Álgebra Lineal

Ruth Cueva - Felipe Navas - José Luis Toro
Profesores de la Escuela Politécnica Nacional

Edición general:
Juan Carlos Trujillo

Editores:
Fabián Barba, Juan Carlos Trujillo
Profesores de la Escuela Politécnica Nacional

Quito - Febrero 2009

Matrices
Determinantes
Sistemas lineales

por Ruth Cueva

Edición: Fabián Barba

Revisión: Juan Carlos Trujillo

Objetivos

El objetivo general de esta parte es:

Resolver problemas sobre sistemas de ecuaciones lineales con coeficientes reales de orden 4 *por* 4 *a lo más, utilizando las propiedades básicas de las matrices y de los determinantes, a un nivel reproductivo.*

Para alcanzar el objetivo general, se proponen los siguientes objetivos específicos:

1. Identificar los diferentes tipos de matrices a partir de su definición a un nivel de familiarización.
2. Calcular una matriz escalonada reducida por filas a partir de una matriz dada, mediante operaciones elementales de fila, a un nivel reproductivo.
3. Resolver problemas con elementos matriciales usando las operaciones de suma de matrices, multiplicación de un escalar con una matriz y la multiplicación entre matrices, a un nivel productivo.
4. Determinar la inversa de una matriz de orden 4 por 4 a lo más, a partir de las propiedades básicas de las operaciones elementales de fila a un nivel reproductivo.
5. Identificar las propiedades fundamentales del determinante de una matriz de orden n, mediante las operaciones elementales de fila y la definición a un nivel reproductivo.
6. Calcular el valor de un determinante de una matriz de orden n, usando sus propiedades y la definición a un nivel reproductivo.
7. Calcular la matriz inversa de una matriz de orden n a partir de su definición y de las propiedades de los determinantes a un nivel reproductivo.
8. Resolver sistemas de ecuaciones lineales con coeficientes constantes reales hasta de orden 4×4, a partir de las operaciones elementales de fila o de las propiedades de los determinantes a un nivel reproductivo.

Matrices | 1

1.1 Definiciones

En este libro, utilizaremos I_n para representar el conjunto de todos los números naturales desde 1 hasta n. Es decir:

$$I_n = \{k \in \mathbb{N} : 1 \leq k \leq n\} = \{1, 2, 3, \ldots, n\}.$$

Definición 1.1 (Matriz) Sean $m \in \mathbb{N}$ y $n \in \mathbb{N}$. Una **matriz** A **sobre un campo** $\mathbb{K}$ es una función A definida por:

$$\begin{aligned} A\colon I_m \times I_n &\longrightarrow \mathbb{K} \\ (i,j) &\longmapsto A(i,j) = a_{ij}. \end{aligned} \tag{1.1}$$

Cada $a_{ij} \in \mathbb{K}$ se denomina **elemento de la matriz** A.

Como el conjunto $I_m \times I_n$ tiene $m \times n$ elementos, la matriz A tiene $m \times n$ elementos, los mismos que se disponen en un arreglo rectangular de sus elementos, dispuestos en m filas y en n columnas de la siguiente forma:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \hdotsfor{4} \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{1.2}$$

Por esta razón, se dice que la matriz A es de *orden* m por n y se la representa por:

$$A = (a_{ij})_{m \times n}.$$

El conjunto de las matrices de orden m por n es notado $M_{m\times n}$. Si se quiere especificar que el conjunto de las matrices es sobre el campo de los números reales $\mathbb{R}$ o de los números complejos $\mathbb{C}$, se puede escribir $M_{m\times n}(\mathbb{R})$ o $M_{m\times n}(\mathbb{C})$, respectivamente. Así, la expresión $A \in M_{m\times n}(\mathbb{C})$ nos dice que la matriz A es una matriz de orden m por n (m filas y n columnas) y que sus elementos son números complejos.

Definición 1.2 (Igualdad de matrices) Dos matrices $A = (a_{ij})_{m\times n}$ y $B = (b_{ij})_{p\times q}$ son **iguales** si y solo si:

1. $m = p,\ n = q$, y
2. $a_{ij} = b_{ij}$ para todo $i = 1, 2, \ldots, m$ y todo $j = 1, 2, \ldots, n$.

Es decir, dos matrices son iguales si y solo si son del mismo orden y sus correspondientes elementos son iguales.

Ejemplo 1 Las matrices

$$A = \begin{pmatrix} 5 & -4 & 2 \\ 0 & 2 & -1 \\ -3 & -2 & 7 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 5 & x & 2 \\ 0 & 2 & y \\ z & -2 & 7 \end{pmatrix}$$

son iguales si $x = -4$, $y = -1$ y $z = -3$, ya que son del mismo orden y el resto de elementos correspondientes son iguales entre sí. ♦

Los nombres dados a las matrices que se definen a continuación aprovechan el hecho de que se han dispuesto los elementos de la matriz en un arreglo rectangular de los mismos. Posteriormente, tal arreglo rectangular permitirá también definir operaciones con las matrices.

Definición 1.3 (Matriz cuadrada) Una matriz es **cuadrada de orden** n si es de orden $n \times n$.

Es decir, una matriz cuadrada tiene igual número de filas y de columnas. Una matriz cuadrada A es notada $A = (a_{ij})_n$. El conjunto de las matrices cuadradas de orden $n \times n$ es notado M_n.

Un cuadrado tiene dos diagonales. En una matriz cuadrada se considera solo una de ellas: la que va desde la esquina superior izquierda a la esquina inferior derecha y es denominada *diagonal principal*, o simplemente *diagonal*. Es decir, la diagonal de la matriz $A = (a_{ij})_n$ es el conjunto de los elementos a_{ii}, con $i = 1, 2, \ldots, n$.

Definición 1.4 (Fila y columna de una matriz) Sea la matriz $A = (a_{ij})_{m \times n}$.

1. Se llama **i-ésima fila** de la matriz A a la matriz A_i de orden $1 \times n$ definida por:

$$A_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{i1} \end{pmatrix}.$$

2. Se llama **j-ésima columna** de la matriz A a la matriz A^j de orden $m \times 1$ definida por:

$$A^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Una matriz $A \in M_{1 \times n}$, que consta de una sola fila, se llama *matriz fila*. Una matriz $A \in M_{m \times 1}$, que consta de una sola columna, se llama *matriz columna*. Una fila (columna) de una matriz es una matriz fila (columna).

Definición 1.5 (Matriz transpuesta) La **matriz transpuesta** de la matriz $A = (a_{ij})_{m \times n}$ es la matriz $A^t = (\hat{a}_{ij})_{n \times m}$ donde

$$\hat{a}_{ij} = a_{ji}.$$

Ejemplo 2 Hallar la matriz transpuesta de la matriz

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -1 \end{pmatrix}.$$

Al aplicar la definición se obtiene que la matriz transpuesta de A es

$$A^t = \begin{pmatrix} 1 & -4 \\ -2 & 5 \\ 3 & -1 \end{pmatrix}.$$

♦

El resultado de aplicar la definición es que las filas de A se convierten en las columnas de A^t, es decir las filas y las columnas se transponen y de allí el nombre de transpuesta para la matriz que se obtiene.

Definición 1.6 (Matriz simétrica) Se dice que la matriz cuadrada A es **simétrica** si y solo si $A^t = A$.

Ejemplo 1.1

Sea la matriz

$$A = \begin{pmatrix} 3 & -1 & 4 \\ -1 & 7 & -2 \\ 4 & -2 & 9 \end{pmatrix}.$$

¿Es A una matriz simétrica?

Solución. Para responder a esta pregunta, hallemos su matriz transpuesta:

$$A^t = \begin{pmatrix} 3 & -1 & 4 \\ -1 & 7 & -2 \\ 4 & -2 & 9 \end{pmatrix}.$$

Como $A^t = A$ concluimos que A es una matriz simétrica. ♦

Observemos que los elementos que están a uno y otro lado de la diagonal son iguales. De ahí proviene el nombre de matriz simétrica.

Definición 1.7 (Matriz antisimétrica) Se dice que la matriz cuadrada $A = (a_{ij})_n$ es **antisimétrica** si y solo si $A^t = (-a_{ij})_n$.

Cuando veamos la operación suma de matrices, podremos definir el *inverso aditivo* de una matriz $A = (a_{ij})_{m\times n}$, notado $-A$, como la matriz $-A = (-a_{ij})_{m\times n}$. Allí podremos decir que una matriz cuadrada es antisimétrica si $A^t = -A$.

Ejemplo 1.2

Sea la matriz

$$A = (a_{ij})_3 = \begin{pmatrix} 0 & 1 & -4 \\ -1 & 0 & -2 \\ 4 & 2 & 0 \end{pmatrix}.$$

¿Es A una matriz antisimétrica?

Solución. Hallemos las matrices A^t y $(-a_{ij})_3$:

$$A^t = \begin{pmatrix} 0 & -1 & 4 \\ 1 & 0 & 2 \\ -4 & -2 & 0 \end{pmatrix}, \quad (-a_{ij})_3 = \begin{pmatrix} 0 & -1 & 4 \\ 1 & 0 & 2 \\ -4 & -2 & 0 \end{pmatrix}.$$

Encontramos que $A^t = (-a_{ij})_3$ y podemos afirmar que A es una matriz antisimétrica. ♦

Observemos que los elementos a uno y otro lado de la diagonal son el uno el negativo del otro y que la diagonal está formada por elementos que necesariamente son ceros.

Definición 1.8 (Matriz nula) Una matriz $A = (a_{ij})_{m\times n}$ se llama **matriz nula** si y solo si $a_{ij} = 0$ para todo $i = 1, 2, \ldots, m$ y todo $j = 1, 2, \ldots, n$.

En otras palabras, los elementos de una matriz nula son todos ceros. Hay una matriz nula para cada orden de matrices. La matriz nula se nota $0_{m\times n}$, o simplemente 0 cuando el orden está sobre entendido. Por el contexto quedará claro que se trata de la matriz nula y no del número cero. La matriz $0 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ es la matriz nula de las matrices fila de orden 1×3, y la matriz $0 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^t$ es la matriz nula de las matrices columna de orden 3×1.

Definición 1.9 (Matriz identidad) Una matriz $A = (a_{ij})_n$ se llama **matriz identidad** si y solo si $a_{ij} = 0$ para $i \neq j$, y $a_{ij} = 1$ para $i = j$.

La matriz identidad de orden n se nota I_n, o simplemente I cuando el orden está sobre entendido:

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

La matriz identidad I_n juega el papel de *neutro multiplicativo* en el producto de matrices cuadradas de orden n.

Definición 1.10 (Matriz diagonal) Una matriz $A = (a_{ij})_n$ se llama **matriz diagonal** si y solo si $a_{ij} = 0$ para $i \neq j$.

Por la definición, una matriz es diagonal si los elementos que no están en la diagonal son iguales a cero. Las matrices identidad y nula son matrices diagonales.

Definición 1.11 (Matriz escalar) Una matriz $A = (a_{ij})_n$ se llama **matriz escalar** si y solo si $a_{ij} = 0$ para $i \neq j$ y $a_{ij} = k \in \mathbb{K}$ para $i = j$.

Una matriz escalar es una matriz diagonal con los elementos de la diagonal iguales. Son ejemplos de matrices diagonales las matrices identidad y nulas.

Definición 1.12 (Matrices triangulares superior e inferior) Sea una matriz $A = (a_{ij})_n$.

1. Se dice que A es una **matriz triangular superior** si y solo si $a_{ij} = 0$ para $i > j$.
2. Se dice que A es una **matriz triangular inferior** si y solo si $a_{ij} = 0$ para $i < j$.

Las matrices triangulares superior e inferior tienen todos sus elementos iguales a cero bajo y sobre la diagonal, respectivamente. Las matrices escalares son triangulares superiores e inferiores a la vez.

Ejemplo 3 Las matrices

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

son triangulares superior e inferior, respectivamente. ♦

1.2 Operaciones con matrices

En lo que sigue nos referiremos como *escalares* a los números reales y complejos. Veremos las siguientes operaciones con matrices:

1. Suma de matrices.
2. Producto de un escalar por una matriz.
3. Producto de matrices.

1.2.1 Suma de matrices

Definición 1.13 (Suma de matrices) Sean las matrices $A = (a_{ij})_{m\times n}$ y $B = (b_{ij})_{m\times n}$. Se llama **suma** de A y B a la matriz $C = (c_{ij})_{m\times n}$ definida por

$$c_{ij} = a_{ij} + b_{ij}, \quad 1 \leq i \leq m,\ 1 \leq i \leq n.$$

La suma de matrices está, pues, definida para matrices de igual orden. La suma de A y B se nota $A+B$, y es una matriz de igual orden que el orden de A y B. De aquí en adelante, si escribimos $A+B$ se entiende que las matrices A y B son del mismo orden.

Ejemplo 4 Sean las matrices

$$A = \begin{pmatrix} 2 & -1 & 3 \\ -2 & 3 & -4 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 0 & 1 & -5 \\ 5 & -7 & 1 \end{pmatrix}.$$

Entonces

$$A + B = \begin{pmatrix} 2+0 & -1+1 & 3+(-5) \\ -2+5 & 3+(-7) & -4+1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 3 & -4 & -3 \end{pmatrix}.$$

♦

Teorema 1.1 (Propiedades de la suma de matrices) Para las matrices A, B y C de orden $m \times n$ se cumplen las siguientes propiedades:

1. **Propiedad clausurativa**: $A + B$ es una matriz de orden $m \times n$.
2. **Propiedad asociativa**: $(A + B) + C = A + (B + C)$.
3. **Propiedad conmutativa**: $A + B = B + A$.
4. **Existencia del neutro para la suma**: para toda matriz $A \in M_{m\times n}$, existe una matriz $0 \in M_{m\times n}$ tal que
$$A + 0 = A.$$
La matriz 0 se llama *neutro aditivo* o *matriz cero*.
5. **Existencia del opuesto para la suma**: para toda matriz $A \in M_{m\times n}$, existe una matriz $D \in M_{m\times n}$ tal que
$$A + D = 0.$$
La matriz D se representa mediante $-A$ y se llama *inverso aditivo* o *negativo* de A. Esta propiedad puede escribirse de la siguiente manera:
$$A + (-A) = 0.$$

La demostración de estas cinco propiedades se basa en las propiedades de campo de los números reales y complejos y se deja como ejercicio. Baste decir que como matriz cero (neutro aditivo) se toma la matriz nula definida anteriormente, y como negativo (inverso aditivo) de la matriz $A = (a_{ij})_{m\times n}$ se toma la matriz $-A = (-a_{ij})_{m\times n}$. Otras propiedades se propondrán como ejercicios.

El conjunto de las matrices $M_{m\times n}(\mathbb{K})$ dotado de la operación suma de matrices definida anteriormente, $(M_{m\times n}(\mathbb{K}), +)$ y las propiedades 1, 2, 4 y 5 del teorema anterior, constituye un grupo. Sí además se toma en cuenta la propiedad conmutativa de la suma, tal grupo es un grupo conmutativo o grupo abeliano.

Se puede demostrar que la matriz neutro aditivo es única y también que la matriz inverso aditivo es única. Estas demostraciones también se basan en la unicidad del neutro aditivo y del inverso aditivo de los escalares, es decir de los números reales o complejos.

La existencia del inverso aditivo permite definir la resta de matrices. La resta de las matrices A y B de igual orden, representada por $A - B$ se define como

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

1.2.2 Producto escalar

Definición 1.14 (Producto escalar) El **producto del escalar** $\alpha \in \mathbb{K}$ **y la matriz** $A = (a_{ij})_{m\times n}$, notado αA, se define como

$$\alpha A = \alpha(a_{ij})_{m\times n} = (\alpha\, a_{ij})_{m\times n}, \quad 1 \leq i \leq m,\ 1 \leq j \leq n.$$

Ejemplo 5

$$(-2+i)\begin{pmatrix} i & 1-i \\ -2 & 3+i \end{pmatrix} = \begin{pmatrix} -1-2i & -1+3i \\ -4+2i & -7+i \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & -2 \\ -2 & 3 \\ -1 & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & -4 \\ -4 & 6 \\ -2 & 0 \end{pmatrix}.$$

♦

Teorema 1.2 (Propiedades del producto escalar) Para los escalares α y β y las matrices A y B de orden $m \times n$ se cumplen las siguientes propiedades:

1. **Propiedad clausurativa**: αA es una matriz de orden $m \times n$.
2. **Propiedad asociativa**: $(\alpha\beta)A = \alpha(\beta A)$.
3. **Propiedad distributiva con respecto a la suma de matrices**: $\alpha(A + B) = \alpha A + \alpha B$.
4. **Propiedad distributiva con respecto a la suma de escalares**: $(\alpha + \beta)A = \alpha A + \beta A$.
5. **Propiedad asociativa mixta**: $A(\alpha B) = \alpha(AB) = (\alpha A)B$.
6. **El número** 1 **es el elemento identidad del producto escalar**: $1A = A$.

Estas son las principales propiedades del producto escalar, las que permitirán posteriormente hablar del “espacio lineal de matrices”. Otras propiedades se propondrán como ejercicios.

1.2.3 Multiplicación de matrices

Definición 1.15 (Multiplicación de matrices) Sean las matrices $A = (a_{ij})_{m\times n}$ y $B = (b_{ij})_{n\times p}$. El **producto de** A **y** B, notado AB, es la matriz $C = (c_{ij})_{m\times p}$ definida por

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}, \quad 1 \leq i \leq m,\ 1 \leq j \leq p.$$

El producto de A y B no está definido si el número de columnas de A no es igual al número de filas de B.

El elemento c_{ij} de AB es el elemento único de la matriz de orden 1×1 que resulta del producto de la i-ésima fila de A, A_i, y la j-ésima columna de B, B^j:

$$A_i B^j = \begin{pmatrix} a_{i1} & a_{i2} \cdots a_{in} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} = \begin{pmatrix} a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \end{pmatrix} = (c_{ij})_{1\times 1}.$$

La multiplicación de matrices es una operación muy especial. Los siguientes ejemplos muestran sus peculiaridades.

Ejemplo 6 Sean

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 3 & -21 & \end{pmatrix} \in M_{2\times 3} \quad \text{y} \quad B = \begin{pmatrix} 1 & 1 \\ 2 & -3 \\ 3 & 1 \end{pmatrix} \in M_{3\times 2}.$$

Entonces

$$\begin{aligned} AB &= \begin{pmatrix} 1 & -2 & 3 \\ 3 & -21 & \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -3 \\ 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1\cdot 1 + (-2)2 + 3\cdot 3 & 1\cdot 1 + (-2)(-3) + 3\cdot 1 \\ 3\cdot 1 + (-2)2 + 1\cdot 3 & 3\cdot 1 + (-2)(-3) + 1\cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 10 \\ 2 & 10 \end{pmatrix}. \end{aligned}$$

De manera similar se calcula

$$BA = \begin{pmatrix} 1 & 1 \\ 2 & -3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & 3 \\ 3 & -21 & \end{pmatrix} = \begin{pmatrix} 4 & -4 & 4 \\ -7 & 2 & 3 \\ 6 & -8 & 10 \end{pmatrix}.$$

Comentario. En este caso, los productos AB y BA están definidos y son matrices de órdenes diferentes, por lo cual $AB \neq BA$. Este ejemplo nos muestra que el producto de matrices A y B no es conmutativo. Nótese además que A y B no son cuadradas pero los dos productos son matrices cuadradas. ♦

Ejemplo 1.3 Sean

$$A = \begin{pmatrix} 2 & -2 \\ -1 & 3 \end{pmatrix} \in M_2 \quad \text{y} \quad B = \begin{pmatrix} 1 & -1 \\ 2 & -3 \end{pmatrix} \in M_2.$$

Entonces

$$AB = \begin{pmatrix} -2 & 4 \\ 5 & -8 \end{pmatrix} \quad \text{y} \quad BA = \begin{pmatrix} 3 & -5 \\ 7 & -13 \end{pmatrix}.$$

Las matrices AB y BA son cuadradas del mismo orden pero, una vez más, diferentes. En general, si las matrices A y B son cuadradas del mismo orden, los productos AB y BA están definidos y son también matrices cuadradas del mismo orden. ♦

Ejemplo 7 Tenemos que:

$$AB = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Las matrices A y B no son nulas, sin embargo su producto sí. ♦

Ejemplo 8 Sean:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \quad \text{y} \quad C = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

Tenemos que

$$AB = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} = AC,$$

pero $B \neq C$ a pesar de que $A \neq 0$. Por lo tanto, la propiedad cancelativa no aplica al producto de matrices. ♦

La multiplicación de matrices no tiene la propiedad conmutativa ni la cancelativa, pero al menos tiene las propiedades asociativa y distributivas como se expresa en el siguiente teorema.

Teorema 1.3 (Propiedades de la multiplicación de matrices) Para matrices A, B y C de órdenes tales que los productos están definidos, se cumplen las siguientes propiedades:

1. **Propiedad asociativa**: $(AB)C = A(BC)$.
2. **Propiedad distributiva por la izquierda**: $A(B + C) = AB + AC$.
3. **Propiedad distributiva por la derecha**: $(A + B)C = AC + BC$.

Demostración. Sean $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{n\times p}$ y $C = (c_{ij})_{p\times q}$.

1. **Propiedad asociativa**:

$$\begin{aligned}
(AB)C &= \big((a_{ij})_{m\times n}(b_{ij})_{n\times p}\big)(c_{ij})_{p\times q} \\
&= \left(\sum_{r=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kr}\right)c_{rj}\right) \\
&= \left(\sum_{r=1}^{p}\sum_{k=1}^{n} a_{ik}b_{kr}c_{rj}\right) \\
&= \left(\sum_{k=1}^{n}\sum_{r=1}^{p} a_{ik}b_{kr}c_{rj}\right) \\
&= \left(\sum_{k=1}^{n} a_{ik}\left(\sum_{r=1}^{p} b_{kr}c_{rj}\right)\right) = A(BC).
\end{aligned}$$

2. **Propiedad distributiva por la izquierda**:

$$\begin{aligned}
A(B+C) &= (a_{ij})_{m\times n}\big((b_{ij})_{n\times p} + C = (c_{ij})_{n\times p}\big)\\
&= (a_{ij})_{m\times n}((b_{ij}) + c_{ij})_{n\times p}\\
&= \sum_{k=1}^{n} a_{ij}(b_{kj} + c_{kj})\\
&= \sum_{k=1}^{n} a_{ij}b_{kj} + \sum_{k=1}^{n} a_{ij}c_{kj}\\
&= AB + AC.
\end{aligned}$$

3. La demostración de la propiedad distributiva por la derecha es similar a la anterior y se deja al lector como ejercicio.

□

Ejemplo 9 Sea

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

Tenemos que

$$I_3A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & -3 \\ -1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -3 \\ -1 & 2 & 4 \end{pmatrix}$$

y

$$AI_2 = \begin{pmatrix} 2 & 0 & -3 \\ -1 & 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -3 \\ -1 & 2 & 4 \end{pmatrix}.$$

Es decir, en este caso,

$$I_2A = AI_3 = A.$$

♦

En general, si $A \in M_{m\times n}$ la igualdad siguiente es verdadera:

$$I_nA = AI_m = A.$$

La matriz I_n suele ser llamada *neutro multiplicativo de A por la izquierda* y la matriz I_m, *neutro multiplicativo de A por la derecha*. En el caso particular de que A es una matriz cuadrada de orden n, la igualdad anterior se ve así:

$$I_nA = AI_n = A.$$

La matriz identidad I_n es denominada simplemente *el elemento neutro para el producto de matrices cuadradas de orden n*.

El conjunto de las matrices cuadradas de orden $M_n(\mathbb{K})$ dotado de las operaciones suma y producto de matrices es un anillo no conmutativo con elemento neutro.

1.3 Operaciones elementales de fila de una matriz

Para cualquier matriz se pueden definir operaciones entre sus filas y también operaciones entre sus columnas. Aquí definiremos las operaciones, que se denominarán *elementales*, de fila.

Definición 1.16 (Operaciones elementales de fila) Sean la matriz $A \in M_{m\times n}(\mathbb{K})$ y el escalar $c \in \mathbb{K}$. Cada una de las siguientes tres operaciones se denomina **operación elemental de fila sobre la matriz** A:

1. **Multiplicar la fila** r-ésima de A por una constante c con $c \neq 0$ y $c \neq 1$. Es decir, *reemplazar* A_r por cA_r.

2. **Sumar** c **veces la fila** s-ésima, cA_s, con la fila r-ésima de A. Es decir, *reemplazar* A_r por $cA_s + A_r$.

3. **Intercambiar las filas distintas** s-ésima y r-ésima de A. Es decir, *reemplazar* A_s por A_r y A_r por A_s con $s \neq r$.

El significado exacto de *reemplazar* es el siguiente: dada la matriz A, se obtiene, en realidad, una nueva matriz A', la cual es idéntica a la matriz A salvo en su fila número r que es cA_r. En otras palabras, la operación *multiplicar una fila por una constante* es una función que va desde el conjunto $M_{m\times n}(\mathbb{K})$ en sí mismo tal que a cada A le hace corresponder A'. Significados similares corresponden a las otras dos operaciones. Más adelante, precisaremos estos significados.

Estas tres operaciones elementales de fila pueden ser representadas así:

1. $cA_r \rightharpoonup A_r$, o simplemente cA_r, que quiere decir: "obtenga una nueva matriz al quitar la fila A_r y, en su lugar, poner la fila cA_r".

2. $A_r + cA_s \rightharpoonup A_r$, o simplemente $A_r + cA_s$, que quiere decir: "obtenga una nueva matriz al quitar la fila A_r y, en su lugar, poner la fila $A_r + cA_s$".

3. $A_s \rightleftharpoons A_r$, que quiere decir: "obtenga una nueva matriz al intercambiar las filas A_s y A_r entre sí".

A las operaciones elementales de fila en el orden en que fueron definidas las denominaremos del *tipo* 1, 2 y 3, respectivamente.

Como ya se dijo anteriormente, al efectuar una de las operaciones elementales de fila sobre una matriz A, se obtiene otra matriz. Formalmente, se puede definir cada una de las operaciones elementales de fila como una función.

Definición 1.17 (Operaciones elementales de fila) Una **operación elemental de fila** sobre una matriz $A = (a_{ij}) \in M_{m\times n}$ es una de las tres funciones

$$\begin{aligned} e\colon M_{m\times n} &\longrightarrow M_{m\times n} \\ A &\longmapsto e(A) = (\hat{a}_{ij}) \end{aligned}$$

tales que:

1. **Tipo 1**: si $c \neq 0$ y $c \neq 1$,
$$\hat{a}_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r, \\ ca_{ij} & \text{si } i = r \end{cases}$$
para todo $j \in I_n$.

   A esta función la vas a representar con e_r^c para indicar que se multiplica la fila número r por el número c.

2. **Tipo 2**: si $c \neq 0$,
$$\hat{a}_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r \\ ca_{sj} + a_{ij} & \text{si } i = r \end{cases}$$
para todo $j \in I_n$.

   A esta función le representaremos con $e_{s,r}^c$ para indicar que se multiplica la fila número s por c y la fila resultante es sumada esta fila a la fila número r.

3. **Tipo 3**: si $r \neq s$,
$$\hat{a}_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r \text{ y } i \neq s, \\ a_{sj} & \text{si } i = r, \\ a_{rj} & \text{si } i = s \end{cases}$$
para todo $j \in I_n$.

   A esta función le representaremos con $e_{r,s}$ para indicar que se intercambiaron las filas número s y número r.

Las operaciones de tipo 1 y tipo 2 afectan solo a una fila de la matriz; las demás quedan idénticas. La operación de tipo 3 afecta a dos filas; las demás quedan inalteradas.

El siguiente teorema establece que las operaciones elementales son funciones biyectivas, por lo que poseen inversa.

Teorema 1.4 Cada operación elemental de fila es biyectiva, y su inversa es del mismo tipo. Es decir, a cada operación elemental de fila e, le corresponde una operación elemental de fila $\widehat{e}$, del mismo tipo que e, tal que:
$$e(\widehat{e}(A)) = \widehat{e}(e(A)) = A.$$
para todo $A \in M_{m\times n}$.

Demostración. Sea e una operación elemental de fila de tipo 1; es decir, e es de la forma e_r^c con $c \neq 0$ y $r \in I_m$. Vamos a probar que es e_r^c es inyectiva y sobreyectiva.

Para ver la inyectividad, supongamos que A y B dos matrices (del mismo orden) tales que $e_r^c(A) = e_r^c(B)$. Entonces:

1. para $i \neq r$, la fila i-ésima de $e_r^c(A)$ es A_i;

2. para $i \neq r$, la fila i-ésima de $e_r^c(B)$ es B_i.

Dado que $e_r^c(A) = e_r^c(B)$, podemos concluir que $A_i = B_i$ para todo $i \neq r$.

Por otro lado, la fila r-ésima de $e_r^c(A)$ es cA_r y la de $e_r^c(B)$ es cB_r. Otra vez, como $e_r^c(A) = e_r^c(B)$, concluimos que $cA_r = cB_r$, de donde, ya que $c \neq 0$, deducimos que $A_r = B_r$.

En resumen, cada fila de A es igual a la correspondiente fila de B, por lo que $A = B$.

Razonamientos similares nos permiten demostrar que las operaciones elementales de los otros tipos son inyectiva, por lo que se dejan las demostraciones al lector.

Probemos ahora la sobreyectividad de la operación e_r^c. Para ello, dada una matriz B, debemos encontrar una matriz A tal que $e_r^c(A) = B$.

Para hallar A, observemos que B deberá tener las filas correspondiente de A, salvo la r-ésima fila, que será cA_r. Es decir:

$$B_i = \begin{cases} A_i & \text{si } i \neq r, \\ cA_r & \text{si } i = r. \end{cases}$$

Esto nos sugiere, inmediatamente, que la matriz A que buscamos, debe definirse así:

$$A_i = \begin{cases} B_i & \text{si } i \neq r, \\ \frac{1}{c}B_r & \text{si } i = r, \end{cases}$$

lo que es posible, ya que $c \neq 0$.

Definida de esta manera la matriz A, es fácil comprobar que $e_r^c(A) = B$, con lo que queda demostrado que e_r^c es sobreyectiva. Adicionalmente, esta demostración nos ha dado la inversa de e_r^c:

$$(e_r^c)^{-1} = e_r^{\frac{1}{c}}.$$

Demostraciones similares se pueden hacer para los otros dos tipos de funciones elementales por filas, por lo que se las deja para que el lector las realice por sí mismo. □

Si no se presta a confusión, evitaremos el uso de los paréntesis en $e(A)$; en su lugar, escribiremos simplemente eA. Esta nueva notación será de mucha utilidad en lo que viene a continuación.

Supongamos que $e_1, e_2, \ldots, e_n$ son n operaciones elementales por filas, y A y B son dos matrices tales que

$$B = e_n e_{n-1} \cdots e_2 e_1 A.$$

En otras palabras, la matriz B se obtuvo, a partir de la matriz A, luego de aplicar sucesivamente, n operaciones elementales.

Es fácil ver que, como existen las inversas de cada e_i, entonces A se puede obtener a partir de B, también por la aplicación de n operaciones elementales:

$$A = e_1 e_2 \cdots e_{n-1} e_n B.$$

Esta relación entre las matrices A y B es, en realidad, una relación simétrica en el conjunto de matrices.

En efecto, si indicamos con $A \sim B$ cuando A se obtiene a partir de B por la aplicación sucesiva de un número finito de operaciones elementales por fila, entonces, es inmediato que $A \sim B$ implica $B \sim A$.

Más aún, se verifica también que $A \sim A$; es decir, toda matriz A se puede obtener a partir de sí misma por la aplicación de un número finito de operaciones elementales. Por ejemplo, si e es una operación elemental, como es inversible, tenemos que

$$A = e^{-1}eA.$$

En otras palabras, la relación $\sim$ es reflexiva.

Finalmente, es fácil probar que también es transitiva. Esto significa que la relación $\sim$ es, en realidad, una relación de equivalencia, lo que motiva la siguiente definición.

Definición 1.18 (Matrices equivalente por filas) La matriz $B \in M_{m\times n}$ es **equivalentes por filas** a la matriz $A \in M_{m\times n}$, notado $B \sim A$, si B se obtiene de A por la aplicación sucesiva un número finito de operaciones elementales por fila.

Dicho de otra manera, si $e_1, e_2, \ldots e_n$ es una sucesión finita de operaciones elementales por filas, tenemos que:

$$B \sim A \quad \Longleftrightarrow \quad B = e_n e_{n-1} \cdots e_2 e_1 A.$$

Teorema 1.5 La relación $\sim$ es de equivalencia sobre el conjunto $M_{m\times n}$:

1. **Reflexiva**: $A \sim B$.
2. **Simétrica**: Si $A \sim B$, entonces $B \sim A$.
3. **Transitiva**: Si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

La verificación de la propiedad transitiva se deja para que el lector la realice.

Ejemplo 10 Sea la matriz

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

Sobre A efectuamos una secuencia de operaciones elementales por filas:

$$A = \begin{pmatrix} 2 & -6 \\ 3 & 4 \\ -2 & 5 \\ 5 & -2 \\ -1 & -2 \end{pmatrix}_{\frac{1}{2}F_1} \sim \begin{pmatrix} 1 & -3 \\ 3 & 4 \\ -2 & 5 \\ 5 & -2 \\ -1 & -2 \end{pmatrix}_{F_2-3F_1} \sim \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ -2 & 5 \\ 5 & -2 \\ -1 & -2 \end{pmatrix}_{F_3+2F_1}$$

$$\sim \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ 0 & -1 \\ 5 & -2 \\ -1 & -2 \end{pmatrix}_{F_4-5F_1} \sim \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ 0 & -1 \\ 0 & 13 \\ -1 & -2 \end{pmatrix}_{F_5+F_1} \sim \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ 0 & -1 \\ 0 & 13 \\ 0 & -5 \end{pmatrix} = B.$$

Como la matriz B fue obtenida de A por la aplicación de una secuencia finita de operaciones elementales por filas, podemos decir que B es equivalente por filas a A:

$$B \sim A.$$

Aunque sabemos que $A \sim B$, ilustremos esta propiedad al aplicar sobre la matriz B, en orden inverso, la secuencia de operaciones inversas anterior, para obtener la matriz A:

$$B = \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ 0 & -1 \\ 0 & 13 \\ 0 & -5 \end{pmatrix}_{F_5 - F_1} \sim \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ 0 & -1 \\ 0 & 13 \\ -1 & -2 \end{pmatrix}_{F_4 + 5F_1} \sim \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ 0 & -1 \\ 5 & -2 \\ -1 & -2 \end{pmatrix}_{F_3 - 2F_1}$$
$$\sim \begin{pmatrix} 1 & -3 \\ 0 & 13 \\ -2 & 5 \\ 5 & -2 \\ -1 & -2 \end{pmatrix}_{F_2 + 3F_1} \sim \begin{pmatrix} 1 & -3 \\ 3 & 4 \\ -2 & 5 \\ 5 & -2 \\ -1 & -2 \end{pmatrix}_{2F_1} \sim \begin{pmatrix} 2 & -6 \\ 3 & 4 \\ -2 & 5 \\ 5 & -2 \\ -1 & -2 \end{pmatrix} = A,$$

que no podía ser de otra manera. ♦

Definición 1.19 (Matriz reducida por filas) Una matriz $A \in M_{m\times n}$ se llama matriz **reducida por filas** si cumple con las siguientes condiciones:

1. El primer elemento no nulo de cada fila no nula,visto de izquierda a derecha, es 1.
2. Cada columna de A que tiene el primer elemento no nulo de alguna fila tiene sus otros elementos nulos.

Ejemplo 11 Las matrices

$$\begin{pmatrix} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 4 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 2 & 7 & 0 \\ 0 & 1 & 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

son matrices reducidas por filas. En cambio, las matrices

$$\begin{pmatrix} 1 & 2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 4 & 0 & 3 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 0 & 1 & 0 & 0 & -2 & 5 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

no son matrices reducidas por filas. ♦

Teorema 1.6 Toda matriz $A \in M_{m\times n}$ es equivalente por filas a una matriz $R \in M_{m\times n}$ reducida por filas.

Dicho de otra forma, por medio de una secuencia apropiada de operaciones elementales de fila efectuadas sobre una matriz A, es siempre posible obtener una matriz reducida por filas equivalente por filas a la matriz A. Tal matriz no es única.

Demostración. Supongamos que la matriz A es no nula. Entonces al menos una fila es diferente de la matriz 0 de orden $1 \times n$. Sean $p \in I_m$ tal que $A_p \neq 0$. Por lo tanto, existe el menor $r \in I_n$ tal que $a_{pr} \neq 0$ (es decir, a_{pr} es el primer elemento no nulo de la fila p visto desde la izquierda hacia la derecha).

Sea $R[1] = e_p^{\frac{1}{a_{pr}}} A$. Esto significa que $R[1]$ se ha obtenido a partir de A al dividir por $\frac{1}{a_{pr}}$ cada uno de los elementos de la fila número p de A. Entonces, el primer elemento de la p-ésima fila de la matriz $R[1]$ es igual a 1.

Ahora bien, para cada $i \in I_m - \{p\}$, definamos $R[i+1] = e_{p,i}^{-a_{ir}}$. Es decir, para cada $i \neq p$, se multiplica la fila p-ésima por a_{ir} y la fila resultante se suma a la fila i-ésima, la que es reemplazada por el resultado obtenido. En otras palabras, la fila i-ésima de $R[i+1]$ es sustituida por la fila $R[i]_i - a_{ir}R[i]_p$. El elemento de la columna r de la matriz $I[i+1]$ obtenida es, entonces 0.

Una vez que se realiza este proceso para cada $i \neq p$ (es decir, se realizan $m-1$ veces), obtenemos la matriz $R[m]$ con la propiedad de que todos los elementos de la columna r son iguales a 0, excepto el que está en la fila p que es 1. Debe quedar claro que $R[m] \sim A$.

Si repetimos este procedimiento con otra fila que no sea nula, que no tenga el primer elemento igual a 1 y que en la columna en la que está dicho elemento, el resto de elementos no son todos iguales a 0, obtendremos, luego de m pasos, una matriz $R[2m]$ equivalente por filas a $R[m]$, lo que significa que será también equivalente por filas a A. Si la fila no nula involucrada en este procedimiento es la q-ésima y la columna que contiene el primer elemento no nula de esta fila es la número s, la matriz $R[2m]$ tendrá la característica de que todos los elementos de la columna número s son iguales a 0, excepto el de la fila q que es 1.

Está claro que si repetimos este proceso con todas las filas no nulas cuyo primer elemento no sea igual a 1 y que en la columna en la que está dicho elemento, no todos los elementos son 0, finalmente tendremos una matriz reducida por filas que será, obviamente, equivalente a la matriz A. $\square$

Ejemplo 1.4

Sea la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \end{pmatrix}.$$

Hallar una matriz reducida por filas equivalente por filas a la matriz A.

Solución.

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \end{pmatrix}_{\substack{F_2-2F_1 \\ F_3-3F_1 \\ F_4-4F_1}} \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -2 & -3 & -4 \\ 0 & -2 & -4 & -6 & -8 \\ 0 & -3 & -6 & -9 & -12 \end{pmatrix}_{\substack{F_1+2F_2 \\ F_3+2F_2 \\ F_4+3F_2}}$$

$$\sim \begin{pmatrix} 1 & 0 & -1 & -2 & -3 \\ 0 & -1 & -2 & -3 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}_{-F_2} \sim \begin{pmatrix} 1 & 0 & -1 & -2 & -3 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}_{F_1 \rightleftharpoons F_4}$$

$$\sim \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & -2 & -3 \end{pmatrix}.$$

Las matrices

$$R_1 = \begin{pmatrix} 1 & 0 & -1 & -2 & -3 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad R_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & -2 & -3 \end{pmatrix}$$

son matrices reducidas por filas y, además, equivalentes por filas a la matriz A:

$$R_1 \sim A \quad \text{y} \quad R_2 \sim A.$$

♦

Este ejemplo muestra que una matriz reducida por filas equivalente por filas a una matriz no es única.

Definición 1.20 (Rango de una matriz) El **rango** de una matriz A, notado $\operatorname{ran}(A)$, es el número de filas no nulas de una matriz reducida por filas equivalente por filas a la matriz A.

Ejemplo 12 Del ejemplo anterior, la matriz

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

es una matriz reducida por filas equivalente por filas a la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \end{pmatrix}.$$

El número de filas no nulas de la matriz R es 2. Por lo tanto,

$$\operatorname{ran}(A) = 2.$$

♦

Definición 1.21 (Matriz escalonada reducida por filas) Una matriz $A \in M_{m\times n}$ se llama matriz **escalonada reducida por filas** si cumple con las siguientes condiciones:

1. La matriz A es reducida por filas.
2. Toda fila nula, si la hay, está por debajo de toda fila no nula.
3. Si las filas $1, 2, \dots, r$, $r \leq m$, son las filas no nulas de la matriz A y si el primer elemento no nulo, visto de izquierda a derecha, de la fila i, $i = 1, 2, \dots, r$ está en la columna k_i, $i = 1, 2, \dots, r$, entonces $k_1 < k_2 < \cdots < k_r$.

Ejemplo 1.5

¿Es la matriz

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

una matriz escalonada reducida por filas?

Solución. La matriz A es reducida por filas. Las filas nulas son la 5 y la 6 y se encuentran por debajo de las filas no nulas 1, 2, 3 y 4. El primer elemento no nulo de la fila 1 está en la columna 2: $k_1 = 2$. El primer elemento no nulo de la fila 2 está en la columna 4: $k_2 = 4$. De manera similar encontramos: $k_3 = 5$ y $k_4 = 8$. Se cumple que $k_1 = 2 < k_2 = 4 < k_3 = 5 < k_4 = 8$. Por todo ello, la matriz A es una matriz escalonada reducida por filas. ♦

Teorema 1.7 Toda matriz $A \in M_{m\times n}$ es equivalente por filas a una matriz escalonada reducida por filas. Tal matriz es única.

Demostración. Del teorema 1.6, existe una matriz reducida por filas R equivalente a A. Si esta matriz R no es escalonada, entonces significa una de dos situaciones, o las dos, suceden:

1. hay filas nulas sobre filas no nulas;

2. hay al menos dos filas no nulas en las que el primer elemento de la que está más arriba corresponde a una columna que está a la derecha de la columna que contiene el primer elemento de la otra fila.

Si ocurre la primera situación, por la aplicación de operaciones elementales del tercer tipo, obtenemos, a partir de R, una matriz S en la que toda fila nula está por debajo de toda fila no nula (intercambiamos filas nulas con no nulas cuando las primeras están arriba de la segundas).

Si luego de este proceso, ocurre la segunda situación, es decir, hay filas en las que la que está arriba tiene su primer elemento a la derecha de la que está abajo, aplicamos operaciones del tercer tipo nuevamente para intercambiar dichas filas. La matriz así obtenida es escalonada reducida y equivalente a S, por lo que es equivalente a R y, finalmente, equivalente a la matriz A.

Es fácil ver que la matriz obtenida al final de proceso descrito es única. □

En otras palabras, dada una matriz A cualquiera, por medio de una secuencia apropiada de operaciones elementales de fila se obtiene siempre una única matriz escalonada reducida por filas equivalente por filas a la matriz A.

De aquí en adelante, vamos a utilizar la siguiente notación alternativa para las operaciones elementales:

1. Para e_r^c, usaremos cF_r; es decir, la operación de multiplicar la fila r-ésima por el escalar $c \neq 0$.
2. Para $c_{s,r}^c$, usaremos $F_r + cF_s$; es decir, para sustituir la fila r-ésima por la suma de esta fila con la fila producto de c y la fila s-ésima.
3. Para $c_{s,r}$, usaremos $F_r \rightleftharpoons F_s$; es decir, para el intercambio entre las filas número r y s.

1.4 Matrices elementales

Definición 1.22 (Matriz elemental) Se dice que una matriz $E \in M_n$ es una **matriz elemental** si se obtiene de la matriz identidad $\mathbb{I}_n \in M_n$ por medio de solo una operación elemental de fila.

Ejemplo 13 Las matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

son matrices elementales pues todas ellas se obtuvieron a partir de la matriz I_3 por medio de la única operación elemental de fila $F_2 \rightleftharpoons F_3$, $F_1 - 2F_3$ y $3F_2$, respectivamente.

Las matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{y} \quad \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

no son matrices elementales. La segunda, por ejemplo, se obtuvo de I_3 por medio de dos operaciones elementales de fila: $F_1 - 2F_3$ y $3F_2$. ♦

Teorema 1.8 Sea e una operación elemental de fila y sea $E \in M_m$ la matriz elemental tal que $e(\mathbb{I}_m) = E$. Entonces, para toda matriz $A \in M_{m\times n}$:

$$e(A) = EA.$$

La demostración debe hacerse para cada tipo de operación elemental de fila. Haremos la demostración para una operación elemental de fila del tipo 2.

Demostración. Sea $\mathbb{I}$ la matriz identidad de orden m. Omitiremos el subíndice en $\mathbb{I}$ para distinguir de la notación que utilizamos para las filas de una matriz. Así, en este caso, $\mathbb{I}_i$ indica la fila i-ésima de la matriz $\mathbb{I}$.

Con esta notación, recordemos que $\mathbb{I} = (\delta_{ij})$ donde

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

Procedamos con la demostración.

Sean $r \in I_m$ y $s \in I_m$ tales que $r \neq s$, y $E = e^c_{s,r}(\mathbb{I})$. Entonces:

$$E_i = \begin{cases} \mathbb{I}_i & \text{si } i \neq r, \\ cA_s + \mathbb{I}_r & \text{si } i = r. \end{cases}$$

Por lo tanto, si $E = (e_{ij})$, entonces:

$$e_{ik} = \begin{cases} \delta_{ij} & \text{si } i \neq r, \\ \delta_{ij} + c\delta_{sj} & \text{si } i = r \end{cases}$$

para todo $j \in I_n$.

Por otro lado, la matriz las filas de la matriz $e^c_{s,r}(A) = (\alpha_{ij})$ son:

$$A_i = \begin{cases} A_i & \text{si } i \neq r, \\ A_i + cA_s & \text{si } i = r, \end{cases}$$

es decir:

$$\alpha_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r, \\ a_{rj} + ca_{sj} & \text{si } i = r \end{cases}$$

para todo $j \in I_n$.

Finalmente, sea $C = EA = (c_{ij})$. Entonces:

$$c_{ij} = \sum_{k=1}^{m} e_{ik}a_{kj} = \begin{cases} \displaystyle\sum_{k=1}^{m} \delta_{ik}a_{kj} & \text{si } i \neq r, \\ \displaystyle\sum_{k=1}^{m} (\delta_{rk} + c\delta_{sk})a_{kj} & \text{si } i = r \end{cases}$$

para todo $j \in I_n$.

Pero, como $\delta_{ik} = 0$ si $i \neq k$, entonces:

$$\sum_{k=1}^{m} e_{ik}a_{kj} = \delta_{ii}a_{ij} = a_{ij}.$$

Con un razonamiento similar, obtenemos que:

$$\sum_{k=1}^{m} (\delta_{rk} + c\delta_{sk})a_{kj} = \delta_{rr}a_{rj} + c\delta_{ss})a_{sj} = a_{rj} + ca_{sj}.$$

Por lo tanto:

$$c_{ij} = \begin{cases} a_{ij} & \text{si } i \neq r, \\ a_{rj} + ca_{sj} & \text{si } i = r \end{cases}$$

para todo $j \in I_n$.

Por lo tanto, $c_{ij} = \alpha_{ij}$ para todo $i \in I_m$ y todo $j \in I_n$. Es decir, $EA = e^c_{s,r}(A)$, que es lo que queríamos demostrar. □

Veamos ahora el mismo teorema ilustrado con un ejemplo.

Ejemplo 14 Sea la matriz

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

Ilustremos el teorema para una operación elemental de fila tipo 1. Sea e_1 la operación elemental de fila $3F_1$. Tenemos que

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{3F_1} \sim \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = e_1(I_2) = E_1$$

y

$$A = \begin{pmatrix} 2 & -3 & 1 \\ 4 & 3 & -2 \end{pmatrix}_{3F_1} \sim \begin{pmatrix} 6 & -9 & 3 \\ 4 & 3 & -2 \end{pmatrix} = e_1(A).$$

Por otro lado

$$E_1 A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ 4 & 3 & -2 \end{pmatrix} = \begin{pmatrix} 6 & -9 & 3 \\ 4 & 3 & -2 \end{pmatrix},$$

de donde

$$e_1(A) = E_1\, A.$$

♦

Corolario 1.9 Sean A y B matrices de orden $m \times n$. La matriz B es equivalente por filas a la matriz A si y solo si

$$B = PA,$$

donde P es una matriz producto de matrices elementales de orden n.

Demostración. Sabemos que $B \sim A$ es equivalente a que exista una sucesión e_1, e_2, ..., e_q operaciones elementales tales que $B = e_q \cdots e_2 e_1(A)$.

Por el teorema 1.8, existe una matriz elemental E_1 tal que $e_1(A) = E_1 A$. Por lo tanto, tenemos que:

$$B \sim A \Longleftrightarrow e_q e_{q-1} \cdots e_2(E_1 A).$$

Una vez más, por el teorema 1.8, existe una matriz elemental E_2 tal que $e_2(E_1 A) = E_2(E_1 A) = (E_2 E_1)A$. Entonces, tenemos que:

$$B \sim A \Longleftrightarrow e_q e_{q-1} \cdots (E_2 E_1)A.$$

Si procedemos de manera similar hasta llegar a la operación e_q, garantizamos la existencia de q matrices elementales $E_1, E_2, \ldots, E_q$ matrices elementales tales que:

$$B \sim A \Longleftrightarrow (E_q \cdots E_2 E_1)A.$$

Sea $P = E_q \cdots E_2 E_1$. Entonces:

$$B \sim A \Longleftrightarrow PA,$$

donde P es el producto de matrices elementales. □

La demostración del corolario nos muestra que la misma secuencia de operaciones elementales de fila $e_1, e_2, \ldots e_q$ que nos conducen de la matriz A a la matriz B, conducen también la matriz identidad $\mathbb{I}_m$ a la matriz P. Es decir, si $(A|I_m)$ representa la matriz $A \in M_{m\times n}$ aumentada con la matriz identidad I_m (la matriz I escrita a continuación de la matriz A) se tiene que:

$$(A|I_m) \sim (B|P).$$

Ejemplo 15 Sea la matriz

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

Obtengamos una matriz B equivalente por filas a la matriz A y, al mismo tiempo, obtengamos una matriz P tal que $B = PA$:

$$\begin{aligned}(A|I_3) = \left(\begin{array}{cc|ccc} 2 & 1 & 1 & 0 & 0 \\ 2 & 4 & 0 & 1 & 0 \\ -1 & -2 & 0 & 0 & 1 \end{array}\right) & \underset{F_1+F_3}{\sim} \left(\begin{array}{cc|ccc} 1 & -1 & 1 & 0 & 0 \\ 2 & 4 & 0 & 1 & 0 \\ -1 & -2 & 0 & 0 & 1 \end{array}\right) \\ \underset{\frac{1}{2}F_2}{\sim} \left(\begin{array}{cc|ccc} 1 & -1 & 1 & 0 & 1 \\ 1 & 2 & 0 & \frac{1}{2} & 0 \\ -1 & -2 & 0 & 0 & 1 \end{array}\right) & \underset{F_2 \rightleftharpoons F_3}{\sim} \left(\begin{array}{cc|ccc} 1 & -1 & 1 & 0 & 1 \\ -1 & -2 & 0 & 0 & 1 \\ 1 & 2 & 0 & \frac{1}{2} & 0 \end{array}\right) = (B|P),\end{aligned}$$

con

$$B = \begin{pmatrix} 1 & -1 \\ -1 & -2 \\ 1 & 2 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Verifiquemos que se cumple el corolario 1.9, es decir, que la matriz P hallada satisface la expresión $B = PA$:

$$PA = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 2 & 4 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & -2 \\ 1 & 2 \end{pmatrix} = B.$$

♦

1.5 La inversa de una matriz

Definición 1.23 (Inversa de una matriz) Una matriz cuadrada A de orden n se llama **invertible** si existe una matriz cuadrada B de orden n tal que

$$AB = BA = I_n.$$

La matriz B se llama una **inversa** de A. Si no existe tal matriz B, se dice que la matriz A es **no inversible**.

Una matriz invertible también se llama matriz *no singular*, y una matriz no invertible, matriz *singular*.

Es inmediato de la definición de matriz inversible que la matriz nula no es inversible. Si lo fuera, digamos que B es su inversa, entonces $0B = I$. Pero $0B = 0$, de donde obtendríamos que $I = 0$, lo que es falso.

Ejemplo 16 Sean las matrices

$$A = \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} -1 & -2 \\ -1 & -1 \end{pmatrix}.$$

Entonces
$$AB = \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & -2 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$
y
$$BA = \begin{pmatrix} -1 & -2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$
Como
$$AB = BA = I_2$$
podemos decir que la matriz B es una inversa de A y que la matriz A es no singular. ♦

Teorema 1.10 (Unicidad de la inversa) Si una matriz tiene una inversa, entonces la inversa es única.

Demostración. Supongamos que B y C son inversas de la matriz A. Entonces
$$AB = BA = I_n \quad \text{y} \quad AC = CA = I_n,$$
de donde
$$BA = AC = I_n.$$
Por otro lado,
$$B = BI_n = B(AC) = (BA)C = I_nC = C.$$
Es decir, A solo puede tener una inversa. □

Como la inversa de una matriz es única podemos representar la inversa de la matriz A, si existe, con A^{-1}. Con ello, podemos escribir lo siguiente:
$$AA^{-1} = A^{-1}A = I_n.$$

Teorema 1.11 (Propiedades de la inversa) Sean A y B matrices de orden n. Entonces:

1. Si la matriz A es inversible, entonces A^{-1} es inversible y
$$(A^{-1})^{-1} = A.$$
2. Si las matrices A y B son inversibles, entonces AB es inversible y
$$(AB)^{-1} = B^{-1}A^{-1}.$$

Demostración. Sea A y B dos matrices de orden n.

1. Si la matriz A es inversible, se cumple que
$$AA^{-1} = A^{-1}A = I_n.$$
Estas igualdades se pueden escribirse de la siguiente manera:
$$A^{-1}A = AA^{-1} = I_,$$
de donde concluimos que A^{-1} es inversible y que su inversa es
$$(A^{-1})^{-1} = A.$$

2. Tenemos que
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$
y que
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n,$$
de donde
$$(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I_n.$$
Esta igualdad muestra que la matriz AB es inversible y que
$$(AB)^{-1} = B^{-1}A^{-1}.$$

□

Mediante inducción sobre el número de matrices, se obtiene el siguiente corolario.

Corolario 1.12 El producto de más de dos matrices inversibles del mismo orden es inversible.

Para demostrar que la matriz B es la inversa de la matriz A, usando la definición de inversa, es necesario demostrar que $AB = I_n$ y que $BA = I_n$. Podemos reducir el procedimiento de probar que B es la inversa de A con tan solo demostrar una de las dos igualdades. Enunciamos esto en el siguiente teorema que no será demostrado, pues requiere de otros resultados que no estudiaremos aquí.

Teorema 1.13 Sean A y B matrices cuadradas de orden n e inversibles. Entonces, si $AB = I_n$, entonces $BA = I_n$.

Teorema 1.14 Una matriz elemental es inversible.

Demostración. Sea D una matriz elemental y sea e la operación elemental tal que $E = e(I)$. Vamos a demostrar que E es inversible.

Por el teorema 1.4, e es una función biyectiva. Por lo tanto, tiene inversa. Sea $F = e^{-1}(I)$. Entonces F es una matriz elemental. Probemos que esta matriz es la matriz inversa de E. Para ello, es suficiente con demostrar, por el teorema 1.13, que $EF = I$.

Por el teorema 1.8, sabemos que $EF = e(F)$. Por lo tanto:
$$EF = e(F) = e(e^{-1}(I)) = ee^{-1}I = I.$$

□

Ejemplo 17 Se tiene la matriz elemental E_1:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{3F_1} \sim \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = E_1.$$

La inversa de la matriz E_1 es E_1^{-1}:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{\frac{1}{3}F_1} \sim \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix} = E_1^{-1}.$$

En efecto:

$$E_1 E_1^{-1} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

♦

Teorema 1.15 Sea la matriz $A \in M_n$. Los siguientes enunciados son equivalentes:

1. La matriz A es inversible.
2. La matriz A es equivalente por filas a la matriz identidad I_n.
3. La matriz A es el producto de matrices elementales de orden n.

Demostración. El teorema 1.6 nos garantiza la existencia de una matriz R escalonada reducida por filas equivalente por filas a la matriz A. Esto significa que existe una sucesión finita de operaciones elementales $e_1, e_2, \ldots, e_p$ y sus correspondientes matrices elementales $E_1, E_2, \ldots, E_p$ tales que

$$R = e_p \ldots e_2 e_1 A = E_p \cdots E_2 E_1 A. \tag{1.3}$$

Puesto que cada E_i es invertible (teorema 1.4), la última igualdad es equivalente a la siguiente:

$$A = E_1^{-1} E_2^{-1} \cdots E_p^{-1} R. \tag{1.4}$$

Ahora bien, supongamos que A. Entonces, como el producto de matrices inversibles es inversible, de la igualdad 1.3 concluimos que la matriz E también es inversible. Pero esto significa que E no puede tener filas nulas, porque de lo contrario, de la igualdad 1.4 colegiríamos que A tiene filas nulas, lo que es imposible, pues A es inversible. Entonces

$$R = I.$$

De la igualdad (1.4) también se concluye que A es el producto de matrices elementales. De (1.3), que A es equivalente por filas a la identidad.

En resumen, queda demostrado que:

(a) Si A es inversible entonces: $A \sim I$, y A es el producto de matrices elementales.

(b) Si $A \sim I$ ($R = I$ en (1.3)), entonces A es el producto de matrices elementales y por lo tanto A es inversible.

(c) Si A es el producto de matrices elementales, en (1.4) se puede tomar $R = I$, de donde se tiene que $A \sim I$ y por lo tanto es inversible.

□

En otras palabras, el teorema nos da las condiciones necesarias y suficientes para que una matriz A sea inversible. Como toda matriz A es equivalente por filas a una matriz R escalonada reducida por filas, si A es inversible, entonces R debe ser igual a la matriz identidad I.

Corolario 1.16 Sea $A \in M_n$ una matriz inversible. Si una sucesión finita de operaciones elementales de fila reducen la matriz A a la matriz identidad, entonces la sucesión de las operaciones elementales de fila que fueron aplicadas a A, pero ahora aplicadas a la matriz identidad, dan como resultado la matriz A^{-1}, la inversa de A.

Demostración. Como A inversible, es equivalente por filas a la matriz identidad: $A \sim I$. Es decir, existen operaciones elementales e_j tales que

$$I = e_p \cdots e_2 e_1 A.$$

Si E_i son las correspondientes matrices elementales para e_i, entonces:

$$I = E_p \cdots E_2 E_1 A,$$

de donde obtenemos que

$$A = E_1^{-1} E_2^{-1} \cdots E_p^{-1} I = e_1^{-1} e_2^{-1} \cdots e_p^{-1} I,$$

que es lo que se quería demostrar. □

Este corolario nos indica la forma como hallar la matriz inversa de una matriz A inversible, por medio de operaciones elementales de fila. Para ello construimos la matriz ampliada $(A \mid I_n)$, luego realizamos sobre ella una secuencia de operaciones elementales de fila de modo que obtengamos la matriz ampliada $(I_n \mid A^{-1})$:

$$(A \mid I_n) \sim \cdots \sim (I_n \mid A^{-1}).$$

En caso de no saber con anterioridad que la matriz A es inversible, el mismo procedimiento nos puede decir si la matriz A es o no inversible. Para ello, construimos la matriz ampliada $(A \mid I)$, hallamos la matriz escalonada reducida por filas de la matriz ampliada:

$$(A \mid I_n) \sim \cdots \sim (R \mid P).$$

Si $R \neq I$, entonces la matriz A no es inversible. Si $R = I_n$, entonces la matriz A es inversible y su inversa es $A^{-1} = P$. Veamos la aplicación de este procedimiento en el siguiente ejemplo.

Ejemplo 1.6

Sea la matriz

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix}.$$

Si la matriz A es inversible, encontrar A^{-1}.

Solución. Construimos la matriz ampliada $A \mid I$ y hallamos su matriz escalonada reducida por filas:

$$\begin{aligned}
(A|I_3) = &\begin{pmatrix} 1 & 0 & 1 & | & 1 & 0 & 0 \\ 0 & -1 & 1 & | & 0 & 1 & 0 \\ -1 & 1 & -1 & | & 0 & 0 & 1 \end{pmatrix}_{F_3+F_1} \sim \begin{pmatrix} 1 & 0 & 1 & | & 1 & 0 & 0 \\ 0 & -1 & 1 & | & 0 & 1 & 0 \\ 0 & 1 & 0 & | & 1 & 0 & 1 \end{pmatrix}_{F_2+F_3} \\
\sim &\begin{pmatrix} 1 & 0 & 1 & | & 1 & 0 & 0 \\ 0 & 0 & 1 & | & 1 & 1 & 1 \\ 0 & 1 & 0 & | & 1 & 0 & 1 \end{pmatrix}_{F_1-F_2} \sim \begin{pmatrix} 1 & 0 & 0 & | & 0 & -1 & -1 \\ 0 & 0 & 1 & | & 1 & 1 & 1 \\ 0 & 1 & 0 & | & 1 & 0 & 1 \end{pmatrix}_{F_2 \rightleftharpoons F_3} \\
\sim &\begin{pmatrix} 1 & 0 & 0 & | & 0 & -1 & -1 \\ 0 & 1 & 0 & | & 1 & 0 & 1 \\ 0 & 0 & 1 & | & 1 & 1 & 1 \end{pmatrix} = (I_3|A^{-1}).
\end{aligned}$$

Por lo tanto, la matriz A es inversible y su inversa es

$$A^{-1} = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Verifiquemos que la matriz A^{-1} es la inversa de la matriz A:

$$AA^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$$

y

$$A^{-1}A = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

♦

Corolario 1.17 Sean $A, B \in M_{m\times n}$. La matriz B es equivalente por filas a la matriz A si y solo si $B = PA$, donde P es una matriz inversible de orden m.

Demostración. $B \sim A$ es equivalente a la existencia de matrices elementales E_i tales que $B = E_p \cdots E_2E_1A$. Si $P = E_p \cdots E_2E_1$, entonces P es inversible.

Esto significa que la misma secuencia de operaciones elementales de fila que aplicadas a A producen B, al ser aplicada a I_m produce P. □

Para ejemplos prácticos, se utiliza el siguiente esquema:

$$(A \mid I_m) \sim \cdots \sim (B \mid P)$$

Ejemplo 1.7

Sean las matrices

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ -1 & -2 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Demostrar que las matrices A y B son equivalentes por filas.

Solución. Encontremos una matriz inversible P tal que $B = PA$:

$$(A \mid I_3) = \begin{pmatrix} 1 & 0 & 1 & \mid & 1 & 0 & 0 \\ 0 & -1 & 1 & \mid & 0 & 1 & 0 \\ -1 & -2 & 1 & \mid & 0 & 0 & 1 \end{pmatrix}_{F_3+F_1} \sim \begin{pmatrix} 1 & 0 & 1 & \mid & 1 & 0 & 0 \\ 0 & -1 & 1 & \mid & 0 & 1 & 0 \\ 0 & -2 & 2 & \mid & 1 & 0 & 1 \end{pmatrix}_{-F_2}$$
$$\sim \begin{pmatrix} 1 & 0 & 1 & \mid & 1 & 0 & 0 \\ 0 & 1 & -1 & \mid & 0 & -1 & 0 \\ 0 & -2 & 2 & \mid & 1 & 0 & 1 \end{pmatrix}_{F_3-2F_2} \sim \begin{pmatrix} 1 & 0 & 1 & \mid & 1 & 0 & 0 \\ 0 & 1 & -1 & \mid & 0 & -1 & 0 \\ 0 & 0 & 0 & \mid & 1 & -2 & 1 \end{pmatrix} = (B \mid P).$$

La matriz B es equivalente por filas a la matriz A. La matriz P es inversible pues es equivalente por filas a la matriz identidad I_3. Hallemos la matriz PA:

$$PA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ -1 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = B.$$

♦

Para saber si dos matrices A y B son o no equivalentes por filas, no siempre se puede llegar fácilmente de A por medio de una secuencia de operaciones elementales de fila a B. En ese caso se procede de la siguiente manera: se hallan las matrices escalonadas reducidas por fila de A y B, R_A y R_B, respectivamente. Por el corolario anterior existen matrices P_A y P_B tales que

$$R_A = P_A A \quad \text{y} \quad R_B = P_B B.$$

Si $P_A \neq P_B$, entonces la matriz A no es equivalente por filas a la matriz B. Pero si $R_A = R_B$, entonces las matrices A y B son equivalentes por filas, y se puede hallar la matriz P inversible tal que $B = PA$:

$$\begin{aligned} P_B B &= P_A A, \\ P_B^{-1} P_B B &= P_B^{-1} P_A A, \\ IB &= P_B^{-1} P_A A, \\ B &= PA, \end{aligned}$$

con

$$P = P_B^{-1} P_A.$$

1.6 Otras matrices

Definición 1.24 (Matriz nilpotente de orden r) Una matriz $A \in M_n$ se denomina **nilpotente de orden** r si r es el menor entero positivo tal que $A^r = O_n$, donde $A^r = \overbrace{A \cdots A}^{r \text{ veces}}$ y O_n es la matriz nula de orden n.

Ejemplo 1.8

Sea la matriz
$$A = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$
Mostrar que es nilpotente y encontrar el orden.

Solución. Multipliquemos la matriz A por sí misma el número de veces que sea necesario hasta obtener la matriz nula O_3:
$$A^2 = AA = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$A^3 = A^2A = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = O_3.$$
Es evidente que para $n > 3$, $A^n = O_3$. Por lo tanto, la matriz A es nilpotente de orden 3. ♦

Se asume que las matrices cuadradas nulas son nilpotentes de orden 1.

Definición 1.25 (Matriz idempotente) Una matriz $A \in M_n$ se denomina **idempotente** si $A^2 = A$.

Ejemplo 18 Las matrices
$$I_n, \quad O_n, \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$
son matrices idempotentes. ♦

Definición 1.26 (Matriz involutiva) Una matriz $A \in M_n$ se denomina **involutiva** si $A^2 = I_n$.

Ejemplo 19 Las matrices
$$I_n, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
son matrices involutivas. ♦

De la definición de matriz involutiva y del teorema 1.13 se sigue si $A \in M_n$ es involutiva, A es inversible y $A^{-1} = A$.

Definición 1.27 (Matriz ortogonal) Una matriz $A \in M_n$ se denomina **ortogonal** si $A^t A = I_n$.

Ejemplo 20 Las matrices

$$I_n, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

son matrices ortogonales. ♦

Una vez más, del teorema 1.13, se deduce que si la matriz $A \in M_n$ es ortogonal, entonces es inversible y $A^{-1} = A^t$.

Definición 1.28 (Matriz conjugada) Sea la matriz $A = (a_{ij})_{m\times n}$. Se denomina **conjugada** de A, notada $\bar{A}$, a la matriz $\bar{A} = (\bar{a}_{ij})_{m\times n}$.

Ejemplo 21 La conjugada de la matriz

$$A = \begin{pmatrix} 1-2i & 4 \\ 1+i & 3i \\ -5i & -2 \end{pmatrix}$$

es la matriz

$$\bar{A} = \begin{pmatrix} 1+2i & 4 \\ 1-i & -3i \\ 5i & -2 \end{pmatrix}.$$

♦

Si los elementos de la matriz A son números reales, se tiene que $\bar{A} = A$.

Definición 1.29 (Matriz hermitiana) Una matriz $A \in M_n$ se denomina **hermitiana** si $A^t = \bar{A}$.

Ejemplo 1.9

Sea la matriz

$$A = \begin{pmatrix} 1 & i & 0 \\ -i & 2 & 1+2i \\ 0 & 1-2i & 3 \end{pmatrix}.$$

¿Es hermitiana?

Solución. Hallemos la transpuesta y la conjugada de la matriz A:

$$A^t = \begin{pmatrix} 1 & -i & 0 \\ i & 2 & 1-2i \\ 0 & 1+2i & 3 \end{pmatrix}, \quad \bar{A} = \begin{pmatrix} 1 & -i & 0 \\ i & 2 & 1-2i \\ 0 & 1+2i & 3 \end{pmatrix}.$$

Como $A^t = \bar{A}$, A es una matriz hermitiana. ♦

Obsérvese que si una matriz es hermitiana, los elementos de su diagonal son números reales. También que toda matriz simétrica es hermitiana.

Definición 1.30 (Matriz unitaria) Una matriz $A \in M_n$ se denomina **unitaria** si $A^t\bar{A} = I_n$.

Ejemplo 1.10

Sea la matriz
$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$
¿Es unitaria?

Solución. Hallemos la transpuesta y la conjugada de la matriz A:
$$A^t = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \bar{A} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$
Calculemos A^tA:
$$A^tA = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$
Entonces la matriz A es unitaria. ♦

Una matriz ortogonal cuyos elementos son números reales es una matriz unitaria.

Definición 1.31 (Traza) Sea la matriz $A = (a_{ij}) \in M_n(\mathbb{K})$. Se llama **traza** de A, notada $\operatorname{tr} A$, al número
$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}.$$

En otras palabras, la traza de una matriz cuadrada es la suma de sus elementos que se encuentran en la diagonal.

1.7 Ejercicios resueltos

1. Sean $A, B \in M_n$ dos matrices triangulares inferiores. Demostrar que AB es una matriz triangular inferior.

   Demostración. Sean $A = (a_{ij})_n$ y $B = (b_{ij})_n$ tales que $a_{ij} = 0$ y $b_{ij} = 0$ para todo i y para todo j con $i < j$.

   Sean $C = AB$ y $C = (c_{ij})_n$. Vamos a demostrar que $c_{ij} = 0$ para todo i y para todo j con $i < j$.

   Como la matriz C es el producto de A por B, tenemos que
$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$
   para todo i y para todo j.

   Sea $i < j$. Entonces:
$$\sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{i} a_{ik}b_{kj} + \sum_{k=i+1}^{n} a_{ik}b_{kj}.$$
   Pero,

(a) si $k < i$, se verifica que $k < j$ y $b_{kj} = 0$, pues la matriz B es triangular inferior; entonces:

$$\sum_{k=1}^{i} a_{ik} b_{kj} = 0;$$

(b) si $k > i$, se tiene que $a_{ik} = 0$, porque la matriz A es triangular inferior; entonces:

$$\sum_{k=i+1}^{i} a_{ik} b_{kj} = 0.$$

Por lo tanto, tenemos que:

$$c_{ij} = \sum_{k=1}^{i} a_{ik} b_{kj} + \sum_{k=i+1}^{n} a_{ik} b_{kj} = 0 + 0 = 0$$

si $i < j$. En otras palabras, la matriz $C = AB$ es triangular inferior. $\square$

2. Demostrar que

   (a) $(\forall A \in M_{m\times n})[(A^t)^t = A]$.

   (b) $(\forall A \in M_{m\times n})(\forall B \in M_{m\times n})[(A + B)^t = A^t + B^t]$.

   (c) $(\forall \alpha \in \mathbb{K})(\forall A \in M_{m\times n})[(\alpha A^t)^t = \alpha A^t]$.

   (d) $(\forall A \in M_{m\times n})\ (\forall B \in M_{n\times p})[(AB)^t = B^t A^t]$.

Demostración. Se demostrarán únicamente las partes (b) y (d). Las otras se dejan como ejercicio para el lector.

(b) Sean $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{m\times n}$. Entonces:

$$A^t = (\hat{a}_{ij})_{n\times m} \quad \text{y} \quad B^t = (\hat{b}_{ij})_{n\times m},$$

donde

$$\hat{a}_{ij} = a_{ji} \quad \text{y} \quad \hat{b}_{ij} = b_{ji}$$

para todo $i \in I_n = \{1, 2, \ldots, n\}$ y todo $j \in I_m = \{1, 2, \ldots, m\}$.
Por otro lado, sea $C = (c_{ij})_{m\times n} = A + B$. Esto significa que:

$$c_{ij} = a_{ij} + b_{ij}$$

para todo $i \in I_n$ y todo $j \in I_m$. Por lo tanto, $C^t = (\hat{c}_{ij})_{n\times m}$, donde:

$$\begin{aligned} \hat{c}_{ij} &= c_{ji} \\ &= a_{ji} + b_{ji} \\ &= \hat{a}_{ij} + \hat{b}_{ij} \end{aligned}$$

para todo $i \in I_n$ y $j \in I_m$. Es decir, por la definición de suma de matrices, podemos concluir que:

$$(A + B)^t = C^t = A^t + B^t.$$

(d) Sean $A = (a_{ij})_{m\times n}$, $B = (b_{jk})_{n\times p}$. Entonces:

$$A^t = (\hat{a}_{ij})_{n\times m} \quad \text{y} \quad B^t = (\hat{b}_{jk})_{p\times n},$$

donde

$$\hat{a}_{ij} = a_{ji} \quad \text{y} \quad \hat{b}_{jk} = b_{kj}$$

para todo $i \in I_p = \{1, 2, \ldots, p\}$, todo $j \in I_n = \{1, 2, \ldots, n\}$ y $k \in I_m = \{1, 2, \ldots, m\}$.

Por otro lado, sea $C = (c_{ik})_{m\times p} = AB$. Esto significa que:

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

para todo $i \in I_m$ y todo $k \in I_p$. Por lo tanto, $C^t = (\hat{c}_{ik})_{p\times m}$, donde:

$$\begin{aligned}\hat{c}_{ik} &= c_{ki}\\ &= \sum_{j=1}^{n} a_{kj} b_{ji}\\ &= \sum_{j=1}^{n} b_{ji} a_{kj} = \sum_{j=1}^{n} \hat{b}_{ij} \hat{a}_{jk}\end{aligned}$$

para todo $i \in I_p$ y todo $k \in I_m$. Entonces, por la definición de producto de matrices, tenemos que:

$$C^t = (AB)^t = B^t A^t.$$

□

3. (a) $(\forall A \in M_n)[\operatorname{tr} A^t = \operatorname{tr} A]$.
   (b) $(\forall A, B \in M_n)[\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B]$.
   (c) $(\alpha \in \mathbb{K})$ $(\forall A \in M_n)[\operatorname{tr}(\alpha A) = \alpha \operatorname{tr} A]$.
   (d) $(\forall A \in M_{m\times n})(\forall B \in M_{n\times m})[\operatorname{tr}(AB) = \operatorname{tr}(BA)]$.

Demostración. Se demostrarán las partes (c) y (d).

(c) Sean $\alpha \in \mathbb{K}$ y $A = (a_{ij})_n$. Entonces:

$$\operatorname{tr}(\alpha A) = \operatorname{tr}(\alpha(a_{ij})) = \operatorname{tr}(\alpha a_{ii}) = \sum_{i=1}^{n} \alpha a_{ii} = \alpha \sum_{i=1}^{n} a_{ii} = \alpha \operatorname{tr} A.$$

(d) Sean $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{n\times m}$ y $C = (c_{ij})_m = AB$. Entonces

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Tenemos que

$$\operatorname{tr}(AB) = \operatorname{tr} C = \sum_{i=1}^{m} c_{ii} = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{k=1}^{n} \sum_{i=1}^{m} b_{ki} a_{ik} = \operatorname{tr}(BA).$$

□

Ejemplo 1.11

Sean las matrices
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 3 & -5 \\ 4 & -2 \\ 6 & 1 \end{pmatrix}.$$
Hallar $\operatorname{tr}(AB)$ y $\operatorname{tr}(BA)$.

Solución. Encontremos los productos
$$AB = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ 4 & -2 \\ 6 & 1 \end{pmatrix} = \begin{pmatrix} 29 & -6 \\ 68 & -24 \end{pmatrix},$$
$$BA = \begin{pmatrix} 3 & -5 \\ 4 & -2 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} -17 & -19 & -21 \\ -4 & -2 & 0 \\ 10 & 17 & 24 \end{pmatrix}.$$
Entonces
$$\operatorname{tr}(AB) = 29 + (-24) = 5, \quad \operatorname{tr}(BA) = -17 + (-2) + 24 = 5.$$
Este ejemplo ilustra la propiedad $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. ♦

4. Demostrar que no existen matrices A, B y C de orden n tales que
$$ABC - CAB + I_n = O_n.$$

Demostración. Supongamos que sí existen matrices A, B y C de orden n tales que $ABC - CAB + I_n = O_n$.

Entonces
$$\operatorname{tr}(ABC - CAB + I_n) = \operatorname{tr} O_n,$$
$$\operatorname{tr}(ABC) - \operatorname{tr}(CAB) + \operatorname{tr} I = 0.$$
Como
$$\operatorname{tr}\big((AB)C\big) = \operatorname{tr}\big(C(AB)\big) = \operatorname{tr}(ABC),$$
podemos escribir:
$$\operatorname{tr}(ABC) - \operatorname{tr}(ABC) + n = 0, \quad n \geq 1,$$
de donde $n = 0$ lo cual es una falsedad. □

5. Demostrar que toda matriz cuadrada es la suma de una matriz simétrica y una matriz antisimétrica.

Demostración. Sea la matriz $A \in M_n$. Vamos a demostrar que $A = S + T$, donde S es una matriz simétrica y T es una matriz antisimétrica.

Sean
$$S = \frac{A + A^t}{2} \quad \text{y} \quad T = \frac{A - A^t}{2}.$$
Entonces la matriz S es simétrica y la matriz T es antisimétrica. La demostración se deja como ejercicio para lector. □

6. Sea A una matriz cuadrada. Demostrar que si $A^2 = A$, $A^n = A$ para todo $n \geq 2$.

Demostración. La demostración es por inducción sobre n.

i) La demostración es verdadera para $n = 2$, pues, si $A^2 = A$, entonces $A^2 = A$ es una proposición verdad.

ii) Supongamos que $k \geq$ es verdad que $A^k = A$. Es inmediato que $A^{k+1} = A$ también es verdadera, pues:

$$A^{k+1} = A^k A = AA = A^2 = A.$$

Por lo tanto, $A^n = A$ para todo $n \geq 2$. $\square$

7. Sea A una matriz cuadrada. Si $A^2 = A$, determinar una fórmula para $(A + I)^n$.

Desarrollo. Como el producto de las matrices A e I conmuta se puede aplicar el binomio de Newton:

$$\begin{aligned}
(A+I)^n &= \sum_{k=0}^{n} \binom{n}{k} I^{n-k} A^k \\
&= \binom{n}{0} A^0 + \sum_{k=1}^{n} \binom{n}{k} A^k \\
&= I + \sum_{k=1}^{n} \binom{n}{k} A \\
&= I + A \sum_{k=1}^{n} \binom{n}{k} \\
&= I + \left(\sum_{k=0}^{n} \binom{n}{k} - \binom{n}{0} \right) \\
&= I + \left(\sum_{k=0}^{n} \binom{n}{k} 1^{n-k} 1^k - 1 \right) \\
&= I + A\big((1+1)^n - 1\big) \\
&= I + (2^n - 1)A.
\end{aligned}$$

Es decir,

$$(A + I)^n = I * (2^n - 1)A \quad \text{siempre que} \quad A^2 = A.$$

8. Sea

$$G = \left\{ \begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix} : a \neq b; \quad a, b \in \mathbb{R} \right\}.$$

¿Cuál de las siguientes proposiciones es verdadera y cuál es falsa?

(a) $(\forall A \in G)(\forall B \in G)(AB \in G)$.

(b) $(\forall A \in G)(\forall B \in G)(\forall C \in G)[(AB)C = A(BC)]$.

(c) $(\exists E \in G)(\forall A \in G)[AE = EA = A]$.

(d) $(\forall A \in G)(\exists B \in G)[AB = BA = I]$.

(e) $(\forall A \in G)(\forall B \in G)(AB = BA)$.

Desarrollo.

(a) Sean las matrices

$$A = \begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix}$$

donde $a \neq b$ y $p \neq q$.
Hallemos el producto

$$\begin{aligned} AB &= \begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix} \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \\ &= \begin{pmatrix} ap + q(1-a) & a(1-p) + (1-a)(1-q) \\ bp + q(1-b) & b(1-p) + (1-b)(1-q) \end{pmatrix} \\ &= \begin{pmatrix} ap + q - aq & 1 - (ap + q - aq) \\ bp + q - bq & 1 - (bp + q - bq) \end{pmatrix} \\ &= \begin{pmatrix} c & 1-c \\ d & 1-d \end{pmatrix}, \end{aligned}$$

donde

$$c = ap + q - aq \quad \text{y} \quad d = bp + q - bq$$

Si $c \neq d$, el producto AB estaría en G. Veamos si es así. Si fuera cierto que $c = d$, entonces:

$$\begin{aligned} ap + q - aq &= bp + q - bq, \\ ap - bp &= aq - bq, \\ (a-b)p &= (a-b)q. \end{aligned}$$

Como $a \neq b$, $p = q$, lo cual es una contradicción ya que $p \neq q$. Por lo tanto $c \neq d$. Se concluye, entonces, que $AB \in G$.

(b) El producto de matrices tiene la propiedad asociativa. Luego esta proposición es evidentemente verdadera, independientemente de que las matrices A, B y C sean elementos de G.

(c) Sea E la matriz identidad de orden 2:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1-1 & 1-0 \end{pmatrix}.$$

Como $1 \neq 0$, $E \in G$. Además, $AE = EA = A$. Entonces esta proposición es verdadera.

(d) Si esta proposición fuera verdadera, la matriz B debería ser la inversa de $A \in G$. Se deja al lector constatar que si

$$A = \begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix},$$

A es inversible y que:

$$A^{-1} = \begin{pmatrix} \frac{1-b}{a-b} & \frac{a-1}{a-b} \\ \frac{-b}{a-b} & \frac{a}{a-b} \end{pmatrix}.$$

Finalmente, veamos si $A^{-1} \in G$. Para ello, reescribamos A^{-1} de la siguiente manera:

$$A^{-1} = \begin{pmatrix} \frac{1-b}{a-b} & 1 - \frac{1-b}{a-b} \\ \frac{-b}{a-b} & 1 - \frac{-b}{a-b} \end{pmatrix}.$$

Si fuera cierto que

$$\frac{1-b}{a-b} = \frac{-b}{a-b},$$

tendríamos que $1 - b = -b$, lo que equivale a $1 = 0$. Por lo tanto, $A^{-1} \in G$, y esta proposición es verdadera.

(e) Demostremos con un contraejemplo que esta proposición es falsa. Sean las matrices

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

que pertenecen al conjunto G. Hallemos los productos AB y BA:

$$AB = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix},$$

$$BA = \begin{pmatrix} 1 & 0 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 6 & -5 \end{pmatrix}.$$

Es decir: $AB \neq BA$.

El conjunto G cumple las propiedades (a), (b), (c) y (d). Por lo tanto, el conjunto G con la operación producto de matrices constituye un grupo no abeliano, pues no tiene la propiedad conmutativa.

9. (a) Sea $A \in M_n$ y nilpotente de orden m. Entonces la matriz $(I - A)$ es inversible y se tiene que

$$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots + A^{m-1}.$$

Demostración. Tenemos que

$$I^m - A^m = (I - A)(I^{m-1} + I^{m-2}A + I^{m-3}A^2 + \cdots + IA^{m-2} + A^{m-1}).$$

Pero $A^m = O_m$, por lo cual

$$(I - A)(I + A + A^2 + \cdots + A^{m-2} + A^{m-1}) = I,$$

de donde concluimos que $(I - A)$ es inversible y que

$$(I - A)^{-1} = (I + A + A^2 + \cdots + A^{m-2} + A^{m-1}).$$

(b) Apliquemos el resultado anterior para hallar la inversa de la matriz

$$B = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Desarrollo. La matriz B la podemos escribir como

$$B = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = I - A,$$

con

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

Por otro lado, tenemos que:

$$A^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Entonces la matriz A es nilpotente de orden 3. Luego podemos aplicar el resultado anterior:

$$\begin{aligned} B^{-1} = (I - A)^{-1} &= I + A + A^2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

10. Demostrar que si $A \in M_n$ es simétrica e inversible, entonces A^{-1} es simétrica.

Demostración. Si la matriz A es inversible, entonces $AA^{-1} = I$. Podemos escribir

$$\begin{aligned} (AA^{-1})^t &= I^t, \\ (A^{-1})^t A^t &= I, \\ (A^{-1})^t A &= I. \end{aligned}$$

La última expresión significa que la matriz $(A^{-1})^t$ es la inversa de la matriz A, es decir,

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

lo cual, a su vez, significa que A^{-1} es una matriz simétrica. □

11. Dada la matriz

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

determinar A^n.

Desarrollo. Encontremos A^n. Para ello, empecemos con A^2:

$$A^2 = \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix} = 3A.$$

Entonces:

$$A^3 = A^2A = 3AA = 3A^2 = 3(3A) = 3^2A.$$

Conjeturamos que $A^n = 3^{n-1}A$. Demostremos por inducción sobre n que esta fórmula es efectivamente verdadera.

i) Si $n = 1$,

$$A^1 = A = 3^0A.$$

ES decir, la fórmula es verdadera.

ii) Si es verdadera para n, demostremos que también lo es para $n+1$:

$$A^{n+1} = A^nA = (3^{n-1}A)A = 3^{n-1}A^2 = 3^{n-1}(3A) = 3^nA.$$

Por lo tanto, si $A = \left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{smallmatrix}\right)$, entonces $A^n = 3^{n-1}A$ para todo $n \in \mathbb{N}$.

12. Dada la matriz

$$A = \begin{pmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{pmatrix},$$

determinar A^n.

Desarrollo. En primer lugar, tenemos que $A = aI + B$ con

$$B = \begin{pmatrix} 0 & b & c \\ 0 & 0 & d \\ 0 & 0 & 0 \end{pmatrix}.$$

Tenemos que

$$B^2 = \begin{pmatrix} 0 & 0 & bd \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B^3 = O.$$

Si aplicamos el binomio de Newton, obtenemos que

$$\begin{aligned} A^n &= (aI + B)^n \\ &= \sum_{k=0}^{n} \binom{n}{k} (aI)^{n-k} B^k \\ &= a^nI + \binom{n}{1} a^{n-1}B + \binom{n}{2} a^{n-2}B^2 + \binom{n}{3} a^{n-3}B^3 + O + \cdots + O \\ &= a^nI + na^{n-1}B + \frac{n(n-1)}{2} a^{n-2}B^2. \end{aligned}$$

Reemplazamos la matriz B:

$$\begin{aligned} A^n &= a^n \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + na^{n-1} \begin{pmatrix} 0 & b & c \\ 0 & 0 & d \\ 0 & 0 & 0 \end{pmatrix} + \frac{n(n-1)}{2} a^{n-2} \begin{pmatrix} 0 & 0 & bd \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a^n & na^{n-1}b & \frac{n(n-1)}{2}a^{n-2}bd \\ 0 & a^n & na^{n-1}d \\ 0 & 0 & a^n \end{pmatrix}. \end{aligned}$$

13. Sean las matrices

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Determinar

(a) AB y BA,

(b) $A^n + B^n$.

Desarrollo.

(a) Calculemos lo pedido:

$$AB = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = O,$$

$$BA = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = O.$$

(b) Como $AB = BA = O$ se puede utilizar el binomio de Newton:

$$\begin{aligned} (A+B)^n &= \sum_{k=0}^{n} \binom{n}{k} A^{n-k} B^k \\ &= A^n + \sum_{k=1}^{n-1} \binom{n}{k} A^{n-k} B^k + B^n \\ &= A^n + O + B^n \\ &= A^n + B^n. \end{aligned}$$

Entonces tenemos que

$$A^n + B^n = (A+B)^n.$$

Como

$$(A+B) = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = 3I,$$

entonces

$$A^n + B^n = (3I)^n = 3^n I^n = 3^n I.$$

Determinantes | 2

2.1 Definición

Definición 2.1 (Determinante) El **determinante** de una matriz $A = (a_{ij}) \in M_n(\mathbb{K})$, notado $\det A$ o $|A|$ es la función

$$\begin{aligned} \det\colon M_n &\longrightarrow \mathbb{K} \\ A &\longmapsto \det A \end{aligned}$$

definida por:

1. Si $n = 1$, $\det(a_{11}) = a_{11}$.

2. Si $n > 1$, las siguientes son dos formas equivalentes de definir $\det(A)$:

   (a) **El desarrollo del determinante por menores por la r-ésima fila de $A = (a_{ij})_n$:**

   $$|A| = \sum_{j=1}^{n} (-1)^{r+j} a_{rj} |A_r^j|,$$

   donde A_r^j es la matriz de orden $n-1$ que resulta de quitar de la matriz A la fila r y la columna j para $j = 1, 2, \ldots n$.

   (b) **El desarrollo del determinante por menores por la s-ésima columna de $A = (a_{ij})_n$:**

   $$|A| = \sum_{i=1}^{n} (-1)^{i+s} a_{is} |A_i^s|,$$

   donde A_i^s es la matriz de orden $n-1$ que resulta de quitar de la matriz A la fila i y la columna s para $i = 1, 2, \ldots n$.

Observaciones

1. Por ser det una función, $\det(A)$ arroja un valor único, por lo que sea que se utilice el desarrollo por menores por una fila o por una columna, se obtendrá el mismo valor.

2. La definición de determinante ofrecida es recursiva, pues se define el determinante de una matriz de orden n a partir de los determinantes de algunas matrices de orden $n-1$. Por ejemplo, para calcular el determinante de una matriz cuadrada de orden 4, se deben calcular 4 determinantes de matrices de orden 3 previamente: A_1^1, A_1^2, A_1^3 y A_1^4 (si se utilizó el desarrollo por menores por la primera fila). A su vez, para calcular cada uno de los determinantes de las matrices de orden 3, se deberán calcular 3 determinantes de matrices de orden 2 previamente. Y, finalmente, para calcular cada uno de los determinantes de las matrices de orden 2, se deberá calcular un determinante de una las matrices de orden 1.

Ejemplo 2.12

Sea

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Hallar $|A|$.

Solución. El desarrollo del determinante $|A|$ por menores, por la segunda fila es el siguiente:

$$\begin{aligned} |A| &= \sum_{j=1}^{2} (-1)^{2j} a_{2j} |A_2^j| \\ &= (-1)^{2+1} a_{21} |A_2^1| + (-1)^{2+2} a_{22} |A_2^2| \\ &= (-1)^3 a_{21} |a_{12}| + (-1)^4 a_{22} |a_{11}| \\ &= -a_{21} a_{12} + a_{22} a_{11} \\ &= a_{11} a_{22} - a_{21} a_{12}. \end{aligned}$$

En cambio, el desarrollo del determinante $|A|$ por menores, por la primera columna es el siguiente:

$$\begin{aligned} |A| &= \sum_{i=1}^{2} (-1)^{i+1} a_{i1} |A_i^1| \\ &= (-1)^{1+1} a_{11} |A_1^1| + (-1)^{2+1} a_{21} |A_2^1| \\ &= (-1)^2 a_{11} |a_{22}| + (-1)^3 a_{21} |a_{12}| \\ &= a_{11} a_{22} - a_{21} a_{12}, \end{aligned}$$

que es igual al valor calculado anteriormente. ♦

Ejemplo 2.13

Hallar el valor del determinante

$$\begin{vmatrix} -1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & -3 & -4 \end{vmatrix}.$$

Solución. Hagamos el desarrollo del determinante por menores por la segunda columna:

$$\begin{aligned}
\begin{vmatrix} -1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & -3 & -4 \end{vmatrix} &= (-1)^{(1+2)}2\begin{vmatrix} 4 & 6 \\ 2 & -4 \end{vmatrix} + (-1)^{(2+2)}5\begin{vmatrix} -1 & 3 \\ 2 & -4 \end{vmatrix} + (-1)^{(3+2)}(-3)\begin{vmatrix} -1 & 3 \\ 4 & 6 \end{vmatrix} \\
&= -2\Big(4(-4) - 2\cdot 6\Big) + 5\Big((-1)(-4) - 2\cdot 3\Big) - (-3)\Big((-1)6 - 4\cdot 3)\Big) \\
&= -2(-28) + 5\cdot(-2) - (-3)(-18) \\
&= 56 + (-10) - 54 \\
&= -8.
\end{aligned}$$

♦

Para agilizar el cálculo de un determinante, es preferible hacer el desarrollo por menores por la fila o la columna que tenga mayor cantidad de ceros, tal como se ilustra en el siguiente ejemplo.

Ejemplo 2.14

Calcular el valor del determinante

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

Solución. En este caso conviene hacer el desarrollo del determinante por menores por la 3-ésima columna:

$$\begin{aligned}
\begin{vmatrix} -1 & 2 & 0 & 3 \\ 2 & -1 & 0 & 1 \\ 1 & 2 & 3 & 1 \\ 4 & 1 & 0 & -2 \end{vmatrix} &= 0 + 0 + (-1)^{(3+3)}3\begin{vmatrix} -1 & 2 & 3 \\ 2 & -1 & 1 \\ 4 & 1 & -2 \end{vmatrix} + 0 \\
&= 3\Big(-(2-1) - 2(-4-4) + 3(2+4)\Big) \\
&= 3(-1 + 16 + 18) \\
&= 99.
\end{aligned}$$

Como puede observarse, si se hubiera elegido una columna diferente o cualesquiera de las filas, el número de operaciones utilizadas para calcular el determinado habría sido mayor significativamente. ♦

Si el orden del determinante es grande y si, además, la mayoría de los elementos de la matriz son diferentes de cero, el cálculo del determinante haciendo uso de la definición puede ser tedioso y propenso a error. No solo eso, el tiempo que, incluso una computadora invierta en el cálculo, podría superar lo aceptable para este cálculo (horas, días). Por ello conviene estudiar las propiedades de los determinantes, las cuales facilitan enormemente su cálculo.

2.2 Propiedades

Teorema 2.1 Si $A = (a_{ij})_n$ es una matriz triangular inferior o superior, entonces

$$|A| = \prod_{i=1}^{n} a_{ii} = a_{11}a_{22}\cdots a_{nn}$$

En otras palabras, el determinante de una matriz triangular inferior o superior es igual al producto de los elementos de su diagonal.

Demostración. Es inmediata si para el cálculo de $\det(A)$ se elige el desarrollo por menores por la primera fila o columna, según A sea triangular inferior o superior. En ese caso, solo hay que calcular un determinante de orden $n-1$, cuyo valor se multiplicará por el primer elemento de la diagonal: a_{11}. La matriz de orden $n-1$ cuyo determinante se calculará, también es una triangular inferior o superior.

□

El carácter recursivo de la definición de determinante, permite utilizar la recursión para la demostración de varias de sus propiedades, como la que acabamos de realizar. Luego del ejemplo, el lector podrá aplicar el método para demostrar otra propiedad de los determinantes.

Ejemplo 2.15

$$\begin{vmatrix} -4 & 0 & 0 \\ 7 & 2 & 0 \\ -34 & 8 & -1 \end{vmatrix} = (-4)\cdot 2\cdot(-1) = 8.$$

♦

Teorema 2.2 Sean las matrices $A = (a_{ij})_n$ y $B = (b_{ij})_n$ tales que sus filas son iguales excepto la fila r-ésima. Sea la matriz $C = (c_{ij})_n$ tal que $C_i = A_i = B_i$ para todo i distinto de r y $C_r = A_r + B_r$. Entonces, $|C| = |A| + |B|$.

Demostración. Calculemos $\det(C)$ mediante el desarrollo por menores de la r-ésima fila:

$$\det(C) = \sum_{j=1}^{n} (-1)^{r+j} c_{rj} \det(C_r^j).$$

Por la definición de C, tenemos que $c_{rj} = a_{rj} + b_{rj}$, por lo tanto:

$$\det(C) = \sum_{j=1}^{n} (-1)^{r+j} a_{rj} \det(C_r^j) + \sum_{j=1}^{n} (-1)^{r+j} b_{rj} \det(C_r^j).$$

Por otro lado, las matrices C_r^j, A_r^j y B_r^j son iguales entre sí, porque la única diferencia entre las matrices A, B y C es la fila r-ésima que, al ser eliminada de cada una, producen las matrices C_r^j, A_r^j y B_r^j respectivamente. Por lo tanto, tenemos que:

$$\det(C) = \sum_{j=1}^{n}(-1)^{r+j}a_{rj}\det(A_r^j) + \sum_{j=1}^{n}(-1)^{r+j}b_{rj}\det(B_r^j) = \det(A) + \det(B).$$

$\square$

Ejemplo 2.16

Sean las matrices

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 6 & 8 \\ 6 & 8 & 10 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & -2 & 3 \\ -4 & -8 & -6 \\ 6 & 8 & 10 \end{pmatrix}.$$

Utilice el teorema anterior para hallar $|A| + |B|$.

Solución. Las matrices A y B tienen iguales sus filas, excepto la segunda. Escribamos una matriz C de orden 3 tal que su primera y tercera filas sean iguales a la primera y tercera filas de A y B, respectivamente, y su segunda fila sea igual a la suma de las segundas filas de A y B:

$$C = \begin{pmatrix} 1 & -2 & 3 \\ 0 & -2 & 2 \\ 6 & 8 & 10 \end{pmatrix}.$$

Las matrices A, B y C cumplen con las condiciones del teorema y podemos escribir:

$$\begin{aligned} |A| + |B| = |C| &= \begin{vmatrix} 1 & -2 & 3 \\ 0 & -2 & 2 \\ 6 & 8 & 10 \end{vmatrix} = -2\begin{vmatrix} 1 & 3 \\ 6 & 10 \end{vmatrix} \\ &= -2\begin{vmatrix} 1 & -2 \\ 6 & 8 \end{vmatrix} = -2(10-18) - 2(8+12) = 16 - 40 = -24. \end{aligned}$$

Verifiquemos el resultado hallando los determinantes de las matrices A y B y luego sumándolos:

$$|A| = \begin{vmatrix} 1 & -2 & 3 \\ 4 & 6 & 10 \\ 6 & 8 & 8 \end{vmatrix} = -32, \quad |B| = \begin{vmatrix} 1 & -2 & 3 \\ -4 & -8 & -6 \\ 6 & 8 & 10 \end{vmatrix} = 8,$$

$$|A| + |B| = -32 + 8 = -24 = |C|.$$

♦

Teorema 2.3 Sean las matrices $A = (a_{ij})_n$ y $B = (b_{ij})_n$ tales que $B_i = A_i$ para todo $i \neq r$ y $B_r = \alpha A_r$, con $\alpha \in \mathbb{K}$. Entonces, $|B| = \alpha|A|$.

Esta propiedad nos indica que si existe un múltiplo de alguna fila este múltiplo se puede sacar fuera del determinante. La propiedad también es válida para cuando existe un múltiplo de una columna.

Demostración. Para calcular $|B|$, utilicemos el desarrollo por menores de la fila r. Entonces:

$$\begin{aligned}|B| &= \sum_{j=1}^{n}(-1)^{r+j}b_{rj}|B_r^j| = \sum_{j=1}^{n}(-1)^{r+j}(\alpha a_{rj})|A_r^j| \\ &= \alpha\left(\sum_{j=1}^{n}(-1)^{r+j}a_{rj}|A_r^j|\right) = \alpha|A|.\end{aligned}$$

Por lo tanto, $|B| = \alpha|A|$. □

Ejemplo 2.17

Sea la matriz

$$A = \begin{pmatrix}1 & 3 & 3\\ 4 & 12 & 4\\ 2 & 6 & 1\end{pmatrix}.$$

Utilice la propiedad anterior para calcular el determinante de la matriz A.

Solución.

$$|A| = \begin{vmatrix}-1 & 3 & 3\\ 4 & 12 & 4\\ 2 & 6 & 1\end{vmatrix} = 4\begin{vmatrix}-1 & 3 & 3\\ 1 & 3 & 1\\ 2 & 6 & 1\end{vmatrix} = 4\cdot 3\begin{vmatrix}-1 & 1 & 3\\ 1 & 1 & 1\\ 2 & 2 & 1\end{vmatrix} = 12(-(-1)-(-1)+0) = 24.$$

Verifiquemos el resultado con un cálculo directo:

$$|A| = \begin{vmatrix}-1 & 3 & 3\\ 4 & 12 & 4\\ 2 & 6 & 1\end{vmatrix} = -(12-24) - 3(4-8) + 3(24-24) = 12 + 12 = 24.$$

♦

Teorema 2.4 Sea la matriz $A = (a_{ij})_n$ tal que alguna fila o columna es nula. Entonces $|A| = 0$.

Demostración. Supongamos que la fila r-ésima de la matriz A es nula. Si desarrollamos $|A|$ por menores sobre la fila r, tenemos que:

$$|A| = \sum_{j=1}^{n}(-1)^{r+j}a_{rj}|A_r^j| = \sum_{j=1}^{n}(-1)^{r+j}0|A_r^j| = 0,$$

pues $a_{rj} = 0$ para todo $j \in I_m$, pues A_r es nula. □

La siguiente propiedad afirma que si existe un múltiplo de todos los elementos de una matriz, el determinante de la matriz es el producto de dicho múltiplo elevado a la potencia n y el determinante de la matriz obtenida al dividir cada uno de los elementos de la matriz original por el múltiplo. En la jerga matemática, se suele decir que el múltiplo "ha sido sacado n veces como factor fuera del determinante de la matriz".

Teorema 2.5 Sean la matriz $A = (a_{ij})_n$ y $\alpha \in \mathbb{K}$. Entonces $|\alpha A| = \alpha^n |A|$.

Demostración. Utilicemos la inducción matemática sobre n, el orden de la matriz A.

1. Si $n = 1$, entonces $A = (a_{11})$. Por lo tanto, $\alpha A = (\alpha a_{11})$. Luego, por la definición de determinante, tenemos que:

$$|\alpha A| = \alpha a_{11} = \alpha |A|.$$

2. Ahora supongamos que $|\alpha A| = \alpha^k |A|$ para $k \geq 1$. Vamos a probar que si el orden de A es $k+1$, entonces $|\alpha A| = \alpha^{k+1}|A|$.

   La primera fila de la matriz αA es

$$\begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \end{pmatrix}.$$

   Además, cada fila de la matriz αA es el producto de α por la fila correspondiente de A. Entonces, si desarrollamos el determinante de A por menores respecto a la primera fila, obtendremos lo siguiente:

$$\begin{aligned} |\alpha A| &= \sum_{j=1}^{k+1} (-1)^{1+j} (\alpha a_{1j}) |\alpha A_1^j| \\ &= \alpha \sum_{j=1}^{k+1} (-1)^{1+j} a_{1j} |\alpha A_1^j|. \end{aligned}$$

   Como A_1^j es una matriz de orden k, entonces, por la hipótesis de inducción, tenemos que

$$|\alpha A_1^j| = \alpha^k |A_1^j|.$$

   Por lo tanto:

$$|\alpha A| = \alpha \sum_{j=1}^{k+1} (-1)^{1+j} a_{1j} \alpha^k |A_1^j| = \alpha^{k+1} \sum_{j=1}^{k+1} (-1)^{1+j} a_{1j} |A_1^j| = \alpha^{k+1}|A|,$$

   como se quería demostrar.

□

Ejemplo 2.18

Sea la matriz

$$A = \begin{pmatrix} 8 & 4 \\ 20 & 24 \end{pmatrix}.$$

Use la propiedad anterior para calcular el determinante de la matriz A.

Solución.

$$|A| = \begin{vmatrix} 8 & 4 \\ 20 & 24 \end{vmatrix} = \begin{vmatrix} 4\cdot 2 & 4\cdot 1 \\ 4\cdot 5 & 4\cdot 6 \end{vmatrix} = 4^2 \begin{vmatrix} 2 & 1 \\ 5 & 6 \end{vmatrix} = 16(12-5) = 16\cdot 7 = 112.$$

Un cálculo directo con la definición da, obviamente, igual resultado:

$$|A| = \begin{pmatrix} 8 & 4 \\ 20 & 24 \end{pmatrix} = 8\cdot 24 - 20\cdot 4 = 192 - 80 = 112.$$

♦

Teorema 2.6 Si se intercambian dos filas (o columnas) distintas de un determinante, entonces el determinante cambia de signo.

Demostración. Sea la matriz $A = (a_{ij})_n$. Sea B la matriz que se obtiene de A al intercambiar la fila r con la fila s en la matriz A. Vamos a demostrar que $|B| = -|A|$. Supongamos además que $r < s$.

En primer lugar, demostremos el caso en que $s = r+1$. Esto significa que el intercambio de dos filas consecutivas cambia el signo del determinante. Luego probaremos que para intercambiar dos filas cualesquiera diferentes se requiere un número impar de intercambios de filas consecutivas, lo que demuestra el teorema.

De la definición de B tenemos que:

$$B_i = \begin{cases} A_i & i \neq r,\ i \neq r+1, \\ A_{r+1} & i = r, \\ A_r & i = r+1. \end{cases}$$

Ahora, calculemos el determinante de B por menores de la fila $r+1$:

$$|B| = \sum_{j=1}^{n} (-1)^{(r+1)+j} b_{(r+1)j} |B_{r+1}^j|.$$

Pero $B_{r+1}^j = A_r^j$ y $b_{r+1}j = a_{rj}$, como se puede ver en la definición de B. Entonces:

$$\begin{aligned} |B| &= \sum_{j=1}^{n} (-1)^{(r+1)+j} a_{rj} |A_r^j| \\ &= \sum_{j=1}^{n} (-1)^{(r+j)} (-1) a_{rj} |A_r^j| \\ &= (-1) \sum_{j=1}^{n} (-1)^{(r+j)} a_{rj} |A_r^j|. \end{aligned}$$

Es decir, $|B| = -|A|$.

Ahora mostremos que se requiere un número impar de intercambios consecutivos para lograr el intercambio entre las filas r y s. Procedamos así:

1. Vamos a colocar la fila número r en la posición s. Para ello, intercambios A_r con A_{r+1}, luego con A_{r+2}, y así hasta intercambiar con A_s. En este proceso se han realizado $(s-r)$ intercambios consecutivos.
2. Ahora hay que llevar la fila A_s, que en este momento se encuentra en la fila número $s-1$, ya que la fila número s está ahora ocupada por A_r, a la fila número r. Esto significa que hay que intercambiar las filas número r y número $s-1$. El paso anterior nos dice que requeriremos $(s-r-1)$ (uno menos que el anterior).

Por lo tanto, se requiere $(s-r)+(s-r)-1 = 2(s-r)-1$ que es, como se puede ver, un número impar. □

Teorema 2.7 Si dos filas (o columnas) distintas de un determinante son iguales, entonces el determinante es igual a cero.

Demostración. Sea A una matriz que tiene dos filas iguales. Sea B la matriz obtenida de A al intercambiar dichas filas iguales. Entonces, $B = A$ y, por el teorema 2.6, $|B| = -|A|$. Por lo tanto $|A| = -|A|$, de donde $|A| = 0$. □

Ejemplo 2.19

$$|A| = \begin{vmatrix} -1 & 3 & 3 \\ 4 & 12 & 4 \\ 2 & 6 & 1 \end{vmatrix}_{F_2 \rightleftharpoons F_3} = -\begin{vmatrix} -1 & 3 & 3 \\ 2 & 6 & 1 \\ 4 & 12 & 4 \end{vmatrix}_{C_1 \rightleftharpoons C_2} = \begin{vmatrix} 3 & -1 & 3 \\ 6 & 2 & 1 \\ 12 & 4 & 4 \end{vmatrix}.$$

♦

Teorema 2.8 Si una fila (o columna) de un determinante es múltiplo de otra fila (o columna, respectivamente), entonces el determinante es igual a cero.

Demostración. Sean A una matriz, r y s, con $r < s$, tales que la fila número r es múltiplo de la fila número s; es decir, existe un escalar $\alpha \neq 0$ tal que

$$A_r = \alpha A_s.$$

Vamos a probar que el determinante de A es igual a 0.

Sea B la matriz definida por:

$$B_i = \begin{cases} A_i & \text{si } i \neq r, \\ A_s & \text{si } i = r. \end{cases}$$

Por lo tanto, B tiene dos filas iguales: la r-ésima y la s-ésima, ya que la matriz B se obtuvo de A al dividir toda la fila r-ésima por α (lo que es posible, ya que α es diferente de 0); es decir, B se obtuvo de A "al quitar el escalar α de cada término de la fila r-ésima". Por lo tanto, tenemos que $|B| = 0$.

Por otro lado, por el teorema 2.3, sabemos que el determinante de A es el producto del escalar por la matriz sin el escalar; es decir: $|A| = \alpha|B|$. Por lo tanto, $|A| = 0$. □

Teorema 2.9 Si a una fila (o columna) de un determinante se le suma un múltiplo constante de otra fila (o columna, respectivamente), entonces el determinante no cambia.

Demostración. Sean A una matriz de orden n y B la matriz cuya fila r-ésima es el resultado de sumar la fila r-ésima de A con el producto de un escalar α por la fila s-ésima de A $(r \neq s)$.

Entonces, si C es la matriz obtenida a partir de A al multiplicar la fila s-ésima por α, tenemos que $|C| = 0$, ya que las fila número s es un múltiplo de número s (teorema 2.8).

Por otro lado, la matriz B es idéntica a las matrices A y C salvo en la fila r-ésima, la cual es la suma de las correspondientes filas r-ésimas de A y de C. Por lo tanto, por el teorema 2.2, tenemos que:

$$|B| = |A| + |C| = |A| + 0 = |A|,$$

que es lo que queríamos demostrar. □

Ejemplo 2.20

$$\begin{vmatrix} 1 & 2 & 3 \\ -2 & 4 & 5 \\ 1 & 10 & 4 \end{vmatrix}_{F_2+3F_1} = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 10 & 14 \\ 1 & 10 & 14 \end{vmatrix}_{F_3-F_2} = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 10 & 14 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

♦

Ejemplo 2.21

Hallar el valor del determinante

$$\begin{vmatrix} \alpha & \alpha+1 & -2 \\ \alpha & \alpha+2 & \alpha+7 \\ 2\alpha & 2\alpha+2 & \alpha+6 \end{vmatrix}.$$

Solución.

$$\begin{vmatrix} \alpha & \alpha+1 & -2 \\ \alpha & \alpha+2 & \alpha+7 \\ 2\alpha & 2\alpha+2 & \alpha+6 \end{vmatrix}_{\substack{F_2-F_1 \\ F_3-2F_1}} = \begin{vmatrix} \alpha & \alpha+1 & -2 \\ 0 & 1 & \alpha+9 \\ 0 & 0 & \alpha+10 \end{vmatrix} = \alpha(\alpha+10).$$

♦

Teorema 2.10 Si A es una matriz cuadrada, entonces $|A^t| = |A|$.

En otras palabras, el determinante de una matriz y el de su transpuesta son iguales.

Demostración. Utilicemos inducción matemática sobre n, el orden de la matriz.

1. Supongamos que $n = 1$. Sea A la matriz
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$
Entonces
$$A^t = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}.$$
y
$$|A^t| = a_{11}a_{22} - a_{12}a_{21} = |A|.$$

2. Supongamos ahora que si A es de orden k, entonces $|A^t| = |A|$. Vamos a probar que si A es una matriz de orden $k+1$, también se verifica que $|A^t| = |A|$.

   Para ello, calculemos el determinante de A^t desarrollando el menor por la primera fila de A^t:
$$\begin{aligned} |A^t| &= \sum_{j=1}^{n} (-1)^{1+j} \hat{a}_{1j} |\hat{A}_1^j| \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{j1} |(A_j^1)^t|. \end{aligned}$$
   Pero la matriz A_j^1 es de orden k. Entonces, por la hipótesis de inducción, tenemos que $|(A_j^1)^t| = |A_1^j|$. Por lo tanto:
$$|A^t| = \sum_{j=1}^{n} (-1)^{1+j} a_{j1} |A_j^1|.$$
   Pero el término de la derecha de esta igualdad, es el determinante de A al desarrollar el menor por la primera columna de A. Es decir:
$$\sum_{j=1}^{n} (-1)^{1+j} a_{j1} |A_j^1| = |A|.$$
   Por lo tanto, $|A^t| = |A|$.

□

Definición 2.2 (Matriz semidescompuesta) Una matriz $A \in M_n$ se denomina **matriz semidescompuesta** si es de la forma
$$\begin{pmatrix} B & O \\ C & D \end{pmatrix},$$
donde $B \in M_{r\times r}$, $D \in M_{s\times s}$, $\mathbb{O} \in M_{r\times s}$ (matriz nula), $C \in M_{s\times r}$ y $r + s = n$.

Una matriz descompuesta también es de la forma
$$\begin{pmatrix} B & C \\ O & D \end{pmatrix},$$
con $B \in M_{r\times r}$, $D \in M_{s\times s}$, $\mathbb{O} \in M_{s\times r}$, $C \in M_{r\times s}$ y $r + s = n$.

En los dos casos a las matrices B y D se les llama *células diagonales* de la matriz A.

Teorema 2.11 El determinante de una matriz semidescompuesta es igual al producto de los determinantes de sus células diagonales.

Es decir, si

$$\begin{pmatrix} B & C \\ O & D \end{pmatrix},$$

es una matriz semidescompuesta, entonces su determinante es igual a $|B|\,|D|$.

Demostración. Utilicemos inducción sobre n, el orden de la matriz A.

1. Si $n = 2$ y
$$A = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix},$$
tenemos que
$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22}.$$

2. Supongamos que el teorema es verdadero para cada matriz semidescompuesta de orden k. Vamos a probar que si A una matriz semidescompuesta de orden $k+1$, es decir, si
$$A = \begin{pmatrix} B & O \\ C & D \end{pmatrix},$$
donde $B = (b_{ij})_r$ y $D = (d_{ij})_s$, con $r + s = k + 1$, se verifica que
$$|A| = |B|\,|D|.$$
Puesto que $a_{1j} = b_{1j}$ para $j = 1, 2, \ldots, r$, y $a_{1j} = 0$ para $j = r+1, \ldots, k+1$, se tiene que
$$|A| = \sum_{j=1}^{k+1} (-1)^{1+j} a_{1j} |A_1^j| = \sum_{j=1}^{r} (-1)^{1+j} b_{1j} |A_1^j|.$$
Por otro lado, A_1^j son matrices descompuestas de orden k cuyas células diagonales son B_1^j, $j = 1, 2, \ldots, r$ y D. Entonces, por la hipótesis de inducción, tenemos que:
$$|A_1^j| = |B_1^j|\,|D|,$$
lo que produce que
$$|A| = \sum_{j=1}^{r} (-1)^{1+j} b_{1j} |B_1^j||D| = \left(\sum_{j=1}^{r} (-1)^{1+j} b_{1j} |B_1^j| \right) |D| = |B|\,|D|.$$

□

El siguiente teorema es, en realidad, un corolario de este último.

Teorema 2.12 Si $A, B \in M_n$, entonces $|AB| = |A|\,|B|$.

Es decir, el determinante de un producto de matrices es igual al producto de los determinantes de las matrices. No existe una propiedad similar para la suma de matrices.

Demostración. La aplicación del teorema anterior a la siguiente matriz semidescompuesta:

$$C = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix},$$

donde $A = (a_{ij})_n$ y $B = (b_{ij})_n$ produce el resultado buscado. □

Corolario 2.13 Sean $A_1, A_2, \ldots, A_m$ m matrices cuadradas de orden n. Entonces se tiene que

$$|A_1 A_2 \cdots A_m| = |A_1|\,|A_2| \cdots |A_m|$$

Demostración. Por inducción sobre m. □

De este corolario se deduce inmediatamente el siguiente teorema.

Teorema 2.14 Sean $A \in M_n$ y $m \in \mathbb{N}$. Entonces $|A^m| = |A|^m$.

Se entiende que A^m es la abreviación de $AA \cdots A$, donde A aparece m veces.

Ejemplo 2.22

Sea la matriz

$$A = \begin{pmatrix} 1 & 3 & -4 \\ 2 & 7 & -6 \\ -3 & -8 & 12 \end{pmatrix}.$$

Halle $|A^n|$.

Solución. Primero encontremos $|A|$ usando las propiedades vistas anteriormente:

$$|A| = \begin{vmatrix} 1 & 3 & -4 \\ 2 & 7 & -6 \\ -3 & -8 & 12 \end{vmatrix}_{\substack{F_2 - 2F_1 \\ F_3 + 3F_1}} = \begin{vmatrix} 1 & 3 & -4 \\ 0 & 1 & 2 \\ 0 & 1 & 0 \end{vmatrix}_{F_3 - F_2} = \begin{vmatrix} 1 & 3 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{vmatrix} = 1 \cdot 1 \cdot (-2) = -2$$

Aplicamos la última propiedad:

$$|A^n| = |A|^n = (-2)^n.$$

♦

2.3 Determinantes y matrices inversibles

Teorema 2.15 Si $A \in M_n$ es una matriz inversible, entonces

$$|A| \neq 0 \quad \text{y} \quad |A^{-1}| = \frac{1}{|A|}.$$

Demostración. Como A es inversible, existe A^{-1} y $A^{-1}A = I$. Entonces:

$$|A^{-1}A| = |I|.$$

Pero

$$|A^{-1}A| = |A^{-1}|\,|A| \quad \text{y} \quad |I| = 1.$$

Por lo tanto

$$|A^{-1}|\,|A| = 1,$$

de donde $|A^{-1}| \neq 0$ y $|A^{-1}| = \dfrac{1}{|A|}$. □

Ejemplo 2.23

Se sabe que la matriz

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

es inversible. Calcular $|A^{-1}|$.

Solución. Como

$$|A| = \begin{vmatrix} 1 & 0 & -1 \\ 3 & 1 & 4 \\ 0 & 2 & 0 \end{vmatrix} = -2\begin{vmatrix} 1 & -1 \\ 3 & 4 \end{vmatrix} = -2(4+3) = -14,$$

tenemos que

$$|A^{-1}| = \frac{1}{|A|} = \frac{1}{-14} = -\frac{1}{14}.$$

♦

Definición 2.3 (Cofactor de una matriz) El **ij-ésimo cofactor de la matriz** $A = (a_{ij}) \in M_n$, notado A_{ij}, se define como

$$A_{ij} = (-1)^{i+j}|A_i^j|.$$

La **matriz de los cofactores** de la matriz $A \in M_n$, notada $\operatorname{cof} A$, se define como $\operatorname{cof} A = (A_{ij})_n$.

Con esta definición, el desarrollo del determinante por menores de la matriz $A = (a_{ij})_n$ por la r-ésima fila puede escribirse como

$$|A| = \sum_{k=1}^{n} a_{rk}A_{rk} \quad r \in \{1, 2, \ldots, n\},$$

o por la s-ésima columna como

$$|A| = \sum_{k=1}^{n} a_{ks}A_{ks} \quad s \in \{1, 2, \ldots, n\}.$$

Definición 2.4 (Matriz adjunta) Se denomina **adjunta** de la matriz $A \in M_n$, notada $\operatorname{adj} A$, a la matriz transpuesta de la matriz de los cofactores de A, es decir, $\operatorname{adj}(A) = (\operatorname{cof} A)^t$.

Teorema 2.16 Si $A = (a_{ij})_n$, entonces $\sum_{k=1}^{n} a_{ij}A_{jk} = 0$ para todo i y para todo j tal que $i \neq j$.

Esto significa que el producto de la i-ésima fila de la matriz A, A_i, por la j-ésima fila, con $j \neq i$, de la matriz $\operatorname{cof} A$ es igual a cero.

Demostración. Sea B la matriz tal que $B_s = A_r$ y el resto de sus filas son iguales a las correspondientes de A. Entonces $|B| = 0$.

Por otro lado, de la definición de B también tenemos que:

$$|B| = \sum_{j=1}^{n} (-1)^{j+s} b_{sj} |B_s^j| = \sum_{j=1}^{n} (-1)^{j+a} a_{rj} |A_s^j|.$$

Entonces

$$|B| = \sum_{j=1}^{n} a_{rk}A_s^j.$$

Entonces, por la definición de cofactor, tenemos que

$$\sum_{j=1}^{n} a_{rj}A_{sj} = |B| = 0.$$

Esto significa que el producto de una fila r de A por una fila s $(s \neq r)$ de $\operatorname{cof} A$ es cero. Además, obsérvese que

$$\sum_{j=1}^{n} a_{rj}A_{sj} = |A|$$

para $r = 1, 2, \ldots, n$. □

El siguiente teorema nos permitirá obtener un método alternativo para calcular la inversa de una matriz.

Teorema 2.17 Sea la matriz $A = (a_{ij}) \in M_n$. Entonces $A \operatorname{adj}(A) = |A|\, I_n$.

Demostración. Sea $C = A \operatorname{adj}(A)$. Entonces:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} A_{jk}.$$

Entonces, por el teorema 2.16, tenemos que $c_{ij} = 0$ si $i \neq j$ y $c_{ii} = |A|$. Por lo tanto:

$$A \operatorname{adj} A = \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |A| \end{pmatrix} = |A| I,$$

por el teorema 2.11. □

Ejemplo 2.24

Sea la matriz

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

Calcular $A \operatorname{adj}(A)$.

Solución. Calculemos

$$|A| = \begin{vmatrix} 1 & 0 & -1 \\ 3 & 1 & 4 \\ 0 & 1 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & -1 \\ 3 & 4 \end{vmatrix} = -7$$

Por el teorema anterior:

$$A \operatorname{adj}(A) = |A|\, I_3 = -7 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -7 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & -7 \end{pmatrix}.$$

Verificación:

$$\operatorname{cof} A = \begin{pmatrix} -4 & 0 & 3 \\ -1 & 0 & -1 \\ 1 & -7 & 1 \end{pmatrix}, \quad \operatorname{adj}(A) = (\operatorname{cof} A)^t = \begin{pmatrix} -4 & -1 & 1 \\ 0 & 0 & -7 \\ 3 & -1 & 1 \end{pmatrix},$$

$$A \operatorname{adj}(A) = \begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 4 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -4 & -1 & 1 \\ 0 & 0 & -7 \\ 3 & -1 & 1 \end{pmatrix} = \begin{pmatrix} -7 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & -7 \end{pmatrix} = -7 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = |A|\, I_3.$$

♦

Teorema 2.18 Sea la matriz $A = (a_{ij}) \in M_n$. La matriz A es inversible si y solo si $|A| \neq 0$ y, además,

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

Este teorema nos permite saber si una matriz es inversible o no con tan solo calcular su determinante. Si éste es igual a cero, la matriz no es inversible; en el caso contrario, la matriz es inversible. Adicionalmente, podemos calcular la matriz inversa como el producto del inverso multiplicativo del determinante de la matriz por la adjunta de la matriz.

Demostración. Sea A una matriz de orden n.

Si la matriz A es inversible, entonces $|A| \neq 0$. Pero, por el teorema 2.17, tenemos que:

$$A \operatorname{adj} A = |A| I,$$

de donde

$$A\left(\frac{1}{|A|}\operatorname{adj} A\right) = I.$$

Por lo tanto:

$$A^{-1} = \frac{1}{|A|}\operatorname{adj} A.$$

2. Recíprocamente, supongamos que $|A| \neq 0$ y que $A \operatorname{adj} A = |A| I$, entonces

$$A\left(\frac{1}{|A|}\operatorname{adj} A\right) = I.$$

Entonces la matriz A es inversible y se tiene

$$A^{-1} = \frac{1}{|A|}\operatorname{adj} A.$$

□

Ejemplo 2.25

Sea la matriz

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

1. ¿Qué condición deben cumplir a, b, c y d para que la matriz A sea invertible?
2. Si la matriz A es invertible, encontrar A^{-1}.

Solución.

1. El determinante de la matriz A es

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb.$$

Entonces la matriz A es inversible si y solo si $ad - cb \neq 0$.

2. Sea la matriz A tal que $ad - cb \neq 0$. Tenemos que

$$|A| = ad - cb, \quad \operatorname{cof} A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}, \quad \operatorname{adj} A = (\operatorname{cof} A)^t = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

   Entonces tenemos que

$$A^{-1} = \frac{1}{|A|} \operatorname{adj} A = \frac{1}{ad - cb} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

♦

Ejemplo 2.26

Sea la matriz

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

Si la matriz A es inversible, calcular su inversa A^{-1}.

Solución. Hallemos el determinante de la matriz:

$$|A| = \begin{vmatrix} 1 & 0 & -1 \\ 3 & 1 & 4 \\ 0 & 1 & 0 \end{vmatrix} = -7 \neq 0.$$

Entonces la matriz A es inversible. De un ejercicio anterior tenemos su adjunta:

$$\operatorname{adj} A = \begin{pmatrix} -4 & -1 & 1 \\ 0 & 0 & -7 \\ 3 & -1 & 1 \end{pmatrix}$$

La inversa de la matriz A es

$$A^{-1} = \frac{1}{|A|} \operatorname{adj} A = -\frac{1}{7} \begin{pmatrix} -4 & -1 & 1 \\ 0 & 0 & -7 \\ 3 & -1 & 1 \end{pmatrix}.$$

Verificación:

$$A^{-1}A = -\frac{1}{7} \begin{pmatrix} -4 & -1 & 1 \\ 0 & 0 & -7 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 4 \\ 0 & 1 & 0 \end{pmatrix} = -\frac{1}{7} \begin{pmatrix} -7 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & -7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

♦

2.4 Ejercicios resueltos

1. Demostrar que si AB es inversible, con A y B matrices cuadradas, entonces A y B son inversibles.

Demostración. Supongamos que AB es inversible. Entonces, $|AB| \neq 0$. Por lo tanto, $|A| \neq 0$ y y $|B| \neq 0$. Esto significa que A y B son inversibles. □

2. Sean $A, B \in M_n$, con A y B inversibles. Demostrar que:

$$|AB - \lambda I| = |BA - \lambda I|.$$

Demostración. Puesto que:

$$\begin{aligned}|AB - \lambda I| &= |AB - \lambda AA^{-1}| = |A||B - \lambda A^{-1}| \\ &= |B - \lambda A^{-1}||A| = |(B - \lambda A^{-1})A| \\ &= |BA - \lambda A^{-1}A| = |BA - \lambda I|,\end{aligned}$$

concluimos que

$$|AB - \lambda I| = |BA - \lambda I|.$$

□

3. Hallar el valor del determinante

$$\Delta_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 2 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix}.$$

Desarrollo. Utilicemos el desarrollo por menores de la primera columna. Obtenemos lo siguiente:

$$\begin{aligned}\Delta_n &= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 2 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix} \\ &= \begin{vmatrix} 2 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix} - (-1) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 2 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix} \\ &= 2^{n-1} + \Delta_{n-1} \\ &= 2^{n-1} + 2^{n-2} + \cdots 2^2 + \Delta_2 \\ &= 2^{n-1} + 2^{n-2} + \cdots 2^2 + 2^1 + 2^0 \\ &= \sum_{i=0}^{n-1} 2^i = \frac{2^n - 1}{2 - 1} = 2^n - 1.\end{aligned}$$

□

4. Sea la matriz $A \in M_n$ antisimétrica. Demostrar que si n es impar, $|A| = 0$.

Demostración. Supongamos que A es antisimétrica. Entonces $A^t = -A$. Por lo tanto, $|A^t| = |-A|$. Pero, $|A| = (-1)^n|A|$. Así que $|A| = -|A|$, ya que n es impar. Esto significa que $|A| = 0$. $\square$

5. ¿Para qué valores de $a \in \mathbb{R}$ la matriz

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

es inversible? Hallar la inversa para dichos valores.

Desarrollo. Hallemos el determinante de la matriz A:

$$|A| = \begin{vmatrix} 1 & -2 & 3 \\ a & 1 & -3 \\ 1 & -1 & a \end{vmatrix}_{\substack{F_2 - aF_1 \\ F_3 - 1F_1}} = \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1+2a & -3-3a \\ 0 & 1 & a-3 \end{vmatrix}$$
$$= (1+2a)(a-3) + (3+3a) = 2a^2 - 2a = 2a(a-1).$$

Si $a \in \mathbb{R} - \{0, 1\}$, $|A| \neq 0$ y, por lo tanto, la matriz A es inversible.

Ahora calculemos la inversa para $a \in \mathbb{R} - \{0, 1\}$. En primer lugar, calculemos los cofactores de la matriz A:

$$A_{11} = \begin{vmatrix} 1 & -3 \\ -1 & a \end{vmatrix} = a - 3, \quad A_{12} = -\begin{vmatrix} a & -3 \\ 1 & a \end{vmatrix} = -a^2 - 3,$$

$$A_{13} = \begin{vmatrix} a & 1 \\ 1 & -1 \end{vmatrix} = -a - 1, \quad A_{21} = -\begin{vmatrix} -2 & 3 \\ -1 & a \end{vmatrix} = 2a - 3,$$

$$A_{22} = \begin{vmatrix} 1 & 3 \\ 1 & a \end{vmatrix} = a - 3, \qquad A_{23} = -\begin{vmatrix} 1 & -2 \\ 1 & -1 \end{vmatrix} = -1,$$

$$A_{31} = \begin{vmatrix} -2 & 3 \\ 1 & -3 \end{vmatrix} = 3, \qquad A_{32} = -\begin{vmatrix} 1 & 3 \\ a & -3 \end{vmatrix} = 3a + 3,$$

$$A_{33} = \begin{vmatrix} 1 & -2 \\ a & -1 \end{vmatrix} = 2a + 1.$$

Ya podemos hallar la inversa:

$$A^{-1} = \frac{1}{|A|} \operatorname{adj} A = \frac{1}{|A|} (\operatorname{cof} A)^t = \frac{1}{2a(a-1)} \begin{pmatrix} a-3 & 2a-3 & 3 \\ -a^2-3 & a-3 & -1 \\ 3 & 3a+3 & 2a+1 \end{pmatrix}.$$

$\square$

Sistemas de ecuaciones lineales 3

3.1 Ecuaciones lineales

La ecuación

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

donde los coeficientes $a_1, a_2, \ldots, a_n$ y el término independiente b son constantes conocidas, se llama *lineal en las incógnitas* $x_1, x_2, \ldots, x_n$. Se llama lineal porque todas las incógnitas están elevadas a la potencia 1 y porque los coeficientes $a_1, a_2, \ldots, a_n$ no contienen a ninguna de las incógnitas.

Ejemplo 3.27 La ecuación

$$2x + y - 5z + 3w = -3$$

es una ecuación lineal en las incógnitas x, y, z y w.

Ninguna de las ecuaciones

$$\begin{gathered} 2x + y - 5z^2 + 3w = -3, \\ 2x + y - 5z + 3xw = -3, \\ 2x^{-1} + y - 5z + 3w = -3 \\ 2x + \cos y - 5z + 3w = -3, \\ 2x + y - 5z + 3\ln w = -3, \end{gathered}$$

es una ecuación lineal. ♦

Resolver una ecuación lineal es encontrar los valores de las incógnitas que satisfacen la ecuación; es decir, encontrar los valores que al sustituir las incógnitas, reducen la ecuación a una identidad. El conjunto de valores que satisfacen la ecuación se llama *solución* de la ecuación. En concreto, una sucesión de números $s_1, s_2, \ldots, s_n$ se llama solución de la ecuación lineal

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

si al sustituir $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ en la ecuación, la proposición

$$a_1s_1 + a_2s_2 + \cdots + a_sx_n = b,$$

obtenida por la sustitución, es verdadera.

Ejemplo 3.28 La sucesión $s_1 = 2, s_2 = 4, s_3 = 4, s_4 = 3$ es una solución de la ecuación lineal

$$2x_1 + x_2 - 5x_3 + 3x_4 = -3.$$

En efecto, al hacer las sustituciones $x_1 = s_1, x_2 = s_2, x_3 = s_3, x_4 = s_4$ en la ecuación, ella se transforma en identidad:

$$\begin{aligned} 2x_1 + x_2 - 5x_3 + 3x_4 &= 2s_1 + s_2 - 5s_3 + 3s_4 \\ &= 2 \cdot 2 + 4 - 5 \cdot 4 + 3 \cdot 3 \\ &= 4 + 4 - 20 + 9 \\ &= -3. \end{aligned}$$

La sucesión $r_1 = -2, r_2 = -4, r_3 = -1, r_4 = 0$ es otra solución de la ecuación lineal porque esos valores también satisfacen la ecuación:

$$\begin{aligned} 2x_1 + x_2 - 5x_3 + 3x_4 &= 2r_1 + r_2 - 5r_3 + 3r_4 \\ &= 2 \cdot (-2) + (-4) - 5 \cdot (-1) + 3 \cdot 0 \\ &= -4 - 4 + 5 + 0 \\ &= -3. \end{aligned}$$

La sucesión $t_1 = 4, t_2 = 3, t_3 = 0, t_4 = 2$ no es solución de la ecuación lineal porque no la reduce a una identidad:

$$\begin{aligned} 2x_1 + x_2 - 5x_3 + 3x_4 &= 2t_1 + t_2 - 5t_3 + t_4 \\ &= 2 \cdot 4 + 3 - 5 \cdot 0 + 3 \cdot 2 \\ &= 8 + 3 + 0 + 6 = 17 \neq -3. \end{aligned}$$

♦

3.2 Sistemas de ecuaciones lineales

Definición 3.1 (Sistema de ecuaciones lineales) Se llama sistema de m ecuaciones lineales en n incógnitas, o simplemente **sistema de ecuaciones lineales**, a una colección de ecuaciones de la forma

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2, \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \tag{3.1}$$

Los x_i son las n **incógnitas**. Los coeficientes a_{ij} y los términos independientes b_i son constantes conocidas.

El número de ecuaciones m puede ser menor, igual o mayor que el número de incógnitas n. Por brevedad, nos referiremos como *sistema lineal* a un sistema de ecuaciones lineales.

De la definición de multiplicación de matrices, podemos realizar una representación matricial, concisa y elegante, de un sistema de ecuaciones lineales:

$$AX = B,$$

donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}. \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

La matriz $A \in M_{m\times n}$ está formada con los coeficientes a_{ij} del sistema; la matriz $X \in M_{n\times 1}$, por las incógnitas x_i del sistema; y $B \in M_{m\times 1}$, por los términos independientes b_i del sistema.

Como se indicó, con la definición de multiplicación de matrices, es fácil ver que la representación matricial $AX = B$ se reduce a la representación explícita del sistema de ecuaciones lineales:

$$AX = B,$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

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

3.3 Solución de un sistema lineal

Definición 3.2 (Solución de un sistema lineal) Una matriz

$$S = \begin{pmatrix} s_1 & s_2 & \cdots & s_n \end{pmatrix}^t$$

es **solución** del sistema lineal $AX = B$ si y solo si $AS = B$.

Dicho de forma explícita, la sucesión $s_1, s_2, \ldots, s_n$ es una solución del sistema lineal $AX = B$ si al realizar las sustituciones $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ en cada una de las ecuaciones

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad i = 1, 2, \ldots, m$$

del sistema lineal, cada una de ellas se reduce a una identidad:

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = b_i. \quad i = 1, 2, \ldots, m$$

Solo existen tres posibilidades de solución de un sistema lineal:

No existe solución: No existe un conjunto de valores para las incógnitas x_i que satisfaga todas las ecuaciones simultáneamente.

Solución única: Existe uno y solo un conjunto de valores para las incógnitas x_i que satisface todas las ecuaciones simultáneamente.

Infinitas soluciones: Existen infinitos conjuntos de valores para las incógnitas x_i que satisfacen todas las ecuaciones simultáneamente. Es sencillo probar que si un sistema tiene más de una solución, tiene, en realidad, un número infinito de soluciones. Esto significa que un sistema no puede tener un número finito de soluciones que no sea el número 1.

Definición 3.3 (Ecuación combinación lineal) Al multiplicar cada ecuación de un sistema lineal por un escalar c_i, $i = 1, 2, \ldots, m$ (no todos nulos) y luego sumar los miembros izquierdos y los miembros derechos, respectivamente, se obtiene la ecuación

$$\left(\sum_{i=1}^{m} c_i a_{i1}\right) x_1 + \left(\sum_{i=1}^{m} c_i a_{i2}\right) x_2 + \cdots + \left(\sum_{i=1}^{m} c_i a_{in}\right) x_n = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} c_i a_{ij}\right) x_j = \sum_{i=1}^{m} c_i b_i,$$

la cual se llama *combinación lineal* de las ecuaciones del sistema lineal.

Ejemplo 3.29 Sea el sistema lineal

$$\begin{aligned} 2x - 3y &= 1, \\ -x + 2y &= -4. \end{aligned}$$

Multipliquemos la primera ecuación por -1 y la segunda por 2:

$$\begin{aligned} -2x + 3y &= -1, \\ -2x + 4y &= -8. \end{aligned}$$

Sumamos miembro a miembro las dos ecuaciones anteriores. La nueva ecuación

$$-4x + 7y = -9$$

es una combinación lineal de las ecuaciones del sistema lineal. ♦

Definición 3.4 (Sistemas lineales equivalentes) Se dice que dos sistemas lineales $AX = B$ y $CX = D$ son **equivalentes** si cada una de las ecuaciones del un sistema es una combinación lineal de las ecuaciones del otro sistema.

El siguiente teorema muestra la importancia de los sistemas equivalentes.

Teorema 3.1 Dos sistemas lineales equivalentes tiene exactamente las mismas soluciones.

Demostración. Supongamos que $AX = B$ y $CX = D$ son dos sistemas lineales equivalentes. Vamos a probar que tienen las mismas soluciones. Para ello, vamos a probar que s es solución de $AX = B$ si y solo si es solución de $CX = D$.

En primer lugar, supongamos que s es una solución de $AX = B$, donde A es de orden $m \times n$. Esto significa que para cada $i \in I_m$, la ecuación i-ésima de este sistema,

$$\sum_{j=1}^{n} a_{ij} x_j = b_i,$$

es satisfecha por s; es decir, se verifica la igualdad:

$$\sum_{j=1}^{n} a_{ij} s_j = b_i, \tag{3.2}$$

donde $s = (s_1, s_2, \cdots, s_n)^t$.

Vamos a probar que s es solución del sistema $CX = D$. Para ello, recordemos que cada ecuación de este sistema, es una combinación lineal del otro sistema. Esto significa que, para cada $i \in I_m$, la ecuación i-ésima de este sistema es de la forma

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} c_i a_{ij} \right) x_j = \sum_{i=1}^{m} c_i b_i. \tag{3.3}$$

Mostremos a continuación que s satisface cada una de estas ecuaciones. Para ello, reemplacemos x_j por s_j en lado izquierdo de la ecuación (3.3):

$$\begin{aligned} \sum_{j=1}^{n} \left(\sum_{i=1}^{m} c_i a_{ij} \right) s_j &= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} c_i a_{ij} s_j \right) \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} c_i a_{ij} s_j \right) = \sum_{i=1}^{n} c_i \left(\sum_{j=1}^{m} a_{ij} s_j \right). \end{aligned}$$

Es decir:

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} c_i a_{ij} \right) s_j = \sum_{i=1}^{n} c_i \left(\sum_{j=1}^{m} a_{ij} s_j \right). \tag{3.4}$$

Pero la igualdad (3.2) produce en la igualdad (3.4) lo siguiente:

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} c_i a_{ij} \right) s_j = \sum_{i=1}^{n} c_i b_i.$$

Es decir, s satisface cada ecuación (3.3) del sistema $CX = D$.

Se deja al lector la demostración de que s es solución del sistema $AX = B$ si es solución del sistema $CX = D$. □

Una forma rápida de saber si dos sistemas lineales $AX = B$ y $CX = D$ son o no equivalentes es hallando las matrices escalonadas reducidas por filas de las matrices ampliadas $(A \mid B)$ y $(C \mid D)$. Si dichas matrices ampliadas son iguales, los sistema son equivalentes. Caso contrario no son equivalentes.

Definición 3.5 (Sistemas lineales homogéneos y no homogéneos) Un sistema de ecuaciones lineales $AX = B$ se llama **homogéneo** si $B = \mathbb{O}$. Si $B \neq \mathbb{O}$ el sistema se llama **no homogéneo**.

Teorema 3.2 Sean A y B dos matrices de orden n y equivalentes por filas. Entonces los sistemas homogéneos $AX = \mathbb{O}$ y $BX = \mathbb{O}$ son equivalentes; es decir, tienen las mismas soluciones.

Demostración. Supongamos que la matriz A_1 se obtiene a partir de A por la aplicación de una operación elemental por fila. Es decir, $A_1 = eA$, donde e es una operación elemental. Se tiene, entonces, que $A_1X = \mathbb{O}$, pues el sistema $A_1X = \mathbb{O}$ es equivalente al sistema $AX = \mathbb{O}$, ya que cualquier operación elemental por filas hace que las ecuaciones del segundo sistema sean combinación lineal de las del primero.

En efecto, si e es del primer tipo, todas las ecuaciones del segundo sistema, salvo una, son idénticas a las del primer sistema (en este caso, todas esas ecuaciones se han obtenido del primer sistema tomando todos los escalares iguales a 0, excepto uno de ellos que ha tomado el valor 1). La ecuación diferente del segundo sistema es el producto de la correspondiente del primer sistema por un escalar diferente de 0; luego, también es una combinación lineal.

Ahora bien, como B es equivalente con A, existe una sucesión de matrices $A_1, \ldots,$ $A_n = B$, y una sucesión de operaciones elementales por fila $e_1, \ldots, e_n$ tales que

$$A_{i+1} = e_iA_i,$$

$A_1 = e_1A$ y $B = e_nA_n$. Por lo demostrado al inicio, tenemos que $A_iX = \mathbb{O}$ para todo i. Entonces, $BX = \mathbb{O}$. □

Todavía no sabemos cuando un sistema lineal tiene o no solución, y en caso de tenerla, no sabemos si ella es única. Pero con lo visto hasta aquí, ya podemos resolver sistemas lineales.

3.3.1 Métodos de solución

La idea general del método de solución de un sistema lineal es hallar un sistema equivalente más simple de resolver que el original. Existen dos métodos que se diferencian en el tipo de sistema equivalente al que se llega.

Método de Gauss

En el método de Gauss se halla una matriz escalonada de la matriz ampliada $(A \mid B)$ del sistema lineal $AX = B$. Veamos cómo funciona con un ejemplo.

Ejemplo 3.30

Sea el sistema lineal

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 6,\\ x_1 + x_2 + 3x_3 + 3x_4 &= 12,\\ x_1 + x_2 + 2x_3 + 3x_4 &= 12,\\ x_1 + 3x_2 + 3x_3 + 3x_4 &= 14. \end{aligned}$$

Resuelva el sistema por el método de Gauss.

Solución. Sea $(A \mid B)$ la matriz ampliada del sistema. Hallemos una matriz escalonada de la matriz ampliada:

$$
\begin{aligned}
(A \mid B) &\sim \begin{pmatrix} 1 & 1 & 1 & 1 & \mid & 6 \\ 1 & 1 & 3 & 3 & \mid & 12 \\ 1 & 1 & 2 & 3 & \mid & 12 \\ 1 & 3 & 3 & 3 & \mid & 14 \end{pmatrix}_{\substack{F_2-F_1 \\ F_3-F_1 \\ F_4-F_1}} \sim \begin{pmatrix} 1 & 1 & 1 & 1 & \mid & 6 \\ 0 & 0 & 2 & 2 & \mid & 6 \\ 0 & 0 & 1 & 2 & \mid & 6 \\ 0 & 2 & 2 & 2 & \mid & 8 \end{pmatrix}_{\substack{\frac{1}{2}F_2 \\ \frac{1}{2}F_4}} \\
&\sim \begin{pmatrix} 1 & 1 & 1 & 1 & \mid & 6 \\ 0 & 0 & 1 & 1 & \mid & 3 \\ 0 & 0 & 1 & 2 & \mid & 6 \\ 0 & 1 & 1 & 1 & \mid & 4 \end{pmatrix}_{F_2 \rightleftharpoons F_4} \sim \begin{pmatrix} 1 & 1 & 1 & 1 & \mid & 6 \\ 0 & 1 & 1 & 1 & \mid & 4 \\ 0 & 0 & 1 & 2 & \mid & 6 \\ 0 & 0 & 1 & 1 & \mid & 3 \end{pmatrix}_{F_4-F_3} \\
&\sim \begin{pmatrix} 1 & 1 & 1 & 1 & \mid & 6 \\ 0 & 1 & 1 & 1 & \mid & 4 \\ 0 & 0 & 1 & 2 & \mid & 6 \\ 0 & 0 & 0 & -1 & \mid & -3 \end{pmatrix}_{-F_4} \sim \begin{pmatrix} 1 & 1 & 1 & 1 & \mid & 6 \\ 0 & 1 & 1 & 1 & \mid & 4 \\ 0 & 0 & 1 & 2 & \mid & 6 \\ 0 & 0 & 0 & 1 & \mid & 3 \end{pmatrix} = (C \mid D).
\end{aligned}
$$

El sistema $CX = D$

$$
\begin{aligned}
x_1 + x_2 + x_3 + \ \ x_4 &= 6, \\
x_2 + x_3 + \ \ x_4 &= 4, \\
x_3 + 2x_4 &= 6, \\
x_4 &= 3,
\end{aligned}
$$

es equivalente al original $AX = B$, pero es más fácil de resolver. Se lo resuelve al empezar a encontrar el valor de las incógnitas desde la última ecuación hacia la primera:

$$
\begin{aligned}
&x_4 = 3, \\
&x_3 = 6 - 2x_4 = 6 - 2 \cdot 3 = 0, \\
&x_2 = 4 - x_3 - x_4 = 4 - 0 - 3 = 1, \\
&x_1 = 6 - x_2 - x_3 - x_4 = 6 - 1 - 0 - 3 = 2.
\end{aligned}
$$

Hemos encontrado que el sistema tiene solución y que tal solución es única. Verifiquemos que la solución encontrada del sistema $CX = D$ es también solución del sistema original $AX = B$. Es decir, si la matriz

$$
S = \begin{pmatrix} 2 & 1 & 0 & 3 \end{pmatrix}^t
$$

es solución del sistema $CX = D$, verifiquemos que también es solución del sistema $AX = B$:

$$
AS = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 3 \\ 1 & 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 12 \\ 12 \\ 14 \end{pmatrix} = B.
$$

♦

Método de Gauss-Jordan

En el método de Gauss-Jordan, se halla la matriz escalonada reducida por filas de la matriz ampliada del sistema lineal.

Ejemplo 3.31

Resolver el sistema lineal del ejemplo anterior por el método de Gauss-Jordan.

Solución. Hallemos la matriz escalonada reducida por filas de la matriz ampliada del sistema partiendo desde la ya encontrada matriz escalonada:

$$(A \mid B) \sim \begin{pmatrix} 1 & 1 & 1 & 1 & | & 6 \\ 1 & 1 & 3 & 3 & | & 12 \\ 1 & 1 & 2 & 3 & | & 12 \\ 1 & 3 & 3 & 3 & | & 14 \end{pmatrix} \sim \cdots \sim \begin{pmatrix} 1 & 1 & 1 & 1 & | & 6 \\ 0 & 1 & 1 & 1 & | & 4 \\ 0 & 0 & 1 & 2 & | & 6 \\ 0 & 0 & 0 & 1 & | & 3 \end{pmatrix}_{F_1 - F_2}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 & | & 2 \\ 0 & 1 & 1 & 1 & | & 4 \\ 0 & 0 & 1 & 2 & | & 6 \\ 0 & 0 & 0 & 1 & | & 3 \end{pmatrix}_{F_2 - F_3} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & | & 2 \\ 0 & 1 & 0 & -1 & | & -2 \\ 0 & 0 & 1 & 2 & | & 6 \\ 0 & 0 & 0 & 1 & | & 3 \end{pmatrix}_{\substack{F_2 + F_4 \\ F_3 - 2F_4}}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 & | & 2 \\ 0 & 1 & 0 & 0 & | & 1 \\ 0 & 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & 1 & | & 3 \end{pmatrix} = (D \mid F)$$

El sistema lineal $DX = F$, equivalente al sistema original $AX = B$, puede escribirse explícitamente como

$$x = 2, \; y = 1, \; z = 0, \; w = 3,$$

el cual, a su vez, es ya la solución del sistema.

Obsérvese que en este caso el sistema lineal tiene solución única y que:

1. El rango de la matriz de coeficientes es igual al rango de la matriz ampliada:

$$\operatorname{ran} A = 4 = \operatorname{ran}(A \mid B)$$

2. Dicho rango es igual al número de incógnitas:

$$\operatorname{ran} A = \operatorname{ran}(A \mid B) = 4 = n$$

♦

Ejemplo 3.32

Resolver por el método de Gauss-Jordan el sistema lineal

$$\begin{aligned} x + 2y \phantom{{}+z} \quad + \; w &= 7, \\ x + \; y + \; z - \; w &= 3, \\ 3x + \; y + 5z - 7w &= 1. \end{aligned}$$

Solución.

Sea $(A \mid B)$ la matriz ampliada del sistema lineal. Hallemos su matriz escalonada reducida por filas:

$$(A \mid B) = \begin{pmatrix} 1 & 2 & 0 & 1 & | & 7 \\ 1 & 1 & 1 & -1 & | & 3 \\ 3 & 1 & 5 & -7 & | & 1 \end{pmatrix}_{\substack{F_2-F_1 \\ F_3-3F_1}} \sim \begin{pmatrix} 1 & 2 & 0 & 1 & | & 7 \\ 0 & -1 & 1 & -2 & | & -4 \\ 0 & -5 & 5 & -10 & | & -20 \end{pmatrix}_{-F_2}$$

$$\sim \begin{pmatrix} 1 & 2 & 0 & 1 & | & 7 \\ 0 & 1 & -1 & 2 & | & 4 \\ 0 & -5 & 5 & -10 & | & -20 \end{pmatrix}_{\substack{F_1-2F_2 \\ F_3+5F_2}} \sim \begin{pmatrix} 1 & 0 & 2 & -3 & | & -1 \\ 0 & 1 & -1 & 2 & | & 4 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix} = (C \mid D)$$

La tercera ecuación

$$0x + 0y + 0z + 0w = 0$$

del sistema $CX = D$ tiene como solución cualesquiera conjunto de valores de las incógnitas x, y, z y w. Por lo tanto, cualesquiera solución de las dos primeras ecuaciones será también solución de la tercera. Por ello el sistema lineal $CX = D$ lo escribimos así:

$$\begin{aligned} x \qquad + 2z - 3w &= -1, \\ y - \quad z + 2w &= 4. \end{aligned}$$

Para empezar a dar la solución del sistema, pensemos solamente en la segunda ecuación. Parece que hay una gran amplitud para los valores de y, z y w. Intentemos escogiendo valores arbitrarios para z y w. Sean $z = a \in \mathbb{R}$ y $w = b \in \mathbb{R}$. Para la incógnita y ya no podemos escoger cualquier valor, sino que lo encontramos despejando y de la segunda ecuación:

$$y = 4 + z - 2w = 4 + a - 2b.$$

Procedamos de igual manera con la primera ecuación. Como los valores de z y w ya están fijados, despejamos x:

$$x = -1 - 2z + 3w = -1 - 2a + 3b$$

Escribamos la solución del sistema $CX = D$ en forma de la matriz

$$S = \begin{pmatrix} -1 - 2a + 3b \\ 4 + a - 2b \\ a \\ b \end{pmatrix}, \quad a, b \in \mathbb{R}.$$

Como los parámetros a y b pueden asumir cualquier valor en los números reales tenemos que el sistema $CX = D$ tiene número infinito de soluciones. Por ejemplo, S_1 y S_2 son dos de dichas soluciones:

$$S_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 2 \end{pmatrix} \text{ si } a = 1,\ b = 2 \quad \text{y} \quad S_2 = \begin{pmatrix} 6 \\ 0 \\ -2 \\ 1 \end{pmatrix} \text{ si } a = -2,\ b = 1.$$

Podríamos verificar que S_1 y S_2 son soluciones del sistema $CX = D$ y también del sistema $AX = B$. Más bien, verifiquemos de una vez que la matriz S es solución del sistema original $AX = B$ que se nos dio a resolver. Sean $a, b \in \mathbb{R}$. Entonces:

$$AS = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & -1 \\ 3 & 1 & 5 & -7 \end{pmatrix} \begin{pmatrix} -1 - 2a + 3b \\ 4 + a - 2b \\ a \\ b \end{pmatrix}$$

$$= \begin{pmatrix} (-1-2a+3b)+2(4+a-2b)+0+b \\ (-1-2a+3b)+(4+a-2b)+a-b \\ 3(-1-2a+3b)+(4+a-2b)+5a-7b \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \\ 1 \end{pmatrix} = B.$$

Observemos que en este ejemplo:

$$\operatorname{ran} A = 2 = \operatorname{ran}(a \mid B)$$

y que el rango es menor que el número de incógnitas:

$$\operatorname{ran} A = \operatorname{ran}(a \mid B) = r = 2 < 4 = n.$$

Observemos también que el número de parámetros es igual al número de incógnitas menos el rango: $n - r = 4 - 2 = 2$. ♦

Ejemplo 3.33

Si es posible, resolver el sistema lineal

$$\begin{aligned} x + 2y + 3z &= 6, \\ -2x + 3y - 5z &= 4, \\ x + 9y + 4z &= 10. \end{aligned}$$

Solución. Sea $(A \mid B)$ la matriz ampliada del sistema lineal. Hallemos su matriz escalonada reducida por filas:

$$(A \mid B) = \begin{pmatrix} 1 & 2 & 3 & \mid & 6 \\ -2 & 3 & -5 & \mid & 4 \\ 1 & 9 & 4 & \mid & 10 \end{pmatrix}_{\substack{F_2+2F_1 \\ F_3-F_1}} \sim \begin{pmatrix} 1 & 2 & 3 & \mid & 6 \\ 0 & 7 & 1 & \mid & 16 \\ 0 & 7 & 1 & \mid & 4 \end{pmatrix}_{F_3-F_2}$$
$$\sim \begin{pmatrix} 1 & 2 & 3 & \mid & 6 \\ 0 & 7 & 1 & \mid & 16 \\ 0 & 0 & 0 & \mid & -12 \end{pmatrix} = (C \mid D).$$

El sistema equivalente $(CX = D)$ puede escribirse así:

$$\begin{aligned} x + 2y + 3z &= 6, \\ 7y + z &= 16, \\ 0x + 0y + 0z &= -12. \end{aligned}$$

Claramente se ve que no existen números x, y y z que satisfagan la tercera ecuación. Si la tercera ecuación no tiene solución, entonces el sistema lineal $AX = B$ tampoco tiene solución.

Observemos que, en este caso, el rango de la matriz de los coeficientes es menor que el rango de la matriz ampliada del sistema:

$$\operatorname{ran} A = 2 < 3 = \operatorname{ran}(A \mid B).$$

♦

En los tres ejemplos anteriores hemos encontrado que en los que existe solución, el rango de la matriz de los coeficientes es igual al rango de la matriz ampliada del sistema. Si, además, ese rango es igual al número de incógnitas la solución es única; pero si el rango es menor al número de incógnitas, el sistema lineal tiene número infinito de soluciones. También hemos encontrado que si el rango de la matriz de los coeficientes es menor que el rango de la matriz ampliada del sistema, entonces el sistema no tiene solución. Por supuesto, los ejemplos estudiados no son suficientes para concluir que esto siempre es así. Sin embargo, se puede demostrar que esto ocurre así efectivamente, por lo que enunciamos el siguiente teorema, e invitamos al lector a que realice la demostración por sí mismo.

Teorema 3.3 (Soluciones de un sistema lineal) Sea el sistema lineal de m ecuaciones con n incógnitas cuya representación matricial es $AX = B$.

1. Si $\operatorname{ran} A < \operatorname{ran}(A \mid B)$, el sistema no tiene solución.
2. Si $\operatorname{ran} A = \operatorname{ran}(A \mid B) = r$, el sistema tiene solución.
   (a) Si, además, $r = n$, tal solución es única.
   (b) Si, además, $r < n$, el sistema tiene número infinito de soluciones con $n - r$ parámetros.

Cuando el sistema lineal tiene número infinito de soluciones, estas se consiguen asignando parámetros a cualesquiera $n - r$ incógnitas. Por facilidad se asignan los parámetros a aquellas variables que corresponden a las columnas de la matriz de los coeficientes que no contienen el primer elemento no nulo de una fila.

En un sistema lineal homogéneo la matriz de términos independientes B es nula. Por tal razón, en un sistema homogéneo, siempre se cumple que $\operatorname{ran} A = \operatorname{ran}(A \mid B)$. El siguiente corolario expresa dicho resultado.

Corolario 3.4 Un sistema lineal homogéneo $AX = \mathbb{O}$ siempre tiene solución.

En los sistemas homogéneos tenemos otro criterio para determinar el número de soluciones que tiene el sistema.

Teorema 3.5 Si en el sistema $AX = \mathbb{O}$ hay menos ecuaciones que incógnitas; es decir, si la matriz A es de orden $m \times n$ con $m < n$, el sistema tiene un número infinito de soluciones.

Demostración. Sea R la matriz escalonada reducida por filas equivalente a A. Entonces el sistema $RX = \mathbb{O}$ tiene las mismas soluciones que el sistema $AX = \mathbb{O}$. Analicemos, pues, las soluciones del sistema con R.

Sea r el número de filas no nulas en R. Entonces $r \leq m < n$. Supongamos que el elemento principal no nulo de R de la fila i está en la columna k_i. Además, supongamos que las incógnitas x_{k_i} aparecen solamente en la i-ésima ecuación $RX = \mathbb{O}$ (con $i \in I_r$).

Ahora bien, sean $u_1, u_2, \ldots, u_{n-r}$ las incógnitas restantes, que son diferentes de x_{k_i}. Las ecuaciones del sistema tienen, entonces, la forma siguiente:

$$x_{k_i} + \sum_{j=1}^{n-r} \alpha_{ij} u_j = 0$$

para $i \in I_r$.

De esta última igualdad, para cada valor que se asignen a los u_j con $j = 1, 2, \ldots, n-r$, se obtienen soluciones no nulas del sistema $RX = \mathbb{O}$. Por lo tanto, este sistema tiene un número infinito de soluciones. $\square$

Si el sistema lineal $AX = B$ tiene igual número de ecuaciones que de incógnitas es más fácil saber si el sistema tiene o no solución con base en el conocimiento del determinante $|A|$. El siguiente teorema nos lo dice cómo.

Teorema 3.6 (Soluciones de un sistema cuadrado) Sea el sistema lineal de n ecuaciones con n incógnitas cuya representación matricial es $AX = B$.

1. Si $|A| \neq 0$, el sistema tiene solución única.
2. Si $|A| = 0$, el sistema o no tiene solución o tiene número infinito de soluciones.

Demostración. 1. Si $|A| \neq 0$, A es inversible. Entonces, existe A^{-1}, por lo que $AX = B$ implica $X = A^{-1}B$. Es decir, el sistema tiene solución. Por la unicidad de la inversa, si Y es una solución de $AX = B$, necesariamente se tiene que $Y = A^{-1}B$. Esto significa que el sistema tiene una sola solución.

2. Si $|A| = 0$, existe una matriz C, equivalente a A tal que todos los elementos de una fila son iguales al número 0. Supongamos que i sea el índice de esa fila. Sea D tal que $CX = D$. Si $d_i = 0$, estamos en el caso de que el sistema tiene menos ecuaciones que incógnitas, lo que significa que el sistema (cualquiera de los dos, ya que son equivalentes) tienen un número infinito de soluciones. En el caso de que $d_i \neq 0$, el sistema no tiene solución, pues el rango de C es menor que el rango de la matriz ampliada $(C \mid D)$.

$\square$

Si el sistema lineal es homogéneo tenemos el siguiente corolario.

Corolario 3.7 Sea el sistema lineal homogéneo de n ecuaciones con n incógnitas cuya representación matricial es $AX = O$.

1. Si $|A| = 0$, el sistema tiene número infinito de soluciones.
2. Si $|A| \neq 0$, el sistema tiene solución única, la trivial; es decir, la matriz columna nula.

Si tenemos un sistema lineal cuadrado que tiene solución única, se cuenta con un procedimiento que utiliza los determinantes, llamado regla de Cramer, para hallar su solución. No se presenta la demostración.

Teorema 3.8 (Regla de Cramer) La solución de un sistema lineal $AX = B$, con $A \in M_n$ inversible, viene dada por

$$x_i = \frac{1}{|A|}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{22} & \cdots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \cdots & a_{nn} \end{vmatrix}, \quad i = 1, 2, \ldots, n.$$

Ejemplo 3.34

Usar la regla de Cramer para encontrar la solución del sistema lineal

$$\begin{aligned} x + 2y + 3z &= 13, \\ 2x + 3y + z &= 10, \\ 3x + y + 2z &= 13. \end{aligned}$$

Solución. Sea $AX = B$ la representación matricial del sistema. Hallemos el determinante de la matriz de coeficientes:

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = -18 \neq 0.$$

Entonces el sistema tiene solución única que la hallamos con la regla de Cramer:

$$\begin{aligned} x &= \frac{1}{-18}\begin{vmatrix} 13 & 2 & 3 \\ 10 & 3 & 1 \\ 13 & 1 & 2 \end{vmatrix} = \frac{1}{-18}(-36) = 2, \\ y &= \frac{1}{-18}\begin{vmatrix} 1 & 13 & 3 \\ 2 & 10 & 1 \\ 3 & 13 & 2 \end{vmatrix} = \frac{1}{-18}(-18) = 1, \\ z &= \frac{1}{-18}\begin{vmatrix} 1 & 2 & 13 \\ 2 & 3 & 10 \\ 3 & 1 & 13 \end{vmatrix} = \frac{1}{-18}(-54) = 3. \end{aligned}$$

♦

3.4 Ejercicio resuelto

1. Sea el sistema lineal

$$\begin{aligned} -x + y - z &= m - p + 2, \\ (p+1)x + y + z &= 2m + 2, \\ x + z &= m, \\ (p+1)x - y + z &= p - 2. \end{aligned}$$

   Hallar los valores de $m, p \in \mathbb{R}$ para que el sistema

   (a) Tenga solución única.

   (b) Tenga un número infinito de soluciones.

   (c) No tenga solución.

   En los dos primeros casos, encuentre las soluciones.

Desarrollo. Sea $AX = B$ la representación matricial del sistema lineal. Hallemos la matriz escalonada reducida por filas de la matriz ampliada $(A \mid B)$:

$$(A \mid B) = \begin{pmatrix} -1 & 1 & -1 & \mid & m-p+2 \\ p+1 & 1 & 1 & \mid & 2m+2 \\ 1 & 0 & 1 & \mid & m \\ p+1 & -1 & 1 & \mid & p-2 \end{pmatrix}_{F_1+F_4} \sim \begin{pmatrix} p & 0 & 0 & \mid & m \\ p+1 & 1 & 1 & \mid & 2m+2 \\ 1 & 0 & 1 & \mid & m \\ p+1 & -1 & 1 & \mid & p-2 \end{pmatrix}_{\substack{F_2-F_1 \\ F_4-F_1}}$$

$$\sim \begin{pmatrix} p & 0 & 0 & \mid & m \\ 1 & 1 & 1 & \mid & m+2 \\ 1 & 0 & 1 & \mid & m \\ 1 & -1 & 1 & \mid & p-m-2 \end{pmatrix}_{F_2-F_3} \sim \begin{pmatrix} p & 0 & 0 & \mid & m \\ 0 & 1 & 0 & \mid & 2 \\ 1 & 0 & 1 & \mid & m \\ 1 & -1 & 1 & \mid & p-m-2 \end{pmatrix}_{F_4+F_2}$$

$$\sim \begin{pmatrix} p & 0 & 0 & \mid & m \\ 0 & 1 & 0 & \mid & 2 \\ 1 & 0 & 1 & \mid & m \\ 1 & 0 & 1 & \mid & p-m \end{pmatrix}_{F_4-F_3} \sim \begin{pmatrix} p & 0 & 0 & \mid & m \\ 0 & 1 & 0 & \mid & 2 \\ 1 & 0 & 1 & \mid & m \\ 0 & 0 & 0 & \mid & p-2m \end{pmatrix} = (C \mid D).$$

Dependiendo de los valores de m y p el rango de la matriz A sera igual o menor que el rango de la matriz ampliada $(A \mid B)$ y, por lo tanto, el sistema $AX = B$ tendrá o no solución. Estudiemos cada caso.

- Sea $p \neq 2m$. Entonces

$$\operatorname{ran} A = 2 < 3 = \operatorname{ran}(A \mid B)$$

  y el sistema $AX = B$ no tiene solución.

- Sea $p = 2m$. El sistema $AX = B$ se convierte en el sistema

$$\begin{aligned} 2mx &= m, \\ y &= 2, \\ x + z &= m. \end{aligned}$$

  Sea $EX = F$ su representación matricial. El determinante de la matriz de los coeficientes es

$$|E| = \begin{vmatrix} 2m & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{vmatrix} = 2m.$$

  El valor del determinante depende del valor de m. Entonces:

  - si $m = 0$, el determinate $|E|$ es igual a cero; el sistema $EX = F$ se convierte en el sistema

$$\begin{aligned} y &= 2, \\ x + z &= m. \end{aligned}$$

    el cual tiene número infinito de soluciones. Sea $z = r \in \mathbb{R}$. Entonces $x = -r$. El número infinito de soluciones viene dado por la matriz

$$S = \begin{pmatrix} -r \\ 2 \\ r \end{pmatrix}, \quad r \in \mathbb{R}.$$

– Si $m \neq 0$, el determinate $|E|$ es distinto de cero. El sistema $EX = F$ tiene, pues, una sola solución:

$$S = \begin{pmatrix} 1/2 \\ 2 \\ m - 1/2 \end{pmatrix}.$$

En resumen:

a) Sean $p, m \in \mathbb{R}$ tales que $p = 2m$ y $m \neq 0$. Entonces el sistema dado tiene la solución única

$$S = \begin{pmatrix} 1/2 \\ 2 \\ m - 1/2 \end{pmatrix}.$$

b) Sean $m = 0$ y $p = 0$. Entonces el sistema dado tiene número infinito de soluciones dados por la matriz

$$S = \begin{pmatrix} -r \\ 2 \\ r \end{pmatrix}, \quad r \in \mathbb{R}.$$

c) Sean $p, m \in \mathbb{R}$ tales que $p \neq 2m$. Entonces el sistema dado no tiene solución.

$\square$

Espacios vectoriales

por Felipe Navas

Edición y revisión: Juan Carlos Trujillo

Espacios vectoriales 4

4.1 Objetivos

El objetivo general de este capítulo es:

> *Fundamentar la dimensión de un subespacio vectorial ortogonal real a partir de una base ortonormal o de los elementos esenciales que lo tipifican en un tiempo máximo de* 30 *minutos para un subespacio vectorial de dimensión a los más* 4.

Para alcanzar el objetivo general, se proponen los siguientes objetivos específicos:

1. Identificar los espacios vectoriales (e.v.) usuales y sus propiedades a partir de la definición a un nivel de familiarización.

2. Identificar un subespacio vectorial de dimensión finita real a partir de las propiedades básicas que lo tipifican a un nivel reproductivo.

3. Identificar la dependencia lineal de un conjunto finito de *vectores* en un e.v. real, utilizando sistemas de ecuaciones lineales a un nivel reproductivo.

4. Caracterizar la cápsula lineal $\langle S \rangle$ a partir de un conjunto de *vectores* S finito y no vacío, a un nivel reproductivo. Los elementos de S puede ser vectores de $\mathbb{R}^n$ (con $n \leq 4$); polinomios de grado menor que o igual a 3; o matrices de orden $m \times n$ (con $m \leq 4$ y $n \leq 4$).

5. Fundamentar una base finita de un subespacio vectorial (s.e.v.) que tiene como elementos a vectores de $\mathbb{R}^n$, polinomios o matrices, a partir de los elementos esenciales que le tipifican, a un nivel reproductivo.

6. Determinar la intersección y la suma de s.e.v. reales finitos a partir de las propiedades básicas de conjuntos, a un nivel productivo.

7. Fundamentar la dimensión de uno o más s.e.v. reales finitos a partir de las propiedades esenciales que lo tipifican, a un nivel de asimilación productivo.

8. Clasificar las funciones a partir de la definición de producto interno (p.i.) en los espacios vectoriales más importantes, a un nivel reproductivo.

9. Calcular un conjunto de *vectores* ortogonales a partir de un conjunto linealmente independiente, usando el proceso de Gram-Schmidt, a un nivel reproductivo.

10. Fundamentar la suma directa de subespacios vectoriales a partir de las propiedades de los subespacios vectoriales ortogonales y sus bases, a un nivel constructivo.

4.2 Definición de *espacio vectorial*

Definición 4.1 (Espacio vectorial) Sean:

V: un conjunto no vacío.
$(\mathbb{K}, \oplus, \odot)$: un campo;
$+$: una operación interna definida sobre V;
$\cdot$: una operación externa definida de K en V.

Se dice que $(V, \mathbb{K}, +, \cdot)$ es un **espacio vectorial** si y solamente si:

1. **Conmutativa.** Para todo $u \in V$ y todo $v \in V$:
$$u + v = v + u.$$

2. **Asociativa.** Para todo $u \in V$, $v \in V$ y $w \in V$:
$$(u + v) + w = u + (v + w).$$

3. **Existencia del neutro aditivo.** Existe $e \in V$ tal que para todo $v \in V$:
$$e + v = v + e = v.$$

4. **Existencia del inverso aditivo.** Para todo $v \in V$, existe $\hat{v} \in V$ tal que:
$$\hat{v} + v = v + \hat{v} = e.$$

5. **Asociativa mixta.** Para todo $\alpha \in \mathbb{K}$, $\beta \in \mathbb{K}$ y todo $v \in V$:
$$(\alpha \odot \beta) \cdot v = \alpha \cdot (\beta \cdot v).$$

6. **Neutro multiplicativo.** Sea $1 \in \mathbb{K}$, el neutro multiplicativo de $\mathbb{K}$. Para todo $v \in V$:
$$1 \cdot v = v.$$

7. **Distributiva de $\cdot$ respecto de $+$.** Para todo $\alpha \in \mathbb{K}$ y todo $u \in V$ y $v \in V$:
$$\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v.$$

8. **Distributiva de $\cdot$ respecto de $\oplus$.** Para todo $\alpha \in \mathbb{K}$ y $\beta \in \mathbb{K}$, y todo $v \in V$:
$$(\alpha \oplus \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v).$$

Se suele decir que el conjunto V, **con las operaciones $+$ y $\cdot$, es un espacio vectorial sobre el campo $\mathbb{K}$.**

4.2.1 Observaciones

1. A los elementos del espacio vectorial V se les denomina *vectores* y a los elementos de campo $\mathbb{K}$, *escalares*.

2. En la definición de espacio vectorial hay cuatro operaciones distintas: dos del campo $\mathbb{K}$ y dos del conjunto V, a las que denominaremos las operaciones del espacio vectorial $(V, \mathbb{K}, +, \cdot)$.

   A pesar de ello, en la práctica, se suele utilizar $+$ también para $\oplus$. Al respecto de $\cdot$ y de $\odot$, ambos símbolos se suelen omitir (al igual que el punto de la multiplicación en los números reales).

   La razón para estas simplificaciones de notación se debe a que las operaciones $+$ y $\oplus$ pueden ser distinguidas una de otra fácilmente. Lo mismo sucede entre $\cdot$ y $\odot$.

   Considerando estas simplificaciones de notación, por ejemplo, las dos distributivas se escribirán de la siguiente manera:

$$\alpha(u+v) = (\alpha u) + (\alpha v) \quad \text{y} \quad (\alpha+\beta)v = \alpha v + \beta v.$$

3. Que la operación $+$ sea una *operación interna* definida sobre V significa que para todo $u \in V$ y todo $v \in V$, debe ocurrir que:

$$(u+v) \in V.$$

   Esto se indica diciendo que la operación $+$ es *cerrada* y que satisface la propiedad *clausurativa*.

   Que la operación $\cdot$ sea una *operación externa* definida de $\mathbb{K}$ en V quiere decir que para todo $\alpha \in \mathbb{K}$ y todo $v \in V$, debe ocurrir que:

$$\alpha v \in V.$$

   Esto se expresa diciendo que la operación $\cdot$ es *cerrada* y que satisface la propiedad *clausurativa*.

4. Para la verificación de que un cierto conjunto es un espacio vectorial, hay que cerciorarse antes que las operaciones $+$ y $\cdot$ son, efectivamente, cerradas.

5. Si en $(V, \mathbb{K}, +, \cdot)$, están sobreentendidas las operaciones $+$ y $\cdot$, pero no el campo, se suele decir que V es un espacio vectorial sobre K.

4.2.2 Ejemplos

Los espacios vectoriales con los cuales trabajaremos principalmente son lo siguientes.

1. **El conjunto de los números reales sobre sí mismo:** $(\mathbb{R}, \mathbb{R}, +, \cdot)$.

   Si las operaciones $+$ y $\cdot$ son la suma y producto usual entre números reales, el conjunto $\mathbb{R}$ es un espacio vectorial sobre sí mismo, ya que $\mathbb{R}$ es un campo.

2. **El conjunto $\mathbb{R}^2$ sobre $\mathbb{R}$: $(\mathbb{R}^2, \mathbb{R}, +, \cdot)$.**

   El conjunto de pares ordenados de números reales

   $$\mathbb{R}^2 = \{(x_1, x_2) : x_1 \in \mathbb{R} \wedge x_2 \in \mathbb{R}\}$$

   es un espacio vectorial sobre el campo $\mathbb{R}$ si se definen las operaciones $+$ y $\cdot$ de la siguiente manera.

   Sean $u = (x_1, x_2) \in \mathbb{R}^2$, $v = (y_1, y_2) \in \mathbb{R}^2$, y $\alpha \in \mathbb{R}$. Entonces, se define:

   (a) $u + v = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$; y
   (b) $\alpha u = \alpha(x_1, x_2) = (\alpha x_1, \alpha x_2)$.

3. **El conjunto $\mathbb{R}^3$ sobre $\mathbb{R}$: $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$.**

   El conjunto de triadas ordenadas de números reales

   $$\mathbb{R}^3 = \{(x_1, x_2, x_3) : x_1 \in \mathbb{R} \wedge x_2 \in \mathbb{R} \wedge x_3 \in \mathbb{R}\}$$

   es un espacio vectorial sobre el campo $\mathbb{R}$ si se definen las operaciones $+$ y $\cdot$ de la siguiente manera.

   Sean $u = (x_1, x_2, x_3) \in \mathbb{R}^3$, $v = (y_1, y_2, y_3) \in \mathbb{R}^3$, y $\alpha \in \mathbb{R}$. Entonces, se define:

   (a) $u + v = (x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$; y
   (b) $\alpha u = \alpha(x_1, x_2, x_3) = (\alpha x_1, \alpha x_2, \alpha x_3)$.

4. **El conjunto $\mathbb{R}^n$ sobre $\mathbb{R}$: $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$ con $n > 3$.**

   El conjunto de n-tuplas ordenadas de números reales

   $$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R},\ 1 \leq i \leq n\}$$

   es un espacio vectorial sobre el campo $\mathbb{R}$ si se definen las operaciones $+$ y $\cdot$ de la siguiente manera.

   Sean $u = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $v = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, y $\alpha \in \mathbb{R}$. Entonces, se define:

   (a) $u + v = (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$; y
   (b) $\alpha u = \alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$.

5. **El conjunto $\mathbb{K}^n$ sobre $\mathbb{K}$, donde K es un campo: $(\mathbb{K}^n, \mathbb{K}, +, \cdot)$ con $n \geq 1$.**

   El conjunto de n-tuplas ordenadas de elementos del campo $\mathbb{K}$

   $$\mathbb{K}^n = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R},\ 1 \leq i \leq n\}$$

   es un espacio vectorial sobre el campo $\mathbb{K}$ si se definen las operaciones $+$ y $\cdot$ de la siguiente manera.

   Sean $u = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, $v = (y_1, y_2, \ldots, y_n) \in \mathbb{K}^n$, y $\alpha \in \mathbb{K}$. Entonces, se define:

   (a) $u + v = (x_1 \oplus y_1, x_2 \oplus y_2, \ldots, x_n \oplus y_n)$; y
   (b) $\alpha u = (\alpha \odot x_1, \alpha \odot x_2, \ldots, \alpha \odot x_n)$,

donde $\oplus$ y $\odot$ son las operaciones internas del campo $\mathbb{K}$.

6. **El conjunto $\mathbb{M}_{m\times n}$ sobre $\mathbb{R}$: $(\mathbb{M}_{m\times n}, \mathbb{R}, +, \cdot)$.**

   El conjunto de las matrices de números reales de orden m por n

   $$\mathbb{M}_{m\times n} = \left\{A = (a_{ij})_{m\times n} : a_{ij} \in \mathbb{R},\ 1 \leq i \leq m,\ 1 \leq j \leq n\right\}$$

   es un espacio vectorial sobre $\mathbb{R}$ si la operación $+$ es l suma usual entre matrices y $\cdot$ es el producto usual entre matrices.

   Es decir, si $A = (a_{ij}) \in \mathbb{M}_{m\times n}$, $B = (b_{ij}) \in \mathbb{M}_{m\times n}$, y $\alpha \in \mathbb{R}$, entonces:

   (a) $A + B = (a_{ij} + b_{ij})$; y
   (b) $\alpha A = (\alpha a_{ij})$.

7. **El conjunto de polinomios $\mathscr{P}_n[x]$ sobre $\mathbb{R}$: $(\mathscr{P}_n[x], \mathbb{R}, +, \cdot)$.**

   El conjunto de los polinomios con coeficientes reales de grado menor que o igual a n

   $$\mathscr{P}_n[x] = \left\{p : p(x) = \sum_{i=0}^{k} a_j x^j, x \in \mathbb{R},\ a_j \in \mathbb{R},\ 1 \leq j \leq k,\ k \leq n\right\}$$

   es un espacio vectorial sobre $\mathbb{R}$ si la operación $+$ es la suma usual entre polinomios y $\cdot$ es el producto usual entre un número real y un polinomio.

   Es decir, si $p \in \mathscr{P}_n[x]$ y $q \in \mathscr{P}_n[x]$, y $\alpha \in \mathbb{R}$, entonces:

   (a) $(p + q)(x) = p(x) + q(x)$; y
   (b) $(\alpha p)(x) = \alpha p(x)$

   para todo $x \in \mathbb{R}$.

8. **El conjunto de las funciones reales $\mathscr{F}$ sobre $\mathbb{R}$: $(\mathscr{F}, \mathbb{R}, +, \cdot)$.**

   El conjunto de las funciones reales

   $$\mathscr{F} = \{f : f \colon \mathbb{R} \longrightarrow \mathbb{R}\}$$

   es un espacio vectorial sobre $\mathbb{R}$ si la operación $+$ es la suma usual entre funciones y $\cdot$ es el producto usual entre un número real y una función.

   Es decir, si $f \in \mathscr{F}$ y $g \in \mathscr{F}$, y $\alpha \in \mathbb{R}$, entonces:

   (a) $(f + g)(x) = f(x) + g(x)$; y
   (b) $(\alpha f)(x) = \alpha f(x)$

   para todo $x \in \mathbb{R}$.

Utilizaremos estos espacios vectoriales para ilustrar cómo se demuestran cada una de las propiedades que un conjunto, sobre el cual se han definido dos operaciones, debe satisfacer para que sea un espacio vectorial.

1. **La operación $+$ en $(\mathbb{R}^2, \mathbb{R}, +, \cdot)$ es asociativa; es decir, se verifica que para todo $\boldsymbol{u} \in \mathbb{R}^2, \boldsymbol{v} \in \mathbb{R}^2$ y $\boldsymbol{w} \in \mathbb{R}^2$:**

   $$(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}).$$

Demostración. Supongamos que $u = (x_1, x_2)$, $v = (y_1, y_2)$ y $w = (z_1, z_2)$. Entonces:

$$
\begin{aligned}
(u+v)+w &= [(x_1,x_2)+(y_1,y_2)]+(z_1,z_2) && \text{definiciones de } u,\ v \text{ y } w,\\
&= (x_1+y_1, x_2+y_2)+(z_1,z_2) && \text{definición de + en } \mathbb{R}^2,\\
&= ([x_1+y_1]+z_1, [x_2+y_2]+z_2) && \text{definición de + en } \mathbb{R}^2,\\
&= (x_1+[y_1+z_1], x_2+[y_2+z_2]) && \text{asociativa de } \mathbb{R},\\
&= (x_1,x_2)+(y_1+z_1, y_2+z_2) && \text{definición de + en } \mathbb{R}^2,\\
&= (x_1,x_2)+[(y_1,y_2)+(z_1,z_2)] && \text{definición de + en } \mathbb{R}^2,\\
&= u+(v+w) && \text{definiciones de } u,\ v \text{ y } w.
\end{aligned}
$$

Por lo tanto, para todo u, v y w en $\mathbb{R}^2$:

$$(u+v)+w = u+(v+w).$$

□

2. **Existencia del neutro aditivo en $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$; es decir, debemos probar que la siguiente proposición es verdadera:**

$$(\exists \boldsymbol{e} \in \mathbb{R}^3)(\forall \boldsymbol{v} \in \mathbb{R}^3)(\boldsymbol{e}+\boldsymbol{v} = \boldsymbol{v}+\boldsymbol{e} = \boldsymbol{v}).$$

La demostración que viene a continuación es de *existencia*. Esto quiere decir que debemos exhibir un elemento de $e \in \mathbb{R}^3$ que satisfaga la definición de elemento neutro. Por el orden de los cuantificadores, el elemento buscado e es independiente de todo elemento de $\mathbb{R}^3$.

Demostración. Sea $e = (0,0,0)$. Vamos a probar que para todo $v = (x_1, x_2, x_3)$, con $x_i \in \mathbb{R}$ para $i = 1, 2, 3$, se verifica la igualdad $e + v = v$.

En efecto:

$$
\begin{aligned}
e+v &= (0,0,0)+(x_1,x_2,x_3) && \text{definiciones de } e \text{ y } v,\\
&= (0+x_1, 0+x_2, 0+x_3) && \text{definición de + en } \mathbb{R}^3,\\
&= (x_1,x_2,x_3) = v && \text{inverso aditivo en } \mathbb{R}.
\end{aligned}
$$

Por lo tanto, $e + v = v$.

Por la propiedad conmutativa, tenemos que $e + v = v + e$. Por lo tanto, se ha demostrado también que $v + e = v$.

Entonces, se ha demostrado que la proposición

$$(\exists e \in \mathbb{R}^3)(\forall v \in \mathbb{R}^3)(e+v = v+e = v)$$

es verdadera. □

No es difícil probar, además, que el elemento neutro en cualquier espacio vectorial es único. En este texto, lo representaremos con 0_v.

3. **Existencia del inverso aditivo en $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$; es decir, debemos probar que la siguiente proposición es verdadera:**

$$(\forall \boldsymbol{v} \in \mathbb{R}^n)(\exists \widehat{\boldsymbol{v}} \in \mathbb{R}^n)(\widehat{\boldsymbol{v}} + \boldsymbol{v} = \boldsymbol{v} + \widehat{\boldsymbol{v}} = \boldsymbol{0_v}).$$

Esta también es una demostración de existencia. En este caso, por el orden de los cuantificadores, el elemento $\widehat{v}$ depende de v. Esto quiere decir que, para valores diferentes de v, $\widehat{v}$ también puede ser diferente.

Es obvio que para mostrar esta igualdad, ya debemos conocer a 0_v en $\mathbb{R}^n$. No es difícil probar que $0_v = (0, 0, \cdot, 0)$ (n ceros).

Demostración. Sea $v = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$. Vamos a probar que el vector buscado $\widehat{v}$ se define del siguiente modo:

$$\widehat{v} = (-x_1, -x_2, \cdots, -x_n),$$

donde cada $-x_i$ es el inverso aditivo de $x_i \in \mathbb{R}$, para $i = 1, 2, \ldots, n$.

Entonces, tenemos que:

$$\begin{aligned}
v + \widehat{v} &= (x_1, x_2, \cdots, x_n) + (-x_1, -x_2, \cdots, -x_n) && \text{definiciones de } v \text{ y } \widehat{v}, \\
&= (x_1 + (-x_1), x_2 + (-x_2), \ldots, x_n + (-x_n)) && \text{definición de } + \text{ en } \mathbb{R}^n, \\
&= (0, 0, \ldots, 0) && \text{inverso aditivo en } \mathbb{R}, \\
&= 0_v && \text{definición de } 0_v \text{ en } \mathbb{R}.
\end{aligned}$$

Por lo tanto, $v + \widehat{v} = 0_v$. Y, por la propiedad conmutativa de $+$, se tiene también que $\widehat{v} + v = 0_v$.

En resumen, hemos demostrado que

$$(\forall v \in \mathbb{R}^n)(\exists \widehat{v} \in \mathbb{R}^n)(\widehat{v} + v = v + \widehat{v} = 0_v).$$

□

Se puede demostrar fácilmente que el elemento inverso es único. Con $-v$ se suele representar el inverso aditivo de v.

4. **La operación $+$ en $(\mathbb{K}^n, \mathbb{K}, +, \cdot)$ es conmutativa; es decir, se verifica para todo $u \in \mathbb{K}^n$ y todo $v \in \mathbb{K}^n$ que:**

$$\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}.$$

Demostración. Sean $u = (x_1, x_2, \cdots, x_n) \in \mathbb{K}^n$ y $v = (y_1, y_2, \cdots, y_n) \in \mathbb{K}$, donde cada x_i y y_i son elementos del campo $\mathbb{K}$ para todo $i = 1, 2, \ldots, n$. Entonces, como la operación $+$ en el campo $\mathbb{K}$ es conmutativa, tenemos que:

$$\begin{aligned}
u + v &= (x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) && \text{definiciones de } u \text{ y } v, \\
&= (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) && \text{definición de } + \text{ en } \mathbb{R}^n, \\
&= (y_1 + x_1, y_2 + x_2, \cdots, y_n + x_n) && \text{conmutativa de } + \text{ en } \mathbb{R}, \\
&= (y_1, y_2, \cdots, y_n) + (x_1, x_2, \cdots, x_n) && \text{definición de } + \text{ en } \mathbb{R}^n, \\
&= v + u && \text{definiciones de } v \text{ y } u.
\end{aligned}$$

Por lo tanto, $u + v = v + u$. □

5. **Las operaciones $+$ y $\cdot$ en el espacio $(\mathbb{M}_{m\times n}, \mathbb{R}, +, \cdot)$ satisface la propiedad asociativa mixta; es decir, se verifica para todo $\alpha \in \mathbb{R}$, todo $\beta \in \mathbb{R}$ y toda $A \in \mathbb{M}_{m\times n}$ que:**
$$\boldsymbol{(\alpha\beta)A = \alpha(\beta A)}.$$

Demostración. Sean $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$ y $A = \big(a_{ij}\big)_{m\times n} \in \mathbb{M}_{m\times n}$, con $a_{ij} \in \mathbb{R}$. Entonces:

$$\begin{aligned}
(\alpha\beta)A &= (\alpha\beta)\big(a_{ij}\big)_{m\times n} && \text{definición de } A,\\
&= ((\alpha\beta)a_{ij})_{m\times n} && \text{definición de } \cdot \text{ en } \mathbb{M}_{m\times n},\\
&= (\alpha(\beta a_{ij}))_{m\times n} && \text{asociativa del producto en } \mathbb{R},\\
&= \alpha\,(\beta a_{ij})_{m\times n} && \text{definición de } \cdot \text{ en } \mathbb{M}_{m\times n},\\
&= \alpha\left(\beta\,(a_{ij})_{m\times n}\right) && \text{definición de } \cdot \text{ en } \mathbb{M}_{m\times n},\\
&= \alpha(\beta A) && \text{definición de } A.
\end{aligned}$$

Por lo tanto, $(\alpha\beta)A = \alpha(\beta A)$. □

Esta propiedad es denominada *asociativa mixta* porque tiene que ver con dos operaciones: la multiplicación de un escalar por un elemento del espacio vectorial y la multiplicación entre dos escalares.

6. **El neutro multiplicativo del campo $\mathbb{R}$, el número 1 satisface la siguiente propiedad:**
$$\boldsymbol{(\forall p \in \mathscr{P}_n[x])(1 \cdot p = p)}.$$

Hay que demostrar que la función $1 \cdot p$ es igual a la función p. Por ello, recordemos que dos funciones f y g son iguales si y solo si tienen el mismo dominio y $f(x) = g(x)$ para todo x elemento del dominio.

Demostración. Sea $p \in \mathscr{P}_n[x]$. El dominio de p es $\mathbb{R}$. Entonces, debemos demostrar que
$$(1 \cdot p)(x) = p(x)$$
para todo $x \in \mathbb{R}$.

Sea $x \in \mathbb{R}$. Entonces:

$$\begin{aligned}
(1 \cdot p)(x) &= 1 \cdot p(x) && \text{definición de } \cdot \text{ en } \mathscr{P}_n[x],\\
&= p(x) && \text{el 1 es el neutro multiplicativo en } \mathbb{R}.
\end{aligned}$$

Por lo tanto, $(1 \cdot p)(x) = p(x)$ para todo $x \in \mathbb{R}$, lo que prueba, a su vez, que es verdadera la siguiente proposición:
$$(\forall p \in \mathscr{P}_n[x])(1 \cdot p = p).$$

□

7. **Las operaciones $+$ y $\cdot$ del espacio vectorial de funciones reales $(\mathscr{F}, \mathbb{R}, +, \cdot)$ satisfacen la propiedad distributiva; es decir, se verifica para todo** $\alpha \in \mathbb{R}$, **todo** $f \in \mathscr{F}$ **y todo** $g \in \mathscr{F}$ **que**:

$$\boldsymbol{\alpha(f+g) = \alpha f + \alpha g}.$$

Demostración. Sean $\alpha \in \mathbb{R}$, $f \in \mathscr{F}$, $g \in \mathscr{F}$. Debemos demostrar que:

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

para todo $x \in \mathbb{R}$, ya que el el conjunto $\mathbb{R}$ es el dominio de f, de g y de $f+g$, respectivamente.

Sea $x \in \mathbb{R}$. Entonces:

$$\begin{aligned}
(\alpha(f+g))(x) &= \alpha(f+g)(x) && \text{definición de } \cdot \text{ en } \mathscr{F},\\
&= \alpha(f(x)+g(x)) && \text{definición de } + \text{ en } \mathscr{F},\\
&= \alpha f(x) + \alpha g(x) && \text{distributiva en } \mathbb{R},\\
&= (\alpha f)(x) + (\alpha g)(x) && \text{definición de } \cdot \text{ en } \mathscr{F},\\
&= (\alpha f + \alpha g)(x) && \text{definición de } + \text{ en } \mathscr{F}.
\end{aligned}$$

Por lo tanto, hemos probado que

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

para todo $x \in \mathbb{R}$, con lo cual hemos probado que:

$$(\forall \alpha \in \mathbb{R})(\forall f \in \mathscr{F})(\forall g \in \mathscr{F})(\alpha(f+g) = \alpha f + \alpha g).$$

□

En una expresión vectorial, los signos de agrupación determinan el orden que se deben realizar las operaciones. En ausencia de ellos, se conviene en efectuar las operaciones externas en primer lugar, y luego las operaciones internas, siempre de izquierda a derecha.

Por ejemplo:

$$\begin{aligned}
2(3,4) + (-1,5) + 3[(-1,0)+(0,-1)] &= (6,8) + (-1,5) + 3(-1,-1)\\
&= (6,8) + (-1,5) + (-3,-3)\\
&= (2,10).
\end{aligned}$$

4.3 Propiedades de los espacios vectoriales

Teorema 4.1 (Propiedades aditiva y multiplicativa) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $\alpha \in \mathbb{K}$, $u \in V$, $v \in V$ y $w \in V$. Se verifican las siguientes propiedades:

1. **Aditiva**: si $u = v$, entonces $u + w = v + w$.

2. **Multiplicativa**: si $u = v$, entonces $\alpha u = \alpha v$.

Demostración. Sean $\alpha \in \mathbb{K}$, $u \in V$, $v \in V$ y $w \in V$.

1. **Aditiva**:

u	$=$	v	hipótesis,
$u + w$	$=$	$u + w$	axioma de identidad de la igualdad,
$u + w$	$=$	$v + w$	sustitución de v por u en la igualdad anterior ya que $u = v$.

2. **Multiplicativa**:

u	$=$	v	hipótesis,
αu	$=$	αu	axioma de identidad de la igualdad,
αu	$=$	αv	v se sustituye por u en la igualdad anterior ya que $u = v$.

□

Teorema 4.2 (Propiedades cancelativas) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $\alpha \in \mathbb{K}$, $u \in V$, $v \in V$ y $w \in V$. Se verifican las siguientes propiedades:

1. **Cancelativa de la suma**: si $u + w = v + w$, entonces $u = v$.
2. **Cancelativa de la multiplicación**: si $\alpha \neq 0$ y $\alpha \cdot u = \alpha \cdot v$, entonces $u = v$.

Demostración. Sean $\alpha \in \mathbb{K}$, $u \in V$, $v \in V$ y $w \in V$.

1. **Cancelativa de la suma**:

1.	$u + w = v + w$	hipótesis,
2.	$(u + w) + (-w) = (v + w) + (-w)$	propiedad aditiva,
3.	$u + (w + (-w)) = v + (w + (-w))$	propiedad asociativa,
4.	$u + 0_v = v + 0_v$	definición de inverso aditivo,
5.	$u = v$	definición de 0_v.

2. **Cancelativa de la multiplicación**: sea $\alpha \neq 0$; entonces, existe α^{-1}:

1.	$\alpha \cdot u = \alpha \cdot v$	hipótesis,
2.	$\alpha^{-1} \cdot [\alpha \cdot u] = \alpha^{-1} \cdot [\alpha \cdot v]$	propiedad multiplicativa,
3.	$(\alpha^{-1}\alpha) \cdot u = (\alpha^{-1}\alpha) \cdot v$	propiedad asociativa mixta,
4.	$1 \cdot u = 1 \cdot v$	definición de inverso multiplicativo en $\mathbb{K}$,
5.	$u = v$	definición de 1.

□

Teorema 4.3 (Propiedades de los espacios vectoriales) Sea $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial. Entonces, para todo $\alpha \in \mathbb{K}$ y todo $v \in V$, se tiene que:

1. $\alpha \cdot 0_v = 0_v$.
2. $0 \cdot v = 0_v$.
3. $a \cdot v = 0_v \Leftrightarrow \alpha = 0 \vee v = 0_v$.
4. $(-1) \cdot v = -v$.

Demostraciones. En todo lo que sigue, α representará un elemento cualquiera de $\mathbb{K}$ y v uno cualquiera de V.

1. Ya que 0_v es el neutro aditivo de V, tenemos que:

$$\alpha \cdot 0_v = \alpha \cdot 0_v + 0_v. \tag{4.1}$$

Por otro lado, la definición de 0_v y la propiedad distributiva en V nos permite escribir lo siguiente:

$$\alpha \cdot 0_v = \alpha \cdot (0_v + 0_v) = \alpha \cdot 0_v + \alpha \cdot 0_v. \tag{4.2}$$

Entonces, por la propiedad transitiva de la igualdad aplicada a (4.1) y (4.2), tenemos que:

$$\alpha \cdot 0_v + \alpha \cdot 0_v = \alpha \cdot 0_v + 0_v.$$

De esta igualdad se colige que $\alpha \cdot 0_v$ también es un neutro aditivo en V. Pero 0_v es único. Entonces:

$$\alpha \cdot 0_v = 0_v.$$

2. Como 0_v es el neutro aditivo de V, tenemos que:

$$0 \cdot v = 0 \cdot v + 0_v. \tag{4.3}$$

Por otro lado, la definición de $0 \in \mathbb{K}$ y la propiedad distributiva en V nos permite escribir lo siguiente:

$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v. \tag{4.4}$$

Entonces, por la propiedad transitiva de la igualdad aplicada a (4.3) y (4.4), tenemos que:

$$0 \cdot v + 0 \cdot v = 0 \cdot v + 0_v.$$

De esta igualdad se concluye que $0 \cdot v$ también es un neutro aditivo en V. Pero 0_v es único. Entonces:

$$0 \cdot v = 0_v.$$

3. En primer lugar, demostremos que si $\alpha \cdot v = 0_v$, entonces o bien $\alpha = 0$ o bien $v = 0_v$ (o ambos). Utilicemos el método por contradicción.

Para ello, supongamos que $\alpha \cdot v = 0_v$, y que $\alpha \neq 0$ y $v \neq 0_v$. Entonces, existe α^{-1} y:

$$\begin{array}{rcll}
\alpha^{-1} \cdot (\alpha \cdot v) & = & \alpha^{-1} \cdot 0_v & \text{multiplicativa,} \\
\alpha^{-1}\alpha)v & = & \alpha^{-1} \cdot 0_v & \text{asociativa mixta,} \\
1 \cdot v & = & \alpha^{-1} \cdot 0_v & \text{definición de } \alpha^{-1}, \\
v & = & \alpha^{-1} \cdot 0_v & \text{definición de } 1, \\
v & = & 0_v & \text{definición de } 0_v.
\end{array}$$

Por lo tanto $v = 0_v$. Pero esta igualdad contradice la hipótesis de que $v \neq 0_v$.

Entonces, lo supuesto es falso; es decir, es verdad que $\alpha = 0$ o $v = 0_v$ siempre que $\alpha \cdot v = 0_v$.

Ahora debemos probar la recíproca: si $\alpha = 0$ o $v = 0_v$, entonces $\alpha \cdot v = 0_v$. Pero esta implicación no es más que la disyunción de las dos primeras propiedades que ya hemos demostrado que son verdaderas.

4. Vamos a demostrar que $(-1) \cdot v = -v$:

$$
\begin{aligned}
(-1) \cdot v &= (-1) \cdot v + 0_v && \text{definición de } 0_v, \\
&= (-1) \cdot v + (v + (-v)) && \text{definición de } -v, \\
&= [(-1) \cdot v + v] + (-v) && \text{asociativa en } V, \\
&= [(-1) \cdot v + 1 \cdot v] + (-v) && \text{axioma del neutro multiplicativo en } V, \\
&= (-1 + 1) \cdot v + (-v) && \text{distributiva en } V, \\
&= 0 \cdot v + (-v) && \text{axioma del inverso aditivo en } \mathbb{K}, \\
&= 0_v + (-v) && \text{segunda propiedad de este teorema,} \\
&= -v && \text{definición de } 0_v.
\end{aligned}
$$

Por lo tanto, hemos demostrado que $(-1) \cdot v = -v$.

□

4.4 Subespacios vectoriales

Definición 4.2 (Subespacio vectorial) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $W \subset V$. Se dice que W *es un subespacio vectorial* (s.e.v.) de V si y solamente si $(W, \mathbb{K}, \oplus, \odot)$ es un espacio vectorial, donde $\oplus$ y $\odot$ son las restricciones de $+$ y $\cdot$ a W.

Puesto que

$$u \oplus v = u + v \quad \text{y} \quad u \odot v = u \cdot v$$

para todo $u \in W$ y $v \in W$, decir que W es un subespacio vectorial de V es equivalente a decir que W es un espacio vectorial sobre $\mathbb{K}$ con las "mismas operaciones" de V. Por esta razón, se acostumbra utilizar la misma notación para las operaciones de W que de V, aunque, en sentido estricto, son diferentes porque poseen dominios diferente. En adelante, seguiremos esta convención.

Ejemplo 4.35

El conjunto

$$W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 2x_1 + x_2\}$$

es un subespacio vectorial de $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$.

Demostración. En primer lugar, debemos asegurarnos de que la restricciones de las operaciones $+$ y $\cdot$ de $\mathbb{R}^3$ a W, respectivamente, sean una operación interna y una externa en W. En otras palabras, debemos asegurarnos de que las restricciones sean cerradas:

1. si $u \in W$ y $v \in W$, entonces $u + v \in W$; y
2. si $\alpha \in \mathbb{K}$ y $u \in W$, entonces $\alpha u \in W$.

En efecto:

1. **Clausurativa de $+$ en W.**
   Sean $u = (x_1, x_2, x_3) \in W$ y $v = (y_1, y_2, y_3) \in W$. Vamos a probar que $u + v = (x_1 + y_1, x_2 + y_2, x_3 + y_3) \in W$; es decir, vamos a probar que
   $$x_3 + y_3 = 2(x_1 + y_1) + (x_2 + y_2).$$
   Como $u \in W$ y $v \in W$, tenemos que
   $$x_3 = 2x_1 + x_2 \quad \text{y} \quad y_3 = 2y_1 + y_2.$$
   Si sumamos los lados izquierdos entre sí y los derechos entre sí de estas dos igualdades, obtenemos, al aplicar las propiedades asocitiva y conmutativa de $+$ y $\cdot$ en $\mathbb{R}$, que:
   $$x_3 + y_3 = (2x_1 + x_2) + (2y_1 + y_2) = (2x_1 + 2y_1) + (y_1 + y_2).$$
   Por lo tanto, gracias a la propiedad distributiva en $\mathbb{R}$, concluimos finalmente que:
   $$x_3 + y_3 = 2(x_1 + y_1) + (x_2 + y_2).$$
   Es decir, hemos probado que $u + v \in W$ para todo $u \in W$ y todo $v \in W$.
2. **Clausurativa de $\cdot$ en W.**
   Sean $u = (x_1, x_2, x_3) \in W$ y $\alpha \in \mathbb{R}$. Probemos que $\alpha u = (\alpha x_1, \alpha x_2, \alpha x_3) \in W$:
   $$\begin{aligned} \alpha x_3 &= \alpha(2x_1 + x_2) && u \in W, \\ &= 2(\alpha x_1) + \alpha x_2 && \text{distributiva y conmutativa en } \mathbb{R}. \end{aligned}$$
   Por lo tanto $\alpha u \in W$ para todo $u \in W$ y todo $\alpha \in \mathbb{R}$.

Una vez que hemos constatado que las restricciones de $+$ y $\cdot$ a W son cerradas, podemos proceder a verificar el cumplimiento de los ocho axiomas que definen un espacio vectorial. Como van a interactuar elementos de V y de W, utilizaremos $\oplus$ y $\odot$ para las operaciones de W. El lector puede probar a escribir las mismas demostraciones que están a continuación sin realizar la distinción entre ambos grupos de operaciones, y podrá constatar por sí mismo o por sí misma la pérdida de claridad.

1. **Conmutativa**: $(\forall u \in W)(\forall v \in W)(u \oplus v = v \oplus u)$:
   $$\begin{aligned} u \oplus v &= u + v && \oplus \text{ es restricción de } + \text{ a } W \text{ y } u \in W \text{ y } v \in W, \\ &= v + u && + \text{ es conmutativa}, \\ &= v \oplus u && \oplus \text{ es restricción de } + \text{ a } W \text{ y } u \in W \text{ y } v \in W. \end{aligned}$$
2. **Asociativa**: $(\forall u \in W)(\forall v \in W)(\forall v \in W)(u \oplus v = v \oplus u)$:
   $$\begin{aligned} u \oplus (v \oplus w) &= u + (v + w) && \oplus \text{ es restricción de } + \text{ a } W \text{ y } u \in W,\ v \in W \text{ y } w \in W, \\ &= (u + v) + w && + \text{ es asociativa}, \\ &= (u \oplus v) \oplus w && \oplus \text{ es restricción de } + \text{ a } W \text{ y } u \in W,\ v \in W \text{ y } w \in W. \end{aligned}$$

3. **Existencia del neutro aditivo**: $(\exists e \in W)(\forall w \in W)(w \oplus e = w)$:
Definamos $e = 0_{\mathbb{R}^3} = (0,0,0)$. Veamos que $e \in W$:

$$0 = 2 \cdot 0 + 0 = 0 + 0 = 0.$$

Por lo tanto, si $w \in W$, tenemos que $w \in V$, de donde:

$$w \oplus e = w + e = w + 0_{\mathbb{R}^3} = w.$$

4. **Existencia del inverso aditivo**: $(\forall w \in W)(\exists \hat{w} \in W)(w + \hat{w} = 0_W)$:
Definamos $\hat{w} = -w = (-x_1, -x_2, -x_3)$, donde $w = (x_1, x_2, x_3)$. Veamos que $-w \in W$:

$$-x_3 = -(2x_1 + x_2) = 2(-x_1) + (-x_2).$$

Además, por la definición de $-w$, se verifica que:

$$w \oplus \hat{w} = w + (-w) = 0_{\mathbb{R}^3}.$$

5. **Asociativa mixta:** $(\forall \alpha \in \mathbb{R})(\forall \beta \in \mathbb{R})(\forall w \in W)(\alpha \odot (\beta \odot w) = (\alpha\beta) \odot w)$**:**

$$\begin{aligned} \alpha \odot (\beta \odot w) &= \alpha(\beta w) && \odot \text{ es restricción de } \cdot \text{ a } W,\ w \in W \text{ y } \beta w \in W, \\ &= (\alpha\beta)w && \text{Asociativa mixta en } \mathbb{R}^3, \\ &= (\alpha\beta) \odot w && \odot \text{ es restricción de } \cdot \text{ a } W \text{ y } w \in W. \end{aligned}$$

6. **Neutro multiplicativo**: $(\forall w \in W)(1 \odot w = w)$, donde $1 \in \mathbb{K}$:

$$\begin{aligned} 1 \odot w &= 1 \cdot w && \odot \text{ es restricción de } \cdot \text{ a } W \text{ y } w \in W, \\ &= w && 1 \text{ es el neutro multiplicativo en } V \text{ y } w \in V. \end{aligned}$$

7. **Distributiva**: $(\forall u \in W)(\forall v \in W)(\forall \alpha \in \mathbb{K})(\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot v)$.
La demostración se deja al lector como ejercicio.

8. **Distributiva**: $(\forall \alpha \in \mathbb{K})(\forall \beta \in \mathbb{K})(\forall w \in W)(\alpha + \beta) \odot w = \alpha \odot u \oplus \beta \odot w)$.
La demostración se deja al lector como ejercicio.

♦

Teorema 4.4 (Subespacio vectorial) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $W \subset V$. Entonces, W es un subespacio vectorial de V si y solamente si:

1. $W \neq \emptyset$;

2. $(\forall u \in W)(\forall v \in W)(u + v \in W)$ (la restricción de $+$ a W es cerrada); y

3. $(\forall \alpha \in \mathbb{K})(\forall w \in W)(\alpha w \in W)$ (la restricción de $\cdot$ a W es cerrada).

Demostración. Es claro que si W es un subespacio vectorial de V, entonces las tres propiedades enunciadas son verdaderas, ya W es un espacio vectorial en sí mismo con las operaciones de V restringidas a W.

Recíprocamente, probemos que si las tres condiciones son verdaderas, entonces $(W, \mathbb{K}, +, \cdot)$ es un espacio vectorial. Para ello, debemos verificar las ocho condiciones de la definición de espacio vectorial. Estas verificaciones siguen el mismo patrón que el realizado en el ejemplo anterior, así que no las vamos a realizar de manera detallada. Sin embargo, ver más de cerca la existencia del neutro y del inverso aditivo en W, ilustra el por qué de la primera condición.

En efecto, dado que $W \neq \emptyset$, sea w un elemento de W. Por la segunda condición, si tomamos $\alpha = 0$, entonces, como $w \in V$ (ya que $W \subset V$):

$$\alpha w = 0 \cdot w = 0_v \in W.$$

En otras palabras, en W debe estar, al menos, el vector nulo de V. Y este elemento es también es vector nulo de W.

Ahora, si aplicamos nuevamente la segunda condición al elemento w de W, pero esta vez con $\alpha = -1$, podemos concluir (con ayuda de la cuarta propiedad enunciada en el teorema 4.3) que:

$$\alpha w = -1 \cdot w = -w \in W.$$

Es decir, en W también está, a más de w, su inverso aditivo, $-w$. Por lo tanto, para cada elemento de W, su inverso aditivo es el mismo inverso aditivo de este elemento considerado un elemento de V. $\square$

Observaciones

1. La segunda y la tercera condiciones del teorema 4.4 pueden ser reemplazadas por la siguiente:
$$(\forall \alpha \in \mathbb{K})(\forall u \in W)(\forall v \in W)(\alpha u + v \in W).$$

2. En la demostración del teorema 4.4, vimos que 0_v siempre debe estar en W. En otras palabras, una condición necesaria para que un subconjunto de un espacio vectorial sea un subespacio vectorial es que contenga al vector cero del espacio vectorial.

   Por esta razón, para demostrar el cumplimiento de la primera propiedad, probaremos que el vector nulo está en el conjunto, ya que si no estuviera, el conjunto no sería un subespacio vectorial.

3. De la demostración del teorema también obtenemos otra condición necesaria para que un subconjunto de un espacio vectorial sea un subespacio vectorial de éste: el inverso aditivo de un elemento del subconjunto debe estar en el conjunto.

Ejemplo 4.36

¿Son los siguientes subconjuntos de $\mathbb{R}^3$ subespacios vectorial de $\mathbb{R}^3$?

1. $W_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = x_3\}$.
2. $W_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 2x_1 > x_3\}$.
3. $W_3 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : |x_1 + x_2| = x_3\}$.

Solución.

1. Vamos a utilizar el teorema 4.4 para determinar si W_1 es un subespacio vectorial de $\mathbb{R}^3$. Para ello, verifiquemos si se cumplen las tres condiciones exigidas en el teorema (la segunda y la tercera expresada por una sola, como se indica en la segunda observación al teorema).

   (a) Veamos si $W_1 \neq \emptyset$. Para ello, veamos si $0_v = (0,0,0) \in W_1$.
   Puesto que $0+0=0$, entonces 0_v sí está en W_1.

   (b) Ahora veamos si se verifica que $\alpha u + v \in W_1$ para todo $\alpha \in \mathbb{R}$ y todo u y v en W_1.
   Para ello, supongamos que $u = (x_1, x_2, x_3)$ y $v = (y_1, y_2, y_3)$. Entonces:

   $$\begin{aligned} \alpha u + v &= \alpha(x_1, x_2, x_3) + (y_1, y_2, y_3) \\ &= (\alpha x_1, \alpha x_2, \alpha x_3) + (y_1, y_2, y_3). \end{aligned}$$

   Por lo tanto:
   $$\alpha u + v = (\alpha x_1 + y_1, \alpha x_2 + y_2, \alpha x_3 + y_3).$$

   Ahora veamos si $\alpha u + v$ está en W_1. Para verlo, sumemos sus dos primeras coordenadas:

   $$\begin{array}{rcll} (\alpha x_1 + y_1) + (\alpha x_2 + y_2) &=& \alpha(x_1 + x_2) + (y_1 + y_2) & \\ &=& \alpha x_3 + y_3 & u \in W_1 \text{ y } v \in W_1. \end{array}$$

   Es decir, la suma de las dos primeras es igual a la tercera. Por lo tanto, $\alpha u + v \in W_1$.

   Las condiciones del teorema 4.4 son verdaderas. Por lo tanto, el conjunto W_1 sí es un subespacio vectorial de $\mathbb{R}^3$.

2. Empecemos por verificar si W_2 contiene o no al vector nulo de $\mathbb{R}^3$. Para ello, debemos determinar si el doble de la primera coordenada de $(0,0,0)$ es mayor que la tercera: $2 \cdot 0 > 0$. Pero $2 \cdot 0 = 0$ no es mayor que 0, sino igual (y no puede ser mayor por el axioma de tricotomía de $\mathbb{R}$. Por lo tanto $0_v \notin W_2$, de donde podemos concluir que W_2 no es un subespacio vectorial de $\mathbb{R}^3$.

3. El vector nulo 0_v sí es elemento de W_3, ya que $|0+0| = |0| = 0$. Por lo tanto, $W_3 \neq \emptyset$.

   Sean $\alpha \in \mathbb{R}^3$, $u \in W_3$ y $v \in W_3$. Veamos si $\alpha u + v \in W_3$.

   Observemos que el vector $(1, -2, 1)$ está en W_3, ya que $|1+(-2)| = |-1| = 1$. Sin embargo, su inverso aditivo $-(1,-2,1) = (-1, 2, -1)$ no está en W_3, porque $|-1+2| = |1| = 1 \neq -1$.

   Esto significa que W_3 no puede ser un subespacio vectorial de $\mathbb{R}^3$.

♦

Teorema 4.5 (Intersección de subespacios vectoriales) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y W_1 y W_2 dos subespacios vectoriales de V. Entonces $W_1 \cap W_2$ es un subespacio vectorial de V.

Demostración. Vamos a utilizar el teorema 4.4.

1. Probemos que $W_1 \cap W_2 \neq \emptyset$.
   Para ello, veamos que $0_v \in W_1 \cap W_2$, ya que 0_v está en W_1 y W_2, porque son subespacios vectoriales de V.

2. Sean $\alpha \in \mathbb{K}$, $u \in W_1 \cap W_2$ y $v \in W_1 \cap W_2$. Debemos probar que

$$\alpha u + v \in W_1 \cap W_2.$$

Puesto que W_1 y W_2 son subespacios vectoriales de V, por el teorema 4.4, tenemos que:

$$\alpha u + v \in W_1 \quad \text{y} \quad \alpha u + v \in W_2.$$

Por lo tanto, $\alpha u + v$ debe estar en la intersección de W_1 y W_2.

□

Corolario 4.6 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y W_1, W_2, ..., W_n subespacios vectoriales de V. Entonces

$$\bigcap_{i=1}^{n} W_i = W_1 \cap W_2 \cap \cdots \cap W_n$$

es un subespacio vectorial de V.

Demostración. Por inducción sobre n. □

Ejemplo 4.37

Los conjuntos $W_1 = \{(x_1, x_2, x_3) : x_2 - x_3 = 0\}$ y $W_2 = \{(x_1, x_2, x_3) : x_1 = x_2 - x_3\}$ son subespacios vectoriales de $\mathbb{R}^3$ con las operaciones usuales. Calcular $W_1 \cap W_2$. ¿El conjunto $W_1 \cup W_2$ es un subespacio vectorial de $\mathbb{R}^3$?

Solución.

1. Tenemos que:

$$W_1 \cap W_2 = \{(x_1, x_2, x_3) : x_2 - x_3 = 0 \wedge x_1 = x_2 - x_3\}.$$

Como se verifican las siguientes equivalencias:

$$\begin{aligned} x_2 - x_3 = 0 \wedge x_1 = x_2 - x_3 &\Longleftrightarrow x_2 - x_3 = 0 \wedge x_1 = 0 \\ &\Longleftrightarrow x_2 = x_3 \wedge x_1 = 0, \end{aligned}$$

podemos escribir que:

$$W_1 \cap W_2 = \{(0, x, x) : x \in \mathbb{R}\}.$$

2. Dado que 0_v están tanto en W_1 como en W_2, también está en $W_1 \cup W_2$. Sin embargo, existen u y v en $W_1 \cup W_2$ tales que $u + v \notin W_1 \cup W_2$.

En efecto, esos dos elementos son $u = (1, 2, 2) \in W_1$ y $v = (1, 2, 1) \in W_2$, como el lector(a) puede constatar por sí mismo(a).

♦

La segunda parte de este ejemplo muestra que, en general, la unión de subespacios vectoriales no es un subespacio vectorial. Sin embargo, se puede demostrar que la unión es un subespacio vectorial solo cuando uno de los dos subespacios está contenido en el otro.

4.5 Combinaciones lineales

Definición 4.3 (Combinación lineal de vectores) Sean $(E, \mathbb{K}, +, \cdot)$ un espacio vectorial y v_1, v_2, …, v_n elementos de V. Una *combinación lineal de los vectores* $v_1, v_2, \ldots, v_n$ es cualquier elemento de V de la forma

$$\sum_{j=1}^{n} \alpha_j v_j = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n,$$

donde $\alpha_1, \alpha_2, \ldots, \alpha_n$ son elementos de $\mathbb{K}$.

Definición 4.4 (Un vector es combinación lineal de vectores) Sean $(E, \mathbb{K}, +, \cdot)$ un espacio vectorial y $v_1, v_2, \ldots, v_n$ elementos de V. Un vector $u \in V$ *es una combinación lineal de los vectores* $v_1, v_2, \ldots, v_n$ si existen n escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que

$$u = \sum_{j=1}^{n} \alpha_j v_j = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Ejemplo 4.38

Sea p un vector del espacio $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$ definido por

$$p(x) = 1 + x - x^2$$

para todo $x \in \mathbb{R}$. ¿Es p una combinación lineal de $q \in \mathscr{P}_2[x]$ y $r \in \mathscr{P}_2[x]$ donde

$$q(x) = 1 + 2x^2 \quad \text{y} \quad r(x) = 2 + x + x^2$$

para todo $x \in \mathbb{R}$?

Solución. Queremos saber si existen $\alpha \in \mathbb{R}$ y $\beta \in \mathbb{R}$ tales que

$$p = \alpha q + \beta r.$$

Para averiguarlo, supongamos que sí existen. Entonces, para todo $x \in \mathbb{R}$, se verifica la siguiente igualdad:

$$p(x) = \alpha q(x) + \beta r(x). \tag{4.5}$$

Es decir, es verdadera la igualdad:

$$1 + x - x^2 = \alpha(1 + 2x^2) + \beta(2 + x + x^2)$$

para todo $x \in \mathbb{R}$.

Ahora bien, si realizamos la suma de los dos polinomio en el lado derecho de la igualdad, obtenemos que:

$$1 + x - x^2 = (\alpha + 2\beta) + \beta x + (2\alpha + \beta)x^2$$

para todo $x \in \mathbb{R}$. Entonces, por la definición de igualdad de polinomios, tenemos que α y β, en el caso de que existieran, satisfarían el siguiente sistema de tres ecuaciones:

$$\begin{aligned} \alpha + 2\beta &= 1 \\ \beta &= 1 \\ 2\alpha + \beta &= -1. \end{aligned} \tag{4.6}$$

Veamos qué valores de α y β satisfacen este sistema.

Para ello, trabajemos con la matriz ampliada de este sistema:

$$\left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 1 & -1 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right)$$

Como el rango de la matriz es 2, entonces el sistema tiene solución única: $\alpha = -1$ y $\beta = 1$. Veamos, entonces, si estos valores satisfacen la igualdad (4.5):

$$\begin{aligned} \alpha p(x) + \beta q(x) &= (-1)(1 + 2x^2) + (1)(2 + x + x^2) \\ &= (-1 + 2) + x + (-2 + 1)x^2 \\ &= 1 + x - x^2 = p(x). \end{aligned}$$

Es decir, existen dos escalares α y β tales que $p = \alpha q + \beta r$. En otras palabras, p sí es una combinación lineal de q y r. ♦

Ejemplo 4.39

¿Para qué valores del número real a el vector $u = (a, 0, 0) \in \mathbb{R}^3$ es una combinación lineal de los vectores $v = (0, 1, 1)$ y $w = (1, -1, 1)$?

Solución. Buscamos el valor de a para que u sea una combinación lineal de v y w. La pregunta es, entonces, ¿qué valor o valores debe tomar a para que existan dos escalares α y β tales que $u = \alpha v + \beta w$?

Supongamos que esos dos escalares existen. Entonces, debe satisfacerse la siguiente igualdad:

$$(a, 0, 0) = \alpha(0, 1, 1) + \beta(1, -1, 1).$$

Como

$$(a, 0, 0) = \alpha(0, 1, 1) + \beta(1, -1, 1) = (\beta, \alpha - \beta, \alpha + \beta),$$

entonces α y β deben satisfacer el siguiente sistema de tres ecuaciones lineales:

$$\begin{aligned} \beta &= a \\ \alpha - \beta &= 0 \\ \alpha + \beta &= 0. \end{aligned}$$

Resolvamos este sistema:

$$\left(\begin{array}{cc|c} 0 & 1 & a \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 0 & 1 & a \\ 0 & -2 & 0 \\ 1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 0 & 1 & a \\ 0 & 0 & 2a \\ 1 & 1 & 0 \end{array}\right).$$

Para que el sistema tenga solución, debemos tener que $2a = 0$; es decir, que $a = 0$.

Por lo tanto, el vector nulo de $\mathbb{R}^3$ es una combinación lineal de $(0, 1, 1)$ y $(1, -1, 1)$ (los escalares respectivos son $\alpha = 1$ y $\beta = 0$). ♦

4.6 Cápsula lineal de un conjunto

Definición 4.5 (Cápsula) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $S \subset V$ tal que $S \neq \emptyset$. La *cápsula de* S es el conjunto formado por todos los elementos de V que se pueden expresar como una combinación lineal de elementos de S. En otras palabras, la cápsula de S es el conjunto de todas las combinaciones lineales de elementos de S.
La cápsula de S es representada por $\langle S \rangle$. Por lo tanto:

$$\langle S \rangle = \{v \in V : v = \sum_j \alpha_{j=1}^n v_j,\ \alpha_j \in \mathbb{K},\ v_j \in V\}.$$

Ejemplo 4.40

Sea $S = \{1 + x, 1 + x^2, 1 + x^3\}$ subconjunto de $\mathscr{P}_3[x]$ con las operaciones usuales de suma de polinomios y producto de un escalar por un polinomio. Calcular la cápsula de S.

Solución. Sea $p \in \langle S \rangle$. Entonces $p \in \mathscr{P}_3[x]$, por lo que:

$$p(x) = a + bx + cx^2 + dx^3$$

para todo $x \in \mathbb{R}$.

Por otro lado, como $p \in \langle S \rangle$, existen α, β, γ y δ en $\mathbb{R}$ tales que

$$p(x) = a + bx + cx^2 + dx^3 = \alpha(1 + x) + \beta(1 + x^2) + \gamma(1 + x^3)$$

para todo $x \in \mathbb{R}$.

Puesto que:

$$a + bx + cx^2 + dx^3 = \alpha(1 + x) + \beta(1 + x^2) + \gamma(1 + x^3) = (\alpha + \beta + \gamma) + \alpha x + \beta x^2 + \gamma x^3,$$

por la definición de igualdad entre polinomios, podemos concluir que:

$$\langle S \rangle = \{p(x) = a + bx + cx^2 + dx^3 \in \mathscr{P}_3[x] : \alpha + \beta + \gamma = a,\ \alpha = b,\ \beta = c,\ \gamma = d\}.$$

Para determinar cómo son los elementos de $\langle S \rangle$, debemos determinar los coeficientes a, b, c y d (esto es lo que significa calcular $\langle S \rangle$). Para ello, resolvamos el sistema de cuatro ecuaciones lineales:

$$\begin{aligned} \alpha + \beta + \gamma &= a \\ \alpha &= b \\ \beta &= c \\ \gamma &= d. \end{aligned}$$

Para ello, trabajemos con la matriz ampliada:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 1 & 0 & 0 & b \\ 0 & 1 & 0 & c \\ 0 & 0 & 1 & d \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & -1 & -1 & -a + b \\ 0 & 1 & 0 & c \\ 0 & 0 & 1 & d \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & -1 & -1 & -a+b \\ 0 & 0 & -1 & -a+b+c \\ 0 & 0 & 1 & d \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & -1 & -1 & -a+b \\ 0 & 0 & -1 & -a+b+c \\ 0 & 0 & 0 & -a+b+c+d \end{array}\right).$$

Por lo tanto, el sistema tiene solución si y solo si se verifica que:

$$-a+b+c+d=0.$$

De esta manera, la cápsula de S está formada por todos los vectores p tal que

$$p(x) = a+bx+cx^2+dx^3 \quad \text{y} \quad -a+b+c+d=0.$$

Entonces, tenemos que:

$$\langle S\rangle = \{a+bx+cx^2+dx^3 \in \mathscr{P}_3[3] : -a+b+c+d=0\}.$$

♦

Ejemplo 4.41

Si $S = \{(1,-1),\ (0,3)\} \in \mathbb{R}^2$, calcular $\langle S\rangle$.

Solución. De la definición de $\langle S\rangle$ tenemos:

$$\begin{aligned} \langle S\rangle &= \{(a,b) \in \mathbb{R}^2 : (a,b) = \alpha(1,-1) + \beta(0,3),\ \alpha,\beta \in \mathbb{R}\} \\ &= \{(a,b) \in \mathbb{R}^2 : (a,b) = (\alpha, -\alpha+3\beta),\ \alpha,\beta \in \mathbb{R}\} \\ &= \{(a,b) \in \mathbb{R}^2 : \alpha = a,\ -\alpha+3\beta = b,\ \alpha,\beta \in \mathbb{R}\}. \end{aligned}$$

Por lo tanto, para caracterizar los elementos de $\langle S\rangle$, debemos determinar las condiciones bajos las cuales el sistema

$$\begin{aligned} \alpha \quad\quad &= a \\ -\alpha + 3\beta &= b \end{aligned} \tag{4.7}$$

tenga solución.

Para ello, observemos que:

$$\left(\begin{array}{cc|c} 1 & 0 & a \\ -1 & 3 & b \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & a \\ 0 & 3 & a+b \end{array}\right).$$

De manera que el sistema (4.7) tiene solución para cualquier par de valores para a y b, ya que el determinante de la matriz de coeficientes es diferente de 0 independientemente de los valores de a y de b. En otras palabras:

$$(a,b) \in \langle S\rangle \iff a \in \mathbb{R} \text{ y } b \in \mathbb{R}.$$

Por lo tanto:

$$\langle S\rangle = \{(a,b) \in \mathbb{R}^2 : a \in \mathbb{R},\ b \in \mathbb{R}\} = \mathbb{R}^2.$$

♦

Teorema 4.7 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $S \subset V$ tal que $S \neq \emptyset$. Entonces la cápsula de S es un subespacio vectorial de V.

Demostración. Para probar que $\langle S \rangle$ es un subespacio vectorial de V, vamos a utilizar el teorema 4.4 (página 99). Es decir, debemos probar que:

1. $\langle S \rangle \neq \emptyset$.
2. $(\forall \lambda \in \mathbb{K})(\forall u \in \langle S \rangle)(\forall v \in \langle S \rangle)(\lambda u + v \in \langle S \rangle)$.

Procedamos así.

1. Para probar que $\langle S \rangle \neq \emptyset$, demostremos que $0_v \in \langle S \rangle$. Para ello, debemos probar que 0_v es una combinación de elementos de S.

   Y es así, pues, como $S \neq \emptyset$, existe $v \in S$. Por otro lado, sabemos que

   $$0_v = 0 \cdot v.$$

   Entonces 0_v es una combinación lineal de un elemento de S, por lo que $0_v \in \langle S \rangle$.

2. Sean $\lambda \in \mathbb{K}$, $u \in \langle S \rangle$ y $v \in \langle S \rangle$. Existen, entonces, escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, $\beta_1, \beta_2, \beta_m$ y elementos de S $u_1, u_2, \ldots, u_m$, $v_1, v_2, \ldots, v_n$ tales que

   $$u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m \quad \text{y} \quad v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n.$$

   Por lo tanto:

   $$\begin{aligned} \lambda u + v &= \lambda\left(\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m\right) + \left(\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n\right) \\ &= (\lambda\alpha_1)u_1 + (\lambda\alpha_2)u_2 + \cdots + (\lambda\alpha_m)u_m + \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n. \end{aligned}$$

   Es decir, $\lambda u + v$ es una combinación lineal de elementos de S, por lo que $\lambda u + v \in \langle S \rangle$.

Hemos demostrado, entonces, que $\langle S \rangle$ es un subespacio vectorial de V. □

4.7 Subespacio vectorial generado por un subconjunto

Definición 4.6 (Subespacio generado) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $S \subset V$. El *subespacio vectorial generado por* S es el subespacio vectorial que se obtiene al intersecar todos los subespacios vectoriales de V que contienen a S.

Dado que la intersección de dos conjuntos siempre está contenida en cualquiera de los dos conjuntos, el subespacio generado por un conjunto será el subespacio vectorial "más pequeño" que contenga al conjunto.

De manera más precisa, el subespacio generado por S es el conjunto

$$\text{s.e.v.}(S) = \bigcap \{W : W \,\text{s.e.v.}\, V,\ S \subset W\}.$$

Sabemos que la intersección de un número finito de subespacios vectoriales es un subespacio vectorial. No es difícil demostrar que esta afirmación es cierta para cualquier número de subespacios. Por lo tanto, el conjunto s.e.v.(S) es un subespacio vectorial de V. Además, por la definición de intersección, tenemos que es única y que

$$\text{s.e.v.}(S) \subset W$$

para todo W subespacio vectorial de V.

En otras palabras, si W es un subespacio vectorial de V tal que $S \subset W$, debe suceder que s.e.v.$(S) \subset W$. Esto significa que el subespacio vectorial generado por S es el subespacio vectorial "más pequeño" que contiene a S.

Ejemplo 4.42

¿Qué subespacio genera el conjunto vacío?

Solución. Sea $S = \emptyset$. Como el conjunto $N = \{0_v\}$ es un subespacio vectorial de V (es no vacío y las dos operaciones heredadas de V son cerradas), tenemos que es el subespacio vectorial de V "más pequeño" que contiene a S, ya que $\emptyset \subset N$. Por lo tanto, el espacio generado por el conjunto vacío es el subespacio vectorial con un solo elemento, el vector nulo de V. A este espacio se lo denomina *subespacio nulo* de V. ♦

Teorema 4.8 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $S \subset V$ tal que $S \neq \emptyset$. Entonces el subespacio generado por S es igual a $\langle S \rangle$.

Demostración. Vamos a probar que $\langle S \rangle$ es el subespacio vectorial más pequeño que contiene a S. Para ello, demostraremos que:

1. $S \subset \langle S \rangle$; y
2. si W es un subespacio vectorial de V tal que $S \subset W$, entonces $\langle S \rangle \subset W$.

Procedamos.

1. Sea $v \in S$. Debemos probar que $v \in \langle S \rangle$. Pero esto se sigue inmediatamente del hecho de que $v = 1 \cdot v$, por lo que v es una combinación lineal de un elemento de S; es decir, v está en la cápsula de S.

2. Sea W un subespacio vectorial de V que contiene a S. Vamos a probar que la cápsula de S está contenida en W. Para ello, sea $v \in \langle S \rangle$. Probemos que $v \in W$.

   Por ser v un elemento de la cápsula de S, existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ y vectores de S $v_1, v_2, \ldots, v_n$ tales que

   $$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n. \tag{4.8}$$

   Por otro lado, cada v_i también es un elemento de W ya que $S \subset W$. Por lo tanto, lo que expresa la igualdad (4.8) es lo siguiente: el vector v es una combinación lineal de elementos de W. Pero W es un subespacio vectorial, por lo que las operaciones $+$ y $\cdot$ son cerradas. Entonces, cualquier combinación lineal de elementos de W está en W; es decir, $v \in W$.

   Hemos demostrado, entonces, que todo elemento de $\langle S \rangle$ es un elemento de W, lo que prueba que $\langle S \rangle \subset W$.

En resumen, hemos probado que el menor subespacio vectorial que contiene a S es su cápsula. Esto significa, entonces, que el subespacio vectorial generado por S es, precisamente, su cápsula. ☐

Ejemplo 4.43

Si $p(x) = x^2 + 1$, el conjunto $S = \{p(x), p'(x), p''(x)\}$ genera el espacio vectorial $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$.

Solución. Por el teorema anterior, lo que debemos mostrar es que $\langle S \rangle = \mathscr{P}_2[x]$. Para ello, calculemos $\langle S \rangle$:

$$\begin{aligned}
\langle S \rangle &= \{q(x) = a + bx + cx^2 \in \mathscr{P}_:[x]q(x) = \alpha p(x) + \beta p'(x) + \gamma p''(x)\} \\
&= \{a + bx + cx^2 \in \mathscr{P}_:[x]a + bx + cx^2 = \alpha(x^2 + 1) + \beta(2x) + \gamma(2)\} \\
&= \{a + bx + cx^2 \in \mathscr{P}_:[x]a + bx + cx^2 = (\alpha + 2\gamma) + 2\beta x + \alpha x^2\} \\
&= \{a + bx + cx^2 \in \mathscr{P}_:[x]\alpha + 2\gamma = a,\ 2\beta = b, \alpha = c\}.
\end{aligned}$$

Entonces, para determinar los elementos a, b y c, debemos averiguar bajo qué condiciones necesarias y suficientes el sistema

$$\begin{aligned}
\alpha \quad\quad + 2\gamma &= a \\
2\beta \quad\quad &= b \\
\alpha \quad\quad\quad\quad &= c
\end{aligned} \tag{4.9}$$

tiene solución.

Para ello, miremos su matriz ampliada:

$$\left(\begin{array}{ccc|c} 1 & 0 & 2 & a \\ 0 & 2 & 0 & b \\ 1 & 0 & 0 & c \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 2 & a \\ 0 & 2 & 0 & b \\ 0 & 0 & -2 & -a + c \end{array}\right).$$

Como el determinante de la matriz de coeficientes es diferente de 0, independientemente de a, b y c, entonces el sistema (4.9) tiene solución para todo a, b y c. Por lo tanto $\langle S \rangle = \mathscr{P}_2[x]$. En otras palabras, el conjunto

$$\{p, p', p''\}$$

genera todo $\mathscr{P}_2[x]$. ♦

4.8 Bases de espacios vectoriales

Definición 4.7 (Dependencia e independencia lineal) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $S = \{v_1, v_2, \ldots, v_n\}$ un subconjunto finito de V.

1. El conjunto S es *linealmente independiente*(li) si y solamente si la combinación lineal

   $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_v,$$

   denominada *combinación lineal nula*, tiene como solución única la solución trivial: $\alpha_j = 0$ para todo $j \in \{1, 2, \ldots, n\}$. En otras palabras, el conjunto S es li si:

   $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_v \Longrightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

2. El conjunto S es *linealmente dependiente* (ld) si y solo si no es linealmente independiente. Es decir, si la combinación lineal nula tiene un número infinito de soluciones. En otras palabras, el conjunto S es ld si:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_v \ \wedge \ \exists j \in \{1, 2, \ldots, n\} \text{ tal que } \alpha_j \neq 0.$$

De la definición es claro que un conjunto no puede ser linealmente independiente y dependiente al mismo tiempo. Esto lo podemos ver de otra manera: la combinación lineal nula produce un sistema de ecuaciones lineales homogéneo que, como sabemos, o tiene una sola solución —la trivial—, o tiene un número infinito de soluciones. En el primer caso, el conjunto es li; en el segundo, es ld.

Ejemplo 4.44

El conjunto $S = \{p, p', p''\}$ es li en $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$, donde

$$p(x) = 1 + x^2$$

para todo $x \in \mathbb{R}$.

Solución. Supongamos que

$$\alpha_1 p + \alpha_2 p' + \alpha_3 p'' = 0_v. \tag{4.10}$$

Vamos a probar que $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Para todo $x \in \mathbb{R}$, la igualdad (4.10) implica la siguiente:

$$\alpha_1 p(x) + \alpha_2 p'(x) + \alpha_3 p''(x) = 0;$$

es decir, es verdad que:

$$\alpha_1(1 + x^2) + \alpha_2(2x) + \alpha_3(2) = (\alpha_1 + 2\alpha_3) + 2\alpha_2 x + \alpha_1 x^2 = 0$$

para todo $x \in \mathbb{R}$. Por lo tanto, por la definición de igualdad entre polinomios, tenemos que:

$$\alpha_1 + 2\alpha_3 = 0,\ 2\alpha_2 = 0,\ \alpha_1 = 0,$$

de donde obtenemos que $\alpha_1 = \alpha_2 = \alpha_3 = 0$. En otras palabras, el conjunto S es linealmente independiente. ♦

Ejemplo 4.45

Sea $S = \{(1, a, 1),\ (-1, 0, 1),\ (1 + a, 1, 0)\} \subset \mathbb{R}^3$. ¿Para qué valores de a el conjunto S es li? ¿Para qué valores de a es ld?

Solución. Sean α, β y γ escalares tales que

$$\alpha(1, a, 1) + \beta(-1, 0, 1) + \gamma(1 + a, 1, 0) = (0, 0, 0). \tag{4.11}$$

Entonces, se verifica que:

$$(\alpha - \beta + (1 + a)\gamma, a\alpha + \gamma, \alpha + \beta) = (0, 0, 0).$$

Por lo tanto, los números α, β y γ que satisfacen la igualdad (4.11) satisfacen el siguiente sistema de tres ecuaciones lineales:

$$\begin{aligned} \alpha - \beta + (1+a)\gamma &= 0 \\ a\alpha \quad + \qquad\quad \gamma &= 0 \\ \alpha + \beta \qquad\qquad &= 0. \end{aligned} \tag{4.12}$$

Ahora debemos determinar para qué valores de a, este sistema tiene como única solución la trivial y para qué valores tiene un número infinito de soluciones.

Para ello, calculemos el determinante de la matriz de coeficientes del sistema:

$$\begin{vmatrix} 1 & -1 & 1+a \\ a & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1+a \\ a & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = -\begin{vmatrix} 2 & 1+a \\ a & 1 \end{vmatrix} = a^2 + a - 2 = (a+2)(a-1).$$

Por lo tanto, el determinante es igual a 0 si y solo si $a = -2$ o $a = 1$.

Esto significa que el sistema (4.12) tiene un número infinito de soluciones si y solo si $a = -2$ o $a = 1$. Por lo tanto, conjunto S es linealmente dependiente cuando $a = -2$ o $a = 1$; y es linealmente independiente si $a \neq -2$ y $a \neq 1$.

♦

Observaciones

1. Todo conjunto finito que contiene el vector nulo es linealmente dependiente, pues $0_v = 1 \cdot 0_v$; es decir, hay al menos un escalar diferente de 0 en la combinación nula.
2. Si un conjunto contiene un subconjunto linealmente dependiente, entonces el conjunto también es linealmente dependiente.
3. Todo subconjunto de un conjunto linealmente independiente también es linealmente independiente.

Definición 4.8 (Base) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $B \subset V$. El conjunto B es una *base* del espacio vectorial V si y solamente si B es linealmente independiente y B *genera* V; es decir, el subespacio vectorial generado por B es el espacio vectorial V.
En otras palabras, B es una base de B si y solo si:

1. B es linealmente independiente; y
2. $\langle B \rangle = V$.

Ejemplo 4.46

El conjunto $B = \{p, p', p''\}$, donde $p(x) = 1 + x^2$ para todo $x \in \mathbb{R}$, es una base del espacio vectorial $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$.

Solución. En un ejemplo anterior, probamos que el conjunto B es linealmente independiente; también se demostró que $\langle B \rangle = \mathscr{P}_2[x]$. Por lo tanto B es una base de $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$. ♦

Ejemplo 4.47

El conjunto
$$B = \{1, x-1, (x-1)^2, (x-1)^3\}$$
es una base del espacio vectorial $(\mathscr{P}_3[x], \mathbb{R}, +, \cdot)$.

Solución. Debemos probar que la cápsula de B es $\mathscr{P}_3[x]$ y que B es un conjunto linealmente independiente.

1. Calculemos la cápsula de B:

$$\begin{aligned}\langle B\rangle &= \{p \in \mathscr{P}_3[x] : p(x) = \alpha(1) + \beta(x-1) + \gamma(x-1)^2 + \delta(x-1)^3\}\\ &= \{p \in \mathscr{P}_3[x] : p(x) = \alpha(1) + \beta(x-1) + \gamma(x^2-2x+1) + \delta(x^3-3x^2+3x-1)\}\\ &= \{p \in \mathscr{P}_3[x] : p(x) = (\alpha-\beta+\gamma-\delta) + (\beta-2\gamma+3\delta)x + (\gamma-3\delta)x^2 + \delta x^3\}\end{aligned}$$

Si $p(x) = a + bx + cx^2 + dx^3$, entonces que $p \in \langle B\rangle$ es equivalente a

$$a + bx + cx^2 + dx^3 = (\alpha-\beta+\gamma-\delta) + (\beta-2\gamma+3\delta)x + (\gamma-3\delta)x^2 + \delta x^3,$$

lo que es equivalente al siguiente sistema de cuatro ecuaciones lineales en α, β, γ y δ:

$$\begin{aligned}\alpha - \beta + \gamma - \delta &= a\\ \beta - 2\gamma + 3\delta &= b\\ \gamma - 3\delta &= c\\ \delta &= d.\end{aligned} \tag{4.13}$$

Por lo tanto:

$$\langle B\rangle = \{a + bx + cx^2 + dx^3 \in \mathscr{P}_3[x] : a, b, c \text{ y } d \text{ satisfacen el sistema (4.13)}\}.$$

Ahora bien, la matriz de los coeficientes del sistema (4.13) es la siguiente:

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

Como se puede ver, su determinante es diferente de 0 para cualquier a, b, c y d. Por lo tanto, el sistema (4.13) tiene solución para todo a, b, c y d. Esto significa que

$$\langle B\rangle = \mathscr{P}_3[x].$$

En resumen, B genera $\mathscr{P}_3[x]$.

2. Ahora probemos que el conjunto B es linealmente independiente. Para ello, sean α, β, γ, escalares tales que:

$$\alpha(1) + \beta(x-1) + \gamma(x-1)^2 + \delta(x-1)^3 = 0 + 0x + 0x^2 + 0x^3. \tag{4.14}$$

Esta igualdad es equivalente a la siguiente:

$$(\alpha-\beta+\gamma-\delta) + (\beta-2\gamma+3\delta)x + (\gamma-3\delta)x^2 + \delta x^3 = 0 + 0x + 0x^2 + 0x^3.$$

Por lo tanto, α, β, γ y δ satisfacen el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} \alpha - \beta + \ \ \gamma - \ \ \delta &= 0 \\ \beta - 2\gamma + 3\delta &= 0 \\ \gamma - 3\delta &= 0 \\ \delta &= 0 \end{aligned}$$

cuya matriz de coeficientes es la misma que la del sistema (4.13), de la cual sabes que su determinante es diferente de 0, por lo que, al tratarse este segundo sistema de uno homogéneo, la única solución que tiene es la trivial. Por lo tanto:

$$\alpha = \beta = \gamma = \delta = 0.$$

En resumen, la igualdad (4.14) implica que los cuatro escalares son iguales a 0, por lo que el conjunto B es linealmente independiente.

Hemos probado, entonces, que el conjunto B, al generar $\mathscr{P}_3[x]$ y ser linealmente independiente, es una base del espacio vectorial $(\mathscr{P}_3[x], \mathbb{R}, +, \cdot)$. ♦

Ejemplo 4.48

El conjunto

$$W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : \begin{vmatrix} a & b \\ 1 & 2 \end{vmatrix} = 0 \right\}$$

es un subespacio vectorial de $(\mathbb{M}_{2\times 2}, \mathbb{R}, +, \cdot)$. Encontrar una base para W.

Solución. Para ello, procedamos de la siguiente manera:

$$\begin{aligned} W &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : \begin{vmatrix} a & b \\ 1 & 2 \end{vmatrix} = 0 \right\} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : 2a - b = 0 \right\} \\ &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : 2a = b \right\} = \left\{ \begin{pmatrix} a & 2a \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : a, c, d \in \mathbb{R} \right\}. \end{aligned}$$

Ahora bien, como:

$$\begin{pmatrix} a & 2a \\ c & d \end{pmatrix} = \begin{pmatrix} a & 2a \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} = a \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

tenemos que:

$$\begin{aligned} W &= \left\{ a \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} : a, c, d \in \mathbb{R} \right\} \\ &= \left\langle \left\{ \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \right\rangle. \end{aligned}$$

Por lo tanto, el conjunto

$$B = \left\{ \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

genera W, ya que $\langle B \rangle = W$.

Ahora probemos que B es linealmente independiente. Para ello, supongamos que α, β y γ escalares tales que:

$$\alpha\begin{pmatrix}1 & 2\\ 0 & 0\end{pmatrix}+\beta\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}+\gamma\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}. \tag{4.15}$$

Esta igualdad implica que

$$\begin{pmatrix}\alpha & 2\alpha\\ \beta & \gamma\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},$$

la que, a su vez, implica, por la definición de igualdad entre matrices, que se verifiquen las siguientes igualdades:

$$\alpha = 0,\ 2\alpha = 0, \beta = 0, \gamma = 0.$$

En resumen, la igualdad (4.15) implica que $\alpha = \beta = \gamma = 0$. Esto significa que B es un conjunto linealmente independiente.

Hemos probado, entonces, que B genera $\mathbb{M}_{2\times 2}$ y es linealmente independiente; entonces, B es una base del espacio $(\mathbb{M}_{2\times 2}, \mathbb{R}, +, \cdot)$. ♦

¿Cuál es el papel de una base? Para contestar esta pregunta, supongamos que

$$B = \{v_i : i \in I_n = \{1, 2, \dots, n\}\}$$

es una base de un espacio vectorial V. Por un lado, como B genera V, cada elemento de V se expresa como una combinación lineal de elementos de B. Pero, por ser B linealmente independiente, esa combinación lineal es única.

En efecto. Sea $v \in V$. Existen entonces n escalares α_i con $i \in I_n$ tales que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Supongamos, además, que β_i con $i \in I_n$ son escalares tales que

$$v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n.$$

Por lo tanto:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n,$$

de donde obtenemos que:

$$(\alpha_1 \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n = 0_v.$$

Pero, como B es un conjunto linealmente independiente, entonces esta última igualdad implica que:

$$(\alpha_1 - \beta_1) = (\alpha_2 - \beta_2) = \cdots = (\alpha_n - \beta_n) = 0;$$

es decir, tenemos que

$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2,\ \dots,\ \alpha_n = \beta_n.$$

En otras palabras, v se expresa de manera única como combinación lineal de los elementos de la base.

A los escalares α_i con $i \in I_n$ se les conoce con el nombre de *coordenadas* del vector v respecto de la base B.

Las coordenadas de un vector de una base son utilizadas para representar e identificar a un vector respecto de una base. Muchas veces, en lugar de hablar de v respecto de la base, hablaríamos de la n-upla $(\alpha_1, \alpha_2, \dots, \alpha_n)$.

Teorema 4.9 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $S = \{v_1, v_2, \ldots, v_m\} \subset V$. Si S genera el espacio V, entonces cualquier conjunto de V linealmente independiente tiene m elementos a lo más.

Demostración. Es suficiente probar que cualquier conjunto que tenga más de m elementos es linealmente dependiente.

En efecto, una vez probado esto, supongamos que W fuera linealmente independiente. Si tuviera más de m elementos, sería linealmente dependiente; es decir, no sería linealmente independiente, lo que entraría en contradicción con la hipótesis de que sí lo es.

Sea, entonces, un conjunto $W = \{w_1, w_2, \ldots, w_n\} \subset V$ tal que $n > m$. Vamos a demostrar que este conjunto es linealmente dependiente.

Para ello, sean λ_j, con $j \in \{1, 2, \ldots, n\}$, escalares tales que:

$$\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n = 0_v. \tag{4.16}$$

Vamos a demostrar esta ecuación tiene un número infinito de soluciones.

Como S genera el espacio V, cualquier elemento de V es una combinación lineal de elementos de S. En particular, cada elemento de W es una combinación lineal de elementos de S. Por lo tanto, cada elemento de W se puede expresar de la siguiente manera:

$$\begin{aligned} w_1 &= \alpha_{11} v_1 + \alpha_{21} v_2 + \cdots + \alpha_{m1} v_m \\ w_2 &= \alpha_{12} v_1 + \alpha_{22} v_2 + \cdots + \alpha_{m2} v_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ w_n &= \alpha_{1n} v_1 + \alpha_{2n} v_2 + \cdots + \alpha_{mn} v_m. \end{aligned}$$

Si ahora sustituimos los w_j por sus combinaciones lineales de elementos de S, la igualdad (4.16) implica la siguiente:

$$\lambda_1(\alpha_{11} v_1 + \cdots + \alpha_{m1} v_m) + \lambda_2(\alpha_{12} v_1 + \cdots + \alpha_{m2} v_m) + \cdots + \lambda_n(\alpha_{1n} v_1 + \cdots + \alpha_{mn} v_m) = 0_v.$$

Si ahora agrupamos respecto de cada v_j, obtenemos la siguiente igualdad:

$$(\alpha_{11}\lambda_1 + \cdots + \alpha_{1n}\lambda_n) v_1 + (\alpha_{21}\lambda_1 + \cdots + \alpha_{2n}\lambda_n) v_2 + \cdots + (\alpha_{m1}\lambda_1 + \cdots + \alpha_{mn}\lambda_n) v_m = 0_v.$$

Esta ecuación tiene, al menos, la solución trivial; es decir, todos los escalares de cada v_j iguales a 0 son una solución de esta ecuación. Por lo tanto, se obtiene:

$$\begin{array}{ccccccc} \alpha_{11}\lambda_1 + & \alpha_{12}\lambda_2 & + \cdots + & \alpha_{1n}\lambda_n & = & 0 \\ \alpha_{21}\lambda_1 + & \alpha_{22}\lambda_2 & + \cdots + & \alpha_{2n}\lambda_n & = & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{m1}\lambda_1 + & \alpha_{m2}\lambda_2 & + \cdots + & \alpha_{mn}\lambda_n & = & 0, \end{array}$$

un sistema homogéneo de m ecuaciones y n incógnitas (cada λ_i).

Puesto que $n > m$ (más incógnitas que ecuaciones), podemos concluir que hay un número infinito de soluciones $\{\lambda_j\}$. Esto significa que el conjunto W es linealmente dependiente. $\square$

Ejemplo 4.49

El conjunto

$$B = \left\{ \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

es una base del subespacio vectorial

$$W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : \begin{vmatrix} a & b \\ 1 & 2 \end{vmatrix} = 0 \right\}.$$

Demuestre que el conjunto

$$S = \left\{ \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 4 & -3 \end{pmatrix}, \begin{pmatrix} -2 & -4 \\ 1 & 0 \end{pmatrix} \right\}$$

es linealmente dependiente.

Solución. Como B es una base de $(\mathbb{M}_{2\times 2}, \mathbb{R}, +, \cdot)$, entonces es un conjunto linealmente independiente que tiene 3 elementos. Entonces, por el teorema 4.9, cualquier conjunto de $\mathbb{M}_{2\times 2}$ con cuatro o más elementos debe ser linealmente dependiente. Ese es el caso de S, que tiene cuatro elementos. ♦

La principal consecuencia del último teorema es que todas las bases tienen el mismo número de elementos. En efecto, supongamos que B y B' son dos bases de un espacio vectorial V, y m y n son el número de elementos de B y B' respectivamente.

Ahora bien, B genera el espacio V y tiene m elementos. Entonces, como B' es linealmente independiente, por el teorema 4.9, se debe cumplir que $n \leq m$. Si ahora intercambiamos B y B' en el razonamiento anterior, tenemos que $m \leq n$. Es decir, $m = n$.

Hemos demostrado así el siguiente teorema.

Teorema 4.10 Todas las bases de un espacio vectorial tienen el mismo número de elementos.

4.9 Dimensión de un espacio vectorial

Definición 4.9 (Dimensión) Sea $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial. El número de elementos de cualquier base de B se denomina *dimensión* del espacio vectorial V. En el caso de que V sea el espacio nulo, se asigna 0 a la dimensión de V. Se acostumbra a representar la dimensión de V con $\dim(V)$.

Ejemplo 4.50 Dado que el conjunto

$$B = \left\{ \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

es una base del subespacio vectorial

$$W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : \begin{vmatrix} a & b \\ 1 & 2 \end{vmatrix} = 0 \right\},$$

podemos concluir que $\dim(W) = 3$. ♦

Cuando el número de elementos de una base es finito, se dice que el espacio vectorial es de dimensión finita.

A continuación, exhibiremos las bases de los principales espacios vectoriales que estamos estudiando en este libro. Estas bases son denominadas *canónicas* en el sentido de que *se ajustan a un canon o modelo ideal*[1]

Ejemplo 4.51

1. En el espacio vectorial $(\mathbb{R}, \mathbb{K}, +, \cdot)$, el conjunto $C = \{1\}$ es linealmente independiente, pues la igualdad $\alpha \cdot 1 = 0$ implica necesariamente que $\alpha = 0$, ya que $1 \neq 0$. Adicionalmente, la cápsula de C es $\mathbb{R}$, pues

$$\langle C \rangle = \{\alpha \cdot 1 \in \mathbb{R} : \alpha \in \mathbb{R}\} = \{\alpha \in \mathbb{R} : \alpha \in \mathbb{R}\} = \mathbb{R}.$$

   Por lo tanto, C es una base para $(\mathbb{R}, \mathbb{K}, +, \cdot)$, y es la canónica. Además, $\dim(\mathbb{R}) = 1$.

2. En el espacio $(\mathbb{R}^2, \mathbb{K}, +, \cdot)$, el conjunto

$$C = \{(1,0),\ (0,1)\}$$

   es la base canónica y $\dim(\mathbb{R}^2) = 2$.

   En efecto:

   (a) C es linealmente independiente, porque la igualdad

$$\alpha(1,0) + \beta(0,1) = (0,0)$$

   implica que:

$$(\alpha, \beta) = (0,0),$$

   de donde: $\alpha = \beta = 0$.

   (b) C genera $\mathbb{R}^2$, pues:

$$\langle C \rangle = \{\alpha(1,0) + \beta(0,1) : \alpha \in \mathbb{R},\ \beta \in \mathbb{R}\} = \{(\alpha, \beta) : \alpha \in \mathbb{R},\ \beta \in \mathbb{R}\} = \mathbb{R}^2.$$

3. En el espacio $(\mathbb{R}^3, \mathbb{K}, +, \cdot)$, la base canónica es:

$$C = \{(1,0,0),\ (0,1,0),\ (0,0,1)\}.$$

   Por lo tanto, $\dim(\mathbb{R}^3) = 3$.

4. En el espacio $(\mathbb{R}^n, \mathbb{K}, +, \cdot)$, la base canónica es:

$$C = \{(1,0,0,\ldots,0),\ (0,1,0,\ldots,0), \ldots, (0,0,\ldots,0,1)\}.$$

   Entonces, $\dim(\mathbb{R}^n) = n$.

[1]Diccionario panhispánico de dudas de la Real Academia Española, Edición 2005.

5. En el espacio $(\mathbb{C}, \mathbb{K}, +, \cdot)$, la base canónica es:

$$C = \{1, \imath\}.$$

Por lo tanto, $\dim(\mathbb{C}) = 2$.

6. En el espacio $(\mathbb{C}^2, \mathbb{K}, +, \cdot)$, la base canónica es:

$$C = \{(1,0),\ (\imath,0),\ (0,1),\ (0,\imath)\}.$$

Por lo tanto, $\dim(\mathbb{C}^2) = 4$.

7. En el espacio $(\mathscr{P}_n[x], \mathbb{R}, +, \cdot)$, la base canónica es:

$$C = \{1, x, x^2, \ldots, x^n\}.$$

Entonces, $\dim(\mathscr{P}_n[x]) = n + 1$.

8. En el espacio $(\mathbb{M}_{2\times 2}, \mathbb{R}, +, \cdot)$, la base canónica es:

$$C = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

♦

Del teorema 4.9, se deducen dos corolarios adicionales.

Corolario 4.11 Si la dimensión de $(V, \mathbb{K}, +, \cdot)$ es n, entonces cualquier subconjunto de V con más de n elementos es linealmente dependiente. En otras palabras, la base de un espacio vectorial es el conjunto linealmente independiente máximo en dicho espacio.

Demostración. En este caso, en el teorema 4.9, el conjunto que genera V es una base, la misma que tiene n elementos. □

Corolario 4.12 Si la dimensión de $(V, \mathbb{K}, +, \cdot)$ es n, entonces ningún subconjunto de V con menos elementos de n puede generar V. En otras palabras, la base de un espacio vectorial es el conjunto mínimo que puede generar el espacio.

Demostración. Supongamos lo contrario: existe un subconjunto W de V con m elementos y $m < n$ tal que W genera V. Sea B una base de V. Entonces, B tiene n elementos. Por lo tanto, por el teorema 4.9, B debe ser linealmente dependiente, lo que es imposible. □

En resumen, la base de un espacio vectorial es el conjunto con el mayor número de elementos que pueden ser linealmente independientes, y el menor número de elementos suficientes para generar el espacio.

En el siguiente teorema, se expresa de otra manera el carácter maximal de una base como conjunto linealmente independiente.

Teorema 4.13 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $B = \{v_1, v_2, \ldots, v_m\} \subset V$ y $u \in V$. Si B es linealmente independiente y $u \notin \langle B \rangle$, entonces el conjunto $B \cup \{u\}$ es linealmente independiente.

En otras palabras, si un conjunto linealmente independiente no genera el espacio (es decir, no es una base), entonces si se añade a dicho conjunto un elemento que no ha sido generado por él, el conjunto resultante sigue siendo linealmente independiente.

Demostración. Sean $\alpha_i \in \mathbb{R}$ con $i \in I_{m+1} = \{1, 2, \ldots, m+1\}$ tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v + \alpha_{m+1} u = 0_v. \tag{4.17}$$

Vamos a demostrar que $\alpha_i = 0$ para todo $i \in I_{m+1}$.

Si $\alpha_{m+1} \neq 0$, de la igualdad (4.17), obtenemos que:

$$u = -\frac{\alpha_1}{\alpha_{m+1}} v_1 - \frac{\alpha_2}{\alpha_{m+1}} v_2 - \cdots - \frac{\alpha_m}{\alpha_{m+1}} v_m,$$

lo que implica que $u \in \langle B \rangle$. Pero esto no es posible, por lo que concluimos que $\alpha = 0$.

Entonces, la igualdad (4.17) se transforma en la siguiente:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v = 0_v,$$

y de aquí se desprende que $\alpha_i = 0$ para todo $i \in I_m$, ya que el conjunto B es linealmente independiente.

En resumen, $\alpha_i = 0$ para todo $i \in I_{m+1}$, lo que significa que $B \cup \{u\}$ es linealmente independiente. □

Teorema 4.14 Sea $(W, \mathbb{K}, +, \cdot)$ un espacio vectorial de dimensión n, entonces todo subconjunto linealmente independiente de W es finito y es parte de una base de W.

Demostración. Puesto que $\dim(W) = n$, cualquier base de W tiene n elementos. Además, como todo conjunto linealmente independiente de W debe tener n elementos a lo más, ese conjunto es finito.

Por otro lado, sea $S \subset W$ linealmente independiente. Si S generara W, entonces S sería una base de W (por lo tanto, parte de una base). Supongamos, entonces, que $S \neq \langle W \rangle$. Existe, entonces, $w_1 \in W$ tal que $w_1 \neq \langle W \rangle$. Entonces, por el teorema anterior, el conjunto $S_1 = S \cup \{w_1\}$ es linealmente independiente.

Si S_1 generara W, entonces S, que es subconjunto de S_1, sería parte de una base de W. Pero, si $S_1 \neq \langle W \rangle$, entonces volvería a realizar el proceso anterior hasta obtener un conjunto B de n elementos y linealmente independiente, lo que no llevaría más allá de $n - m$ pasos, donde m es el número de elementos de S.

Pero este conjunto sería una base para W, porque si no generaría W, entonces, ejecutando un paso más el proceso anterior, obtendríamos un conjunto linealmente independiente con $n + 1$ elementos, lo que es imposible, pues $\dim(W) = n$. □

Consecuencia inmediata del proceso utilizado en la demostración de este teorema es que si $\dim(W == n$, todo conjunto linealmente independiente de n elementos es una base de W. Resumamos este resultado en el siguiente corolario.

Corolario 4.15 Sea $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial tal que $\dim(V) = n$. Entonces, un conjunto B linealmente independiente es una base de V si y solo si tiene n elementos.

La utilidad de este resultado es evidente: si $\dim(V) = n$ es conocido, para determinar que un conjunto con n elementos es base, únicamente hay que constatar que es linealmente independiente, ya no es necesario mostrar que también genera V.

También es verdad que si el conjunto tiene n elementos y genera el espacio, es una base. Para probar esta afirmación, requerimos algunos probar aún algunos resultados adicionales. Antes de hacerlo, veamos un ejemplo.

Ejemplo 4.52

Determinar si el conjunto

$$B = \{(0,1,1),\ (1,0,1),\ (1,1,0)\}$$

es una base del espacio vectorial $(\mathbb{R}^3, \mathbb{K}, +, \cdot)$.

Solución. Puesto que B tiene tres elementos y la dimensión de $\mathbb{R}^3$ es 3, para determinar si B es una base, es suficiente determinar si es un conjunto linealmente independiente.

Para ello, sean α, β y γ escalares tales que:

$$\alpha(0,1,1) + \beta(1,0,1) + \gamma(1,1,0) = (0,0,0). \tag{4.18}$$

Entonces:

$$(\beta+\gamma,\ \alpha+\gamma,\ \alpha+\beta) = (0,0,0).$$

Por lo tanto, si la igualdad (4.18) fuera verdadera, entonces α, β y γ satisfarían el siguiente sistema de ecuaciones:

$$\begin{aligned} \beta + \gamma &= 0 \\ \alpha \quad + \gamma &= 0 \\ \alpha + \beta \quad &= 0, \end{aligned}$$

cuya matriz de coeficientes:

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

tiene determinante es diferente de 0, pues:

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

Por lo tanto, la solución del sistema (4.18) es la trivial: $\alpha = \beta = \gamma = 0$.

En otras palabras, el conjunto B es linealmente independiente y, por lo tanto, es una base de $\mathbb{R}^3$. ♦

Ejemplo 4.53

El conjunto de $\mathbb{R}^3$
$$B = \{(1,0,1),\ (1,-1,0)\}$$
es linealmente independiente. Obtener una base B' para $\mathbb{R}^3$ de manera que B' contenga a B.

Solución. El procedimiento a utilizar es el dado en la demostración del teorema 4.14. Consiste en añadir un elemento $u \in \mathbb{R}^3$ tal que $u \notin \langle B \rangle$. Sabemos que este elemento existe, porque B no puede ser una base de $\mathbb{R}^3$, ya que B tiene dos elementos, mientras que $\dim(\mathbb{R}^3) = 3$.

Empecemos, entonces, obteniendo la cápsula de B:

$$\begin{aligned}\langle B \rangle &= \{(a,b,c) \in \mathbb{R}^3 : (a,b,c) = \alpha(1,0,1) + \beta(1,-1,0)\}\\ &= \{(a,b,c) \in \mathbb{R}^3 : (a,b,c) = (\alpha+\beta, -\beta, \alpha)\}\\ &= \{(a,b,c) \in \mathbb{R}^3 : \alpha+\beta = a,\ -\beta = b,\ \alpha = c)\}.\end{aligned}$$

Ahora determinemos las condiciones que deben cumplir a, b y c para que el sistema

$$\begin{aligned}\alpha + \beta &= a\\ -\beta &= b\\ \alpha &= c\end{aligned} \tag{4.19}$$

tenga solución. Para ello, trabajemos con su matriz ampliada:

$$\left(\begin{array}{cc|c}1 & 1 & a\\ 0 & -1 & b\\ 1 & 0 & c\end{array}\right) \sim \left(\begin{array}{cc|c}1 & 1 & a\\ 0 & -1 & b\\ 0 & -1 & -a+c\end{array}\right) \sim \left(\begin{array}{cc|c}1 & 1 & a\\ 0 & -1 & b\\ 0 & -1 & -a-b+c\end{array}\right).$$

Por lo tanto, el sistema 4.19 tiene solución si y solo si

$$-a+b+c = 0.$$

Entonces:
$$\langle B \rangle = \{(a,b,c) \in \mathbb{R}^3 : c = a+b\}.$$

Ahora busquemos un elemento de $\mathbb{R}^3$ que no esté en $\langle B \rangle$. Por ejemplo, el vector $(0,0,1)$ (la tercera coordenada no es la suma de las dos primeras) no está en $\langle B \rangle$. Por lo tanto, el conjunto $B' = B \cup \{(0,0,1)\}$ es linealmente independiente (teorema 4.13). Como tiene tres elementos, B' es una base de $\mathbb{R}^3$ y contiene a B. ♦

El siguiente es otro corolario del teorema (4.9).

Teorema 4.16 Sea $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial de dimensión n. Si W es un subespacio vectorial propio de V (es decir, $W \subset V$ y $W \neq V$), entonces W tiene dimensión finita y $\dim(W) < \dim(V)$.

Demostración. Sea B una base de W. Entonces $B \subset V$. Como B es linealmente independiente y V tiene dimensión finita, B tiene n elementos a lo más (teorema 4.14). Por lo tanto:
$$\dim(W) \leq \dim(V) = n.$$

Esto significa que la dimensión de W es finita.

Ahora probemos que n no puede ser la dimensión de W, con lo que quedará demostrado que $\dim(W) < \dim(V)$.

Ya que W es un subconjunto propio de V, entonces $\langle B \rangle = W \neq V$. Sea, entonces, $u \in V$ tal que $u \notin \langle B \rangle$. Como B es linealmente independiente, por el teorema 4.13 podemos concluir que $B \cup \{u\}$ es linealmente independiente. Entonces, por el teorema 4.14, $B \cup \{u\}$ es parte de una base de V. Esto significa que:

$$\dim(W) < \dim(W) + 1 \leq \dim(V).$$

□

Teorema 4.17 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y S un subconjunto finito de V. Si S es linealmente dependiente y $v \in S$ es combinación lineal de los demás elementos de S, entonces:

$$\langle S \rangle = \langle S - \{v\} \rangle .$$

Demostración. Supongamos que $S = \{v_1, v_2, \ldots, v_n\}$ y que $v = v_j$ con $j \in I_n = \{1, 2, \ldots, n\}$. Vamos a probar que $\langle S \rangle \subset \langle S - \{v\} \rangle$ y que $\langle S - \{v\} \rangle \subset \langle S \rangle$.

1. Sea $u \in \langle S \rangle$. Entonces existen n escalares α_i con $i \in I_n$ tales que

   $$u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_j v_j + \cdots + \alpha_n v_n.$$

   Por otro lado, por hipótesis, sabemos que existen $n-1$ escalares β_i con $i \in I_n - \{j\}$ tales que:

   $$v_j = \beta_1 v_1 + \cdots + \beta_{j-1} v_{j-1} + \beta_{j+1} v_{j+1} + \cdots + \beta_n v_n.$$

   Si sustituimos este valor de v_j en la primera igualdad, obtenemos que:

   $$u = \gamma_1 v_1 + \cdots + \gamma_{j-1} v_{j-1} + \gamma_{j+1} v_{j+1} + \cdots + \gamma_n v_n,$$

   donde $\gamma_i = \alpha_i + \alpha_i \beta_j$ para cada $i \in I_n - \{j\}$.

   Pero esta igualdad nos dice que u es combinación lineal de los elementos de $S - \{v_j\}$. Es decir, $u \in \langle S - \{v_j\} \rangle$.

   En otras palabras, hemos probado que todo elemento de $\langle S \rangle$ es elemento de $\langle S \rangle - \{v_j\}$.

2. Sea $u \in \langle S \rangle - \{v_j\}$. Mostremos que $u \in \langle S \rangle$. Pero esto es verdadero, porque, al ser u combinación lineal de elementos de $S - \{v_j\}$, también es combinación lineal de S, ya que $S \subset S - \{v_j\}$.

□

Ejemplo 4.54

Sea
$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} 2 & 2 \\ 1 & -3 \end{pmatrix} \right\}.$$

Obtener:

1. la cápsula de S, $\langle S \rangle$; y
2. una base y la dimensión de $\langle S \rangle$.

Solución.

1. Observemos que el conjunto S es linealmente dependiente, ya que
$$2\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + (-1)\begin{pmatrix} 0 & -2 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1 & -3 \end{pmatrix}.$$

Por el teorema anterior, al eliminar el vector
$$\begin{pmatrix} 2 & 2 \\ 1 & -3 \end{pmatrix}$$
de S, la cápsula de S menos ese vector sigue siendo igual a la cápsula de S:
$$\begin{aligned} \langle S \rangle &= \left\langle \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ -1 & -1 \end{pmatrix} \right\} \right\rangle \\ &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \beta \begin{pmatrix} 0 & -2 \\ -1 & -1 \end{pmatrix} \right\}. \end{aligned}$$

Ahora veamos qué condiciones deben satisfacer a, b, c y d para que el sistema
$$\begin{aligned} \alpha \phantom{{}-2\beta} &= a \\ -2\beta &= b \\ -\beta &= c \\ -\alpha + \beta &= d \end{aligned} \tag{4.20}$$
tenga solución. De las matrices equivalentes de la matriz ampliada:
$$\left(\begin{array}{rr|c} 1 & 0 & a \\ 0 & -2 & b \\ 0 & -1 & c \\ -1 & 1 & d \end{array}\right) \sim \left(\begin{array}{rr|c} 1 & 0 & a \\ 0 & -2 & b \\ 0 & -1 & c \\ 0 & 1 & a+d \end{array}\right) \sim \left(\begin{array}{rr|c} 1 & 0 & a \\ 0 & 1 & a+d \\ 0 & -1 & c \\ 0 & -2 & b \end{array}\right) \sim \left(\begin{array}{rr|c} 1 & 0 & a \\ 0 & 1 & a+d \\ 0 & 0 & a+c+d \\ 0 & 0 & 2a+b+2d \end{array}\right)$$
concluimos que el sistema (4.20) tiene solución si
$$a + c + d = 0 \quad \text{y} \quad 2a + b + 2d = 0.$$

Por lo tanto:
$$\langle S \rangle = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2} : a + c + d = 0,\ 2a + b + 2d = 0 \right\}.$$

2. El conjunto
$$\left\{\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}0 & -2\\ -1 & -1\end{pmatrix}\right\}$$
es el ideal para ser la base de $\langle S\rangle$. Para confirmarlo, lo único que resta por hacer es probar que es linealmente independiente. Para ello, sean α y β escalares tales que:
$$\alpha\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}+\beta\begin{pmatrix}0 & -2\\ -1 & -1\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}.$$
De esta igualdad, se obtiene que
$$\begin{pmatrix}\alpha & -2\beta\\ -\beta & -\alpha+\beta\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},$$
de donde tenemos que $\alpha=0$ y $\beta=0$. Por lo tanto, este conjunto es linealmente independiente y, por lo tanto, una base de $\langle S\rangle$. Finalmente, podemos afirmar que $\dim(\langle S\rangle)=2$.

♦

Teorema 4.18 Sean $(V,\mathbb{K},+,\cdot)$ un espacio vectorial de dimensión n y $B\subset V$ con n elementos. Entonces, B es una base de V si y solo si B genera V; es decir, si y solo si $\langle B\rangle=V$.

Demostración. Sea B un subconjunto de V con n elementos. Si B es una base, entonces B genera el espacio V.

Recíprocamente, supongamos que B genera el espacio V y que no es una base de V. Entonces, B es linealmente dependiente. Existe, entonces, $u\in B$ que es combinación lineal de los otros elementos de B. Por lo tanto, por el teorema 4.17, el conjunto $B-\{u\}$, que tiene $n-1$ elementos, también genera V. Pero esto es imposible, ya que la dimensión de V es n por lo que ningún conjunto con menos de n elementos puede generar el espacio (corolario 4.12). De modo que, B es linealmente independiente. Es decir, es una base de V. □

Ejemplo 4.55

El conjunto
$$W=\{(a,b,c,d)\in\mathbb{R}^4 : b-c=0\}$$
es un subespacio vectorial de $\mathbb{R}^4$ y su dimensión es igual a 3. Encontrar una base para W.

Solución. El conjunto W puede ser caracterizado de la siguiente manera:
$$W=\{(a,c,c,d)\in\mathbb{R}^4 : a,\ c,\ d\in\mathbb{R}\}.$$
Por lo tanto:
$$\begin{aligned}W&=\{(a,0,0,0)+(0,c,c,0)+(0,0,0,d)\in\mathbb{R}^4 : a,\ c,\ d\in\mathbb{R}\}\\ &=\{a(1,0,0,0)+c(0,1,1,0)+d(0,0,0,1)\in\mathbb{R}^4 : a,\ c,\ d\in\mathbb{R}\}.\end{aligned}$$
Si definimos el conjunto B por:
$$B=\{(1,0,0,0),\ (0,1,1,0),\ (0,0,0,1)\},$$
entonces $W=\langle B\rangle$ y, como $\dim(W)=3$ y B tiene tres elementos, entonces B es una base de W.

♦

Ejemplo 4.56

Sea $S = \{f, \operatorname{sen}, \cos\}$ donde $f(x) = x$ para todo $x \in \mathbb{R}$. Probar que S es linealmente independiente.

Solución. Sean α, β y γ tres escalares tales que

$$\alpha f + \beta \operatorname{sen} + \gamma \cos = 0_v. \tag{4.21}$$

Vamos a probar que $\alpha = \beta = \gamma = 0$.

La igualdad (4.21) implica que

$$\alpha f(x) + \beta \operatorname{sen} x + \gamma \cos x = 0_v(x)$$

para todo $x \in \mathbb{R}$. Es decir, la igualdad

$$\alpha x + \beta \operatorname{sen} x + \gamma \cos x = 0 \tag{4.22}$$

para todo $x \in \mathbb{R}$.

Por lo tanto, si sustituimos 0 por x en (4.22), tenemos que

$$\alpha(0) + \beta(0) + \gamma(1) = 0,$$

de donde obtenemos que $\gamma = 0$.

Ahora, si sustituimos π por x, resulta que

$$\alpha(\pi) + \beta(0) = 0,$$

de donde se comprueba que $\alpha = 0$.

Finalmente, si hacemos $x = \frac{\pi}{2}$, entonces

$$\beta(1) = 0,$$

de donde $\beta = 0$.

En resumen, si se verifica la igualdad (4.21), necesariamente $\alpha = \beta = \gamma = 0$. Entonces, S es linealmente independiente. ♦

Existe un procedimiento para verificar que un conjunto de funciones diferenciables son linealmente independientes: el wronskiano.

Definición 4.10 (Wronskiano) Sean $f_1, f_2, \ldots, f_n$ funciones reales $n-1$ veces diferenciables. El *wronskiano* de estas funciones es la función con dominio la intersección de los dominios de todas las f_i definida por

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

para todo $x \in \bigcap_{i=1}^{n} \bigcap_{j=1}^{n-1} \operatorname{Dm}(f_i^{(j)})$.

Teorema 4.19 Sea $S = \{f_1, f_2, \ldots, f_n\}$ subconjunto del espacio vectorial $(\mathscr{F}, \mathbb{R}, +, \cdot)$. Si cada elemento de S es una $n-1$ veces diferenciables, entonces S es linealmente independiente si y solo existe al menos $x \in \bigcap_{i=1}^{n} \bigcap_{j=1}^{n-1} \mathrm{Dm}(f_i^{(j)})$ tal que $W(x) \neq 0$.

Ejemplo 4.57 Sea $S = \{f, \ \mathrm{sen}, \cos\}$ donde $f(x) = x$ para todo $x \in \mathbb{R}$. Con ayuda del wronskiano también podemos probar que S es linealmente independiente, pues, para todo $x \in \mathbb{R}$ se tiene que:

$$\begin{aligned}
W(x) &= \begin{vmatrix} f(x) & \mathrm{sen}\, x & \cos x \\ f'(x) & \cos x & -\mathrm{sen}\, x \\ f''(x) & -\mathrm{sen}\, x & -\cos x \end{vmatrix} \\
&= \begin{vmatrix} x & \mathrm{sen}\, x & \cos x \\ 1 & \cos x & -\mathrm{sen}\, x \\ 0 & -\mathrm{sen}\, x & -\cos x \end{vmatrix} \\
&= \begin{vmatrix} x & 0 & 0 \\ 1 & \cos x & -\mathrm{sen}\, x \\ 0 & -\mathrm{sen}\, x & -\cos x \end{vmatrix} = x \begin{vmatrix} \cos x & -\mathrm{sen}\, x \\ -\mathrm{sen}\, x & -\cos x \end{vmatrix}.
\end{aligned}$$

Por lo tanto:

$$W(x) = x(-\cos^2 x - \mathrm{sen}^2 x) = -x(\mathrm{sen}^2 x + \cos^2 x) = -x.$$

Luego, para $x \neq 0$, tenemos que $W(x) \neq 0$. ♦

4.10 Suma de subespacios vectoriales

Definición 4.11 (Suma de subespacios) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y W_1 y W_2 dos subespacios vectoriales de V. Se define $W_1 + W_2$ por:

$$W_1 + W_2 = \{v \in V : v = w_1 + w_2, \ w_1 \in W_1, w_2 \in W_2\}.$$

Es decir, $W_1 + W_2$ es el conjunto de todos los elementos del espacio vectorial V que se pueden expresar como la suma de un elemento del subespacio vectorial W_1 con otro elemento del subespacio vectorial W_2.

Teorema 4.20 Sean $(V, \mathbb{K}, +, \cdot)$ y W_1 y W_2 dos subespacios vectoriales de V. Entonces, $W_1 + W_2$ es un subespacio vectorial de V.

Demostración. 1. Es inmediato que $W_1 + W_2$ es diferente del vacío ya que $0_v \in W_1 + W_2$, pues $0_v = 0_v + 0_v$ y 0_v es tanto elemento de W_1 como de W_2 (por ser subespacios vectoriales de V).

2. Sean $\alpha \in \mathbb{K}$, $u \in W_1 + W_2$ y $v \in W_1 + W_2$. Probemos que $\alpha u + v$ también es un elemento de $W_1 + W_2$.

   Existen $u_1 \in W_1$, $u_2 \in W_2$, $v_1 \in W_1$ y $v_2 \in W_2$ tales que

$$u = u_1 + u_2 \quad \text{y} \quad v = v_1 + v_2.$$

   Entonces:

$$\begin{aligned} \alpha u + v &= \alpha(u_1 + u_2) + (v_1 + v_2) \\ &= (\alpha u_1 + v_1) + (\alpha u_2 + v_2) \in W_2 + W_2, \end{aligned}$$

   ya que $\alpha u_1 + v_1 \in W_1$ y $\alpha u_2 + v_2 \in W_2$ (W_1 y W_2 son subespacios vectoriales).

□

Teorema 4.21 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y W_1 y W_2 subespacios vectoriales de V de dimensión finita. Si B_1 y B_2 son bases de W_1 y W_2, respectivamente, entonces $W_1 + W_2 = \langle B1 \cap B_2 \rangle$.

Demostración. Vamos a probar que $W_1 + W_2 \subset \langle B_1 \cap B_2 \rangle$ y que $\langle B_1 \cap B_2 \rangle \subset W_1 + W_2$.

1. Sea $w \in W_1 + W_2$. Mostremos que $w \in \langle B_1 \cap B_2 \rangle$; es decir, probemos que w es una combinación lineal de elementos de la unión de las dos bases.

   Supongamos que $B_1 = \{u_1, u_2, \ldots, u_m\}$ y $B_2 = \{v_1, v_2, \ldots, v_n\}$. Sean $u \in W_1$ y $v \in W_2$ tales que $w = u + v$. Existen, entonces escalares α_i con $i \in I_m = \{1, 2, \ldots, m\}$ y β_j con $j \in I_n = \{1, 2, \ldots, n\}$ tales que

$$u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m \quad \text{y} \quad v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n. \tag{4.23}$$

   Por lo tanto:

$$w = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m + \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n, \tag{4.24}$$

   donde los u_i y v_j pertenecen a $W_1 \cap W_2$. Esto significa que $u + v \in \langle B_1 \cap B_2 \rangle$.

2. Sea $w \in \langle B_1 \cap B_2 \rangle$. Entonces existen escalares α_i y β_j que satisfacen la igualdad (4.24). Es inmediato entonces que $w = u + v$ con u y v dados en la igualdad (4.23). Por lo tanto, $w \in W_1 + W_2$.

□

Teorema 4.22 (de la dimensión) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y W_1 y W_2 dos subespacios vectoriales de V de dimension finita. Entonces, se verifica la siguiente igualdad:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

Demostración. Si uno de los dos subespacios vectoriales es el espacio nulo, la demostración es inmediata. En efecto, si fuera el caso que $W_1 = \{0_v\}$, entonces $W_1 + W_2 = W_2$ y $W_1 \cap W_2 = \{0_v\}$, y $\dim(W_1) = \dim(W_1 \cap W_2) = 0$. Por lo tanto:

$$\begin{aligned}\dim(W_1 + W_2) &= \dim(W_2) = 0 + \dim(W_2) - 0 \\ &= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).\end{aligned}$$

Supongamos ahora que tanto W_1 como W_2 son ambos diferentes del espacio nulo. Realizaremos la demostración en dos casos.

1. Supongamos que $W_1 \cap W_2 = \{0_v\}$. Debemos probar, entonces, que la dimensión de $W_1 + W_2$ es la suma de las dimensiones de W_1 y W_2 respectivamente.

   Para ello, vamos a probar que una base para $W_1 + W_2$ se obtiene mediante la unión entre una base de W_1 y otra de W_2. Como $W_1 \cap W_2 = \{0_v\}$, entonces, si B_1 y B_2 son bases de W_1 y W_2 respectivamente, se tiene que $B_1 \cap B_2 = \emptyset$ (ya que el vector nulo no puede ser nunca un elemento de una base). Por lo tanto, la unión de las dos bases contiene $m + n$ elementos donde m es el número de elementos de B_1 y n el número de elementos de B_2. Con ello se prueba que

   $$\begin{aligned}\dim(W_1 + W_2) &= m + n = \\ &= \dim(W_1) + \dim(W_2) \\ &= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2),\end{aligned}$$

   ya que $W_1 \cap W_2 = \{0_v\}$.

   Probemos, entonces, que si $B_1 = \{u_1, u_2, \ldots, u_m\}$ es una base de W_1 y $B_2 = \{v_1, v_2, \ldots, v_n\}$ es una base de W_2, entonces $B_1 \cup B_2$ es una base para $W_1 + W_2$.

   Por el teorema 4.21, sabemos que $B_1 \cup B_2$ genera $W_1 + W_2$. Por lo tanto, solo resta probar que la unión de las bases es linealmente independiente.

   Sean α_i con $i \in I_m = \{1, 2, \ldots, m\}$ y β_j con $j \in I_n = \{1, 2, \ldots, n\}$ escalares tales que

   $$\sum_{i=1}^{m} \alpha_i u_i + \sum_{j=1}^{n} \beta_j v_j = 0_v.$$

   Entonces:

   $$\sum_{i=1}^{m} \alpha_i u_i = -\sum_{j=1}^{n} \beta_j v_j.$$

   Por un lado, tenemos que

   $$\sum_{i=1}^{m} \alpha_i u_i \in W_1 \quad \text{y} \quad -\sum_{j=1}^{n} \beta_j v_j \in W_2,$$

   y por otro, tenemos que $W_1 \cap W_2 = \{0_v\}$. Entonces:

   $$\sum_{i=1}^{m} \alpha_i u_i = -\sum_{j=1}^{n} \beta_j v_j = 0_v.$$

   Por lo tanto, ya que B_1 y B_2 son ambos linealmente independientes, tenemos que $\alpha_i = 0$ y $\beta_j = 0$ para todo $i \in I_m$ y todo $j \in I_n$. Esto demuestra que $B_1 \cup B_2$ es linealmente independiente.

2. Supongamos que $W_1 \cap W_2 \neq \{0_v\}$. Como $W_1 \cap W_2$ es subconjunto tanto de W_1 como de W_2, que son de dimensión finita, por el teorema 4.16, la intersección también es un espacio de dimensión finita y, además, se verifica que:

$$\dim(W_1 \cap W_2) \leq \dim(W_1) \quad \text{y} \quad \dim(W_1 \cap W_2) \leq \dim(W_2).$$

Sea $B = \{w_i : i \in I_p = \{1, 2, \ldots, p\}\}$ una base de $W_1 \cap W_2$. Por el teorema 4.14, B es parte de una base de B_1, una de W_1, y parte de B_2, una base de W_2.

Sean $B_1 = \{w_i, u_j : i \in I_p,\ j \in I_m = \{1, 2, \ldots, m\}\}$ y $B_2 = \{w_i, v_j : i \in I_p,\ k \in I_n = \{1, 2, \ldots, m\}\}$. Entonces

$$\dim(W_1) = p + m \quad \text{y} \quad \dim(W_2) = p + n. \tag{4.25}$$

Observemos que ningún u_j puede ser igual a algún v_k, pues si fuera así, digamos que $u_1 = v_1$, por ejemplo, entonces u_1 y v_1 estarían en la intersección de $W_1 \cap W_2$ y serían uno de los w_i, pues, de lo contrario, ya no sería p la dimensión de $W_1 \cap W_2$.

Por esto, basta que probemos que

$$B_1 \cup B_2 = \{w_i, u_j, v_k : i \in I_p,\ j \in I_m,\ k \in I_n\}$$

es una base de $W_1 + W_2$ para demostrar el teorema, pues, la unión de las bases contendría exactamente

$$p + m + n$$

elementos, de donde tendríamos, con ayuda de la igualdad (4.25), que

$$\begin{aligned}
\dim(W_1 + W_2) &= p + m + n \\
&= (p + m) + (p + n) - p \\
&= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).
\end{aligned}$$

Demostremos, entonces, que $B_1 \cup B_2$ es una base, para lo cual, es suficiente con probar que es linealmente independiente, ya que, por el teorema 4.21, genera el espacio $W_1 + W_2$.

Sean, entonces, α_i con $i \in I_m$, β_j con $j \in I_n$ y $\gamma_k \in k \in I_p$ escalares tales que:

$$\sum_{k=1}^{p} \gamma_k w_k + \sum_{i=1}^{m} \alpha_i u_i + \sum_{j=1}^{n} \beta_j v_j = 0_v. \tag{4.26}$$

Entonces:

$$\sum_{k=1}^{p} \gamma_k w_k + \sum_{i=1}^{m} \alpha_i u_i = -\sum_{j=1}^{n} \beta_j v_j. \tag{4.27}$$

Pero

$$\sum_{k=1}^{p} \gamma_k w_k + \sum_{i=1}^{m} \alpha_i u_i \in W_1 \quad \text{y} \quad -\sum_{j=1}^{n} \beta_j v_j \in W_2,$$

junto con la igualdad (4.27), implica que:

$$-\sum_{j=1}^{n} \beta_j v_j \in W_1 \cap W_2.$$

Entonces, como B es una base de $W_1 \cap W_2$, existen p escalares δ_k con $k \in I_p$ tales que

$$-\sum_{j=1}^{n} \beta_j v_j = \sum_{k=1}^{p} \delta_k w_k,$$

de donde obtenemos que:

$$\sum_{k=1}^{p} \delta_k w_k + \sum_{j=1}^{n} \beta_j v_j = 0_v.$$

Pero esta igualdad implica que $\delta_k = 0$ y $\beta_j = 0$ para todo $k \in I_p$ y todo $j \in I_n$, ya que B_2 es un conjunto linealmente independiente.

Entonces, la igualdad (4.27) se expresa de la siguiente manera:

$$\sum_{k=1}^{p} \gamma_k w_k + \sum_{i=1}^{m} \alpha_i u_i = 0_v,$$

de donde se obtiene que $\gamma_k = 0$ y $\alpha = 0$ para todo $k \in I_p$ y todo $i \in I_m$, ya que B_1 es un conjunto linealmente independiente.

En resumen, hemos probado que la igualdad (4.26) implica que todos los escalares α_i, β_j y γ_k son iguales a 0. Es decir, hemos probado que $B_1 \cup B_2$ es linealmente independiente, y con ello, que es una base de $W_1 + W_2$.

□

4.11 Suma directa de subespacios vectoriales

Definición 4.12 (Suma directa) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y W_1 y W_2 dos subespacios vectoriales de V. La *suma directa* de W_1 y W_2, representada por $W_1 \oplus W_2$ se define de la siguiente manera:

$$W_1 \oplus W_2 = \{w \in V : (\exists! u \in W_1)(\exists! v \in W_2)(w = u + v)\}.$$

Teorema 4.23 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y U, W_1 y W_2 subespacios de dimensión finita de V. Entonces:

$$U = W_1 \oplus W_2 \text{ si y solo si } U = W_1 + W_2 \text{ y } W_1 \cap W_2 = \{0_v\}.$$

Demostración. 1. Si $U = W_1 \oplus W_2$, es inmediato, de la definición de suma directa que $U = W_1 + W_2$. Así que probemos que la intersección de W_1 y W_2 es el espacio nulo.

Supongamos lo contrario. Esto significa que existe $v \neq 0_v$ tal que $v \in W_1 \cap W_2$. Entonces, por un lado, tenemos que

$$2v = v + v$$

está en $W_1 \oplus W_2$. Por otro, tenemos que

$$2v = 2v + 0_v.$$

Como la cada elemento de la suma directa se expresa de manera única como la suma de dos elementos, uno de W_1 y otro de W_2, necesariamente tiene que suceder que $v = 2v$ y que $v = 0_v$, ambas afirmaciones imposibles. Por lo tanto, $W_1 \cap W_2 = \{0_v\}$.

2. Supongamos ahora que $U = W_1 + W_2$ y que $W_1 \cap W_2 = \{0_v\}$. Para probar que U es la suma directa de W_1 y W_2, vamos a probar que cada elemento de $W_1 + W_2$ se expresa de manera única como la suma un elemento de W_1 con otro de W_2.

   Sean $u = v + w \in U$ tal que $v \in W_1$ y $w \in W_2$. Supongamos que existen $v' \in W_1$ y $w' \in W_2$ tales que $u = v' + w'$. Esto significa que

   $$v + w = v' + w',$$

   de donde obtenemos que

   $$v - v' = w' - w. \tag{4.28}$$

   Pero $v - v' \in W_1$ y $w' - w \in W_2$. Por lo tanto, la igualdad (4.28) implica que

   $$v - v' \in W_1 \cap W_2 \quad \text{y} \quad w' - w \in W_1 \cap W_2.$$

   De donde, como $W_1 \cap W_2 = \{0_v\}$, tenemos que

   $$v - v' = 0_v \quad \text{y} \quad w' - w = 0_v.$$

   Por lo tanto $v = v'$ y $w = w'$, lo que implica que $u \in W_1 \oplus W_2$.

□

De este teorema y del teorema de la dimensión, ya que $\dim(\{0_v\}) = 0$, se deduce el siguiente corolario.

Corolario 4.24 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y W_1 y W_2 subespacios vectoriales de V de dimensión finita. Entonces:

$$\dim(W_1 \oplus W_2) = \dim(W_1) + \dim(W_2).$$

Teorema 4.25 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial de dimensión finita n y U un subespacio vectorial de V. Entonces, existe un subespacio vectorial W de V tal que $V = U \oplus V$.

Demostración. Si U es el espacio nulo, entonces $W = V$. Si U es todo V, es claro que W debe ser el espacio nulo.

Supongamos, entonces, que U ni es el espacio nulo ni es V. Como la dimensión de U es finita, ya que V es de dimensión finita, sea $m = \dim(U)$. Por lo tanto $m \leq n$.

Sea $B_1 = \{u_1, u_2, \ldots, u_m$ una base de U. Por el teorema 4.14, existe B una base de V que contiene a B_1. Por ello, B es de la forma:

$$B = \{u_1, u_2, \ldots, u_m, u_{m+1}, \ldots, u_n\}.$$

El candidato perfecto para W es, entonces, $\langle B - B_1 \rangle$.

En efecto, probemos que:

1. $V = U + W$; y
2. $U \cap W = \{0_v\}$.

1. Es inmediato que $U + W \subset V$. Recíprocamente, si $v \in V$, entonces existen escalares α_i con $i \in I_n = \{1, 2, \ldots, n\}$ tales que:

$$\begin{aligned} v &= \sum_{i=1}^{n} \alpha_i u_i \\ &= \sum_{i=1}^{m} \alpha_i u_i + \sum_{i=m+1}^{n} \alpha_i u_i \\ &= u + v, \end{aligned}$$

donde

$$u = \sum_{i=1}^{m} \alpha_i u_i \quad \text{y} \quad v = \sum_{i=m+1}^{n} \alpha_i u_i.$$

Entonces $u \in U$ y $v \in W$. Por lo tanto, $v \in U + W$; es decir, $V \subset U + W$.

En resumen, hemos probado que $V = U + W$.

2. Sea $v \in U \cap W$. Vamos a probar que $v = 0_v$, con lo cual quedaría demostrado que $U \cap W = \{0_v\}$.

Como $u \in U$ y $u \in V$, existen escalares α_i tales que

$$v = \sum_{i=1}^{m} \alpha_i u_i = \sum_{i=m+1}^{n} \alpha_i u_i.$$

Por lo tanto:

$$0_v = v - v = \sum_{i=1}^{m} \alpha_i u_i - \sum_{i=m+1}^{n} \alpha_i u_i,$$

de donde se tiene que

$$\sum_{i=1}^{m} \alpha_i u_i + \sum_{i=m+1}^{n} (-\alpha_i) u_i = 0_v.$$

Como B es linealmente independiente, tenemos que todos los $\alpha_i = 0$. Por lo tanto, $v = 0_v$.

□

Ejemplo 4.58

Sea $U = \{a + bx + cx^2 + dx^3 \in \mathscr{P}_3[x] : a - 2b - c = 0,\ b - c - d = 0\}$ un subespacio vectorial de $(\mathscr{P}_3[x], \mathbb{R}, +, \cdot)$. Encontrar un subespacio vectorial W tal que $\mathscr{P}_3[x] = U \oplus W$.

Solución. Vamos a utilizar la demostración del teorema 4.25, ya que en ella se describe el siguiente procedimiento para encontrar W:

1. Encontrar una base para U.
2. Si U es el espacio nulo, entonces $W = \mathscr{P}_3[x]$.
3. Si U es el V, entonces W es el espacio nulo.
4. Si U no es ni el espacio nulo ni V, entonces $W = \langle B - B_1 \rangle$, donde B es una base de V obtenida a partir de B_1, una base de U.

En este caso, es fácil ver que U ni es el espacio nulo ni $\mathscr{P}_3[x]$. Entonces, lo que tenemos que hacer es: encontrar una base para U y luego completar esa base a una para $\mathscr{P}_3[x]$.

1. *Una base para U.*

$$\begin{aligned} U &= \{a + bx + cx^2 + dx^3 \in \mathscr{P}_3[x] : a = 2b + c,\ d = b - c\} \\ &= \{(2b + c) + bx + cx^2 + (b - c)x^3 \in \mathscr{P}_3[x] : b \in \mathbb{R}, c \in \mathbb{R}\} \\ &= \{(2b + bx + bx^3) + (c + cx^2 - cx^3) \in \mathscr{P}_3[x] : b \in \mathbb{R}, c \in \mathbb{R}\} \\ &= \{b(2 + x + x^3) + c(1 + x^2 - x^3) \in \mathscr{P}_3[x] : b \in \mathbb{R}, c \in \mathbb{R}\}. \end{aligned}$$

Sea $B_1 = \{2 + x + x^3,\ 1 + x^2 - x^3\}$. Entonces $U = \langle B_1 \rangle$. Para que B_1 sea una base de U, es suficiente que sea linealmente independiente. Veamos si es así efectivamente.

Sean α y β dos escalares tales que

$$\alpha(2 + x + x^3) + \beta(1 + x^2 - x^3) = 0 + 0x + 0x^2 + 0x^3.$$

Por lo tanto:

$$(2\alpha + \beta)\alpha x + \beta x^2 + (\alpha - \beta)x^3 = 0 + 0x + 0x^2 + 0x^3,$$

de donde es inmediato que $\alpha = 0$ y $\beta = 0$ son los únicos valores que satisface esta igualdad. Por lo tanto, B_1 es linealmente independiente.

En resumen, hemos encontrado una base para U, el conjunto B_1.

2. *Una base para V que contenga a B_1.* El procedimiento a seguir es el utilizado en la demostración del teorema 4.14, y consiste en lo siguiente:

   (a) Si B_1 genera V (es decir, $U = V$), entonces B_1 es una base de V.

   (b) Si U es diferente de V, entonces se determina si $B_2 = B_1 \cup \{v_1\}$, donde $v_1 \in V - \langle B_1 \rangle$, genera V. En caso afirmativo, B_2 es la base buscada. En el caso contrario, se repite este paso hasta encontrar B.

Como $\dim(\mathscr{P}_3[x]) = 4$, entonces U, cuya dimensión es 2, no es igual a $\mathscr{P}_3[x]$. Por lo tanto, debemos aplicar el segundo paso del procedimiento descrito. ¿Cuántas veces? Exactamente dos.

En primer lugar, busquemos un polinomio de $\mathscr{P}_3[x]$ que no esté en U. Por ejemplo, x, ya que $a \neq 2b - c$, pues $a = 0, b = 1$ y $c = 0$. Sea $B_2 = B_1 \cup \{x\}$. Como tiene tres elementos, entonces no genera $\mathscr{P}_3[x]$, cuya dimensión es 4.

Volvemos al segundo paso del procedimiento: buscar un vector que esté en $\mathscr{P}_3[x]$ pero no en $\langle B_2 \rangle$. Para esto, determinemos $\langle B_2 \rangle$:

$$\begin{aligned} \langle B_2 \rangle &= \{p(x) = a + bx + cx^2 + dx^3 \in \mathscr{P}_3[x] : p(x) = \alpha(2 + x + x^3) + \beta(1 + x^2 - x^3) + \gamma x\} \\ &= \{p(x) = a + bx + cx^2 + dx^3 \in \mathscr{P}_3[x] : p(x) = (2\alpha + \beta)(\alpha + \gamma)x + \beta x^2 + (\alpha - \beta)x^3\} \end{aligned}$$

Entonces, α, β y γ satisfacen el siguiente sistema de ecuaciones:

$$\begin{aligned} 2\alpha + \beta \quad\quad &= a \\ \alpha \quad\quad + \gamma &= b \\ \beta \quad\quad &= c \\ \alpha - \beta \quad\quad &= d. \end{aligned} \tag{4.29}$$

Ahora determinemos las condiciones que deben cumplir a, b, c y d para que este sistema tenga solución. Para ello, trabajemos con la matriz ampliada del sistema:

$$\begin{aligned} &\left(\begin{array}{ccc|c} 2 & 1 & 0 & a \\ 1 & 0 & 1 & b \\ 0 & 1 & 0 & c \\ 1 & -1 & 0 & d \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & d \\ 1 & 0 & 1 & b \\ 0 & 1 & 0 & c \\ 2 & 1 & 0 & a \end{array}\right) \\ &\sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & d \\ 0 & 1 & 1 & b-d \\ 0 & 1 & 0 & c \\ 0 & 3 & 0 & a-2d \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & d \\ 0 & 1 & 1 & b-d \\ 0 & 1 & 0 & c \\ 0 & 0 & 0 & a-3c-2d \end{array}\right). \end{aligned}$$

Por lo tanto, para que el sistema (4.29) tenga solución, se debe satisfacer la igualdad:

$$a - 3c - 2d = 0.$$

Entonces:

$$\langle B_2 \rangle = \{a + bx + cx^2 + dx^3 \in \mathscr{P}_3[x] : a - 3c - 2d = 0\}.$$

El polinomio x^2 no es elemento de $\langle B_2 \rangle$, pues, como $a = 0$, $b = 0$, $c = 1$ y $d = 0$, entonces $a - 3c - 2d = -3 \neq 0$. Entonces

$$B' = B_1 \cup \{x^2\}$$

es una base de $\mathscr{P}_3[x]$ y $W = \langle B' - B_1 \rangle = \left\langle x, x^2 \right\rangle$. Por lo tanto:

$$W = \{ax + bx^2 : a \in \mathbb{R},\ b \in \mathbb{R}\}.$$

♦

Ejemplo 4.59

Sean

$$W_1 = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\} \text{ y } W_2 = \{(x, y, z) \in \mathbb{R}^3 : 2x + y = 0,\ y - z = 0\}$$

dos espacios vectoriales de $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$. Calcular $W_1 + W_2$ y probar que:

$$\mathbb{R}^3 = W_1 \oplus W_2.$$

Solución. Por el teorema 4.21, sabemos que $W_1 + W_2 = \langle B_1 \cup B_2 \rangle$, donde B_1 y B_2 son bases de W_1 y W_2 respectivamente. Por lo tanto, para calcular la suma de W_1 y W_2, encontremos una base para cada uno.

1. *Una base para* W_1:

$$\begin{aligned} W_1 &= \{(x,y,z) \in \mathbb{R}^3 : x+y+z=0\} \\ &= \{(-y-z,y,z) \in \mathbb{R}^3 : y \in \mathbb{R},\ z \in \mathbb{R}\} \\ &= \{y(-1,1,0)+z(-1,0,1) \in \mathbb{R}^3 : y \in \mathbb{R},\ z \in \mathbb{R}\}. \end{aligned}$$

   Por lo tanto, el conjunto $B_1 = \{(-1,1,0),\ (-1,0,1)\}$ genera W_1.

   Veamos si B_1 es linealmente independiente. Para ello, sean α y β escalares tales que:

$$\alpha(-1,1,0)+\beta(-1,0,1)=(0,0,0).$$

   De esta igualdad, obtenemos que:

$$-\alpha-\beta=0,\ \alpha=0,\ \beta=0,$$

   de donde $\alpha=\beta=0$ es la única solución. Por lo tanto, B_1 es linealmente independiente, y, por lo tanto, es una base de W_1.

2. *Una base para* W_2:

$$\begin{aligned} W_2 &= \{(x,y,z) \in \mathbb{R}^3 : 2x+y=0,\ y-z=0\} \\ &= \{(x,y,z) \in \mathbb{R}^3 : y=-2x,\ z=y=-2x\} \\ &= \{(x,-2x,-2x) \in \mathbb{R}^3 : x \in \mathbb{R}\} \\ &= \{x(1,-2,-2) \in \mathbb{R}^3 : x \in \mathbb{R}\}. \end{aligned}$$

   Por lo tanto, el conjunto $B_2 = \{(1,-2,-2)\}$ genera W_2. Puesto que también es linealmente independiente, pues tiene un solo elemento y ese elemento es diferente de 0_v[2], B_2 es una base de W_2.

3. *Determinación de* W_1+W_2. Recordemos que: $W_1+W_2=\langle B_1\cup B_2\rangle$. Vamos a probar que $B_1\cup B_2$ es linealmente independiente. Como tiene tres elementos, entonces va a generar $\mathbb{R}^3$, es decir, $W_1+W_2=\mathbb{R}^3$. Pero esto implica, por el teorema de la dimensión, que $\dim(W_1\cap W_2)=0$, de donde $W_1\cap W_2=\{0_v\}$. Es decir, $W_1\oplus W_2=\mathbb{R}^3$.

4. *El conjunto* $B_1\cup B_2$ *es linealmente independiente*. En efecto, sean α, β y γ escalares tales que

$$\alpha(-1,1,0)+\beta(-1,0,1)+\gamma(1,-2,-2)=(0,0,0).$$

   Entonces los escalares satisfacen el siguiente sistema de ecuaciones lineales

$$\begin{aligned} -\alpha-\beta+\ \gamma&=0 \\ \alpha\qquad-2\gamma&=0 \\ \beta-2\gamma&=0 \end{aligned}$$

   que tiene solución única, la trivial, pues la matriz de coeficientes es equivalente a una triangular superior cuyo rango es igual a 3:

$$\left(\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 1 & 0 & -2 & 0 \\ 0 & 1 & -2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 1 & -2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -3 & 0 \end{array}\right).$$

   Por lo tanto, $B_1\cup B_2$ es un conjunto linealmente independiente.

♦

[2] Si $v\neq 0_v$, entonces la igualdad $\alpha v=0_v$ implica que $\alpha=0$ (ver el teorema 4.3). Es decir, el conjunto $\{v\}$ es linealmente independiente.

Ejemplo 4.60

Sean W_1, W_2 y W_3 tres subespacios vectoriales de un espacio vectorial V de dimensión finita tales que:

1. $W_1 \subset W_2$;
2. $W_1 + W_3 = W_2 + W_3$; y
3. $W_1 \cap W_3 = W_2 \cap W_3$.

Entonces $W_1 = W_2$.

Solución. Del teorema de la dimensión, podemos afirmar que

$$\dim(W_1 \cap W_3) = \dim(W_1) + \dim(W_3) - \dim(W_1 + W_3)$$

y que

$$\dim(W_2 \cap W_3) = \dim(W_2) + \dim(W_3) - \dim(W_2 + W_3).$$

La segunda hipótesis implica que

$$\dim(W_1 + W_3) = \dim(W_2 + W_3)$$

y la tercera,

$$\dim(W_1 \cap W_3) = \dim(W_2 \cap W_3).$$

Por lo tanto:

$$\begin{aligned} \dim(W_1) + \dim(W_3) - \dim(W_1 + W_3) &= \dim(W_2) + \dim(W_3) - \dim(W_2 + W_3) \\ \dim(W_1) + \dim(W_3) &= \dim(W_2) + \dim(W_3) \\ \dim(W_1) &= \dim(W_2). \end{aligned}$$

De la primera hipótesis, W_1 es un subespacio vectorial de W_2 y una base de W_1 tiene el mismo número de elementos que una base de W_2. Entonces la base de W_1 genera W_2. Esto significa que $W_1 = W_2$. ♦

4.12 Producto interno

Definición 4.13 (Producto interno) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y K el campo de los números reales o complejos. Un *producto interno* sobre V es una función $\langle \cdot\, , \cdot \rangle : V \times V \longrightarrow \mathbb{K}$ tal que:

1. **Linealidad**: $(\forall u \in V)(\forall v \in V)(\forall w \in V)(\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle)$.
2. **Homogeneidad**: $(\forall \lambda \in \mathbb{K})(\forall u \in V)(\forall v \in V)(\langle \lambda u, v\rangle = \lambda \langle u, v\rangle)$.
3. **Simetría de Herminte**: $(\forall u \in V)(\forall v \in V)(\langle u, v\rangle = \overline{\langle v, u\rangle})$.
4. **Positividad**: $(\forall v \in V - \{0_v\})(\langle v, v\rangle > 0)$.

Observaciones

1. Por la simetría de Hermite, tenemos que $\langle u, u\rangle = \overline{\langle u, u\rangle}$; es decir, el conjugado de $\langle u, u\rangle$ es el mismo. Esto significa que $\langle u, u\rangle$ siempre es un número real, y es positivo cuando $u \neq 0_v$.

2. Si $u = 0_v$, entonces $\langle u, v\rangle = 0$, pues $u = 0 \cdot u$, de donde, por la homogeneidad, se tiene que $\langle u, v\rangle = \langle 0 \cdot u, v\rangle = 0\,\langle u, v\rangle = 0$.

3. De la observación anterior, si $u = 0_v$, entonces $\langle u, u\rangle = 0$. Recíprocamente, si $\langle u, u\rangle = 0$, entonces $u = 0_v$, pues, si $u \neq 0_v$, por la positividad, tendríamos que $\langle u, u\rangle > 0$.

4. A la estructura de espacio vectorial junto con un producto interno se le denomina *espacio con producto interior*.

A continuación, presentamos los productos internos más usuales para los espacios que aparecen frecuentemente en este texto.

1. En el espacio vectorial $(\mathbb{R}^2, \mathbb{K}, +, \cdot)$:

$$\langle u, v\rangle = x_1y_1 + x_2y_2,$$

donde $u = (x_1, x_2)$ y $v = (y_1, y_2)$.

2. En el espacio vectorial $(\mathbb{R}^3, \mathbb{K}, +, \cdot)$:

$$\langle u, v\rangle = x_1y_1 + x_2y_2 + x_3y_3,$$

donde $u = (x_1, x_2, x_3)$ y $v = (y_1, y_2, y_3)$.

3. En el espacio vectorial $(\mathbb{R}^n, \mathbb{K}, +, \cdot)$:

$$\langle u, v\rangle = \sum_{j=1}^{n} x_iy_i = x_1y_1 + x_2y_2 + \cdots + x_ny_n,$$

donde $u = (x_1, x_2, \cdots, x_n)$ y $v = (y_1, y_2, y_3)$.

4. En el espacio vectorial $(\mathbb{C}^n, \mathbb{C}, +, \cdot)$:

$$\langle u, v\rangle = \sum_{j=1}^{n} x_i\overline{y_i},$$

donde $u = (x_1, x_2, \cdots, x_n)$ y $v = (y_1, y_2, y_3)$.

5. En el espacio vectorial $(\mathscr{P}_n[x], \mathbb{R}, +, \cdot)$:

$$\langle p, q\rangle = \sum_{j=1}^{n} a_jb_j,$$

donde $p(x) = \sum_{j=1}^{n} a_jx^j$ y $q(x) = \sum_{j=1}^{n} b_jx^j$.

6. En el espacio vectorial $(\mathbb{M}_{n\times n}, \mathbb{R}, +, \cdot)$:

$$\langle A, B\rangle = \operatorname{tr}(AB^t) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij},$$

donde $A = (a_{ij})$ y $B = (b_{ij})$.

Sobre un conjunto, se pueden definir varios productos internos. En el siguiente ejemplo, se ilustra esta situación.

Ejemplo 4.61

Definimos la función $\langle \cdot\ , \cdot\rangle$ en $\mathscr{P}_2[x] \times \mathscr{P}_2[x]$ por:

$$\langle p, q\rangle = \sum_{j\in J} p(j)q(j),$$

donde $J = \{-1, 0, 1\}$. Probar que esta función es un producto interno sobre $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$.

Solución.

1. **Linealidad**. Sean p, q y r en $\mathscr{P}_2[x]$. Debemos probar que

$$\langle p+q, r\rangle = \langle p, r\rangle + \langle q, r\rangle .$$

Demostración.

$$\begin{aligned}\langle p+q, r\rangle &= \sum_{j\in J}(p+q)(j)r(j)\\ &= \sum_{j\in J}[(p(j)+q(j)]r(j)\\ &= \sum_{j\in J}[(p(j)r(j)+q(j)r(j)]\\ &= \sum_{j\in J}(p(j)r(j) + \sum_{j\in J} q(j)r(j) = \langle p, r\rangle + \langle q, r\rangle .\end{aligned}$$

$\square$

2. **Homogeneidad**. Sean $\lambda \in \mathbb{R}$ y p y q en $\mathscr{P}_2[x]$. Vamos a probar que

$$\langle \lambda p, q\rangle = \lambda \langle p, q\rangle .$$

Demostración.

$$\begin{aligned}\langle \lambda p, q\rangle &= \sum_{j\in J}[\lambda p(j)]q(j)\\ &= \sum_{j\in J}\lambda[p(j)q(j)]\\ &= \lambda\sum_{j\in J} p(j)q(j) = \lambda\langle p, q\rangle .\end{aligned}$$

$\square$

3. **Simetría de Hermite**. Como $\mathbb{K} = \mathbb{R}$, entonces el conjugado de cualquier elemento de $\mathbb{K}$ es él mismo. Por ello, lo que debemos probar es:

$$\langle p, q \rangle = \langle q, p \rangle$$

para todo p y q en $\mathscr{P}_2[x]$.

Demostración.

$$\begin{aligned}\langle p, q \rangle &= \sum_{j \in J} p(j)q(j) \\ &= \sum_{j \in J} q(j)p(j) = \langle q, p \rangle .\end{aligned}$$

□

4. **Positividad**. Sea $p \in \mathscr{P}_2[x] - \{0_v\}$. Vamos a demostrar que:

$$\langle p, p \rangle > 0.$$

Demostración. Tenemos que

$$\langle p, p \rangle = \sum_{j \in J} p^2(j).$$

Ahora bien, como p no es nulo y su grado es a lo más 2, entonces no puede tener más de dos raíces. Esto significa que -1, 0 y 1 no pueden ser, los tres, raíces de p. En otras palabras, al menos uno de los tres números $p(-1)$, $p(0)$ y $p(1)$ es distinto de 0.; es decir, existe $k \in J$ tal que $p(k) \neq 0$, por lo que $p^2(k) > 0$.

Por otro lado, para todo $j \in J - \{k\}$, tenemos que $p^2(j) \geq 0$. Entonces:

$$\langle p, p \rangle = p^2(k) + \sum_{j \in J - \{k\}} p^2(j) > 0.$$

□

En resumen, la función $\langle \cdot \, , \cdot \rangle$ es un producto interno sobre $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$. ♦

Teorema 4.26 (Propiedades del producto interno) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $\langle \cdot \, , \cdot \rangle$ un producto interno. Entonces:

1. $(\forall u \in V)(\forall v \in V)(\forall w \in V)(\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle)$.
2. $(\forall \lambda \in \mathbb{K})(\forall u \in V)(\forall v \in V)(\langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle)$.
3. $(\forall u \in V)(\langle 0_v, u \rangle = 0)$.
4. Si $(\forall u \in V)(\langle u, w \rangle = 0)$, entonces $w = 0_v$.
5. Si $\langle v, v \rangle = 0$, entonces $v = 0_v$.

Demostración. Sean u, v, w elementos de V y λ de $\mathbb{K}$.

1.

$$\begin{aligned}
\langle u, v+w\rangle &= \overline{\langle v+w, u\rangle} && \text{Simetría de Hermite.}\\
&= \overline{\langle v, u\rangle + \langle w, u\rangle} && \text{Linealidad.}\\
&= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} && \text{El conjugado de una suma es la suma de los conjugados.}\\
&= \langle u, v\rangle + \langle u, w\rangle && \text{Simetría de Hermite.}
\end{aligned}$$

2.

$$\begin{aligned}
\langle u, \lambda v\rangle &= \overline{\langle \lambda v, u\rangle} && \text{Simetría de Hermite.}\\
&= \overline{\lambda \langle v, u\rangle} && \text{Homogeneidad.}\\
&= \overline{\lambda}\ \overline{\langle v, u\rangle} && \text{El conjugado de un producto es el producto de los conjugados de cada factor.}\\
&= \overline{\lambda} \langle u, v\rangle && \text{Simetría de Hermite.}
\end{aligned}$$

3. Se demostró en la segunda observación a la definición de producto interno.

4. La demostración se deja para el lector.

5. Se demostró en la tercera observación a la definición de producto interno.

□

Teorema 4.27 (Desigualdad de Cauchy-Schwarz) Sea $\langle \cdot\,,\cdot\rangle$ un producto escalar definido sobre un espacio vectorial $(V, \mathbb{K}, +, \cdot)$. Entonces se satisface la desigualdad

$$|\langle u, v\rangle| \le \sqrt{\langle u, u\rangle}\sqrt{\langle v, v\rangle}$$

para todo $u \in V$ y todo $v \in V$.

Aunque el teorema es verdadero para cualquier campo $\mathbb{K}$, la demostración que presentamos a continuación se la realiza para el caso $\mathbb{K} = \mathbb{R}$. Esto significa que $\langle u, v\rangle = \langle v, u\rangle$ y no aparecerá el conjugado de un número en ninguna expresión.

Demostración. Si $u = 0_v$, entonces, por un lado:

$$|\langle u, v\rangle| = |0| = 0.$$

Por otro lado:

$$\sqrt{\langle u, u\rangle}\sqrt{\langle v, v\rangle} = 0 \times \sqrt{\langle v, v\rangle} = 0.$$

Por lo tanto:

$$|\langle u, v\rangle| = \sqrt{\langle u, u\rangle}\sqrt{\langle v, v\rangle}$$

cuando $u = 0_v$.

Supongamos ahora que $u \neq 0_v$. Definamos w así:

$$w = v - \frac{\langle u, v\rangle}{\langle u, u\rangle} u.$$

Entonces:

$$
\begin{aligned}
\langle w, w\rangle &= \left\langle v - \frac{\langle u, v\rangle}{\langle u, u\rangle}u, v - \frac{\langle u, v\rangle}{\langle u, u\rangle}u\right\rangle \\
&= \left\langle v, v - \frac{\langle u, v\rangle}{\langle u, u\rangle}u\right\rangle - \left\langle \frac{\langle u, v\rangle}{\langle u, u\rangle}u, v - \frac{\langle u, v\rangle}{\langle u, u\rangle}u\right\rangle \\
&= \langle v, v\rangle - \frac{\langle u, v\rangle}{\langle u, u\rangle}\langle v, u\rangle - \frac{\langle u, v\rangle}{\langle u, u\rangle}\langle u, v\rangle + \frac{\langle u, v\rangle}{\langle u, u\rangle}\frac{\langle u, v\rangle}{\langle u, u\rangle}\langle u, u\rangle \\
&= \langle v, v\rangle - \frac{\langle u, v\rangle}{\langle u, u\rangle}\langle u, v\rangle - \frac{\langle u, v\rangle}{\langle u, u\rangle}\langle u, v\rangle + \frac{\langle u, v\rangle}{\langle u, u\rangle}\langle u, v\rangle \\
&= \langle v, v\rangle - \frac{\langle u, v\rangle}{\langle u, u\rangle}\langle u, v\rangle \\
\langle w, w\rangle &= \frac{\langle v, v\rangle\langle u, u\rangle - \langle u, v\rangle^2}{\langle u, u\rangle}.
\end{aligned}
$$

Puesto que $\langle w, w\rangle \geq 0$, entonces

$$\frac{\langle v, v\rangle\langle u, u\rangle - \langle u, v\rangle^2}{\langle u, u\rangle} \geq 0,$$

y, como $\langle u, u\rangle > 0$, obtenemos que

$$\langle v, v\rangle\langle u, u\rangle - \langle u, v\rangle^2 \geq 0,$$

de donde se verifica que

$$\langle u, v\rangle^2 \leq \langle v, v\rangle\langle u, u\rangle .$$

Por lo tanto:

$$|\langle u, v\rangle| \leq \sqrt{\langle u, u\rangle}\sqrt{\langle v, v\rangle}.$$

□

Definición 4.14 (Norma) Sea $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial. Una *norma sobre V* es una función $\|\cdot\| : V \longrightarrow \mathbb{R}$ que satisface las siguientes propiedades:

1. **Positividad**: $(\forall v \in V - \{0_v\})(\|v\| > 0)$.
2. **Homogeneidad**: $(\forall \lambda \in \mathbb{K})(\forall v \in V)(\|\lambda v\| = |\lambda|\,\|v\|)$.
3. **Desigualdad triangular**: $(\forall u \in V)(\forall v \in V)(\|u + v\| \leq \|u\| + \|v\|)$.

A un espacio vectorial junto con una norma definida sobre él se le conoce con el nombre de *espacio normado*.

La desigualdad de Cauchy-Schwarz puede ser reformulada de la siguiente manera mediante la norma:

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|$$

para todo u y todo v en un espacio normado V.

Teorema 4.28 (Norma inducida) Sea $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial sobre el que se ha definido un producto interno $\langle \cdot\ , \cdot \rangle$. Entonces la función de $\|\cdot\| : V \longrightarrow \mathbb{R}$ definida por

$$\|v\| = \sqrt{\langle v, v\rangle}$$

para todo $v \in V$ es una norma sobre V.
A esta norma se le conoce con el nombre de *norma inducida* sobre V por el producto interno sobre V.

Demostración. Sean u y v elementos de V y λ un escalar.

1. **Positividad**. Supongamos que $v \neq 0_v$. Entonces, por la positividad del producto interno, tenemos que $\langle v, v\rangle > 0$. Por lo tanto, $\|v\|^2 > 0$, de donde $\|v\| > 0$.

2. **Homogeneidad**.

$$\begin{aligned} \|\lambda v\|^2 &= \langle \lambda v, \lambda v\rangle \\ &= \lambda\overline{\lambda}\,\langle v, v\rangle \\ &= |\lambda|^2\,\|v\|^2 . \end{aligned}$$

   Por lo tanto, $\|v\| = |\lambda|\,\|v\|$.

3. **Desigualdad triangular**. La demostración de esta propiedad se realiza únicamente para el caso en que $\mathbb{K} = \mathbb{R}$.

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \|u\|^2 + \langle u, v\rangle + \langle u, v\rangle + \|v\|^2 \\ &= \|u\|^2 + 2\,\langle u, v\rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\,\langle u, v\rangle\,| + \|v\|^2 \\ &\leq \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2. \end{aligned}$$

   Por lo tanto, $\|u+v\| \leq \|u\| + \|v\|$.

□

Ejemplo 4.62 En el espacio $(\mathbb{R}^4, \mathbb{R}, +, \cdot)$, con el producto interno usual, tenemos que

$$\begin{aligned} \|(4,-2,0,3)\|^2 &= \langle (4,-2,0,3), (4,-2,0,3)\rangle \\ &= 4*4 + (-2)(-2) + (0)(0) + (3)(3) \\ &= 16 + 4 + 9. \end{aligned}$$

Por lo tanto, $\|(4,-2,0,3)\| = \sqrt{29}$. ♦

4.13 Ortogonalidad

Definición 4.15 (Vectores ortogonales) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $\langle \cdot\,, \cdot \rangle$ un producto interno definido sobre V.

1. Dos vectores u y v de V son *ortogonales* si y solamente si $\langle u, v\rangle = 0$.
2. Un conjunto S de V es *ortogonal* si y solamente si todo par de elementos distintos de S son ortogonales.

Observaciones

1. El vector nulo 0_v es ortogonal a cualquier vector, pues $\langle 0_v, v\rangle = 0$ para todo $v \in V$.
2. Si $S \subset V$ es distinto de $\{0_v\}$, para que sea un conjunto ortogonal, debe tener al menos dos elementos.

Ejemplo 4.63 En el espacio $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$, con el producto interno usual, el conjunto

$$S = \{(1, 0, -1),\ (2, 1, 2),\ (-1, 4, -1)\}$$

es ortogonal pues:

$$\begin{aligned} \langle (1, 0, -1), (2, 1, 2)\rangle &= 2 + 0 - 2 = 0; \\ \langle (1, 0, -1), (-1, 4, -1)\rangle &= -1 + 0 + 1 = 0; \\ \langle (2, 1, 2), (-1, 4, -1)\rangle &= -2 + 4 - 2 = 0. \end{aligned}$$

En otras palabras, hemos mostrado que $\langle u, v\rangle = 0$ para todo $u \in S$ y todo $v \in S$ tal que $u \neq v$. ♦

Teorema 4.29 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $\langle \cdot\,, \cdot \rangle$ un producto interno sobre V y $S \subset V$ tal que $0_v \notin S$. Si S es un conjunto ortogonal, entonces S es linealmente independiente.

Demostración. Supongamos que $S = \{v_i : i \in I_n = \{1, 2, \ldots, n\}\}$. Sean α_i con $i \in I_n$ escalares tales que:

$$\sum_{i=1}^{n} \alpha_i v_i = 0_v. \tag{4.30}$$

Sea $j \in I_n$. Entonces:

$$\begin{aligned} 0 &= \langle 0_v, v_j\rangle \\ &= \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_j \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i \langle v_i, v_j\rangle \end{aligned}$$

$$= \alpha_j \langle v_j, v_j \rangle ,$$

pues, por ser S ortogonal, $\langle v_i, v_j \rangle = 0$ para todo $i \in I_m$ tal que $i \neq j$. Por lo tanto:

$$\alpha_j \langle v_j, v_j \rangle = 0. \tag{4.31}$$

Puesto que $v_j \neq 0_v$ (ya que $0_v \notin S$), entonces $\langle v_j, v_j \rangle > 0$. Esto significa que la igualdad (4.31) implica que $\alpha_j = 0$.

En resumen, hemos probado que la igualdad (4.30) implica que $\alpha_j = 0$ para todo $j \in I_n$; es decir, hemos probado que S es un conjunto linealmente independiente. □

Ejemplo 4.64 En el espacio $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$, el conjunto

$$s = \{(1, 0, -1),\ (2, 1, 2),\ (-1, 4, -1)\}$$

es una base, ya que es linealmente independiente, pues es ortogonal —como se probó en el ejemplo anterior.

Recuérdese que es suficiente probar que S es linealmente independiente para que sea una base de $\mathbb{R}^3$, ya que S tiene tres elementos y la dimensión de $\mathbb{R}^3$ es 3. ♦

Teorema 4.30 (Proceso de Gram-Schmidt) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $\langle \cdot\, , \cdot \rangle$ un producto interno sobre V. Si el conjunto $S_1 = \{v_i \in V : i \in I_n = \{1, 2, \ldots, n\}\}$ es linealmente independiente, entonces existe un conjunto ortogonal $S_2 = \{w_i : i \in I_n\}$, donde:

1. $w_1 = v_1$; y
2. $w_i = v_i - \displaystyle\sum_{j=1}^{i-1} \frac{\langle v_i, w_j \rangle}{\langle w_j, w_j \rangle} w_j$ para cada $i \in I_n - \{1\}$;

y tal que $\langle S_2 \rangle = \langle S_1 \rangle$.

Definición 4.16 (Base ortogonal) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y $\langle \cdot\, , \cdot \rangle$ un producto interno sobre V. Una base B de V es ortogonal su B es un conjunto ortogonal.

Ejemplo 4.65 En $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$, el conjunto

$$B = \{(1, 0, -1),\ (2, 1, 2),\ (-1, 4, -1)\}$$

es una base ortogonal. ♦

Teorema 4.31 Todo espacio vectorial de dimension finita mayor que o igual a dos sobre el cual se ha definido un producto interno tiene una base ortogonal.

Demostración. Sea B_1 una base de V, un espacio vectorial tal que $\dim(V) \geq 2$. Entonces B_1 tiene al menos dos elementos. Si se aplica el proceso de Gram-Schmidt, a partir de B_1 se obtiene un conjunto ortogonal B con el mismo número de elementos que B_1 y tal que $\langle B \rangle = \langle B_1 \rangle$. Como B_1 genera V, entonces B también genera V. Entonces B es una base pues tiene el mismo número de elementos que la dimensión de V. En resumen, B es una base ortogonal de V. $\square$

Teorema 4.32 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $\langle \cdot\,, \cdot \rangle$ un producto interno sobre V y $S \subset V$. Si $S = \{v_i : i \in I_n\}$ es ortogonal, entonces el conjunto

$$S' = \{\lambda_i v_i : i \in I_n\}$$

es ortogonal, donde $\lambda_i \neq 0$ para todo $i \in I_n$.

Demostración. Sean $i \in I_n$ y $j \in I_n$ tales $i \neq j$. Vamos a probar que $\lambda_i v_i$ y $\lambda_j v_j$ son ortogonales:

$$\begin{aligned} \langle \lambda_i v_i, \lambda_j v_j \rangle &= \lambda_i \overline{\lambda_j}\, \langle v_i, v_j \rangle \\ &= \lambda_i \overline{\lambda_j} \times 0 = 0, \end{aligned}$$

pues S es un conjunto ortogonal. $\square$

Definición 4.17 (Conjunto ortogonal) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $\langle \cdot\,, \cdot \rangle$ un producto interno sobre V y $S \subset V$. El conjunto S es *ortonormal* si es ortogonal y la norma de cada uno de sus elementos es igual a 1. En particular, una base es ortonormal si es un conjunto ortonormal.

Corolario 4.33 Todo espacio vectorial de dimensión finita mayor que o igual a dos sobre el cual se ha definido un producto interno tiene una base ortonormal.

Demostración. Sea B' una base ortogonal para el espacio vectorial, que existe por el teorema 4.31. Definamos entonces B de la siguiente manera:

$$B = \left\{ \frac{1}{\|v\|} v : v \in B' \right\}.$$

Entonces B es ortogonal, ya que B' lo es (teorema 4.32). Además, cada elemento de B tiene norma igual a 1. En efecto:

$$\left\| \frac{1}{\|v\|} v \right\| = \frac{1}{\|v\|} \|v\| = \frac{\|v\|}{\|v\|} = 1.$$

Finalmente, B es linealmente independiente, ya que es ortogonal (teorema 4.29). Por lo tanto, B es una base ortonormal de V. $\square$

Corolario 4.34 Si $B = \{v_i : i \in I_n\}$ es una base ortogonal de un espacio vectorial V, entonces cada elemento u de V se escribe en la forma

$$u = \sum_{i=1}^{n} \frac{\langle u, v_i \rangle}{\|v_i\|^2} v_i.$$

Demostración. Dado que B es una base de V, cada u de V es combinación lineal única de elementos de B. Esto significa que existen n escalares únicos α_i con $i \in I_n$ tales que

$$u = \sum_{i=1}^{n} \alpha_i v_i.$$

Sea $j \in I_n$. Entonces:

$$\begin{aligned} \langle u, v_j \rangle &= \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_j \right\rangle \\ &= \sum_{i=1}^{n} \langle \alpha_i v_i, v_j \rangle \\ &= \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle \\ &= \alpha_j \langle v_j, v_j \rangle, \end{aligned}$$

pues B es un conjunto ortogonal, por lo que $\langle v_i, v_j \rangle = 0$ para todo $i \neq j$. Por lo tanto:

$$\langle u, v_j \rangle = \alpha_j \langle v_j, v_j \rangle = \alpha_j \|v_j\|^2.$$

para cada $j \in I_n$.

Ahora bien, como B es una base, ningún v_j puede ser igual a 0_v. Por lo tanto, tenemos que $\|v_j\| \neq 0$. Entonces, tenemos que las coordenadas de u respecto de B están dadas por:

$$\alpha_j = \frac{\langle u, v_j \rangle}{\|v_j\|^2}$$

para cada $j \in I_n$, lo que implica que:

$$u = \sum_{i=1}^{n} \frac{\langle u, v_i \rangle}{\|v_i\|^2} v_i.$$

□

Ejemplo 4.66

Sea

$$W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a + d = b + c \right\}$$

un subespacio vectorial de $(\mathbb{M}_{2\times 2}, \mathbb{R}, +, \cdot)$ dotado con el producto interno usual. Encontrar una base ortonormal de W.

Solución. En primer lugar, vamos a encontrar una base para W. A continuación, utilizaremos el proceso de Gram-Schmidt para obtener, a partir de dicha base, una ortonormal.

1. Puesto que:

$$\begin{aligned} W &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a = b + c - d \right\} \\ &= \left\{ \begin{pmatrix} b+c-d & b \\ c & d \end{pmatrix} : b \in \mathbb{R},\ c \in \mathbb{R},\ d \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} b & b \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} c & 0 \\ c & 0 \end{pmatrix} + \begin{pmatrix} -d & 0 \\ 0 & d \end{pmatrix} : b \in \mathbb{R},\ c \in \mathbb{R},\ d \in \mathbb{R} \right\} \\ &= \left\{ b\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} : b \in \mathbb{R},\ c \in \mathbb{R},\ d \in \mathbb{R} \right\}, \end{aligned}$$

   el conjunto

$$B_1 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

   genera el subespacio W. Es fácil probar que B_1 es linealmente independiente, por lo que se deja esta tarea al lector. Por lo tanto B_1 es una base de W.

2. Ahora apliquemos el proceso de Gram-Schmidt (teorema 4.30) para obtener, a partir de B_1, una base ortonormal para W.

   Nombremos con v_1, v_2 y v_3 los elementos de B_1 (en el orden de izquierda a derecha en el que están escritos arriba). Sea $B_2 = \{w_1, w_2, w_3\}$ la base ortogonal producida por el método de Gram-Schmidt. Entonces:

$$w_1 = v_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

   El elemento w_2 se define así:

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1.$$

   Ahora bien, tenemos que:

   (a) $\langle v_2, w_1 \rangle = \operatorname{tr}(w_2 w_1^t) = \operatorname{tr}\left[\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^t\right] = \operatorname{tr}\left[\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\right] = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = 1.$

   (b) $\langle w_1, w_1 \rangle = \operatorname{tr}(w_1 w_1^t) = \operatorname{tr}\left[\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^t\right] = \operatorname{tr}\left[\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\right] = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} = 2.$

   Entonces:

$$w_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 \end{pmatrix}.$$

   Por último, el elemento w_3 se define de la siguiente manera:

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2.$$

   Para encontrarlo, realicemos los siguientes cálculos:

   (a) $\langle v_3, w_1 \rangle = \operatorname{tr}(v_3 w_1^t) = \operatorname{tr}\left[\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\right] = \operatorname{tr}\begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} = -1.$

(b) $\langle v_3, w_2\rangle = \operatorname{tr}(v_3 w_2^t) = \operatorname{tr}\left[\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\frac{1}{2} & 1\\ -\frac{1}{2} & 0\end{pmatrix}\right] = \operatorname{tr}\begin{pmatrix}-\frac{1}{2} & -1\\ -\frac{1}{2} & 0\end{pmatrix} = -\dfrac{1}{2}.$

(c) $\langle w_2, w_2\rangle = \operatorname{tr}(v_2 w_2^t) = \operatorname{tr}\left[\begin{pmatrix}\frac{1}{2} & -\frac{1}{2}\\ 1 & 0\end{pmatrix}\begin{pmatrix}\frac{1}{2} & 1\\ -\frac{1}{2} & 0\end{pmatrix}\right] = \operatorname{tr}\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & 1\end{pmatrix} = \dfrac{3}{2}.$

Por lo tanto:

$$w_3 = \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix} + \frac{1}{3}\begin{pmatrix}\frac{1}{2} & -\frac{1}{2}\\ 1 & 0\end{pmatrix} = \begin{pmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & 1\end{pmatrix}.$$

Ya tenemos, entonces, una base ortogonal de W: el conjunto $B_2 = \{w_1, w_2, w_3\}$.

3. Nos falta únicamente obtener la base ortonormal. Eso lo hacemos a partir de B_2 al multiplicar a cada vector de B_2 por el recíproco de su norma. Sean e_i tales que:

$$e_i = \frac{1}{\|w_i\|} w_i.$$

Ya conocemos las normas de w_1 y w_2. Calculemos la de w_3:

$$\langle w_3, w_3\rangle = \operatorname{tr}\left[\begin{pmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & 1\end{pmatrix}\begin{pmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & 1\end{pmatrix}^t\right] = \operatorname{tr}\left[\begin{pmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & 1\end{pmatrix}\begin{pmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & 1\end{pmatrix}\right] = \operatorname{tr}\begin{pmatrix}\frac{2}{9} & \frac{2}{9}\\ \frac{2}{9} & \frac{10}{9}\end{pmatrix} = \frac{4}{3}.$$

Entonces, $B = \{e_1, e_2, e_3\}$ es la base ortonormal de W:

$$B = \left\{\sqrt{2}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}, \sqrt{\frac{2}{3}}\begin{pmatrix}\frac{1}{2} & -\frac{1}{2}\\ 1 & 0\end{pmatrix}, \frac{\sqrt{3}}{2}\begin{pmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & 1\end{pmatrix}\right\}$$

♦

Definición 4.18 (El ortogonal de un conjunto) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $\langle\cdot\,,\cdot\rangle$ un producto interno sobre V y W un subespacio vectorial de V. El *ortogonal* de W, representado con $W^\perp$, se define por:

$$W^\perp = \{u \in V : (\forall w \in W)(\langle u, w\rangle = 0)\}.$$

Es decir, el ortogonal del subespacio vectorial W está constituido por todos los vectores de V que son ortogonales a cada uno de los elementos de W.

Es sencillo probar que el ortogonal de un subespacio vectorial también es un subespacio vectorial. Así que esta demostración se deja como ejercicio para el lector. En el siguiente teorema, se enuncia esta propiedad y, además, se enuncia una caracterización del ortogonal de un subespacio: para saber si un vector es elemento del ortogonal, basta comprobar que ese elemento es ortogonal a cada uno de los elementos de la base del subespacio vectorial.

Precisemos estas ideas en el siguiente teorema.

Teorema 4.35 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $\langle\cdot\,,\cdot\rangle$ un producto interno sobre V y W un subespacio vectorial de V. Entonces:

1. $W^\top$ es un subespacio vectorial de W.

2. Si B es una base de W, entonces:

$$W^\top = \{u \in W : (\forall w \in B)(\langle u, w\rangle = 0)\}.$$

Demostración. Demostraremos únicamente el segundo punto. Es claro que

$$W^\top \subset \{u \in W : (\forall w \in B)(\langle u, w\rangle = 0)\}.$$

Por lo tanto, para establecer la igualdad, supongamos que u es tal que

$$\langle u, w\rangle = 0$$

para todo $w \in B = \{v_1, v_2, \ldots, v_n\}$, una base de W, y probemos que $u \in W^t$.

Sea $v \in W$. Existen, entonces, n escalares α_i con $i \in I_n$ tales que

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Por lo tanto:

$$\langle u, v\rangle = \left\langle u, \sum_{i=1}^{n} \alpha_i v_i \right\rangle = \sum_{i=1}^{n} \overline{\alpha_i}\,\langle u, v_i\rangle = 0,$$

pues $\langle u, v_i\rangle = 0$ para todo $i \in I_n$, ya que $v_i \in B$. Por lo tanto, $v \in W^t$, con lo que queda el teorema demostrado. ☐

Teorema 4.36 Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial de dimensión finita, $\langle \cdot\,, \cdot\rangle$ un producto interno sobre V y W un subespacio vectorial de V. Entonces $V = W \oplus W^\top$.

Demostración. En primer lugar, vamos a demostrar que $V = W + W^\top$y luego que $W \cap W^\top = \{0_v\}$.

1. Como V es de dimensión finita, W también lo es. Sea $B = \{e_i : i \in I_n\}$ una base ortonormal de V (teorema 4.31 y proceso de Gram-Schmidt).

   Dado $v \in V$, vamos a construir un $w \in W$ y un elemento $w' \in W^\top$ tal que $v = w + w'$. Está claro que, si encontráramos, por ejemplo $w \in W^\top$, entonces w' se definiría por $w' = v - w$.

   El $w \in W$ que buscamos, debe ser de la forma

   $$w = \sum_{i=1}^{n} \alpha_i e_i, \tag{4.32}$$

   ya que B es una base de W. Por lo tanto, todo el trabajo que vamos a realizar a continuación está dirigido a saber cómo son esos α_i.

   En primer lugar, $v - w$ debe ser un elemento de $W^\top$. Esto significa que

   $$\langle v - w, u\rangle = 0$$

para todo $u \in W$.

En particular, esta igualdad debe ser verdadera para cada elemento de B:

$$\langle v - w, e_j \rangle = 0 \tag{4.33}$$

para todo $j \in I_n$.

Entonces, a partir de las igualdades (4.32) y 4.33, tenemos las siguientes equivalencias:

$$\begin{aligned}
\langle v - w, e_j \rangle = 0 &\iff \left\langle v - \sum_{i=1}^{n} \alpha_i e_i, e_j \right\rangle = 0 \\
&\iff \langle v, e_j \rangle - \left\langle \sum_{i=1}^{n} \alpha_i e_i, e_j \right\rangle = 0 \\
&\iff \langle v, e_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i e_i, e_j \right\rangle \\
&\iff \langle v, e_j \rangle = \sum_{i=1}^{n} \alpha_i \langle e_i, e_j \rangle \\
&\iff \langle v, e_j \rangle = \alpha_j \langle e_j, e_j \rangle = \alpha_j.
\end{aligned}$$

La penúltima igualdad en la última equivalencia se verifica porque $\langle e_i, e_j \rangle = 0$ si $i \neq j$, ya que B es ortogonal. Y la última igualdad es verdadera ya que la base es ortonormal, por lo que la norma de cada vector e_j es igual a 1.

En resumen, tenemos que

$$\langle v - w, e_j \rangle = 0 \iff \alpha_j = \langle v, e_j \rangle .$$

Por lo tanto, si definimos w como

$$w = \sum_{j=1}^{n} \langle v, e_j \rangle e_j,$$

entonces $v - w$ es ortogonal a cada elemento de la base B, por lo que, por el teorema 4.35, tenemos que $v - w$ es ortogonal a todo elemento de W. Esto significa que $v - w \in W^\top$. Sea, entonces, $w' = v - w$. Entonces $w' \in W^\top$ y, por lo tanto, $v = w + w'$ con $w \in W$ y $w' \in W^\top$. En otras palabras, tenemos que $V = W + W^\top$.

2. Probemos ahora que $W \cap W^\top = \{0_v\}$. Para ello, sea $u \in W \cap W^\top$. Entonces $u \in W^\top$. Por lo tanto:

$$\langle u, v \rangle = 0$$

para todo $v \in W$. En particular, para $v = u$, ya que $u \in W$. Así, tenemos que:

$$\langle u, u \rangle = 0.$$

Pero esto implica que $u = 0_v$. Es decir, $W \cap W^\top = \{0_v\}$.

□

Ejemplo 4.67

Sea
$$W = \{p \in \mathscr{P}_2[x] : (\forall x \in \mathbb{R})[p(x) = p(x+1)]\}.$$
Entonces W es un subespacio vectorial de $(\mathscr{P}_2[x], \mathbb{R}, +, \cdot)$. La función $\langle \cdot\ , \cdot \rangle : \mathscr{P}_2[x] \times \mathscr{P}_2[x] \longrightarrow \mathbb{R}$ definida por
$$\langle p, q \rangle = \int_0^1 p(x)q(x)dx$$
es un producto interno sobre $\mathscr{P}_2[x]$.

1. Obtener una base para W.
2. Encontrar $W^\top$.
3. Ilustrar que $\mathscr{P}_2[x] = W \oplus W^\top$.

Solución.

1. $$\begin{aligned} W &= \{a + bx + cx^2 \in \mathscr{P}_2[x] : a + bx + cx^2 = a + b(x+1)^2 + c(x+1)^2\} \\ &= \{a + bx + cx^2 \in \mathscr{P}_2[x] : a + bx + cx^2 = (a+b+c) + (b+2c)x + cx^2\} \\ &= \{a + bx + cx^2 \in \mathscr{P}_2[x] : a = a+b+c,\ b = b+2c,\ c = c\} \\ &= \{a + bx + cx^2 \in \mathscr{P}_2[x] : b + c = 0,\ 2c = 0\} \\ &= \{a + bx + cx^2 \in \mathscr{P}_2[x] : a \in \mathbb{R},\ b = 0,\ c = 0\} \\ &= \{a \in \mathscr{P}_2[x] : a \in \mathbb{R}\}. \end{aligned}$$

   Por lo tanto, el conjunto $B = \{p\}$ donde $p(x) = 1$ para todo $x \in \mathbb{R}$ genera el subespacio W. Es, además, base de W porque es linealmente independiente (tiene un solo elemento diferente del vector nulo).

2. Por el teorema 4.35, tenemos que
   $$W^\top = \{q \in \mathscr{P}_2[x] : \langle q, p \rangle = 0\},$$
   donde $p(x) = 1$ para todo $x \in \mathbb{R}$. Entonces, si $q \in W^\top$, tenemos que:
   $$\langle q, p \rangle = \int_0^1 (a + bx + cx^2)(1)dx = 0,$$
   donde $q(x) = a + bx + cx^2$. Esto significa que:
   $$a + \frac{1}{2}b + \frac{1}{3}c = 0.$$
   En resumen:
   $$W^\top = \{a + bx + cx^2 \in \mathscr{P}_2[x] : 6a + 3b + 2c = 0\}.$$

3. En primer lugar, si $p \in W \cap W^\top$ tal que $p(x) = a + bx + cx^2$, tenemos que:
   $$6a + 3b + 2c = 0,\ b = 0,\ c = 0.$$
   Entonces $a = 0$. Es decir, $p = 0_v$. Por lo tanto, $W \cap W^\top = \{0_v\}$.

Por otro lado, como $W^\top$ puede expresarse de la siguiente manera:

$$\begin{aligned} W^\top &= \left\{a + bx + cx^2 \in \mathscr{P}_2[x] : a = -\frac{b}{2} - \frac{c}{3}\right\} \\ &= \left\{-\frac{b}{2} - \frac{c}{3} + bx + cx^2 \in \mathscr{P}_2[x] : b \in \mathbb{R},\ c \in \mathbb{R}\right\} \\ &= \left\{b\left(-\frac{1}{2} + x\right) + c\left(-\frac{1}{3} + x^2\right) \in \mathscr{P}_2[x] : b \in \mathbb{R},\ c \in \mathbb{R}\right\}, \end{aligned}$$

tenemos que el conjunto

$$\left\{-\frac{1}{2} + x, \frac{1}{3} + x^2\right\}$$

genera $W^\top$.

Además, es sencillo probar que este conjunto es linealmente independiente, por lo que también es una base de $W^\top$. Así que $\dim(W^\top) = 2$. Sabemos que $\dim(W) = 1$, de donde:

$$\dim(W + W^\top) = \dim(W) + \dim(W^\top) - \dim(W \cap W^\top) = 1 + 2 - 0 = 3,$$

lo que prueba que $W \oplus W^\top = \mathscr{P}_2[x]$.

♦

Definición 4.19 (Base ordenada) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial y B una base de V. Se dice que B es *ordenada* si sus elementos son considerados como una sucesión.

Al decir que los elementos de B son considerados como una sucesión, estamos suponiendo que existe una función $f\colon I_n \longrightarrow V$, donde n es el número de elementos de B, tal que a cada $j \in I_n$, $f(n)$ es un elemento de B. El orden que hay entre los elementos de I_n es el que impone el orden en B. Así, el 1 es el primer elemento de I_n, entonces, $v_1 = f(1)$ será el primer elemento de la base; $v_2 = f(2)$ será el segundo elemento de la base, pues 2 es el segundo elemento de I_n, etcétera.

La siguiente definición impone la noción de orden a las coordenadas de un vector respecto de una base. Así, podemos hablar de la *primera coordenada*, la *segunda coordenada*, etcétera, de un vector respecto de una base ordenada. Precisemos esta idea.

Definición 4.20 (Coordenadas de un vector respecto de una base) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial de dimensión n y $B = \{v_i : i \in I_n\}$ una base ordenada de V. Para cada $v \in V$, el elemento $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ en $\mathbb{K}^n$ tal que

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

se denomina *coordenadas* de v respecto de B. El escalar α_j es la j–ésima coordenada de v.

Dado el vector v, sus coordenadas respecto a la base ordenada B, se representa por $[v]_B$. Es decir, $[v]_B \in \mathbb{K}^n$.

Es común representar también las coordenadas de un vector v como una matriz de una sola columna. En efecto, si $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ son las coordenadas de v respecto de una base B, entonces las coordenadas también se representan por:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Ejemplo 4.68

Sea

$$B = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \right\}$$

una base ordenada del espacio vectorial $(\mathbb{M}_{2\times 2}, \mathbb{R}, +, \cdot)$. Encontrar las coordenadas del vector

$$v = \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix}$$

respecto de B.

Solución. Debemos encontrar el vector $(\alpha, \beta, \gamma, \delta) \in \mathbb{R}^4$ tal que

$$\begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix} + \delta \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Entonces, los escalares α, β, γ y δ satisfacen el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} \alpha \qquad\qquad - \delta &= 1 \\ -\beta \qquad + \delta &= 2 \\ \beta - \gamma \qquad &= -1 \\ \alpha \qquad + \gamma \qquad &= 2. \end{aligned} \tag{4.34}$$

Resolvamos este sistema:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & 0 & 2 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 1 & 2 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right)$$

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 1 & 2 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right).$$

Por lo tanto:

$$\alpha = 2,\ \beta = -1,\ \gamma = 0,\ \delta = 1.$$

Entonces, las coordenadas del v es el vector de $\mathbb{R}^4$: $(2, -1, 0, 1)$. ♦

Definición 4.21 (Ángulo entre dos vectores) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $u \in V$ y $v \in V$. El ángulo formado entre u y v es el número θ tal que $0 \leq \theta \leq \pi$ y

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\,\|v\|}.$$

Ejemplo 4.69 Sean $p(x) = x + x^2$ y $q(x) = -1 + x$ elementos de $\mathscr{P}_2[x]$. Entonces, el ángulo θ formado entre u y v cumple la siguiente igualdad:

$$\cos\theta = \frac{\langle x + x^2, -1 + x \rangle}{\|x + x^2\| \, \|-1 + x\|} == \frac{(0)(-1) + (1)(1) + (1)(0)}{\sqrt{0^2 + 1^2 + 1^2}\sqrt{(-1)^2 + 1^2 + 0^2}} = \frac{1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}.$$

Por lo tanto, $\theta = \frac{\pi}{3}$ (es decir, $30°$). ♦

Definición 4.22 (Proyección ortogonal) Sean $(V, \mathbb{K}, +, \cdot)$ un espacio vectorial, $u \in V$ y $v \in V$. La *proyección ortogonal* de u respecto a v es el vector αv donde

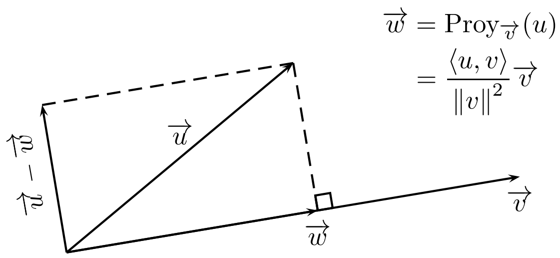
$$\alpha = \frac{\langle u, v \rangle}{\|v\|^2}.$$

La proyección αv es representado por $\mathrm{Proy}_v(u)$.

Nótese que el vector $u - \alpha v$ es ortogonal a v, pues

$$\langle u - \alpha v, v \rangle = \langle u, v \rangle - \alpha \langle v, v \rangle = \langle u, v \rangle - \frac{\langle u, v \rangle}{\langle v, v \rangle} \langle v, v \rangle = \langle u, v \rangle - \langle u, v \rangle = 0.$$

Esta situación se ilustra en el siguiente dibujo:



Ejemplo 4.70 En el espacio $(\mathbb{R}^3, \mathbb{R}, +, \cdot)$, el vector $u = (1, 0, 1)$ tiene como proyección respecto de $v = (-1, 1, 2)$ al vector:

$$\alpha v = \frac{\langle u, v \rangle}{\|v\|^2} v = \frac{-1 + 2}{1 + 1 + 4}(-1, 1, 2) = \left(-\frac{1}{6}, \frac{1}{6}, \frac{1}{3}\right).$$

♦

Aplicaciones lineales

por José Luis Toro

Edición y revisión: Juan Carlos Trujillo

Aplicaciones lineales 5

El objetivo general de este capítulo es:

> Resolver problemas sobre transformaciones lineales, definidas sobre un espacio vectorial de dimensión finita, utilizando la matriz asociada respecto de cualquier base, a un nivel reproductivo.

Para alcanzar el objetivo general, se proponen los siguientes objetivos específicos:

1. Identificar una aplicación lineal a partir de su definición y familiarizarse con sus propiedades.

2. Calcular bases para el núcleo y la imagen de una aplicación lineal usando la definición y las propiedades a un nivel reproductivo.

3. Resolver problemas de biyectividad a partir de las definiciones y propiedades de una aplicación lineal a un nivel reproductivo.

4. Resolver problemas con las operaciones entre las aplicaciones lineales y la inversa de una aplicación lineal usando las definiciones y sus propiedades a un nivel reproductivo.

5. Calcular la matriz asociada a una aplicación lineal respecto de dos bases, una del dominio y otra del espacio de llegada, respectivamente, donde las dimensiones de ambos espacios no sea mayor que 4, y utilizando los elementos esenciales que los tipifican a un nivel reproductivo.

6. Resolver problemas sobre las matrices asociadas a una aplicación lineal mediante las propiedades que la tipifican a un nivel reproductivo.

5.1 Introducción

Las funciones continuas son aquellas que conservan ciertas propiedades topológicas de los números reales. En los espacios vectoriales interesa, en cambio, conservar las estructuras algebraicas (operaciones como la suma y el producto por un escalar); es decir, se trata de que la aplicación sea tal que "conserve" las dos operaciones fundamentales que definen la estructura de espacio vectorial.

En el conjunto de todas las funciones, las **lineales** son las más importantes. Una buena parte de las Matemáticas está dedicada a resolver interrogantes relacionadas con las aplicaciones lineales. Por otra parte, éstas son de interés en sí mismas, y muchas aplicaciones importantes son lineales. Adicionalmente, es posible, a menudo, aproximar una aplicación arbitraria mediante una aplicación lineal, cuyo estudio es mucho más sencillo que el de la aplicación original.

Definición 5.1 (Aplicación lineal) Sean $(V, \mathbb{K}, +, \cdot)$ y $(W, \mathbb{K}, \oplus, \odot)$ dos espacios vectoriales (ambos están definidos sobre el mismo campo). Una función de V en W, $f\colon V \longrightarrow W$, es una **aplicación lineal** si y solo si para todo $\alpha \in \mathbb{K}$, $u \in V$ y $v \in V$ se verifica que:

1. *Conservación de* $+$: $f(u+v) = f(u) \oplus f(v)$.
2. *Conservación de* $\cdot$: $f(\alpha \cdot u) = \alpha \odot f(u)$.

Obsérvese que una aplicación es lineal cuando "conserva" las dos operaciones del espacio vectorial de salida.

En efecto, si u y v están en V, y $f(u)$ y $f(v)$ son sus imágenes en W bajo f, entonces, la primera condición de la definición asegura que la imagen de la suma de $u+v$ tiene que ser, necesariamente, la suma de las imágenes de u y de v.

Lo mismo sucede con el producto por un escalar: si $\alpha \in \mathbb{K}$, la segunda condición afirma que la imagen de αu bajo f debe ser el producto de α por la imagen de u bajo f. Éste es el significado de "conservación" de las dos operaciones.

Es fácil demostrar el siguiente teorema, por lo que se deja al lector su prueba, que permite simplificar la demostración de que una aplicación es lineal:

Teorema 5.1 Sean $(V, \mathbb{K}, +, \cdot)$ y $(W, \mathbb{K}, \oplus, \odot)$ dos espacios vectoriales sobre el mismo campo $\mathbb{K}$. Una función $f\colon V \longrightarrow W$ es lineal si y solo si

$$f(\alpha \cdot u + v) = \alpha \odot f(u) \oplus f(v)$$

para todo $u \in V$, $v \in V$ y $\alpha \in \mathbb{K}$.

Vamos a representar con $\mathscr{L}(V, W)$ al conjunto de todas las aplicaciones lineales de V en W. Es decir:

$$\mathscr{L}(V, W) = \{f\colon V \longrightarrow W : f \text{ es lineal.}\}$$

Un nombre alternativo para las aplicaciones lineales es el de **transformaciones lineales**. En este texto, utilizaremos preferentemente el primero. En la definición de aplicación lineal, cuando los espacios $V = W$, se suele utilizar el nombre **operador lineal sobre** V en lugar de aplicación lineal; y, cuando $W = \mathbb{K}$, el nombre que se utiliza es **forma lineal sobre** V.

De aquí en adelante, no diferenciaremos las operaciones de V y de W, pues, en el contexto, siempre se podrá distinguir si el uso de $+$ se refiere a la operación en V o en W. También omitiremos la mayoría de las veces el símbolo para el producto por un escalar; así, escribiremos solamente αu en lugar de $\alpha \cdot u$.

Los dos teoremas siguientes confirman el hecho de que una aplicación lineal conserva la estructura de espacio vectorial del espacio de salida. En efecto, el primer teorema muestra que una aplicación lineal hace corresponder el vector nulo del espacio de salida con el vector nulo del espacio de llegada.

Teorema 5.2 Para toda $f \in \mathscr{L}(V, W)$, se verifica que $f(0_V) = 0_W$.

Demostración. Dado que $0_V = 0 \cdot v$ para todo $v \in V$, entonces, ya que f es una aplicación lineal, tenemos que:

$$f(0_V) = f(0 \cdot v) = 0f(v) = 0_W.$$

□

Si no hay lugar a confusión, utilizaremos 0 para indicar tanto el elemento neutro de $\mathbb{K}$ como el vector nulo de cualquier espacio vectorial.

El siguiente teorema muestra que el “inverso aditivo” se conserva; es decir, el inverso aditivo de la imagen de un vector es la imagen del inverso aditivo del vector.

Teorema 5.3 Para todo $v \in V$ y toda $f \in \mathscr{L}(V, W)$, se verifica que $f(-v) = -f(v)$.

Demostración. f(-v) = f((-1)v) = (-1)f(v) = -f(v). □

De este teorema, es inmediato que si f es lineal, entonces $f(u - v) = f(u) - f(v)$.

5.1.1 Ejemplos

1. **Aplicación lineal nula**: es la función $\mathrm{O_f} \colon V \longrightarrow W$ definida por

   $$\mathrm{O_f}(v) = 0$$

   para todo $v \in V$. Es fácil ver que esta aplicación es lineal.

2. **Aplicación lineal identidad**: es la función $I \colon V \longrightarrow V$ definida por

   $$I(v) = v$$

   para todo $v \in V$. Es fácil constatar que esta aplicación es lineal.

3. **Reflexión con respecto al eje horizontal $\boldsymbol{x}$**: es la función $\rho \colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por[1]

   $$\rho(x, y) = (x, -y).$$

   Probemos que ρ es una aplicación lineal.

[1]Escribiremos $f(x, y)$ en lugar de $f((x, y))$ siempre que no haya lugar a confusión.

Demostración. Sean $u = (x_1, x_2)$ y $v = (y_1, y_2)$ elementos de $\mathbb{R}^2$ y $\alpha \in \mathbb{R}$. Entonces:

$$\begin{aligned}
\rho(\alpha u + v) &= \rho(\alpha(x_1, x_2) + (y_1, y_2)) && \text{definición de } u \text{ y } v,\\
&= \rho(\alpha x_1 + y_1, \alpha x_2 + y_2) && \text{suma en } \mathbb{R}^2,\\
&= (\alpha x_1 + y_1, -(\alpha x_2 + y_2)) && \text{definición de } \rho,\\
&= (\alpha x_1, -\alpha x_2) + (y_1, -y_2) && \text{suma en } \mathbb{R}^2,\\
&= \alpha(x_1, -x_2) + (y_1, -y_2) && \text{producto por un escalar en } \mathbb{R}^2,\\
&= \alpha\rho(x_1, x_2) + \rho(y_1, y_2) && \text{definición de } \rho,\\
&= \alpha\rho(u) + \rho(v) && \text{definición de } u \text{ y de } v.
\end{aligned}$$

□

4. La función $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ definida por $f(x) = x^2$ no es una aplicación lineal.

   Demostración. Sean $x \in \mathbb{R}$ y $y \in \mathbb{R}$. Por un lado, tenemos que

   $$f(x + y) = (x + y)^2 = x^2 + 2xy + y^2.$$

   Por otro lado, se cumple que

   $$f(x) + f(y) = x^2 + y^2.$$

   Entonces, $f(x + y) = f(x) + f(y)$ si y solo si $0 = 2xy$; es decir, si y solo si $x = 0$ o $y = 0$. Por lo tanto, si $x \neq 0$ y $y \neq 0$, se tiene que

   $$f(x + y) \neq f(x) + f(y).$$

   Esto significa que f no es una aplicación lineal. □

5. ¿Es la función $g\colon \mathbb{C} \longrightarrow \mathbb{C}$, definida por $g(z) = \overline{z}$, una aplicación lineal si se considera $\mathbb{C}$ un espacio vectorial sobre $\mathbb{C}$? ¿En el caso de que $\mathbb{C}$ sea un espacio vectorial sobre $\mathbb{R}$?

   Desarrollo.

   (a) Si $\mathbb{C}$ se considera como un espacio sobre $\mathbb{C}$, g no es lineal, pues, por un lado se tiene que:

   $$g(\imath(1 + \imath)) = g(\imath - 1) = -1 - \imath,$$

   pero por el otro lado, que:

   $$\imath g(1 + \imath) = 1 - \imath;$$

   es decir:

   $$g(\imath(1 + \imath)) \neq \imath g(1 + \imath).$$

   Por lo tanto g no es lineal.

(b) Si $\mathbb{C}$ es el espacio vectorial sobre $\mathbb{R}$, tenemos que para cualquier escalar α, se cumple que:

$$\begin{aligned} g(\alpha u + v) &= \overline{\alpha u + v} \\ &= \overline{\alpha u} + \overline{v} \\ &= \overline{\alpha u} + \overline{v} \\ &= \alpha \overline{u} + \overline{ve}, \\ &= \alpha g(u) + g(v), \end{aligned}$$

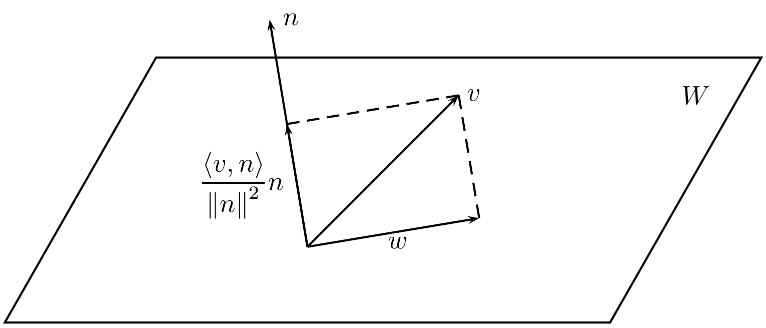
ya que $\overline{\alpha} = \alpha$ (ya que α es un número real). Por lo tanto, g sí es una aplicación lineal.

□

6. Sea W un subespacio vectorial de un espacio vectorial V provisto de un producto interno. Dado $v \in V$, existe un único vector proyección de v sobre W dado por:

$$w = v - \frac{\langle v, n\rangle}{\|n\|^n} n,$$

donde n es un vector ortogonal a W. En el siguiente dibujo se ilustra la relación entre v, w y n:



La aplicación $p\colon V \longrightarrow W$ definida por

$$p(v) = v - \frac{\langle v, n\rangle}{\|n\|^n} n,$$

es una aplicación lineal, lo que es sencillo probar pues el producto interno es lineal respecto a la primera componente.

7. Sea $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por

$$f(x, y, z) = (x - z, y + z).$$

Entonces f es una aplicación lineal.

Demostración. Sean $u = (x_1, x_2, x_3)$ y $v = (y_1, y_2, y_3)$ elementos de $\mathbb{R}^3$ y $\alpha \in \mathbb{R}$. Entonces:

$$f(\alpha u + v) = f(\alpha(x_1, x_2, x_3) + (y_1, y_2, y_3))$$

$$
\begin{aligned}
&= f(\alpha x_1 + y_1, \alpha x_2 + y_2, \alpha x_3 + y_3) \\
&= (\alpha x_1 + y_1 - (\alpha x_3 + y_3), (\alpha x_2 + y_2) + (\alpha x_3 + y_3)) \\
&= (\alpha(x_1 - x_3) + (y_1 - y_3), \alpha(x_2 + x_3) + (y_2 + y_3)) \\
&= (\alpha(x_1 - x_3), \alpha(x_2 + x_3)) + (y_1 - y_3, y_2 + y_3) \\
&= \alpha(x_1 - x_3, x_2 + x_3) + (y_1 - y_3, y_2 + y_3) \\
&= \alpha f(x_1, x_2, x_3) + f(y_1, y_2, y_3).
\end{aligned}
$$

Por lo tanto, la función f es una aplicación lineal. $\square$

5.2 Isomorfismo entre espacios vectoriales

Definición 5.2 (Isomorfismo) Una aplicación $f\colon V \longrightarrow W$ es un **isomorfismo** entre V y W si y solo si f es lineal y biyectiva. Dos espacios vectoriales V y W son **isomorfos** si y solo si existe un isomorfismo de V sobre W. Se utilizará la notación $V \cong W$ para indicar que V y W son isomorfos.

Ejemplo 5.71

Sea $f\colon \mathscr{P}_n[x] \longrightarrow \mathbb{R}^{n+1}$ definida por

$$f(p) = (a_0, a_1, \ldots, a_n)$$

para todo $p \in \mathscr{P}_n[x]$ tal que

$$p(x) = \sum_{j=1}^{n} a_j x^j.$$

¿Es f un isomorfismo entre $\mathscr{P}_n[x]$ y R^{n+1}?

Solución.

1. Veamos si f es una aplicación lineal. Para ello, sean $\alpha \in \mathbb{R}$, p y q en $\mathscr{P}_n[x]$, donde

$$p(x) = \sum_{j=1}^{n} a_j x^j \quad \text{y} \quad q(x) = \sum_{j=1}^{n} b_j x^j.$$

Entonces:

$$(\alpha p + q)(x) = \alpha p(x) + q(x) = \sum_{j=1}^{n} (\alpha a_j + b_j) x^j.$$

Por lo tanto:

$$
\begin{aligned}
f(\alpha p + q) &= (\alpha a_0 + b_0, \alpha a_1 + b_1, \ldots, \alpha a_n + b_n) \\
&= (\alpha a_0, \alpha a_1, \ldots, \alpha a_n) + (b_0, b_1, \ldots, b_n) \\
&= \alpha(a_0, a_1, \ldots, a_n) + (b_0, b_1, \ldots, b_n) \\
&= \alpha f(p) + f(q).
\end{aligned}
$$

Es decir, f es una aplicación lineal.

2. Veamos ahora si f es inyectiva. Para ello, supongamos que $f(p) = f(q)$, donde

$$p(x) = \sum_{j=1}^{n} a_j x^j \quad \text{y} \quad q(x) = \sum_{j=1}^{n} b_j x^j.$$

   Entonces

$$(a_0, a_1, \ldots, a_n) = (b_0, b_1, \ldots, b_n),$$

   de donde, por la igualdad en $\mathbb{R}^{n+1}$, tenemos que $a_j = b_j$ para todo $j \in \{0, 1, \ldots, n\}$. Por lo tanto $p = q$. Es decir, f es inyectiva.

3. Finalmente, veamos si f es sobreyectiva. Para ello, sea $v = (a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1}$. Vamos a averiguar si existe $p \in \mathscr{P}_n[x]$ tal que $f(p) = v$.

   Si se define

$$p(x) = \sum_{j=1}^{n} a_j x^j$$

   para todo $x \in \mathbb{R}$, entonces es inmediato que $f(p) = v$. Es decir, f es sobreyectiva.

Tenemos, entonces, que la aplicación f es lineal y sobreyectiva. Por lo tanto, es un isomorfismo entre $\mathscr{P}_n[x]$ y $\mathbb{R}^{n+1}$. ♦

Teorema 5.4 El conjunto $\mathscr{L}(V, W)$ es un subespacio vectorial del espacio de todas las funciones de V en W con las operaciones usuales. Es decir:

1. si $f \in \mathscr{L}(V, W)$ y $g \in \mathscr{L}(V, W)$, entonces $f + g \in \mathscr{L}(V, W)$; y
2. si $f \in \mathscr{L}(V, W)$ y $\alpha \in \mathbb{K}$, entonces $\alpha f \in \mathscr{L}(V, W)$.

En otras palabras, la suma de aplicaciones lineales es una aplicación lineal, y el producto de un escalar y una aplicación lineal es una aplicación lineal.

Demostración. Sean $\alpha \in \mathbb{K}$, $f \in \mathscr{L}(V, W)$ y $g \in \mathscr{L}(V, W)$. Entonces, para todo $u \in V$ y todo $v \in V$:

$$\begin{aligned}
(f+g)(\alpha u + v) &= f(\alpha u + v) + g(\alpha u + v) && \text{definición de suma de funciones,}\\
&= [\alpha f(u) + f(v)] + [\alpha g(u) + g(v)] && f \text{ y } g \text{ son aplicaciones lineales,}\\
&= \alpha[f(u) + g(u)] + [f(v) + g(v)] && \text{asociativa, conmutativa y distributiva en } W,\\
&= \alpha(f+g)(u) + (f+g)(v) && \text{definición de suma de funciones.}
\end{aligned}$$

Por lo tanto, $f + g \in \mathscr{L}(V, W)$.

Ahora probemos que $\lambda f \in \mathscr{L}(V,W)$ si $\lambda \in \mathbb{K}$. Sean $\alpha \in \mathbb{K}$, $u \in V$ y $v \in V$. Entonces:

$$\begin{aligned}
(\lambda f)(\alpha u+v) &= \lambda f(\alpha u+v) && \text{definición del producto de un escalar y una función,}\\
&= \lambda[\alpha f(u)+f(v)] && f \text{ es una aplicación lineal,}\\
&= \alpha\lambda f(u)+\lambda f(v) && \text{distributiva en } W \text{ y conmutativa en } \mathbb{K},\\
&= \alpha(\lambda f)(u)+(\lambda f)(v) && \text{definición del producto de un escalar y una función.}
\end{aligned}$$

Por lo tanto, λf es una aplicación lineal de V en W. □

La composición de dos aplicaciones lineales también es una aplicación lineal, como lo probamos a continuación.

Teorema 5.5 Sean $f \in \mathscr{L}(V,W)$ y $g \in \mathscr{L}(W,U)$. Entonces $g \circ f \in \mathscr{L}(V,U)$.

Demostración. Sean $\alpha \in \mathbb{K}$, $u \in V$ y $v \in V$. Entonces:

$$\begin{aligned}
(g\circ f)(\alpha u+v) &= g(f(\alpha u+v)) && \text{definición de composición de funciones,}\\
&= g(\alpha f(u)+f(v)) && f \text{ es lineal,}\\
&= \alpha g(f(u))+g(f(v)) && g \text{ es lineal,}\\
&= \alpha(g\circ f)(u)+(g\circ f)(v) && \text{definición de composición de funciones.}
\end{aligned}$$

Por lo tanto $g \circ f \in \mathscr{L}(V,U)$. □

Otra operación que preserva la linealidad es la inversión de funciones.

Teorema 5.6 Sea $f \in \mathscr{L}(V,W)$ y biyectiva. Entonces $f^{-1} \in \mathscr{L}(W,V)$.

Demostración. Sean $\alpha \in \mathbb{K}$, $u \in W$ y $v \in W$. Vamos a probar que

$$f^{-1}(\alpha u+v) = \alpha f^{-1}(u)+f^{-1}(v).$$

Para ello, sea $w = f^{-1}(\alpha u+v)$. Entonces, por la definición de inversa, tenemos que

$$f(w) = \alpha u+v.$$

Ahora bien, como f es sobreyectiva, existen únicos $\hat{u} \in V$ y $\hat{v} \in V$ tales que

$$u = f(\hat{u}) \quad \text{y} \quad v = f(\hat{v}).$$

Por lo tanto,

$$f(w) = \alpha f(\hat{u}) + f(\hat{v}),$$

de donde, como f es lineal, tenemos lo siguiente:

$$f(w) = f(\alpha\hat{u} + \hat{v}).$$

Y, como f es inyectiva, entonces:

$$w = \alpha\hat{u} + \hat{v},$$

que, como $\hat{u} = f^{-1}(u)$ y $\hat{v} = f^{-1}(v)$, equivale a:

$$f^{-1}(\alpha u + v) = \alpha f^{-1}(u) + f^{-1}(v).$$

Es decir, f^{-1} es una aplicación lineal. □

Ejemplo 5.72

Sean $f\colon V \longrightarrow U$ y $g\colon U \longrightarrow W$ tales que $g \circ f$ es lineal y g es lineal e inyectiva. Probemos que f es lineal.

Solución. Sean $\alpha \in \mathbb{K}$, $u \in V$ y $w \in V$. Vamos a probar que

$$f(\alpha u + v) = \alpha f(u) + f(v).$$

Como $g \circ f$ es lineal, tenemos que

$$(g \circ f)(\alpha u + v) = \alpha(g \circ f)(u) + (g \circ f)(v).$$

De la definición de compuesta, esta igualdad se escribe así:

$$(g \circ f)(\alpha u + v) = \alpha g(f(u)) + g(f(v)).$$

Como g es lineal, tenemos que:

$$g(f(\alpha u + v)) = g(\alpha f(u) + f(v)).$$

Y, esta igualdad, junto con la inyectividad de g, implica que

$$f(\alpha u + v) = \alpha f(u) + f(v).$$

Es decir, f es lineal. ♦

El siguiente ejemplo es similar al anterior, salvo que se elimina la hipótesis de que g sea inyectiva. ¿Se mantiene la conclusión de que f es lineal?

Ejemplo 5.73

Sean $f\colon V \longrightarrow U$ y $g\colon U \longrightarrow W$ tales que $g \circ f$ es lineal y g es lineal. ¿Es f es lineal?

Solución. La condición de inyectividad para g es inevitable para que f sea lineal. En efecto, si $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por

$$f(x,y) = (x^2, y),$$

entonces f no es lineal, como el lector puede verificar fácilmente.

Por otro lado, sea g la aplicación lineal de $\mathbb{R}^2$ en sí mismo definida por

$$g(x,y) = (0,y).$$

Esta función es lineal pero no inyectiva. Lo primero es de demostración sencilla. Lo segundo se obtiene al observar que

$$g(1,1) = (0,1) = g(-1,1),$$

pero $(1,1) \neq (-1,1)$.

Finalmente, es fácil probar también que $g \circ f$ es lineal. ♦

Teorema 5.7 Una aplicación lineal de un espacio lineal de dimensión finita queda completamente determinada por lo valores que tome en la base. De manera más precisa, sean V es un espacio de dimensión finita n y W un espacio vectorial, ambos sobre un campo $\mathbb{K}$, $B = \{e_i\}$ una base de V y $f \in \mathscr{L}(V,W)$. Entonces, f queda determinada por los valores $f(e_i)$.

Demostración. Sea $v \in V$. Como B es una base de V, existen escalares α_i tales que

$$v = \sum_{j=1}^{n} \alpha_j e_j.$$

Entonces, como f es lineal, tenemos que

$$f(v) = \sum_{j=1}^{n} \alpha_j f(e_j).$$

Es decir, el valor de $f(v)$ para todo $v \in V$ queda determinado por $f(e_j)$, a más de los α_j propios de v. □

En otras palabras, para definir una aplicación lineal en un espacio de dimensión finita, es suficiente con definir la aplicación para cada uno de los elementos de una base del espacio.

Ejemplo 5.74

Sea $f \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^3)$ tal que

$$f(1,-1) = (1,2,-1) \quad \text{y} \quad f(1,1) = (2,-1,3).$$

Determinar f.

Solución. Dado que $B = \{(1,-1),\ (1,1)\}$ es una base de $\mathbb{R}^2$, podemos obtener $f(x,y)$ para todo $(x,y) \in \mathbb{R}^2$.

Para ello, encontremos, en primer lugar, las coordenadas de (x,y) respecto de B. Sean $\alpha \in \mathbb{R}$ y $\beta \in \mathbb{R}$ tales que

$$(x,y) = \alpha(1,-1) + \beta(1,1).$$

Entonces, α y β satisfacen el siguiente sistema de dos ecuaciones lineales:

$$\begin{aligned} \alpha + \beta &= x \\ -\alpha + \beta &= y, \end{aligned}$$

cuya única solución es fácil hallar:

$$\alpha = \frac{x-y}{2} \quad \text{y} \quad \beta = \frac{x+y}{2}.$$

Entonces:

$$\begin{aligned} f(x,y) &= f(\alpha(1,-1) + \beta(1,1)) \\ &= \alpha f(1,-1) + \beta f(1,1) \\ &= \alpha(1,2,-1) + \beta(2,-1,3) \\ &= (\alpha + 2\beta, 2\alpha - \beta, -\alpha + 3\beta). \end{aligned}$$

Como:

$$\begin{aligned} \alpha + 2\beta &= \frac{x-y}{2} + 2\frac{x+y}{2} = \frac{3x+y}{2} \\ 2\alpha - \beta &= 2\frac{x-y}{2} - \frac{x+y}{2} = \frac{x-3y}{2} \\ -\alpha + 3\beta &= -\frac{x-y}{2} + 3\frac{x+y}{2} = x + 2y, \end{aligned}$$

se tiene que

$$f(x,y) = \left(\frac{3x+y}{2},\ \frac{x-3y}{2},\ 2+2y\right).$$

♦

Teorema 5.8 Sean V y W dos espacios vectoriales sobre un campo $\mathbb{K}$, $B = \{v_1, v_2, \ldots, v_n\}$ una base de V, y sean w_1, w_2, ..., w_n elementos distintos de W. Entonces, existe una sola aplicación $f \in \mathscr{L}(V,W)$ tal que $f(v_i) = w_i$ para todo i con $1 \leq i \leq n$.

Demostración. Sea $v \in V$. Entonces, existe un único conjunto de escalares α_1, α_2, ..., α_n tales que

$$v = \sum_{j=1}^{n} \alpha_j v_j.$$

Definamos la aplicación f de V en W de la siguiente manera:

$$f(v) = \sum_{j=1}^{n} \alpha_j w_j.$$

Dada la unicidad de los α_j, f está bien definida.

Ahora probemos que f es lineal y que $f(v_j) = w_j$ para cada j con $1 \leq j \leq n$.

1. Sean $u \in V$ y $v \in V$. Entonces existe escalares α_j y β_j tales que:

$$u = \sum_{j=1}^{n} \alpha_j v_j \quad \text{y} \quad v = \sum_{j=1}^{n} \beta_j v_j.$$

Entonces, tenemos que

$$f(u) = \sum_{j=1}^{n} \alpha_j w_j \quad \text{y} \quad f(v) = \sum_{j=1}^{n} \beta_j w_j.$$

Para $\lambda \in \mathbb{K}$ tenemos que:

$$\lambda u + v = \sum_{j=1}^{n} (\lambda\alpha_j + \beta_j) v_j.$$

Por lo tanto:

$$\begin{aligned} f(\lambda u + v) &= f\left(\sum_{j=1}^{n} (\lambda\alpha_j + \beta_j) v_j\right) \\ &= \sum_{j=1}^{n} (\lambda\alpha_j + \beta_j) w_j \\ &= \lambda \sum_{j=1}^{n} \alpha_j w_j + \sum_{j=1}^{n} \beta_j w_j \\ &= \lambda f(u) + f(v). \end{aligned}$$

Es decir, f es lineal.

2. Para cada i con $1 \leq i \leq n$, se tiene que

$$v_i = \sum_{j=1}^{n} \delta_j^i v_j,$$

donde

$$\delta_j^i = \begin{cases} 1 & \text{si } j = i, \\ 0 & \text{si } j \neq i. \end{cases}$$

Por lo tanto:

$$f(v_i) = \sum_{j=1}^{n} \delta_j^i w_j = w_i.$$

Finalmente, el lector puede probar fácilmente que f es única. □

Ejemplo 5.75

Sea

$$W = \left\{(x, y, z) \in \mathbb{R}^3 : x - y + z = 0\right\}.$$

Definir una aplicación lineal de $\mathbb{R}^2$ en W.

Solución. Consideremos la base canónica en $\mathbb{R}^2$ y los vectores $w_1 = (1,1,0)$ y $w_2 = (0,1,1)$ son elementos de W. Entonces, por teorema 5.8, existe un única aplicación lineal f de $\mathbb{R}^2$ en W tal que

$$f(1,0) = w_1 \quad \text{y} \quad f(0,1) = w_2.$$

Ahora bien, dado $(x,y) \in \mathbb{R}^2$, tenemos que

$$\begin{aligned} f(x,y) &= f(x(1,0) + y(0,1)) \\ &= xf(1,0) + yf(0,1) \\ &= xw_1 + yw_2 \\ &= x(1,1,0) + y(0,1,1) \\ &= (x, x+y, y). \end{aligned}$$

En resumen, f es una aplicación de $\mathbb{R}^2$ en W definida por

$$f(x,y) = (x, x+y, y).$$

♦

Teorema 5.9 Sean V y W dos espacios vectoriales de igual dimensión finita,

$$B_v = \{v_1, v_2, \ldots, v_n\} \quad \text{y} \quad B_w = \{w_1, w_2, \ldots, w_n\}$$

bases para V y W respectivamente, y $f \in \mathscr{L}(V,W)$ tal que

$$f(v_i) = w_i$$

para todo i con $1 \leq i \leq n$. Entonces f es un isomorfismo.

Demostración. Debemos probar que f es biyectiva.

1. *Inyectividad*: sean $u \in V$ y $v \in V$ tales que $f(u) = f(v)$. Vamos a demostrar que $u = v$.

   Existen escalares α_i y β_i tales que

   $$u = \sum_{j=1}^{n} \alpha_j v_j \quad \text{y} \quad v = \sum_{j=1}^{n} \beta_j v_j.$$

   Por lo tanto, de $f(u) = f(v)$, obtenemos que

   $$\sum_{j=1}^{n} \alpha_j w_j = \sum_{j=1}^{n} \beta_j w_j,$$

   de donde:

   $$\sum_{j=1}^{n} (\alpha_j - \beta_j) w_j = 0.$$

   Puesto que B_w es linealmente independiente, entonces

   $$\alpha_j - \beta_j = 0$$

   para todo j tal que $1 \leq j \leq n$.

   Pero esto significa que $\alpha_j = \beta_j$ para todo j de donde $u = v$.

2. *Sobreyectividad*: sea $w \in W$; debemos hallar $v \in V$ tal que $w = f(w)$.

   Como B_w es una base de W, existen escalares α_j tales que

$$w = \sum_{j=1}^{n} \alpha_j w_j.$$

   Dado que $f(v_j) = w_j$, el v que buscamos es

$$v = \sum_{j=1}^{n} \alpha_j v_j.$$

   En efecto, tenemos que

$$f(v) = \sum_{j=1}^{n} \alpha_j f(v_j) = \sum_{j=1}^{n} \alpha_j w_j = w.$$

□

Ejemplo 5.76

Sea $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$f(x, y, z) = (x - y + z, x + y + z, x + y - z).$$

Probar que f es biyectiva.

Solución. Es fácil probar que f es lineal. Ahora, apliquemos el teorema 5.9 para mostrar que f es un isomorfismo, con lo cual probaremos que f es biyectiva.

Probemos, entonces, que f aplica la base canónica de $\mathbb{R}^3$ en una base. De manera más precisa, probemos que el conjunto

$$B = \{f(1, 0, 0), f(0, 1, 0), f(0, 0, 1)\}$$

es linealmente independiente (con lo cual ya sabríamos que es una base de $\mathbb{R}^3$).

El conjunto B es

$$B = \{(1, 1, 1),\ (-1, 1, 1),\ (1, 1, -1)\}.$$

Sean α, β y γ escalares tales que

$$\alpha(1, 1, 1) + \beta(-1, 1, 1) + \gamma(1, 1, -1) = 0.$$

Entonces estos escalares satisfacen el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} \alpha - \beta + \gamma &= 0 \\ \alpha + \beta + \gamma &= 0 \\ \alpha + \beta - \gamma &= 0. \end{aligned}$$

Ahora bien, el determinante de este sistema es igual a:

$$\begin{aligned} \begin{vmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{vmatrix} &= \begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & -1 \end{vmatrix} \\ &= 2 \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -4 \neq 0. \end{aligned}$$

Por lo tanto, el sistema tiene una única solución y esa solución es la trivial: $\alpha = \beta = \gamma = 0$.

En resumen, f es un isomorfismo. ♦

5.3 Núcleo de una aplicación lineal

Definición 5.3 (Núcleo) Sea $f \in \mathscr{L}(V, W)$. El **núcleo** de f, representado por $\mathcal{N}(f)$ es el conjunto de todos los elementos v de V tales que $f(v) = 0$; es decir:

$$\mathcal{N}(f) = \{v \in V : f(v) = 0\}.$$

En la caracterización de una aplicación lineal, el núcleo, junto la imagen, juegan un rol fundamental. Recordemos que la imagen de una función, en particular de una aplicación lineal f es el conjunto

$$\mathrm{Im}(f) = \{w \in W : (\exists v \in V)(f(v) = w)\}.$$

Teorema 5.10 Sea $f \in \mathscr{L}(V, W)$. Entonces $\mathcal{N}(f)$ es un subespacio vectorial de V y $\mathrm{Im}(f)$ son subespacio vectorial de W.

Demostración.

1. En primer lugar, $0 \in \mathcal{N}(f)$ pues $f(0) = 0$. Por otro lado, si $\alpha \in \mathbb{K}$, u y v en $\mathcal{N}(f)$, entonces:
$$f(\alpha u + v) = \alpha f(u) + f(v) = \alpha \cdot 0 + 0 = 0,$$
ya que $f(u) = f(v) = 0$. Por lo tanto $\alpha u + v \in \mathcal{N}(f)$. Hemos demostrado, entonces, que $\mathcal{N}(f)$ es un subespacio vectorial de V.

2. Dado que $f(0) = 0$, se tiene que $0 \in \mathrm{Im}(f)$. Además, w_1 y w_2 en $\mathrm{Im}(f)$, entonces existen v_1 y v_2 en V tales que
$$w_1 = f(v_1) \quad \text{y} \quad w_2 = f(v_2).$$
Ahora, si $\alpha \in \mathbb{K}$, entonces
$$\alpha w_1 + w_2 = \alpha f(v_1) + f(v_2) = f(\alpha v_1 + v_2).$$
Por lo tanto $\alpha w_1 + w_2 \in \mathrm{Im}(f)$, ya que $\alpha v_1 + v_2 \in V$. Esto prueba que $\mathrm{Im}(f)$ es un subespacio vectorial de W. □

Ejemplo 5.77

Sea $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$f(x, y, z) = (x - y + z, x + y + z, x - 3y + z).$$

Encontrar la dimensión de $\mathcal{N}(f)$ y de $\mathrm{Im}(f)$.

Solución.

1. Para determinar la dimensión de $\mathcal{N}(f)$ caractericemos, en primer lugar, sus elementos; y luego encontremos una base para él.

   Dado que

$$\begin{aligned}\mathcal{N}(f) &= \{(x,y,z) : f(x,y,z) = 0 = (0,0,0)\}\\ &= \{(x,y,z) : (x-y+z, x+y+z, x-3y+z) = (0,0,0)\},\end{aligned}$$

   el núcleo de f es el conjunto solución del sistema de tres ecuaciones con tres incógnitas cuya matriz ampliada es la siguiente:

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

   Ahora bien, tenemos que:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0\\ 1 & 1 & 1 & 0\\ 1 & -3 & 1 & 0\end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right).$$

   Por lo tanto:

$$\begin{aligned}\mathcal{N}(f) &= \{(x,y,z) : x+z = 0 \text{ y } y = 0\}\\ &= \{(x,0,-x) : x \in \mathbb{R}\}.\end{aligned}$$

   Ahora, una base para $\mathcal{N}(f)$ la obtenemos así:

$$\begin{aligned}\mathcal{N}(f) &= \{(x,0,-x) : x \in \mathbb{R}\}\\ &= \{x(1,0,-1) : x \in \mathbb{R}\}\\ &= \langle\{(1,0,-1)\}\rangle.\end{aligned}$$

   Y como el conjunto $B = \{(1,0,-1)\}$ es linealmente independiente (ya que tiene un solo elemento diferente de 0), tenemos que B es una base para $\mathcal{N}(f)$, y, por lo tanto, $\dim(\mathcal{N}(f)) = 1$.

2. Hagamos un procedimiento similar para obtener la dimensión de la imagen de f. Por un lado, tenemos que:

$$\begin{aligned}\operatorname{Im}(f) &= \{(r,s,t) : (\exists(x,y,z))(f(x,y,z) = (r,s,t))\}\\ &= \{(r,s,t) : (\exists(x,y,z))((x-y+z, x+y+z, x-3y+z) = (r,s,t))\}.\end{aligned}$$

   Entonces, la imagen de f es el conjunto solución del sistema de ecuaciones cuya matriz ampliada es:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & r\\ 1 & 1 & 1 & s\\ 1 & -3 & 1 & t\end{array}\right).$$

   Con el mismo procedimiento realizado en el caso del núcleo, obtenemos que

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & r\\ 1 & 1 & 1 & s\\ 1 & -3 & 1 & t\end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & r\\ 0 & 2 & 0 & s-r\\ 0 & 0 & 0 & t-2r+s\end{array}\right).$$

Entonces:

$$\begin{aligned}
\operatorname{Im}(f) &= \{(r,s,t) : t - 2r + s = 0\} \\
&= \{(r,s,2r-s) : r \in \mathbb{R},\ s \in \mathbb{R}\} \\
&= \{r(1,0,2) + s(0,1,-1) : r \in \mathbb{R},\ s \in \mathbb{R}\}.
\end{aligned}$$

Esto significa que

$$B = \{(1,0,2),\ (0,1,-1)\}$$

genera $\operatorname{Im}(f)$. Además, como B es linealmente independiente (algo que es fácil demostrar, por lo que se deja al lector la tarea), tenemos que $\dim(\operatorname{Im}(f)) = 2$.

♦

Teorema 5.11 Sea $f \in \mathscr{L}(V,W)$. Entonces, f es inyectiva si y solo si $\mathcal{N}(f) = \{0\}$.

Demostración.

1. Supongamos que $\mathcal{N}(f) = \{0\}$. Vamos a probar que f es inyectiva. Para ello, sean u y v en V tales que $f(u) = f(v)$. Entonces, dado que f es lineal, tenemos que:

   $$f(u) - f(v) = 0 \Rightarrow f(u-v) = 0.$$

   Y, como el único elemento de V cuya imagen bajo f es 0, ya que $\mathcal{N}(f) = \{0\}$, tenemos que $u - v = 0$, de donde $u = v$. Por lo tanto, f es inyectiva.

2. Supongamos que f es inyectiva. Dado que $0 \in \mathcal{N}(f)$, solamente nos resta por probar que $\mathcal{N}(f) \subseteq \{0\}$. Para ello, sea $v \in \mathcal{N}(f)$. Entonces $f(v) = 0$. Como $f(0) = 0$, entonces $f(v) = f(0)$, de donde, por la inyectividad de f, concluimos que $v = 0$. Por lo tanto $\mathcal{N}(f) \subseteq \{0\}$. □

Teorema 5.12 Sean $f \in \mathscr{L}(V,W)$, $w \in W$ y $v_0 \in V$ tales que $f(v_0) = w$. Entonces, la solución general de la ecuación

$$f(x) = w \tag{5.1}$$

es de la forma

$$x = v_0 + u,$$

donde $u \in \mathcal{N}(f)$.

Demostración. Que $v_0 + u$ sea una solución de la ecuación (5.1) se deduce de:

$$f(v_0 + u) = f(v_0) + f(u) = w + 0 = w.$$

Por otro lado, supongamos que v es una solución de la ecuación (5.1). Vamos a probar que existe $u \in \mathcal{N}(f)$ tal que $v = v_0 + u$.

Es claro que u debe ser $v - v_0$. En efecto:

$$f(u) = f(v - v_0) = f(v) - f(v_0) = w - w = 0.$$ □

Ejemplo 5.78 Sea $f \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^2)$ tal que

$$F(x, y, z) = (x - y + z, x + y - z).$$

Es claro que $f(1, 1, 1) = (1, 1)$. Entonces la solución general de la ecuación

$$(x - y + z, x + y - z) = (1, 1) \tag{5.2}$$

es

$$(1, 1, 1) + u,$$

donde $u \in \mathcal{N}(f)$. Por lo tanto, para encontrar la solución general de 5.2, encontremos el núcleo de f.

Para ello, debemos hallar la solución de la ecuación

$$(x - y + z, x + y - z) = (0, 0).$$

O, de manera equivalente, debemos resolver el sistema de dos ecuaciones con tres incógnitas cuya matriz ampliada es:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right).$$

Ahora bien, dado que:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 2 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right),$$

tenemos que

$$\mathcal{N}(f) = \{(x, y, z) \in \mathbb{R}^3 : x = 0,\ y = z\} = \{(0, z, z) : z \in \mathbb{R}\}.$$

Por lo tanto, la solución general del sistema 5.2 es:

$$(1, 1, 1) + (0, z, z) = (1, 1 + z, 1 + z)$$

para todo $z \in \mathbb{R}$. ♦

El siguiente teorema afirma que la independencia lineal es conservada por las aplicaciones lineales inyectivas.

Teorema 5.13 Sea $f \in \mathscr{L}(V, W)$ tal que $\mathcal{N}(f) = \{0\}$. Entonces, si $B = \{v_1, v_2, \ldots, v_n\} \subseteq V$ es linealmente independiente, entonces $S = \{f(v_1), f(v_2), \ldots, f(v_n)\}$ es linealmente independiente.

Demostración. Sean α_i escalares tales que

$$\sum_{j=1}^{n} \alpha_i f(v_i) = 0.$$

Vamos a probar que $\alpha_i = 0$ para todo $i \in I_n$.

Dado que f es lineal, tenemos que

$$\sum_{j=1}^{n} \alpha_i f(v_i) = f\left(\sum_{j=1}^{n} \alpha_i v_i\right) = 0.$$

Como el núcleo de f contiene únicamente a 0, esta igualdad implica que

$$\sum_{j=1}^{n} \alpha_i v_i = 0.$$

A su vez, esta igualdad implica que todos los α_i son iguales a 0, pues B es un conjunto linealmente independiente. Por lo tanto, S también es linealmente independiente. $\square$

El siguiente teorema afirma que la imagen por una aplicación lineal de un conjunto generador del espacio de salida es igual a la imagen de la aplicación.

Teorema 5.14 Sea $f \in \mathscr{L}(V, W)$. Si $B = \{v_1, v_2, \ldots, v_n\}$ genera V, entonces $f(B)$ genera la imagen $\operatorname{Im}(f)$ de f.

Demostración.

1. Probemos que $\langle f(B)\rangle \subseteq \operatorname{Im}(f)$. Para ello, sea $w \in \langle f(B)\rangle$. Entonces existe escalares α_j tales que

   $$w = \sum_{j=1}^{n} \alpha_i f(v_i),$$

   pues $f(B) = \{f(v_i) : i \in I_n\}$.

   De la linealidad de f, obtenemos que:

   $$w = \sum_{j=1}^{n} \alpha_i f(v_i) = f\left(\sum_{j=1}^{n} \alpha_i v_i\right).$$

   Si definimos

   $$v = \sum_{j=1}^{n} \alpha_i v_i,$$

   entonces $v \in V$, ya que B genera V, y $w = f(v)$. Por lo tanto, existe $v \in V$ tal que $w = f(v)$. Esto significa que $w \in \operatorname{Im}(f)$.

2. Ahora probemos que $\operatorname{Im}(f) \subseteq \langle f(B)\rangle$. Para ello, sea $w \in \operatorname{Im}(f)$. Existe, entonces, $v \in V$ tal que $w = f(v)$. Como B genera V, entonces existen escalares α_j tales que

   $$v = \sum_{j=1}^{n} \alpha_j v_j.$$

   Entonces

   $$w = f(v) = f\left(\sum_{j=1}^{n} \alpha_j v_j\right) = \sum_{j=1}^{n} \alpha_j f(v_j).$$

   Es decir, $w \in \langle f(B)\rangle$. $\square$

Corolario 5.15 Sea $f \in \mathscr{L}(V,W)$ tal que $\mathcal{N}(f) = \{0\}$. Entonces, $B = \{v_j : j \in I_n\}$ es una base de V si y solo si $S = f(B)$ es una base de $\mathrm{Im}(f)$.

En otras palabras, la dimensión de la imagen de f es igual a la dimensión de V.

Demostración. Si el conjunto B es una base de V, entonces B es linealmente independiente y genera V. Por lo tanto, dado que f es inyectiva, ya que $\mathcal{N}(f) = \{0\}$ (teorema 5.11), por el teorema 5.13, $f(B)$ es linealmente independiente; y por el teorema 5.14, $f(B)$ genera $\mathrm{Im}(f)$. Por lo tanto, $f(B)$ es una base de $\mathrm{Im}(f)$.

Recíprocamente, si $f(B)$ es una base de $\mathrm{Im}(f)$, entonces es un conjunto linealmente independiente y genera $\mathrm{Im}(f)$. Probemos que B es una base de V.

Para ello, sea $v \in V$. Entonces $f(v) \in \mathrm{Im}(f)$. Por lo tanto, existe escalares α_j tales que

$$f(v) = \sum_{j=1}^{n} \alpha_j f(v_j).$$

Por la linealidad de f, de esta igualdad obtenemos que:

$$f(v) = f\left(\sum_{j=1}^{n} \alpha_j v_j\right).$$

Y por la inyectividad de f, esta igualdad implica que:

$$v = \sum_{j=1}^{n} \alpha_j v_j.$$

Por lo tanto, B genera V.

Ahora probemos que B es linealmente independiente. Sean α_j escalares tales que

$$\sum_{j=1}^{n} \alpha_j v_j = 0.$$

Entonces, por la linealidad de f, obtenemos que:

$$f\left(\sum_{j=1}^{n} \alpha_j v_j\right) = \sum_{j=1}^{n} \alpha_j f(v_j) = f(0) = 0.$$

Como $f(B)$ es linealmente independiente, concluimos que todos los α_j son iguales a 0. Entonces, $f(B)$ es linealmente independiente. □

Teorema 5.16 Sea $f \in \mathscr{L}(V,W)$ con V un espacio de dimensión finita. Entonces:

$$\dim(V) = \dim(\mathcal{N}(f)) + \dim(\mathrm{Im}(f)).$$

Demostración.

1. Supongamos que $\mathcal{N}(f) = V$. Entonces $f(u) = 0$ para todo $v \in V$. Por lo tanto, $\operatorname{Im}(f) = \{0\}$. Entonces
$$\dim(\mathcal{N}(f)) + \dim(\operatorname{Im}(f)) = \dim(V) + 0 = \dim(V).$$

2. Supongamos que $\mathcal{N}(f) \neq V$. Sea $u \in V$ tal que $f(u) \neq 0$. Entonces $\operatorname{Im}(f) \neq \{0\}$.

   Si $\mathcal{N}(f) = \{0\}$, por el corolario 5.15, la dimensión de $\operatorname{Im}(f)$ es igual a la dimensión de V. Entonces:
$$\dim(V) = \dim(\operatorname{Im}(f)) = \dim(\operatorname{Im}(f)) + \dim(\mathcal{N}(f)),$$
   ya que $\dim(\mathcal{N}(f)) = 0$.

   Ahora supongamos que $\mathcal{N}(f) \neq \{0\}$. Sean $W = \{w_j : j \in I_s\}$ una base de $\operatorname{Im}(f)$ y $B = \{v_j : j \in I_s\}$ donde $f(v_j) = w_j$. De la definición de función, se colige que la cardinalidad de B es s.

   Sea $C = \{u_k : k \in I_r\}$ con $r < s$ una base para $\mathcal{N}(f)$. Probemos que
$$T = \{v_j, u_k : j \in I_s,\ k \in I_r\}$$
   es una base para V, lo que probaría que la dimensión de V es igual a la suma de las dimensiones del núcleo y de la imagen.

   En primer lugar, si $v \in V$, entonces $f(v)$ está en $\operatorname{Im}(f)$. Existen, entonces, s escalares α_j tales que
$$f(v) = \sum_{j=1}^{s} \alpha_j w_j = \sum_{j=1}^{s} \alpha_j f(v_j).$$
   De esta igualdad, por la linealidad de f, obtenemos que:
$$f\left(v - \sum_{j=1}^{s} \alpha_j v_j\right) = 0.$$
   Entonces
$$v - \sum_{j=1}^{s} \alpha_j v_j \in \mathcal{N}(f)\,.$$
   Por lo tanto, existen r escalares β_j tales que
$$v - \sum_{j=1}^{s} \alpha_j v_j = \sum_{j=1}^{r} \beta_j u_j.$$
   Entonces, se verifica que
$$v = \sum_{j=1}^{s} \alpha_j v_j + \sum_{j=1}^{r} \beta_j u_j.$$
   Por lo tanto $v \in \langle T\rangle$; es decir, T genera V.

   Ahora probemos que T es linealmente independiente. Para ello, sean α_j s escalares y β_j r escalares tales que
$$\sum_{j=1}^{r} \alpha_j v_j + \sum_{j=1}^{s} \beta_j u_j = 0. \tag{5.3}$$

Entonces:

$$\begin{aligned} f\left(\sum_{j=1}^{r}\alpha_j v_j+\sum_{j=1}^{s}\beta_j u_j\right) &= \sum_{j=1}^{r}\alpha_j f(v_j)+\sum_{j=1}^{s}\beta_j f(u_j)\\ &= \sum_{j=1}^{r}\alpha_j w_j = 0, \end{aligned}$$

ya que $u_j \in \mathcal{N}(f)$.

Como B es linealmente independiente, tenemos que los $\alpha_j = 0$ para todo $j \in I_s$.

Entonces, la igualdad 5.3 se expresa de la siguiente manera:

$$\sum_{j=1}^{s}\beta_j u_j = 0.$$

Pero C es linealmente independiente. Entonces, esta última igualdad implica que $\beta_j = 0$ para todo $j \in I_r$, lo que termina por demostrar que T es linealmente independiente y, por lo tanto, es una base para V.

□

Teorema 5.17 Sean V y W dos espacios vectoriales de dimensión n y $f \in \mathscr{L}(V, W)$. Entonces f es inyectiva si y solo si es sobreyectiva.

Demostración.

1. Supongamos que f es inyectiva. Entonces $\dim(V) = \dim(\mathrm{Im}(f))$, ya que $\mathcal{N}(f) = \{0\}$. Como $\dim W = n$, entonces $\dim(\mathrm{Im}(f)) = n$, de donde $\mathrm{Im}(f) = W$; es decir, f es sobreyectiva.

2. Ahora supongamos que f es sobreyectiva. Por el teorema de la dimensión, esto significa que $\dim(\mathcal{N}(f)) = 0$, de donde $\mathcal{N}(f) = \{0\}$; es decir, f es inyectiva. □

Teorema 5.18 Sea $f \in \mathscr{L}(V, W)$ con V y W espacios de dimensión finita. Entonces, si f es biyectiva, entonces $\dim(V) = \dim(W)$.

Demostración. Como f es inyectiva, $\mathcal{N}(f) = \{0\}$, de donde, por el teorema de la dimensión tenemos que

$$\dim(V) = \dim(\mathrm{Im}(f)).$$

Como f es sobreyectiva, $\mathrm{Im}(f) = W$; entonces, $\dim(V) = \dim(W)$. □

Ejemplo 5.79

Sea $f \in \mathscr{L}(V, V)$ tal que
$$f^2 - 2f + I_V = 0,$$
donde I_V es la identidad en V y $f \neq I_V$. Demostrar que existe $u \in V$ tal que $u \neq 0$ y $f(u) = u$. El vector u se denomina *punto fijo* de f.

Solución. Sea $v \in V$. Entonces
$$(f^2 - 2f + I_V)(v) = 0.$$
Por lo tanto:
$$\begin{aligned} f(f(v)) - 2f(v) + I_V(v) &= f(f(v)) - f(v) - f(v) + v \\ &= [f(f(v)) - f(v)] - [f(v) - v] \\ &= f(f(v) - v) - [f(v) - v] = 0. \end{aligned}$$
De modo que
$$f(f(v) - v) = f(v) - v.$$
Como $f \neq I_V$, existe $w \in V$ tal que $f(w) \neq w$. Sea, entonces, $u = f(w) - w$. Esto significa que $u \neq 0$ y que
$$f(u) = f(f(w) - w) = f(w) - w = u.$$
♦

Ejemplo 5.80

Sean f y g elementos de $\mathscr{L}(V, V)$. Demuestre que:

1. Si $f = f \circ g$ y $g = g \circ f$, entonces $\mathcal{N}(f) = \mathcal{N}(g)$.
2. Si $f = f \circ f$, entonces $\mathcal{N}(I_V - f) = \operatorname{Im}(f)$.
3. $\mathcal{N}(f) \cap \operatorname{Im}(f) = \{0\}$ si y solo si $f(f(v)) = 0 \Rightarrow f(v) = 0$.

Solución.

1. Sea $v \in \mathcal{N}(f)$. Entonces $f(v) = 0$. Por lo tanto:
$$g(v) = g(f(v)) = g(0) = 0.$$
Es decir, $v \in \mathcal{N}(g)$.
Hemos demostrado, entonces, que $\mathcal{N}(f) \subseteq \mathcal{N}(g)$. El mismo razonamiento, pero aplicando la hipótesis $f = f \circ g$ nos permite establecer que $\mathcal{N}(g) \subseteq \mathcal{N}(f)$.
2. Sea $v \in \mathcal{N}(I_V - f)$. Entonces $(I_v - f)(v) = 0$; es decir, se verifica que $f(v) = v$. Por lo tanto, $v \in \operatorname{Im}(f)$.
Recíprocamente, si $v \in \operatorname{Im}(f)$, existe $u \in V$ tal que $v = f(u)$. Entonces
$$f(v) = f(f(u)) = (f \circ f)(u) = f(u) = v.$$
Por lo tanto, $f(v) = I_V(v)$, de donde $v \in \mathcal{N}(I_V - f)$.

3. (a) Supongamos que $\mathcal{N}(f) \cap \operatorname{Im}(f) = \{0\}$. Sea $v \in V$ tal que $f(f(v)) = 0$. Entonces $f(v) \in \mathcal{N}(f)$. Pero $f(v) \in \operatorname{Im}(f)$; es decir, $v \in \mathcal{N}(f) \cap \operatorname{Im}(f)$. Como el único elemento común entre $\mathcal{N}(f)$ y $\operatorname{Im}(f)$ es el vector 0, entonces, necesariamente, $f(v) = 0$.

   (b) Supongamos que $f(f(v)) = 0 \Rightarrow f(v) = 0$. Sea $u \in \mathcal{N}(f) \cap \operatorname{Im}(f)$. Vamos a demostrar que $u = 0$.

   Como $u \in \operatorname{Im}(f)$, existe $v \in V$ tal que $f(v) = u$. Por lo tanto, $f(f(v)) = f(u) = 0$, pues $u \in \mathcal{N}(f)$. Entonces, por hipótesis, $f(v) = 0$; es decir, $u = 0$, ya que $u = f(v)$.

♦

Ejemplo 5.81

¿Existe una aplicación lineal f de $\mathbb{R}^2$ en $\mathbb{R}^3$ tal que

$$\mathcal{N}(f) = \langle\{2,1\}\rangle \quad \text{y} \quad \operatorname{Im}(f) = \langle\{1,1,-1\}\rangle\,?$$

En el caso de que exista tal aplicación, describirla. ¿Es única?

Solución. El teorema de la dimensión (teorema 5.16) asegura que si f es lineal, entonces

$$\dim(\mathcal{N}(f)) + \dim(\operatorname{Im}(f)) = \dim(\mathbb{R}^2).$$

Como las dimensiones de los conjuntos $A = \langle\{2,1\}\rangle$ y $B = \langle\{1,1,-1\}\rangle$ son 1 y 1, respectivamente, entonces, en caso de existir tal f, su existencia no contradiría con el teorema de la dimensión. De modo que, podemos intentar encontrar una tal f.

Para la construcción de f, podemos utilizar la idea de la demostración del teorema de la dimensión para encontrar una base del espacio de salida, $\mathbb{R}^2$ a partir de las bases de $\mathcal{N}(f)$ y de $\operatorname{Im}(f)$, respectivamente.

Si el lector lee con atención la demostración citada, encontrará que, si f es una aplicación lineal con núcleo del conjunto A y con imagen el conjunto B, entonces una base para $\mathbb{R}^2$ es la siguiente: $C = \{(2,1),\ (x_1,x_2)\}$, donde $f(x_1,x_2) = (1,1,-1)$. Por lo tanto, para definir f basta elegir (x_1,x_2) tal que C sea linealmente independiente (con lo cual ya será una base de $\mathbb{R}^2$) y asignar $(1,1,-1)$ como imagen de (x_1,x_2) bajo f.

Por ejemplo, (x_1,x_2) puede ser el vector $(1,0)$. Entonces, $f(1,0) = (1,1,-1)$. Ahora, para cualquier (x,y), ya que C es una base de $\mathbb{R}^2$, existen escalares α y β tales que

$$(x,y) = \alpha(1,0) + \beta(2,1).$$

Es fácil mostrar que $\alpha = x - 2y$ y que $\beta = y$.

Ya podemos determinar f:

$$\begin{aligned} f(x,y) &= f(\alpha(1,0) + \beta(2,1)) = \alpha f(1,0) + \beta f(2,1) \\ &= \alpha(1,1,-1) + \beta \cdot 0 = (x-2y)(1,1,-1) \\ &= (x-2y, x-2y, 2y-x). \end{aligned}$$

El lector puede comprobar fácilmente que f es lineal y que su núcleo e imagen son los indicados en el enunciado del ejemplo. Además puede constatar que f no es única, porque depende de el elemento que se elija para definir la base C; por ejemplo, determine el lector la aplicación f que se obtiene si se elige $(0,1)$ para obtener C. ♦

Aplicaciones lineales y matrices | 6

En este capítulo, los elementos de $\mathbb{K}^n$ serán vistos tanto como vectores que se escriben horizontalmente:

$$X = (x_1, x_2, \ldots, x_n)$$

o como matrices de n filas y una columna:

$$X = (x_1, x_2, \ldots, x_n)^t = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

6.1 Aplicación lineal asociada a una matriz

Teorema 6.1 Dada la matriz $A = (a_{ij}) \in \mathbb{M}_{m\times n}$, la aplicación f_A definida por

$$\begin{aligned} f_A\colon \mathbb{K}^n &\longrightarrow \mathbb{K}^m \\ X &\longmapsto f_A(X) = AX \end{aligned}$$

es una aplicación lineal; es decir, $f_A \in \mathscr{L}(\mathbb{K}^n, \mathbb{K}^m)$. Esta aplicación se denomina **aplicación lineal asociada a la matriz** A.

Demostración. Las propiedades de espacio vectorial del espacio de matrices nos permite probar que f_A es lineal. Para ello, sean X y Y en $\mathbb{K}^n$ y α en $\mathbb{K}$. Entonces:

$$\begin{aligned} f_A(\alpha X + Y) &= A(\alpha X + Y) \\ &= A(\alpha X) + AY \\ &= \alpha AX + AY = \alpha f_A(X) + f_A(Y). \end{aligned}$$ □

Ejemplo 6.82 Dada la matriz

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

la aplicación lineal asociada es $f_A\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ tal que

$$f_A(x,y,z) = A\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Puesto que

$$AX = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x-y+z \\ 2x-y \end{pmatrix},$$

tenemos que

$$f_A(x,y,z) = (x-y+z, 2x-y).$$

♦

Teorema 6.2 Sean A y B dos matrices de $\mathbb{M}_{m\times n}$. Entonces $f_A = f_B$ si y solo si $A = B$

Demostración. Tenemos $f_A = f_B$ si y solo si para todo $X \in \mathbb{K}^n$, se verifica que

$$f_A(X) = AX = f_B(X) = BX.$$

Pero esta igualdad se verifica si y solo si para todo $X \in \mathbb{K}^n$, se verifica que $(A-B)X = 0$. A su vez, dado que esta última igualdad es verdadera para todo $X \in \mathbb{K}$, se tiene que $(A-B) = 0$, lo que es equivalente a $A = B$. □

La aplicación lineal asociada a una matriz permite expresar la solución de un sistema de ecuaciones lineales en términos de aplicaciones lineales. En efecto, el conjunto solución del sistema homogéneo

$$AX = 0,$$

se expresa como el conjunto de todos los $X \in \mathbb{K}^n$ tales que

$$f_A(X) = 0.$$

Es decir, el conjunto solución del sistema homogéneo no es más que el núcleo de la aplicación lineal asociada a la matriz de coeficientes del sistema lineal homogéneo.

6.2 Matriz asociada a una aplicación lineal

Este estudio se va a dividir en dos casos: cuando la aplicación lineal es de $\mathbb{K}^n$ en $\mathbb{K}^m$ y cuando es entre dos espacios vectoriales de dimensión finita.

Teorema 6.3 Sea $f \in \mathscr{L}(\mathbb{K}^n, \mathbb{K}^m)$. Entonces existe una matriz $A = (a_{ij}) \in \mathbb{M}_{m\times n}$ tal que

$$f(x) = Ax^t$$

para todo $x \in \mathbb{K}^n$.

Demostración. Sean $C_n = \{e_j : 1 \leq j \leq n\}$ y $C_m = \{f_j : 1 \leq j \leq m\}$ las bases ordenadas canónicas de $\mathbb{K}^n$ y $\mathbb{K}^m$ respectivamente. Entonces, para todo $x \in \mathbb{K}^n$, tenemos que

$$x = \sum_{j=1}^{n} x_j e_j.$$

Por lo tanto, de la linealidad de f, tenemos que

$$f(x) = \sum_{j=1}^{n} x_j f(e_j).$$

Pero, como $f(e_j) \in \mathbb{K}^m$, existen escalares a_{ij} tales que

$$f(e_i) = \sum_{i=1}^{m} a_{ij} f_i.$$

Por lo tanto:

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

Por lo tanto:

$$\begin{aligned} f(x) &= \left(\sum_{j=1}^{n} a_{ij} x_j \right) \\ &= Ax^t, \end{aligned}$$

donde

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

Las columnas de la matriz A son, entonces, las coordenadas de las imágenes de los elementos de la base canónica de $\mathbb{K}^n$ bajo f respecto de la base canónica de $\mathbb{K}^m$. $\square$

Definición 6.1 (Matriz asociada a una aplicación lineal) La matriz A se denomina **matriz asociada a la aplicación lineal f respecto de las bases canónicas C_m y C_n**. Se la representa así:

$$A = [f]_{C_m}^{C_n}.$$

Si las bases están sobre entendidas, entonces se escribirá simplemente $[f]$ para indicar la matriz A.

Ejemplo 6.83 Sea $f \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^2)$ definida por

$$f(x, y, z) = (x + z, y - 2z).$$

Vamos a encontrar la matriz asociada a f respecto de las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente.

Recordemos que los conjuntos

$$B_1 = \{(1,0,0),\ (0,1,0),\ (0,0,1)\} \quad \text{y} \quad B_1 = \{(1,0),\ (0,1),\}$$

son las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente. De la demostración del teorema 6.3, la matriz asociada a f es de orden 2×3 y tiene por columnas las coordenadas de las imágenes de los vectores de la base canónica de $\mathbb{R}^3$ respecto de la base canónica de $\mathbb{R}^2$. Por ello, el procedimiento que vas a seguir es el siguiente:

1. Calcular $f(C_1)$, la imagen de cada uno de los elementos de la base canónica de $\mathbb{R}^3$.
2. Calcular las coordenadas de cada elemento de $f(C_1)$ respecto de C_2.

Procedamos.

1. $f(1,0,0) = (1+0, 0-0) = (1,0)$, $f(0,1,0) = (0, 1-0) = (0,1)$ y $f(0,0,1) = (0+1, 0-2) = (1,-2)$.
2. $[f(1,0,0)]_{C_2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $[f(,1,0)]_{C_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ y $[f(,0,1)]_{C_2} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

Por lo tanto, la matriz asociada a f es:

$$A = [f]_{C_2}^{C_1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix}.$$

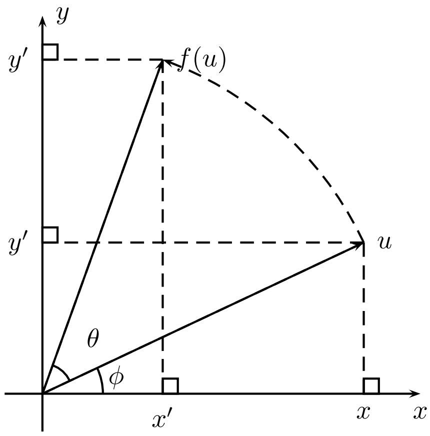
Por la construcción de A, debe verificarse que $f(x) = AX^t$, lo que sucede efectivamente:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 0 + z \\ 0 + y - 2z \end{pmatrix} = \begin{pmatrix} x + z \\ y - 2z \end{pmatrix}.$$

♦

Ejemplo 6.84 Supongamos que $f \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^2)$ representa la rotación de un vector u un ángulo θ. Encontremos la matriz asociada a esta rotación respecto de la base canónica de $\mathbb{R}^2$.

Para ello, encontremos la ley de asignación de f en primer lugar. En el dibujo, está la información necesaria para calcular $f(u)$ dado u. Tenemos, entonces que si $u = (x, y)$, entonces



$$x = ||u|| \cos\phi \quad \text{y} \quad y = ||u|| \operatorname{sen}\phi.$$

Además, si $(x', y') = f(u)$, entonces

$$x' = ||f(u)|| \cos(\phi + \theta) \quad \text{y} \quad y' = ||f(u)|| \operatorname{sen}(\phi + \theta).$$

Por lo tanto,

$$x' = ||f(u)||[\cos\phi\cos\theta - \operatorname{sen}\phi\operatorname{sen}\theta]$$

y

$$y' = ||f(u)||[\operatorname{sen}\phi\cos\theta + \operatorname{sen}\theta\cos\phi].$$

Por otro lado, se tiene que $||f(u)|| = ||u||$. Entonces:

$$x' = [||u||\cos\phi]\cos\theta - [||u||\operatorname{sen}\phi]\operatorname{sen}\theta \quad \text{y} \quad y' = [||u||\operatorname{sen}\phi]\cos\theta + [||u||\cos\phi]\operatorname{sen}\theta,$$

de donde obtenemos que

$$x' = x\cos\theta - y\operatorname{sen}\theta \quad \text{y} \quad y' = y\cos\theta + x\operatorname{sen}\theta.$$

Por lo tanto, para cada $(x, y) \in \mathbb{R}^2$, f se define así:

$$f(x, y) = (x\cos\theta - y\operatorname{sen}\theta,\ y\cos\theta + x\operatorname{sen}\theta).$$

Ahora bien, calculemos las coordenadas de $f(1, 0)$ y $f(0, 1)$ respecto a la base canónica. Puesto que

$$\begin{aligned} f(1, 0) &= (\cos\theta - 0, 0 + \operatorname{sen}\theta) = (\cos\theta, \operatorname{sen}\theta) \\ f(0, 1) &= (0 - \operatorname{sen}\theta, \cos\theta + 0) = (-\operatorname{sen}\theta, \cos\theta), \end{aligned}$$

tenemos que:

$$[f(1, 0)] = (\cos\theta, \operatorname{sen}\theta) \quad \text{y} \quad [f(0, 1)] = (-\operatorname{sen}\theta, \cos\theta).$$

Por lo tanto, la matriz asociada a f es:

$$A = [f] = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Por ejemplo, el vector que resulta de rotar 45 grados al vector $(2, 1)$ tiene por coordenadas

$$\begin{pmatrix} \cos 45 & -\operatorname{sen} 45 \\ \operatorname{sen} 45 & \cos 45 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{2}}{2} \\ \dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{2}}{2} \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \dfrac{\sqrt{2}}{2} \\ \dfrac{3\sqrt{2}}{2} \end{pmatrix}.$$

♦

La matriz asociada a una composición de aplicaciones lineales resulta ser el producto de las matrices asociadas a las aplicaciones que se componen. De manera más precisa, tenemos el siguiente teorema.

Teorema 6.4 Sean $f \in \mathscr{L}(\mathbb{K}^n, \mathbb{K}^m)$ y $g \in \mathscr{L}(\mathbb{K}^m, \mathbb{K}^r)$, entonces $[g \circ f] = [g][f]$.

Demostración. Sean A la matriz asociada a f y B la asociada a g. Entonces:

$$f(x) = Ax \quad \text{y} \quad g(y) = By$$

para todo $x \in \mathbb{K}^n$ y todo $y \in \operatorname{Im}(g) \subset \mathbb{K}^m$. Por lo tanto, para todo $x \in \mathbb{K}^n$, tenemos que:

$$(g \circ f)(x) = g(f(x)) = B(f(x)) = B(Ax) = BAx.$$

es decir, $[g \circ f] = BA = [g][f]$. □

Ahora estudiemos el caso general.

Teorema 6.5 Sean V y W dos espacios vectoriales de dimensión finita sobre un campo $\mathbb{K}$. Sean $B_1 = \{v_j : j \in I_n\}$ y $B_2 = \{w_j : j \in I_m\}$ bases ordenadas de V y W, respectivamente. Sea también $f \in \mathscr{L}(V, W)$. Entonces existe una matriz $A \in \mathbb{M}_{m\times n}$ tal que

$$[f(v)]_{B_2} = A[v]_{B_1}$$

para todo $v \in V$.

Demostración. Sean $v \in V$ y x_j para $j \in I_n$ las coordenadas de v respecto de B_1. Entonces

$$f(v) = \sum_{j=1}^{n} x_j f(v_j).$$

Existen, entonces, escalares α_{ij} tales que

$$f(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i$$

para cada $j \in I_n$.

Por lo tanto:

$$f(v) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} \alpha_{ij} w_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \alpha_{ij} x_j \right) w_i.$$

Tenemos, entonces, que:

$$[f(v)]_{B_2} = \begin{pmatrix} \sum_{j=1}^{n} \alpha_{1j} x_j \\ \sum_{j=1}^{n} \alpha_{2j} x_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{mj} x_j \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 & & \\ x_2 & \vdots & x_n \end{pmatrix}.$$

De manera que:

$$[f(v)]_{B_2} = Av.$$

□

Como se puede ver, la matriz A es la matriz cuyas columnas son las coordenadas de las imágenes de los vectores de la base B_1 bajo f respecto a la base B_2. La matriz A es denominada **matriz asociada a la aplicación lineal f respecto a las bases ordenadas B_1 y B_2**. Se la suele representar con $[f]_{B_2}^{B_1}$.

Con esta notación, la propiedad de la matriz asociada se expresa de la siguiente manera:

$$[f(v)]_{B_2} = [f]_{B_2}^{B_1} [v]_{B_1}.$$

Ejemplo 6.85 La aplicación $f \in \mathscr{L}(\mathbb{M}_{2\times 2}, \mathscr{P}_2[x])$ se define por

$$f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + (b-c)x + dx^2.$$

Encontremos la matriz asociada a f respecto a las bases canónicas:

$$B_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \quad \text{y} \quad B_2 = \{1,\ x,\ x^2\}.$$

Para ello, es suficiente con calcular las imágenes de los elementos de la base B_1 bajo f:

$$\begin{aligned}
f\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} &= 1 + (0-0)x + 0x^2 = 1, \\
f\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} &= 0 + (1-0)x + 0x^2 = x, \\
f\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} &= 0 + (0-1)x + 0x^2 = -x, \\
f\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} &= 0 + (0-0)x + 1x^2 = x^2.
\end{aligned}$$

Entonces, las coordenadas de esta cuatro imágenes respecto a la base canónica dan la siguiente matriz:

$$[f] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ahora calculemos la matriz de f, pero respecto de las siguientes bases:

$$B_1 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \quad \text{y} \quad B_2 = \{1-x,\ x^2,\ x+x^2\}.$$

Para ello, debemos obtener las imágenes de los elementos de la base B_1 bajo f y, luego, las coordenadas de estas imágenes respecto de la base B_2. Empecemos con el cálculo de las imágenes:

$$\begin{aligned}
f\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} &= 1 + (1-0)x + 0x^2 = 1 + x, \\
f\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} &= 0 + (-1-1)x + 0x^2 = -2x, \\
f\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} &= 0 + (0-1)x + x^2 = -x + x^2, \\
f\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= 1 + (0-0)x + x^2 = 1 + x^2.
\end{aligned}$$

Ahora encontremos las coordenadas de estos cuatro vectores respecto de la base B_2. Para ello, encontremos una fórmula general para encontrar las coordenadas cualquier elemento de $\mathscr{P}_2[x]$.

Sea $p \in \mathscr{P}_2[x]$ tal que $p(x) = a + bx + cx^2$. Buscamos los escalares α_j tales que

$$a + bx + cx^2 = \alpha_1(1-x) + \alpha_2(x^2) + \alpha_3(x + x^2).$$

De esta igualdad, obtenemos que los escalares α_j deben satisfacer el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} \alpha_1 \qquad\quad &= a \\ -\alpha_1 \qquad + \alpha_3 &= b \\ \alpha_2 + \alpha_3 &= c. \end{aligned}$$

Las siguientes equivalencias:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & a \\ -1 & 0 & 1 & b \\ 0 & 1 & 1 & c \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & a \\ 0 & 0 & 1 & a+b \\ 0 & 1 & 1 & c \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & a \\ 0 & 0 & 1 & a+b \\ 0 & 1 & 0 & -a-b+c \end{array}\right),$$

implican que $\alpha_1 = a$, $\alpha_2 = a+b$ y $\alpha_3 = -a-b+c$. Entonces:

$$[a + bx + cx^2]_{B_2} = (a,\ a+b,\ -a-b+c).$$

Con ayuda de esta fórmula, obtenemos las coordenadas de las cuatro imágenes calculadas:

$$\begin{aligned} \left[f\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\right]_{B_2} &= [1+x]_{B_2} = (1,\ -2,\ 2), \\ \left[f\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right]_{B_2} &= [-2x]_{B_2} = (0,\ 2,\ -2), \\ \left[f\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}\right]_{B_2} &= [-x+x^2]_{B_2} = (0,\ 2,\ -1), \\ \left[f\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]_{B_2} &= [1+x^2]_{B_2} = (1,\ 0,\ 1). \end{aligned}$$

Y ya podemos obtener la matriz asociada a f respecto de las bases B_1 y B_2:

$$[f]_{B_2}^{B_1} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -2 & 2 & 2 & 0 \\ 2 & -2 & -1 & 1 \end{pmatrix}.$$

Con ayuda de esta matriz podemos calcular las coordenadas del vector

$$\left[f\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\right]_{B_2}.$$

Para ello, usaremos la matriz $[f]_{B_2}^{B_1}$, pues se verifica la igualdad siguiente:

$$[f(x)]_{B_2} = [f]_{B_2}^{B_1}[x]_{B_1}^t.$$

Entonces, lo que primero tenemos que hacer es obtener las coordenadas de x respecto de B_1.

Estas coordenadas son la solución del sistema cuya matriz ampliada es:

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

Y, puesto que

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{array}\right),$$

tenemos que

$$[x]_{B_1} = (\frac{1}{2},\ \frac{1}{2},\ -\frac{1}{2},\ \frac{1}{2}).$$

Por lo tanto:

$$\left[f\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\right]_{B_2} = \begin{pmatrix}1 & 0 & 0 & 1\\ -2 & 2 & 2 & 0\\ 2 & -2 & -1 & 1\end{pmatrix}\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\\ -\frac{1}{2}\\ \frac{1}{2}\end{pmatrix} = \begin{pmatrix}1 & -1 & 1\end{pmatrix}.$$

♦

Definición 6.2 (Rango de una aplicación lineal) Sea $f \in \mathscr{L}(V, W)$ donde V y W tienen dimensión finita. Sea A la matriz asociada a f respecto a dos bases cualesquiera en V y W respectivamente. Entonces, el **rango** de f es la diferencia entre la dimensión de V y el núcleo de f:

$$\operatorname{ran}(f) = \dim(V) - \dim \mathcal{N}(f).$$

Por lo tanto, el rango de f es igual a la dimensión de la dimensión de la imagen de f.

Cuando el espacio de salida y el de llegada sean el mismo y la matriz asociada sea respecto de la misma base en ambos espacios, por ejemplo si B es la base, entonces usaremos la notación $[f]_B$ para la matriz asociada.

Ejemplo 6.86

Sea $f \in \mathscr{L}(\mathbb{C}^2, \mathbb{C}^2)$ definida por

$$f(z_1, z_2) = (\imath z_1,\ (1+\imath)z_2 - z_1).$$

El conjunto

$$B = \{(\imath, 0),\ (0, 1)\}$$

es una base de $\mathbb{C}^2$ sobre $\mathbb{C}$. Calculemos la matriz asociada a f respecto de B.

Solución. En primer lugar, calculemos las imágenes de los elementos de la base B:

$$f(\imath, 0) = (\imath\imath,\ (1+\imath)\cdot 0 - \imath) = (-1,\ -\imath),$$
$$f(0, 1) = (\imath \cdot 0,\ (1+\imath)\cdot 1 - 0) = (0,\ 1+\imath).$$

Ahora bien, las coordenadas de cualquier elemento (z_1, z_2) de $\mathbb{C}^2$ respecto de B se obtienen de la siguiente equivalencia:

$$\left(\begin{array}{cc|c}\imath & 0 & z_1\\ 0 & 1 & z_2\end{array}\right) \sim \left(\begin{array}{cc|c}1 & 0 & (-\imath)z_1\\ 0 & 1 & z_2\end{array}\right).$$

Entonces

$$[(z_1, z_2)]_B = ((-\imath)z_1,\ z_2).$$

En particular, para calcular $[f(\imath, 0)]_B$, hacemos $z_1 = -1$ y $z_2 = -\imath$. Entonces:

$$[f(\imath, 0)]_B = ((-\imath)(-1),\ -\imath) = (\imath,\ -\imath).$$

Y, para calcular $[f(0,1)]_B$, hacemos $z_1 = 0$ y $z_2 = 1 + \imath$. Entonces:

$$[f(0,1)]_B = ((-\imath)(0),\ 1+\imath) = (0,\ 1+\imath).$$

Por lo tanto:

$$[f]_B = \begin{pmatrix} \imath & 0 \\ -\imath & 1+\imath \end{pmatrix}.$$

♦

Ejemplo 6.87

Sea $f \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^3)$ definida por

$$f(x,y,z) = (y-z, x+z, -x+y).$$

El conjunto

$$B = \{(1,1,0),\ (1,0,1),\ (0,1,1)\}$$

es una base para $\mathbb{R}^3$.

1. Calcular la matriz asociada a f respecto de B y C (la primera vista como base del espacio de salid y la segunda, como base del espacio de llegada).
2. Si f es invertible, determinar f^{-1}.

Solución.

1. Las imágenes bajo f de cada elemento de B son:

$$\begin{aligned} f(1,1,0) &= (1,1,0), \\ f(1,0,1) &= (-1,2,-1), \\ f(0,1,1) &= (0,1,1). \end{aligned}$$

Por lo tanto:

$$\begin{aligned} [f(1,1,0)]_C &= (1,1,0), \\ [f(1,0,1)]_C &= (-1,2,-1), \\ [f(0,1,1)]_C &= (0,1,1). \end{aligned}$$

Entonces:

$$[f]_C^B = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & 1 \end{pmatrix}.$$

2. Se tiene que $|[f]_C^B| = 4$. Por lo tanto, la matriz asociada es invertible. Esto significa que

$$[f]_C^B [x]_B = 0 \Longrightarrow [x]_B = 0. \tag{6.1}$$

Por lo tanto, si $x \in \mathcal{N}(f)$, entonces $f(x) = 0$, de donde, por la implicación (6.1), todas las coordenadas de x respecto de B son iguales a 0; es decir $x = 0$. Por lo tanto $\mathcal{N}(f) = \{0\}$. Entonces, f es invertible.

Para hallar f^{-1} debemos resolver la ecuación

$$f(x,y,z) = (a,b,c) \tag{6.2}$$

para todo $(a, b, c) \in \mathbb{R}^3$, pues
$$f^{-1}(a, b, c) = (x, y, z).$$

La igualdad (6.2) implica que (x, y, z) deben satisfacer el siguiente sistema de ecuaciones lineales:
$$\begin{aligned} y - z &= a \\ x + \quad + z &= b. \\ -x + y \qquad &= c \end{aligned}$$

Como la solución de este sistema es
$$x = \frac{1}{2}(a + b - c), \quad y = \frac{1}{2}(a + b + c) \quad \text{y} \quad z = \frac{1}{2}(-a + b + c),$$

entonces
$$f^{-1}(a, b, c) = \frac{1}{2}(a + b - c,\ a + b + c,\ -a + b + c).$$

♦

6.3 Matriz de cambio de base

Sea V un espacio vectorial sobre $\mathbb{K}$ tal que $\dim V = n$. Sean $B = \{v_j\}$ y $S = \{w_j\}$ dos bases ordenadas de V. ¿Cuál es la relación entre las coordenadas de cualquier $v \in V$ respecto de la base B y las coordenadas de v respecto de la base S? Este es el problema que pretendemos resolver a continuación.

Sea f la función identidad definida sobre V. Es decir, para todo $v \in V$, se tiene que
$$f(v) = v.$$

Por lo tanto, $f \in \mathscr{L}(V, V)$.

Las coordenadas de v respecto de S pueden ser vistas como las coordenadas de $f(v)$ respecto de S. Entonces, la relación buscada es
$$[v]_S = [f(v)]_S = [f]_S^B [v]_B.$$

En otras palabras, la relación entre ambos conjuntos de coordenadas está dada por la matriz asociada a f respecto de las bases B y S. A esta matriz se la conoce con el nombre de **matriz de cambio de base de S a B**. También se la conoce con el nombre de **matriz de transición de B a S**.

Está claro que si las matrices B y S son iguales, entonces la matriz de cambio de base es la matriz identidad.

Teorema 6.6 Sean $f \in \mathscr{L}(V, W)$ y $g \in \mathscr{L}(W, U)$ tales que $\dim(V) = n$, $\dim(W) = m$ y $\dim(U) = r$. Sean B_1, B_2 y B_3 bases de V, W y U, respectivamente. Entonces
$$[g \circ f]_{B_3}^{B_1} = [g]_{B_3}^{B_2} [f]_{B_2}^{B_1}.$$

Demostración. Sea $v \in V$. Por un lado, tenemos que:

$$[(g \circ f)(v)]_{B_3} = [g \circ f]_{B_3}^{B_1} [v]_{B_1}. \tag{6.3}$$

Por otro lado, como $(g \circ f)(v) = g(f(v))$ y $f(v) \in W$, tenemos que:

$$\begin{aligned}[(g \circ f)(v)]_{B_3} &= [g(f(v))]_{B_3} \\ &= [g]_{B_3}^{B_2} [f]_{B_2}^{B_1} [v]_{B_1} \\ &= \left([g]_{B_3}^{B_2} [f]_{B_2}^{B_1}\right) [v]_{B_1}.\end{aligned}$$

Es decir:

$$[(g \circ f)(v)]_{B_3} = \left([g]_{B_3}^{B_2} [f]_{B_2}^{B_1}\right) [v]_{B_1}$$

para todo $v \in V$. Y esta igualdad, junto con (6.3), nos permite concluir que

$$[g \circ f]_{B_3}^{B_1} = [g]_{B_3}^{B_2} [f]_{B_2}^{B_1}.$$

□

En otras palabras, la matriz asociada a una composición de aplicaciones lineales es igual al producto de las matrices asociadas a las aplicaciones de la composición, multiplicación que tiene el mismo orden que el de la composición.

Corolario 6.7 Sean B y S dos bases de un espacio vectorial V. Entonces la matriz de cambio de base B a S es invertible y su inversa es igual a la matriz de cambio de base de S a B.

Demostración. La matriz de cambio de base de B a S está asociada a la identidad de V. Sea $[I_V]_S^B$ la matriz. Con esta misma notación, $[I_V]_B^S$ representa la matriz de cambio de base de S a B. Ahora bien, como $I_V = I_V \circ I_V$, entonces

$$[I_V]_B^B = [I_V]_S^B [I_V]_B^S \quad \text{y} \quad [I_V]_S^S = [I_V]_B^S [I_V]_S^B.$$

Pero, como

$$[I_V]_B^B = [I_V]_S^S = I,$$

entonces

$$I = [I_V]_S^B [I_V]_B^S = [I_V]_B^S [I_V]_S^B.$$

Lo que demuestra el corolario. □

Corolario 6.8 Sean $f \in \mathscr{L}(V, W)$ un isomorfismo y B y S bases ordenadas de V y W respectivamente. Entonces la matriz asociada a f^{-1} respecto de S y B es la matriz inversa de la matriz asociada a f respecto a B y S.

Demostración. Dado que f es un isomorfismo, existe f^{-1}. Puesto que $f^{-1} \circ f = I_v$, entonces

$$I = [f^{-1}]_B^S [f]_S^B,$$

con lo cual el corolario queda demostrado. □

Ejemplo 6.88

Sea $f \in \mathscr{L}(\mathbb{R}^3, \mathscr{P}_2[x])$ tal que su matriz asociada respecto de las bases canónicas en ambos espacios es

$$[f] = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Encontrar:

1. la aplicación f,
2. el valor de $f(1,-1,2)$, y
3. bases para $\mathcal{N}(f)$ y para $\operatorname{Im}(f)$.

Solución.

1. Dado que
$$[f(u,v,w)] = [f][(u,v,w)]^t,$$
entonces
$$[f(u,v,w)] = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} u-v \\ u+w \\ -v+w \end{pmatrix}.$$
Y, dado que la bases consideradas son la canónicas, tenemos que
$$f(u,v,w) = (u-v) + (u+w)x + (-u+z)x^2$$
para todo $(u,v,w) \in \mathbb{R}^3$.

2. $f(1,-1,2) = (1+1) + (1+2)x + (-1+2)x^2 = 2 + 3x + x^2$.

3. Tenemos que
$$\begin{aligned} \mathcal{N}(f) &= \{(u,v,w) \in \mathbb{R}^3 : f(u,v,w) = 0\} \\ &= \{(u,v,w) \in \mathbb{R}^3 : (u-v) + (u+v)x + (-u+w)x^2 = 0\} \\ &= \{(u,v,w) \in \mathbb{R}^3 : u-v=0,\ u+v=0,\ -u+w=0\} \\ &= \{(u,v,w) \in \mathbb{R}^3 : u=v=w=0\} = \{(0,0,0)\}. \end{aligned}$$
Por lo tanto, la única base posible para $\mathcal{N}(f)$ es el conjunto vacío. Esto significa que la dimensión de $\mathcal{N}(f) = 0$, de donde:
$$\dim \operatorname{Im}(f) = \dim(\mathbb{R}^3) - \dim(\mathcal{N}(f)) = 3.$$
Luego, como $\dim(\mathscr{P}_2[x]) = 3$, $\operatorname{Im}(f) = \mathscr{P}_2[x]$, de donde la base canónica de $\mathscr{P}_2[x]$ es una base de $\operatorname{Im} f$.

♦

Ejemplo 6.89

Los conjuntos

$$S = \{((1,1),\ (-1,1))\} \quad \text{y} \quad T = \{(2,-1),\ (1,2)\}$$

son bases de $\mathbb{R}^2$. Determinar:

1. la matriz de transición de S a T y de T a S; y
2. las coordenadas del vector $v = (-3,4)$ con respecto a la base T si se sabe que las coordenadas de v respecto a la base S son $\frac{1}{2}$ y $\frac{7}{2}$.

Solución.

1. Para encontrar la matriz de transición de S a T, debemos hallar las coordenadas de los elementos de S respecto de T. Para ello, encontremos las coordenadas respecto de T de cualquier elemento de $\mathbb{R}^2$.

   Sea $(a,b) \in \mathbb{R}^2$. De las siguientes equivalencias, podemos obtener las coordenadas de este elemento:

$$\left(\begin{array}{cc|c} 2 & 1 & a \\ -1 & 2 & b \end{array}\right) \sim \left(\begin{array}{cc|c} 0 & 5 & a+2b \\ -1 & 2 & b \end{array}\right) \sim \left(\begin{array}{cc|c} 0 & 1 & \frac{a+2b}{5} \\ 1 & -2 & -b \end{array}\right) \sim \left(\begin{array}{cc|c} 0 & 1 & \frac{a+2b}{5} \\ 1 & 0 & \frac{2a-b}{5} \end{array}\right).$$

   Por lo tanto:

$$[(a,b)]_T = \left(\frac{2a-b}{5},\ \frac{a+2b}{5}\right).$$

   En particular:

$$\begin{aligned} [(1,1)]_T &= \left(\frac{2-1}{5},\ \frac{1+2}{5}\right) = \left(\frac{1}{5},\ \frac{3}{5}\right) \\ [(-1,1)]_T &= \left(\frac{-2-1}{5},\ \frac{-1+2}{5}\right) = \left(-\frac{3}{5},\ \frac{1}{5}\right). \end{aligned}$$

   Tenemos ya la matriz de transición de S a T:

$$[I_{\mathbb{R}^3}]_T^S = \begin{pmatrix} \frac{1}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{1}{5} \end{pmatrix}.$$

2. La matriz de transición de T a S es la matriz inversa de la matriz de transición de S a T. Por lo tanto:

$$\begin{aligned} [I_{\mathbb{R}^3}]_S^T &= \left([I_{\mathbb{R}^3}]_T^S\right)^{-1} \\ &= \begin{pmatrix} \frac{1}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{1}{5} \end{pmatrix}^{-1} \\ &= \left[\frac{1}{5}\begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}\right]^{-1} \\ &= 5\begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}^{-1} \\ &= 5\frac{1}{10}\begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix}. \end{aligned}$$

3. Dado que
$$[(-3,4)]_S = \begin{pmatrix} \frac{1}{2} \\ \frac{7}{2} \end{pmatrix},$$
tenemos que
$$\begin{aligned} [(-3,4)]_B &= \begin{pmatrix} \frac{1}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{1}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{7}{2} \end{pmatrix} \\ &= \begin{pmatrix} -2 \\ 1 \end{pmatrix}. \end{aligned}$$

♦

Ejemplo 6.90

Sea $\rho_\theta \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^2)$ una rotación de θ unidades. Entonces:

1. Una rotación de θ unidades seguida de una de α unidades es igual a una única rotación de $\theta + \alpha$ unidades; es decir:
$$\rho_\alpha \circ \rho_\theta = \rho_{\theta+\rho}.$$
2. Deshacer una rotación de θ unidades es igual a realizar una rotación de $-\theta$ unidades. Es decir:
$$\rho_\theta^{-1} = \rho_{-\theta}.$$

Solución.

1. Dado que $[\rho_\theta \circ \rho_\alpha] = [\rho_\theta][\rho_\alpha]$ y que tenemos que:
$$\begin{aligned} [\rho_\theta \circ \rho_\alpha] &= [\rho_\theta][\rho_\alpha] \\ &= \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\alpha - \operatorname{sen}\theta\operatorname{sen}\alpha & -\cos\theta\operatorname{sen}\alpha - \operatorname{sen}\theta\cos\alpha \\ \operatorname{sen}\theta\cos\alpha + \cos\theta\operatorname{sen}\alpha & \cos\theta\cos\alpha - \operatorname{sen}\theta\operatorname{sen}\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta+\alpha) & -\operatorname{sen}(\theta+\alpha) \\ \operatorname{sen}(\theta+\alpha) & \cos(\theta+\alpha) \end{pmatrix} = [\rho_{\theta+\alpha}]. \end{aligned}$$
Por lo tanto, se verifica que $\rho_\theta \circ \rho_\alpha = \rho_{\theta+\alpha}$.
2. Dado que $[\rho_\theta^{-1}] = [\rho_\theta]^{-1}$ y que:
$$\begin{aligned} [\rho_\theta^{-1}] &= [\rho_\theta]^{-1} \\ &= \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos(-\theta) & \operatorname{sen}(-\theta) \\ -\operatorname{sen}(-\theta) & \cos(-\theta) \end{pmatrix} = [\rho_{-\theta}]. \end{aligned}$$
Por lo tanto, se verifica que $\rho_\theta^{-1} = \rho_{-\theta}$.

♦

Teorema 6.9 Sean V y W dos espacios vectoriales tales que $\dim(V) = n$ y $\dim(W) = m$, $f \in \mathscr{L}(V, W)$, S_1 y S_2 dos bases ordenadas de V, y T_1 y T_2 dos bases ordenadas de W. Entonces:

$$[f]_{T_2}^{S_2} = Q^{-1}[f]_{T_1}^{S_1}P,$$

donde $P = [I_V]_{S_1}^{S_2}$ y $Q = [I_W]_{T_1}^{T_2}$.

Demostración. Con ayuda de siguiente diagrama, podemos visualizar las relaciones existentes entre las coordenadas de v y $f(v)$ según las bases que se elijan en los espacios V y W:

$$\begin{array}{ccc} (V, S_1) & \xrightarrow{f} & (W, T_1) \\ \uparrow I_V & & \downarrow I_W \\ (V, S_2) & \xrightarrow{f} & (W, T_2) \end{array}$$

Por ejemplo, si leemos la segunda línea horizontal, tenemos que para todo $v \in V$:

$$[f(v)]_{T_2} = [f]_{T_2}^{S_2}[v]_{S_2}. \tag{6.4}$$

Ahora, si leemos la columna de la derecha, tenemos que

$$[f(v)]_{T_2} = [I_W]_{T_2}^{T_1}[f(v)]_{T_1}. \tag{6.5}$$

En cambio, la fila superior nos dice que

$$[f(v)]_{T_1} = [f]_{T_1}^{S_1}[v]_{S_1}. \tag{6.6}$$

Entonces, de las igualdades (6.6) y (6.5), obtenemos que

$$[f(v)]_{T_2} = [I_W]_{T_2}^{T_1}[f]_{T_1}^{S_1}[v]_{S_1}. \tag{6.7}$$

Por lo tanto, de las igualdades (6.4) y (6.7), resulta que

$$[f]_{T_2}^{S_2}[v]_{S_2} = [I_W]_{T_2}^{T_1}[f]_{T_1}^{S_1}[v]_{S_1}. \tag{6.8}$$

Ahora, la lectura de la columna de la izquierda nos informa que

$$[v]_{S_1} = [I_V]_{S_1}^{S_2}[v]_{S_2},$$

que al reemplazar el lado derecho de esta igualdad en (6.8) nos permite proceder de la siguiente manera:

$$\begin{aligned} [f]_{T_2}^{S_2}[v]_{S_2} &= [I_W]_{T_2}^{T_1}[f]_{T_1}^{S_1}[v]_{S_1} \\ &= [I_W]_{T_2}^{T_1}[f]_{T_1}^{S_1}[I_V]_{S_1}^{S_2}[v]_{S_2} \\ &= \left([I_W]_{T_1}^{T_2}\right)^{-1}[f]_{T_1}^{S_1}[I_V]_{S_1}^{S_2}[v]_{S_2}. \end{aligned}$$

Por lo tanto:

$$[f]_{T_2}^{S_2} = \left([I_W]_{T_1}^{T_2}\right)^{-1}[f]_{T_1}^{S_1}[I_V]_{S_1}^{S_2},$$

que es lo que queríamos demostrar. □

Ejemplo 6.91

Dada

$$[f]_S^T = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix},$$

donde

$$B = \{(1,1,0),\ (1,0,1),\ (0,1,1)\}$$

y S es una base de $\mathbb{R}^3$.

1. Determinar $[I_{\mathbb{R}^3}]_B^C$, donde C es la base canónica de $\mathbb{R}^3$.
2. Determinar $[f]_S^C$.

Solución. En primer lugar, del teorema (6.9), tenemos que

$$[f]_S^B = [I_{\mathbb{R}^3}]_S^S [f]_S^B [I_{\mathbb{R}^3}]_B^C = [f]_S^B [I_{\mathbb{R}^3}]_B^C, \tag{6.9}$$

ya que $[I_{\mathbb{R}^3}]_S^S = I$.

1. Como C es la base canónica, la matriz $[I_{\mathbb{R}^3}]_C^B$ tiene por columnas las coordenadas de los elementos de B respecto de C. Por lo tanto:

   $$[I_{\mathbb{R}^3}]_C^B = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

   Entonces

   $$[I_{\mathbb{R}^3}]_B^C = \left([I_{\mathbb{R}^3}]_C^B\right)^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

2. Ahora ya podemos calcular $[f]_S^C$ a partir de la igualdad (6.9):

   $$\begin{aligned} [f]_S^C &= [f]_S^B [I_{\mathbb{R}^3}]_B^C \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \cdot \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}. \end{aligned}$$

♦

Corolario 6.10 Sean $f \in \mathscr{L}(V,V)$, B_1 y B_2 dos bases de V. Entonces existe una matriz inversible P tal que

$$[f]_{B_2}^{B_2} = P^{-1}[f]_{B_1}^{B_1} P.$$

Demostración. Del teorema 6.9, tenemos que:

$$[f]_{B_2}^{B_2} = [I_V]_{B_1}^{B_2}[f]_{B_1}^{B_1}[I_V]_{B_2}^{B_1}.$$

Sea $P = [I_V]_{B_2}^{B_1}$. Entonces $P^{-1} = [I_V]_{B_1}^{B_2}$. Por lo tanto:

$$[f]_{B_2}^{B_2} = P^{-1}[f]_{B_1}^{B_1}P.$$ □

El propósito de representar mediante una matriz una aplicación lineal es trabajar con la matriz en lugar de la aplicación. En el siguiente capítulo, se estudia cómo obtener una matriz que tiene propiedades que hacen el “trabajo” más sencillo que hacerlo directamente con la aplicación.

Ejemplo 6.92 Sean $f \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^2)$ definida por

$$f(x, y) = (x + 6y, 3x + 4y)$$

y el conjunto

$$B = \{(2, -1),\ (1, 1)\}$$

una base de $\mathbb{R}^2$. Encontremos las matrices asociadas a f respecto, en primer lugar, de la base canónica tanto en la salida como en la llegada y, en segundo lugar, respecto de la B tanto en la salida como en la llegada.

1. Puesto que

$$\begin{pmatrix} x+6y \\ 3x+4y \end{pmatrix} = \begin{pmatrix} 1 & 6 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

   tenemos que

$$[f] = \begin{pmatrix} 1 & 6 \\ 3 & 4 \end{pmatrix}.$$

2. Dado que

$$\begin{aligned} f(2,-1) &= (-4, 2) = (-2)(2,-1) + (0)(1,1) \\ f(1,1) &= (7,7) = 0(2,-1) + 7(1,1). \end{aligned}$$

   tenemos que

$$[f]_B^B = \begin{pmatrix} -2 & 0 \\ 0 & 7 \end{pmatrix}.$$

♦

En este último ejemplo, podemos ver que las dos matrices que representan a f, la segunda es diagonal. Esto la hace “más sencilla” en el sentido de que en las operaciones que se hagan con ella, el trabajo requiere menos esfuerzo (por la presencia de los ceros, excepto en la diagonal).

Definición 6.3 (Matrices semejantes) Dos matrices cuadradas A y B de orden n son **semejantes** o **similares** si existe una matriz P inversible tal que $B = P^{-1}AP$.

De esta definición, es inmediato que dos matrices asociadas a una misma aplicación lineal f de un espacio vectorial en sí mismo son siempre semejantes, gracias al corolario 6.10 en la página 197.

Definición 6.4 (Diagonizable) Sean V un espacio vectorial de dimensión finita y $f \in \mathscr{L}(V, V)$. Una base B de V **diagonaliza** a f si y solo si la matriz asociada a f respecto a B, es decir, la matriz $[f]_B^B$, es una matriz diagonal. En este caso, se dice que la aplicación lineal f es **diagonizable**.

No toda matriz es diagonizable. En el siguiente capítulo, se establecerán condiciones necesarias y suficientes para que una matriz sea diagonizable.

Teorema 6.11 Sean V un espacio vectorial real de dimensión n, $B_1 = \{u_j\}$ y $B_2 = \{v_j\}$ dos bases ortonormales de V. Entonces la matriz $[I_V]_{B_2}^{B_1}$ es ortogonal.

Demostración. Sea $P = [I_V]_{B_2}^{B_1}$. Debemos probar que $P^{-1} = P^t$. Para ello, calculemos en primer lugar P y luego P^{-1}.

1. Dado que B_2 es una base ortonormal de V, para cada j, podemos escribir u_j de la siguiente manera:
$$u_j = \sum_{i=1}^{n} \langle u_j, v_i \rangle \, v_i.$$
Por lo tanto, las coordenadas de u_j respecto de B_2 son:
$$[u_j]_{B_2} = (\langle u_j, v_1 \rangle, \langle u_j, v_2 \rangle, \ldots, \langle u_j, v_i \rangle).$$
Entonces:
$$[I_V]_{B_2}^{B_1} = \begin{pmatrix} \langle u_1, v_1 \rangle & \langle u_2, v_1 \rangle & \ldots & \langle u_n, v_1 \rangle \\ \langle u_1, v_2 \rangle & \langle u_2, v_2 \rangle & \ldots & \langle u_n, v_2 \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle u_1, v_n \rangle & \langle u_2, v_n \rangle & \ldots & \langle u_n, v_n \rangle \end{pmatrix}.$$
En otras palabras, el término general de la matriz P es $\langle u_j, v_i \rangle$.

2. Recordemos que la inversa de $[I_V]_{B_2}^{B_1}$ es $[I_V]_{B_1}^{B_2}$ (ver el corolario 6.7 en la página 192). Ahora bien, tenemos que
$$P = [I_V]_{B_2}^{B_1} = (\langle u_j, v_i \rangle).$$
Si intercambiamos B_2 con B_1 en el cálculo de P, obtenemos que
$$P^{-1} = [I_V]_{B_1}^{B_2} = (\langle v_j, u_i \rangle).$$

Ahora solo falta probar que $P^{-1} = P^t$. Pero esto es inmediato, ya que
$$\langle u_j, v_i \rangle = \overline{\langle v_i, u_j \rangle} = \langle v_i, u_j \rangle,$$
ya que V es un espacio vectorial sobre $\mathbb{R}$. Por lo tanto, la inversa de P es la transpuesta de P. □

Ejemplo 6.93

Sean

$$B_1 = \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \right\} \quad \text{y} \quad B_2 = \left\{ \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \end{pmatrix}, \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \end{pmatrix} \right\}$$

dos bases ortonormales de $\mathbb{R}^2$. Encontrar la matriz de transición de B_1 a B_2.

Solución. Sean u_k los elementos de B_1 y v_k los de B_2. Entonces, por el teorema 6.11, la matriz de transición de B_1 a B_2 es:

$$[I_{\mathbb{R}^2}]_{B_2}^{B_1} = (\langle u_j, v_i \rangle).$$

Así que, calculemos todos los productos escalares $\langle u_j, v_i \rangle$:

$$\begin{aligned}
\langle u_1, v_1 \rangle &= \left\langle \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \end{pmatrix} \right\rangle = \frac{7}{5\sqrt{2}}, \\
\langle u_1, v_2 \rangle &= \left\langle \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \end{pmatrix} \right\rangle = -\frac{1}{5\sqrt{2}}, \\
\langle u_2, v_1 \rangle &= \left\langle \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \end{pmatrix} \right\rangle = -\frac{1}{5\sqrt{2}}, \\
\langle u_2, v_2 \rangle &= \left\langle \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \end{pmatrix} \right\rangle = -\frac{7}{5\sqrt{2}}.
\end{aligned}$$

Por lo tanto:

$$[I_{\mathbb{R}^2}]_{B_2}^{B_1} = \frac{1}{5\sqrt{2}} \begin{pmatrix} 7 & -1 \\ -1 & -7 \end{pmatrix}.$$

♦

El siguiente ejemplo utiliza la siguiente propiedad general del producto escalar:

Popiedad 6.1 Sean A una matriz de orden n por m y u y v elementos de $\mathbb{R}^n$. Entonces

$$\langle Av, u \rangle = \langle v, A^t u \rangle .$$

Demostración. La igualdad es fácilmente obtenida de la siguiente manera:

$$\begin{aligned}
\langle Av, u \rangle &= (Av)^t u \\
&= (v^t A^t) u \\
&= v(A^t u) = \langle v, A^t u \rangle .
\end{aligned}$$

□

Ejemplo 6.94

Sea $f \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n)$ tal que

$$\|f(u)\| = \|u\|$$

para todo $u \in \mathbb{R}^n$. Demuestre que:

1. Para todo $u \in \mathbb{R}^n$ y todo $v \in \mathbb{R}^n$, se verifica que
$$\langle f(u), f(v) \rangle = \langle u, v \rangle .)$$
2. La matriz $[f]$ es ortogonal.

Solución.

1. Dado que
$$\|f(u-v)\| = \|u-v\|,$$
y que f es lineal, tenemos que
$$\|f(u) - f(v)\|^2 = \|u-v\|^2. \tag{6.10}$$
Por otro lado, tenemos que
$$\begin{aligned}\|f(u) - f(v)\|^2 &= \langle f(u) - f(v), f(u) - f(v)\rangle \\ &= \langle f(u), f(u)\rangle - 2\langle f(u), f(v)\rangle + \langle f(v), f(v)\rangle \\ &= \|f(u)\|^2 - 2\langle f(u), f(v)\rangle + \|f(v)\|^2.\end{aligned}$$
De donde, concluimos que:
$$\|f(u) - f(v)\|^2 = \|f(u)\|^2 - 2\langle f(u), f(v)\rangle + \|f(v)\|^2. \tag{6.11}$$
Por otro lado, también tenemos que
$$\begin{aligned}\|u-v\|^2 &= \langle u-v, u-v\rangle \\ &= \langle u, u\rangle - 2\langle u, v\rangle + \langle v, v\rangle,\end{aligned}$$
de donde obtenemos que:
$$\|u-v\| = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2. \tag{6.12}$$
Finalmente, la propiedad transitiva aplicada a las igualdades (6.10), 6.11 y 6.12 muestra que
$$\|f(u)\|^2 - 2\langle f(u), f(v)\rangle + \|f(v)\|^2 = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2,$$
de donde obtenemos que
$$-2\langle f(u), f(v)\rangle = -2\langle u, v\rangle;$$
es decir:
$$\langle f(u), f(v)\rangle = \langle u, v\rangle,$$
que es lo que queríamos demostrar.

2. Sea $A = [f]$. Debemos probar que $A^{-1} = A^t$. Para ello, vamos a utilizar la propiedad 6.1 y la igualdad establecida en el numeral anterior.

   Sea $v \in \mathbb{R}^n$. Entonces
$$f(v) = Av,$$
y, dado que $\langle f(u), f(v)\rangle = \langle u, v\rangle$, se verifica la siguiente igualdad:
$$\langle Au, Av\rangle = \langle u, v\rangle.$$
Ahora bien, por la propiedad 6.1, la última igualdad implica que
$$\left\langle u, A^t Av\right\rangle = \langle u, v\rangle,$$
de donde
$$\left\langle u, A^t Av - v\right\rangle = 0$$
para todo $u \in \mathbb{R}^n$.

Esto significa que $A^tAv - v \in \{\mathbb{R}^n\}^\perp = \{0\}$, de donde

$$A^tAv - v = 0$$

para todo $v \in \mathbb{R}^n$; es decir:

$$A^tAv = v.$$

Por lo tanto, $A^tA = I$, de donde $A^{-1} = A^t$.

♦

Ejemplo 6.95

Sea $f \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^3)$ tal que

$$f(1,1,0) = (1,1,-1), \quad f(1,0,1) = (1,-1,1), \quad \text{y} \quad f(0,1,1) = (-1,1,1).$$

Sean

$$B_1 = \{(1,0,1),\ (0,1,1),\ (1,1,1)\} \quad \text{y} \quad B_2 = \{(1,1,1),\ (1,1,0),\ (1,-1,1)\}$$

dos bases de $\mathbb{R}^3$. Determine:

$$[I_{\mathbb{R}^3}]_{B_1}^{B_2}, \quad [f]_{B_1}^{B_1}, \quad [f]_{B_1}^{B_2} \quad \text{y} \quad [f]_{B_2}^{B_1}.$$

Solución. Sea C la base canónica de $\mathbb{R}^n$.

1.

$$\begin{aligned}
[I_{\mathbb{R}^3}]_{B_1}^{B_2} &= [I_{\mathbb{R}^3}]_{B_1}^{C}[I_{\mathbb{R}^3}]_{C}^{B_2} \\
&= \left([I_{\mathbb{R}^3}]_{C}^{B_1}\right)^{-1}[I_{\mathbb{R}^3}]_{C}^{B_2} \\
&= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

Entonces

$$[I_{\mathbb{R}^3}]_{B_1}^{B_2} = \begin{pmatrix} 0 & -1 & 2 \\ 0 & -1 & 0 \\ 1 & 2 & -1 \end{pmatrix}.$$

2.

$$\begin{aligned}
[f]_{B_1}^{B_1} &= [I_{\mathbb{R}^3}]_{B_1}^{C}[f]_{C}^{B_1} \\
&= \left([I_{\mathbb{R}^3}]_{C}^{B_1}\right)^{-1}[f]_{C}^{B_1} \\
&= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 & \frac{1}{2} \\ -1 & 1 & \frac{1}{2} \\ 1 & 1 & \frac{1}{2} \end{pmatrix}.
\end{aligned}$$

Por lo tanto:

$$[f]_{B_1}^{B_1} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & -1 & \frac{1}{2} \end{pmatrix}.$$

3.

$$\begin{aligned} [f]_{B_1}^{B_2} &= [f]_{B_1}^{B_1}[f]_{B_1}^{B_2} \\ &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & -1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & -1 & 2 \\ 0 & -1 & 0 \\ 1 & 2 & -1 \end{pmatrix}. \end{aligned}$$

Por lo tanto:

$$[f]_{B_1}^{B_2} = \begin{pmatrix} 0 & -2 & 4 \\ 0 & -2 & 0 \\ \frac{1}{2} & 3 & -\frac{5}{2} \end{pmatrix}.$$

4.

$$\begin{aligned} [f]_{B_2}^{B_1} &= [I_{\mathbb{R}^3}]_{B_2}^{B_1}[f]_{B_1}^{B_1} \\ &= \begin{pmatrix} 0 & -1 & 2 \\ 0 & -1 & 0 \\ 1 & 2 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & -1 & \frac{1}{2} \end{pmatrix}. \end{aligned}$$

Por lo tanto:

$$[f]_{B_2}^{B_1} = \begin{pmatrix} 0 & 2 & \frac{1}{2} \\ 0 & -2 & 0 \\ 1 & -1 & 0 \end{pmatrix}.$$

♦

Valores y vectores propios | 7

El objetivo general de este capítulo es

> Resolver problemas de diagonalización de matrices a partir de las propiedades de los valores y vectores propios de las matrices, a un nivel productivo.

Para alcanzar el objetivo general se proponen los siguientes objetivos específicos.

1. Identificar los valores y vectores propios de un operador lineal y sus propiedades a partir de las definiciones.
2. Calcular el polinomio característico y los valores y vectores propios de una matriz de orden n.
3. Calcular los valores propios de una matriz simétrica usando sus propiedades a un nivel productivo.
4. Fundamentar la diagonalización de una matriz de orden máximo 4 a partir de las propiedades básicas de los valores y vectores propios, a un nivel reproductivo.
5. Resolver problemas de diagonalización de una matriz de orden máximo 4 a partir de las propiedades básicas de los valores y vectores propios, a un nivel reproductivo.
6. Calcular la matriz ortogonal semejante a una matriz diagonalizable dada de orden máximo 4, a un nivel productivo.

7.1 Introducción

Una de las aplicaciones de las ecuaciones homogéneas se da en el análisis de problemas de ingeniería en los cuales se usan los valores y vectores propios en áreas tales como análisis de vibraciones, análisis de circuitos eléctricos, teoría de elasticidad, etc. Cuando se desarrollan los modelos matemáticos de problemas que aparecen en las áreas indicadas, aparecen sistemas de ecuaciones que generalmente tienen la forma

$$\begin{aligned}
(a_{11}-\lambda)x_1 + \quad a_{12}x_2 + \cdots + \quad a_{1n}x_n &= 0,\\
a_{21}x_1 + (a_{22}-\lambda)x_2 + \cdots + \quad a_{2n}x_n &= 0,\\
&\vdots\\
a_{n1}x_1 + \quad a_{n2}x_2 + \cdots + (a_{nn}-\lambda)x_4 &= 0,
\end{aligned}$$

en donde los coeficientes a_{ij} son números reales, los x_i son las variables del sistema y λ es un parámetro particular del sistema que toma valores desconocidos.

En el modelado matemático de problemas en los cuales se usan las ecuaciones de Lagrange o las leyes de Newton aparecen ecuaciones diferenciales de la forma $X'(t) = AX(t)$. Por ejemplo, en un sistema vibratorio con varias cuerpos sujetos a varios resortes, los valores de los coeficientes a_{ij} dependerán de las masas m_i de los cuerpos y las constantes elásticas k_i de los resortes, las variables x_i serán los desplazamientos de los cuerpos, y los valores λ resultan ser los cuadrados de las frecuencias de oscilación del sistema vibratorio.

Los valores y vectores propios también se usan en la solución de sistemas de ecuaciones diferenciales lineales homogéneas, en estadística, en las cadenas de Markov que se usan para la predicción de eventos futuros o para analizar modelos de crecimiento de una población, en geometría analítica para identificar y graficar cónicas y cuádricas, etc. El problema original radica en modelar el sistema, analizarlo y sacar conclusiones sobre su aplicabilidad.

7.2 Nociones fundamentales

Definición 7.1 (Valor propio y vector propio de una aplicación lineal) Sea f una aplicación lineal de un espacio vectorial V sobre un campo $\mathbb{K}$ en V. Se dice que un vector $v \in V$ no nulo es un **vector propio** de la aplicación lineal f asociado con **valor propio** $\lambda \in \mathbb{K}$ si se cumple que

$$f(v) = \lambda v.$$

En la literatura matemática es común llamar a los valores y vectores propios *valores y vectores característicos*, *autovalores y autovectores*, o *eigenvalores y eigenvectores*, respectivamente.

Ejemplo 7.96 Sea la aplicación lineal $f \in \mathscr{L}(V, V)$ tal que $f(v) = v$. Entonces $\lambda = 1$ es un valor propio de f y todo vector $v \in V$ no nulo es un vector propio de f asociado con valor propio $\lambda = 1$. ♦

Ejemplo 7.97 Sea la aplicación lineal $f \in \mathscr{L}(V, V)$ tal que $f(v) = 0_V$. Entonces $\lambda = 0$ es un valor propio de f y todo vector $v \in V$ no nulo es un vector propio de f asociado con valor propio $\lambda = 0$. ♦

Teorema 7.1 Sean $f \in \mathscr{L}(V, V)$ y $V_\lambda = \{v \in V \mid f(v) = \lambda v\}$. Entonces V_λ es un subespacio vectorial del espacio vectorial V.

Demostración.

1. $0_V \in V_\lambda$ puesto que $f(0_V) = 0_V = \lambda 0_V$ ya que f es aplicación lineal.

2. Por demostrarse que

$$(\forall \alpha \in \mathbb{K})(\forall u, v \in V)\colon \alpha u + v \in V_\lambda.$$

En efecto,

$$\begin{aligned} f(\alpha u + v) &= \alpha f(u) + f(v) \\ &= \alpha(\lambda u) + (\lambda v) \\ &= \lambda(\alpha u + v), \end{aligned}$$

por lo cual $\alpha u + v \in V_\lambda$. □

El subespacio vectorial V_λ es el conjunto de todos los vectores propios de la aplicación lineal f asociados con valor propio λ, unido con el conjunto que contiene el vector nulo 0_V. Al conjunto V_λ se le llama el **subespacio vectorial propio de la aplicación lineal f asociado con valor propio λ**.

Recordemos que a una aplicación lineal se le puede asociar una matriz. Sea $[f]_B^B$ la matriz asociado a la aplicación lineal $f\colon V \to V$ con respecto a la base ordenada B del espacio vectorial V. Sea $[v]_B$ el vector coordenado de un vector $v \in V$ con respecto a la base ordenada B. Entonces podemos escribir

$$[f(v)]_B = [f]_B^B [v]_B,$$

donde $[f(v)]_B$ es el vector coordenado de la imagen del vector v por la aplicación f con respecto a la base ordenada B. Por otro lado el vector coordenado del vector λv con respecto a la base ordenada B es

$$[\lambda v]_B = \lambda [v]_B.$$

Supongamos que v es un vector propio de la aplicación lineal f asociado con valor propio λ. Entonces podemos escribir las siguientes expresiones:

$$\begin{aligned} f(v) &= \lambda v, \\ [f(v)]_B &= [\lambda v]_B, \\ [f]_B^B [v]_B &= \lambda [v]_B. \end{aligned}$$

Si la dimensión del espacio vectorial V es n, entonces $[f]_B^B \in M_n$ y $[v]_B \in M_{n\times 1}$. Notemos $A = [f]_B^B$ y $u = [v]_B$. La expresión $[f]_B^B[v]_B = \lambda[v]_B$ la podemos escribir así:

$$Au = \lambda u.$$

Este resultado muestra que los valores propios de la aplicación lineal f y su matriz asociada son los mismos. Además, motiva la siguiente definición.

Definición 7.2 (Valor propio y vector propio de una matriz) Sea A una matriz cuadrada de orden n sobre el campo de los números complejos. Se dice que un vector $v \in M_{n\times 1}$ no nulo es un **vector propio** de la matriz A asociado con **valor propio** $\lambda \in \mathbb{K}$ si se cumple que

$$Au = \lambda u.$$

Independientemente de que una matriz cuadrada sea o no la representación matricial de alguna aplicación lineal, con la definición anterior podemos hablar de los valores y vectores propios de esa matriz.

Los valores propios y los elementos de los vectores propios de una matriz sobre el campo de los números reales pueden ser números complejos. Esto no es ningún inconveniente pues toda matriz sobre el campo de los números reales es una matriz sobre el campo de los números complejos.

Si A es una matriz cuadrada de orden n diremos que $V_\lambda \subset M_{n\times 1}$ es el subespacio vectorial propio de la matriz A, asociado con valor propio λ.

Teorema 7.2 Sean $f \in \mathscr{L}(V, V)$ y $\lambda \in \mathbb{K}$. Sea A la matriz asociada a f. Las siguientes afirmaciones son equivalentes:

a) λ es un valor propio de A (o de f).

b) La matriz $(A - \lambda I)$ no es inversible (o $f - \lambda I_V$ no es inversible.

c) $|A - \lambda I| = 0$.

Demostración.

1. **a) $\Rightarrow$ b)**: si λ es un valor propio de f, entonces existe un vector v no nulo tal que $f(v) = \lambda v$ que se puede escribir como

   $$(f - \lambda I_V)(v) = 0$$

   y como $v \neq 0$ podemos escribir

   $$\mathcal{N}(f - \lambda I_V) \neq \{0_V\},$$

   de donde concluimos que $f - \lambda I_V$ no es inyectiva y, por lo tanto, no es inversible.

2. **b) $\Rightarrow$ c)**: si $f - \lambda I_V$ no es inversible, entonces $A - \lambda I$ no es inversible, razón por la cual $\det(A - \lambda I) = 0$.

3. **c)** $\Rightarrow$ **a)**: si $\det(A-\lambda I)=0$, entonces existe un vector v no nulo tal que

$$(A-\lambda I)v = O.$$

Esta última expresión es un sistema lineal homogéneo que tiene soluciones no nulas y lo podemos escribir así:

$$\begin{gathered} Av-\lambda I v = O, \\ Av = \lambda v + O, \\ Av = \lambda v. \end{gathered}$$

Esta última expresión muestra que λ es un valor propio de la matriz A o de la aplicación lineal f. $\square$

El teorema anterior permite encontrar los valores propios λ de una aplicación lineal f o de su matriz asociada A por medio de la solución de la ecuación

$$\det(A-\lambda I)=0.$$

Los vectores propios v de la matriz A se encuentran resolviendo el sistema lineal homogéneo

$$(A-\lambda I)v = O.$$

Ejemplo 7.98

Hallar los valores y vectores propios de la matriz

$$A = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix}.$$

Solución. Para hallar los valores propios de la matriz A resolvemos la ecuación $|A-\lambda I|=0$:

$$\begin{gathered} \left| \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0, \\ \begin{vmatrix} 10-\lambda & -18 \\ 6 & -11-\lambda \end{vmatrix} = 0, \\ (10-\lambda)(-11-\lambda)+108 = 0, \\ \lambda^2+\lambda-2=0, \\ (\lambda-1)(\lambda+2)=0. \end{gathered}$$

Luego los valores propios de la matriz A son $\lambda_1=-2$ y $\lambda_2=1$.

Calculemos ahora los vectores propios de la matriz A.

- Sea el valor propio $\lambda_1=-2$. Para hallar los vectores propios v asociados con valor propio $\lambda_1=-2$ resolvemos el sistema lineal

$$\begin{gathered} (A-\lambda_1 I)v = O, \\ \begin{pmatrix} 10-\lambda_1 & -18 \\ 6 & -11-\lambda_1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{gathered}$$

$$\begin{pmatrix} 10-(-2) & -18 \\ 6 & -11-(-2) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$\begin{pmatrix} 12 & -18 \\ 6 & -9 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Para resolver el sistema lineal homogéneo

$$\begin{aligned} 12u_1 - 18u_2 &= 0, \\ 6u_1 - \ \ 9u_2 &= 0, \end{aligned}$$

hallamos su matriz escalonada reducida por filas:

$$(A - \lambda_1 I \mid O) = \begin{pmatrix} 12 & -18 & \mid 0, \\ 6 & -9 & \mid 0, \end{pmatrix} \sim \begin{pmatrix} 2 & -3 & \mid 0, \\ 0 & 0 & \mid 0, \end{pmatrix}.$$

Sea $u_2 = 2r \in \mathbb{R} \setminus \{0\}$. Entonces $u_1 = 3r$ y

$$v = \begin{pmatrix} 3r \\ 2r \end{pmatrix}, \quad r \in \mathbb{R} \setminus \{0\}$$

nos da el conjunto de vectores propios de la matriz A asociados al valor propio $\lambda_1 = -2$. En particular, si $r = 1$ obtenemos

$$v_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix},$$

un vector propio de la matriz A asociado al valor propio $\lambda_1 = -2$.

El subespacio vectorial propio asociado con valor propio $\lambda_1 = -2$ es

$$V_{\lambda_1 = -2} = \left\langle \left\{ \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\} \right\rangle.$$

- Sea el valor propio $\lambda_2 = 1$. Solo indicamos los pasos de obtención del vector propio asociado a λ_2:

$$(A - \lambda_2 I)v = O,$$
$$\begin{pmatrix} 10-\lambda_2 & -18 \\ 6 & -11-\lambda_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$\begin{pmatrix} 9 & -18 \\ 6 & -12 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$(A - \lambda_1 I \mid O) = \begin{pmatrix} 9 & -18 & \mid 0 \\ 6 & -12 & \mid 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & \mid 0 \\ 0 & 0 & \mid 0 \end{pmatrix}.$$

Sea $u_2 = r \in \mathbb{R} \setminus \{0\}$. Entonces $u_1 = 2r$ y

$$v = \begin{pmatrix} 2r \\ r \end{pmatrix}, \quad r \in \mathbb{R} \setminus \{0\}$$

nos da el conjunto de vectores propios de la matriz A asociados al valor propio $\lambda_2 = 1$. En particular, si $r = 1$ obtenemos

$$v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

un vector propio de la matriz A asociado al valor propio $\lambda_2 = 1$.

El subespacio vectorial propio asociado con valor propio $\lambda_2 = 1$ es

$$V_{\lambda_2 = 1} = \left\langle \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\} \right\rangle.$$

Verifiquemos que $\binom{3r}{2r}$, $r \in \mathbb{R} \setminus \{0\}$ da el conjunto de vectores propios de la matriz A asociados con valor propio $\lambda_1 = -2$:

$$Av = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix} \begin{pmatrix} 3r \\ 2r \end{pmatrix} = \begin{pmatrix} -6r \\ -4r \end{pmatrix} = -2 \begin{pmatrix} 3r \\ 2r \end{pmatrix} = \lambda_1 v, \quad \text{para todo } r \in \mathbb{R}.$$

♦

Definición 7.3 (Polinomio y ecuación característicos) Sea la matriz $A \in M_n$. A $P_A(\lambda) = |A - \lambda I|$ se le llama **polinomio característico** de la matriz A. La ecuación $P_A(\lambda) = 0$ recibe el nombre de **ecuación característica** de la matriz A.

Las raíces del polinomio característico son los valores propios de la matriz A.

Teorema 7.3 Los valores propios de la aplicación lineal $f \in \mathscr{L}(V, V)$ no dependen de la elección de la matriz asociada a f.

Demostración. Sean S y T dos bases del espacio vectorial V. Sean $A = [f]_S^S$ y $V = [f]_T^E$ las matrices asociadas a la aplicación lineal f, respectivamente. Si $P = [I_V]_S^T$ es la matriz de cambio de base de T a S, entonces $B = P^{-1}AP$. Por demostrar que $P_B(\lambda) = P_A(\lambda)$.

En efecto, tenemos las siguientes expresiones:

$$\begin{aligned} P_B(\lambda) &= |B - \lambda I| = |P^{-1}AP - \lambda I| \\ &= |P^{-1}AP - \lambda P^{-1}P| = |(P^{-1}A - \lambda P^{-1})P| \\ &= |(P^{-1}A - \lambda P^{-1}I)P| = |(P^{-1}A - P^{-1}\lambda I)P| \\ &= |P^{-1}(A - \lambda I)P| = |P^{-1}|\,|A - \lambda I|\,|P| \\ &= |P^{-1}|\,|P|\,|A - \lambda I| = |P^{-1}P|\,|A - \lambda I| \\ &= |I|\,|A - \lambda I| = 1\,|A - \lambda I| \\ &= |A - \lambda I| \\ &= P_A(\lambda). \end{aligned}$$

Como los polinomios característicos de las matrices A y B son iguales, entonces los valores propios de A y B son los mismos. □

Los valores propios de una matriz asociada a una aplicación lineal f con respecto a cualquier base son los valores propios de la aplicación lineal f.

Estamos en condiciones de delinear el siguiente procedimiento para hallar los valores y vectores propios de una aplicación lineal $f : V \to V$:

1. Hallar la representación matricial de la aplicación lineal con respecto a cualquier base B del espacio vectorial V.

2. Hallar los valores y vectores propios de esa representación matricial de f.

3. Los valores propios de la representación matricial son también los valores propios de la aplicación lineal.

4. Los vectores propios de la representación matricial son los vectores coordenados de los vectores propios de la aplicación lineal.

Ejemplo 7.99

Sea la aplicación lineal $f \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^2)$ tal que $f(x, y) = (-x - 2y, -7x + 8y)$. Hallar los valores y vectores propios de la aplicación lineal f.

Solución. Primero hallamos los valores y vectores propios de una matriz A asociada a f con respecto a cualquier base. Por ser más cómodo escogemos la base canónica $C = \{e_1, e_2\} = \{(1, 0), (0, 1)\}$. No es difícil hallar la matriz

$$A = [f]_C^C = \begin{pmatrix} -1 & 2 \\ -7 & 8 \end{pmatrix}.$$

Hallemos los valores propios de la matriz A:

$$\begin{gathered} |A - \lambda I| = 0, \\ \begin{vmatrix} -1-\lambda & 2 \\ -7 & 8-\lambda \end{vmatrix} = 0, \\ (\lambda - 1)(\lambda - 6) = 0. \end{gathered}$$

Los valores propios de la matriz A son $\lambda_1 = 1$ y $\lambda_2 = 6$.

Hallemos los vectores propios de la matriz A.

- Sea el valor propio $\lambda_1 = 1$.

$$\begin{gathered} (A - \lambda_1 I)v = O, \\ \begin{pmatrix} -1-1 & 2 \\ -7 & 8-1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \begin{pmatrix} -2 & 2 \\ -7 & 7 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{gathered}$$

  El sistema lineal se reduce a la ecuación $u_1 - u_2 = 0$. Sea $u_2 = r$, $r \in \mathbb{R} \setminus \{0\}$. Entonces $u_1 = r$. Si $r = 1$ tenemos que

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

  es un vector propio de la matriz A asociado con valor propio $\lambda_1 = 1$.

- Sea el valor propio $\lambda_2 = 6$.

$$\begin{gathered} (A - \lambda_2 I)v = O, \\ \begin{pmatrix} -1-6 & 2 \\ -7 & 8-6 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \begin{pmatrix} -7 & 2 \\ -7 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{gathered}$$

  El sistema lineal se reduce a la ecuación $-7u_1 + 2u_2 = 0$. Sea $u_2 = 7r$, $r \in \mathbb{R} \setminus \{0\}$. Entonces $u_1 = 2r$. Si $r = 1$ tenemos que

$$v_2 = \begin{pmatrix} 2 \\ 7 \end{pmatrix}$$

  es un vector propio de la matriz A asociado con valor propio $\lambda_2 = 6$.

Los vectores propios v_1 y v_2 de la matriz A son los vectores coordenados de los vectores propios u_1 y u_2, respectivamente, de la aplicación lineal f. Es decir, $v_1 = [w_1]_C$ y $v_2 = [w_1]_C$. Por lo tanto,

$$w_1 = 1 \cdot e_1 + 1 \cdot e_2 = 1(1,0) + 1(0,1) = (1,1) \in \mathbb{R}^2$$

es un vector propio de la aplicación lineal f asociado con valor propio $\lambda_1 = 1$, y

$$w_2 = 2 \cdot e_1 + 7 \cdot e_2 = 2(1,0) + 7(0,1) = (2,7) \in \mathbb{R}^2$$

es un vector propio de la aplicación lineal f asociado con valor propio $\lambda_2 = 6$.

Los conjuntos

$$W_{\lambda_1=1} = \langle\{(1,1)\}\rangle \quad \text{y} \quad W_{\lambda_2=6} = \langle\{(2,7)\}\rangle$$

son los subespacios vectoriales propios de la aplicación lineal f asociados a los valores propios $\lambda_1 = 1$ y $\lambda_2 = 2$, respectivamente. ♦

7.2.1 Polinomio característico

Hay una forma interesante para escribir el polinomio característico de una matriz.

1. Si tenemos la matriz

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

   entonces el polinomio característico de la matriz A puede escribirse como

$$\begin{aligned} p_A(\lambda) &= |A - \lambda I| \\ &= \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} \\ &= \lambda^2 - (a+d)\lambda + (ad - bc) \\ &= \lambda^2 - \operatorname{tr}(A)\lambda + |A|. \end{aligned}$$

2. Si tenemos la matriz

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

   entonces el polinomio característico de la matriz A puede escribirse como

$$\begin{aligned} p_A(\lambda) &= |A - \lambda I| \\ &= \begin{vmatrix} a_1-\lambda & b_1 & c_1 \\ a_2 & b_2-\lambda & c_2 \\ a_3 & b_3 & c_3-\lambda \end{vmatrix} \\ &= -\lambda^3 + \operatorname{tr}(A)\lambda^2 - k\lambda + |A|, \end{aligned}$$

   donde

$$k = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix}.$$

3. Si tenemos la matriz
$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$
entonces el polinomio característico de la matriz A puede escribirse como
$$\begin{aligned} p_A(\lambda) &= |A - \lambda I| \\ &= \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} \\ &= (-1)^n\big(\lambda^n - \sigma_1\lambda^{n-1} + \sigma_2\lambda^{n-2} - \sigma_3\lambda^{n-3} + \cdots + (-1)^n\sigma_n\big), \end{aligned}$$
donde
$$\begin{aligned} \sigma_1 &= \operatorname{tr}(A), \\ \sigma_2 &= \sum_{i<j} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}, \\ \sigma_3 &= \sum_{i<j<r} \begin{vmatrix} a_{ii} & a_{ij} & a_{ir} \\ a_{ji} & a_{jj} & a_{jr} \\ a_{ri} & a_{rj} & a_{rr} \end{vmatrix}, \\ &\vdots \\ \sigma_n &= |A|. \end{aligned}$$
Podemos decir que σ_2 es la suma de todos los menores diagonales de segundo orden de la matriz A, σ_3 es la suma de todos los menores diagonales de tercer orden de la matriz A, etc.

Antes de dar la siguiente definición recordemos que un polinomio de grado n tiene n raíces no necesariamente distintas.

Definición 7.4 (Multiplicidad algebraica y geométrica) Sea $p_A(\lambda)$ el polinomio característico de grado n de la matriz A de orden n. el polinomio tiene n raíces y puede escribirse como
$$p_A(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_m)^{r_m},$$
donde $r_1 + r_2 + \cdots r_m = n$.

1. Se llama **multiplicidad algebraica** del valor propio λ_i al número r_i, $i = 1, 2, \ldots, m$.
2. Se llama **multiplicidad geométrica** del valor propio λ_i a la dimensión del subespacio vectorial V_{λ_i}, $i = 1, 2, \ldots, m$.

En otras palabras, la multiplicidad algebraica de un valor propio es el número de veces que se repite dicho valor propio, y la multiplicidad geométrica es el número de vectores propios linealmente independientes que genera el valor propio. Si la multiplicidad algebraica de un valor propio es 1, entonces la multiplicidad geométrica también es 1. Pero si la multiplicidad algebraica es $p > 1$, $p \leq n$, entonces la multiplicidad geométrica puede ser 1 o 2 o hasta p. Todo valor propio tiene asociado con menos un vector propio.

Ejemplo 7.100

Hallar las multiplicidades algebraica y geométrica de los valores propios de la matriz

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

Solución. Hallemos los valores propios de la matriz A:

$$p_A(\lambda) = \begin{vmatrix} 2-\lambda & -1 & 1 \\ 0 & 3-\lambda & -1 \\ 2 & 1 & 3-\lambda \end{vmatrix} = -\lambda^3 + 8\lambda^2 - 20\lambda + 16 = -(\lambda-2)^2(\lambda-4) = 0.$$

Entonces $\lambda_1 = 2$ y $\lambda_2 = 4$ son los valores propios de la matriz A, de multiplicidades algebraicas 2 y 1, respectivamente. La multiplicidad geométrica de λ_2 es 1.

Hallemos los vectores propios de la matriz A.

- Sea el valor propio $\lambda_1 = 2$.

$$(A - \lambda_1 I)v = O,$$

$$(A - \lambda_1 I \mid O) = \left(\begin{array}{ccc|c} 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 2 & 1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Sea $u_3 = r \in \mathbb{R} \setminus \{0\}$. Entonces, $u_1 = -r$ y $u_2 = r$ y

$$v = \begin{pmatrix} -r \\ r \\ r \end{pmatrix} \qquad r \in \mathbb{R} \setminus \{0\}$$

es el conjunto de vectores propios de la matriz A asociados con valor propio $\lambda_1 = 2$. Si $r = 1$, tenemos

$$v_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

que es un vector propio de la matriz A asociados con valor propio $\lambda_1 = 2$, y

$$V_{\lambda_1=2} = \left\langle \left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\} \right\rangle$$

es el subespacio vectorial propio de la matriz A asociado con valor propio $\lambda_1 = 2$. Como la dimensión del subespacio vectorial propio $V_{\lambda_1=2}$ es 1, entonces la multiplicidad geométrica del valor propio $\lambda_1 = 2$ es 1.

- Sea el valor propio $\lambda_2 = 4$.

$$(A - \lambda_2 I)v = O,$$

$$(A - \lambda_2 I \mid O) = \left(\begin{array}{ccc|c} -2 & -1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 2 & 1 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Entonces

$$v = \begin{pmatrix} r \\ -r \\ r \end{pmatrix} \qquad r \in \mathbb{R} \setminus \{0\}$$

es el conjunto de vectores propios de la matriz A asociados con valor propio $\lambda_2 = 4$. El vector

$$v_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

es un vector propio de la matriz A asociado con valor propio $\lambda_2 = 4$, y

$$V_{\lambda_2=4} = \left\langle \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\} \right\rangle$$

es el subespacio vectorial propio de la matriz A asociado con valor propio $\lambda_2 = 4$. La multiplicidad geométrica del valor propio $\lambda_2 = 4$ es 1.

♦

Teorema 7.4 Sea la aplicación lineal $f \in \mathscr{L}(V, V)$. Sean $\lambda_1, \lambda_1, \ldots, \lambda_k$, k valores propios de f distintos dos a dos. Entonces los k vectores propios asociados con los valores propios λ_i, $i = 1, 2, \ldots, k$, son linealmente independientes.

Demostración. Sean v_i los vectores propios asociados con los valores propios λ_i para $i = 1, 2, \ldots, k$. Es decir $f(v_i) = \lambda_i v_i$ para $i = 1, 2, \ldots, k$. Por demostrar que si

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0_V,$$

entonces $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$.

La demostración es por inducción.

i. Demostración para $k = 1$. si $\alpha_1 v_1 = 0_V$ entonces $\alpha_1 = 0$, puesto que $v_1 \neq 0_V$.

ii. Suponemos que es verdad para j y demostraremos para $j + 1$. Sea

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_j v_j + \alpha_{j+1} v_{j+1} = 0_V. \tag{7.1}$$

Entonces

$$\alpha_1 \lambda_{j+1} v_1 + \alpha_2 \lambda_{j+1} v_2 + \cdots + \alpha_j \lambda_{j+1} v_j + \alpha_{j+1} \lambda_{j+1} v_{j+1} = 0_V. \tag{7.2}$$

De (7.1):

$$\alpha_1 f(v_1) + \alpha_2 f(v_2) + \cdots + \alpha_j f(v_j) + \alpha_{j+1} f(v_{j+1}) = 0_V. \tag{7.3}$$

Entonces

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_j \lambda_j v_j + \alpha_{j+1} \lambda_{j+1} v_{j+1} = 0_V. \tag{7.4}$$

Hacemos (7.2) - (7.4):

$$\alpha_1 (\lambda_{j+1} - \lambda_1) v_1 + \alpha_2 (\lambda_{j+1} - \lambda_2) v_2 + \cdots + \alpha_j (\lambda_{j+1} - \lambda_j) v_j = 0_V. \tag{7.5}$$

Como los valores propios son distintos dos a dos tenemos que $\lambda_{j+1} - \lambda_j$ para $j = 1, 2, \ldots, k-1$. Luego de (4) se concluye que

$$\alpha_1 = \alpha_2 = \cdots = \alpha_j = 0,$$

por hipótesis inductiva. De (7.1) se tiene $a_{j+1}v_{j+1} = 0_V$, por lo cual $a_{j+1} = 0$ pues $v_{j+1} \neq 0_V$. Por i. y ii. se concluye el teorema.

$\square$

Corolario 7.5 Sea la aplicación lineal $f \in \mathscr{L}(V, V)$ tal que $\dim(V) = n$. Si $\lambda_1, \lambda_1, \ldots, \lambda_n$ son n valores propios de f distintos dos a dos, entonces los n vectores propios $v_1, v_2, \ldots, v_n$ asociados forman una base del espacio vectorial V.

7.3 Matrices simétricas

Teorema 7.6 Toda matriz simétrica $A \in M_n(\mathbb{R})$ posee n valores propios reales, no necesariamente distintos.

Demostración. Sea λ un valor propio de la matriz A. Entonces existe una matriz v no nula tal que $Av = \lambda v$. Por demostrar que $\lambda = \bar{\lambda}$. Tenemos que

$$\langle Av, v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2, \tag{7.6}$$
$$\langle Av, v\rangle = \langle v, A^t v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle = \bar{\lambda}\|v\|^2. \tag{7.7}$$

De (7.6) y (7.7) tenemos que

$$\lambda\|v\|^2 = \bar{\lambda}\|v\|^2.$$

Como $v \neq 0_V$, $\lambda = \bar{\lambda}$. Es decir, λ es un número real. $\square$

Teorema 7.7 Los vectores propios de una matriz simétrica real, asociados a valores propios distintos, son ortogonales.

Demostración. Sean λ_i y λ_j dos valores propios de una matriz simétrica $A \in M(\mathbb{R})$ tales que $\lambda_i \neq \lambda_j$. Sean v_i y v_j los vectores propios asociados de la matriz A asociados a los valores propios λ_i y λ_j, respectivamente, es decir:

$$Av_i = \lambda_i v_i \quad \text{y} \quad Av_j = \lambda_j v_j.$$

Por demostrar que $\langle v_i, v_j\rangle = 0$ para $i \neq j$. En efecto:

$$\langle Av_i, v_j\rangle = \langle \lambda_i v_i, v_j\rangle = \lambda_i\langle v_i, v_j\rangle, \tag{7.8}$$
$$\langle Av_i, v_j\rangle = \langle v_i, A^t v_j\rangle = \langle v_i, Av_j\rangle = \langle v_i, \lambda_j v_j\rangle = \bar{\lambda}_j\langle v_i, v_j\rangle = \lambda_j\langle v_i, v_j\rangle. \tag{7.9}$$

De (7.8) y (7.9):

$$(\lambda_i - \lambda_j)\langle v_i, v_j\rangle = 0.$$

Pero $(\lambda_i - \lambda_j) \neq 0$, por lo cual $\langle v_i, v_j\rangle = 0$. $\square$

Corolario 7.8 Sean la aplicación lineal $f \in \mathscr{L}(V, V)$ y la matriz simétrica $A \in M_n(\mathbb{R})$ asociada a f con respecto a cualquier base. Si la matriz A tiene n valores propios distintos dos a dos, entonces los vectores propios de f forman una base ortogonal del espacio vectorial V.

Ejemplo 7.101

Sean $f \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^3)$ y

$$A = [f]_C^C = \begin{pmatrix} 9 & -3 & 0 \\ -3 & 12 & -3 \\ 0 & -3 & 9 \end{pmatrix}.$$

Calcular los valores y vectores propios de la matriz A y de la aplicación lineal f.

Solución. Hallemos los valores propios de la matriz A:

$$\begin{aligned} p_A(\lambda) &= |A - \lambda I| \\ &= \begin{vmatrix} 9-\lambda & -3 & 0 \\ -3 & 12-\lambda & -3 \\ 0 & -3 & 9-\lambda \end{vmatrix} \\ &= -(\lambda^3 - 30\lambda^2 + 279\lambda - 810) \\ &= -(\lambda - 6)(\lambda - 9)(\lambda - 15) \\ &= 0, \end{aligned}$$

de donde encontramos

$$\lambda_1 = 6, \quad \lambda_2 = 9, \quad \text{y} \quad \lambda_3 = 15,$$

que son los valores propios de la matriz A y la aplicación lineal f.

Encontremos los vectores propios de la matriz A:

- Sea el valor propio $\lambda_1 = 6$:

$$(A - \lambda_1 I \mid O = \begin{pmatrix} 3 & -3 & 0 & | & 0 \\ -3 & 6 & -3 & | & 0 \\ 0 & -3 & 3 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & | & 0 \\ 0 & 1 & -1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

  de donde tenemos que

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

  es un vector propio de la matriz A asociado al valor propio $\lambda_1 = 6$. Fácilmente se encuentra que $u_1 = (1, 1, 1)$ es un vector propio de la aplicación lineal f asociado al valor propio $\lambda_1 = 6$.

- Sea el valor propio $\lambda_2 = 9$:

$$(A - \lambda_1 I \mid O = \begin{pmatrix} 0 & -3 & 0 & | & 0 \\ -3 & 3 & -3 & | & 0 \\ 0 & -3 & 0 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

de donde tenemos que

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

es un vector propio de la matriz A asociado al valor propio $\lambda_2 = 9$ y $u_2 = (1, 0, -1)$ es un vector propio de la aplicación lineal f asociado al valor propio $\lambda_2 = 9$.

- Sea el valor propio $\lambda_3 = 15$:

$$(A - \lambda_1 I \mid O = \begin{pmatrix} -6 & -3 & 0 & | & 0 \\ -3 & -3 & -3 & | & 0 \\ 0 & -3 & -6 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & | & 0 \\ 0 & 1 & 2 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

de donde tenemos que

$$v_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

es un vector propio de la matriz A asociado al valor propio $\lambda_3 = 15$ y $u_3 = (1, -2, 1)$ es un vector propio de la aplicación lineal f asociado al valor propio $\lambda_3 = 15$.

La matriz A de este ejemplo es simétrica y podemos verificar las propiedades enunciadas anteriormente para las matrices simétricas:

- Los valores propios de la matriz, $\lambda_1 = 6$, $\lambda_2 = 9$, y $\lambda_3 = 15$, son números reales.

- Como los valores propios de la matriz A son distintos dos a dos, los vectores propios de la matriz A también son distintos dos a dos y, por ser matriz simétrica, son ortogonales:

$$\begin{aligned} \langle v_1, v_2 \rangle &= \operatorname{tr}(v_1 v_2^t) = 0, \\ \langle v_1, v_3 \rangle &= \operatorname{tr}(v_1 v_3^t) = 0, \\ \langle v_2, v_3 \rangle &= \operatorname{tr}(v_2 v_3^t) = 0. \end{aligned}$$

Igual cosa sucede con los vectores propios de la aplicación lineal f:

$$\begin{aligned} \langle u_1, u_2 \rangle &= \langle (1,1,1), (1,0,-1) \rangle = 0, \\ \langle u_1, u_3 \rangle &= \langle (1,1,1), (1,-2,1) \rangle = 0, \\ \langle u_2, u_3 \rangle &= \langle (1,0,-1), (1,-2,1) \rangle = 0. \end{aligned}$$

- Los vectores propios de la matriz A constituyen una base

$$B_1 = \{v_1, v_2, v_3\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$$

ortogonal del espacio vectorial $M_{3\times 1}$, y los vectores propios de la aplicación lineal f constituyen una base

$$B_2 = \{u_1, u_2, u_3\} = \{(1,1,1), (1,0,-1), (1,-2,1)\}$$

del espacio vectorial $\mathbb{R}^3$.

♦

Teorema 7.9 Sean la aplicación lineal $f \in \mathscr{L}(V, V)$ y una matriz simétrica $A \in M_n(\mathbb{R})$ asociada a f con respecto a una base B del espacio vectorial V. Entonces existe una base ortogonal de V formada por vectores propios de la aplicación lineal f.

Demostración. Se hace la demostración por inducción sobre la dimensión del espacio vectorial V. Sea $\dim(V) = n$.

1. Si $n = 1$, no hay dada que demostrar pues para hablar de base ortogonal se necesita que la base tenga al menos dos vectores.

2. Sea $n > 1$. La matriz A tiene al menos un vector propio. Sea v_1 ese vector propio el cual satisface la ecuación $Av_1 = \lambda_1 v_1$.

   Sea
   $$W = \langle \{v_1\} \perp \rangle = \{u \in V \mid \langle u, v_1 \rangle = 0\}.$$

   Es claro que $\dim(W) = n - 1$.

   Sea $f|_W \in \mathscr{L}(W, W)$ la aplicación lineal f restringida al subespacio vectorial W. La matriz asociada a esta aplicación lineal sigue siendo simétrica. Luego, por hipótesis inductiva, se tiene que existe una base ortogonal de vectores propios de f para el subespacio vectorial W. Sea $S = \{u_1, u_2, \ldots, u_{n-1}\}$ dicha base.

   Como $\langle u_i, v_1 \rangle = 0$ para todo $i = 1, 2, \ldots, n-1$, el conjunto $B = \{v_1, u_1, u_2, \ldots, u_{n-1}\}$ es una base ortogonal para el espacio vectorial V, formada por vectores propios de f. $\square$

De esta propiedad se deduce que de toda matriz simétrica real de orden n se pueden obtener n vectores propios ortogonales. Esto significa que la multiplicidad geométrica de un un valor propio de una matriz simétrica real es igual a la multiplicidad algebraica de dicho valor propio.

Ejemplo 7.102

Sea la matriz
$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$
asociada a la aplicación lineal $f \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^3)$ con respecto a la base canónica.

a. Calcular los valores propios de la matriz A y la aplicación lineal f, y los vectores propios de f.

b. Construir una base ortogonal del espacio vectorial $\mathbb{R}^3$, formada por vectores propios de f.

Solución.

a. Hallemos los valores propios de A y f:

$$\begin{aligned} p_A(\lambda) &= |A - \lambda I| \\ &= \begin{vmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix} \\ &= -(\lambda^3 - 6\lambda^2 - 15\lambda - 8) \\ &= -(\lambda+1)^2(\lambda - 8) \\ &= 0, \end{aligned}$$

de donde encontramos

$$\lambda_1 = -1, \quad \text{y} \quad \lambda_2 = 8,$$

que son los valores propios de la matriz A y la aplicación lineal f. El valor $\lambda_1 = -1$ es de multiplicidad algebraica 2 y $\lambda_2 = 8$, de multiplicidad algebraica 1. Como A es una matriz simétrica, la multiplicidad geométrica de los dos valores es necesariamente igual a la algebraica: esto significa que λ_1 genera dos vectores propios y λ_2 genera un vector propio.

Encontremos los vectores propios de la matriz A:

- Sea el valor propio $\lambda_1 = -1$:

$$(A - \lambda_1 I \mid O = \begin{pmatrix} 4 & 2 & 4 & | & 0 \\ 2 & 1 & 2 & | & 0 \\ 4 & 2 & 4 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 2 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

de donde tenemos que

$$v_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \quad \text{y} \quad v_2 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$$

son dos vectores propios de la matriz A asociados al valor propio $\lambda_1 = -1$. De donde

$$u_1 = (1, -2, 0) \quad \text{y} \quad u_2 = (0, -2, 1)$$

son dos vectores propios de la aplicación lineal f asociados al valor propio $\lambda_1 = -1$.

- Sea el valor propio $\lambda_2 = 8$:

$$(A - \lambda_1 I \mid O = \begin{pmatrix} -5 & 2 & 4 & | & 0 \\ 2 & -8 & 2 & | & 0 \\ 4 & 2 & -5 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & | & 0 \\ 0 & 2 & -1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

de donde tenemos que

$$v_3 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

es un vector propio de la matriz A asociado al valor propio $\lambda_2 = 8$ y

$$u_3 = (2, 1, 2)$$

$u_3 = (2, 1, 2)$ es un vector propio de la aplicación lineal f asociado al valor propio $\lambda_2 = 8$.

b. Debido a que A es matriz simétrica, el vector propio u_3 es ortogonal a u_1 y u_2:

$$\langle u_3, u_1\rangle = \langle (2,1,2),(1,-2,0)\rangle = 0,$$
$$\langle u_3, u_2\rangle = \langle (2,1,2),(0,-2,1)\rangle = 0,$$

pero u_1 y u_2 no son necesariamente ortogonales, como en efecto es el caso:

$$\langle u_1, u_2\rangle = \langle (1,-2,0),(0,-2,1)\rangle = 4 \neq 0.$$

Sin embargo el conjunto $\{u_1, u_2\}$ se puede ortogonalizar con el proceso de Gram-Schmidt; el nuevo conjunto está formado por vectores que siguen siendo vectores propios de la aplicación lineal f asociados al valor propio $\lambda_1 = -1$. Ortogonalicemos, pues el conjunto $\{u_1, u_2\}$:

Sean

$$\begin{aligned} w_1 &= u_1 = (1,-2,0),\\ w_2 &= u_2 - \frac{\langle u_2, w_1\rangle}{\|w_1\|^2} w_1\\ &= (0,-2,1) - \frac{\langle (0,-2,1),(1,-2,0)\rangle}{\|(1,-2,0)\|^2}(1,-2,0)\\ &= (0,-2,1) - \frac{4}{5}(1,-2,0)\\ &= (0,-2,1) - \left(\frac{4}{5}, \frac{-8}{5}, 0\right)\\ &= \left(-\frac{4}{5}, -\frac{2}{5}, 1\right). \end{aligned}$$

Si llamamos $w_3 = u_3 = (2,1,2)$,

$$B = \{w_1, w_2, w_3\} = \left\{(1,-2,0), \left(-\frac{4}{5}, -\frac{2}{5}, 1\right), (2,1,2)\right\}$$

es una base ortogonal del espacio vectorial $\mathbb{R}^3$ formada por vectores propios de la aplicación lineal f.

♦

7.4 Diagonalización

Definición 7.5 (Diagonalizable) Sea $f \in \mathscr{L}(V,V)$. Se dice que f es **diagonalizable** si y solo si existe una base B tal que la representación matricial de f respecto de B es una matriz diagonal.

Observaciones

1. Una matriz $A = [f]_C^C$ es diagonalizable si y solo si existe una base B tal que

$$[f]_B^B = [Id]_B^C [f]_C^C [Id]_C^B$$

es una matriz diagonal.

2. Una matriz A es diagonalizable si y solo si existe una matriz P inversible tal que la matriz
$$D = P^{-1}AP$$
es una matriz diagonal.

3. Si una matriz A es diagonalizable, entonces A es semejante a una matriz diagonal D.

4. Dos matrices semejantes tienen los mismos valores propios.

Teorema 7.10 Sea $f \in \mathscr{L}(V, V)$ tal que $\dim(V) = n$. Entonces f es diagonalizable si y solo si f tiene n vectores linealmente independientes.

Demostración. Sea $A = [f]_C^C$.

1. Supongamos que f tiene n vectores propios linealmente independientes. Demostremos que f es diagonalizable. Para ello, mostraremos que hay una base B de V respecto de la cual la matriz A es diagonal.

   Sean $v_1 = (v_{11}\ v_{21}\ \cdots\ v_n 1)^t$, $v_2 = (v_{12}\ v_{22}\ \cdots\ v_{n2})^t$, $\ldots$, $v_n = (v_{1n}\ v_{2n}\ \cdots\ v_{nn})^t$ los n vectores propios linealmente independientes y sean λ_1, λ_2, $\ldots$, λ_n los valores propios asociados, respectivamente (no todos los λ_i son necesariamente distintos). Sea
$$P = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}.$$
   Como el conjunto $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente, entonces la matriz P es inversible. Además, se tiene que:
$$\begin{aligned} AP &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} \\ &= \begin{pmatrix} Av_1 & Av_2 & \cdots & Av_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \cdots & \lambda_n v_{1n} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \cdots & \lambda_n v_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_1 v_{n1} & \lambda_2 v_{n2} & \cdots & \lambda_n v_{nn} \end{pmatrix} \\ &= \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \end{aligned}$$

Por lo tanto

$$AP = PD, \tag{7.10}$$

donde $D = \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ es una matriz diagonal cuya diagonal tiene a los valores propios de A.

Finalmente, de (7.10), tenemos que

$$D = P^{-1}AP;$$

es decir, la matriz A es semejante a una matriz diagonal; por lo tanto, A es diagonalizable, como se quería demostrar.

2. Supongamos, ahora, que f es diagonalizable. Probemos que f tiene n vectores propios linealmente independientes.

   Como f es diagonalizable, la matriz A también lo es; por lo tanto, existe una matriz P, inversible, tal que la matriz

   $$D = P^{-1}AP \tag{7.11}$$

   es una matriz diagonal. Sean $v_1, v_2, \ldots, v_n$ las columnas de la matriz P; entonces, cada v_i es un vector de V. Sean λ_i los elementos de la diagonal de D.

   Por un lado, tenemos que:

   $$\begin{aligned} PD &= \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{pmatrix}. \end{aligned}$$

   Por otro,

   $$AP = A\begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} = \begin{pmatrix} Av_1 & Av_2 & \cdots & Av_n \end{pmatrix}.$$

   Por lo tanto, por la definición de igualdad de matrices aplicado a (7.11), podemos concluir que

   $$Av_i = \lambda_i v_i;$$

   es decir, los vectores v_i son los vectores propios de A y λ_i los valores propios asociados. Finalmente, como P es inversible, el conjunto $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente. □

De la demostración, podemos ver que si una aplicación lineal f es diagonalizable, existe siempre una matriz P inversible tal que $D = P^{-1}[f]_C^C P$ que es diagonal, donde los elementos de esta diagonal son, precisamente, los valores propios de $[f_C^C]$, y las columnas de P son los vectores propios correspondientes; es decir, la matriz asociada a f respecto a la base formada por estos vectores propios tiene una representación diagonal.

Corolario 7.11 Si $A \in \mathbb{M}_n$ tiene n valores propios distintos dos a dos, entonces A es diagonalizable.

Corolario 7.12 Toda matriz simétrica $A \in \mathbb{M}_n(\mathbb{R})$ es diagonalizable.

Teorema 7.13 Sea $f \in \mathscr{L}(V, V)$ y $\dim(V) = n$. Si f tiene una representación matricial A simétrica, entonces existe una matriz ortogonal P tal que la matriz $P^t AP$ es una matriz diagonal donde los valores propios de A son los elementos de la diagonal.

Demostración. Como A es simétrica, es diagonalizable. Entonces, por el teorema 7.10, la matriz P de los vectores propios de A tal que

$$D = P^{-1}AP \tag{7.12}$$

es la matriz diagonal de los valores propios de A.

Tenemos, además, que $P^t P = I$, puesto que los vectores columnas de I son ortonormales, de manera que $P^{-1} = P^t$. Así, por (7.12), tenemos que $D = P^t AP$. □

Definición 7.6 (Matriz diagonalizable ortogonalmente) Una matriz $A \in \mathbb{M}_n$ se dice que es *diagonalizable ortogonalmente* si y solo si existe una matriz ortogonal P tal que la matriz $D = P^t AP$ es diagonal.

Ejemplo 7.103

Determinar una matriz P ortogonal tal que $D = P^t AP$ sea diagonal, donde A es la matriz simétrica real

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

Solución. Por el teorema 7.13, la matriz buscada es la matriz de los vectores propios ortonormales.

La matriz tiene dos valores propios: $\lambda_1 = -1$ y $\lambda_2 = 8$. Hay dos vectores propios asociados a λ_1 y son:

$$v_1 = \frac{1}{\sqrt{5}}(1, -2, 0) \quad \text{y} \quad v_2 = \frac{\sqrt{5}}{3}\left(-\frac{4}{5}, -\frac{2}{5}, 1\right).$$

El vector propio asociado a λ_2 es:

$$v_3 = \frac{1}{3}(2, 1, 2).$$

Por lo tanto, la matriz P buscada es:

$$\frac{1}{3\sqrt{5}} \begin{pmatrix} 3 & -4 & 2\sqrt{5} \\ -6 & -2 & \sqrt{5} \\ 0 & 5 & 2\sqrt{5} \end{pmatrix}.$$

A pesar de que no es necesario, pues el teorema 7.13 nos lo garantiza, verifiquemos que P^tAP es la matriz diagonal de los valores propios:

$$\begin{aligned} P^tAP &= \frac{1}{45}\begin{pmatrix} -3 & 6 & 0 \\ 4 & 2 & -5 \\ 16\sqrt{5} & 8\sqrt{5} & 16\sqrt{5} \end{pmatrix}\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}\begin{pmatrix} 3 & -4 & 2\sqrt{5} \\ -6 & -2 & \sqrt{5} \\ 0 & 5 & 2\sqrt{5} \end{pmatrix} \\ &= \frac{1}{45}\begin{pmatrix} -45 & 0 & 0 \\ 0 & -45 & 0 \\ 0 & 0 & 360 \end{pmatrix} \\ &= \operatorname{diag}\{-1,-1,8\}. \end{aligned}$$

♦

7.5 Cadenas de Markov

Las cadenas de Markov se emplean para analizar sistemas que en un momento dado pueden estar en un estado de número finito de estados; así, una persona puede tener una deuda o no tenerla; el clima puede estar seco o húmedo; un sistema mecánico puede estar en equilibrio o bien fuera de equilibrio. En cada caso, el sistema puede cambiar de un estado a otro. Además, existe una probabilidad de transcición de un estado a otro, entre tiempos de observación sucesivos. El objetivo del análisis de Markov es calcular la probabilidad de que un sistema se encuentre en un estado en particular en un tiempo futuro, y determinar el comportamiento del sistema a largo plazo.

Ejemplo 7.104 Se considera el sistema de un estudiante E quien puede estar en los estados:

$$\begin{cases} s_1: & \text{calificación } p < 7.5 \\ s_2: & \text{calificación } p \geq 7.5. \end{cases}$$

Se supone que la calificación máxima es 10 y que p es la calificación de una prueba. El resultado de las observaciones en el nivel de enseñanza media es que si E se halla en estado s_1 en una prueba, entonces estudiará más para la próxima prueba y logrará el estado s_2 con una probabilidad de 0.8, y, en consecuencia, la probabilidad de que un estudiante esté en el estado s_1 es de 0.2.

Pero, si E se encuentra en s_2 en una prueba, para la siguiente prueba se relajará y caerá en el estado s_1 con una probabilidad de 0.3, y, en consecuencia, estará en el estado s_2 con una probabilidad de 0.7.

En este ejemplo, se ha supuesto que el desempeño de E en una prueba está determinado únicamente por su motivación debida a su calificación de la prueba anterior.

Las probabilidades anteriores se llaman *probabilidades de transición*, y se las dispone en una matriz:

$$\begin{array}{c} \text{Estado en el tiempo } T_{k-1} \\ M = \begin{pmatrix} 0.2 & 0.3 \\ 0.8 & 0.7 \end{pmatrix} \begin{matrix} s_1 \\ s_2 \end{matrix} \end{array} \quad \text{Estado en el tiempo } T_k \; .$$

Obsérvese que la suma de cada columna es igual a 1.

La matriz M es denominada *matriz de transición* para la cadena de Markov.

Como ejemplo, supongamos que el estudiante tiene una nota de $p = 7$ en la primera prueba. Entonces, E se halla en el estado s_1, y éste se representa mediante un vector de estado:

$$S = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \begin{matrix} E \text{ se encuentra en } s_1 \\ E \text{ se encuentra en } s_2 \end{matrix}$$

y

$$MS = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$$

representa las probabilidades de estar en s_1 y s_2 después de un prueba con $p < 7.5$ inicialmente. Las probabilidades de estar en s_1 y s_2 después de n pruebas será $M^n S$.

Ahora, el problema radica en calcular M^n. Para ello, se calculan los valores propios y vectores propios de M que, en este caso, son $\lambda_1 = 1$ y $v_1 = (3, 8)$, y $\lambda_2 = -0.1$ con $v_2 = (1, -1)$. Por lo tanto:

$$\begin{aligned} M^n &= PD^nP^{-1} \\ &= \frac{1}{11} \begin{pmatrix} 3 & 1 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (-0.1)^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 8 & -3 \end{pmatrix} \\ &= \frac{1}{11} \begin{pmatrix} 3 + 8(-0.1)^n & 3 - 3(-0.1)^n \\ 8 - 8(-0.1)^n & 8 + 3(-0.1)^n \end{pmatrix}. \end{aligned}$$

Las probabilidades de que E esté en los estados s_1 y s_2 después de 4 pruebas sucesivas es

$$M^4 S = \begin{pmatrix} 0.272\,8 \\ 0.727\,2 \end{pmatrix};$$

es decir, la probabilidad de que E tenga una nota p superior a 7.5 en la cuarta prueba es de 0.727 2. ♦

Ejemplo 7.105

Se considera el sistema de ecuaciones diferenciales:

$$\begin{cases} x_1'(t) = & 3x_1(t) - x_2(t) \\ x_2'(t) = & -2x_1(t) + 2x_2(t); \end{cases}$$

es decir: $X'(t) = AX(t)$, que es un modelo matemático, el cual representa el comportamiento que se da entre dos especies.
Si se supone que las poblaciones iniciales para cada especie (para el tiempo $t = 0$) son $x_1(0) = 90$ y $x_2(0) = 150$,

a) calcule la población de cada especie para un tiempo $t > 0$; y

b) calcule el tiempo para el cual una de las dos especies se elimina (si esto sucediera) y calcule la población de la otra especie.

Solución. La solución general de un sistema $X'(t) = AX(t)$ es

$$X(t) = \sum_{i=1}^{2} c_i e^{\lambda_i t} v_i,$$

siempre que A sea diagonalizable, donde c_i son constantes arbitrarias, que se determinan a partir de las condiciones iniciales del sistema, los λ_i son los valores propios de A y los v_i son los vectores propios asociados a los valores λ_i, respectivamente.

a) Los valores y vectores propios de A son $\lambda_1 = 1$ con $v_1 = (1, 2)$, $\lambda_2 = 4$ con $v_2 = (1, -1)$. Entonces la solución general del sistema es:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 e^{4t} \begin{pmatrix} 1 \\ -1 \end{pmatrix};$$

es decir:

$$\begin{cases} x_1(t) = & c_1 e^t + c_2 e^{4t} \\ x_2(t) = & 2c_2 e^t - c_2 e^{4t}. \end{cases}$$

Para hallar las constantes c_i, se utilizan las condiciones iniciales:

$$x_1(0) = c_1 + c_2 = 90$$
$$x_2(0) = 2c_1 - c_2 = 150.$$

Luego de resolver este sistema de ecuaciones lineales, se obtiene que $c_1 = 80$ y $c_2 = 10$. Por lo tanto, la solución particular del sistema de ecuaciones diferenciales, bajo las condiciones iniciales es:

$$\begin{cases} x_1(t) = & 80e^t + 10e^{4t} \\ x_2(t) = & 160e^t - 10e^{4t}. \end{cases}$$

Si t se mide en años, al cabo de 6 meses ($t = 1/2$), se tiene que las poblaciones son $x_1(t) = 206$ ejemplares y $x_2(t) \approx 190$ ejemplares.

b) Según la solución particular, se tiene que $x_1(t)$ no se anula para ningún valor de t, y que $x_2(t)$ sí se anula. En efecto, $160e^t - 10e^{4t} = 0$ implica que $t \approx 11$ meses; es decir, al cabo de aproximadamente 11 meses, la especie x_2 se elimina.

♦

7.6 Ejercicios resueltos

1. Demostrar que los valores propios de una matriz A y de su transpuesta A^t son los mismos.

   Demostración. Es suficiente demostrar que el polinomio característico de A es el igual al polinomio característico de A^t; es decir, vamos a demostrar que $p_A = p_{A^t}$:

$$\begin{aligned} p_A(\lambda) &= \left|A - \lambda I\right| \\ &= \left|(A - \lambda I)^t\right| \\ &= \left|A^t - \lambda I\right| = p_{A^t}(\lambda). \end{aligned}$$ □

2. Sea A una matriz de orden n diagonalizable, y sean λ_i con $i \in \{1, 2, \ldots, n\}$, los valores propios de A (no necesariamente distintos). Demuestre que:

   (a) $\operatorname{Tr}(A) = \displaystyle\sum_{i=1}^{n} \lambda_i.$ (b) $|A| = \displaystyle\prod_{i=1}^{n} \lambda_i.$

   Demostración. Como A es diagonalizable, existe una matriz P inversible tal que

$$\operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = P^{-1}AP.$$

(a) Tenemos que $A = PDP^{-1}$. Por lo tanto:

$$\begin{aligned}\operatorname{Tr}(A) &= \operatorname{Tr}((PD)P^{-1}) \\ &= \operatorname{Tr}(P^{-1}(PD)) \\ &= \operatorname{Tr}(P^{-1}PD) \\ &= \operatorname{Tr}(ID) \\ &= \operatorname{Tr}(D) = \sum_{i=1}^{n} \lambda_i.\end{aligned}$$

(b) De $A = PDP^{-1}$, tenemos que:

$$\begin{aligned}|A| &= |(PD)P^{-1}| \\ &= |P||D||P^{-1}| \\ &= |D| = \prod_{i=1}^{n} \lambda_i.\end{aligned}$$

□

Los resultados demostrados son válidos, aunque la matriz A no sea diagonalizable; sin embargo, la demostración requiere de la teoría de las formas canónicas de Jordan, tema que no se trata en este libro.

Además, estos resultados son útiles ara calcular los valores propios de una matriz de orden 2 sin hallar el polinomio característico, pues se tiene dos ecuaciones no lineales con dos incógnitas. También pueden ser útiles para verificar que los valores propios encontrados en una matriz son correctos.

3. Sea $f \in \mathscr{L}(V, V)$ una aplicación biyectiva. Demuestre que

   a) todos los valores propios de f son no nulos; y

   b) si λ es un valor propio de f, entonces λ^{-1} es un valor propio de f^{-1}.

Demostración.

(a) Supongamos que 0 es un valor propio de f; entonces, existe $v \in V$ tal que $v \neq 0$ y $f(v) = 0v = 0$. Por lo tanto, $v = 0$, pues f es biyectiva, lo que es imposible. La contradicción también puede verse con el hecho de que si 0 fuera una valor propio, entonces el núcleo de f sería distinto del espacio nulo, lo que es imposible por ser f inyectiva.

(b) Si λ es un valor propio de f, entonces $\lambda \neq 0$. Sea $v \in V$ tal que $v \neq 0$ y $f(v) = \lambda v$. Por lo tanto, por ser f^{-1} una función, tenemos que

$$f^{-1}(f(v)) = f^{-1}(\lambda v),$$

de donde:

$$v = \lambda f^{-1}(v),$$

lo que implica que

$$f^{-1}(v) = \lambda^{-1} v;$$

es decir, λ^{-1} es un valor propio de f^{-1}.

□

Cuando $f \in \mathscr{L}(V,V)$ es una función biyectiva, la matriz asociada f con respecto a cualquier base es inversible. Luego, si λ es un valor propio de una matriz inversible, entonces $\lambda \neq 0$ y λ^{-1} es un valor propio de A^{-1}. Nótese también que los vectores propios de A y de su inversa son los mismos.

4. Sean $f \in \mathscr{L}(V,V)$ y λ un valor propio de f. Determine los valores propios de f^2, f^3 y f^n con $n \in \mathbb{N}$.

Solución. Si λ es un valor propio de la aplicación lineal f, existe un vector no nulo $v \in V$ tal que $f(v) = \lambda v$.

Entendemos f^2 como $f \circ f$. Tomando en cuenta que f es aplicación lineal, tenemos que

$$f^2(v) = f(f(v)) = f(\lambda v) = \lambda f(v) = \lambda(\lambda v) = \lambda^2 v;$$

es decir, $f^2(v) = \lambda^2 v$, por lo cual λ^2 es valor propio de f^2.

Entendemos f^3 como $f \circ f \circ f$. De manera similar, y usando el resultado anterior, tenemos que

$$f^3(v) = f(f^2(v)) = f(\lambda^2 v) = \lambda^2 f(v) = \lambda^2(\lambda v) = \lambda^3 v,$$

por lo que concluimos que λ^3 es un valor propio de la aplicación lineal f^3.

Demostraremos por inducción que si λ es valor propio de la aplicación lineal f, entonces λ^n, con $n \in \mathbb{N}$, es valor propio de f^n.

i. Anteriormente demostramos para $n = 2$.
ii. Suponiendo que es verdad para n, demostraremos que es verdad para $n + 1$.

 Si λ es un valor propio de la aplicación lineal f, existe un vector no nulo $v \in V$ tal que $f(v) = \lambda v$.

 Por hipótesis inductiva, el número λ^n es valor propio de la aplicación lineal f^n. Por lo anterior, y tomando en cuenta que la aplicación f^n es lineal, podemos escribir

$$f^{n+1}(v) = f(f^n(v)) = f(\lambda^n v) = \lambda^n f(v) = \lambda^n(\lambda v)) = \lambda^{n+1} v;$$

 es decir,

$$f^{n+1}(v) = \lambda^{n+1} v,$$

 que es equivalente a decir que λ^{n+1} es valor propio de la aplicación lineal f^{n+1}. □

5. Los vectores v_1 y v_2 son dos vectores propios de una aplicación lineal f, asociados a los valores propios distintos λ_1 y λ_2, respectivamente. Si $\alpha v_1 + \beta v_2$ (con $\alpha \in K$ y $\beta \in K$) es un vector propio de f, entonces $\alpha = 0$ o $\beta = 0$, pero no ambos.

Demostración. Si v_1 y v_2 son vectores propios, son vectores no nulos. Decir que v_1 y v_2 son dos vectores propios de una aplicación lineal f es equivalente a decir que

$$f(v_1) = \lambda_1 v_1 \quad \text{y} \quad f(v_2) = \lambda_2 v_2.$$

Por ser f una aplicación lineal, podemos escribir:

$$f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2) = \alpha(\lambda_1 v_1) + \beta(\lambda_2 v_2).$$

Pero, por hipótesis

$$f(\alpha v_1 + \beta v_2) = \lambda(\alpha v_1 + \beta v_2).$$

Por transitividad, tenemos que

$$\alpha(\lambda_1 v_1) + \beta(\lambda_2 v_2) = \lambda(\alpha v_1 + \beta v_2) = (\alpha\lambda)v_1 + (\beta\lambda)v_2,$$

de donde

$$\alpha\lambda_1 = \alpha\lambda \quad \text{y} \quad \beta\lambda_2 = \beta\lambda.$$

Supongamos que $\alpha \neq 0$ y $\beta \neq 0$. Entonces $\lambda_1 = \lambda$ y $\lambda_2 = \lambda$, es decir $\lambda_1 = \lambda_2$, lo cual es una contradicción. Por lo tanto, o bien $\alpha = 0$ o bien $\beta = 0$.

Tampoco es posible que $\alpha = \beta = 0$, pues, en ese caso, el vector 0 sería un vector propio, ya que

$$\alpha v_1 + \beta v_2 = 0$$

es un vector propio por hipótesis. □

6. Sea la aplicación lineal $f \in (V, V)$ tal que $\dim(\mathrm{Im}(f)) = 1$ y $\dim V \geq 2$. Demostrar que

   a. existe un vector $v \in V$ tal que $f(v)$ es un vector propio de la aplicación lineal f,
   b. $\lambda = 0$ es un valor propio de f.

Demostración.

a. Como $V = N_f \oplus \mathrm{Im}(f)$, entonces existe un vector $v \in V$ tal que $v = w + f(v)$, donde $f(v) \neq 0_v$, pues $\mathrm{Im}(f) \neq \{0_v\}$ y $w \in N_f$. Es decir,

$$f(v) = v - w.$$

La aplicación f es lineal y podemos escribir

$$f(f(v)) = f(v - w) = f(v) - f(w).$$

Pero como $w \in N_f$, $f(w) = 0_v$ y

$$f(f(v)) = f(v),$$

que significa que $f(v)$ es un vector propio de la aplicación lineal f asociado al valor propio $\lambda = 1$.

b. Como $\dim N_f = \dim V - \dim(\mathrm{Im}(f))$ y, además, $\dim(\mathrm{Im}(f)) = 1$ y $\dim V \geq 2$, entonces

$$\dim N_f \geq 1,$$

lo cual significa que

$$N_f \neq \{0_v\}.$$

Entonces, existe un vector $w \in N_f$ no nulo tal que $f(w) = 0_v$, y podemos escribir

$$f(w) = 0w,$$

lo cual significa que w es un vector propio de la aplicación lineal f asociado al valor propio $\lambda = 0$.

□

7. Sea la aplicación lineal $f \in \mathcal{L}(P_2(t), P_2(t))$ tal que

$$f(a + bt + ct^2) = c + (b - a)t + (a - b + c)t^2.$$

   a. Encontrar los valores y vectores propios de la aplicación lineal f.
   b. ¿Es diagonalizable la aplicación lineal f?

Solución.

a. Para encontrar los valores y vectores propios de la aplicación lineal f necesitamos primero hallar una representación matricial de ella, y encontrar los valores y vectores propios de esa matriz. Los valores propios de la matriz son también los valores propios de la aplicación lineal.
   Sea el polinomio $p(t) = a + bt + ct^2$. Por facilidad escogemos la base canónica $C = \{e_1, e_2, e_3\} = \{1, t, t^2\}$ del espacio vectorial $P_2(t)$. Claramente, tenemos los vectores coordenados

$$[p(t)]_C = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad [f(p(t))]_C = \begin{pmatrix} a \\ -a + b \\ a - b + c \end{pmatrix}.$$

   Tenemos que

$$[f(e_1)]_C = [f(1)]_C = [-t + t^2]_C = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix},$$

$$[f(e_2)]_C = [f(t)]_C = [t - t^2]_C = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

$$[f(e_3)]_C = [f(t^2)]_C = [1 + t^2]_C = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

   Con los tres vectores coordenados anteriores, escritos como columnas, obtenemos la representación matricial $[f]_C^C$ de la aplicación lineal f con respecto a la base canónica C:

$$[f]_C^C = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

   Sea $A = [f]_C^C$. Verifiquemos que la Matriz A representa a la aplicación lineal f:

$$[f]_C^C\ [p(t)]_C = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ -a + b \\ a - b + c \end{pmatrix} = [f(p(t))]_C.$$

   Hallemos los valores propios de la matriz A, los cuales también son los valores propios de la aplicación lineal f:

$$p_A(\lambda) = |A - \lambda I| = \begin{vmatrix} -\lambda & 0 & 1 \\ -1 & 1 - \lambda & 0 \\ 1 & -1 & 1 - \lambda \end{vmatrix} = -\lambda^2(\lambda - 2) = 0.$$

   Los valores propios de la matriz A y la aplicación lineal f son $\lambda_1 = 0$ de multiplicidad algebraica 2, y $\lambda_2 = 2$ de multiplicidad algebraica 1 y geométrica 1.
   Hallemos los vectores propios de la matriz A y de la aplicación lineal f.

– Sea el valor propio $\lambda_1 = 0$. Para hallar los vectores propios v asociados con el valor propio $\lambda_1 = 0$ resolvemos el sistema lineal

$$\big(A - \lambda_1 I\big)\, v = O, \quad \text{donde} \quad v = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

Para hallar el vector v resolvemos el sistema de ecuaciones lineales

$$\big(A - \lambda_1 I \mid O\big) = \begin{pmatrix} 0 & 0 & 1 & | & 0 \\ -1 & 1 & 0 & | & 0 \\ 1 & -1 & 1 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 & 1 & | & 0 \\ -1 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}$$

y encontramos que $u_3 = 0$ y $u_2 = u_1$. Si $u_1 = 1$, $u_2 = 1$.
De donde $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ es un vector propio de la matriz A asociado al valor propio $\lambda_1 = 0$. El vector propio v_1 de la matriz A es el vector coordenado del vector propio $p_1(t)$ de la aplicación lineal f. Por lo tanto, si $[p_1(t)]_C = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, entonces $p_1(t) = 1 + t$.

– Sea el valor propio $\lambda_2 = 2$. Hallemos rápidamente su vector propio asociado:

$$\big(A - \lambda_2 I \mid O\big) = \begin{pmatrix} -1 & 0 & 1 & | & 0 \\ -1 & -1 & 0 & | & 0 \\ 1 & -1 & -1 & | & 0 \end{pmatrix} \sim \begin{pmatrix} -2 & 0 & 1 & | & 0 \\ 1 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

de donde $u_3 = 2u_1$ y $u_2 = -u_1$. Si $u_1 = 1$, $u_2 = -1$ y $u_3 = 2$.
Por lo cual $v_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ es un vector propio de la matriz A asociado al valor propio $\lambda_2 = 2$.
Entonces $p_2(t) = 1 - t + 2t^2$ es un vector propio de la aplicación lineal f asociado al valor propio $\lambda_2 = 2$.

Verifiquemos que $p_1(t)$ y $p_2(t)$ son vectores propios de la aplicación lineal f:

$$f(p_1(t)) = f(1+t) = 0 + (1-1)t + (1-1+0)t^2 = 0 = 0(1+t) = \lambda_1\, p_1(t),$$
$$f(p_2(t)) = f(1-t+2t^2) = 2 + (-1-1)t + (1+1+2)t^2 = 2(1-t+2t^2) = \lambda_2\, p_2(t).$$

b. Debido a que la representación matricial de la aplicación lineal f tiene un valor propio repetido, ella no es diagonalizable. En consecuencia, la aplicación lineal f tampoco es diagonalizable. □

8. Demostrar que si $A = PDP^{-1}$, entonces $A^n = PD^nP^{-1}$ para todo $n \in \mathbb{N}$.

Demostración. Demostraremos por inducción.

(i) Para $n = 1$.
Si $A = PDP^{-1}$, entonces $A^1 = PD^1P^{-1}$, pues $A = A^1$ y $D = D^1$

(ii) Suponemos que es verdad para n y demostraremos que es verdad para $n+1$. Es decir, si $A^n = PD^nP^{-1}$ es verdadero, por demostrar que $A^{n+1} = PDP^{-1}$. En efecto,

$$\begin{aligned} A^{n+1} &= A^n A \\ &= PD^nP^{-1}PDP^{-1} \quad \text{por hipótesis inductiva} \\ &= PD^nIDP^{-1} \\ &= PD^nDP^{-1} \\ &= PD^{n+1}P^{-1}. \end{aligned}$$

Por (i) y (ii) se tiene que si $A = PDP^{-1}$, entonces $A^n = PD^nP^{-1}$ para todo $n \in \mathbb{N}$. □

9. Dada la matriz $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix}$, hallar la matriz A^n.

Solución. Primero hallemos los valores propios de la matriz A.

$$p_A(\lambda) = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 1 \\ 2 & -\lambda & 2 \\ 1 & 2 & 1-\lambda \end{vmatrix} = -\lambda(\lambda+2)(\lambda-4) = 0$$

Los valores propios de la matriz A son

$$\lambda_1 = -2, \quad \lambda_2 = 0, \quad \lambda_3 = 4.$$

Debido a que los tres valores propios son distintos entre sí la matriz A es diagonalizable. Ahora encontremos los vectores propios de la matriz A.

- Sea el valor propio $\lambda_1 = -2$. Para hallar los vectores propios v asociados con el valor propio $\lambda_1 = -2$ resolvemos el sistema lineal

$$\big(A - \lambda_1 I\big)\, v = O, \quad \text{donde} \quad v = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

Para hallar el vector v resolvemos el sistema de ecuaciones lineales

$$\big(A - \lambda_1 I \mid O\big) = \begin{pmatrix} 3 & 2 & 1 & | & 0 \\ 2 & 2 & 2 & | & 0 \\ 1 & 2 & 3 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & | & 0 \\ 1 & 1 & 1 & | & 0 \\ -1 & 0 & 1 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & | & 0 \\ 2 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}$$

y encontramos que $u_3 = u_1$ y $u_2 = -2u_1$. Si $u_1 = 1$, $u_2 = -2$ y $u_3 = 1$. De donde $v_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ es un vector propio de la matriz A asociado al valor propio $\lambda_1 = -2$.

- Sea el valor propio $\lambda_2 = 0$. Resumimos el hallazgo del vector propio asociado:

$$\big(A - \lambda_2 I \mid O\big) = \begin{pmatrix} 1 & 2 & 1 & | & 0 \\ 2 & 0 & 2 & | & 0 \\ 1 & 2 & 1 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 2 & 0 & | & 0 \\ 1 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

de donde $u_2 = 0$ y $u_3 = -u_1$. Si $u_1 = 1$, $u_3 = -1$. De donde $v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ es un vector propio de la matriz A asociado al valor propio $\lambda_2 = 0$.

- Sea el valor propio $\lambda_3 = 4$. Hallemos el vector propio asociado:

$$\big(A - \lambda_3 I \mid O\big) = \begin{pmatrix} -3 & 2 & 1 & | & 0 \\ 2 & -4 & 2 & | & 0 \\ 1 & 2 & -3 & | & 0 \end{pmatrix} \sim \begin{pmatrix} -4 & 0 & 4 & | & 0 \\ 2 & -4 & 2 & | & 0 \\ 3 & -2 & -1 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & | & 0 \\ 0 & -1 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix},$$

de donde $u_3 = u_2 = u_1$. Si $u_1 = 1$, $u_3 = u_2 = 1$. Así, pues, $v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ es un vector propio de la matriz A asociado al valor propio $\lambda_3 = 4$.

Por el Teorema 4 de la página 160 los 3 vectores propios, por ser distintos entre sí, son linealmente independientes. Por el Corolario 1 de la página 169 la matriz A es diagonalizable. Es decir, existe una matriz P inversible tal que

$$D = P^{-1}AP$$

donde las columnas de la matriz P son los vectores propios de la matriz A y los elementos de la diagonal de la matriz diagonal D son los valores propios de la matriz A. Sea la matriz P formada con los vectores propios v_1, v_2 y v_3 de la matriz A:

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

Su matriz inversa es

$$P^{-1} = \frac{1}{6}\begin{pmatrix} 1 & -2 & 1 \\ 3 & 0 & -3 \\ 2 & 2 & 2 \end{pmatrix}.$$

Se deja como ejercicio verificar que la matriz $P^{-1}AP$ es la matriz diagonal

$$D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Por ser D una matriz diagonal tenemos que

$$D^n = 2^n \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2^n \end{pmatrix}.$$

Por el ejercicio anterior tenemos que

$$\begin{aligned} A^n &= PD^nP - 1 \\ &= \frac{2^n}{6}\begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2^n \end{pmatrix}\begin{pmatrix} 1 & -2 & 1 \\ 3 & 0 & -3 \\ 2 & 2 & 2 \end{pmatrix} \\ &= \frac{2^n}{6}\begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} (-1)^n & -2(-1)^n & (-1)^n \\ 0 & 0 & 0 \\ 2^{n+1} & 2^{n+1} & 2^{n+1} \end{pmatrix} \\ &= \frac{2^n}{6}\begin{pmatrix} 2^{n+1} + (-1)^n & 2^{n+1} - 2(-1)^n & 2^{n+1} + (-1)^n \\ 2^{n+1} - 2(-1)^n & 2^{n+1} + 4(-1)^n & 2^{n+1} - 2(-1)^n \\ 2^{n+1} + (-1)^n & 2^{n+1} - 2(-1)^n & 2^{n+1} + (-1)^n \end{pmatrix}. \end{aligned}$$

□

Tabla de contenidos