{11} \cdot M_{11} - a_{12} \cdot M_{12} + a_{13} \cdot M_{13}$$

2.4.4 Determinante de una matriz de n x n

Sea A una matriz de orden n x n, entonces el determinante de A está dado indistintamente por:

$$|A| = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = \sum_{k=0}^{n} a_{ik}C_{ik} \qquad \textit{Desarrollo por la fila } i$$

$$|A| = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} = \sum_{k=0}^{n} a_{kj}C_{kj} \qquad \textit{Desarrollo por la columna } j$$

Nótese que cuando desarrollamos en cofactores no es necesario evaluar los cofactores de los elementos iguales a cero, porque cero multiplicado por el cofactor dará cero. Por esa razón suele desarrollarse por la fila o columna que tiene mayor número de ceros como vemos en el siguiente ejemplo.

Ejemplo. Determinante de una matriz de orden 4

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

$$\det(A) = 3 \cdot C_{13} + 0 \cdot C_{23} + 0 \cdot C_{33} + 0 \cdot C_{43}$$

Por tanto solo debemos calcular:

$C_{13} = (-1)^{(1+3)} \cdot \begin{vmatrix} -1 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 4 & -2 \end{vmatrix}$ = (desarrollando en cofactores por la segunda fila) =

= 0·(-1)$^{(2+1)}$·$\begin{vmatrix} 1 & 2 \\ 4 & -2 \end{vmatrix}$ + 2·(-1)$^{(2+2)}$·$\begin{vmatrix} -1 & 2 \\ 3 & -2 \end{vmatrix}$ + 3·(-1)$^{(2+3)}$·$\begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} = 0 + 2 \cdot (-4) + 3 \cdot (-1) \cdot (-7) = 13$

Por tanto $det\,(A) = 3 \cdot (13) = 39$

2.4.5 Cálculo de determinantes.

2.4.5.1 Teorema. Determinante de matrices triangulares

Sea $A = [a_{ij}]$ una matriz cuadrada de orden n triangular superior o inferior. Entonces:

$$\boxed{|A| = a_{11} \cdot a_{22} \cdot a_{33} \cdot \ \ldots\ldots \ \cdot a_{nn}}$$

Esto es: *el determinante de una matriz triangular es igual al producto de sus componentes en la diagonal.*

Demostración.

Se demostrará la parte triangular superior por inducción matemática comenzando con $n = 2$. Si A es una matriz triangular superior de $2 \ x \ 2$, entonces $A = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}$ y $\det(A) = a_{11} \cdot a_{22} - a_{12} \cdot 0 = a_{11} \cdot a_{22}$, de manera que el teorema se cumple para $n = 2$. Se supondrá que se cumple para $k = n - 1$ y se demostrará para $k = n$. El determinante de una matriz triangular superior de $n \ x \ n$ es:

$$\begin{vmatrix} a_{11} & a_{11} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} 0 & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}$$

$$+ a_{13} \begin{vmatrix} 0 & a_{22} & \cdots & a_{2n} \\ 0 & 0 & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} 0 & a_{22} & \cdots & a_{2,n-1} \\ 0 & 0 & \cdots & a_{3,n-1} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{vmatrix}$$

Cada uno de estos determinantes es el determinante de una matriz triangular superior de $(n - 1) \ x \ (n - 1)$ que, de acuerdo con la hipótesis de inducción, es igual al producto de las componentes en la diagonal. Todas las matrices, excepto la primera, tienen una columna de ceros, por lo que por lo menos una de sus componentes diagonales es cero. De este modo, todos los determinantes, excepto el primero, son cero. Por último,

$$\det A = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}(a_{22}a_{33} \cdots a_{nn})$$

lo que prueba que el teorema se cumple para matrices de $n \ x \ n$.

2.4.5.2 Cálculo de determinantes mediante operaciones elementales

Operaciones elementales y determinantes

Sean A y B matrices cuadradas.

1. Si B se obtiene intercambiando dos filas (columnas) de A

$$det(B) = -det(A)$$

2. Si B se obtiene sumando un múltiplo de una fila (columna) de A a otra fila de A,

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

3. Si B se obtiene multiplicando una fila (columna) de A por una constante $c \neq 0$,

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

Esta última propiedad permite sacar un factor común de una fila. Por ejemplo:

$$\begin{vmatrix} 4 & 12 \\ 1 & 5 \end{vmatrix} = 4 \cdot \begin{vmatrix} 1 & 3 \\ 1 & 5 \end{vmatrix}$$

Estas tres propiedades nos permiten calcular determinantes efectuando operaciones elementales por filas hasta obtener una matriz triangular equivalente. En cada paso del proceso las propiedades anteriores nos indican el efecto de la operación elemental sobre el determinante. Finalmente el determinante se calcula haciendo el producto de los elementos de su diagonal principal, tal como vemos en el siguiente ejemplo:

Calcular el determinante de:

$$A = \begin{bmatrix} 2 & -3 & 10 \\ 1 & 2 & -2 \\ 0 & 1 & -3 \end{bmatrix}$$

Mediante operaciones elementales por filas, llevamos A a forma triangular:

$\begin{vmatrix} 2 & -3 & 10 \\ 1 & 2 & -2 \\ 0 & 1 & -3 \end{vmatrix}$ = (Intercambiar las dos primeras filas)= - $\begin{vmatrix} 1 & 2 & -2 \\ 2 & -3 & 10 \\ 0 & 1 & -3 \end{vmatrix}$ =

= (Sumar -2 veces la primera fila a la segunda produce una nueva 2ª fila)= - $\begin{vmatrix} 1 & 2 & -2 \\ 0 & -7 & 14 \\ 0 & 1 & -3 \end{vmatrix}$ =

= (Sacar factor -7 de la segunda fila) = 7· $\begin{vmatrix} 1 & 2 & -2 \\ 0 & 1 & -2 \\ 0 & 1 & -3 \end{vmatrix}$ = (sustituir la 3ª fila por la diferencia entre la 2ª y 3ª filas) =7· $\begin{vmatrix} 1 & 2 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{vmatrix}$ = (como la matriz es triangular) $= 7 \cdot 1 \cdot 1 \cdot (-1) = -7$

2.4.5.3 Condiciones que producen determinante cero

Si A es una matriz cuadrada y cumple cualquiera de las condiciones siguientes, entonces $det(A) = 0$.

- Hay una fila (o columna) completa de ceros
- Dos filas (o columnas) son iguales
- Una fila (o columna) es múltiplo de otra fila (o columna)

Estas condiciones no son las únicas que hacen que un determinante sea cero. Muchas veces, se aplican indirectamente, es decir, podemos comenzar con una matriz que no cumple ninguna de esas tres condiciones y llegar, efectuando operaciones elementales por filas o columnas, a otra que sí satisface alguna de ellas. En tales casos, podemos concluir que la matriz tiene determinante cero.

Al calcular a mano un determinante, se puede ahorrar esfuerzo si se consigue llevar la matriz, mediante operaciones elementales, a una forma en la que una de sus filas (o columnas) tenga todos sus elementos cero menos uno. El desarrollo en cofactores por esa fila (o columna) reduce el orden de la matriz en 1, tal como se ilustra en el siguiente ejemplo:

Ejemplo.

Calcular el determinante de

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

Solución. La matriz A ya tiene un cero en la tercera fila. Podemos crear otro sumando 2 veces la primera columna a la tercera:

$$\det(A) = \begin{vmatrix} -3 & 5 & 2 \\ 2 & -4 & -1 \\ -3 & 0 & 6 \end{vmatrix} = \begin{vmatrix} -3 & 5 & -4 \\ 2 & -4 & 3 \\ -3 & 0 & 0 \end{vmatrix}$$

Desarrollando en cofactores por la tercera fila, obtenemos:

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

2.4.6 Propiedades de los determinantes

2.4.6.1 Determinante de un producto de matrices

Teorema. Si A y B son matrices cuadradas de orden n, se cumple que:

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

Demostración.

Para empezar, observemos que si E es una matriz elemental (ver 2.3.4) por 2.4.5.2 son ciertas las siguientes afirmaciones:

- Si E se obtiene de I por intercambio de dos filas, $det(E) = -1$
- Si E se obtiene de I multiplicando una fila de I por una constante $c \neq 0$, $det\,(E) = c$
- Si E se obtiene de I sumando un múltiplo de una fila de I a otra fila, $det\,(E) = 1$

Además por el teorema de 2.3.4.1, si se aplica a B la misma operación elemental que convierte I en E, se obtiene la matriz EB. De todo ello se sigue que:

$$det\,(EB) = det\,(E) \cdot det\,(B)$$

Esto se puede generalizar para concluir que:

$$\det(E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot B) = \det(E_k) \cdot \ldots \det(E_2) \cdot \det(E_1) \cdot \det\,(B)$$

donde E_i es una matriz elemental. Consideremos ahora la matriz AB. Si A no es singular, por el segundo teorema de 2.3.4.3 podemos expresarla como producto de matrices elementales:

$A = E_k \cdot \ldots \cdot E_2 \cdot E_1$ y escribir:

$$\det(AB) = \det(E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot B) = \det(E_k) \cdot \ldots \det(E_2) \cdot \det(E_1) \cdot \det(B)$$
$$= \det(E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot B) = det(A) \cdot det(B).$$

Si A es singular, A es equivalente a una matriz con toda la fila de ceros. Del teorema 2.4.5.3 se deduce que $det(A) = 0$. Además como A es singular, también AB lo es. (Si AB no fuera singular, $A[B(AB)^{-1}] = I$ implicaría que A es no singular). Así pues, $det(AB) = 0$, luego $det(AB) = det(A) \cdot det(B)$.

___

Nota. El determinante de la suma no siempre es igual a la suma de los determinantes. Es decir, para la mayoría de los pares de matrices, A y B,

$$det\ (A + B) \neq det\ (A) + det\ (B)$$

Utilizando la factorización LU de una matriz cuadrada A de $n \: x \: n$ se tiene $A = LU$ (véase 2.3.5). Entonces, por el teorema anterior,

$$det\ (A) = det\ (LU) = det\ (L)\ det\ (U)$$

Pero L es una matriz triangular inferior con unos en la diagonal, así

$det\ (L)$ = producto de los elementos en la diagonal = 1

De manera similar, como U es triangular superior,

$det\ (U)$ = producto de los elementos en la diagonal

Entonces se tiene el siguiente teorema:

Si una matriz cuadrada A tiene la factorización LU, $A = LU$ donde L tiene unos en la diagonal, entonces

$det\ (A) = det\ (U)$ = producto de los elementos de la diagonal de U

2.4.6.2 Determinante de un múltiplo escalar de un matriz

Teorema. Si A es una matriz cuadrada de orden n y c un escalar,

$$det(cA) = c^n \cdot det(A)$$

Esta fórmula se puede demostrar aplicando repetidamente la propiedad 3 de 2.4.5.2. Es decir, sacando factor c de cada una de las n filas de $|cA|$.

2.4.6.3 Determinante de una matriz invertible

Teorema. Una matriz cuadrada A es invertible (no singular) si y solo sí

$$\det\ (A) \neq 0$$

Demostración.

a). Supongamos que A es invertible. Entonces $\exists\ A^{-1}$ tal que $AA^{-1} = I$, tomando determinantes a ambos lados de la igualdad y aplicando las propiedades del producto: $det(A) \cdot det(A^{-1}) = det(I) = 1$, de lo que se concluye que $\det(A) \neq 0$.

b) Supongamos que $\det(A) \neq 0$. Por eliminación de Gauss-Jordan, podemos hallar una matriz B, en forma escalonada reducida, equivalente por filas a A. Como B está en forma escalonada reducida, o es la matriz identidad (I) o tiene al menos una fila completa de ceros. Pero si tuviera una fila entera de ceros, $det(B) = 0$, lo que implicaría que $det(A) = 0$. Puesto que hemos partido de que $\det(A) \neq 0$, podemos concluir que $B = \mathrm{I}$, por lo que A es equivalente por filas a la matriz identidad y, por tanto, invertible.

2.4.6.4 Determinante de la matriz inversa

> Teorema. Si A es invertible, $\det(A^{-1}) = \frac{1}{\det(A)}$

Demostración.

Si A es invertible, existe A^{-1} tal que $AA^{-1} = I$. Por el teorema de 2.4.6.1

$$det(AA^{-1}) = det(A) \cdot det(A^{-1}) = det(I) = 1$$

Además, por ser A invertible sabemos que $\det(A) \neq 0$, así que podemos dividir ambos términos de la igualdad anterior por $det(A)$ resultando:

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

2.4.6.5 Condiciones equivalentes para matrices no singulares

Si A es una matriz $n\ x\ n$, las siguientes afirmaciones son equivalentes:

1. A es invertible
2. $A\boldsymbol{x} = \boldsymbol{b}$ tiene solución única para toda matriz $\boldsymbol{b}$ de tamaño $n\ x\ 1$
3. $A\boldsymbol{x} = \boldsymbol{0}$ tiene solo la solución trivial
4. A es equivalente por filas a I_n
5. A se puede escribir como producto de matrices elementales
6. $\det(A) \neq 0$

Como se vio en 2.4.5.3, una matriz cuadrada tiene determinante cero sí y sólo sí, es equivalente por filas a una matriz con al menos una fila completa de ceros La validez de esta afirmación se deduce de la equivalencia de las propiedades 4 y 6 anteriores.

2.4.6.6 Determinante de la matriz traspuesta

> Teorema. Si A es una matriz cuadrada,
>
> $$det(A) = det(A^T)$$

Dado que las filas de una matriz son las columnas de su traspuesta, se deduce que los teoremas se cumplen también para las columnas.

2.4.6.7 Otras propiedades de los determinantes

A continuación enumeramos algunas propiedades de los determinantes. La demostración se deja al lector.

1. Si cualquier fila o columna de A es un vector cero, entonces $det\ A = 0$.
2. Si la fila i o columna j de A se multiplica por un escalar c, entonces $det\ A$ se multiplica por c.
3. Si A, B y C son idénticas excepto por la columna j y que la columna j de C es la suma de las j-ésimas columnas de A y B. Entonces, $det\ C = det\ A + det\ B$. La misma afirmación es cierta para filas.
4. El intercambio de cualesquiera dos filas (o columnas) distintos de A tiene el efecto de multiplicar el determinante de A por -1.
5. Si una matriz cuadrada A tiene dos filas (o columnas) iguales, entonces su determinante es cero.
6. Si una fila (columna) de A es un múltiplo escalar de otra fila (columna), entonces $det\ A = 0$.
7. Si se suma un múltiplo escalar de una fila (columna) de A a otra fila (columna) de A, entonces el determinante no cambia.

Las propiedades que se acaban de presentar simplifican la evaluación de determinantes de alto orden. Se "reduce por filas" el determinante, usando la propiedad 7, hasta que tenga una forma en la que se pueda evaluar con facilidad. La meta más común será utilizando la propiedad 7 de manera repetida hasta que:

a) el nuevo determinante tenga un renglón (columna) de ceros o un renglón (columna) que sea múltiplo de otro —en cuyo caso el determinante es cero—,o
b) que la nueva matriz sea triangular, con lo que su determinante será el producto de sus elementos en la diagonal.

Ejemplo

Calcular $det\ (A)$ si $A = \begin{bmatrix} -2 & 1 & 0 & 4 \\ 3 & -1 & 5 & 2 \\ -2 & 7 & 3 & 1 \\ 3 & -7 & 2 & 5 \end{bmatrix}$

Solución

Existen varias formas de proceder en este caso y no es evidente cuál de ellas será la más rápida para llegar a la respuesta. Sin embargo, como ya existe un cero en la primera fila, se comienza la reducción en esa fila.

Se multiplica la segunda columna por 2 y por -4 y se suma a la primera y cuarta columnas, respectivamente:

$$|A| = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 5 & 6 \\ 12 & 7 & 3 & -27 \\ -11 & -7 & 2 & 33 \end{vmatrix}$$

Se intercambian las primeras dos columnas.

$$= -\begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 5 & 6 \\ 7 & 12 & 3 & -27 \\ -7 & -11 & 2 & 33 \end{vmatrix} =$$

Se multiplica la segunda columna por -5 y por -6 y se suma a la tercera y cuarta columnas, respectivamente.

$$= -\begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 7 & 12 & -57 & -99 \\ -7 & -11 & 57 & 99 \end{vmatrix}$$

Como la cuarta columna es ahora un múltiplo de la tercera (columna 4 = $\frac{99}{57}$ x columna 3) se ve que $|A|$ = 0.

2.4.7 Aplicaciones de los determinantes

2.4.7.1 Matriz adjunta

Recordemos que el cofactor C_{ij} de un matriz A ha quedado definido como $(-1)^{i+j}$ veces el determinante de la matriz obtenida al eliminar en A la fila i y la columna j. Si A es una matriz cuadrada, la **matriz de cofactores** de A tiene la forma:

$$\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n1} & \cdots & C_{nn} \end{bmatrix}$$

La traspuesta de esta matriz se llama **matriz adjunta** de A y se denota por $adj(A)$, es decir:

$$adj(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

Teorema. Si A es una matriz $n\ x\ n$ invertible:

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

Demostración.

En primer lugar probaremos que: $A \cdot adj(A) = det(A) \cdot I_n$

Consideremos el producto: $A \cdot [adj(A)]$. En este producto, el elemento de la fila i y la columna j viene dado por:

$$a_{i1} \cdot C_{j1} + a_{i2} \cdot C_{j2} + \ldots\ldots.. + a_{in} \cdot C_{jn} \quad (1)$$

Si $i = j$, esta suma es simplemente el desarrollo en cofactores de A por la fila i, es decir, el determinante de A. Por otra parte, si $i \neq j$, la suma es cero.

En efecto, sea:

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{i1} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \longleftarrow \text{fila j}$$

Entonces, como dos filas de B son iguales, $det\ B = 0$. Pero $B = A$ excepto por la fila *j*. Calculando el determinante de B expandiendo la fila j de B, obtenemos la suma (1) anterior. Obsérvese que al hacer la expansión respecto de la fila j, esta fila se elimina al calcular los cofactores de B. Así $B_{jk} = A_{jk}$ para $k = 1, 2, \dots, n$

En consecuencia:

$$A \cdot [adj(A)] = \begin{bmatrix} \det(A) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \det(A) \end{bmatrix} = \det(A) \cdot I$$

Como A es invertible, $\det(A) \neq 0$, así que podemos escribir:

$$\frac{1}{det(A)} \cdot A \cdot [adj\ (A)] = I_n \quad \Rightarrow \quad A \cdot \left[\frac{1}{\det(A)} \cdot [\text{adj}\ (A)]\right] = I_n$$

Multiplicando a la izquierda en ambos lados de la igualdad por A^{-1}:

$$A^{-1} \cdot A \cdot \left[\frac{1}{\det(A)} \cdot [\text{adj}\ (A)]\right] = A^{-1} I_n \Rightarrow \left[\frac{1}{\det(A)} \cdot [\text{adj}\ (A)]\right] = A^{-1}$$

2.4.7.2 Regla de Cramer

La regla de Cramer, es una fórmula que utiliza determinantes para resolver sistemas de n ecuaciones lineales en n variables. Es aplicable sólo a sistemas lineales con solución única.

> Teorema. Regla de Cramer
>
> Si la matriz de coeficientes de un sistema de n ecuaciones con n variables tiene determinante no nulo, la única solución del sistema viene dada por:
>
> $$X_1 = \frac{\det(A_1)}{\det(A)},\ X_2 = \frac{\det(A_2)}{\det(A)},\ \dots\dots\dots\dots,\ X_n = \frac{\det(A_n)}{\det(A)}$$
>
> Donde la columna i de A_i es la columna de términos independientes del sistema.

Demostración.

Consideremos el sistema representado por $AX = B$. Puesto que $\det(A) \neq 0$, podemos escribir:

$$X = A^{-1}B = \frac{1}{\det(A)} \cdot adj(A) \cdot B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Si $b_1, b_2, \dots \dots, b_n$ son los elementos de B, x_i viene dada por:

$$x_i = \frac{1}{|A|} \cdot (b_1 C_{1i} + b_2 C_{2i} + \cdots + b_n C_{ni})$$

Pero la suma entre paréntesis es justamente el desarrollo en cofactores de A_i, lo cual significa que

$x_i = \frac{|A_i|}{|A|}$, como queríamos demostrar.

Ejemplo.

Usar la regla de Cramer para hallar el valor de x en el sistema lineal:

$$\begin{cases} -x + 2y - 3z = 1 \\ 2x \qquad + \ z = 0 \\ 3x - 4y + 4z = 2 \end{cases}$$

Solución

El determinante de la matriz de coeficientes es:

$$|A| = \begin{vmatrix} -1 & 2 & -3 \\ 2 & 0 & 1 \\ 3 & -4 & 4 \end{vmatrix} = 10 \qquad\qquad b = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

Como $|A| \neq 0$, la solución es única. Podemos hallar x aplicando la regla de Cramer:

$$x = \frac{\begin{vmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 2 & -4 & 4 \end{vmatrix}}{10} = \frac{(-1)^5}{10} \cdot \begin{vmatrix} 1 & 2 \\ 2 & -4 \end{vmatrix} = \frac{(-1)\cdot(-8)}{10} = \frac{4}{5}$$

2.5 Observaciones sobre el cálculo de determinantes mediante ordenador

Hasta ahora, hemos desarrollado dos métodos para resolver un sistema lineal de *n* ecuaciones con *n* incógnitas: la reducción de Gauss-Jordan y la regla de Cramer. También tenemos dos métodos para obtener la inversa de una matriz no singular: el método que involucra determinantes y el método presentado en la sección 2.3.1 Se nos plantea pues la pregunta sobre qué criterios que deben considerarse para elegir uno u otro de estos métodos.

Casi todos los problemas de gran tamaño de álgebra lineal se resuelven usando ordenadores, de modo que es natural comparar dos métodos estimando el tiempo requerido por los cálculos para el mismo problema. Como la suma es mucho más rápida que la multiplicación, con frecuencia se utiliza el número de multiplicaciones como base para comparar dos procedimientos numéricos.

Considere el sistema lineal $A\boldsymbol{x} = \boldsymbol{b}$, donde A es de 25 × 25. Si hallamos $\boldsymbol{x}$ por medio de la regla de Cramer, primero debemos obtener $det(A)$. Podemos hacer esto mediante un desarrollo por cofactores, digamos $det(A) = a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1}$, donde hemos desarrollado $det(A)$ mediante los cofactores de la primera columna. Obsérvese que si se dispone de cada cofactor, se requieren 25 multiplicaciones. Ahora, cada cofactor A_{ij} es más (+) o menos (–) el determinante de una matriz de 24 × 24, el cual puede desarrollarse a lo largo de una fila o una columna dadas, requiriendo para ello 24 multiplicaciones. Entonces, el cálculo de $det(A)$ requiere más de $25 \times 24 \times \ldots \times 2 \times 1 = 25!$ multiplicaciones. Aunque empleáramos un ordenador del futuro (futuro no tan lejano) capaz de realizar un billón (1×10^{12}) de multiplicaciones *por segundo* $(3{,}15 \times 10^{19}$ *por año*), tardaría *cerca de 49.000 años* en evaluar $det(A)$. Por otro lado, la reducción de Gauss-Jordan necesita cerca de $25^3/3$ multiplicaciones, y hallaríamos la solución en menos de *un segundo*.

Por supuesto, podemos calcular $det(A)$ de una manera mucho más eficiente utilizando operaciones elementales por filas para reducir A a su forma triangular, y entonces aplicar el teorema 2.4.5.1. Al usar este método para una matriz de *n* × *n*, la regla de Cramer requiere aproximadamente n^4 multiplicaciones, en comparación con las $n^3/3$ multiplicaciones necesarias para la reducción de Gauss-Jordan. En consecuencia, la reducción de Gauss-Jordan sigue siendo mucho más rápida.

En general, si estamos buscando respuestas numéricas, podemos emplear cualquier método que involucra determinantes, si $n \leq 4$. Para $n \geq 5$, los métodos que utilizan determinantes son menos eficientes que la reducción de Gauss-Jordan y que el método de la sección 2.3.1 para invertir una matriz.

Por supuesto, la importancia de los determinantes no recae en su uso computacional. Ten en cuenta que los métodos con determinantes permiten expresar la inversa de una matriz y la solución de un sistema de *n* ecuaciones lineales con *n* incógnitas por medio de expresiones o fórmulas. La reducción de Gauss-Jordan y el método para determinar A^{-1}, analizado en la sección 2.3.1, no proporcionan una *fórmula* para la respuesta; a fin de hallarla, debemos proceder en forma numérica. A veces no es necesaria una respuesta numérica, sino una expresión para dicha respuesta, pues tal vez se quiera seguir utilizándola. Otra razón importante para el estudio de los determinantes es que éstos desempeñan un papel importante en el estudio de los valores y vectores propios, que será abordado en el capítulo 5.

3 Espacios vectoriales

3.1 Espacios vectoriales

Sea V un conjunto en el que se han definido dos operaciones:

- Suma de vectores
- Producto de un vector por un escalar

Se dice que V tiene estructura de espacio vectorial sobre $\mathbb{R}$ si se verifica para todo $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ y para todo $\lambda, \mu \in \mathbb{R}$ las siguientes propiedades:

1. $\boldsymbol{u}+\boldsymbol{v} \in V$	La suma es cerrada
2. $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$	Propiedad conmutativa
3. $\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})=(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}$	Propiedad asociativa
4. $\exists\, \boldsymbol{o} \in V, \forall\, \boldsymbol{u} \in V, \boldsymbol{u}+\boldsymbol{o}=\boldsymbol{u}$	Elemento neutro de la suma
5. $\forall\, \boldsymbol{u} \in V, \exists -\boldsymbol{u} \in V, \boldsymbol{u}+(-\boldsymbol{v})=\boldsymbol{o}$	Elemento inverso de la suma
6. $\lambda\boldsymbol{u} \in V$	Multiplicación por un escalar es cerrada
7. $\lambda\,(\boldsymbol{u}+\boldsymbol{v})=\lambda\boldsymbol{u}+\lambda\boldsymbol{v}$	Propiedad distributiva
8. $(\lambda+\mu)\,\boldsymbol{u}=\lambda\boldsymbol{u}+\mu\boldsymbol{v}$	Propiedad distributiva
9. $\lambda(\mu\boldsymbol{u})=(\lambda\mu)\,\boldsymbol{u}$	Propiedad asociativa
10. $1\cdot\boldsymbol{u}=\boldsymbol{u}$	Identidad escalar

Ejemplos

1. Puede demostrase que definidos en $\mathbb{R}^2$ los vectores como pares ordenados de números $\boldsymbol{v}=(v_1,v_2)$ y las operaciones:

- Suma de vectores: $\boldsymbol{u}+\boldsymbol{v}=(u_1,u_2)+(v_1,v_2)=(u_1+v_1,u_2+v_2)$
- Producto por un escalar: $\lambda\,\boldsymbol{u}=(\lambda u_1,\ \lambda u_2)$

Dotan a $\mathbb{R}^2$ de estructura de espacio vectorial sobre $\mathbb{R}$ con el vector $\boldsymbol{o}=(0,0)$ como elemento neutor de la suma y $(-u_1,-u_2)$ como elemento inverso.

2. El conjunto de las matrices de tamaño mxn, con las operaciones habituales de suma de matrices y producto por un escalar, posee estructura de espacio vectorial sobre $\mathbb{R}$.

3. El conjunto de los polinomios de coeficientes reales de grado menor o igual que n con la suma y multiplicación por un numero real habituales.

4. El conjunto de los polinomios de coeficientes reales de grado n no es un espacio vectorial ya que, por ejemplo, la suma de x^n+2x+1 y $-x^n+3x$ da como resultado $5x+1$ que no es un polinomio de grado n, es decir la suma no es una operación cerrada.

4. El conjunto de las funciones reales continuas en toda la recta real, con las operaciones:

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

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

3.1.1 Propiedades de la multiplicación por un escalar

Sea V un espacio vectorial sobre $\mathbb{R}$, si $\boldsymbol{u}, \boldsymbol{v} \in V$ y $\lambda, \mu \in \mathbb{R}$, se cumplen las siguientes propiedades:

1. $0\boldsymbol{v}=\boldsymbol{o}$
2. $\lambda\boldsymbol{o}=\boldsymbol{o}$

3. $Si\ \lambda\boldsymbol{v} = \boldsymbol{o} \implies \boldsymbol{\lambda} = 0$ ó $\boldsymbol{v} = \boldsymbol{o}$
4. $-(\lambda\boldsymbol{u}) = (-\boldsymbol{\lambda})\boldsymbol{u} = \boldsymbol{\lambda}(-\boldsymbol{u})$
5. $(\lambda\boldsymbol{u} = \mu\boldsymbol{u}$ y $\boldsymbol{u} \neq \boldsymbol{o}) \Rightarrow \lambda = \mu$
6. $(\lambda\boldsymbol{u} = \boldsymbol{\lambda v}$ y $\boldsymbol{\lambda} \neq 0) \Rightarrow \boldsymbol{u} = \boldsymbol{v}$

Demostración:

1. $0\boldsymbol{v} = (0+0)\boldsymbol{v} = 0\boldsymbol{v} + 0\boldsymbol{v}, luego\ \boldsymbol{v} = \boldsymbol{o}$
2. $\lambda\boldsymbol{o} = \lambda(\boldsymbol{o}+\boldsymbol{o}) = \lambda\boldsymbol{o} + \lambda\boldsymbol{o}, luego\ \lambda\boldsymbol{o} = \boldsymbol{o}$
3. Si fuese $\lambda \neq 0, existiría\ \lambda^{-1}\ y\ podría\ escribirse\ \lambda^{-1}(\lambda\boldsymbol{v}) = \lambda^{-1}\boldsymbol{o} = \boldsymbol{o}, es\ decir, \boldsymbol{o} = (\lambda^{-1}\lambda)\boldsymbol{u} = 1\boldsymbol{u} = \boldsymbol{u}$
4. $(-\lambda)\boldsymbol{u} + \lambda\boldsymbol{u} = (-\lambda+\lambda)\boldsymbol{u} = 0\boldsymbol{u} = \boldsymbol{o}$ y $\lambda(-\boldsymbol{u}) + \lambda\boldsymbol{u} = \lambda(-\boldsymbol{u}+\boldsymbol{u}) = \lambda\boldsymbol{o} = \boldsymbol{o}$
5. Como $\lambda\boldsymbol{u} - \mu\boldsymbol{u} = \boldsymbol{o}$, es decir, $(\lambda-\mu)\boldsymbol{u} = \boldsymbol{o}$**,** por ser $\boldsymbol{u} \neq \boldsymbol{o}$ ha de ser $(\lambda-\mu) = 0$
6. Como $\lambda\boldsymbol{u} - \lambda\boldsymbol{v} = \boldsymbol{o}$, es decir, $\lambda(\boldsymbol{u}-\boldsymbol{v}) = \boldsymbol{o}$, por ser $\lambda \neq 0$, ha de ser $(\boldsymbol{u}-\boldsymbol{v}) = \boldsymbol{o}$.

3.2 Subespacios vectoriales

Un subconjunto no vacío W de un espacio vectorial V, es un subespacio vectorial de V si W es espacio vectorial con las mismas operaciones de suma y producto por un escalar definidas en V.

Para que W pueda ser subespacio de V, las dos operaciones han de ser cerradas en W.

Para concluir que un conjunto es espacio vectorial, hay que comprobar los diez axiomas (3.1). Sin embargo al ser W un subconjunto de un espacio vectorial V basta con verificar que las dos operaciones son cerradas en W, es decir:

Teorema. Un subconjunto no vacío W de un espacio vectorial V es subespacio de V sí, y sólo sí, se satisfacen las dos condiciones:

1. $\boldsymbol{u}+\boldsymbol{v} \in W, \forall\ \boldsymbol{u},\boldsymbol{v} \in W$
2. $\lambda\boldsymbol{u} \in W, \forall\boldsymbol{u} \in W\ y\ \forall\lambda \in \mathbb{R}$

Demostración

a) En una dirección la demostración es inmediata, si W es subespacio de V , entonces W es un espacio vectorial, y deben ser cerradas la suma y la multiplicación por un escalar.

b) Para demostrar el teorema en la otra dirección, supongamos que la suma y la multiplicación por un escalar son cerradas en W. Nótese que si $\boldsymbol{u},\boldsymbol{v}$ y $\boldsymbol{w} \in W$, también pertenecen a V por lo que los axiomas, 2, 3, 7, 8, 9 y 10 de espacio vectorial se cumplen automáticamente. Además como la suma y la multiplicación por un escalar son cerradas en W:

$\forall\ \boldsymbol{w} \in W$ y con $\lambda = 0$, se tiene que:

$$\lambda\boldsymbol{w} = \boldsymbol{o}$$

$$(-1)\boldsymbol{w} = -\boldsymbol{w}$$

ambos pertenecen a W, luego se satisfacen los axiomas 4 y 5.

Si W es un subespacio del espacio vectorial V, ambos deben tener el mismo vector $\boldsymbol{o}$**.** De hecho el subespacio vectorial más pequeño de un espacio vectorial, es aquél que posee como único vector el vector nulo y se denomina como ***subespacio cero***: $W = \{\boldsymbol{o}\}$**.**

Otro subespacio vectorial de V es el propio V. Todo espacio vectorial contiene estos dos subespacios, denominados **triviales**. Los subespacios distintos de estos dos se llaman **subespacios propios o no triviales**.

Ejemplos

1. El conjunto de matrices simétricas de orden 2, es subespacio del espacio vectorial de las matrices de orden 2, con las operaciones usuales de suma de matrices y multiplicación por un escalar.

La suma de matrices simétricas, produce una matriz simétrica y el producto por un escalar también produce una matriz simétrica ya que $\forall\ A, B \in \mathcal{M}_2^s$ y $\forall \lambda \in \mathbb{R}$:

$$(A+B)^T = A^T + B^T = A + B$$

$$(\lambda A)^T = \lambda A^T = \lambda A$$

2. El conjunto de matrices singulares de orden 2 no es subespacio del espacio vectorial de las matrices de orden 2, porque, por ejemplo, la suma de dos matrices:

$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ singulares , da como resultado:

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

que es no singular.

3.2.1 Intersección de subespacios vectoriales.

Si V y W son subespacios del espacio vectorial U, la intersección de V y W (denotada por $V \cap W$) es también un subespacio de U.

Demostración

a) Por ser V y W subespacios de U, ambos contienen el vector cero, luego $V \cap W$ es no vacío.

b) Para ver que la suma es cerrada en $V \cap W$, sean $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ dos vectores de $V \cap W$. Al ser V y W subespacios de U, la suma es cerrada en ambos. Por tanto, como $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ están en V, $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in V$. Análogamente, como$\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ están en W, $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in W$. Lo que implica que $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in V \cap W$, de modo que en $V \cap W$ la suma es cerrada.

c) De forma análoga se demuestra que la mutliplicación por un escalar es cerrada en $V \cap W$.

3.2.2 Subespacios de $\mathbb{R}^n$

¿Cuál de estos conjuntos es subespacio de $\mathbb{R}^2$?

a) El conjunto de los puntos de la recta $x + 2y = 0$
b) El conjunto de puntos de la recta $x + 2y = 1$

Solución

a) Un punto de la recta $x + 2y = 0$ es de la forma $(-2t, t)$, donde t es cualquier número real.

Sean $\boldsymbol{v_1} = (-2t_1, t_1)$ y $\boldsymbol{v_2} = (-2t_2, t_2)$, dos puntos cualesquiera de la recta. Entonces:

$$\boldsymbol{v_1} + \boldsymbol{v_2} = (-2t_1, t_1) + (-2t_2, t_2) = (-2(t_1 + t_2), t_1 + t_2) = (-2t_3, t_3)$$

donde $t_3 = t_1 + t_2$. Así pues, $\boldsymbol{v_1} + \boldsymbol{v_2}$ está en la recta y, por tanto, la suma es cerrada. De forma análoga se demuestra que la multiplicación por un escalar es cerrada en este conjunto y, por tanto es subespacio de $\mathbb{R}^2$.

b) Este subconjunto de $\mathbb{R}^2$ no es subespacio porque no contiene al vector cero.

Un subconjunto W de $\mathbb{R}^2$ es subespacio sí y sólo sí, está formado por alguno de estos conjuntos de puntos:

a) El punto $(0, 0)$ únicamente.
b) Una recta que pasa por el origen
c) Todo $\mathbb{R}^2$.

Otro ejemplo

Probar que el subconjunto de $\mathbb{R}^2$ formado por todos los puntos de la circunferencia unidad $x^2 + y^2 = 1$ no es subespacio de $\mathbb{R}^2$.

Solución

Este subconjunto no es subespacio porque los puntos $(1, 0)$ y $(0, 1)$ pertenecen a él, mientras que su suma (1, 1) no pertenece, por tanto, la suma no es cerrada. También podría llegarse a la misma conclusión, al denotar que el vector cero (0, 0) no pertenece a él.

3.3 Sistemas generadores

3.3.1 Definición

Se dice que los vectores $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ de un espacio vectorial V **generan** V si todo vector de V se puede escribir como combinación lineal de los mismos. Es decir,

$$\forall \boldsymbol{v} \in V, \exists\, \lambda_1, \lambda_2, \dots, \lambda_n; \boldsymbol{v} = \lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2, + \cdots + \lambda_n \boldsymbol{v}_n$$

Ejemplos

a) El conjunto $S = \{(1,0,0), (0,1,0), (0,0,1)\}$ es un sistema de generadores de $\mathbb{R}^3$ porque todo vector $\boldsymbol{u} = (u_1, u_2, u_3)$ de $\mathbb{R}^3$ puede expresarse como combinación lineal de los vectores de S, de la forma:

$$\boldsymbol{u} = u_1(1,0,0) + u_2(0,1,0) + u_3(0,0,1) = (u_1, u_2, u_3)$$

b) El conjunto $S = \{1, x, x^2\}$ genera P_2[3] porque cualquier polinomio $p(x) = a + bx + cx^2$ de P_2 se puede expresar como:

$$p(x) = a(1) + b(x) + c(x^2)$$

Los sistemas de generadores de los ejemplos anteriores se denominan sistemas canónicos de $\mathbb{R}^3$ y P_2 respectivamente.

Ejemplo

Probar que el conjunto $S = \{(1, 2, 3), (0\ 1, 2), (-2, 0, 1)\}$ genera $\mathbb{R}^3$.

[3] P_2 es el conjunto de todos los polinomios de grado menor o igual que 2

Solución

Sea $\boldsymbol{u} = (u_1, u_2, u_3)$ cualquier vector de $\mathbb{R}^3$, necesitamos encontrar escalares $\lambda_1, \lambda_2, \lambda_3$ tales que:

$$(u_1, u_2, u_3) = \lambda_1\,(1, 2, 3) + \lambda_2\,(0, 1, 2) + \lambda_3\,(-2, 0, 1) = (\lambda_1 - 2\lambda_3, 2\lambda_1 + \lambda_2, 3\lambda_1 + 2\lambda_2 + \lambda_3)$$

Igualando componente a componente, obtenemos un sistema lineal de ecuaciones:

$$\begin{cases} \lambda_1 - 2\lambda_3 = u_1 \\ 2\lambda_1 + \lambda_2 = u_2 \\ 3\lambda_1 + 2\lambda_2 + \lambda_3 = u_3 \end{cases}$$

Dado que la matriz de coeficientes tiene determinante no nulo:

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

El sistema tiene solución única, por lo tanto, todo vector de $\mathbb{R}^3$ se puede escribir como combinación lineal de los vectores de S, es decir, S genera $\mathbb{R}^3$.

3.3.2 Espacio generado por un conjunto de vectores.

Sea $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k, k$ vectores de un espacio vectorial V. El espacio generado por $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}$ es el conjunto de combinaciones lineales de $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k$**, es decir:**

$$gen\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\} = \{\boldsymbol{v}: \boldsymbol{v} = \lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2, + \cdots + \lambda_k\boldsymbol{v}_k\}$$

donde $\lambda_1, \lambda_2, \dots, \lambda_k$ son escalares arbitrarios.

A este conjunto también se le denomina como *envolvente lineal* y se escribe $lin\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}$

Demostración

Para probar que el conjunto de todas las combinaciones lineales de $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k$ es un subespacio de V bastará ver que las dos operaciones son cerradas.

Dados dos vectores cualesquiera $\boldsymbol{u}, \boldsymbol{w} \in lin(S)$ tales que:

$$\boldsymbol{u} = \lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_k\boldsymbol{v}_k = \sum_{i=1}^{\boldsymbol{k}} \lambda_i\boldsymbol{v}_i$$

$$\boldsymbol{w} = \mu_1\boldsymbol{v}_1 + \mu_2\boldsymbol{v}_2 + \cdots + \mu_k\boldsymbol{v}_k = \sum_{j=1}^{k} \mu_j\boldsymbol{v}_j$$

con λ_i, μ_j escalares. Entonces:

$$\boldsymbol{u} + \boldsymbol{w} = (\lambda_1 + \mu_1)\boldsymbol{v}_1 + (\lambda_2 + \mu_2)\boldsymbol{v}_2 + \cdots + (\lambda_k + \mu_k)\boldsymbol{v}_k$$

y

$$\alpha\boldsymbol{u} = (\alpha\lambda_1)\boldsymbol{v}_1 + (\alpha\lambda_2)\boldsymbol{v}_2 + \cdots + (\alpha\lambda_k)\boldsymbol{v}_k$$

lo que implica que $\boldsymbol{u} + \boldsymbol{w}$ y $\alpha\boldsymbol{u}$ también pertenecen a $lin(S)$, ya que se pueden escribir como combinaciones lineales de los vectores de S. Por tanto $lin(S)$ es subespacio vectorial de V.

Teorema.

Si $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ son vectores en un espacio vectorial V, entonces $S = gen\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ es un subespacio de V

Ejemplo: Ecuación del espacio vectorial generado por dos vectores en $\mathbb{R}^3$

Sean $\boldsymbol{v_1} = (2, -1, 4)$ y $\boldsymbol{v}_2 = (4, 1, 6)$.

Entonces $H = gen\{\boldsymbol{v_1}, \boldsymbol{v}_2\} = \{\boldsymbol{v}: \boldsymbol{v} = \lambda_1(2, -1, 4) + \lambda_2(4, 1, 6)\}$

¿Qué forma tiene H?. Si $\boldsymbol{v} = (x, y, z) \in H$, entonces se tiene:

$$x = 2\lambda_1 + 4\lambda_2$$

$$y = -\lambda_1 + \lambda_2$$

$$z = 4\lambda_1 + 6\lambda_2$$

Si se piensa que (x, y, z) está fijo, entonces estas ecuaciones se pueden ver como un sistema de tres ecuaciones con dos incógnitas (λ_1, λ_2). Este sistema se resuelve en la forma usual:

$$\left(\begin{array}{cc|c} -1 & 1 & y \\ 2 & 4 & x \\ 4 & 6 & z \end{array}\right) \xrightarrow{R_1 \to -R_1} \left(\begin{array}{cc|c} 1 & -1 & -y \\ 2 & 4 & x \\ 4 & 6 & z \end{array}\right) \xrightarrow[R_3 \to R_3 - 4R_1]{R_2 \to R_2 - 2R_1} \left(\begin{array}{cc|c} 1 & -1 & -y \\ 0 & 6 & x+2y \\ 0 & 10 & z+4y \end{array}\right)$$

$$\xrightarrow{R_2 \to \frac{1}{6}R_2} \left(\begin{array}{cc|c} 1 & -1 & -y \\ 0 & 1 & \frac{x+2y}{6} \\ 0 & 10 & z+4y \end{array}\right) \xrightarrow[R_3 \to R_3 - 10R_2]{R_1 \to R_1 + R_2} \left(\begin{array}{cc|c} 1 & 0 & \frac{x}{6} - \frac{2y}{3} \\ 0 & 1 & \frac{x}{6} + \frac{y}{3} \\ 0 & 0 & \frac{-5x}{3} + \frac{2y}{3+z} \end{array}\right)$$

El sistema tiene una solución únicamente si $\frac{-5x}{3} + \frac{2y}{3} + z = 0 \Longrightarrow 5x - 2y - 3z = 0$, que es la ecuación de un plano en $\mathbb{R}^3$ que pasa por el origen.

Este ejemplo se puede generalizar para probar que el espacio generado por dos vectores diferentes de cero en $\mathbb{R}^3$ que no son paralelos es un plano que pasa por el origen.

3.3.3 Dependencia e independencia lineal

Dado cualquier conjunto de vectores $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ de un espacio vectorial V, la ecuación vectorial:

$$\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_k\boldsymbol{v}_k = \boldsymbol{o}$$

tiene siempre la solución trivial $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$.

Ahora bien, muchas veces existen otras soluciones no triviales. Esta característica se describe diciendo que el conjunto S es **linealmente dependiente**. Si hubiera existido solamente la solución trivial diríamos que el conjunto S es **linealmente independiente**.

Definición. Se dice que un conjunto de vectores $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}$ de un espacio vectorial V es **linealmente independiente** si la ecuación vectorial:

$$\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_k \boldsymbol{v}_k = \boldsymbol{o}$$

tiene sólo la solución trivial $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. Si hay otras soluciones no triviales, se dice que S es **linealmente dependiente**.

Análisis de la independencia lineal

Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}$ un conjunto de vectores de un espacio vectorial V. Para saber si S es linealmente independiente hay que seguir los siguientes pasos:

1. Reescribir la ecuación vectorial $\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_k \boldsymbol{v}_k = \boldsymbol{o}$, como un sistema lineal homogéneo en las variables $\lambda_1, \lambda_2, \dots, \lambda_k$.
2. Averiguar si el sistema tiene solución única mediante eliminación gaussiana.
3. Si tiene solución única (la trivial con $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$), S es linealmente independiente. Si el sistema admite soluciones no triviales, S es linealmente dependiente.

Ejemplo

Determinar si es linealmente independiente en P_2 el conjunto formado por los vectores:

$$S = \{1 + x - 2x^2, 2 + 5x - x^2, x + x^2\}$$

Solución

Desarrollando la ecuación vectorial $\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \lambda_3 \boldsymbol{v}_3 = \boldsymbol{o}$, se tiene:

$$\lambda_1(1 + x - 2x^2) + \lambda_2(2 + 5x - x^2) + \lambda_3(x + x^2) = 0 + 0x + 0x^2$$

$$(\lambda_1 + 2\lambda_2) + (\lambda_1 + 5\lambda_2 + \lambda_3)x + (-2\lambda_1 - \lambda_2 + \lambda_3)x^2 = 0 + 0x + 0x^2$$

Igualando los coeficientes de los polinomios de ambos lados de la ecuación, se llega al sistema lineal:

$$\begin{cases} \lambda_1 + 2\lambda_2 = 0 \\ \lambda_1 + 5\lambda_2 + \lambda_3 = 0 \\ -2\lambda_1 - \lambda_2 + \lambda_3 = 0 \end{cases}$$

Su matriz ampliada se reduce por eliminación gaussiana:

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 1 & 5 & 1 & 0 \\ -2 & -1 & 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 0 & 0 \\ 1 & 1 & {}^1/_3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Esto implica que el sistema tiene infinitas soluciones. Por tanto admite soluciones no triviales, luego el conjunto S es linealmente dependiente.

Una solución no trivial es:

$$\lambda_1 = 2, \;\; \lambda_2 = -1, \;\; \lambda_3 = 3$$

que permite escribir la combinación lineal no trivial:

$$2(1 + x - 2x^2) - (2 + 5x - x^2) + 3(x + x^2) = \boldsymbol{o}$$

Ejemplo

Determinar si es linealmente independiente en $\mathcal{M}_{2,2}$ el conjunto de vectores:

$$S = \left\{ \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \right\}$$

Solución

De la ecuación vectorial:

$$\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \lambda_3 \boldsymbol{v}_3 = \boldsymbol{o}$$

resulta:

$$\lambda_1 \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

que produce el sistema de ecuaciones lineales en $\lambda_1, \lambda_2, \lambda_3$:

$$\begin{cases} 2\lambda_1 + 3\lambda_2 + \lambda_3 = 0 \\ \lambda_1 = 0 \\ 2\lambda_2 + \lambda_3 = 0 \\ \lambda_1 + \lambda_2 = 0 \end{cases}$$

Por eliminación gaussiana, su matriz ampliada se reduce así:

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

Así pues, el sistema tiene sólo la solución trivial de modo que el conjunto S es linealmente independiente.

Teorema. Una propiedad de los conjuntos linealmente dependientes.

Un conjunto $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}, k \geq 2$, es linealmente dependiente sí, y sólo sí, al menos uno de sus vectores $\boldsymbol{v}_j$ se puede escribir como combinación lineal de los demás vectores de S.

Demostración

a) Supongamos que S es linealmente dependiente. En tal caso, existen escalares $\lambda_1, \lambda_2, \dots, \lambda_k$, no todos nulos, tales que:

$$\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_k \boldsymbol{v}_k = \boldsymbol{o}$$

Puesto que alguno de los coeficientes ha de ser no nulo, podemos suponer que $\lambda_1 \neq 0$. Entonces, despejando $\boldsymbol{v}_1$ como combinación lineal de los demás vectores, resulta:

$$\lambda_1 \boldsymbol{v}_1 = -\lambda_2 \boldsymbol{v}_2 - \lambda_3 \boldsymbol{v}_3 \dots - \lambda_k \boldsymbol{v}_k$$

$$\boldsymbol{v}_1 = -\frac{\lambda_2}{\lambda_1}\boldsymbol{v}_2 - \frac{\lambda_3}{\lambda_1}\boldsymbol{v}_3 \dots - \frac{\lambda_k}{\lambda_1}\boldsymbol{v}_k$$

b) Recíprocamente, supongamos que el vector $\boldsymbol{v}_1$ de S es combinación lineal de los demás, o sea:

$$\boldsymbol{v}_1 = \lambda_2\boldsymbol{v}_2 + \lambda_3\boldsymbol{v}_3 + \cdots + \lambda_k\boldsymbol{v}_k$$

Entonces la ecuación:

$$-\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \lambda_3\boldsymbol{v}_3 + \cdots + \lambda_k\boldsymbol{v}_k = \boldsymbol{o}$$

tiene al menos un coeficiente, -1, no nulo. Concluimos que S es linealmente dependiente.

Corolario

Los vectores $\boldsymbol{u}$ y $\boldsymbol{v}$ del espacio vectorial V son linealmente dependientes sí, y sólo sí, uno de ellos es múltiplo escalar del otro.

El vector cero es múltiplo escalar de cualquier otro vector del espacio vectorial.

Teorema. Un conjunto de n vectores en $\mathbb{R}^m$ es siempre linealmente dependiente si $n > m$.

Demostración

Sean $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n, n$ vectores en $\mathbb{R}^m$ e intentemos encontrar constantes $\alpha_1, \alpha_2, \dots, \alpha_n$, no todas cero tales que:

$$\alpha_1\boldsymbol{v}_1 + \ \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_n\boldsymbol{v}_n = \boldsymbol{o}$$

Sean $\boldsymbol{v}_1 = (a_{11}, a_{21}, \dots, a_{m1}), \boldsymbol{v}_2 = (a_{12}, a_{22}, \dots, a_{m2}), \dots, \boldsymbol{v}_n = (a_{1n}, a_{2n}, \dots, a_{mn})$. Entonces la ecuación anterior se convierte en:

$$\begin{cases} a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n = 0 \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2n}\alpha_n = 0 \\ \vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \cdots + a_{mn}\alpha_n = 0 \end{cases} \qquad (1)$$

Pero este sistema tiene un número finito de soluciones si $n \geq m$. De esta forma, existen escalares $\alpha_1, \alpha_2, \dots, \alpha_n$ no todos cero que satisfacen el sistema y, por lo tanto, los vectores $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ son linealmente dependientes.

Corolario

Teorema. Sean $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}, n+1$ vectores que están en un espacio vectorial V. Si $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ genera a V, entonces $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}$ también genera a V. Es decir, si se agregan uno o más vectores a un conjunto generador, se obtiene otro conjunto generador.

Un conjunto de vectores linealmente independiente de $\mathbb{R}^n$ contiene a lo sumo n vectores.

Teorema. Sea

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

Entonces las columnas de A consideradas como vectores son linealmente dependientes si y sólo si el sistema, que se puede escribir como $A\boldsymbol{\lambda} = \mathbf{0}$, tiene soluciones no triviales. Donde $\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$

Ejemplo:

Considérese el sistema homogéneo:

$$\begin{cases} x_1 + 2x_2 - x_3 + 2x_4 = 0 \\ 3x_1 + 7x_2 + x_3 + 4x_4 = 0 \end{cases}$$

Haciendo una reducción de filas:

$$\begin{bmatrix} 1 & 2 & -1 & 2 & 0 \\ 3 & 7 & 1 & 4 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 2 & 0 \\ 0 & 1 & 4 & -2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -9 & 6 & 0 \\ 0 & 1 & 4 & -2 & 0 \end{bmatrix}$$

El último sistema es:

$$\begin{cases} x_1 - 9x_3 + 6x_4 = 0 \\ x_2 + 4x_3 - 2x_4 = 0 \end{cases}$$

Se ve que este sistema tiene un número infinito de soluciones, que se escriben como combinación lineal de los vectores columna, para ello basta tomar como variables libres: $x_3 = \lambda$ y $x_4 = \mu$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{pmatrix} 9\lambda - 6\mu \\ -4\lambda + 2\mu \\ \lambda \\ \mu \end{pmatrix} = \lambda \begin{pmatrix} 9 \\ -4 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

Obsérvese que $\begin{pmatrix} 9 \\ -4 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix}$ son soluciones linealmente independientes para el sistema dado porque ninguno de los dos es múltiplo del otro. Como λ y μ son números reales arbitrarios, se ve que el conjunto de soluciones al sistema es un subespacio de $\mathbb{R}^4$ generado por estos dos vectores solución linealmente independientes.

Teorema. $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n, n$ vectores de $\mathbb{R}^n$ y sea A una matriz de n x n cuyas columnas son $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$. Entonces $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ son linealmente independientes si y sólo si la única solución al sistema homogéneo $A\boldsymbol{x} = \boldsymbol{o}$ es la solución trivial $\boldsymbol{x} = \boldsymbol{o}$

Teorema. Sea A una matriz de $n \times n$. Entonces $\det(A) \neq 0$ si y sólo si, las columnas de A son linealmente independientes.

Teorema. Cualquier conjunto de n vectores de $\mathbb{R}^n$ linealmente independiente, genera $\mathbb{R}^n$

Demostración

Sean $\boldsymbol{v}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ 2 \end{bmatrix}, \dots, \boldsymbol{v}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}$ vectores linealmente independientes y sea $\boldsymbol{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ un vector de $\mathbb{R}^n$. Debemos demostrar qe existen escalares $\lambda_1, \lambda_2, \dots, \lambda_n$ tales que:

$$\boldsymbol{v} = \lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \dots + \lambda_n \boldsymbol{v}_n$$

Es decir:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + \lambda_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ 2 \end{bmatrix} + \dots + \lambda_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}$$

Desarrollando la ecuación vectorial anterior obtenemos el sistema de n ecuaciones con n incógnitas $(\lambda_1, \lambda_2, \dots, \lambda_n)$:

$$\begin{cases} a_{11}\lambda_1 + a_{12}\lambda_2 + \dots + a_{1n}\lambda_n = x_1 \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \dots + a_{2n}\lambda_n = x_2 \\ \vdots \\ a_{n1}\lambda_1 + a_{n2}\lambda_2 + \dots + a_{nn}\lambda_n = x_n \end{cases}$$

Que se puede escribir como $A\boldsymbol{\lambda} = \boldsymbol{v}$, donde:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{y} \quad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$$

Pero $\det(A) \neq 0$ ya que las columnas de A son linealmente independientes. De manera que el sistema tiene una solución única $\boldsymbol{\lambda}$ como queríamos demostrar.

Observación. Esta demostración no sólo muestra que $\boldsymbol{v}$ se puede escribir como una combinación lineal de los vectores independientes $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$, sino también que esto se puede lograr *de una sola manera* (ya que el vector solución $\boldsymbol{\lambda}$ es único).

Ejemplo

Los vectores $(2, -1, 4), (1, 0, 2)$ y $(3, -1, 5)$ generan $\mathbb{R}^3$ porque $\begin{vmatrix} 2 & 1 & 3 \\ -1 & 0 & -1 \\ 4 & 2 & 5 \end{vmatrix} = -1 \neq 0$ y, por lo tanto, son independientes.

Ejemplo

En $\mathcal{M}_{2,3}$, sean $A_1 = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix}$, $A_2 = \begin{bmatrix} -1 & 1 & 4 \\ 2 & 3 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix}$. Determinar si A_1, A_2, A_3 son linealmente independientes o dependientes.

Solución

Sopongamos que $\alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3 = 0$, entonces:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 & 1 & 4 \\ 2 & 3 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} -1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \alpha_1 - \alpha_2 - \alpha_3 & \alpha_2 & 2\alpha_1 + 4\alpha_2 + \alpha_3 \\ 3\alpha_1 + 2\alpha_2 + \alpha_3 & \alpha_1 + 3\alpha_2 + 2\alpha_3 & -\alpha_1 + \alpha_3 \end{bmatrix}$$

Esto nos proporciona un sistema homogéneo de seis ecuaciones con tres incógnitas $\alpha_1, \alpha_2, \alpha_3$ en el cual resulta sencillo verificar que la única solución es $\alpha_1 = \alpha_2 = \alpha_3 = 0$, lo que demuestra que las tres matrices son linealmente independientes.

Ejemplo

En P_3 determinar si los polinomios $1, x, x^2$ y x^3 son linealmente dependientes o independientes.

Solución

Supongamos que $c_1 + c_2 x + c_3 x^2 + c_4 x^3 = 0$. Esto debe cumplirse para todo número real x. En particular, si $x = 0$, se obtiene $c_1 = 0$. Entonces, dando sucesivamente a x los valores $x = 1, -1, 2$ se obtiene:

$$c_2 + c_3 + c_4 = 0$$

$$-c_2 + c_3 - c_4 = 0$$

$$2c_2 + 4c_3 + 8c_4 = 0$$

El determinante de este sistema homogéneo es:

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 2 & 4 & 8 \end{vmatrix} = 12 \neq 0$$

De manera que el sistema tiene una solución única $c_1 = c_2 = c_3 = 0$ y los cuatro polinomios son linealmente independientes.

3.4 Bases y dimensión de un espacio vectorial

3.4.1 Base de un espacio vectorial

Definición de base. Un conjunto de vectores $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ de un espacio vectorial V es una base de V si se cumple:

1. S genera V
2. S es linealmente independiente

Esta definición no asegura que todo espacio vectorial V tenga una base formada por un número finito de vectores. Si un espacio vectorial V admite una base finita, se dice que V es de **dimensión finita**. En el caso contrario, se dice que V es de **dimensión infinita**.

El espacio vectorial P de todos los polinomios y el espacio vectorial $C(-\infty,\infty)$ de las funciones reales continuas son espacios de dimensión infinita. El espacio $\mathbb{R}^3$ es un espacio vectorial de dimensión 3.

Ejemplo

Probar que el espacio vectorial P_3 admite como base el conjunto $S = \{1, x, x^2, x^3\}$

Solución

Es evidente que S genera P_3, porque la envolvente lineal de S contiene todos los polinimios de la forma:

$$a_0 + a_1x + a_2x^2 + a_3x^3 \text{ donde } a_0, a_1, a_2, a_3 \text{ son reales}$$

Para verificar la independencia lineal de S, hay que recordar que en P_3 el vector cero es el polinomio dado por $\mathbf{0}(x) = 0, \forall x$. Así pues el criterio de independencia lineal da la ecuación:

$$a_0 + a_1x + a_2x^2 + a_3x^3 = \mathbf{0}(x) = 0, \forall x$$

De este polinomio de tercer grado se dice que es idénticamente nulo, es decir, que todos sus coeficientes son cero $a_0 = a_1 = a_2 = a_3 = 0$

Por tanto S es linealmente independiente y, como consecuencia, es una base de P_3, que se denomina ***base canónica***.

Teorema. Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ es una base de un espacio vectorial V, cada vector de V se puede escribir como una, y sólo una, forma como combinación lineal de los vectores de S.

Demostración

Como S genera V, cualquier vector $\boldsymbol{u} \in V$ se puede expresar como:

$\boldsymbol{u} = \lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2, + \cdots + \lambda_n\boldsymbol{v}_n$.

Para probar la unicidad, supongamos que $\boldsymbol{u}$ tiene otra expresión de la forma :

$\boldsymbol{u} = \mu_1\boldsymbol{v}_1 + \mu_2\boldsymbol{v}_2, + \cdots + \mu_n\boldsymbol{v}_n$.

Restando ambas expresiones resulta:

$$\boldsymbol{u} - \boldsymbol{u} = (\lambda_1 - \mu_1)\boldsymbol{v}_1 + (\lambda_2 - \mu_2)\boldsymbol{v}_2, + \cdots + (\lambda_n - \mu_n)\boldsymbol{v}_n = \boldsymbol{o}$$

Ahora bien, si S es linealmente independiente, la única solución de esta ecuación es la solución trivial:

$$(\lambda_1 - \mu_1) = 0, \ (\lambda_2 - \mu_2) = 0, \ \ (\lambda_n - \mu_n) = 0$$

Lo cual significa que $\lambda_i = \mu_i, \ \forall i = 1, 2, \dots, n$. Por tanto $\boldsymbol{u}$ tiene un único desarrollo en términos de la base S dada.

Ejemplo

Sea $\boldsymbol{u} = (u_1, u_2, u_3)$ cualquier vector de $\mathbb{R}^3$. Demostrar que la ecuación:

$$\boldsymbol{u} = \alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3$$

tiene una única solución en términos de la base $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\} = \{(1, 2, 3), (0, 1, 2), (-2, 0, 1)\}$

Solución

$(u_1, u_2, u_3) = \alpha_1(1,2,3) + \alpha_2(0,1,2) + \alpha_3(-2,0,1) = (\alpha_1 + 2\alpha_3), (2\alpha_1 + \alpha_2), (3\alpha_1 + 2\alpha_2 + \alpha_3)$.

Da lugar, componente a componente al siguiente sistema de ecuaciones lineales:

$$\begin{cases} \alpha_1 + 2\alpha_3 = u_1 \\ 2\alpha_1 + \alpha_2 = u_2 \\ 3\alpha_1 + 2\alpha_2 + \alpha_3 = u_3 \end{cases} \qquad \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \qquad A\boldsymbol{\alpha} = \boldsymbol{u}$$

Como la matriz A es invertible, el sistema tiene una única solución, dada por $\boldsymbol{\alpha} = A^{-1}\boldsymbol{u}$. Calculando A^{-1} resulta:

$$A^{-1} = \begin{bmatrix} -1 & 4 & -2 \\ 2 & -7 & 4 \\ -1 & 2 & -1 \end{bmatrix}$$

Lo cual implica que:

$$\alpha_1 = -u_1 + 4u_2 - 2u_3$$

$$\alpha_2 = 2u_1 - 7u_2 + 4u_3$$

$$\alpha_3 = -u_1 + 2u_2 - u_3$$

Por ejemplo, el vector $\boldsymbol{u} = (1,0,0)$ se puede representar de forma única como combinación lineal de S, como sigue:

$$(1,0,0) = -\boldsymbol{v}_1 + 2\boldsymbol{v}_2 - \boldsymbol{v}_3$$

3.4.2 Coordenadas y componentes de un vector en una base.

Si el espacio vectorial V admite una base $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ se llaman **coordenadas** de un vector $\boldsymbol{x} \in V$ respecto de la base B, a los únicos escalares $\lambda_1, \lambda_2, \ldots., \lambda_n$ para los que:

$$\boldsymbol{x} = \sum_{i=1}^{n} \lambda_i \boldsymbol{v}_i$$

Los vectores $\lambda_i \boldsymbol{v}_i$ se llaman **componentes** del vector $\boldsymbol{x}$ respecto de la base B.

La matriz de coordenadas (o vector de coordenadas) de $\boldsymbol{x}$ en la base B es la matriz columna en $\mathbb{R}^n$ cuyas componentes son las coordenadas de $\boldsymbol{x}$**:**

$$[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$$

Teorema. Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ es una base de un espacio vectorial V, todo conjunto que contenga más de n vectores es linealmente dependiente.

Demostración

Sea $S_1 = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ un conjunto de m vectores de V, con $m > n$. Para probar que S_1 es linealmente dependiente, debemos encontrar escalares $k_1, k_2, \ldots, k_m$ (no todos cero), tales que:

$$k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2 + \cdots + k_m\boldsymbol{u}_m = \boldsymbol{o} \quad (1)$$

Puesto que $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ es una base de V, cada $\boldsymbol{u}_i$ es combinación lineal de los vectores de S, con lo que podemos escribir:

$$\boldsymbol{u}_1 = c_{11}\boldsymbol{v}_1 + c_{21}\boldsymbol{v}_2 + \cdots + c_{n1}\boldsymbol{v}_n$$

$$\boldsymbol{u}_2 = c_{12}\boldsymbol{v}_1 + c_{22}\boldsymbol{v}_2 + \cdots + c_{n2}\boldsymbol{v}_n$$

$$\vdots$$

$$\boldsymbol{u}_m = c_{1m}\boldsymbol{v}_1 + c_{2m}\boldsymbol{v}_2 + \cdots + c_{nm}\boldsymbol{v}_n$$

Sustituyendo estas expresiones en la ecuación (1) y reagrupando términos, obtenemos:

$$d_1\boldsymbol{v}_1 + d_2\boldsymbol{v}_2 + \cdots + d_n\boldsymbol{v}_n = \boldsymbol{o}$$

donde:

$$d_i = \sum_{j=1}^{m} c_{ij}k_j$$

Sin embargo, como, los $\boldsymbol{v}_i$ forman un conjunto linealmente independiente, concluimos que cada $d_i = 0$, con lo que llegamos al sistema de ecuaciones lineales:

$$c_{11}k_1 + c_{12}k_2 + \cdots + c_{1m}k_m = 0$$

$$c_{21}k_1 + c_{22}k_2 + \cdots + c_{2m}k_m = 0$$

$$\vdots$$

$$c_{n1}k_1 + c_{n2}k_2 + \cdots + c_{nm}k_m = 0$$

Pero este sistema lineal homogéneo tiene menos ecuaciones que variables $(k_1, k_2, \ldots, k_m)$, por lo que, como sabemos, debe tener soluciones no triviales. En consecuencia, S_1 es linealmente dependiente.

> Teorema. Si un espacio vectorial V tiene una base formada por n vectores, toda base de V contiene exactamente n vectores.

Demostración

Sea $S_1 = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ la base de V dada , y sea $S_2 = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ otra base de V. Como S_1 es una base y S_2 es linealmente independiente, por el teorema anterior, implica que $m \leq n$. Análogamente $n \leq m$ porque S_1 es linealmente independiente y S_2 es una base. Por tanto $n = m$.

3.4.3 Dimensión de un espacio vectorial

Definición. Si un espacio vectorial tiene una base formada por n vectores, se dice que V tiene dimensión n (o que es n-dimensional), y se escribe $\dim(V) = n$. Si V contiene sólo el vector cero, se define como de dimensión cero.

Teorema. Criterio para bases de un espacio vectorial de tipo finito. Sea V un espacio vectorial de dimensión n, un sistema de n vectores $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ es base sí, y sólo sí, se cumple una cualquiera de las condiciones:

1. B es un sistema linealmente independiente
2. B es un sistema generador (es decir, genera V).

Demostración

Si B es independiente es base, pues de lo contrario se podría encontrar una base con más de n vectores. Si B genera V no puede ser dependiente, pues si lo fuese habría una base con menos de n vectores y en consecuencia, B es generador y es independiente, es decir, es base. Los recíprocos son evidentes.

Propiedad. Si V es un espacio vectorial de tipo finito y $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_p\}$ es un sistema linealmente independiente de vectores de V, siempre es posible encontrar unos vectores $\boldsymbol{v}_{p+1}, \boldsymbol{v}_{p+2}, \dots, \boldsymbol{v}_n$ de V tales que, siendo $n \geq p$, el sistema $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_p, \boldsymbol{v}_{p+1}, \dots, \boldsymbol{v}_n\}$ sea una base de V.

En efecto, si $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_p\}$ generase V, sería una base y no se hace necesario añadir ningún vector. Si $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_p\}$ no genera V, ha de haber en V un vector, al menos, $\boldsymbol{v}_{p+1}$, independiente de ellos; si $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_p, \boldsymbol{v}_{p+1}\}$ genera V, la propiedad está demostrada, y si no lo es, se procede como en el caso anterior. Razonando de esta forma sucesivamente, se consigue una base, puesto que, de lo contrario, el proceso se prolongaría indefinidamente y V no podría ser de tipo finito.

Propiedad. Si $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ es una base del espacio vectorial V y $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_r\}$ es un sistema linealmente independiente, se puede asegurar que hay $n - r$ vectores de la base que añadidos a los $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_r$, constituyen una nueva base.

En efecto, según la propiedad anterior, se puede construir una base a partir de los vectores $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_r$, añadiendo, de modo adecuado, vectores independientes de ellos y éstos siempre es posible buscarlos de entre los de la base.

3.5 Rango de una matriz y sistemas de ecuaciones lineales

3.5.1 Espacio de filas y espacio de columnas de una matriz

Sea A una matriz $m \times n$, se define:

1. El espacio de filas de A es el subespacio de $\mathbb{R}^n$ generado por los vectores fila de A.
2. El espacio de columnas de A es el subespacio de $\mathbb{R}^m$ generado por los vectores columna de A.

Recordemos que dos matrices son equivalentes por filas si una puede obtenerse de la otra mediante operaciones elementales por filas.

Teorema. Sean A y B matrices $m \times n$. Si A es equivalente por filas a B, el espacio de filas de A es igual al espacio de filas de B.

Demostración

Como las filas de B se obtienen de las de A por operaciones elementales por filas (suma y multiplicación por un escalar), cada uno de los vectores fila de B se puede escribir como combinación lineal de los vectores fila de A. Por tanto los vectores fila de B y por tanto, el subespacio generado por los vectores fila de B, están contenidos en el espacio de filas de A. Pero también es cierto que las filas de A se pueden obtener de las de B mediante operaciones elementales, de modo que cada uno de los espacios de filas es subespacio del otro, lo cual implica que son iguales.

Nota. Según el anterior teorema, el espacio de filas de una matriz no cambian al efectuar operaciones elementales por filas, Sin embargo, las operaciones elementales por filas pueden modificar el espacio de columnas.

Si la matriz B es escalonada por filas, sus vectores fila no nulos forman un conjunto linealmente independiente, por tanto, forman una base del espacio de filas de B y, por el teorema anterior, también del espacio de filas de A, lo que resulta en el siguiente teorema.

Teorema. Si A es una matriz equivalente por filas a una matriz B escalonada por filas, los vectores fila no nulos de B forman una base del espacio de filas de A.

Ejemplo.

Hallar una base del espacio de filas de la matriz:

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

Solución

Mediante operaciones elementales por filas, reescribimos A como:

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

Por el teorema anterior podemos concluir que los vectores fila de B:

$$\boldsymbol{w}_1 = (1,3,1,3), \qquad \boldsymbol{w}_2 = (0,1,1,0), \qquad \boldsymbol{w}_3 = (0,0,0,1)$$

forman una base del espacio de filas de A.

———

Para hallar una base del **espacio de columnas** de una matriz A tenemos dos opciones:

1. Como el espacio de columnas de A es el espacio de filas de A^T, basta aplicar el método del ejemplo anterior a la matriz A^T.
2. Aunque las operaciones por filas cambian el espacio de columnas de la matriz, no cambian las relaciones de dependencia entre las columnas.

Consideremos las dos matrices del ejemplo anterior:

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

Las columnas 1, 2 y 3 de B satisfacen la ecuación $\boldsymbol{b}_3 = -2\boldsymbol{b}_1 + \boldsymbol{b}_2$ y también las correspondientes columnas de A la satisfacen.

Análogamente los vectores columna $\boldsymbol{b}_1, \boldsymbol{b}_2$ y $\boldsymbol{b}_4$ de B son linealmente independientes y también lo son las correspondientes columnas de A.

Ejemplo

Hallar una base del espacio de columnas de la matriz A del ejemplo anterior.

Solución 1

Tomamos A^T, y la reescribimos, mediante operaciones elementales por filas, en forma escalonada:

$$A^T = \begin{bmatrix} 1 & 0 & -3 & 3 & 2 \\ 3 & 1 & 0 & 4 & 0 \\ 1 & 1 & 6 & -2 & -4 \\ 3 & 0 & -1 & 1 & -2 \end{bmatrix} \longrightarrow \begin{array}{c} \begin{bmatrix} 1 & 0 & -3 & 3 & 2 \\ 0 & 1 & 9 & -5 & -6 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}\begin{array}{l} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \\ \boldsymbol{w}_3 \\ \\ \end{array}$$

Así pues, $\boldsymbol{w}_1, \boldsymbol{w}_2$ y $\boldsymbol{w}_3$ forman una base del espacio vectorial de filas de la matriz A^T. Eso equivale a decir que los vectores columna:

$$\begin{bmatrix} 1 \\ 0 \\ -3 \\ 3 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 9 \\ -5 \\ -6 \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

forman una base del espacio de columnas de A.

Solución 2

Es fácil ver que en B los vectores columna 1, 2 y 4 son linealmente independientes, por tanto las correspondientes columnas de A son linealmente independientes, luego una base del espacio de columnas viene dada por los vectores:

$$\begin{bmatrix} 1 \\ 0 \\ -3 \\ 3 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} 3 \\ 1 \\ 0 \\ 4 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 3 \\ 0 \\ -1 \\ 1 \\ -2 \end{bmatrix}$$

Nótese que esta base del espacio de columnas es distinta de la obtenida en la solución anterior. Compruébese que ambas generan el mismo subespacio de $\mathbb{R}^5$.

Nota. En la segunda solución la forma escalonada por filas de B indica que las columnas de A forman una base del espacio de columnas. No utilice los vectores columna de B para construir la base.

> Teorema. El espacio de filas y el espacio de columnas tienen la misma dimensión.
> Si A es una matriz de $m\ x\ n$ su espacio de filas y su espacio de columnas tienen la misma dimensión.

3.5.2 Rango de una matriz

La dimensión del espacio de filas (o de columnas) de una matriz A se denomina rango de A y se denota por $rango(A)$.

Ejemplo

Calcular el rango de la matriz:

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

Solución

La llevamos a forma escalonada por filas:

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

y vemos que B tiene 3 filas no nulas, por lo que el rango de A es 3.

3.5.3 Núcleo de una matriz

Considérese el sistema lineal homogéneo $A\boldsymbol{x} = \boldsymbol{o}$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Demostración

Teorema. Si A es una matriz $m\ x\ n$, el conjunto solución del sistema lineal homogéneo

$$A\boldsymbol{x} = \boldsymbol{o}$$

es un subespacio de $\mathbb{R}^n$ que se llama **núcleo de $\boldsymbol{A}$** y se denota por $N(A)$. Así pues,

$$N(A) = \{x \in \mathbb{R}^n \colon A\boldsymbol{x} = \boldsymbol{o}\}$$

La dimensión de $N(A)$ se llama **nulidad de $\boldsymbol{A}$**.

Como A es una matriz $m\ x\ n$, $\boldsymbol{x}$ tiene tamaño $n\ x\ 1$, de modo que el conjunto solución es un subconjunto de $\mathbb{R}^n$. Este subconjunto es claramente no vacío, ya que $A\boldsymbol{o} = \boldsymbol{o}$ (es decir el vector nulo de $\mathbb{R}^n$ pertenece a dicho subconjunto).

Para comprobar que es un subespacio, basta ver que la suma y la multiplicación por un escalar son cerradas en él. Sean $\boldsymbol{x_1}$ y $\boldsymbol{x_2}$ dos vectores solución del sistema $A\boldsymbol{x} = \boldsymbol{o}$, y sea c un escalar. De $A\boldsymbol{x_1} = \boldsymbol{o}$ y $A\boldsymbol{x}_2 = \boldsymbol{o}$, se sigue que:

$$A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = A\boldsymbol{x}_1 + A\boldsymbol{x}_2 = \boldsymbol{o} + \boldsymbol{o} = \boldsymbol{o}$$

$$A(\boldsymbol{c}\boldsymbol{x}_1) = c(A\boldsymbol{x}_1) = c\boldsymbol{o} = \boldsymbol{o}$$

Así pues, tanto $\boldsymbol{x}_1 + \boldsymbol{x}_2$ como $c\boldsymbol{x}_1$ son soluciones de $A\boldsymbol{x} = \boldsymbol{o}$, lo que permite concluir que el conjunto solución es un subespacio de $\mathbb{R}^n$.

Ejemplo

Hallar el núcleo de la matriz:

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

Solución

El núcleo de A es el conjunto solución del sistema homogéneo $A\boldsymbol{x} = \boldsymbol{o}$.

Para resolver este sistema, hay que escribir su matriz ampliada $[A \vdots O]$ en forma escalonada reducida por filas. Sin embargo, al tratarse de un sistema homogéneo, la columna de la derecha de la matriz ampliada es nula, y no cambia bajo operaciones elementales por filas. Por tanto, es suficiente hallar la forma escalonada reducida por filas de A.

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

El sistema de ecuaciones correspondiente es:

$$\begin{cases} x_1 + 2x_2 + 3x_4 = 0 \\ \qquad x_3 + x_4 = 0 \end{cases}$$

Elegimos $x_2 = s$ y $x_4 = t$ como variables libres para representar las soluciones en forma paramétrica:

$$x_1 = -2s - 3t$$

$$x_3 = -t$$

Eso significa que el espacio de soluciones de $A\boldsymbol{x} = \boldsymbol{o}$ consta de todos los vectores solución $\boldsymbol{x}$ de la forma:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2s - 3t \\ s \\ -t \\ t \end{bmatrix} = s\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

Así pues, el núcleo de A está generado por los vectores:

$$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

Es decir, estos dos vectores son solución del sistema $A\boldsymbol{x} = \boldsymbol{o}$ y cualquier otra solución es combinación lineal de ellos.

Se puede demostrar que cuando se resuelve un sistema homogéneo a partir de la forma escalonada reducida, el sistema generado es siempre linealmente independiente.

En este ejemplo la matriz A tiene 4 columnas, rango 2 (columnas 1 y 3 de la matriz escalonada) y nulidad 2 (columnas 2 y 4 de la matriz escalonada).

<u>Teorema. Dimensión del espacio de soluciones.</u>

Si A es una matriz $m\ x\ n$ de rango r, la dimensión del espacio de soluciones del sistema $A\boldsymbol{x} = \boldsymbol{o}$ es $n - r$, esto es:

$$n = rango(A) + nulidad(A)$$

Demostración

Puesto que A tiene rango r, sabemos que es equivalente por filas a una matriz escalonada reducida B con r filas no nulas. Sin pérdida de generalidad, suponemos que la esquina $r\ x\ r$ superior izquierda de B es la matriz identidad I_r. Como las filas nulas de B no contribuyen a la solución, las descartamos para quedarnos con una matriz B', de tamaño $r\ x\ n$, donde $B' = [I_r \vdots C]$. La matriz C tiene $n-r$ columnas correspondientes a las variables $x_{r+1}, x_{r+2}, \dots, x_n$. Así pues, el espacio de soluciones de $A\boldsymbol{x} = \boldsymbol{o}$ viene representado por el sistema:

$$\begin{cases} x_1 + \qquad + c_{11}x_{r+1} + c_{12}x_{r+2} + \cdots + c_{1n-r}x_n = 0 \\ \quad x_2 \qquad + c_{21}x_{r+1} + c_{22}x_{r+2} + \cdots + c_{2n-r}x_n = 0 \\ \qquad\qquad \vdots \\ \qquad x_r + c_{r1}x_{r+1} + c_{r2}x_{r+2} + \cdots + c_{rn-r}x_n = 0 \end{cases}$$

Resolviendo en las primeras r variables, expresadas en términos de las $n-r$ últimas, se obtienen $n-r$ vectores de la base del espacio de soluciones. En consecuencia, el espacio de soluciones tiene dimensión $n-r$.

Ejemplo

Dada la matriz:

$$A = \begin{bmatrix} 1 & 0 & -2 & 1 & 0 \\ 0 & -1 & -3 & 1 & 3 \\ -2 & -1 & 1 & -1 & 3 \\ 0 & 3 & 9 & 0 & -12 \end{bmatrix}$$

Denotemos los vectores columna por $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5$

a) Calcular el rango y la nulidad de A
b) Hallar un subconjunto de los vectores columna de A que formen base del espacio de columnas de A.
c) Escribir, si es posible, la tercera columna de A como combinación lineal de las dos primeras.

Solución

Sea B la forma escalonada reducida por filas de A:

$$A = \begin{bmatrix} 1 & 0 & -2 & 1 & 0 \\ 0 & -1 & -3 & 1 & 3 \\ -2 & -1 & 1 & -1 & 3 \\ 0 & 3 & 9 & 0 & -12 \end{bmatrix} \longrightarrow B = \begin{bmatrix} 1 & 0 & -2 & 1 & 0 \\ 0 & 1 & 3 & 0 & -4 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. B tiene 3 filas no nulas, luego el rango de A es 3. Además A tiene $n = 5$ columnas, de modo que la nulidad de A es $n - rango(A) = 5 - 3 = 2$
2. Como las columnas 1, 2 y 4 de B son linealmente independientes, las correspondientes columnas de A:

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \end{bmatrix} \quad \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ -1 \\ -1 \\ 3 \end{bmatrix} \quad \boldsymbol{a}_4 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

forman una base del espacio de columnas de A.

3. La tercera columna de B es combinación lineal de las dos primeras: $\boldsymbol{b}_3 = -2\boldsymbol{b}_1 + 3\boldsymbol{b}_2$, por tanto la misma relación es válida para las columnas correspondientes de A:

$$\boldsymbol{a}_3 = \begin{bmatrix} -2 \\ -3 \\ 1 \\ 9 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ -1 \\ -1 \\ 3 \end{bmatrix}$$

3.5.4 Soluciones de sistemas de ecuaciones lineales

¿Es también subespacio vectorial el conjunto solución del sistema no homogéneo $A\boldsymbol{x} = \boldsymbol{b}$? No, ya que el vector cero nunca es solución del sistema.

No obstante existe una relación entre los conjuntos de soluciones de los sistemas $A\boldsymbol{x} = \boldsymbol{o}$ y $A\boldsymbol{x} = \boldsymbol{b}$. Si $\boldsymbol{x}_p$ es una solución particular del sistema no homogéneo, toda solución del sistema es de la forma:

$$\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_h$$

donde $\boldsymbol{x}_h$ es una solución del sistema homogéneo asociado.

Teorema.

Si $\boldsymbol{x}_p$ es una solución particular del sistema $A\boldsymbol{x} = \boldsymbol{b}$, toda solución de este sistema es de la forma

$$\boldsymbol{x}_h$$

donde $\boldsymbol{x}_h$ es una solución del sistema homogéneo asociado.

Demostración

Sea $\boldsymbol{x}$ una solución cualquiera de $A\boldsymbol{x} = \boldsymbol{b}$. Entonces $(\boldsymbol{x} - \boldsymbol{x}_p)$ es solución de $A\boldsymbol{x} = \boldsymbol{o}$ porque:

$$A(\boldsymbol{x} - \boldsymbol{x}_p) = A\boldsymbol{x} - A\boldsymbol{x}_p = \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{o}$$

Llamando $\boldsymbol{x}_h = \boldsymbol{x} - \boldsymbol{x}_p$ obtenemos $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_h$

Ejemplo

Hallar el conjunto de vectores solución del sistema de ecuaciones lineales:

$$\begin{cases} x_1 - 2x_3 + x_4 = 5 \\ 3x_1 + x_2 - 5x_3 = 8 \\ x_1 + 2x_2 - 5x_4 = -9 \end{cases}$$

Solución

La matriz ampliada del sistema $A\boldsymbol{x} = \boldsymbol{b}$ se reduce como sigue:

$$\begin{bmatrix} 1 & 0 & -2 & 1 & 5 \\ 3 & 1 & -5 & 0 & 8 \\ 1 & 2 & 0 & -5 & -9 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -2 & 1 & 5 \\ 0 & 1 & 1 & -3 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

El sistema de ecuaciones lineales correspondiente a esta matriz escalonada reducida es:

$$\begin{cases} x_1 - 2x_3 + x_4 = 5 \\ x_2 + x_3 - 3x_4 = -7 \end{cases}$$

Haciendo $x_3 = s,\ x_4 = t$ podemos representar el vector solución de $A\boldsymbol{x} = \boldsymbol{b}$ como:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2s - t + 5 \\ -s + 3t - 7 \\ s \\ t \end{bmatrix} = s \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 3 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -7 \\ 0 \\ 0 \end{bmatrix}$$

$$\boldsymbol{x} = s\boldsymbol{u}_1 + t\boldsymbol{u}_2 + \boldsymbol{x}_p$$

Es fácil ver que $\boldsymbol{x}_p$ es una solución particular de $A\boldsymbol{x} = \boldsymbol{b}$ y $\boldsymbol{x}_h = s\boldsymbol{u}_1 + t\boldsymbol{u}_2$ representa un vector arbitrario del espacio de soluciones de $A\boldsymbol{x} = \boldsymbol{o}$.

Teorema. Compatibilidad de un sistema lineal

El sistema de ecuaciones lineales $A\boldsymbol{x} = \boldsymbol{b}$ es compatible sí, y sólo sí, $\boldsymbol{b}$ está en el espacio de columnas de A.

Demostración

Sean:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

la matriz de coeficientes, el vector columna de incógnitas y el vector columna de los términos independientes, respectivamente del sistema $A\boldsymbol{x} = \boldsymbol{b}$. Entonces:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} =$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Por tanto, $A\boldsymbol{x} = \boldsymbol{b}$ sí, y sólo sí, $\boldsymbol{b}$ es combinación lineal de las columnas de A. Es decir, el sistema es compatible sí, y sólo sí, $\boldsymbol{b}$ pertenece al subespacio de $\mathbb{R}^m$ generado por las columnas de A.

Ejemplo

Considérese el sistema lineal:

$$\begin{cases} x_1 + x_2 - x_3 = -1 \\ x_1 + x_3 = 3 \\ 3x_1 + 2x_2 - x_3 = 1 \end{cases}$$

La matriz de coeficientes se reduce:

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

Se observa que el rango de la matriz de coeficientes es 2.

Procedemos de igual forma con la matriz ampliada:

$$[A \vdots \boldsymbol{b}] = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 0 & 1 & 3 \\ 3 & 2 & -1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

El rango de la matriz ampliada es 2, por tanto $\boldsymbol{b}$ está en el espacio de columnas de A y el sistema es compatible.

3.5.5 Sistemas lineales con matriz de coeficientes cuadrada

Resumen de condiciones equivalentes para matrices cuadradas:

Si A es una matriz cuadrada de orden n, las siguientes afirmaciones son equivalentes:

1. A es invertible
2. $A\boldsymbol{x} = \boldsymbol{b}$ tiene solución única para cualquier matriz $\boldsymbol{b}$ de tamaño $n\ x\ 1$.
3. $A\boldsymbol{x} = \boldsymbol{o}$ tiene sólo la solución trivial
4. A es equivalente por filas a la matriz identidad de orden n, I_n.
5. $|A| \neq 0$
6. $rango(A) = n$
7. Los vectores fila de A son linealmente independientes.
8. Los n vectores columna de A son linealmente independientes.

3.6 Cambios de base

Sean $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ y $B' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n\}$ dos bases de un espacio vectorial V. Si

$$\begin{cases} \boldsymbol{v}_1 = c_{11}\boldsymbol{u}_1 + c_{21}\boldsymbol{u}_2 + \cdots + c_{n1}\boldsymbol{u}_n \\ \boldsymbol{v}_2 = c_{12}\boldsymbol{u}_1 + c_{22}\boldsymbol{u}_2 + \cdots + c_{n2}\boldsymbol{u}_n \\ \vdots \\ \boldsymbol{v}_n = c_{1n}\boldsymbol{u}_1 + c_{2n}\boldsymbol{u}_2 + \cdots + c_{nn}\boldsymbol{u}_n \end{cases}$$

entonces la matriz de cambio de B a B' es:

$$Q = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

Demostración

Sea $\boldsymbol{v} = d_1\boldsymbol{v}_1 + d_2\boldsymbol{v}_2 + \cdots + d_n\boldsymbol{v}_n$ un vector arbitrario de V. Su matriz de coordenadas en la base B es, por tanto:

$$\boldsymbol{v} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

Así pues:

$$Q[\boldsymbol{v}]_B = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = \begin{bmatrix} c_{11}d_1 + c_{12}d_2 + \cdots + c_{1n}d_n \\ c_{21}d_1 + c_{22}d_2 + \cdots + c_{2n}d_n \\ \vdots \\ c_{n1}d_1 + c_{n2}d_2 + \cdots + c_{nn}d_n \end{bmatrix}$$

Por otra parte, podemos escribir:

$$\boldsymbol{v} = d_1\boldsymbol{v}_1 + d_2\boldsymbol{v}_2 + \cdots + d_n\boldsymbol{v}_n = d_1(c_{11}\boldsymbol{u}_1 + c_{21}\boldsymbol{u}_2 + \cdots + c_{n1}\boldsymbol{u}_n) + \cdots + d_n(c_{1n}\boldsymbol{u}_1 + c_{2n}\boldsymbol{u}_2 + \cdots + c_{nn}\boldsymbol{u}_n)$$
$$= (d_1c_{11} + \cdots + d_nc_{1n})\boldsymbol{u}_1 + \cdots + (d_1c_{n1} + \cdots + d_nc_{nn})\boldsymbol{u}_n$$

lo cual implica que:

$$[\boldsymbol{v}]_{B'} = \begin{bmatrix} c_{11}d_1 + c_{12}d_2 + \cdots + c_{1n}d_n \\ c_{21}d_1 + c_{22}d_2 + \cdots + c_{2n}d_n \\ \vdots \\ c_{n1}d_1 + c_{n2}d_2 + \cdots + c_{nn}d_n \end{bmatrix}$$

Por tanto $Q[\boldsymbol{v}]_B = [\boldsymbol{v}]_{B'}$ y concluimos que Q es la matriz de cambio de B a B'.

Teorema. Inversa de la matriz de cambio de base.

Si P es la matriz de cambio de una base B' a otra base B de $\mathbb{R}^n$, P es invertible, y la matriz del cambio de B a B' viene dada por P^{-1}.

Demostración

Por lo demostrado anteriormente, sea Q la matriz de cambio de base de B a B'. Entonces:

$[\boldsymbol{v}]_B = P[\boldsymbol{v}]_{B'}$ y $[\boldsymbol{v}]_{B'} = Q[\boldsymbol{v}]_B$

Lo cual implica que $[\boldsymbol{v}]_B = PQ[\boldsymbol{v}]_B$ para todo vector $\boldsymbol{v}$ de $\mathbb{R}^n$. De ahí se sigue que $PQ = I_n$. En consecuencia P es invertible y $P^{-1} = Q$, la matriz de cambio de B a B'.

Teorema. Sean $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ y $B' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n\}$ dos bases de $\mathbb{R}^n$. La matriz de cambio de B a B' se puede hallar aplicando el método de eliminación de Gauss-Jordan a la matriz $[B' \vdots B]$ como sigue:

$$[B' \vdots B] \longrightarrow [I_n \vdots P^{-1}]$$

Demostración

Supongamos que:

$$\begin{cases} \boldsymbol{v}_1 = c_{11}\boldsymbol{u}_1 + c_{21}\boldsymbol{u}_2 + \cdots + c_{n1}\boldsymbol{u}_n \\ \boldsymbol{v}_2 = c_{12}\boldsymbol{u}_1 + c_{22}\boldsymbol{u}_2 + \cdots + c_{n2}\boldsymbol{u}_n \\ \vdots \\ \boldsymbol{v}_n = c_{1n}\boldsymbol{u}_1 + c_{2n}\boldsymbol{u}_2 + \cdots + c_{nn}\boldsymbol{u}_n \end{cases}$$

de modo que:

$$c_{1i}\begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{1n} \end{bmatrix} + c_{2i}\begin{bmatrix} u_{21} \\ u_{22} \\ \vdots \\ u_{2n} \end{bmatrix} + \cdots + c_{ni}\begin{bmatrix} u_{n1} \\ u_{n2} \\ \vdots \\ u_{nn} \end{bmatrix} = \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{in} \end{bmatrix} \; para \; i = 1, 2, \dots, n$$

Esta ecuación vectorial da lugar, componente a componente, al sistema lineal:

$$\left.\begin{array}{l} v_{i1} = c_{1i}u_{11} + c_{2i}u_{21} + \cdots + c_{ni}u_{n1} \\ v_{i2} = c_{1i}u_{12} + c_{2i}u_{22} + \cdots + c_{ni}u_{n2} \\ \qquad\vdots \\ v_{in} = c_{1i}u_{1n} + c_{2i}u_{2n} + \cdots + c_{ni}u_{nn} \end{array}\right. \quad para\; i = 1, 2, \dots, n$$

Como los n sistemas tienen la misma matriz de coeficientes los podemos reducir simultánelamente usando la matriz ampliada:

$$\begin{bmatrix} u_{11} & u_{21} & \cdots & u_{n1} & \vdots & v_{11} & v_{21} & \cdots & v_{n1} \\ u_{12} & u_{22} & \cdots & u_{n2} & \vdots & v_{12} & v_{22} & \cdots & v_{n2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} & \vdots & v_{1n} & v_{2n} & \cdots & v_{nn} \end{bmatrix}$$

Aplicando el método de eliminación de Gauss-Jordan a esta matriz, se obtiene:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & \vdots & c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & 1 & \cdots & 0 & \vdots & c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \vdots & c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

Ahora bien, por lo demostrado anteriormente la parte derecha de esta matriz es $Q = P^{-1}$, luego la matriz tiene la forma:

$$[I_n \vdots P^{-1}]$$

lo cual demuestra el teorema.

Ejemplo

Hallar la matriz de cambio de base de B a B' para las bases de $\mathbb{R}^3$:

$B = \{(1,0,0),(0,1,0),(0,0,1)\}$ y $B' = \{(1,0,1),(0,-1,2),(2,3,-5)\}$

Solución

Con los vectores de las dos bases formamos las matrices B y B':

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

Ahora construimos $[B' \vdots B]$ y usamos Gauss-Jordan para reescribirla como $[I \vdots P^{-1}]$:

$$\begin{bmatrix} 1 & 0 & 2 & \vdots & 1 & 0 & 0 \\ 0 & -1 & 3 & \vdots & 0 & 1 & 0 \\ 1 & 2 & -5 & \vdots & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & \vdots & -1 & 4 & 2 \\ 0 & 1 & 0 & \vdots & 3 & -7 & -3 \\ 0 & 0 & 1 & \vdots & 1 & -2 & -1 \end{bmatrix}$$

Por tanto la matriz de cambio de B a B' es:

$$P^{-1} = \begin{bmatrix} -1 & 4 & 2 \\ 3 & -7 & -3 \\ 1 & -2 & -1 \end{bmatrix}$$

Si B es la base canónica, como en el ejemplo anterior, el proceso de transformar $[B' \vdots B]$ a $[I_n \vdots P^{-1}]$ se convierte en:

$$[B' \vdots I_n] \longrightarrow [I_n \vdots P^{-1}]$$

Pero este es el mismo proceso utilizado para calcular la inversa de la matriz B'. Es decir, si B es la base canónica de $\mathbb{R}^n$, la matriz de transición de B a B' viene dada por:

$$P^{-1} = (B')^{-1}$$

El proceso es aún más simple si B' es la base canónica, ya que la matriz $[B' \vdots B]$ está ya en la forma

$$[I_n \vdots B] = [I_n \vdots P^{-1}]$$

En este caso la matriz de transición es simplemente:

$$P^{-1} = B$$

Ejemplo

Hallar la matriz de cambio de B a B' para las siguientes bases de $\mathbb{R}^2$:

$$B = \{(-3,2),(4,-2)\}$$

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

Solución

Empezamos formando la matriz:

$$[B' \vdots B] = \begin{bmatrix} -1 & 2 & \vdots & -3 & 4 \\ 2 & -2 & \vdots & 2 & -2 \end{bmatrix}$$

y usamos el método de eliminación de Gauss-Jordan para obtener la matriz de cambio P^{-1} de B a B':

$$[I_2 \vdots P^{-1}] = \begin{bmatrix} 1 & 0 & \vdots & -1 & 2 \\ 0 & 1 & \vdots & -2 & 3 \end{bmatrix}$$

Por tanto:

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

Podríamos calcular la matriz de cambio de B' a B obteniendo:

$$[B \vdots B'] = \begin{bmatrix} -3 & 4 & \vdots & -1 & 2 \\ 2 & -2 & \vdots & 2 & -2 \end{bmatrix}$$

que se reduce a:

$$[I_2 \vdots P] = \begin{bmatrix} 1 & 0 & \vdots & 3 & -2 \\ 0 & 1 & \vdots & 2 & -1 \end{bmatrix}$$

Por tanto la matriz de cambio de B' a B es:

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

Podemos comprobar que esta matriz es la inversa de la matriz de cambio hallada anteriormente sin más que calcular el producto:

$$PP^{-1} = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

3.6.1 Coordenadas en espacios generales n-dimensionales

Una ventaja de las coordenadas es que nos permiten representar vectores de un espacio n-dimensional arbitrario con la notación de $\mathbb{R}^n$ como vemos en los siguientes ejemplos.

Ejemplo

Hallar la matriz de coordenadas de $p(x) = 3x^3 - 2x^2 + 4$ en la base canónica de P_3, $S = \{1, x, x^2, x^3\}$.

Solución

En primer lugar expresamos $p(x)$ como combinación lineal de los vectores de la base:

$$p(x) = 4(1) + 0(x) + (-2)(x^2) + 3(x^3)$$

Esto quiere decir que la matriz de coordenadas de $p(x)$ en la base S es:

$$[p(x)]_S = \begin{bmatrix} 4 \\ 0 \\ -2 \\ 3 \end{bmatrix}$$

Ejemplo

Hallar la matriz de coordenadas de:

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

en la base canónica de $\mathcal{M}_{3,1}$:

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Solución

Como X se puede escribir:

$$X = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} = (-1) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

su matriz de coordenadas en la base S es:

$$[X]_S = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$$

___

4 Aplicaciones lineales

4.1 Introducción

Definición de aplicación lineal.

Sean V y W espacios vectoriales. La aplicación $f: V \to W$ se dice que es una aplicación lineal de V en W si satisface las dos siguientes condiciones para todo $\boldsymbol{u}, \boldsymbol{v} \in V$ y para todo escalar λ:

1. $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$
2. $f(\lambda \boldsymbol{u}) = \lambda f(\boldsymbol{u})$

Hay que tener en cuenta que aunque se emplean los mismos signos para representar las operaciones vectoriales en V y en W las operaciones pueden ser diferentes.

No confundir el término aplicación lineal entre espacios vectoriales con el de función lineal (aquella cuya representación gráfica es una recta). Así $f(x) = x + 1$ es una función lineal, mientras que considerada como aplicación de $\mathbb{R}$ en $\mathbb{R}$, no es lineal ya que $f(x_1 + x_2) = x_1 + x_2 + 1$ mientras que $f(x_1) + f(x_2) = x_1 + 1 + x_2 + 1 = x_1 + x_2 + 2$.

Ejemplo

Probar que la aplicación $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ que a todo vector $\boldsymbol{v} = (v_1, v_2)$ le hace corresponder el vector:

$$f(v_1, v_2) = (v_1 - v_2, v_1 + 2v_2)$$

es lineal.

Solución

Sean $\boldsymbol{v} = (v_1, v_2)$ y $\boldsymbol{u} = (u_1, u_2)$ dos vectores cualesquiera de $\mathbb{R}^2$. Usando las propiedades de la suma y la multiplicación por un escalar, tenemos:

1. De $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$, se sigue que:

$$f(\boldsymbol{u} + \boldsymbol{v}) = f(u_1 + v_1, u_2 + v_2) = (u_1 + v_1 - (u_2 + v_2), u_1 + v_1 + 2(u_2 + v_2) =$$

$$= (u_1 - u_2, u_1 + 2u_2) + (v_1 - v_2, v_1 + 2v_2) = f(\boldsymbol{u}) + f(\boldsymbol{v})$$

2. De $\lambda \boldsymbol{u} = \lambda(u_1, u_2) = (\lambda u_1, \lambda u_2)$, se sigue que:

$$f(\lambda \boldsymbol{u}) = f(\lambda u_1, \lambda u_2) = (\lambda u_1 - \lambda u_2, \lambda u_1 + 2\lambda u_2) = \lambda(u_1 - u_2, u_1 + 2u_2) = \lambda f(\boldsymbol{u})$$

Por tanto f es una aplicación lineal.

Dos aplicaciones lineales muy simples son la **aplicación cero** y la **aplicación identidad**, definidas como sigue:

1. Aplicación cero: $f(\boldsymbol{v}) = \boldsymbol{o}, \forall \boldsymbol{v} \in V$
2. Aplicación identidad: $f(\boldsymbol{v}) = \boldsymbol{v}, \forall \boldsymbol{v} \in V$

4.2 Propiedades de las aplicaciones lineales

Sea f una aplicación lineal de V en W y sean $\boldsymbol{u}$ y $\boldsymbol{v}$ vectores de V. Se cumplen las siguientes propiedades:

1. $f(\boldsymbol{o}) = \boldsymbol{o}$
2. $f(-\boldsymbol{v}) = -f(\boldsymbol{v})$
3. $f(\boldsymbol{u} - \boldsymbol{v}) = f(\boldsymbol{u}) - f(\boldsymbol{v})$
4. Si $\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n$, entonces: $f(\boldsymbol{v}) = f(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n) = c_1 f(\boldsymbol{v}_1) + c_2 f(\boldsymbol{v}_2) + \cdots + c_n f(\boldsymbol{v}_n)$

Demostración

1. Recordemos que $0 \cdot \boldsymbol{v} = \boldsymbol{o}$, de donde $f(\boldsymbol{o}) = f(0 \cdot \boldsymbol{v}) = 0 \cdot f(\boldsymbol{v}) = \boldsymbol{o}$
2. Dado que $-\boldsymbol{v} = (-1) \cdot v$ se tiene que: $f(-\boldsymbol{v}) = \boldsymbol{f}\big((-1) \cdot v\big) = (-1) \cdot f(\boldsymbol{v}) = -f(\boldsymbol{v})$
3. Dado que $\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{u} + (-1) \cdot \boldsymbol{v}$, se tiene que: $f(\boldsymbol{u} - \boldsymbol{v}) = \boldsymbol{f}(\boldsymbol{u} + (-1) \cdot \boldsymbol{v}) = f(\boldsymbol{u}) + f\big((-1) \cdot \boldsymbol{v}\big) = f(\boldsymbol{u}) + (-1)f(\boldsymbol{v}) = f(\boldsymbol{u}) - f(\boldsymbol{v})$
4. Se demuestra aplicando al definición de aplicación lineal.

La propiedad 4 afirma que una aplicación lineal $f: V \to W$ queda completamente determinada por su acción sobre una base de V. En otras palabras, si $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ es una base del espacio vectorial V y se conocen las imágenes por f de dichos vectores ($f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \dots, f(\boldsymbol{v}_n)$), entonces $f(\boldsymbol{v})$ está determinada para todo $\boldsymbol{v}$.

Las propiedades anteriores nos proporcionan un criterio rápido para detectar aplicaciones que no son lineales. Como toda aplicación lineal debe satisfacer las cuatro propiedades, en cuanto una aplicación no satisface alguna de ellas ya podemos concluir que no es lineal.

Así, por ejemplo, la aplicación dada por $f(x_1, x_2) = (x_1 + 1, x_2)$ no es una aplicación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ ya que $f(0,0) = (1,0) \neq (0,0)$.

Teorema. Sea A una matriz $m\ x\ n$. La aplicación definida por:

$$f(\boldsymbol{v}) = A\boldsymbol{v}$$

es una aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Demostración

Para todo par de vectores $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, la propiedad distributiva del producto de matrices respecto de la suma nos lleva a concluir que:

$$f(\boldsymbol{u} + \boldsymbol{v}) = A(\boldsymbol{u} + \boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v} = f(\boldsymbol{u}) + f(\boldsymbol{v})$$

Análogamente, para todo vector $\boldsymbol{v} \in \mathbb{R}^n$ y todo escalar λ, teniendo en cuenta las propiedades conmutativa del producto matricial y de la multiplicación por un escalar , concluimos que:

$$f(\lambda\boldsymbol{v}) = A(\lambda\boldsymbol{v}) = \lambda A(\boldsymbol{v}) = \lambda f(\boldsymbol{v})$$

Con el fin de ajustar la notación a la del producto por una matriz $m\ x\ n$, los vectores de $\mathbb{R}^n$ se representan mediante matrices $n\ x\ 1$ y los vectores de $\mathbb{R}^m$ mediante matrices $m\ x\ 1$.

La matriz cero $m\ x\ n$ corresponde a la aplicación cero de $\mathbb{R}^n$ en $\mathbb{R}^m$. La matriz identidad I_n corresponde a la aplicación identidad de $\mathbb{R}^n$ en $\mathbb{R}^m$.

$$A\boldsymbol{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}$$

Ejemplo

Sea $g: \mathcal{M}_{m,n} \to \mathcal{M}_{n,m}$ la aplicación que transforma cada matriz A de tamaño $m\ x\ n$ en su traspuesta:

$g(A) = A^T$

Probar que g es una aplicación lineal

Solución

Sean A y B matrices de $m\ x\ n$ y λ un escalar. Por las propiedades de la traspuesta (2.2.6):

$f(A+B) = (A+B)^T = A^T + B^T = f(A) + f(B)$

$f(\lambda A) = (\lambda A)^T = \lambda A^T = \lambda f(A)$

4.3 Núcleo e imagen de una aplicación lineal

4.3.1 Núcleo de una aplicación lineal

Sabemos por las propiedades de las aplicaciones lineales, que toda aplicación lineal $f: V \to W$ transforma el vector cero de V en el vector cero de W, es decir, $f(\boldsymbol{o}) = \boldsymbol{o}$.

El conjunto de todos los vectores de V tales que $f(v) = \boldsymbol{o}$ se llama **núcleo** de f y se denota por $ker\ (f)$.

En el ejemplo anterior en que $g(A) = A^T$ es evidente que el único vector del núcleo es la matriz cero de $\mathcal{M}_{m,n}$.

El núcleo de la aplicación cero es V. El núcleo de la aplicación identidad es el vector cero de V.

Ejemplo

Hallar el núcleo de la aplicación lineal $f: \mathbb{R}^2 \to \mathbb{R}^3$ definida por:

$f(x_1, x_2) = (x_1 - 2x_2, 0, -x_1)$

Solución

Para hallar $ker\ (f)$, necesitamos encontrar todos los vectores de $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$ tales que:

$f(x_1, x_2) = (x_1 - 2x_2, 0, -x_1) = (0, 0, 0)$

Estos nos lleva al sistema lineal homogéneo:

$$\begin{cases} x_1 - 2x_2 = 0 \\ 0 = 0 \\ -x_1 = 0 \end{cases}$$

que admite sólo la solución trivial $(x_1, x_2) = (0,0)$, en consecuencia:

$$ker(f) = \{(0,0)\} = \{\boldsymbol{o}\}$$

Ejemplo

Hallar el núcleo de la aplicación lineal $g: \mathbb{R}^3 \to \mathbb{R}^2$ definida por $g(\boldsymbol{x}) = A\boldsymbol{x}$ donde:

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

Solución

El núcleo de g es el conjunto de todos los $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ tales que:

$g(\boldsymbol{x}) = g(x_1, x_2, x_3) = (0,0)$

Esta ecuación vectorial equivale al sistema homogéneo:

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

Llevando su matriz ampliada a forma escalonada se obtiene:

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \longrightarrow \begin{cases} x_1 = x_3 \\ x_2 = -x_3 \end{cases}$$

En términos del parámetro $x_3 = t$ la familia de soluciones se expresa:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ -t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Por tanto, el núcleo de g viene dado por:

$ker(g) = \{(t, -t, t), t \in \mathbb{R}\}$

Teorema

El núcleo de una aplicación lineal $f: V \to W$ es un subespacio de V.

Demostración

Sabemos que $ker(f)$ es un conjunto no vacío de V. Para probar que $ker(f)$ es subespacio, basta ver que es cerrado bajo la suma y la multiplicación por un escalar.

Sean $\boldsymbol{u}, \boldsymbol{v} \in ker\,(f)$, es decir, $f(\boldsymbol{u}) = f(\boldsymbol{v}) = \boldsymbol{o}$ y λ un escalar. Entonces:

1. $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v}) = \boldsymbol{o}$
2. $f(\boldsymbol{\lambda u}) = \lambda f(\boldsymbol{u}) = \lambda \boldsymbol{o} = \boldsymbol{o}$

Por tanto $ker\,(f)$ es un subespacio vectorial de V.

Ejemplo. Cálculo de una base del núcleo

Sea $f: \mathbb{R}^5 \longrightarrow \mathbb{R}^4$ definida por $f(\boldsymbol{x}) = A\boldsymbol{x}$, donde:

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

Hallar la base de $ker(f)$ coomo subespacio de $\mathbb{R}^5$.

Solución

Reducimos la matriz ampliada $[A \vdots O]$ a forma escalonada:

$$\begin{bmatrix} 1 & 0 & 2 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{cases} x_1 = -2x_3 + x_5 \\ x_2 = x_3 + 2x_5 \\ x_4 = -4x_5 \end{cases}$$

Haciendo $x_3 = s$ y $x_5 = t$, obtenemos:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s + t \\ s + 2t \\ s \\ -4t \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

Así pues, la base del $ker(f)$ viene dada por:

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

En este ejemplo hemos encontrado una base del núcleo de f resolviendo el sistema homogéneo $A\boldsymbol{x} = \boldsymbol{o}$. Este procedimiento es el mismo utilizado para hallar el espacio solución de $A\boldsymbol{x} = \boldsymbol{o}$. En otras palabras, el núcleo de f es el espacio nulo de la matriz A.

Corolario

Sea $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ la aplicación lineal dada por $f(\boldsymbol{x}) = A\boldsymbol{x}$. El núcleo de f es el espacio solución de $A\boldsymbol{x} = \boldsymbol{o}$.

4.3.2 Imagen de una aplicación lineal

La imagen de una aplicación lineal $f: V \to W$, denotado por $imagen(f)$ es el conjunto de todos los vectores de W que son imágenes de algún vector de V, es decir:

$imagen(f) = \{\boldsymbol{w} \in W, \exists\, \boldsymbol{v} \in V, f(\boldsymbol{v}) = \boldsymbol{w}\}$

Teorema

La imagen de una aplicación lineal $f: V \to W$ es un subespacio de W.

Demostración

La imagen de f es no vacía, ya que el vector nulo de W pertenece a ella al ser $f(\boldsymbol{o}) = \boldsymbol{o}$.

Sólo falta probar que es cerrada bajo la suma y la multiplicación por un escalar.

a) Sean $f(\boldsymbol{u}), f(\boldsymbol{v}) \in imagen(f)$. Puesto que $\boldsymbol{u}, \boldsymbol{v} \in V \Longrightarrow \boldsymbol{u} + \boldsymbol{v} \in V$, por tanto $f(\boldsymbol{u}) + f(\boldsymbol{v}) = f(\boldsymbol{u} + \boldsymbol{v}) \in imagen(f)$.
b) Sea $f(\boldsymbol{u}) \in imagen(f)$ y λ un escalar. Como $\boldsymbol{u} \in V \Longrightarrow \lambda\boldsymbol{u} \in V$, por tanto $\lambda f(\boldsymbol{u}) = f(\lambda\boldsymbol{u}) \in imagen(f)$.

4.3.2.1 Cálculo de una base de la imagen

Para hallar una base de la imagen de una aplicación lineal definida como $f(x) = A\boldsymbol{x}$ observemos que la imagen contiene todos aquellos vectores $\boldsymbol{b}$ para los que el sistema $A\boldsymbol{x} = \boldsymbol{b}$ admite solución (es compatible).

Escribiendo el sistema:

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

en la forma:

$$A\boldsymbol{x} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \boldsymbol{b}$$

vemos que $\boldsymbol{b}$ está en la imagen de f sí, y sólo sí, es combinación lineal de los vectores columna de A. Así pues, <u>el espacio de columnas de la matriz A coincide con la imagen de f</u>.

En 3.5.1 vimos dos procedimientos para hallar una base del espacio de columnas de una matriz. A continuación aplicamos el segundo de estos procedimientos para calcular una base de la imagen de una aplicación lineal.

<u>Ejemplo</u>

Hallar una base de la imagen de la aplicación lineal del ejemplo anterior $f: \mathbb{R}^5 \longrightarrow \mathbb{R}^4$, definida por $f(\boldsymbol{x}) = A\boldsymbol{x}$, donde:

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

Solución

Ya calculamos la forma escalonada de la matriz A:

$$\begin{bmatrix} 1 & 2 & 0 & 1 & -1 \\ 2 & 1 & 3 & 1 & 0 \\ -1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Como los pivotes aparecen en las columnas 1, 2 y 4 de la matriz escalonada, los correspondientes vectores columna de A forman una base del espacio de columnas de A. Por tanto, una base de $imagen(f)$ es:

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

4.3.2.2 Rango y nulidad de una aplicación lineal

Sea $f: V \to W$ una aplicación lineal. La dimensión del núcleo de f se llama **nulidad** de f y se denota por $nul(f)$. La dimensión de la imagen de f se llama **rango** de f y se denota por $rango(f)$.

Si f viene definida mediante una matriz A, el rango de f es igual al rango de A definido en 3.5.2.

Teorema

Sea $f: V \to W$ una aplicación lineal de un espacio vectorial V, de dimensión n, en un espacio vectorial W. La suma de las dimensiones de la imagen y el núcleo de f es igual a la dimensión del espacio vectorial V. Esto es:

$$rango(f) + nul(f) = n$$

o sea:

$$dim(imagen(f)) + dim(núcleo(f)) = dim(V)$$

Demostración

Supongamos que f viene representada por una matriz A de tamaño $m\ x\ n$. Supongamos que la matriz A tiene rango r. Entonces:

$$rango(f) = \dim\big(imagen(f)\big) = \dim(espacio\ de\ columnas) = rango(A) = r$$

Por el corolario visto en 4.3.1 sabemos que:

$$nul(f) = \dim\big(núcleo(f)\big) = \dim(espacio\ solución\ de\ A\mathbf{x} = 0) = n - r$$

Por tanto:

$$rango(f) + nul(f) = r + (n - r) = n$$

Ejemplo

Hallar el rango y la nulidad de la aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^3$ definida por la matriz:

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

Solución

Como A está en forma escalonada y tiene dos filas no nulas, su rango es 2. En consecuencia, el rango de f es 2, luego su nulidad es:

$$dim(\mathbb{R}^3) - rango = 3 - 2 = 1$$

La relación entre el rango y la nulidad de una aplicación lineal dada por una matriz se puede calcular observando que el rango lo determina el número de pivotes y la nulidad el número de variables libres (columnas sin pivotes). Su suma es el número total de columnas de la matriz, que es la dimensión del espacio vectorial de partida.

En el ejemplo anterior las dos primeras columnas tienen pivotes, lo cual implica que el rango es 2. La tercera columna corresponde a una variable libre, lo que indica que la nulidad es 1.

Ejemplo

Sea $g: \mathbb{R}^5 \to \mathbb{R}^7$ una aplicación lineal. Hallar:

1. La dimensión del núcleo de g si la dimensión de la imagen es 2.
2. El rango de g si la nulidad de g es 4.
3. El rango de g si $\ker(g) = \{\boldsymbol{o}\}$

Solución

1. Como $dim(imagen(g)) + dim(núcleo(g)) = dim(V)$, sustituyendo los valores conocidos, se tiene $2 - dim(núcleo(g)) = 5$, por lo que $\dim(núcleo(g)) = 5 - 2 = 3$.
2. Como $rango(g) + nul(g) = n$, sustituyendo los valores conocidos, se tiene: $rango(g) + 4 = 5$, por lo que $rango(g) = 5 - 4 = 1$.
3. En este caso la nulidad de g es 0. Por tanto: $rango(g) = n - nul(g) = 5 - 0 = 5$.

4.4 Isomorfismos de espacios vectoriales

4.4.1 Aplicaciones lineales inyectivas y suprayectivas

4.4.1.1 Aplicaciones lineales inyectivas

Una aplicación $f: V \to W$ se dice que es inyectiva sí, y sólo sí, la preimagen de todo $\boldsymbol{w} \in imagen(f)$ consta de un único vector. Eso equivale a decir que f es inyectiva sí, y sólo sí:

$\forall\, \boldsymbol{u}, \boldsymbol{v} \in V, f(\boldsymbol{u}) = f(\boldsymbol{v}) \Longrightarrow \boldsymbol{u} = \boldsymbol{v}$

Teorema

Una aplicación lineal $f: V \to W$ es inyectiva sí, y sólo sí, $\ker(f) = \{\boldsymbol{o}\}$

Demostración

Si f es inyectiva, $f(\boldsymbol{v}) = \boldsymbol{o}$ sólo puede tener una solución: $\boldsymbol{v} = \boldsymbol{o}$. En tal caso $ker(f) = \{\boldsymbol{o}\}$. Recíprocamente, supongamos que $ker(f) = \{\boldsymbol{o}\}$ y $f(\boldsymbol{u}) = f(\boldsymbol{v})$. Por ser f lineal:

$$f(\boldsymbol{u} - \boldsymbol{v}) = f(\boldsymbol{u}) - f(\boldsymbol{v}) = \boldsymbol{o}$$

Esto implica que el vector $\boldsymbol{u} - \boldsymbol{v}$ está en el núcleo de f, de modo que $\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{o}$, luego $\boldsymbol{u} = \boldsymbol{v}$ y concluimos que f es inyectiva.

Teorema. Si V tiene dimensión finita, entonces f es inyectiva sí, y sólo sí, $\dim V = \dim f(V)$

Demostración

De acuerdo con el teorema anterior y sabiendo que $\dim(Ker(f)) + \dim(Imagen(f)) = \dim V$, se obtiene:

$$f\ inyectiva \Leftrightarrow Ker(f) = \boldsymbol{o} \Leftrightarrow \dim(Ker(f)) = 0 \Leftrightarrow \dim V = \dim(f(V)) + 0$$

Teorema. Si $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots., \boldsymbol{e}_n\}$ es una base de V, entonces f es inyectiva sí, y sólo sí, $f(B) = (f(\boldsymbol{e_1}), f(\boldsymbol{e_2}), \dots, f(\boldsymbol{e_n}))$ es una base de $f(V)$, es decir, si y sólo si $f(B)$ es un sistema independiente de vectores de W.

Demostración

Dada una base $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ de V, como $f(B)$ es un sistema generador de n vectores de $f(V)$, se verifica que:

$$f(B) base\ de\ f(V) \Leftrightarrow f(B)\ linealmente\ independiente \Leftrightarrow rang\big(f(B)\big) = n \Leftrightarrow \dim\big(f(V)\big) = n = \dim V \Leftrightarrow f\ inyectiva$$

Observación

Una aplicación lineal $f: V \to W$, entre espacios vectoriales es inyectiva si y sólo si, todo sistema linealmente independientes de vectores de V tiene por imagen a un sistema de vectores de W que también es linealmente independiente.

Comprobación

a) Si f es inyectiva, dado un sistema $(\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_p)$ de vectores independientes de V, el sistema imagen $(f(\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_p)$ también es linealmente independiente, puesto que (para $\lambda_1, \lambda_2, \dots, \lambda_p$ escalares):

$$\sum_{i=1}^{p} \lambda_i f(\boldsymbol{u}_i) = \boldsymbol{o} \Rightarrow \boldsymbol{f}\left(\sum_{i=1}^{p} \lambda_i \boldsymbol{u}_i\right) = \boldsymbol{o} \Rightarrow \sum_{i=1}^{p} \lambda_i \boldsymbol{u}_i = \boldsymbol{o} \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$$

b) Si la imagen de cualquier sistema independiente de V es un sistema independiente de W, se verificará en particular que dado $\boldsymbol{u} \neq \boldsymbol{o}$ en V, ha de ser $f(\boldsymbol{u}) \neq \boldsymbol{o}$**.** De ello resulta, pues, que $Kerf(f) = O$ y, por tanto, f es inyectiva.

Ejemplo

La aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^4$ definida mediante:

$$f(x, y, z) = (x, x + y, y + z, x + y + z)$$

es inyectiva ya que los vectores de la base canónica de $\mathbb{R}^3$ son os vectores:

$$f(1,0,0) = (1,1,0,1)\ ;\ \ f(0,1,0) = (0,1,1,1)\ ;\ \ f(0,0,1) = (0,0,1,1)$$

que forman una base de la imagen, ya que son linealmente independientes pues:

$$rango(f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), f(\boldsymbol{e}_3)) = rango \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = 3$$

La inyectividad de f también se podía haber comprobado viendo que $\ker(f) = O$; así es, ya que:

$$\ker(f): \begin{cases} x = 0 \\ x + y = 0 \\ y + z = 0 \\ x + y + z = 0 \end{cases} \Leftrightarrow \ker(f): \begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases} \Leftrightarrow \ker(f) = O$$

4.4.1.2 Aplicaciones lineales suprayectivas[4]

[4] También denominadas *sobreyectivas*

Una aplicación $f: V \to W$ se dice que es suprayectiva si todo elemento de W tiene alguna preimagen en V. En otras palabras, si la imagen de f es W. De ahí se sigue el siguiente teorema:

Teorema

Se dice que $f: V \to W$, donde W es de dimensión finita, es suprayectiva sí, y sólo sí, el rango de f es igual a la dimensión de W.

En el caso de que V y W sean de la misma dimensión, se verifica que:

Teorema

Una aplicación lineal $f: V \to W$, con $V\ y\ W$ de la misma dimensión n, es inyectiva sí, y sólo sí, es suprayectiva.

Demostración

Si f es inyectiva, sabemos que $ker(f) = \{\boldsymbol{o}\}$, luego $dim\ (ker(f) = 0$. En este caso:

$$dim\big(imagen(f)\big) = n - dim(ker(f)) = n = dim\ (W)$$

Por tanto, por el teorema anterior, f es suprayectiva.

Análogamente, si f es suprayectiva:

$$dim\big(imagen(f)\big) = dim(W) = n$$

lo cual implica que $dim\ (ker\ (f)) = 0$, por lo que f es inyectiva.

Ejemplo

La aplicación lineal $f: \mathbb{R}^n \to \mathbb{R}^m$ viene dada por $f(\boldsymbol{x}) = A\boldsymbol{x}$. Hallar la nulidad y el rango de f. Averiguar si f es inyectiva, suprayectiva o ninguna de ambas cosas en los siguientes casos:

a) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Solución

Cada una de las matrices ya está en forma escalonada, de manera que su rango se sabe a simple vista.

$f: \mathbb{R}^n \to \mathbb{R}^m$	$dim\ (dominio)$	$dim\ (imagen)$ $rango(f)$	$dim\ (núcleo)$ $nul(f)$	$Inyectiva$	$Suprayectiva$
a) $f: \mathbb{R}^3 \to \mathbb{R}^3$	3	2	0	SI	SI
b) $f: \mathbb{R}^2 \to \mathbb{R}^3$	2	2	0	SI	NO
c) $f: \mathbb{R}^3 \to \mathbb{R}^2$	3	2	1	NO	SI
d) $f: \mathbb{R}^3 \to \mathbb{R}^3$	3	2	1	NO	NO

4.4.2 Isomorfismos entre espacios vectoriales

Definición. Un **isomorfismo** entre los espacios vectoriales V y W es una aplicación lineal que es inyectiva y suprayectiva[5]. En tal circunstancia se dice que V y W son **isomorfos**.

Desde el punto de vista de los espacios vectoriales, no hay nada que permita diferenciar a V de W.

Un isomorfismo de un espacio vectorial en sí mismo recibe el nombre de **automorfismo**.

Teorema

Dos espacios vectoriales V y W son isomorfos sí, y sólo sí, tienen la misma dimensión.

Demostración

Supongamos que V, de dimensión n es isomorfo a W. Por definición de isomorfismo, existe una aplicación lineal $f: V \to W$ que es inyectiva y suprayectiva.

Por ser inyectiva: $dim(\ker(f)) = 0$, lo cual implica que

$$dim\big(imagen(f)\big) = dim(W) = n$$

Por ser suprayectiva: $dim\big(imagen(f)\big) = dim(W) = n$

Para demostrarla otra dirección del enunciado, supongamos que V y W tienen dimensión n. Sea $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots., \boldsymbol{v}_n\}$ una base de V y $B' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ una base de W. Un vector arbitrario de V se puede descomponer como:

$$\boldsymbol{v} = \lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_n \boldsymbol{v}_n$$

y podemos definir una aplicación lineal $f: V \to W$ dada por:

$$f(\boldsymbol{v}) = \lambda_1 \boldsymbol{w}_1 + \lambda_2 \boldsymbol{w}_2 + \cdots + \lambda_n \boldsymbol{w}_n$$

Es fácil verificar que f es inyectiva y suprayectiva, por tanto V y W son isomorfos.

Ejemplo

[5] También llamadas *biyectivas*

Los siguientes espacios son isomorfos entre sí:

a) $\mathbb{R}^4$
b) $\mathcal{M}_{4,1}$ = espacio de todas las matrices de tamaño $4\ x\ 1$
c) $\mathcal{M}_{2,2}$ = espacio de todas las matrices de tamaño $2\ x\ 2$
d) P_3 = espacio de todos los polinomios de grado menor o igual que 3.
e) $V = \{(x_1, x_2, x_3, x_4, 0) : x_i \in \mathbb{R}\}$, subespacio de $\mathbb{R}^5$

Los elementos de esos espacios se comportan todos del mismo modo como vectores, aun cuando sean objetos matemáticos muy diversos. Por esta razón, el convenio de usar indistintamente ela notación n-upla o de matriz $n\ x\ 1$ queda plenamente justificada.

Propiedades

1. Una aplicación lineal $f: V \to W$ es un isomorfismo si, y sólo si: $imagen(f) = W$ y $Ker\ (f) = O$.
2. Si V tiene dimensión finita, una aplicación lineal $f: V \to W$ es un isomorfismo si, y sólo si, $\dim V = \dim f(V) = \dim W$.
3. Si V tiene dimensión finita, una aplicación lineal $f: V \to V$ es automorfismo si y sólo si, es inyectiva o si y sólo si es suprayectiva.

Demostración

1. Por definición f es suprayectiva si y sólo si, $imagen(f) = W$. Por otra parte, f es inyectiva si y sólo si $Ker(f) = O$.
2. Esta propiedad se deduce fácilmente de:
   a. $f\ suprayectiva \Leftrightarrow f(V) = W \Leftrightarrow \dim\big(f(V)\big) = \dim W$.
   b. $f\ inyectiva\ \Leftrightarrow \dim\big(Ker(f)\big) = 0 \Leftrightarrow \dim V = \dim f(V)$.
3. Si en el caso anterior se toma $W = V$, resulta que f es isomorfismo si, y sólo si, $\dim(V) = \dim\ (f(V)$, es decir, si y sólo si, f es suprayectiva. La anterior igualdad entre dimensiones equivale a la $\dim\big(Ker(f)\big) = 0$, o sea, a que f es inyectiva.

4.5 Matrices de las aplicaciones lineales

4.5.1 Matriz canónica de una aplicación lineal

Sea $f: \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal y $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ la base canónica de $\mathbb{R}^n$, tal que:

$$f(\boldsymbol{e}_1) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, f(\boldsymbol{e}_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, f(\boldsymbol{e}_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

entonces la matriz $m\ x\ n$ cuyas n columnas vienen dadas por $f(\boldsymbol{e}_i)$:

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

es tal que $f(\boldsymbol{v}) = A\boldsymbol{v}$ para todo $\boldsymbol{v} \in \mathbb{R}^n$. Se dice que A es la matriz canónica de f.

Demostración

Para probar que $f(\boldsymbol{v}) = A\boldsymbol{v}$ para todo $\boldsymbol{v} \in \mathbb{R}^n$, escribimos:

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + \cdots + v_n\boldsymbol{e}_n$$

Como f es lineal, se tiene:

$$f(\boldsymbol{v}) = f(v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + \cdots + v_n\boldsymbol{e}_n) = f(v_1\boldsymbol{e}_1) + f(v_2\boldsymbol{e}_2) + \cdots + f(v_n\boldsymbol{e}_n) =$$

$$= v_1 f(\boldsymbol{e}_1) + v_2 f(\boldsymbol{e}_2) + \cdots + v_n f(\boldsymbol{e}_n)$$

Por otra parte, el producto de matrices $A\boldsymbol{v}$ viene dado por (ver teorema de 4.2) :

$$A\boldsymbol{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix} =$$

$$= v_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = v_1 f(\boldsymbol{e}_1) + v_2 f(\boldsymbol{e}_2) + \cdots + v_n f(\boldsymbol{e}_n)$$

Por tanto, $f(\boldsymbol{v}) = A\boldsymbol{v}$ para todo $\boldsymbol{v} \in \mathbb{R}^n$.

Ejemplo

Hallar la matriz canónica de la aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^2$ definida por:

$$f(x, y, z) = (x - 2y, 2x + y)$$

Calculamos las imágenes de la base canónica de $\mathbb{R}^3$:

$$f(\boldsymbol{e}_1) = f(1,0,0) = (1,2)$$

$$f(\boldsymbol{e}_2) = f(0,1,0) = (-2,1)$$

$$f(\boldsymbol{e}_3) = f(0,0,1) = (0,0)$$

Por tanto la matriz de la aplicación es:

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

Para comprobarlo:

$$A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - 2y \\ 2x + y \end{bmatrix}$$

4.5.2 Bases no canónicas y espacios vectoriales generales

Consideramos ahora el problema más general de hallar una matriz para una aplicación lineal $f: V \to W$, relativa a dos bases ordenadas de V y W, digamos B y B', respectivamente. Recordemos que la matriz de coordenadas de $\boldsymbol{v}$ en la base B se denota por $[\boldsymbol{v}]_B$. Para representar la aplicación lineal f, debemos multiplicar A por una matriz de coordenadas respecto de B. El resultado de ese producto será una matriz de coordenadas respecto de B'. Esto es:

$$[f(\boldsymbol{v})]_{B'} = A[\boldsymbol{v}]_B$$

Se dice que A es la **matriz de f respecto de las bases B y B'.**

Para hallar la matriz A usamos un procedimiento análogo al empleado para hallar la matriz canónica de f, es decir, escribimos las imágenes de los vectores de B como matrices de coordenadas relativas a la base B'. Estas matrices de coordenadas son columnas de A.

Definición.

Sean V y W espacios vectoriales de dimensión finita con bases respectivas B y B', donde

$$B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$$

Si $f: V \to W$ es una aplicación lineal tal que:

$$[f(\boldsymbol{v}_1)]_{B'} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \ [f(\boldsymbol{v}_2)]_{B'} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ \dots\dots, \ [f(\boldsymbol{v}_n)]_{B'} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

La matriz $m\ x\ n$ con columnas $[f(\boldsymbol{v}_i)]_{B'}$,

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

Es tal que

$$[f(\boldsymbol{v})]_{B'} = A[\boldsymbol{v}]_B \ , \ \ \forall\ \boldsymbol{v} \in V$$

Ejemplo

Sea $f: \mathbb{R}^2 \to \mathbb{R}^2$ la aplicación lineal definida por:

$$f(x_1, x_2) = (x_1 + x_2, 2x_1 - x_2)$$

Hallar la matriz de f respecto de las bases:

$$B = \{\boldsymbol{v}_1, \boldsymbol{v}_2\} = \{(1,2), (-1,1)\} \qquad y \qquad B' = \{\boldsymbol{w}_1, \boldsymbol{w}_2\} = \{(1,0), (0,1)\}$$

Solución

Por definición de f:

$$f(\boldsymbol{v}_1) = f(1,2) = (3,0) = 3\boldsymbol{w}_1 + 0\boldsymbol{w}_2$$

$$f(\boldsymbol{v}_2) = f(-1,1) = (0,-3) = 0\boldsymbol{w}_1 - 3\boldsymbol{w}_2$$

Por tanto, las matrices de coordenadas de $f(\boldsymbol{v}_1)$ y $f(\boldsymbol{v}_2)$ respecto de la base B' son:

$$[f(\boldsymbol{v}_1)]_{B'} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \qquad \text{y} \qquad [f(\boldsymbol{v}_2)]_{B'} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$$

La matriz de f respecto de B y B' se halla tomando estas matrices de coordenadas como columnas:

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

En el caso especial en que $V = W$ y $B = B'$, la matriz A se llama matriz de f respecto de la base B. En tal circunstancia, la matriz de la aplicación es, simplemente, I_n.

Ejemplo

Sea $D_x: P_2 \to P_1$ el operador derivada que transforma todo polinomio de grado 2 $p(x)$ en su derivada $p'(x)$. Hallar la matriz de D_x en las bases:

$B = \{1, x, x^2\}$ y $B' = \{1, x\}$

Solución

Las derivadas de los vectores de la base son:

$$D_x(1) = 0(1) + 0(x)$$

$$D_x(x) = 1(1) + 0(x)$$

$$D_x(x^2) = 0(1) + 2(x)$$

Así pues, sus matrices de coordenadas respecto de B' son:

$[D_x(1)]_{B'} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $[D_x(x)]_{B'} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $[D_x(x^2)]_{B'} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

y la matriz de D_x viene dada por:

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

Nótese que esta matriz produce la derivada de todo polinomio de grado 2 $p(x) = a + bx + cx^2$:

$$Ap = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ 2c \end{bmatrix} \Rightarrow b + 2cx = D_x[a + bx + cx^2]$$

4.6 Equivalencia, semejanza y congruencia de matrices

4.6.1 Expresión matricial de un cambio de base

Dadas dos bases distintas $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ y $B' = \{\boldsymbol{e'}_1, \boldsymbol{e'}_2, \dots, \boldsymbol{e'}_n\}$ de un espacio vectorial V de dimensión n, un vector cualquiera de $\boldsymbol{x} \in V$ puede ser expresado de las dos formas siguientes:

$$\boldsymbol{x} = x_i \boldsymbol{e}_i \quad \text{y} \quad \boldsymbol{x} = x'_j \boldsymbol{e'}_j$$

donde $(x_1, x_2, \dots, x_n)$ son las coordenadas de $\boldsymbol{x}$ en la base B y $(x'_1, x'_2, \dots, x'_n)$ son las coordenadas de $\boldsymbol{x}$ en la base B'.

Si la relación que guardan los vectores de las bases entre sí es:

$$\boldsymbol{e'}_j = q_{ij} \boldsymbol{e}_i$$

ambos sistemas de coordenadas dependen el uno del otro y dicha dependencia no es otra que:

$$x_i = q_{ij} x'_j \quad para\, i = 1, 2, \dots, n \;\; j = 1, 2, \dots, n$$

Designando entonces por:

$$Q = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \cdots & \cdots & \cdots \\ q_{n1} & \cdots & q_{nn} \end{bmatrix} = [q_{ij}] \quad , \quad X = [x_i] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad , \quad X' = [x'_i] = \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}$$

resulta evidente que las ecuaciones del cambio de coordenadas se pueden escribir matricialmente de la forma:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \cdots & \cdots & \cdots \\ q_{n1} & \cdots & q_{nn} \end{bmatrix} \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}$$

o abreviadamente:

$$X = QX', \qquad [x_i] = [q_{ij}][x'_j]$$

La expresión del cambio de base puede ser interpretada como la ecuación de una aplicación lineal, la aplicación $f_Q: V \to V$ que para las bases B' del espacio origen y B del espacio de llegada, tiene a Q por matriz asociada.

Como cualquier vector $\boldsymbol{x} \in V$ tiene coordenadas (x_i) y coordenadas (x'_j). la anterior interpretación permite afirmar que cualquier vector $x_i\boldsymbol{e}_i$ del espacio de llegada procede de un vector $x'_j\boldsymbol{e}'_j$ del espacio origen, es decir, que f_Q es suprayectiva y que, por tanto Q es una matriz regular.

En consecuencia, un cambio de base en un espacio vectorial viene dado matricialmente por:

$$X = QX', \qquad [x_i] = [q_{ij}][x'_j]$$

siendo Q matriz regular.

Recíprocamente, dada una matriz regular Q de tamaño n y una base $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ del espacio vectorial V de dimensión n, se puede asegurar que existe otra base $B' = \{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \dots, \boldsymbol{e}'_n\}$ de V tal que la ecuación matricial del cambio de coordenadas es $X = QX'$.

4.6.2 Equivalencia de matrices

4.6.2.1 Definición

Dos matrices A y B, ambas de m filas y n columnas, es decir $A, B \in \mathcal{M}_{mxn}(\mathbb{R})$ se dice que son matrices **equivalentes** siempre que existan dos matrices P y Q, cuadradas y regulares, de tamaños m y n respectivamente, tales que:

$$B = Q^{-1}AP$$

Es decir:

$$\boxed{A_{mxn} \ y \ B_{mxn} \ son \ equivalentes \ \Leftrightarrow Existen \ P_n \ y \ Q_m \ matrices \ regulares, tales \ que: B = Q^{-1}AP}$$

Esta relación entre matrices es una *relación de equivalencia* para el conjunto $\mathcal{M}_{mxn}(\mathbb{R})$ de las matrices de tamaño $m \ x \ n$. En efecto:

a) Toda matriz A_{mxn} es equivalente a sí misma, ya que $A = I_m^{-1}AI_n$ y las matrices identidad I_m e I_n son regulares.
b) Si A_{mxn} es equivalente a B_{mxn}, entonces B es equivalente a A, ya que de ser $B = Q^{-1}AP$ se sigue que $A = QBP^{-1}$ o bien $A = (Q^{-1})^{-1}B(P^{-1})$, donde P^{-1} y Q^{-1} son regulares.

c) Si A_{mxn} es equivalente es equivalente a B_{mxn} y B_{mxn} lo es a C_{mxn}, entonces A es equivalente a C; ya que, de ser $B = Q^{-1}AP$ y $C = R^{-1}BS$, se sigue que $C = R^{-1}(Q^{-1}AP)S = (QR)^{-1}A(PS)$ y PS y QR son regulares.

En términos de aplicaciones lineales, la interpretación de esta definición es como sigue:

4.6.2.2 Matrices asociadas, en distintas bases, a una misma aplicación lineal

Sean V_n y W_m espacios vectoriales de dimensiones n y m respectivamente, y sea f una aplicación lineal de V en W, que transforma todo vector $\boldsymbol{x} \in V$ en otro $\boldsymbol{y} = f(\boldsymbol{x}) \in W$. Si en V se elige una base $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ y en W otra base $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m\}$, respecto de éstas bases hay una, y sólo una, matriz $A_{mxn} = [a_{ij}]$ asociada a la aplicación f. Esta matriz es la que permite escribir la ecuación de f en dichas bases, en la forma:

$$[y_i] = [a_{ij}][x_j]$$

siendo $X = [x_j] \in \mathcal{M}_{nx1}(\mathbb{R})$ la matriz columna de las coordenadas de $\boldsymbol{x}$ e $Y = [y_i] \in \mathcal{M}_{mx1}(\mathbb{R})$ la matriz columna de las coordenadas de $\boldsymbol{y}$.

Supóngase que en V se toma otra base $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \dots, \boldsymbol{e}'_n\}$ respecto de la cual es $\boldsymbol{x} = x'_j \boldsymbol{e}'_j$ y que en W también se elige una nueva base $\{\boldsymbol{u}'_1, \boldsymbol{u}'_2, \dots, \boldsymbol{u}'_m\}$ que permite escribir $\boldsymbol{y} = y'_i \boldsymbol{u}'_i$. Adoptadas estas nuevas bases la aplicación f tiene asociada una nueva matriz $B_{mxn} = [b_{ij}]$, la cual proporciona la ecuación de f en las nuevas bases, en la forma:

$$[y'_i] = [b_{ij}][x'_j]$$

Los cambios de base en V y W quedan biunívocamente determinados por dos matrices cuadradas regulares P_n y Q_m, que son las que permiten expresar los cambios de coordenadas en la siguiente forma matricial:

$$X = PX' \quad \text{e} \quad Y = QY'$$

En estas condiciones, como $[y_i] = A[x_j]$, será:

$$Q[y'_i] = A(P[x'_j])$$

y, por tanto,

$$[y'_i] = (Q^{-1}AP)[x'_j]$$

de donde se infiere que:

$$B = Q^{-1}AP$$

es decir, A y B son equivalentes; de ahí que las matrices asociadas a una determinada aplicación, en distintas bases, son equivalentes.

Es trivial, a la vista de lo anterior, que si dos matrices son equivalentes, entonces son matrices asociadas a una misma aplicación en distintas bases, o, dicho de otro modo, si A y B son matrices equivalentes y A tiene asociada en ciertas bases una aplicación lineal f, existen otras dos bases respecto de las cuales es B la matriz asociada a f.

La correspondencia entre aplicaciones lineales y matrices es pues tal que, al cambiar de bases en los espacios vectoriales, las matrices asociadas a una aplicación lineal dada constituyen una clase de matrices equivalentes.

4.6.3 Semejanza de matrices

La semejanza de matrices puede considerarse como un caso particular de la equivalencia. Así como la equivalencia se definía entre matrices rectangulares cualesquiera, pero ambas de igual tamaño, la semejanza sólo se define entre matrices cuadradas del mismo tamaño.

4.6.3.1 Definición

Dos matrices cuadradas A y B de tamaño n, se dice que son matrices semejantes si existe una matriz P cuadrada de tamaño n y regular, tal que se verifica:

$$B = P^{-1}AP$$

es decir,

$$[A\ y\ B\ son\ matrices\ semejantes] \Leftrightarrow [Existe\ una\ matriz\ P_n, tal\ que\ B = P^{-1}AP]$$

La relación de semejanza es evidentemente, para el conjunto $\mathcal{M}_n(\mathbb{R})$ de matrices cuadradas, una relación de equivalencia. La comprobación se realiza siguiendo los mismos pasos que en 4.6.2.1.

Es frecuente, para expresar que A y B son matrices semejantes, utilizar la expresión: B *es la matriz transformada de* A *mediante una matriz regular*.

La interpretación en términos de aplicaciones lineales de la semejanza de matrices, es ahora más simple que la de equivalencia.

4.6.3.2 Matrices asociadas, al cambiar de base, a un endomorfismo.

Sean V un espacio vectorial de dimensión n y f una aplicación lineal de V en sí mismo, que transforma todo vector $\boldsymbol{x} \in V$ en otro $\boldsymbol{y} = f(\boldsymbol{x})$. Si en V se elige una base $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$, respecto de esta base hay una, y sólo una, matriz $A_n = [a_{ij}]$ asociada al endomorfismo f, respecto de dicha base, en la forma $Y = A \cdot X$, siendo $X = [x_j]$, $Y = [y_i] \in \mathcal{M}_{nx1}(\mathbb{R})$ las matrices columna de las coordenadas de $\boldsymbol{x}$ e $\boldsymbol{y}$ respectivamente.

Adoptando para V una nueva base $\{\boldsymbol{e'}_1, \boldsymbol{e'}_2, \dots, \boldsymbol{e'}_n\}$, respecto de ella, la matriz asociada a f será una cierta $B_n = [b_{ij}]$ tal que, en la nueva base, la ecuación matricial de f es $Y' = B \cdot X'$, donde $X' = [x'_j]$ e $Y' = [y'_i]$ son las matrices columna de las nuevas coordenadas de $\boldsymbol{x}$ e $\boldsymbol{y}$ respectivamente.

El cambio de base en V queda inequívocamente determinado por una matriz cuadrada regular P_n que permite expresar el cambio de coordenadas de la siguiente manera:

$$X = PX', \ \ Y = PY'$$

En estas condiciones, al ser $Y = A \cdot X$ será

$$PY' = A \cdot (PX')$$

y, por tanto,

$$Y' = (P^{-1}AP)X'$$

de ahí que sea

$$B = P^{-1}AP$$

es decir, A y B son semejantes y P es la matriz que transforma A en B.

Resulta evidente que dos matrices cuadradas son semejantes si, y sólo si, son matrices asociadas a un mismo endomorfismo en bases distintas.

4.6.3.3 Propiedades

Para la semejanza entre matrices de $\mathcal{M}_n(\mathbb{R})$ se verifica que:

a) La transformada de una suma es la suma de las transformadas, es decir:

$$P^{-1}(A_1 + A_2)P = P^{-1}A_1P + P^{-1}A_2P$$

b) La transformada de un escalar por una matriz es igual al escalar por la transformada de la matriz, es decir:

$$P^{-1}(kA)P = k(P^{-1}AP)$$

c) La transformada de un producto, es el producto de las transformadas, es decir:

$$P^{-1}(A_1A_2)P = (P^{-1}A_1P) \cdot (P^{-1}A_2P)$$

d) La transformada de la matriz unidad es ella misma, es decir:

$$P^{-1}I_nP = I_n$$

e) La transformada de la inversa de una matriz es la inversa de la transformada de la matriz, es decir:

$$P^{-1}(A^{-1})P = (P^{-1}AP)^{-1}$$

De las propiedades anteriores se deduce que, para el conjunto de las matrices regulares de tamaño n, la aplicación definida por la matriz regular P, mediante $A \longmapsto P^{-1}AP$, es un automorfismo.

4.6.4 Congruencia de matrices

Dos matrices cuadradas A y B de tamaño n, se dice que son matrices congruentes si existe una matriz Q cuadrada de tamaño n y regular, tal que se verifique:

$B = Q^TAQ$

$$[A_n\ y\ B_n\ son\ dos\ matrices\ congruentes] \Leftrightarrow [Existe\ una\ matriz\ Q\ regular, tal\ que\ B = Q^TAQ$$

Para el conjunto $\mathcal{M}_n(\mathbb{R})$ de las matrices cuadradas de tamaño n, la congruencia de matrices es una relación de equivalencia, ya que:

a) Toda matriz es congruente consigo misma, pues $A = I^TAI$, donde I es la matriz identidad, que es regular.
b) Si A es congruente con B, es decir $B = Q^TAQ$, entonces $A = (Q^{-1})^TBQ^{-1}$, es decir, B es congruente con A.
c) Si A es congruente con B, y B lo es con C, es decir, $B = Q^TAQ$ y $C = R^TBR$, entonces

$$C = R^T(Q^TAQ)R = (QR)^TA(QR)$$

o sea, A es congruente con C.

Obsérvese que, si una matriz $A \in \mathcal{M}_n(\mathbb{R})$ es una matriz simétrica, también son simétricas todas las matrices congruentes con ella, es decir, si $A = A^T$, entonces $B = Q^T A Q$ también satisface a $B = B^T$ ya que:

$$B^T = (Q^T A Q)^T = Q^T A^T Q = B$$

4.7 Operaciones con aplicaciones lineales

4.7.1 Espacio vectorial $\mathcal{L}(V, W)$

Para el conjunto $\mathcal{L}(V, W)$ de las aplicaciones lineales del espacio V en W, son operaciones las definidas del siguiente modo:

Suma. Para cualesquiera $f, g \in \mathcal{L}(V, W)$, se define la aplicación $f + g: V \to W$ mediante $(f + g)(\boldsymbol{u}) = f(\boldsymbol{u}) + g(\boldsymbol{u}),\ \forall\, \boldsymbol{u} \in V$.

La aplicación $f + g$ es lineal ya que, para cualesquiera $\boldsymbol{u}, \boldsymbol{v} \in V$ y $\lambda_1, \lambda_2 \in \mathbb{R}$, se verifica:

$$(f + g)(\lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v}) = f(\lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v}) + g(\lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v}) = [\lambda_1 f(\boldsymbol{u}) + \lambda_2 f(\boldsymbol{v})] + [\lambda_1 g(\boldsymbol{u}) + \lambda_2 g(\boldsymbol{v})] =$$

$$= \lambda_1 [f(\boldsymbol{u}) + g(\boldsymbol{u})] + \lambda_2 [f(\boldsymbol{u}) + g(\boldsymbol{v})] = \lambda_1 (f + g)(\boldsymbol{u}) + \lambda_2 (f + g)(\boldsymbol{v})$$

Producto por un escalar. Para cualesquiera que sean el escalar $\lambda \in \mathbb{R}$ y la aplicación lineal $f \in \mathcal{L}(V, W)$, se define la aplicación $\lambda f: V \to W$ mediante $(\lambda f)(\boldsymbol{u}) = \lambda f(\boldsymbol{u}), \forall\, \boldsymbol{u} \in V$.

La aplicación λf es lineal ya que, para cualesquiera $\boldsymbol{u}, \boldsymbol{v} \in V$ y $\lambda_1, \lambda_2 \in \mathbb{R}$, se verifica:

$$(\alpha f)(\lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v}) = \alpha f(\lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v}) = \alpha[\lambda_1 f(\boldsymbol{u}) + \lambda_2 f(\boldsymbol{v})] = \lambda_1 \alpha f(\boldsymbol{u}) + \lambda_2 \alpha f(\boldsymbol{v}) =$$
$$= \lambda_1 (\alpha f)(\boldsymbol{u}) + \lambda_2 (\alpha f)(\boldsymbol{v})$$

Estas operaciones confieren a $\mathcal{L}(V, W)$ de estructura de espacio vectorial, esto es, para cualesquiera $f, g, h \in \mathcal{L}(V, W)$ y $\lambda, \mu \in \mathbb{R}$, se verifica:

$$(f + g) + h = f + (g + h)$$

$$f + o = f$$

$$f + (-f) = o$$

$$f + g = g + f$$

$$\lambda(f + g) = \lambda f + \lambda g$$

$$(\lambda + \mu) f = \lambda f + \mu f$$

$$(\lambda \mu) f = \lambda(\mu f)$$

$$1 f = f$$

donde o es la aplicación nula, La aplicación $-f$, es tal que $(-f)(\boldsymbol{u}) = -f(\boldsymbol{u}),\ \forall\, \boldsymbol{u} \in V$, es lineal.

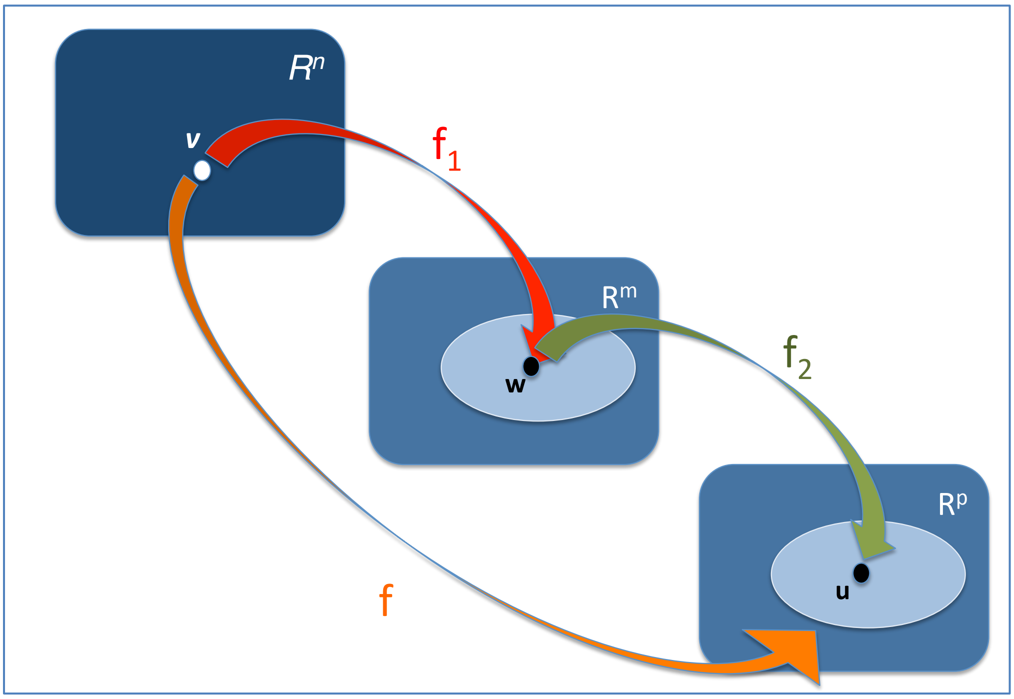
4.7.2 Composición de aplicaciones lineales

La composición f de $f_1: \mathbb{R}^n \to \mathbb{R}^m$ con $f_2: \mathbb{R}^m \to \mathbb{R}^p$ se define como:

$f(\boldsymbol{v}) = f_2(f_1(\boldsymbol{v})),\ \forall\, \boldsymbol{v} \in \mathbb{R}^n$. Esta composición se denota con el símbolo:

$f = f_2 \circ f_1$ y se lee "f_1 *compuesta con* f_2" .

El dominio de f es, por definición, el dominio de f_1. La composición sólo está definida si la imagen de f_1 está contenida en el dominio de f_2:



> ### Teorema.
>
> Sean $f_1: \mathbb{R}^n \to \mathbb{R}^m$ y $f_2: \mathbb{R}^m \to \mathbb{R}^p$ aplicaciones lineales con matrices canónicas A_1 y A_2 respectivamente. La composición $f: \mathbb{R}^n \to \mathbb{R}^p$, definida como $f(\boldsymbol{v}) = f_2(f_1(\boldsymbol{v}))$, $\forall\, \boldsymbol{v} \in \mathbb{R}^n$, es una aplicación lineal. Además, la matriz canónica A de f viene dada por la matriz producto:
>
> $$A = A_2 A_1$$

Demostración

Para probar que f es lineal, sean $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ y λ un escalar. Debido al carácter lineal de f_1 y f_2, se tiene:

$$f(\boldsymbol{u}+\boldsymbol{v}) = f_2\big(f_1(\boldsymbol{u}+\boldsymbol{v})\big) = f_2\big(f_1(\boldsymbol{u}) + f_1(\boldsymbol{v})\big) = f_2\big(f_1(\boldsymbol{u})\big) + f_2\big(f_1(\boldsymbol{v})\big) = f(\boldsymbol{u}) + f(\boldsymbol{v})$$

$$f(\lambda\boldsymbol{u}) = f_2\big(f_1(\lambda\boldsymbol{u})\big) = f_2\big(\lambda f_1(\boldsymbol{u})\big) = \lambda f_2\big(f_1(\boldsymbol{u})\big) = \lambda f(\boldsymbol{u})$$

Para demostrar que A_2A_1 es la matriz canónica de f, usamos la propiedad asociativa del producto de matrices:

$$f(\boldsymbol{v}) = f_2\big(f_1(\boldsymbol{v})\big) = f_2(A_1\boldsymbol{v}) = A_2(A_1\boldsymbol{v}) = (A_2A_1)\boldsymbol{v}$$

Ejemplo

Sean f_1 y f_2 aplicaciones lineales de $\mathbb{R}^3$ en $\mathbb{R}^3$ definidas por:

$f_1(x,y,z) = (2x+y, 0, x+z)$ y $f_2(x,y,z) = (x-y,\ z,\ y)$

Hallar las matrices canónicas de las composiciones $f = f_2 \circ f_1$ y $f' = f_1 \circ f_2$

Solución

Las matrices canónicas de f_1 y f_2 son:

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad A_2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

A consecuencia del teorema anterior se deduce que la matriz canónica de f es:

$$A = A_2 A_1 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Y la de f' es:

$$A' = A_1 A_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

4.7.2.1 Propiedades de la composición de aplicaciones lineales

Dados tres espacios vectoriales $V, W\ y\ U$, toda aplicación lineal $g: V \to W$ se puede componer con cualquier aplicación lineal $f: W \to U$, siendo el resultado $f \circ g: V \to U$ una nueva aplicación lineal.

Siempre que las siguientes composiciones tengan sentido, se verifica que, dados $f, g\ y\ h$ aplicaciones lineales y λ escalar:

1. $(f \circ g) \circ h = f \circ (g \circ h)$
2. $(f + g) \circ h = f \circ h + g \circ h$
3. $h \circ (f + g) = h \circ f + h \circ g$
4. $\lambda(f \circ g) = (\lambda f) \circ g = f \circ (\lambda g)$

Demostración

Ya se ha comprobado que la composición de dos aplicaciones lineales es otra aplicación lineal. Sólo debemos probar, por tanto, las cuatro igualdades del enunciado. Éstas son ciertas ya que, para cualquiera que sea el vector $\boldsymbol{x}$ al que se aplique unas y otras composiciones de funciones, se verifica que (se omite la (3) por analogía con la (2)):

1. $[(f \circ g) \circ h](\boldsymbol{x}) = (f \circ g)\big(h(\boldsymbol{x})\big) = f\Big(g\big(h(\boldsymbol{x})\big)\Big)$
   $[f \circ (g \circ h)](\boldsymbol{x}) = f[(g \circ h)(\boldsymbol{x})] = f\Big(g\big(h(x)\big)\Big)$

2. $[(f + g) \circ h](\boldsymbol{x}) = (f + g)\big(h(\boldsymbol{x})\big) = f\big(h(\boldsymbol{x})\big) + g\big(h(\boldsymbol{x})\big)$
   $[f \circ h + g \circ h](\boldsymbol{x}) = (f \circ h)(\boldsymbol{x}) + (g \circ h)(\boldsymbol{x}) = f\big(h(\boldsymbol{x})\big) + g(h(\boldsymbol{x}))$

3. $[\lambda(f \circ g)](\boldsymbol{x}) = \lambda(f \circ g)(\boldsymbol{x}) = \lambda f\big(g(\boldsymbol{x})\big)$
   $[(\lambda f) \circ g](\boldsymbol{x}) = (\lambda f)\big(g(\boldsymbol{x})\big) = \lambda f(g(\boldsymbol{x}))$

4.7.3 Aplicación lineal inversa

Si $f_1: \mathbb{R}^n \to \mathbb{R}^n$ y $f_2: \mathbb{R}^n \to \mathbb{R}^n$ son dos aplicaciones lineales tales que $\forall\, \boldsymbol{v} \in \mathbb{R}^n$:

$f_2\big(f_1(\boldsymbol{v})\big) = \boldsymbol{v}$ y $f_1\big(f_2(\boldsymbol{v})\big) = \boldsymbol{v}$

se dice que f_1 es **invertible** y que f_2 es la **inversa** de f_1

No toda aplicación lineal es invertible. Ahora bien, si f_1 es invertible, su inversa es única y se denota por ${f_1}^{-1}$.

1. Igual que la inversa de una función de una variable real decimos deshace lo hecho por ella, la inversa de una aplicación lineal f deshace la aplicación efectuada por f.

Teorema.

Sea $f: \mathbb{R}^n \to \mathbb{R}^n$ una aplicación lineal con matriz canónica A. Entonces las siguientes condiciones son equivalentes:

1. f es invertible.
2. f es un isomorfismo.
3. A es invertible

Además. si f es invertible, la matriz canónica de f^{-1} es A^{-1}.

Ejemplo

Probar que la aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^3$ definida por:

$$f(x, y, z) = (2x + 3y + z, 3x + 3y + z, 2x + 4y + z)$$

es invertible y hallar su inversa.

Solución

La matriz canónica de f es:

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

Usando las técnicas de inversión de matrices puede verse que A es invertible y que su inversa es:

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

Por tanto f es invertible y la matriz canónica de f^{-1} es A^{-1}.

5 Valores y vectores propios

5.1 Valores y vectores propios

Sea A una matriz $n\ x\ n$. Se dice que el escalar λ es un **valor propio** (o autovalor) de A si existe un vector $\boldsymbol{x}$, no nulo, tal que:

$$\boldsymbol{Ax} = \boldsymbol{\lambda x}$$

Y se dice que ese vector $\boldsymbol{x}$ es un **vector propio** (o autovector) de A asociado al valor propio λ.

Nota. El vector cero no se admite como vector propio. Si admitiéramos como vector propio al vector cero, la definición carecería de sentido, ya que $A\boldsymbol{o} = \lambda\boldsymbol{o}$ es cierto para todo valor real de λ. Como valor propio $\lambda = 0$ sí es posible.

5.1.1 Subespacios propios

Si λ es un valor propia de A, una matriz de $n\ x\ n$, y $\boldsymbol{x}$ es un vector propio asociado a λ, todo múltiplo escalar no nulo de $\boldsymbol{x}$ es también vector propio de A asociado a λ.

Esto puede probarse llamando c al escalar no nulo, sin más que hacer:

$$A(c\boldsymbol{x}) = c(A\boldsymbol{x}) = c(\lambda\boldsymbol{x}) = \lambda(c\boldsymbol{x})$$

Además, si $\boldsymbol{x}_1, \boldsymbol{x}_2$ son vectores propios asociados a un mismo valor propio λ, su suma es también vector propio asociado a λ, porque:

$$A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = A\boldsymbol{x}_1 + A\boldsymbol{x}_2 = \lambda\boldsymbol{x}_1 + \lambda\boldsymbol{x}_2 = \lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2)$$

En otras palabras, el conjunto de vectores propios de A asociados a un valor propio dado λ, junto con el vector cero, es un subespacio de $\mathbb{R}^n$, denominado **subespacio propio** de λ.

5.1.2 Cálculo de valores y vectores propios

Para calcular los valores y vectores propios de una matriz A, de tamaño $n\ x\ n$, sea I_n la matriz identidad $n\ x\ n$. La ecuación $A\boldsymbol{x} = \lambda\boldsymbol{x}$, escrita en la forma $\lambda I_n\boldsymbol{x} = A\boldsymbol{x}$, produce la ecuación vectorial:

$$(\lambda I - A)\boldsymbol{x} = \boldsymbol{o}$$

Este sistema homogéneo de ecuaciones tiene soluciones no nulas si, y sólo si, la matriz de coeficientes $(\lambda I - A)$ es no invertible, esto es, si y sólo si, el determinante de $(\lambda I - A)$ es cero. Este resultado se enuncia en el siguiente teorema:

<u>Teorema.</u>
Sea A una matriz $n\ x\ n$.

1. Un escalar λ es valor propio de A si, y sólo si

$$\det(\lambda I - A) = 0$$

2. Los valores propios de A asociados a λ son las soluciones no nulas de:

$$(\lambda I - A)\boldsymbol{x} = \boldsymbol{o}$$

La ecuación $\det(\lambda I - A) = 0$ se llama **ecuación característica** de A. Si se desarrolla en forma de polinomio:

$$|\lambda I - A| = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

se llama **polinomio característico** de A.

Esta definición nos dice que los valores propios de una matriz A son las raíces del polinomio característico de A. Si A es $n\ x\ n$, su polinomio característico es de grado n, de lo que se concluye que A tiene, a lo sumo, n valores propios distintos.

Nota. El teorema fundamental del Álgebra afirma que en polinomio de grado n tiene exactamente n raíces. Sin embargo estas n raíces incluyen tanto raíces repetidas como posibles raíces complejas. Nosotros sólo nos ocuparemos de las raíces reales de los polinomios característicos, es decir, de los valores propios reales.

Ejemplo

Hallar los valores propios y los correspondientes vectores propios para la matriz

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{vmatrix} = (\lambda - 2)(\lambda + 5) — 12 = \lambda^2 + 3\lambda - 10 + 12 = \lambda^2 + 3\lambda + 2$$

Igualando el polinomio a cero, obtenemos las dos raíces:

$$\lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2)$$

es decir, los dos valores propios son $\lambda_1 = -1$ y $\lambda_2 = -2$.

Para hallar los vectores propios asociados, usamos eliminación de Gauss-Jordan con el fin de resolver el sistema lineal homogéneo

$$(\lambda I - A)\boldsymbol{x} = \boldsymbol{o}$$

dos veces, primero para $\lambda = \lambda_1 = -1$ y después para $\lambda = \lambda_2 = -2$.

Para $\lambda = \lambda_1 = -1$ la matriz de coeficientes es:

$$(-1)I - A = \begin{bmatrix} -1 - 2 & 12 \\ -1 & -1 + 5 \end{bmatrix} = \begin{bmatrix} -2 & 12 \\ -1 & 4 \end{bmatrix}$$

que se reduce por filas a:

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

Por tanto $x_1 - 4x_2 = 0$. Haciendo $x_2 = t$, concluimos que todo vector propio asociado a $\lambda_1 = -1$ es de la forma:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t\begin{bmatrix} 4 \\ 1 \end{bmatrix}, t \neq 0$$

Para $\lambda = \lambda_2 = -2$, se tiene:

$$(-2)I - A = \begin{bmatrix} -2-2 & 12 \\ -1 & -2+5 \end{bmatrix} = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

Haciendo $x_2 = t$, concluimos que todo vector propio asociado a $\lambda_2 = -2$ es de la forma:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3t \\ t \end{bmatrix} = t \begin{bmatrix} 3 \\ 1 \end{bmatrix}, t \neq 0$$

____

Los sistemas homogéneos que aparecen al buscar vectores propios siempre se reduce, mediante operaciones elementales por filas, a una matriz con al menos una fila de ceros, ya que el sistema debe tener soluciones no triviales.

Procedimiento

Sea A una matriz $n\ x\ n$.

1. Escribir la ecuación característica $\det(\lambda I - A) = 0$, que será una ecuación polinómica de grado n en la variable λ.
2. Hallar las raíces reales de la ecuación característica. Esos son los valores propios de A.
3. Para cada valor propio λ_i, hallar los vectores propios asociados, resolviendo el sistema homogéneo $(\lambda_i I - A)\boldsymbol{x} = \boldsymbol{o}$. Esto exige reducir por filas una matriz $n\ x\ n$. La matriz escalonada resultantes debe tener, al menos, una fila de ceros.

Calcular los valores propios de una matriz suele ser difícil, porque requiere factorizar un polinomio de grado n. Ahora bien, una vez encontrado un valor propio, el cálculo de los vectores propios asociados es un ejercicio sencillo, de aplicación de la reducción de Gauss-Jordan.

Ejemplo

Hallar los valores y vectores propios de:

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

Hallar la dimensión del subespacio propio asociado a cada valor propio.

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{bmatrix} \lambda-2 & 1 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & \lambda-2 \end{bmatrix} = (\lambda-2)^3$$

Por tanto la ecuación característica es $(\lambda - 2)^3 = 0$

Así pues, el único valor propio es $\lambda = 2$. Para hallar los vectores propios asociados, resolvemos el sistema homogéneo $(2I - A)\boldsymbol{x} = \boldsymbol{o}$.

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

Esto implica que $x_2 = 0$. Usando los parámetros $x_1 = s$ y $x_3 = t$, vemos que los vectores propios son de la forma:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ 0 \\ t \end{bmatrix} = \boldsymbol{s} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \;\; no\ simultáneamente\ nulos$$

Como $\lambda = 2$ tiene dos vectores propios linealmente independientes, la dimensión de su subespacio propio es 2.

____

Si un valor propio λ_1 es raíz múltiple (k veces) del polinomio característico, se dice que λ_1 tiene multiplicidad k. Eso significa que $(\lambda - \lambda_1)^k$ es un factor del polinomio característico, pero que $(\lambda - \lambda_1)^{k+1}$ no lo es.

Siempre ocurre que la multiplicidad de un valor propio es mayor o igual que la dimensión de su subespacio propio.

Ejemplo

Hallar los valores propios de:

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

y una base de cada uno de los subespacios propios.

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{bmatrix} \lambda - 1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & -5 & 10 \\ -1 & 0 & \lambda - 2 & 0 \\ -1 & 0 & 0 & \lambda - 3 \end{bmatrix} = (\lambda - 1)^2(\lambda - 2)(\lambda - 3)$$

Por tanto, la ecuación característica es $(\lambda - 1)^2(\lambda - 2)(\lambda - 3) = 0$, luego A tiene como valores propios $\lambda_1 = 1, \lambda_2 = 2$ y $\lambda_3 = 3$, siendo λ_1 de multiplicidad 2.

Podemos hallar una base del subespacio propio asociado a $\lambda_1 = 1$ como sigue:

$$1I - A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 10 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Haciendo $x_2 = s$ y $x_4 = t$, obtenemos:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2t \\ s \\ 2t \\ t \end{bmatrix} = \boldsymbol{s} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

Por consiguiente, una base del subespacio propio de $\lambda_1 = 1$ viene dada por:

$$B_1 = \{(0,1,0,0),(-2,0,2,1)\}$$

Para los subespacios propios de $\lambda_2 = 2$ y $\lambda_3 = 3$, se obtienen, de manera análoga, las bases:

$$B_2 = \{(0,5,1,0)\}$$

$$B_3 = \{(0,-5,0,1)\}$$

Teorema.

Si A es una matriz $n \; x \; n$ triangular, sus valores propios son los elementos de la diagonal principal.

5.1.3 Valores y vectores propios de aplicaciones lineales

Se pueden definir también valores y vectores propios para aplicaciones lineales. Se dice que un número λ es valor propio de la aplicación lineal $f: V \longrightarrow V$ si existe un vector $\boldsymbol{x}$ no nulo tal que $f(\boldsymbol{x}) = \lambda\boldsymbol{x}$. En tal caso, el vector $\boldsymbol{x}$ se dice que es un vector propio de f asociado a λ. El conjunto de todos los vectores propios asociados a λ, junto con el vector cero, forman subespacio propio de λ.

5.1.4 Invarianza de polinomio característico

Teorema

Si A y B son matrices $n \; x \; n$ semejantes, tienen los mismos valores propios.

Demostración

Puesto que A y B son semejantes, existe una matriz invertible P tal que $B = P^{-1}AP$. Por las propiedades de los determinantes:

$$|\lambda I - B| = |\lambda I - P^{-1}AP| = |P^{-1}\lambda IP - P^{-1}AP| = |P^{-1}(\lambda I - A)P| = |P^{-1}||\lambda I - A||P| =$$

$$= |P^{-1}||P||\lambda I - A| = |P^{-1}P||\lambda I - A| = |\lambda I - A|$$

Al tener A y B el mismo polinomio característico, tienen los mismos valores propios.

5.2 Diagonalización

Se dice que una matriz A de tamaño $n \; x \; n$ es **diagonalizable** si es semejante a una matriz diagonal. Es decir, A es diagonalizable si existe una matriz invertible P tal que:

$$P^{-1}AP$$

es una matriz diagonal.

Teorema. Criterio de diagonalización

A es una matriz $n\ x\ n$ diagonalizable si, y sólo si, tiene n vectores propios linealmente independientes.

Demostración

Supongamos que A es diagonalizable. En tal caso, existe una matriz invertible P tal que $D = P^{-1}AP$, es diagonal. Denotando los elementos diagonales de D por $\lambda_1, \lambda_2, \dots, \lambda_n$ y los vectores columna de P por $\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_n$, se tiene:

$$[PD] = [\boldsymbol{p}_1 \vdots \boldsymbol{p}_2 \vdots \cdots \vdots \boldsymbol{p}_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\lambda_1\boldsymbol{p}_1 \vdots \lambda_2\boldsymbol{p}_2 \vdots \cdots \vdots \lambda_n\boldsymbol{p}_n]$$

De $P^{-1}AP = D$ se sigue $AP = PD$, lo cual implica que:

$$[A\boldsymbol{p}_1 \vdots A\boldsymbol{p}_2 \vdots \cdots \vdots A\boldsymbol{p}_n] = [\lambda_1\boldsymbol{p}_1 \vdots \lambda_2\boldsymbol{p}_2 \vdots \cdots \vdots \lambda_n\boldsymbol{p}_n]$$

en otras palabras, $A\boldsymbol{p}_i = \lambda_i\boldsymbol{p}_i$ para cada vector columna $\boldsymbol{p}_i$. Así pues, los vectores columna $\boldsymbol{p}_i$ son vectores propios de A. Además, al ser P invertible, sus vectores columna son linealmente independientes. Por tanto A tiene n vectores propios linealmente independientes.

Recíprocamente, supongamos que A tiene n vectores propios linealmente independientes $\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_n$ con valores propios correspondientes $\lambda_1, \lambda_2, \dots, \lambda_n$. Sea P la matriz cuyas columnas son esos n vectores propios, es decir, $P = [\boldsymbol{p}_1 \vdots \boldsymbol{p}_2 \vdots \cdots \vdots \boldsymbol{p}_n]$. Como cada vector $\boldsymbol{p}_i$ es vector propio de A, se tiene $A\boldsymbol{p}_i = \lambda_i\boldsymbol{p}_i$, así que:

$$AP = [\boldsymbol{p}_1 \vdots \boldsymbol{p}_2 \vdots \cdots \vdots \boldsymbol{p}_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = PD$$

Finalmente, puesto que los vectores $\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_n$, son linealmente independientes, P es invertible, y podemos escribir la ecuación $AP = PD$ como $P^{-1}AP = D$, lo que significa que A es diagonalizable.

Procedimiento para la diagonalización de una matriz $n\ x\ n$

Sea A una matriz $n\ x\ n$.

1. Hallar n vectores propios linealmente independientes $\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_n$ de A asociados a los valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$. Si no existen n vectores propios linealmente independientes, A no es diagonalizable.
2. Si A tiene n vectores propios linealmente independientes, sea P la matriz $n\ x\ n$ formada por esos vectores propios como columnas, o sea:

$$P = [\boldsymbol{p}_1 \vdots \boldsymbol{p}_2 \vdots \cdots \vdots \boldsymbol{p}_n]$$

3. La matriz diagonal $D = P^{-1}AP$ tendrá los valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$en su diagonal principal (y ceros fuera de esa diagonal). El orden de los vectores propios determina el orden en que aparecen los valores propios en la diagonal de D.

Ejemplo

Probar que la matriz:

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

es diagonalizable y hallar una matriz P tal que $P^{-1}AP$ sea diagonal.

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{bmatrix} \lambda - 1 & -1 & -1 \\ 1 & \lambda - 3 & 1 \\ -3 & 1 & \lambda + 1 \end{bmatrix} = (\lambda - 2)(\lambda + 2)(\lambda - 3)$$

luego los valores propios de A son $\lambda_1 = 2, \lambda_2 = -2$, y $\lambda_3 = 3$. Usando estos valores propios, se obtienen, para cada caso, la forma escalonada reducida y los vectores propios asociados:

$$2I - A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 3 & -1 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \longrightarrow \quad vector\ propio: \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$-2I - A = \begin{bmatrix} -3 & 1 & 1 \\ -1 & -5 & -1 \\ 3 & -1 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{bmatrix} \quad \longrightarrow \quad vector\ propio: \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

$$3I - A = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 3 & -1 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad \longrightarrow \quad vector\ propio: \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Formamos la matriz P que tiene por columnas los vectores propios obtenidos:

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

Esta matriz es no singular, de modo que los vectores propios son linealmente independientes y A es diagonalizable. La inversa de P es:

$$P^{-1} = \begin{bmatrix} -1 & -1 & 0 \\ 1/5 & 0 & 1/5 \\ 1/5 & 1 & 1/5 \end{bmatrix}$$

Luego:

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Teorema. Condición suficiente para la diagonalización

Si A es una matriz $n\ x\ n$ con n valores propios *distintos*, los vectores propios asociados son linealmente independientes y A es diagonalizable

Demostración

Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los n valores propios distintos de A y $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ los vectores propios asociados. Para empezar supongamos que estos vectores propios fueran linealmente dependientes. Además tomemos los vectores propios ordenados de manera que los m primeros sean linealmente independientes , pero los $m+1$ primeros sean linealmente dependientes, donde $m < n$.

Entonces $\boldsymbol{x}_{m+1}$ se puede escribir como combinación lineal de los primeros m vectores propios:

$$\boldsymbol{x}_{m+1} = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \cdots + c_m\boldsymbol{x}_m \quad (1)$$

donde no todos los c_i son cero. Aplicando a ambos lados de la igualdad la matriz A resulta:

$$A\boldsymbol{x}_{m+1} = Ac_1\boldsymbol{x}_1 + Ac_2\boldsymbol{x}_2 + \cdots + Ac_m\boldsymbol{x}_m \Longrightarrow \lambda_{m+1}\boldsymbol{x}_{m+1} = \lambda_1 c_1\boldsymbol{x}_1 + \lambda_2 c_2\boldsymbol{x}_2 + \cdots + \lambda_m c_m\boldsymbol{x}_m \quad (2)$$

Mientras que multiplicando la ecuación (1) por λ_{m+1}, se obtiene:

$$\lambda_{m+1}\boldsymbol{x}_{m+1} = c_1\lambda_{m+1}\boldsymbol{x}_1 + c_2\lambda_{m+1}\boldsymbol{x}_2 + \cdots + c_m\lambda_{m+1}\boldsymbol{x}_m \quad (3)$$

restando (2)-(3), obtenemos:

$$c_1(\lambda_{m+1} - \lambda_1)\boldsymbol{x}_1 + c_2(\lambda_{m+1} - \lambda_2)\boldsymbol{x}_2 + \cdots + c_m(\lambda_{m+1} - \lambda_m)\boldsymbol{x}_m = \boldsymbol{o}$$

que al ser linealmente independientes los m primeros vectores propios permite concluir que todos los coeficientes de esta ecuación son cero, esto es:

$$c_1(\lambda_{m+1} - \lambda_1) = c_2(\lambda_{m+1} - \lambda_2) = \cdots = c_m(\lambda_{m+1} - \lambda_m) = 0$$

Como todos los valores propios son distintos, se sigue que $c_i = 0, i = 1, 2, \ldots, m$. Pero este resultado contradice nuestra hipótesis de que $\boldsymbol{x}_{m+1}$ se puede escribir como combinación lineal de los m primeros vectores propios. Por tanto, el conjunto de los vectores propios es linealmente independiente y, por el teorema anterior, A es diagonalizable.

Nota. La condición del teorema anterior es suficiente, pero no necesaria, para que la matriz sea diagonalizable. En otras palabras, los valores propios de una matriz diagonalizable no tienen por qué ser todos distintos.

5.2.1 Diagonalización y aplicaciones lineales

En términos de aplicaciones lineales, el problema de diagonalización se enuncia como sigue.

Dada una aplicación lineal $f: V \to V$ ¿existe una base B de V en la que la matriz de f sea diagonal?

La respuesta es sí, siempre que la matriz canónica de f sea diagonalizable.

Ejemplo

Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ la aplicación lineal definida por:

$$f(x_1, x_2, x_3) = (x_1 - x_2 - x_3, x_1 + 3x_2 + x_3, -3x_1 + x_2 - x_3)$$

Si es posible, hallar una base B de $\mathbb{R}^3$ tal que la matriz de f respecto de esa base sea diagonal.

Solución

La matriz canónica de f es:

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

En el ejemplo anterior hemos visto que A es diagonalizable. Así pues, podemos tomar como base B la formada por los tres vectores propios encontrados, esto es:

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

La matriz de f en esta base es:

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

Algunas propiedades a tener en cuenta:

1. La traza de una matriz simétrica es un invariante, por lo tanto, una vez calculados los autovalores podemos comprobar que su suma es igual a la traza de la matriz.
2. Si todos los autovalores obtenidos tienen multiplicidad uno, entonces el endomorfismo es diagonalizable.
3. Para evitar calcular inversas una vez obtenidas D y P en lugar de calcular $P^{-1}AP = D$ podemos calcular $PD = AP$.
4. Una matriz y su traspuesta tienen los mismos autovalores.
5. Si λ es autovalor de una matriz A, $K\lambda$ será autovalor de la matriz KA
6. Si λ es autovalor de una matriz A, y k un escalar cualquiera, entonces $(\lambda - k)$ será un autovalor de la matriz $(A - kI)$.
7. Si λ es una autovalor de la matriz A, entonces $1/\lambda$ es un autovalor de la matriz A^{-1}
8. $A^n = P^{-1}D^nP$, es decir, si A es diagonalizable, es fácil obtener cualquier potencia de A

5.3 Espacios con producto escalar

5.3.1 Definiciones

A continuación se relacionan una serie de definiciones relativas al espacio vectorial $\mathbb{R}^n$ dotado de un producto escalar, con el fin de aclarar algunos conceptos tratados en el apartado 5.4. El lector puede consultar cualquier libro de Álgebra básica para obtener las demostraciones de las proposiciones presentadas así como profundizar en las propiedades de los espacios vectoriales euclídeos (con producto escalar).

Si $\boldsymbol{v} = (v_1, v_2, \dots, v_n)$ es un vector de $\mathbb{R}^n$:

Su <u>longitud</u>, denotada por $\|\boldsymbol{v}\|$, se define como:

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

El vector $\boldsymbol{u} = \frac{v}{\|v\|}$ tiene longitud 1 y la misma dirección de $\boldsymbol{v}$. Este vector $\boldsymbol{u}$ se llama <u>vector unitario en la dirección de v.</u>

La <u>distancia</u> entre dos vectores $\boldsymbol{u}$ y $\boldsymbol{v}$ de $\mathbb{R}^n$ es: $d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|$

El producto escalar de $\boldsymbol{u} = (u_1, u_2, \dots, u_n)$ y $\boldsymbol{v} = (v_1, v_2, \dots, v_n)$ es la cantidad escalar:

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 \cdot v_1 + u_2 \cdot v_2 + \cdots + u_n \cdot v_n$$

Dos vectores $\boldsymbol{u}$ y $\boldsymbol{v}$ de $\mathbb{R}^n$ son ortogonales si $\boldsymbol{u} \cdot \boldsymbol{v} = 0$

5.3.1 Conjuntos ortogonales y ortonormales

Definición

Un conjunto S de vectores de un espacio V con un producto escalar es ortogonal si todo par de vectores en S es ortogonal. Si además, cada uno de los vectores de S es unitario, se dice que S es ortonormal.

Los conjuntos ortogonales son linealmente independientes

Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ un conjunto de vectores en un espacio con producto escalar. Si S es ortogonal, entonces es linealmente independiente.

Si V es un espacio vectorial con producto escalar de dimensión n, cualquier conjunto ortogonal de n vectores no nulos, es una base de V.

5.3.2 Método de normalización de Gram-Schmidt

Sea $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ una base de un espacio V con producto escalar.

Sea $B' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots., \boldsymbol{w}_n\}$, donde los $\boldsymbol{w}_i$ vienen dados por:

$$\boldsymbol{w}_1 = \boldsymbol{v}_1$$

$$\boldsymbol{w}_2 = \boldsymbol{v}_2 - \frac{\boldsymbol{v}_2 \cdot \boldsymbol{w}_1}{\boldsymbol{w}_1 \cdot \boldsymbol{w}_1}\boldsymbol{w}_1$$

$$\boldsymbol{w}_3 = \boldsymbol{v}_3 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{w}_1}{\boldsymbol{w}_1 \cdot \boldsymbol{w}_1}\boldsymbol{w}_1 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{w}_2}{\boldsymbol{w}_2 \cdot \boldsymbol{w}_2}\boldsymbol{w}_2$$

$$\vdots$$

$$\boldsymbol{w}_n = \boldsymbol{v}_n - \frac{\boldsymbol{v}_n \cdot \boldsymbol{w}_1}{\boldsymbol{w}_1 \cdot \boldsymbol{w}_1}\boldsymbol{w}_1 - \frac{\boldsymbol{v}_n \cdot \boldsymbol{w}_2}{\boldsymbol{w}_2 \cdot \boldsymbol{w}_2}\boldsymbol{w}_2 - \cdots - \frac{\boldsymbol{v}_n \cdot \boldsymbol{w}_{n-1}}{\boldsymbol{w}_{n-1} \cdot \boldsymbol{w}_{n-1}}\boldsymbol{w}_{n-1}$$

Entonces B' es una base **ortogonal** de V.

Llamando $\boldsymbol{u}_i = \frac{\boldsymbol{w}_i}{\|\boldsymbol{w}_i\|}$, el conjunto $B'' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n\}$ es una base **ortonormal** de V.

Ejemplo

Aplicar el método de ortonormalización de Gram-Schmidt a la siguiente base de $\mathbb{R}^2$:

$$B = \{(1,1), (0,1)\}$$

Solución

Tomamos:

$$\boldsymbol{w}_1 = \boldsymbol{v}_1 = (1,1)$$

$$\boldsymbol{w}_2 = \boldsymbol{v}_2 - \frac{\boldsymbol{v}_2 \cdot \boldsymbol{w}_1}{\boldsymbol{w}_1 \cdot \boldsymbol{w}_1}\boldsymbol{w}_1 = (0,1) - \frac{1}{2}(1,1) = \left(-\frac{1}{2}, \frac{1}{2}\right)$$

El conjunto $B' = \{\boldsymbol{w}_1, \boldsymbol{w}_2\}$ es una base ortogonal de $\mathbb{R}^2$. Normalizando cada vector de B' obtenemos:

$$\boldsymbol{u}_1 = \frac{\boldsymbol{w}_1}{\|\boldsymbol{w}_1\|} = \frac{1}{\sqrt{2}}(1,1) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

$$\boldsymbol{u}_2 = \frac{\boldsymbol{w}_2}{\|\boldsymbol{w}_2\|} = \frac{1}{{}^{1}\!/\!_{\sqrt{2}}}\left(-\frac{1}{2}, \frac{1}{2}\right) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

Por tanto, $B'' = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ es una base ortonormal de $\mathbb{R}^2$.

5.4 Matrices simétricas y diagonalización ortogonal

5.4.1 Matrices simétricas

Recordemos que una matriz cuadrada A es simétrica si es igual a su traspuesta: $A = A^T$.

Teorema.

Sea A una matriz $n \times n$ simétrica. Entonces:

1. A es diagonalizable
2. Todos los valores propios de A son reales
3. Si λ es un valor propio de A de multiplicidad k, hay k vectores propios de A asociados a λ y linealmente independientes. En otras palabras, el subespacio propio de λ tiene dimensión k.

Este teorema se conoce como **teorema espectral real**, y el conjunto de todos los valores propios de A se llama el **espectro** de A.

Ejemplo

Hallar los valores propios de la matriz simétrica:

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

y determinar las dimensiones de los subespacios propios correspondientes.

Solución

El polinomio característico de A viene dado por:

$$|\lambda I - A| = \begin{bmatrix} \lambda - 1 & -2 & 0 & 0 \\ -2 & \lambda - 1 & 0 & 0 \\ 0 & 0 & \lambda - 1 & -2 \\ 0 & 0 & -2 & \lambda - 1 \end{bmatrix} = (\lambda + 1)^2 + (\lambda - 3)^2$$

Así pues, los valores propios de A son $\lambda_1 = -1$ y $\lambda_2 = 3$. como ambos tienen multiplicidad 2, sabemos que sus correspondientes subespacios propios tienen dimensión 2.

En concreto, el subespacio propio de $\lambda_1 = -1$ admite como base $B_1 = \{(1,1,0,0),(0,0,1,1)\}$ y el subespacio propio de $\lambda_2 = 3$ admite como base $B_2 = \{(1,-1,0,0),(0,0,1,-1)\}$.

5.4.2 Matrices ortogonales

Una matriz cuadrada P se dice que es ortogonal si, y sólo si, es invertible y $P^{-1} = P^T$.

Teorema. Propiedad de las matrices ortogonales

Una matriz cuadrada P es ortogonal si, y sólo si, sus vectores columna forman un conjunto ortonormal

Demostración

Supongamos que los vectores columna de P forman un conjunto ortonormal:

$$P = [\boldsymbol{p}_1 \vdots \boldsymbol{p}_2 \vdots \cdots \vdots \boldsymbol{p}_n] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

Entonces el producto $P^T P$ tiene la forma:

$$P^T P = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{n1} \\ p_{12} & p_{22} & \cdots & p_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} =$$

$$= \begin{bmatrix} \boldsymbol{p}_1 \cdot \boldsymbol{p}_1 & \boldsymbol{p}_1 \cdot \boldsymbol{p}_2 & \cdots & \boldsymbol{p}_1 \cdot \boldsymbol{p}_n \\ \boldsymbol{p}_2 \cdot \boldsymbol{p}_1 & \boldsymbol{p}_2 \cdot \boldsymbol{p}_2 & \cdots & \boldsymbol{p}_2 \cdot \boldsymbol{p}_n \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{p}_n \cdot \boldsymbol{p}_1 & \boldsymbol{p}_n \cdot \boldsymbol{p}_2 & \cdots & \boldsymbol{p}_n \cdot \boldsymbol{p}_n \end{bmatrix} \quad (1)$$

Como el conjunto $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_n\}$ es ortonormal, se tiene:

$$\boldsymbol{p}_i \cdot \boldsymbol{p}_j = 0, i \neq j \quad \text{y} \quad \boldsymbol{p}_i \cdot \boldsymbol{p}_i = \|\boldsymbol{p}_i\|^2 = 1$$

Por lo tanto la matriz compuesta por los productos escalares (1) es:

$$P^T P = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_n$$

Esto implica que $P^T = P^{-1}$, luego P es ortogonal.

Recíprocamente, si P es ortogonal, podemos invertir los pasos del razonamiento anterior para demostrar que los vectores columna de P forman un conjunto ortogonal.

Ejemplo

Probar que la matriz:

$$P = \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ {}^{-2}/_{\sqrt{5}} & {}^{1}/_{\sqrt{5}} & 0 \\ {}^{-2}/_{3\sqrt{5}} & {}^{-4}/_{3\sqrt{5}} & {}^{5}/_{3\sqrt{5}} \end{bmatrix}$$

es ortogonal verificando que $P^T P = I$. a continuación, demostrar que los vectores columna de P forman un conjunto ortonormal.

Solución

$$PP^T = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ {}^{-2}/_{\sqrt{5}} & {}^{1}/_{\sqrt{5}} & 0 \\ {}^{-2}/_{3\sqrt{5}} & {}^{-4}/_{3\sqrt{5}} & {}^{5}/_{3\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & {}^{-2}/_{\sqrt{5}} & {}^{-2}/_{3\sqrt{5}} \\ \frac{2}{3} & {}^{1}/_{\sqrt{5}} & {}^{-4}/_{3\sqrt{5}} \\ \frac{2}{3} & 0 & {}^{5}/_{3\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

Se sigue que $P^T = P^{-1}$, luego P es ortogonal.

Además, denotando:

$$\boldsymbol{p}_1 = \begin{bmatrix} 1/3 \\ {}^{-2}/_{\sqrt{5}} \\ {}^{-2}/_{3\sqrt{5}} \end{bmatrix}, \quad \boldsymbol{p}_2 = \begin{bmatrix} 2/3 \\ {}^{1}/_{\sqrt{5}} \\ {}^{-4}/_{3\sqrt{5}} \end{bmatrix} \text{ , y } \boldsymbol{p}_3 = \begin{bmatrix} 2/3 \\ 0 \\ {}^{5}/_{3\sqrt{5}} \end{bmatrix}$$

se tiene:

$$\boldsymbol{p}_1 \cdot \boldsymbol{p}_2 = \boldsymbol{p}_1 \cdot \boldsymbol{p}_3 = \boldsymbol{p}_2 \cdot \boldsymbol{p}_3 = 0$$

y

$\|\boldsymbol{p}_1\| = \|\boldsymbol{p}_2\| = \|\boldsymbol{p}_3\| = 1$

En consecuencia $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$ es un conjunto ortonormal, tal como afirma el teorema anterior.

5.4.3 Propiedad de las matrices simétricas

Sea A una matriz simétrica de tamaño $n\ x\ n$. Si λ_1 y λ_2 son autovalores distintos de A, entonces sus autovectores correspondientes $\boldsymbol{x}_1$ y $\boldsymbol{x}_2$ son ortogonales.

Demostración

Sean λ_1 y λ_2 dos valores propios distintos de A, con vectores propios asociados $\boldsymbol{x}_1$ y $\boldsymbol{x}_2$. Eso significa que:

$A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$ y $A\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2$

Para demostrar el teorema usaremos la siguiente forma matricial del producto escalar:

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \end{bmatrix} \cdot \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix} = X_1^T X_2$$

Ahora podemos escribir:

$$\lambda_1(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = (\lambda_1\boldsymbol{x}_1) \cdot \boldsymbol{x}_2 = (A\boldsymbol{x}_1) \cdot \boldsymbol{x}_2 = (A\boldsymbol{x}_1)^{\boldsymbol{T}} X_2 = (X_1^T A^T)X_2 = (X_1^T A)X_2 = X_1^T(AX_2) = X_1^T(\lambda_2 X_2) =$$

$$= \boldsymbol{x}_1 \cdot (\lambda_2\boldsymbol{x}_2) = \lambda_2(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2)$$

Eso implica que $(\lambda_1 - \lambda_2)(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = 0$**.** Y como $\lambda_1 \neq \lambda_2$, concluimos que $\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = 0$. Por tanto $\boldsymbol{x}_1$ y $\boldsymbol{x}_2$ son ortogonales.

Ejemplo

Probar que dos vectores propios cualesquiera de la matriz:

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

asociados a dos valores propios distintos son ortogonales.

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{bmatrix} \lambda - 3 & -1 \\ -1 & \lambda - 3 \end{bmatrix} = (\lambda - 2)(\lambda - 4)$$

luego los valores propios de A son $\lambda_1 = 2$ y $\lambda_2 = 4$.

Todo vector propio asociado a $\lambda_1 = 2$ es de la forma:

$$\boldsymbol{x}_1 = \begin{bmatrix} s \\ -s \end{bmatrix}, \quad s \neq 0$$

y todo vector propio asociado a $\lambda_2 = 4$ es de la forma:

$$\boldsymbol{x}_2 = \begin{bmatrix} t \\ t \end{bmatrix}, \quad t \neq 0$$

Por tanto:

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = \begin{bmatrix} s \\ -s \end{bmatrix} \begin{bmatrix} t \\ t \end{bmatrix} = st - st = 0$$

de donde se deduce que $\boldsymbol{x}_1$ y $\boldsymbol{x}_2$ son ortogonales.

5.4.4 Diagonalización ortogonal

Una matriz A es **ortogonalmente diagonalizable** si existe una matriz ortogonal P tal que $P^{-1}AP = D$ es diagonal. Las matrices con esa propiedad son precisamente las simétricas, tal como establece le siguiente teorema.

Teorema fundamental de las matrices siimétricas.

Una matriz $n\ x\ n$ es ortogonalmente diagonalizable si, y sólo si, es simétrica.

Demostración

a) La demostración en una dirección es sencilla. Si A es ortogonalmente diagonalizable, existe una matriz ortogonal P tal que $P^{-1}AP = D$ es diagonal. Y al ser $P^{-1} = P^T$, se tiene:

$$A = PDP^{-1} = PDP^T$$

lo cual implica que:

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

por tanto A es simétrica.

b) Supongamos que A es simétrica y tiene una valor propio λ de multiplicidad k, debe tener k vectores linealmente independientes (5.4.1) Usando el método de Gram-Schmidt podemos formar con ellos una base ortonormal del subespacio propio de λ. Repetimos esta proceso para cada uno de los valores propios de A. La colección resultante de vectores propios es ortogonal (5.4.3) y, por tanto, ortonormal por construcción.

Sea ahora P la matriz que tiene por columnas esos n vectores propios ortonormales. Por 5.4.3 P es una matriz ortogonal. Finalmente, por el criterio de diagonalización, concluimos que $P^{-1}AP$ es diagonal. Por tanto A es ortogonalmente diagonalizable.

5.4.5 Diagonalización ortogonal de matrices simétricas

Sea A una matriz $n \times n$ simétrica. El proceso de diagonalización ortogonal de A es el siguiente:

1. Calcular sus valores propios y la multiplicidad de cada uno de ellos.
2. Para cada valor propio de multiplicidad 1, tomar un vector propio unitario (Hallar cualquier vector propio y normalizarlo).
3. Para cada valor propio de multiplicidad $k \geq 2$, hallar k vectores propios linealmente independientes (siempre es posible). Si este conjunto no es ortonormal, aplicarle el método de Gram-Schmidt.
4. Los pasos 2 y 3 producen finalmente un conjunto ortonormal de n vectores propios. Formar una matriz P que tenga por columnas a esos n vectores propios. La matriz $P^{-1}AP = P^TAP = D$ será diagonal. (Los elementos de la diagonal principal de D serán los valores propios de A.

Ejemplo

Hallar la matriz ortogonal P que diagonalice:

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

Solución

1. El polinomio característico de A es:

$$|\lambda I - A| = (\lambda - 3)^2(\lambda + 6)$$

da como valores propios $\lambda_1 = -6$ con multiplicidad 1, y $\lambda_2 = 3$ con multiplicidad 2.

2. Un vector propio asociado a λ_1 es $\boldsymbol{v}_1 = (1, -2, 2)$ que, normalizado, se convierte en:

$$\boldsymbol{u}_1 = \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$$

3. Dos vectores propios asociados a λ_2 son $\boldsymbol{v}_2 = (2, 1, 0)$ y $\boldsymbol{v}_3 = (-2, 0, 1)$

Nótese que $\boldsymbol{v}_1$ es ortogonal a $\boldsymbol{v}_2$ y $\boldsymbol{v}_3$. Sin embargo $\boldsymbol{v}_2$ y $\boldsymbol{v}_3$ no son ortogonales entre sí. Para encontrar dos vectores propios ortonormalizados asociados a λ_2 usamos el método de Gram-Schmidt:

$$\boldsymbol{w}_2 = \boldsymbol{v}_2 = (2,1,0)$$

$$\boldsymbol{w}_3 = \boldsymbol{v}_3 - \left(\frac{\boldsymbol{v}_3 \cdot \boldsymbol{w}_2}{\boldsymbol{w}_2 \cdot \boldsymbol{w}_2}\right)\boldsymbol{w}_2 = \left(-\frac{2}{5}, \ \frac{4}{5}, 1\right)$$

Estos vectores normalizados, pasan a ser:

$$\boldsymbol{u}_2 = \frac{\boldsymbol{w}_2}{\|\boldsymbol{w}_2\|} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0\right)$$

$$\boldsymbol{u}_3 = \frac{\boldsymbol{w}_3}{\|\boldsymbol{w}_3\|} = \left(\frac{-2}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}\right)$$

4. La matriz P tiene por columnas los vectores $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$:

$$P = \begin{bmatrix} \frac{1}{3} & {}^{2}/_{\sqrt{5}} & {}^{-2}/_{3\sqrt{5}} \\ \frac{-2}{3} & {}^{1}/_{\sqrt{5}} & {}^{4}/_{3\sqrt{5}} \\ \frac{2}{3} & 0 & {}^{5}/_{3\sqrt{5}} \end{bmatrix}$$

Es fácil comprobar que $P^{-1}AP = P^{T}AP = \begin{bmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

5.5 Factorización QR

Si A es una matriz de $m\ x\ n$ con columnas linealmente independientes (lo que requiere que $m \geq\ n$), entonces, la aplicación del proceso de Gram-Schmidt a dichas columnas produce una factorización de A muy útil que consiste en el producto de una matriz *Q* con columnas ortonormales y una matriz triangular superior *R* denominada **factorización QR**.

Para ver cómo surge la factorización *QR*, sean $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ las columnas (linealmente independientes) de A las cuales forman una base del espacio de columnas de A , mediante el proceso Gram-Schmidt podemos obtener una base ortonormal $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ los vectores ortonormales para el espacio generado por las columnas de A . Recordemos cómo se obtiene esta base.

Primero construimos una base ortogonal $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ como sigue:

$$\boldsymbol{v}_1 = \boldsymbol{u}_1$$

y luego para $i = 2, 3, \dots, n$ tenemos:

$$\boldsymbol{v}_i = \boldsymbol{u}_i - \frac{\boldsymbol{u}_i \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_i \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 - \cdots - \frac{\boldsymbol{u}_i \cdot \boldsymbol{v}_{n-1}}{\boldsymbol{v}_{i-1} \cdot \boldsymbol{v}_{i-1}}\boldsymbol{v}_{i-1} \qquad (1)$$

Por último:

$\boldsymbol{w}_i = \frac{\boldsymbol{v}_i}{\|\boldsymbol{v}_i\|}$ para $i = 1, 2, 3, \ldots, n$. Ahora cada uno de los vectores $\boldsymbol{u}_i$ se escribe como una combinación lineal de los vectores $\boldsymbol{w}$:

$$\begin{aligned} \boldsymbol{u}_1 &= r_{11}\boldsymbol{w}_1 + r_{21}\boldsymbol{w}_2 + \cdots + r_{n1}\boldsymbol{w}_n \\ \boldsymbol{u}_2 &= r_{12}\boldsymbol{w}_1 + r_{22}\boldsymbol{w}_2 + \cdots + r_{n2}\boldsymbol{w}_n \\ &\vdots \\ \boldsymbol{u}_n &= r_{1n}\boldsymbol{w}_1 + r_{2n}\boldsymbol{w}_2 + \cdots + r_{nn}\boldsymbol{w}_n \end{aligned} \qquad (2)$$

Donde $r_{ji} = \boldsymbol{u}_i \cdot \boldsymbol{w}_j$

Además se observa a partir de (1) que $\boldsymbol{u}_i$ están en

$$gen(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = gen(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$$

Como $\boldsymbol{w}_j$ es ortogonal a $gen(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ para $j > i$, es ortogonal a $\boldsymbol{u}_i$. Por lo tanto $r_{ji} = 0$ para $j > i$.

Sea Q la matriz cuyas columnas son $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_j$. Sea:

$$\boldsymbol{r}_j = \begin{bmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{nj} \end{bmatrix}$$

Entonces, las ecuaciones en (2) pueden escribirse en forma matricial como:

$$A = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \ \cdots \boldsymbol{u}_n] = [Q\boldsymbol{r}_1 \quad Q\boldsymbol{r}_2 \ \cdots Q\boldsymbol{r}_n] = QR$$

Donde R es la matriz cuyas columnas son $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n$. En consecuencia:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

A continuación demostremos que R es no singular.

Sea $\boldsymbol{x}$ una solución del sistema lineal $R\boldsymbol{x} = \boldsymbol{0}$. Al multiplicar esta ecuación por Q a la izquierda, tenemos:

$$Q(R\boldsymbol{x}) = (QR)\boldsymbol{x} = A\boldsymbol{x} = Q\boldsymbol{0} = \boldsymbol{0}$$

Como sabemos el sistema homogéneo $A\boldsymbol{x} = \boldsymbol{0}$ puede escribirse como:

$$x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \cdots x_n\boldsymbol{u}_n = \boldsymbol{0}$$

donde $x_1, x_2, \ldots, x_n$ son las coordenadas del vector $\boldsymbol{x}$. Como las columnas de A son linealmente independientes,

$$x_1 = x_2 = \cdots = x_n = 0$$

de modo que $\boldsymbol{x}$ debe ser el vector nulo, por lo que la matriz R es no singular.

Claramente, la matriz Q tiene columnas ortonormales. También los elementos de la diagonal principal de R son todos distintos de cero. Para ver esto, observe que si $r_{ii} = 0$, entonces $\boldsymbol{a}_i$ es una combinación lineal de $\boldsymbol{q}_1, \dots, \boldsymbol{q}_{i-1}$ y, por tanto, está en W_{i-1} . Pero entonces $\boldsymbol{a_i}$ sería una combinación lineal de $\boldsymbol{a_1}, \dots, \boldsymbol{a_{i-1}}$, lo que es imposible, pues $\boldsymbol{a_1}, \dots, \boldsymbol{a_i}$ son linealmente independientes. Se concluye que $r_{ii} \neq 0$ para $i = 1, \dots n$. Dado que R es triangular superior, debe ser invertible.

Acaba de probarse el siguiente:

Teorema Sea A una matriz de $m \times n$ con columnas linealmente independientes. Entonces A puede factorizarse como $A = QR$, donde Q es una matriz de $m \times n$ con columnas ortonormales y R es una matriz triangular superior invertible.

Procedimiento

Dada una matriz A de $m \times n$ con columnas linealmente independientes $(rango\ (A) = n)$:

1. Obtener una base ortonormal de $col(A)$ mediante el método de Gram-Schmidt. Los vectores obtenidos forman las columnas de la matriz Q.
2. Calcular R . Para ello podemos proceder de dos modos:
   a. Mediante: $Q^T A = Q^T QR = IR = R$, es decir, $R = Q^T A$
   b. Calculando los elementos $r_{ji} = \boldsymbol{u}_i \cdot \boldsymbol{w_j}$

Comentarios

- Puede hacer arreglos para que las entradas diagonales de R sean positivas. Si cualquier $r_{ii} < 0$, simplemente sustituya $\boldsymbol{q}_i$ por $-\boldsymbol{q}_i$y r_{ii} por $-r_{ii}$.
- El requisito de que A tenga columnas linealmente independientes es necesario. Para probar esto, suponga que A es una matriz de $m \times n$ que tiene una factorización QR. Entonces, dado que R es invertible, se tiene $Q = AR^{-1}$. Por tanto, $rango(Q) = rango(A)$. Pero $rango(Q) = n$, pues sus columnas son ortonormales y, en consecuencia, linealmente independientes. De modo que $rango(A) = n$ también, y en consecuencia las columnas de A son linealmente independientes por el teorema fundamental.
- La factorización QR puede extenderse a matrices arbitrarias en una forma ligeramente modificada. Si A es de $m \times n$, se puede encontrar una secuencia de matrices ortogonales $Q_1, \dots, Q_{m-1}$ tales que $Q_{m-1} \cdot \dots \cdot Q_2 \cdot Q_1 \cdot A$ es una matriz triangular superior R de $m \times n$. Entonces $A = QR$, donde $Q = (Q_{m-1} \cdot \dots \cdot Q_2 \cdot Q_1)^{-1}$ es una matriz ortogonal.

Ejemplo

Encontrar la factorización QR de la matriz:

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

La base ortonormal para $col(A)$ producida por el proceso de Gram-Schmidt es:

$$\boldsymbol{q}_1 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \quad \boldsymbol{q}_2 = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix}, \quad \boldsymbol{q}_3 = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}$$

de mode que:

$$Q = [\boldsymbol{q}_1 \quad \boldsymbol{q}_2 \ \ \boldsymbol{q}_3] \cdot \begin{bmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{bmatrix}$$

Sabemos que $A = QR$ para alguna matriz triangular superior R. Para encontrar R, use el hecho de que Q tiene columnas ortonormales y, en consecuencia, $Q^T Q = I$. Por tanto:

$$Q^T A = Q^T QR = IR = R$$

Por tanto, calculamos R:

$$R = Q^T A = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix}$$

También podemos calcular los elementos de la matriz R mediante:

$$r_{ji} = \boldsymbol{u}_i \cdot \boldsymbol{w}_j$$

$$r_{11} = \boldsymbol{u}_1 \cdot \boldsymbol{w}_1 = (1, -1, -1, 1) \cdot \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) = 1 + 1 + 1 + 1 = 2$$

$$r_{12} = \boldsymbol{u}_2 \cdot \boldsymbol{w}_1 = (2, 1, 0, 1) \cdot \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) = 1 - \frac{1}{2} + 0 + \frac{1}{2} = 1$$

$$r_{13} = \boldsymbol{u}_3 \cdot \boldsymbol{w}_1 = (2, 2, 1, 2) \cdot \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) = 1 - 1 - \frac{1}{2} + 1 = \frac{1}{2}$$

$$r_{22} = \boldsymbol{u}_2 \cdot \boldsymbol{w}_2 = (2, 1, 0, 1) \cdot \left(\frac{3\sqrt{5}}{10}, \frac{3\sqrt{5}}{10}, \frac{\sqrt{5}}{10}, \frac{\sqrt{5}}{10}\right) = \frac{6\sqrt{5}}{10} + \frac{3\sqrt{5}}{10} + 0 + \frac{\sqrt{5}}{10} = \sqrt{5}$$

$$r_{23} = \boldsymbol{u}_3 \cdot \boldsymbol{w}_2 = (2, 2, 1, 2) \cdot \left(\frac{3\sqrt{5}}{10}, \frac{3\sqrt{5}}{10}, \frac{\sqrt{5}}{10}, \frac{\sqrt{5}}{10}\right) = \frac{6\sqrt{5}}{10} + \frac{6\sqrt{5}}{10} + \frac{\sqrt{5}}{10} + \frac{2\sqrt{5}}{10} = \frac{3\sqrt{5}}{2}$$

$$r_{33} = \boldsymbol{u}_3 \cdot \boldsymbol{w}_3 = (2, 2, 1, 2) \cdot \left(-\frac{\sqrt{6}}{6}, 0, \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right) = -\frac{2\sqrt{6}}{6} + 0 + \frac{\sqrt{6}}{6} + \frac{2\sqrt{6}}{3} = \frac{\sqrt{6}}{2}$$