Introducción al álgebra lineal

Juan **Rada**
Profesor del departamento de Matemática
Facultad de Ciencias de la Universidad de Los Andes

Introducción al álgebra lineal

Universidad de Los Andes
Consejo de Publicaciones
2011

Título de la obra: **Introducción al álgebra lineal**

Autor: Juan **Rada**

Coeditado por la Comisión de Desarrolllo de Pregrado (CODEPRE), Fundación Polar
y el Consejo de Publicaciones de la Universidad de Los Andes
Av. Andrés Bello, antiguo CALA. La Parroquia
Mérida, estado Mérida, Venezuela
Telefax: (+58 274) 2711955, 2713210, 2712034
e-mail: cpula@ula.ve
http://www.ula.ve/cp

Colección: Ciencias Básicas
Serie: Matemática
1ª edición. 2011

Diagramación: Antonio Vizcaya y Edgar Iturriaga
Diseño de portada: Leroy Rojas

Hecho el depósito de ley
Depósito legal: lf23720115102044
ISBN 978-980-11-1378-2

Impreso en Gráficas El Portatítulo
Mérida, Venezuela, 2011

A MAAS.

ÍNDICE GENERAL

PRÓLOGO

El álgebra lineal es una rama de las matemáticas que estudia los espacios vectoriales y las transformaciones lineales. Estos conceptos han contribuido notablemente en el desarrollo del conocimiento dentro de las matemáticas y también en otras ciencias, especialmente en las ciencias básicas, la economía, la informática, la ingeniería y las ciencias sociales. Por eso se justifica el estudio del álgebra lineal en la mayoría de las carreras universitarias.

El material contenido en estas notas tiene como objetivo iniciar al estudiante en los conceptos básicos del álgebra lineal. Se trata de un curso dirigido a estudiantes de segundo año de las carreras de matemática, física e ingeniería, que han estudiado previamente nociones básicas de cálculo diferencial e integral y álgebra. Está escrito en un estilo matemático riguroso, en el cual los teoremas son presentados con precisión y están seguidos por sus demostraciones formales; es posible cubrir el material completo en un semestre.

Estas notas surgieron de varios cursos de álgebra lineal que dicté en la Universidad de Los Andes y en la Universidad Simón Bolívar durante los últimos años. Expreso mi agradecimiento a los colegas José Luis Chacón, Olga Porras, Leonel Mendoza, María González y Aurora Olivieri por usar estas notas y hacer comentarios interesantes que mejoraron las primeras versiones.

CAPÍTULO 1

MATRICES

1.1. Definición y terminología

Comenzamos recordando el álgebra de los números complejos. El conjunto de los números complejos se denota por $\mathbb{C}$ y está formado por pares ordenados (a, b), en los cuales $a, b \in \mathbb{R}$. Decimos que dos números complejos son iguales cuando son iguales coordenada a coordenada:

$$(a, b) = (c, d) \text{ si, y sólo si, } a = c \text{ y } b = d$$

Definimos dos operaciones en $\mathbb{C}$, llamadas suma y producto, de la siguiente manera:

$$\begin{aligned} (a, b) + (c, d) &= (a + c, b + d) \\ (a, b)(c, d) &= (ac - bd, ad + bc) \end{aligned}$$

El teorema que sigue contiene las propiedades más notables de estas operaciones.

Teorema 1.1.1 *Las siguientes propiedades se cumplen:*

1. $(a, b) + (c, d) = (c, d) + (a, b)$;
2. $[(a, b) + (c, d)] + (e, f) = (a, b) + [(c, d) + (e, f)]$;
3. $(0, 0) + (a, b) = (a, b)$;
4. $(a, b) + (-a, -b) = (0, 0)$;
5. $(a, b)(c, d) = (c, d)(a, b)$;
6. $[(a, b)(c, d)](e, f) = (a, b)[(c, d)(e, f)]$;
7. $(1, 0)(a, b) = (a, b)$;

8. *Si* $(a,b) \neq (0,0)$ *entonces* $(a,b)\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = (1,0)$;

9. $(a,b)[(c,d)+(e,f)] = (a,b)(c,d) + (a,b)(e,f)$;

10. $[(a,b)+(c,d)](e,f) = (a,b)(e,f) + (c,d)(e,f)$.

Demostración. Demostramos 8, las demás las dejamos como ejercicio. Como $(a,b) \neq (0,0)$, entonces $a^2+b^2 \neq 0$ y así, $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$ está bien definido. Además,

$$\begin{aligned} (a,b)\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) &= \left(a\frac{a}{a^2+b^2} - b\frac{-b}{a^2+b^2}, a\frac{-b}{a^2+b^2} + b\frac{a}{a^2+b^2}\right) \\ &= \left(\frac{a^2+b^2}{a^2+b^2}, \frac{-ab+ba}{a^2+b^2}\right) = (1,0) \end{aligned}$$

■

Por otra parte se comprueba fácilmente que

$$\begin{aligned} (a,0)+(b,0) &= (a+b,0) \\ (a,0)(b,0) &= (ab,0) \end{aligned}$$

Además, como la función $\iota : \mathbb{R} \to \mathbb{C}$ definida por $\iota(a) = (a,0)$ es inyectiva, entonces es posible considerar a $\mathbb{R}$ como subconjunto de $\mathbb{C}$; simplemente identificamos a $\mathbb{R}$ como los números complejos de la forma $(a,0)$. De esta manera podemos escribir

$$\begin{aligned} (a,b) &= (a,0)+(0,b) = (a,0)+(b,0)(0,1) \\ &= a+bi \end{aligned}$$

donde denotamos a $(0,1)$ por i. Notamos que i satisface

$$i^2 = (0,1)(0,1) = (-1,0) = -1$$

Luego cada número complejo z se representa de manera única como $z = a+bi$. En esta representación llamamos a la **parte real** y b la **parte imaginaria** de z y los denotamos por

$$Re(z) = a \text{ y } Im(z) = b$$

Con esta notación, las operaciones suma y producto se convierten en

$$\begin{aligned} (a+bi)+(c+di) &= (a+c)+(b+d)i \\ (a+bi)(c+di) &= (ac-bd)+(ad+bc)i \end{aligned}$$

A partir de ahora usamos la notación $z = a+bi$ para denotar un número complejo.

Si $z = a+bi$, entonces definimos el **conjugado** de z, denotado por $\overline{z}$, al número complejo

$$\overline{z} = a-bi$$

En este caso tenemos que

$$z\overline{z} = (a + bi)(a - bi) = a^2 + b^2$$

Observamos que $z\overline{z}$ es precisamente la longitud al cuadrado del vector que z representa en el plano real, es decir, el vector que tiene origen en $(0, 0)$ y extremo en (a, b). Denotamos por $|z|$ a esta longitud, es decir,

$$|z| = \sqrt{a^2 + b^2}$$

y la llamamos la **norma** del número complejo z.

Si θ es el ángulo que forma el vector que z representa en el plano real con el eje x, entonces se tiene que

$$a = rcos\theta \text{ y } b = rsen\theta$$

y en consecuencia, $z = a + bi = r(cos\theta + isen\theta)$. Esta es la **forma polar** del número complejo z.

A lo largo de estas notas, $\mathbb{F}$ denotará el conjunto de los números reales o el conjunto de los números complejos. Cualquier afirmación relativa a $\mathbb{F}$ significa que es válida en los dos casos: $\mathbb{F} = \mathbb{R}$ o $\mathbb{F} = \mathbb{C}$. Diremos que un elemento de $\mathbb{F}$ es un **escalar**.

Sean m y n enteros mayores o iguales a 1. Una **matriz** es un arreglo rectangular de elementos de $\mathbb{F}$ de la forma

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Las m *n-uplas*

$$(a_{11}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn})$$

son las **filas** de la matriz y las n *m-uplas*

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \cdots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

son las **columnas** de la matriz. Una matriz A con m filas y n columnas es una matriz de **tamaño** $m \times n$. En este caso denotamos las m filas de A por

$$A_{1*}, A_{2*}, \ldots, A_{m*}$$

y las n columnas de A por

$$A_{*1}, A_{*2}, \ldots, A_{*n}$$

El elemento a_{ij} de la matriz A aparece en la fila i y en la columna j de A. También usamos la notación $[A]_{ij}$ para denotar el elemento ij de la matriz A. Los elementos $[A]_{ii}$,

$$i = 1, \ldots, r = min\{m, n\}$$

son los **elementos de la diagonal** de A.

Dos matrices A y B son **iguales** si tienen el mismo tamaño y $[A]_{ij} = [B]_{ij}$ para todo i , j.

Denotamos por $\mathcal{M}_{m\times n}(\mathbb{F})$ al conjunto de todas las matrices $m \times n$ con elementos en $\mathbb{F}$. Si $m = n$, simplemente escribimos $\mathcal{M}_n(\mathbb{F})$ y sus matrices las llamamos **matrices cuadradas** (de tamaño n).

EJERCICIOS

En los siguientes ejercicios, w y z pertenecen a $\mathbb{C}$. Demuestre que:

1. $\overline{\overline{z}} = z$.
2. $\overline{z+w} = \overline{z} + \overline{w}$.
3. $\overline{zw} = \overline{z}\overline{w}$.
4. Defina $\frac{w}{z}$. Demuestre que $\overline{\frac{w}{z}} = \frac{\overline{w}}{\overline{z}}$.
5. $z \in \mathbb{R}$ si, y sólo si, $\overline{z} = z$.
6. $|z| = 0$ si, y sólo si, $z = 0$.
7. $|z| = |\overline{z}|$.
8. $|zw| = |z|\,|w|$.
9. Si $z \neq 0$ entonces $\left|\frac{1}{z}\right| = \frac{1}{|z|}$.
10. $|z+w| \leq |z| + |w|$.
11. Representando a z y w en sus formas polares, encuentre las fórmulas para zw y $\frac{z}{w}$. En particular, para $\frac{1}{z}$.
12. Si $z = r(cos\theta + isen\theta)$, entonces $z^n = r^n(cos(n\theta) + isen(n\theta))$ para todo entero positivo n.
13. Complete la demostración del teorema 1.1.1.

1.2. Álgebra de matrices

Definimos a continuación dos operaciones sobre el conjunto $\mathcal{M}_{m\times n}(\mathbb{F})$.

Definición 1.2.1 *Sean $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$. La* ***suma*** *de A y B, denotada por $A+B$, es la matriz $m \times n$ definida por*

$$[A+B]_{ij} = [A]_{ij} + [B]_{ij}\,.$$

Si $\alpha \in \mathbb{F}$, la ***multiplicación*** *del escalar α por la matriz A, denotado por αA, es la matriz $m \times n$ definida por*

$$[\alpha A]_{ij} = \alpha\,[A]_{ij}$$

Ejemplo 1.2.2 *Consideremos las matrices*

$$A = \begin{pmatrix} 3 & 2 & i \\ 0 & -2 & 4 \\ 8 & i+1 & 6 \end{pmatrix} \text{ y } B = \begin{pmatrix} 5 & -2 & 4 \\ \pi & -2 & 4 \\ -6 & -1 & 3 \end{pmatrix}$$

Entonces

$$A + B = \begin{pmatrix} 8 & 0 & i+4 \\ \pi & -4 & 8 \\ 2 & i & 9 \end{pmatrix} \text{ y } 3A = \begin{pmatrix} 9 & 6 & 3i \\ 0 & -6 & 12 \\ 24 & 3i+3 & 18 \end{pmatrix}$$

Veamos las propiedades más importantes que cumplen las operaciones recién definidas. El $\mathbf{0}$ de $\mathcal{M}_{m\times n}(\mathbb{F})$ es la matriz definida por $[\mathbf{0}]_{ij} = 0$ para todo i, j. Por otra parte, el inverso aditivo de A, que denotamos por $-A$, es la matriz definida como $[-A]_{ij} = -[A]_{ij}$. Para abreviar, de aquí en adelante escribimos $A - B$ en lugar de $A + (-B)$.

Ejemplo 1.2.3 *Si* $A = \begin{pmatrix} 3 & 2 & i \\ 0 & -2 & 4 \\ 8 & i+1 & -6 \end{pmatrix}$ *entonces* $-A = \begin{pmatrix} -3 & -2 & -i \\ 0 & 2 & -4 \\ -8 & -i-1 & 6 \end{pmatrix}$

Teorema 1.2.4 *Sean* $A, B, C \in \mathcal{M}_{m\times n}(\mathbb{F})$ *y* $\alpha, \beta \in \mathbb{F}$. *Entonces*

1. $A + B = B + A$;
2. $(A + B) + C = A + (B + C)$;
3. $A + \mathbf{0} = A$;
4. $A + (-A) = \mathbf{0}$;
5. $\alpha(A + B) = \alpha A + \alpha B$;
6. $(\alpha + \beta) A = \alpha A + \beta A$;
7. $(\alpha\beta) A = \alpha(\beta A)$;
8. $1A = A$.

Demostración. Vamos a demostrar la propiedad 7, dejamos las demostraciones restantes para que el lector las lleve a cabo. Para todo i, j tenemos

$$[(\alpha\beta) A]_{ij} = (\alpha\beta)[A]_{ij} = \alpha\left(\beta[A]_{ij}\right) = \alpha\left([\beta A]_{ij}\right) = [\alpha(\beta A)]_{ij}$$

Por lo tanto, $(\alpha\beta) A = \alpha(\beta A)$. ■

Llamamos a $\mathcal{M}_{m\times n}(\mathbb{F})$ junto con las operaciones definidas anteriormente el espacio de las matrices $m \times n$ con coeficientes en $\mathbb{F}$.

Ahora introducimos la multiplicación de matrices. En un principio, esta definición parece innecesariamente complicada, pero quedará plenamente justificada en las secciones posteriores.

Definición 1.2.5 *Sean $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $B \in \mathcal{M}_{n\times p}(\mathbb{F})$. La* ***multiplicación*** *de A por B, denotada por AB, es la matriz de $\mathcal{M}_{m\times p}(\mathbb{F})$ definida como*

$$[AB]_{ij} = \sum_{k=1}^{n} [A]_{ik} [B]_{kj}$$

Observamos que la multiplicación de dos matrices está definida cuando el número de columnas de la matriz de la izquierda es igual al número de filas de la matriz de la derecha.

Ejemplo 1.2.6 *Sean* $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & -1 \end{pmatrix}$ *y* $B = \begin{pmatrix} 2 & 1 \\ -2 & 0 \\ 1 & 1 \end{pmatrix}$*. Entonces*

$$AB = \begin{pmatrix} 2 & 3 \\ -3 & 1 \end{pmatrix} \text{ y } BA = \begin{pmatrix} 4 & 5 & 3 \\ -2 & -2 & -4 \\ 3 & 4 & 1 \end{pmatrix}$$

Es claro que la única posibilidad para que AB y BA estén ambas bien definidas es que $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $B \in \mathcal{M}_{n\times m}(\mathbb{F})$. Si $m \neq n$, entonces AB y BA tienen distinto tamaño y por lo tanto $AB \neq BA$.Aun en el caso que A y B sean matrices cuadradas del mismo tamaño, la multiplicación no es conmutativa, como muestra el siguiente ejemplo.

Ejemplo 1.2.7 *Sean* $A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$ *y* $B = \begin{pmatrix} 4 & 3 \\ -1 & 2 \end{pmatrix}$*. Entonces*

$$AB = \begin{pmatrix} 2 & 7 \\ 8 & 6 \end{pmatrix} \neq \begin{pmatrix} 10 & 8 \\ 3 & -2 \end{pmatrix} = BA$$

A continuación presentamos una lista de propiedades básicas que satisface la multiplicación de matrices. Suponemos en el próximo teorema que los tamaños de las matrices son los adecuados.

Teorema 1.2.8 *Sean A, B, C matrices. Entonces*

1. $(AB)C = A(BC)$;
2. $A(B + C) = AB + AC$;
3. $(B + C)A = BA + CA$;
4. $\lambda(AB) = (\lambda A)B = A(\lambda B)$ *para todo* $\lambda \in \mathbb{F}$;

Demostración. Demostramos 1, el resto de las demostraciones las dejamos como ejercicio. Para cada i, j tenemos

$$\begin{aligned}
[(AB)\,C]_{ij} &= \sum_{k=1}^{p} [AB]_{ik}\,[C]_{kj} = \sum_{k=1}^{p}\left(\sum_{l=1}^{n} [A]_{il}\,[B]_{lk}\right)[C]_{kj} \\
&= \sum_{k=1}^{p}\sum_{l=1}^{n} [A]_{il}\,[B]_{lk}\,[C]_{kj} = \sum_{l=1}^{n}\sum_{k=1}^{p} [A]_{il}\,[B]_{lk}\,[C]_{kj} \\
&= \sum_{l=1}^{n} [A]_{il}\sum_{k=1}^{p} [B]_{lk}\,[C]_{kj} = \sum_{l=1}^{n} [A]_{il}\,[BC]_{lj} = [A\,(BC)]_{ij}
\end{aligned}$$

■

La matriz **identidad** de $\mathcal{M}_n(\mathbb{F})$ se denota por I_n y se define como:

$$[I_n]_{ij} = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j \end{cases}, \text{ esto es, } I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Teorema 1.2.9 *Si $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, entonces $I_m A = A$ y $AI_n = A$.*

Demostración. Demostramos que $I_m A = A$. En efecto,

$$[I_m A]_{ij} = \sum_{k=1}^{m} [I_m]_{ik}\,[A]_{kj} = [I_m]_{ii}\,[A]_{ij} = [A]_{ij}$$

para todo $1 \leq i \leq m$ y $1 \leq j \leq n$. ■

Definición 1.2.10 *Una matriz $A \in \mathcal{M}_n(\mathbb{F})$ tiene* ***inversa*** *(o es* ***invertible****) si existe $B \in \mathcal{M}_n(\mathbb{F})$ tal que $AB = BA = I_n$. En este caso decimos que B es una inversa de A.*

Existen matrices que no tienen inversas. Por ejemplo, la matriz $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{F})$ no es invertible porque

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} \neq I_2$$

para todo $a, b, c, d \in \mathbb{F}$.

Teorema 1.2.11 *La inversa de una matriz, si existe, es única.*

Demostración. Supongamos que $A \in \mathcal{M}_n(\mathbb{F})$ y B, C son matrices inversas de A. Entonces $AB = I_n = CA$. Luego, $B = I_n B = (CA)\,B = C\,(AB) = CI_n = C$. ■

En virtud del resultado anterior hablamos de la inversa de una matriz A y la denotamos por A^{-1}. En general, el cálculo de la inversa de una matriz resulta largo y tedioso. En secciones posteriores estudiaremos algunos métodos para calcularla.

Teorema 1.2.12 *Sean $A, B \in \mathcal{M}_n(\mathbb{F})$ invertibles. Entonces*

1. $(A^{-1})^{-1} = A$;
2. *AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.*

Demostración. Demostramos 2. Tenemos que

$$\begin{aligned}(AB)\left(B^{-1}A^{-1}\right) &= A\left(BB^{-1}\right)A^{-1} = I_n \\ \left(B^{-1}A^{-1}\right)(AB) &= B^{-1}\left(A^{-1}A\right)B = I_n\end{aligned}$$

Luego

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

. ■

Dado $A \in \mathcal{M}_n(\mathbb{F})$ definimos $A^0 = I_n$ y para $m \geq 1$, definimos recursivamente $A^m = AA^{m-1}$. Si r, s son enteros no negativos, entonces se demuestra por inducción que $A^rA^s = A^sA^r = A^{r+s}$ (Ejercicio 4).

Definición 1.2.13 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. La **traspuesta** de A, denotada por $A^\top$, es la matriz de $\mathcal{M}_{n\times m}(\mathbb{F})$ que se obtiene al intercambiar las filas por columnas:*

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}^\top = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

Formalmente, $\left[A^\top\right]_{ij} = [A]_{ji}$ para todo i, j.

Ejemplo 1.2.14 *La traspuesta de la matriz $A = \begin{pmatrix} 3 & i & 2-i \\ 1 & 2 & 3 \\ 5 & 2i & 5-i \end{pmatrix}$ es*

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

Las propiedades básicas de la traspuesta las presentamos en el siguiente resultado. Suponemos que los tamaños de las matrices son adecuados para efectuar las operaciones.

Teorema 1.2.15 *Las siguientes condiciones se cumplen:*

1. $\left(A^\top\right)^\top = A$;
2. $(A+B)^\top = A^\top + B^\top$;

3. $(\alpha A)^{\top} = \alpha A^{\top}$;

4. $(AB)^{\top} = B^{\top} A^{\top}$;

5. *Si* A *es invertible, entonces* $A^{\top}$ *es invertible y* $\left(A^{\top}\right)^{-1} = \left(A^{-1}\right)^{\top}$.

Demostración. Vamos a demostrar las propiedades 4. y 5.
4. Para todo i, j tenemos

$$\begin{aligned} \left[(AB)^{\top}\right]_{ij} &= [AB]_{ji} = \sum_k [A]_{jk} [B]_{ki} = \sum_k \left[A^{\top}\right]_{kj} \left[B^{\top}\right]_{ik} \\ &= \sum_k \left[B^{\top}\right]_{ik} \left[A^{\top}\right]_{kj} = \left[B^{\top} A^{\top}\right]_{ij} \end{aligned}$$

5. Supongamos que A es invertible. Entonces, por la parte 4 tenemos

$$\begin{aligned} \left(A^{-1}\right)^{\top} A^{\top} &= \left(AA^{-1}\right)^{\top} = I^{\top} = I \\ A^{\top} \left(A^{-1}\right)^{\top} &= \left(A^{-1}A\right)^{\top} = I^{\top} = I \end{aligned}$$

Por lo tanto, $A^{\top}$ es invertible y la inversa de $A^{\top}$ es $\left(A^{-1}\right)^{\top}$. ■

Recordamos que $\overline{z}$ denota el conjugado del número complejo z.

Definición 1.2.16 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. *La* ***conjugada*** *de* A *la denotamos por* $\overline{A}$ *y se define por* $\left[\overline{A}\right]_{ij} = \overline{[A]_{ij}}$. *La* ***traspuesta conjugada*** *o* ***traspuesta hermitiana****, denotada por* A^**, se define como* $A^* = \overline{A^{\top}}$.

Claramente, si $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ entonces $A^* = A^{\top}$.

Ejemplo 1.2.17 *La matriz* $A = \begin{pmatrix} 3 & i & 2-i \\ 1 & 2 & 3 \\ 5 & 2i & 5-i \end{pmatrix}$ *tiene traspuesta conjugada*

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

Teorema 1.2.18 *Las siguientes condiciones se cumplen:*

1. $\overline{AB} = \overline{A}\,\overline{B}$;

2. $(A^*)^* = A$;

3. $(A+B)^* = A^* + B^*$;

4. $(\alpha A)^* = \overline{\alpha} A^*$;

5. $(AB)^* = B^*A^*$.

6. *Si* A *es invertible, entonces* A^* *es invertible y* $(A^*)^{-1} = (A^{-1})^*$.

Demostración. Demostramos las propiedades 5. y 6.
5. $(AB)^* = \overline{(AB)^\top} = \overline{B^\top A^\top} = \overline{B^\top}\,\overline{A^\top} = B^*A^*$.
6. Supongamos que A es invertible. Entonces por la parte 5 tenemos

$$\begin{aligned} \left(A^{-1}\right)^* A^* &= \left(AA^{-1}\right)^* = I^* = I \\ A^* \left(A^{-1}\right)^* &= \left(A^{-1}A\right)^* = I^* = I \end{aligned}$$

El resto lo dejamos como ejercicio. ■

EJERCICIOS

1. Complete las demostraciones de los teoremas de esta sección.

2. Dadas las matrices

$$A = \begin{pmatrix} 3+i & 5 & -2 \\ 1 & 0 & 2-3i \end{pmatrix} \text{ y } B = \begin{pmatrix} 3 & 1+i & 0 \\ i & 1 & -3i \end{pmatrix}$$

calcule $2iA - 5B$.

3. Dadas las matrices

$$A = \begin{pmatrix} 3+i & 5 & -2 \\ 1 & 0 & 2-3i \end{pmatrix}, B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \text{ y } C = \begin{pmatrix} i & -i \end{pmatrix}$$

calcule ABC y CAB.

4. Si A es una matriz, demuestre que para todo entero no-negativo r y s se tiene que

$$A^rA^s = A^sA^r = A^{r+s}$$

.

5. Sea $B = \begin{pmatrix} 2 & 3 & 1 & -1 \\ 5 & -2 & 0 & 1 \\ 1 & 2 & -2 & 0 \end{pmatrix}$. Encuentre Be_i para $i = 1, \ldots, 4$ donde e_i viene dada por:

$$\text{a) } e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \text{ b) } e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \text{ c) } e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ y c) } e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

6. Encuentre para la matriz B del ejercicio anterior $e_i B$, donde e_i viene dada por: a) $e_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$; b) $e_2 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$ y c) $e_3 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$.

7. Generalice los ejercicios 4 y 5 para matrices en $\mathcal{M}_{m\times n}(\mathbb{F})$.

8. Demuestre que $(AB)_{*j} = AB_{*j}$ para todas las matrices de tamaño adecuado.

9. Demuestre que $(AB)_{i*} = A_{i*}B$ para todas las matrices de tamaño adecuado.

10. Demuestre que si $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $x = (x_1, \ldots, x_n)^\top$, entonces $Ax = x_1A_{*1} + \cdots + x_nA_{*n}$.

11. Demuestre que si $A \in \mathcal{M}_n(\mathbb{F})$ satisface $AB = BA$ para todo $B \in \mathcal{M}_n(\mathbb{F})$, entonces A es un múltiplo escalar de la identidad.

12. Encuentre una condición necesaria y suficiente para que una matriz $B \in \mathcal{M}_2(\mathbb{R})$ sea invertible y calcule la inversa cuando existe.

13. Demuestre que si $A \in \mathcal{M}_n(\mathbb{F})$ y $A^2 = 0$, entonces $I - A$ es invertible.

14. Demuestre que si A es invertible y $AB = 0$, entonces $B = 0$.

15. Una **matriz diagonal** es una matriz de la forma

$$D = \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & d_{nn} \end{pmatrix}$$

¿Cuándo es D invertible?. En ese caso, ¿cuál es la inversa?

16. Considere la matriz diagonal $D = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ y $A = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}$. Calcule $(A^{-1}DA)^{100}$.

17. Dadas las matrices

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

Calcule $(AB)^\top, B^\top A^\top, (AB)^*$ y B^*A^*.

18. Una matriz cuadrada $A \in \mathcal{M}_n(\mathbb{F})$ es **simétrica** si $A^\top = A$ y **antisimétrica** si $A^\top = -A$. Demuestre que para cualquier matriz $B \in \mathcal{M}_n(\mathbb{F})$ se tiene: a) $BB^\top$ y $B+B^\top$ son matrices simétricas; b) $B - B^\top$ es antisimétrica.

19. Sea $A \in \mathcal{M}_n(\mathbb{F})$. Se define la **traza** de A, denotada por $Tr(A)$, como el elemento de $\mathbb{F}$

$$Tr(A) = \sum_{k=1}^{n} [A]_{kk}$$

Dadas $A, B \in \mathcal{M}_n(\mathbb{F})$ y $\alpha \in \mathbb{F}$ demuestre: $i)$ $Tr(A+B) = Tr(A) + Tr(B)$; $ii)$ $Tr(\alpha A) = \alpha Tr(A)$ y $iii)$ $Tr(AB) = Tr(BA)$.

20. Una matriz $T \in \mathcal{M}_n(\mathbb{F})$ es **triangular superior** si $[T]_{ij} = 0$ para $i > j$. Suponga que $S, T \in \mathcal{M}_n(\mathbb{F})$ son triangulares superiores y $\alpha \in \mathbb{F}$. Demuestre que

    $a)$ $S + T$ es triangular superior;

    $b)$ αS es triangular superior;

    $c)$ ST es triangular superior.

1.3. Sistemas de ecuaciones lineales

El objetivo de esta sección es presentar un método para resolver la ecuación matricial $AX = B$, donde $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $B \in \mathcal{M}_{m\times 1}(\mathbb{F})$. En otras palabras, queremos encontrar todas las matrices $X \in \mathcal{M}_{n\times 1}(\mathbb{F})$ que satisfacen la igualdad $AX = B$. Interpretando cada fila A_{i*} de A como un elemento de $\mathcal{M}_{1\times n}(\mathbb{F})$, la ecuación matricial $AX = B$ se traduce en el sistema de m ecuaciones lineales

$$A_{1*}X = b_1, \cdots, A_{m*}X = b_m$$

donde $[B]_{i1} = b_i \in \mathbb{F}$ para todo $1 \leq i \leq m$. También lo podemos interpretar como un sistema de m ecuaciones lineales con **incógnitas** o **variables** $x_1, \ldots, x_n$:

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

donde $[A]_{ij} = a_{ij} \in \mathbb{F}$ para todo $1 \leq i \leq m, 1 \leq j \leq n$. El sistema de ecuaciones lineales $AX = 0$ recibe el nombre de **sistema de ecuaciones lineales homogéneo**.

La técnica de eliminación de Gauss-Jordan que estudiamos más adelante en esta sección, consiste en transformar una ecuación matricial en una más sencilla pero con el mismo conjunto solución. Esto es posible mediante las operaciones elementales aplicadas a una matriz.

Definición 1.3.1 *Una **operación elemental por filas** aplicada a una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ significa llevar a cabo una de las siguientes operaciones a la matriz A :*

1. *Intercambiar la fila p por la fila q, denotada por $f_{pq}(A)$;*

2. *Muliplicar por* $\lambda \neq 0$ *a la fila* p*, denotada por* $f_p(\lambda)(A)$;

3. *Sumar a la fila* p *la fila* q *multiplicada por* λ*, denotada por* $f_{pq}(\lambda)(A)$.

Más aún, si B *se obtiene a partir de* A *a través de una sucesión finita de operaciones elementales por filas, entonces decimos que* A *y* B *son* ***equivalentes por filas****, y lo denotamos por* $A \underset{f}{\sim} B$.

Es fácil ver que la relación "equivalente por filas" es una relación de equivalencia sobre $\mathcal{M}_{m\times n}(\mathbb{F})$.

Usando operaciones elementales por filas es posible transformar una matriz en otra que tiene una forma especialmente sencilla.

Definición 1.3.2 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$*. Decimos que* A *es* ***escalonada reducida*** *si cumple con las siguientes condiciones:*

1. *Si* $A_{j*} = 0$*, entonces* $A_{k*} = 0$ *para todo* $j < k \leq m$*. Es decir, todas las filas* 0 *están en la parte de abajo de la matriz;*

2. *Sean* A_{p*} *y* A_{q*} *filas distintas de* 0 *y* $p < q$*. Si* a_{pr} *es el primer elemento distinto de cero de la fila* A_{p*} *y* a_{qs} *es el primer elemento distinto de cero de la fila* A_{q*}*, entonces* $a_{pr} = 1 = a_{qs}$ *y* $r < s$.

3. *Si* $a_{pr} = 1$ *es el primer elemento distinto de cero de la fila* A_{p*} *(llamado* 1 ***principal****), entonces todos los elementos de* A_{*r} *distinto de* a_{pr} *son ceros.*

Ejemplo 1.3.3 *Consideremos la matriz*

$$\begin{pmatrix} 2 & 4 & 6 & 2 & 6 \\ 6 & 4 & 2 & 2 & 14 \\ 0 & 4 & 8 & 2 & 2 \\ 2 & 2 & 2 & 2 & 8 \end{pmatrix}$$

Entonces, efectuando las operaciones $f_1\left(\frac{1}{2}\right), f_{21}(-6), f_{41}(-2)$ *obtenemos*

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 0 & -8 & -16 & -4 & -4 \\ 0 & 4 & 8 & 2 & 2 \\ 0 & -2 & -4 & 0 & 2 \end{pmatrix}$$

Después de $f_2\left(-\frac{1}{8}\right), f_{12}(-2), f_{32}(-4), f_{42}(2)$ *llegamos a*

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 2 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}$$

y al aplicar $f_{34}, f_{23}\left(-\frac{1}{2}\right)$ *obtenemos la matriz*

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

Esta última matriz es escalonada reducida.

El ejemplo 1.3.3 muestra cómo se lleva una matriz por medio de operaciones elementales por filas a una matriz escalonada reducida. Esto siempre es posible como muestra nuestro próximo resultado.

Teorema 1.3.4 *(Eliminación de Gauss-Jordan) Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. *Entonces existe una matriz escalonada reducida* R *tal que* $A \underset{f}{\sim} R$.

Demostración. Supongamos que $[A]_{ij} = a_{ij}$ para todo i, j y que A_{*r} es la primera columna de A (de izquierda a derecha) diferente de cero. Podemos suponer que $a_{1r} \neq 0$, de lo contrario aplicamos f_{1p} a A si $a_{pr} \neq 0$. Ahora aplicamos las operaciones elementales por filas a A

$$f_1\left(\frac{1}{a_{1r}}\right), f_{21}\left(-a_{2r}\right), \ldots, f_{m1}\left(-a_{mr}\right)$$

para obtener la matriz B que tiene las primeras $r-1$ columnas iguales a cero y los elementos de la columna B_{*r} son todos ceros excepto $b_{1r} = 1$:

$$A \underset{f}{\sim} B = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & * & \cdots & * \end{pmatrix}$$

Sea $\widehat{B}$ la matriz que se obtiene a partir de B eliminando la primera fila de B y sea $\widehat{B}_{*s}$ la primera columna de $\widehat{B}$ diferente de cero. Observamos que $r < s$ e igual que antes, podemos asumir que $\left[\widehat{B}\right]_{1s} = b_{2s} \neq 0$. Ahora aplicamos a B las operaciones elementales por filas

$$f_2\left(\frac{1}{b_{2s}}\right), f_{12}\left(-b_{1s}\right), f_{32}\left(-b_{3s}\right) \ldots, f_{m2}\left(-b_{ms}\right)$$

para obtener una matriz C de la forma

$$A \underset{f}{\sim} B \underset{f}{\sim} C = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & * & * & \cdots & * \end{pmatrix}$$

Continuamos este procedimiento y es claro que después de un número finito de pasos llegamos a una matriz escalonada reducida. ■

Una aplicación de la eliminación de Gauss-Jordan junto con el próximo resultado nos proporciona una técnica para encontrar todas las soluciones de una ecuación matricial $AX = B$, donde $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $B \in \mathcal{M}_{m\times 1}(\mathbb{F})$.

Teorema 1.3.5 *Sean $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, $B \in \mathcal{M}_{m\times 1}(\mathbb{F})$ y f una operación elemental por fila. Entonces, las ecuaciones $AX = B$ y $f(A)X = f(B)$ tienen la misma solución.*

Demostración. Es claro cuando $f = f_{pq}$ o $f = f_p(\lambda)$. Supongamos que $f = f_{pq}(\lambda)$. Si $AX = B$, donde $[B]_{ij} = b_i$, entonces $A_{i*}X = b_i$ para todo i. Luego

$$(A_{p*} + \lambda A_{q*})X = A_{p*}X + \lambda A_{q*}X = b_p + \lambda b_q$$

y, en consecuencia, $f_{pq}(\lambda)(A)X = f_{pq}(\lambda)(B)$. Recíprocamente, supongamos que

$$f_{pq}(\lambda)(A)X = f_{pq}(\lambda)(B).$$

Luego $A_{k*}X = b_k$ para todo $k \neq p$ y

$$(A_{p*} + \lambda A_{q*})X = b_p + \lambda b_q$$

Equivalentemente,

$$A_{p*}X + \lambda A_{q*}X = b_p + \lambda b_q$$

Pero como $q \neq p$, entonces $A_{q*}X = b_q$ y así, $A_{p*}X = b_p$. Esto demuestra que $AX = B$. ■

El problema de resolver una ecuación matricial de la forma $AX = B$ se reduce a resolver una ecuación de la forma $EX = C$, donde E es escalonada reducida. La idea es la siguiente: si tenemos la ecuación $AX = B$, entonces, por la eliminación de Gauss-Jordan, efectuamos operaciones elementales por filas a A hasta convertirla en una matriz escalonada reducida E. Las mismas operaciones llevan la **matriz ampliada** (A, B) a la matriz ampliada (E, C). Por el teorema 1.3.5, las ecuaciones $AX = B$ y $EX = C$ tienen la misma solución.

Teorema 1.3.6 *Consideremos la ecuación matricial*

$$EX = C \tag{1.1}$$

donde $C = (c_1, \ldots, c_m)^\top$ y $E \in \mathcal{M}_{m\times n}(\mathbb{F})$ es una matriz escalonada reducida con elementos $[E]_{ij} = e_{ij}$. Supongamos además que E tiene r filas distintas de cero y que $e_{ik_i} = 1$ es el 1 principal de cada fila E_{i*}, para todo $1 \leq i \leq r$. Entonces:

1. *Si $c_j \neq 0$ para algún $j > r$, entonces la ecuación 1.1 no tiene solución.*

2. *Si $c_j = 0$ para todo $j > r$, entonces*

a) *$r = n$ implica que $E = \begin{pmatrix} I_n \\ \mathbf{0} \end{pmatrix}$, donde $\mathbf{0}$ es la matriz cero de tamaño $(m-n) \times n$, y la solución única de la ecuación 1.1 es $(c_1, \ldots, c_n)^\top$.*

b) *$r < n$ implica que la ecuación 1.1 tiene infinitas soluciones: para cada*

$$p \in \{1, \ldots, n\} \setminus \{k_1, \ldots, k_r\}$$

asignamos cualquier valor a x_p y luego calculamos x_{k_s} para

$$s = r, r-1, \ldots, 1$$

en este orden, mediante la fórmula:

$$x_{k_s} = \left[c_s - \sum_{k_s < j \leq n} e_{sj} x_j \right] \tag{1.2}$$

Demostración. Por ser E escalonada reducida la ecuación matricial (1.1) tiene la forma

$$\begin{array}{ccccc} x_{k_1} & + & \sum\limits_{k_1 < j \leq n} e_{1j} x_j & = & c_1 \\ \vdots & & \vdots & & \vdots \\ x_{k_r} & + & \sum\limits_{k_r < j \leq n} e_{rj} x_j & = & c_r \\ & & 0 & = & c_{r+1} \\ & & \vdots & & \vdots \\ & & 0 & = & c_m \end{array}$$

1. Es claro que si $c_j \neq 0$ para algún $j > r$, entonces el sistema de ecuaciones lineales no tiene solución.

2. Supongamos entonces que $c_j = 0$ para todo $j > r$.

a. Si $r = n$, entonces, como $1 \leq k_1 < k_2 < \cdots < k_n \leq n$ resulta $k_i = i$ para todo $i = 1, \ldots, n$ y así, al ser E escalonada reducida necesariamente

$$E = \begin{pmatrix} I_n \\ \mathbf{0} \end{pmatrix}$$

donde $\mathbf{0}$ es la matriz $(m-n) \times n$ formada por ceros. Se deduce inmediatamente que la única solución es $(c_1, \ldots, c_n)^\top$.

b. Si $r < n$ entonces $\{1, \ldots, n\} \setminus \{k_1, \ldots, k_r\} \neq \emptyset$. Asignemos un valor en $\mathbb{F}$ a x_p para cada $p \in \{1, \ldots, n\} \setminus \{k_1, \ldots, k_r\}$. Como $k_1 < \cdots < k_r$, si $j > k_r$, entonces x_j tiene un valor en $\mathbb{F}$ asignado y podemos calcular

$$x_{k_r} = \left[c_r - \sum_{k_r < j \leq n} e_{rj} x_j \right]$$

Seguimos hacia atrás tomando las ecuaciones que se obtienen para $s = r-1, r-2, \ldots, 1$, en este orden, en cada paso sustituimos los valores de x_j para $j > k_s$ que ya hemos calculado y determinamos el valor de x_{k_s} :

$$x_{k_s} = \left[c_s - \sum_{k_s < j \leq n} e_{sj} x_j \right]$$

■

En el teorema anterior, las variables x_p, donde $p \in \{1, \ldots, n\} \setminus \{k_1, \ldots, k_r\}$, asociadas a la ecuación matricial $EX = C$, reciben el nombre de **variables libres**. Estas son las variables que corresponden a las columnas de E que no contienen un 1 principal; las variables $x_{k_1}, \ldots, x_{k_r}$ se llaman **variables dependientes** de la ecuación matricial $EX = C$.

Ejemplo 1.3.7 *Vamos a resolver el sistema de ecuaciones*

$$\begin{array}{ccccccc} x & + & 2y & - & 3z & = & a \\ 2x & + & 6y & - & 11z & = & b \\ x & - & 2y & + & 7z & = & c \end{array}$$

Primero construimos la matriz ampliada (A, B)

$$\begin{pmatrix} 1 & 2 & -3 & a \\ 2 & 6 & -11 & b \\ 1 & -2 & 7 & c \end{pmatrix}$$

y aplicamos sucesivamente las operaciones elementales

$$f_{21}(-2), f_{31}(-1), f_2\left(\frac{1}{2}\right), f_{12}(-2), f_{32}(4)$$

para llegar a la matriz

$$\begin{pmatrix} 1 & 0 & 2 & 3a-b \\ 0 & 1 & -\frac{5}{2} & \frac{b-2a}{2} \\ 0 & 0 & 0 & c-5a+2b \end{pmatrix}$$

Luego la ecuación $AX = B$ *tiene la misma solución que la ecuación* $EX = C$*, donde*

$$E = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -\frac{5}{2} \\ 0 & 0 & 0 \end{pmatrix} \text{ y } C = \begin{pmatrix} 3a-b \\ \frac{b-2a}{2} \\ c-5a+2b \end{pmatrix}$$

Como E *es escalonada reducida aplicamos la primera parte del teorema 1.3.6 para concluir que si* $c - 5a + 2b \neq 0$*, entonces la ecuación no tiene solución. Si* $c - 5a + 2b = 0$ *entonces, por la parte 2b. del teorema 1.3.6, deducimos que la ecuación tiene infinitas soluciones porque el número de filas distintas de cero de* E *es menor que el número de columnas:* $2 < 3$*. En este caso, la variable* z *es la única variable libre, luego las soluciones de* $EX = C$ *son de la forma*

$$\left(3a - b - 2z_0, \frac{b-2a}{2} + \frac{5}{2}z_0, z_0\right)^\top$$

donde z_0 *toma cualquier valor en* $\mathbb{F}$*.*

Corolario 1.3.8 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ *y* $B \in \mathcal{M}_{m\times 1}(\mathbb{F})$. *Si* $m < n$, *entonces el sistema de ecuaciones lineales homogéneo* $AX = 0$ *tiene infinitas soluciones.*

Demostración. Sea E una matriz escalonada reducida equivalente por filas a A. Entonces, las ecuaciones $AX = 0$ y $EX = 0$ tienen la misma solución. Ahora, el número de filas r distintas de cero de E satisface $r \leq m < n$. Se deduce del teorema 1.3.6 que $EX = 0$ tiene infinitas soluciones. ■

Finalizamos esta sección con el importante hecho de que si $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, entonces existe una *única* matriz escalonada reducida R tal que $A \underset{f}{\sim} R$.

Teorema 1.3.9 *Si* $R, S \in \mathcal{M}_{m\times n}(\mathbb{F})$ *son matrices escalonadas reducidas y* $R \underset{f}{\sim} S$, *entonces* $R = S$.

Demostración. Primero que todo observamos que por el teorema 1.3.5, las ecuaciones matriciales $RX = 0$ y $SX = 0$ tienen la misma solución. Ahora supongamos que R tiene un 1 principal en la columna i. Entonces, una de las ecuaciones del sistema $RX = 0$ es

$$x_i + r_{i+1}x_{i+1} + \cdots + r_n x_n = 0$$

y así, este sistema de ecuaciones no tiene solución con $x_i = 1$, $x_{i+1} = \cdots = x_n = 0$. Ahora, si no hubiera un 1 principal en la columna i de S, entonces x_i sería una variable libre, así, por el teorema 1.3.6 (parte 2b) podemos construir una solución con $x_i = 1$, $x_{i+1} = \cdots = x_n = 0$ para el sistema de ecuaciones $SX = 0$. Esto es una contradicción. Así, los 1 principales de R y S están ubicados en las mismas columnas. Vamos a demostrar ahora que las filas de R y S son iguales. Consideremos la fila de R

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & r_{i+1} & r_{i+2} & \cdots & r_n \end{pmatrix}$$

y la correspondiente fila de S

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & s_{i+1} & s_{i+2} & \cdots & s_n \end{pmatrix}$$

Sea k tal que $i + 1 \leq k \leq n$. Si hay un 1 principal en la columna k de R (o S), entonces, claramente, $r_k = s_k = 0$. En caso contrario, x_k sería una variable libre y de nuevo, por la parte 2b del teorema 1.3.6, la ecuación $RX = 0$ tiene una solución con $x_k = 1$, $x_p = 0$ para todo $k \neq p \geq i + 1$ y $x_i = -r_k$. Esta es también una solución de la ecuación $SX = 0$, lo cual implica

$$1(-r_k) + s_{i+1}0 + \cdots + s_{k-1}0 + s_k 1 + s_{k+1}0 + \cdots + s_n 0 = 0$$

y así $r_k = s_k$. Demostramos de esta manera que R y S tienen filas idénticas y por lo tanto, $R = S$. ■

Corolario 1.3.10 *Si* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, *entonces existe una* única *matriz escalonada reducida* R *tal que* $A \underset{f}{\sim} R$.

Demostración. Es consecuencia inmediata del teorema 1.3.4, el teorema anterior y el hecho de que la relación “ $\underset{f}{\sim}$ ” es una relación de equivalencia sobre $\mathcal{M}_{m\times n}(\mathbb{F})$. ■

A partir de este momento diremos que la matriz escalonada reducida R tal que $A \underset{f}{\sim} R$ es la matriz escalonada reducida de A. Tiene sentido ahora la siguiente definición.

Definición 1.3.11 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. El **rango** de A, denotado por $rgo(A)$, se define como el número de filas distintas de cero en la matriz escalonada reducida de A.*

Ejemplo 1.3.12 *Por el ejemplo 1.3.3,*

$$A = \begin{pmatrix} 2 & 4 & 6 & 2 & 6 \\ 6 & 4 & 2 & 2 & 14 \\ 0 & 4 & 8 & 2 & 2 \\ 2 & 2 & 2 & 2 & 8 \end{pmatrix} \underset{f}{\sim} \begin{pmatrix} 1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por lo tanto, $rgo(A) = 3$.

El rango de una matriz está intimamente ligado a la solución de un sistema de ecuaciones lineales.

Corolario 1.3.13 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, $B \in \mathcal{M}_{m\times 1}(\mathbb{F})$ y $\widehat{A} = (A, B)$ la matriz ampliada. Entonces:*

1. *$AX = B$ tiene una solución si, y sólo si, $rgo(A) = rgo\left(\widehat{A}\right)$;*

2. *$AX = B$ tiene una solución única si, y sólo si, $rgo(A) = rgo\left(\widehat{A}\right) = n$.*

Demostración. Las mismas operaciones elementales por fila que llevan la matriz A a la escalonada reducida E, llevan la matriz $\widehat{A} = (A, B)$ a la matriz $\widehat{E} = (E, C)$. Por el teorema 1.3.5, las ecuaciones $AX = B$ y $EX = C$ tienen exactamente las mismas soluciones. Ahora, por el teorema 1.3.6 tenemos:

1. $EX = C$ tiene solución si, y sólo si, $[C]_{j1} = 0$ para todo $j > rgo(A)$ si, y sólo si, $\widehat{E}$ es la matriz escalonada reducida de $\widehat{A}$ si, y sólo si, $rgo(A) = rgo\left(\widehat{A}\right)$.

2. $EX = C$ tiene solución única si, y sólo si, $[C]_{j1} = 0$ para todo $j > rgo(A) = n$ si, y sólo si, $rgo(A) = rgo\left(\widehat{A}\right) = n$. ■

EJERCICIOS

1. Encuentre la matriz escalonada reducida y el rango de las siguientes matrices:

$$A = \begin{pmatrix} -2 & 6 & 0 \\ -1 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}; \quad B = \begin{pmatrix} 3 & -1 & 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & -3 & 0 & 0 & 0 & 1 \end{pmatrix};$$

$$C=\begin{pmatrix} -i & 1+i & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{pmatrix}$$

2. Resuelva cada uno de los siguientes sistemas de ecuaciones lineales:

$$\text{a)}\ \left.\begin{array}{c} 6x-4y+4z=28 \\ 3x-6y+3z=21 \\ 4x-2y+8z=34 \end{array}\right\} ;\quad \text{b)}\ \left.\begin{array}{c} 4x-10y+6w-8z+4u=8 \\ 5x-10y-5w-4z+7u=22 \\ 3x-7y+2w-5z+4u=9 \end{array}\right\} ;$$

$$\text{c)}\ \left.\begin{array}{c} -x-2y+3w-4z=-2 \\ -6x-15y+6w-3z=-3 \\ -5x-12y+7w-6z=-7 \end{array}\right\} ;\quad \text{d)}\ \left.\begin{array}{c} ix-(1+i)\,y=0 \\ x-2y+z=0 \\ x+2iy-z=0 \end{array}\right\}$$

3. Determine los valores de k para que el sistema de ecuaciones lineales tenga: $i)$ una solución única; $ii)$ infinitas soluciones; y $iii)$ ninguna solución.

$$\begin{array}{rcl} -x+y+z & = & 1 \\ kx+3y+2z & = & 3 \\ -3x-ky-z & = & -2 \end{array}$$

4. Sea

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

¿Para qué matrices $B\in\mathcal{M}_{3\times 1}(\mathbb{R})$ tiene solución el sistema $AX=B$?

5. Demuestre que si P,Q son soluciones del sistema homogéno $AX=0$, entonces $\alpha P+\beta Q$ es una solución de $AX=0$, para todo $\alpha,\beta\in\mathbb{F}$.

6. Sea $A\in\mathcal{M}_{m\times n}(\mathbb{F})$ y $B\in\mathcal{M}_{m\times 1}(\mathbb{F})$. Suponga que P es una solución del sistema de ecuaciones lineales $AX=B$. Demuestre que cualquier otra solución de $AX=B$ es de la forma $P+Q$, donde Q es una solución del sistema homogéneo $AX=0$.

7. Determine exactamente cuándo una matriz triangular superior $n\times n$ tiene rango n.

1.4. Matrices elementales

Las operaciones elementales por filas tienen una interpretación a través de la multiplicación de matrices. La matriz resultante al aplicar una operación elemental por filas a la matriz A se obtiene multiplicando la matriz A por una matriz, que llamaremos matriz elemental.

Definición 1.4.1 *Una **matriz elemental** de $\mathcal{M}_m(\mathbb{F})$ es una matriz de la forma*

$$\begin{aligned} E_{pq} &= f_{pq}(I_m) \\ E_p(\lambda) &= f_p(\lambda)(I_m) \\ E_{pq}(\lambda) &= f_{pq}(\lambda)(I_m) \end{aligned}$$

donde $p, q \in \{1, \ldots, m\}$ y $\lambda \in \mathbb{F}$.

Ejemplo 1.4.2 *Las siguientes matrices son matrices elementales en $\mathcal{M}_3(\mathbb{C})$:*

$$E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; E_2(i) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{pmatrix} \; y \; E_{21}(4) = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La relación que existe entre las operaciones elementales por filas y las matrices elementales viene dada en el siguiente resultado.

Teorema 1.4.3 *Sea $A \in \mathcal{M}_{m \times n}(\mathbb{F})$. Entonces*

1. $E_{pq}A = f_{pq}(A)$;
2. $E_p(\lambda)A = f_p(\lambda)(A)$;
3. $E_{pq}(\lambda)A = f_{pq}(\lambda)(A)$.

Demostración. Vamos a demostrar 3, las demás las dejamos como ejercicio. Sea

$$e_i = \Big(0, \ldots, 0, \underset{i}{1}, 0, \ldots, 0\Big), i = 1, \ldots, m.$$

Si $i \neq p$, entonces

$$(E_{pq}(\lambda)A)_{i*} = (E_{pq}(\lambda))_{i*}A = e_iA = A_{i*} = (f_{pq}(\lambda)(A))_{i*}$$

Por otra parte,

$$\begin{aligned} (E_{pq}(\lambda)A)_{p*} &= (E_{pq}(\lambda))_{p*}A = (e_p + \lambda e_q)A = e_pA + \lambda e_qA \\ &= A_{p*} + \lambda A_{q*} = (f_{pq}(\lambda)(A))_{p*} \end{aligned}$$

■

En otras palabras, el teorema 1.4.3 establece que efectuar una operación elemental por filas a una matriz equivale a multiplicar por la izquierda la matriz por una matriz elemental. Luego la relación $A \underset{f}{\sim} B$ equivale a decir que existen matrices elementales $E_1, \ldots, E_k$ tales que $B = E_k \cdots E_1A$.

Teorema 1.4.4 *Las matrices elementales son invertibles. Además, sus inversas son matrices elementales.*

Demostración. Demostraremos que $E_p(\lambda) E_p\left(\frac{1}{\lambda}\right) = I_m$. Si $i \neq p$, entonces

$$\left(E_p(\lambda) E_p\left(\frac{1}{\lambda}\right)\right)_{i*} = (E_p(\lambda))_{i*} E_p\left(\frac{1}{\lambda}\right) = e_i E_p\left(\frac{1}{\lambda}\right) = \left(E_p\left(\frac{1}{\lambda}\right)\right)_{i*} = e_i$$

Además,

$$\begin{aligned}\left(E_p(\lambda) E_p\left(\frac{1}{\lambda}\right)\right)_{p*} &= (E_p(\lambda))_{p*} E_p\left(\frac{1}{\lambda}\right) = (\lambda e_p) E_p\left(\frac{1}{\lambda}\right) = \lambda\left(E_p\left(\frac{1}{\lambda}\right)\right)_{p*} \\ &= \lambda\left(\frac{1}{\lambda} e_p\right) = e_p\end{aligned}$$

Esto demuestra que $E_p(\lambda) E_p\left(\frac{1}{\lambda}\right) = I_m$. Similarmente se demuestra que

$$E_{pq} E_{pq} = E_{pq}(\lambda) E_{pq}(-\lambda) = I_m$$

. ■

Como consecuencia de este resultado tenemos la siguiente caracterización de las matrices invertibles.

Teorema 1.4.5 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Las siguientes condiciones son equivalentes:*

1. *A es invertible;*
2. *La ecuación $AX = B$ tiene una única solución para todo $B \in \mathcal{M}_{n\times 1}(\mathbb{F})$;*
3. *$rgo(A) = n$;*
4. *$A \underset{f}{\sim} I_n$;*
5. *A es producto de matrices elementales.*

Demostración. 1. $\Rightarrow$ 2. Supongamos que A es una matriz invertible. Entonces,

$$A\left(A^{-1}B\right) = B$$

y así $X = A^{-1}B$ es una solución de la ecuación $AX = B$. Además, si $AC = B$, entonces, multiplicando por la inversa de A ambos lados obtenemos $C = A^{-1}B$. Es decir, la solución es única.

2. $\Rightarrow$ 3. Es consecuencia del corolario 14.

3. $\Rightarrow$ 4. Si $rgo(A) = n$, entonces las n filas de la matriz escalonada reducida R de A son distintas de cero, lo que obliga a que $R = I_n$. Así $A \underset{f}{\sim} I_n$.

4. $\Rightarrow$ 5. Si $A \underset{f}{\sim} I_n$, entonces, por el teorema 1.4.3 existen matrices elementales $E_1, \ldots, E_k$ tales que $A = E_k \cdots E_1 I_n = E_k \cdots E_1$.

5. $\Rightarrow$ 1. Es consecuencia de los teoremas 1.4.4 y 1.2.12. ■

Tenemos ahora un método para calcular la inversa de una matriz. En efecto, si A es invertible, entonces el teorema 1.3.4 nos indica cómo construir matrices elementales $E_1, \ldots, E_k$ tales que $E_k \cdots E_1 A = I_n$. En particular, $E_k \cdots E_1 = A^{-1}$. Además,

$$E_k \cdots E_1 (A, I) = (E_k \cdots E_1 A, E_k \cdots E_1 I) = \left(I, A^{-1}\right)$$

Ejemplo 1.4.6 *Calculemos la inversa de la matriz* $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$.

$$\begin{pmatrix} 1 & 3 & -2 & 1 & 0 & 0 \\ 2 & 8 & -3 & 0 & 1 & 0 \\ 1 & 7 & 1 & 0 & 0 & 1 \end{pmatrix} \overset{f_{21}(-2)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 2 & 1 & -2 & 1 & 0 \\ 1 & 7 & 1 & 0 & 0 & 1 \end{pmatrix} \overset{f_{31}(-1)}{\rightsquigarrow}$$

$$\begin{pmatrix} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 2 & 1 & -2 & 1 & 0 \\ 0 & 4 & 3 & -1 & 0 & 1 \end{pmatrix} \overset{f_{32}(-2)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 2 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{pmatrix} \overset{f_{2}(1/2)}{\rightsquigarrow}$$

$$\begin{pmatrix} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 1/2 & -1 & 1/2 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{pmatrix} \overset{f_{23}(-1/2)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -5/2 & 3/2 & -1/2 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{pmatrix} \overset{f_{13}(2)}{\rightsquigarrow}$$

$$\begin{pmatrix} 1 & 3 & 0 & 7 & -4 & 2 \\ 0 & 1 & 0 & -5/2 & 3/2 & -1/2 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{pmatrix} \overset{f_{12}(-3)}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 0 & 29/2 & -17/2 & 7/2 \\ 0 & 1 & 0 & -5/2 & 3/2 & -1/2 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{pmatrix}$$

Por lo tanto, $A^{-1} = \begin{pmatrix} 29/2 & -17/2 & 7/2 \\ -5/2 & 3/2 & -1/2 \\ 3 & -2 & 1 \end{pmatrix}$.

Corolario 1.4.7 *Sean* $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$. *Entonces,* $A \underset{f}{\sim} B$ *si, y sólo si, existe una matriz invertible* $P \in \mathcal{M}_m(\mathbb{F})$ *tal que* $B = PA$.

Demostración. Si $A \underset{f}{\sim} B$, entonces existen matrices elementales $E_1, \ldots, E_k$ tales que

$$B = E_k \cdots E_1 A.$$

Luego $B = PA$, donde $P = E_k \cdots E_1 \in \mathcal{M}_m(\mathbb{F})$ es invertible. Recíprocamente, supongamos que $B = QA$, donde Q es una matriz invertible. Entonces, por el teorema 1.4.5, $Q = E_r \cdots E_1$, para ciertas matrices elementales $E_1, \ldots, E_r$. En consecuencia,

$$B = E_r \cdots E_1 A$$

y así, $A \underset{f}{\sim} B$. ■

EJERCICIOS

1. Determine si las matrices dadas son invertibles y, en caso de serlo, calcule la inversa.

$$A = \begin{pmatrix} 1 & 2 & -4 \\ -1 & -1 & 5 \\ 2 & 7 & -3 \end{pmatrix}; \quad B = \begin{pmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{pmatrix} \text{ y}$$

$$C = \begin{pmatrix} -i & 1+i & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{pmatrix}$$

2. Dada la matriz

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

encuentre matrices elementales $E_1, \ldots, E_k$ tales que $E_k \cdots E_1 A = I$.

3. Demuestre que la relación “equivalencia por filas” es una relación de equivalencia sobre el espacio de matrices.

4. Demuestre que $(E_{pq})^{\top} = E_{pq}$, $(E_p(\lambda))^{\top} = E_p(\lambda)$ y $(E_{pq}(\lambda))^{\top} = E_{qp}(\lambda)$.

5. Demuestre que una matriz triangular superior es invertible si, y sólo si, todos los elementos de la diagonal son diferentes de cero.

6. Demuestre que si T es una matriz triangular superior invertible, entonces T^{-1} también es triangular superior.

7. Considere la matriz $I_{pq} \in \mathcal{M}_m(\mathbb{F})$ que tiene todos los elementos 0 excepto el elemento pq que es 1. Es decir, $[I_{pq}]_{pq} = 1$ y $[I_{pq}]_{ij} = 0$ si $i \neq p$ o $j \neq q$. Demuestre que

$$I_{pq} I_{rs} = \begin{cases} 0 & \text{si} \quad q \neq r \\ I_{ps} & \text{si} \quad q = s \end{cases}$$

8. Demuestre que una matriz $B \in \mathcal{M}_n(\mathbb{F})$ es escalonada reducida e invertible si, y sólo si, B es la identidad.

CAPÍTULO 2

DETERMINANTES

En este capítulo introducimos el determinante de una matriz y estudiamos algunas conexiones con el cálculo de la matriz inversa y la solución de sistemas de ecuaciones lineales.

2.1. Función determinante

Definición 2.1.1 *Sea $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ una función. Decimos que D es una* ***función determinante*** *si satisface las siguientes condiciones:*

1. *Si B se obtiene a partir de A multiplicando la fila i de A por $\alpha \in \mathbb{F}$, entonces*
$$D(B) = \alpha D(A);$$
2. *Si las matrices A, B, C son idénticas excepto en la fila i, y la fila i de C es la suma de la fila i de A y la fila i de B, entonces $D(C) = D(A) + D(B)$;*
3. *Si A tiene dos filas idénticas entonces $D(A) = 0$ y*
4. *$D(I_n) = 1$.*

Una función $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ que verifica las primeras dos condiciones en la definición 2.1.1 se dice que es n-**lineal**; el nombre lo justificaremos en el capítulo 4. Si además D satisface la tercera condición, entonces se dice que D es **alternada**.

Demostraremos en esta sección que de existir una función determinante $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$, esta es única. Además, desarrollaremos las propiedades más importantes que poseen estas funciones. Dejaremos para la sección 2.3 la construcción explícita de la (única) función determinante $\mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$.

Comenzamos estudiando el comportamiento de una función determinante cuando una operación elemental por fila es aplicada a una matriz.

Teorema 2.1.2 *Sea $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ una función alternada y sea $A \in \mathcal{M}_n(\mathbb{F})$. Entonces:*

1. *Si B se obtiene a partir de A multiplicando una fila de A por $\alpha \in \mathbb{F}$, entonces*
$$D(B) = \alpha D(A);$$
2. *Si B se obtiene a partir de A intercambiando dos filas, entonces*
$$D(B) = -D(A);$$
3. *Si B se obtiene a partir de A sumando un múltiplo de una fila a otra, entonces*
$$D(B) = D(A).$$

Demostración. 1. Es la condición 1. de la definición 2.1.1.

2. Si $A = \begin{pmatrix} \vdots \\ A_{i*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix}$ entonces $B = \begin{pmatrix} \vdots \\ A_{j*} \\ \vdots \\ A_{i*} \\ \vdots \end{pmatrix}$. Sea $C = \begin{pmatrix} \vdots \\ A_{i*} + A_{j*} \\ \vdots \\ A_{i*} + A_{j*} \\ \vdots \end{pmatrix}$. Entonces al ser D una función alternada se obtiene

$$\begin{aligned} 0 &= D(C) = D\begin{pmatrix} \vdots \\ A_{i*} \\ \vdots \\ A_{i*} + A_{j*} \\ \vdots \end{pmatrix} + D\begin{pmatrix} \vdots \\ A_{j*} \\ \vdots \\ A_{i*} + A_{j*} \\ \vdots \end{pmatrix} \\ &= D\begin{pmatrix} \vdots \\ A_{i*} \\ \vdots \\ A_{i*} \\ \vdots \end{pmatrix} + D\begin{pmatrix} \vdots \\ A_{i*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix} + D\begin{pmatrix} \vdots \\ A_{j*} \\ \vdots \\ A_{i*} \\ \vdots \end{pmatrix} + D\begin{pmatrix} \vdots \\ A_{j*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix} \\ &= D(A) + D(B). \end{aligned}$$

3. Si $A = \begin{pmatrix} \vdots \\ A_{i*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix}$ entonces $B = \begin{pmatrix} \vdots \\ A_{i*} + \alpha A_{j*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix}$ y así

$$D(B) = D\begin{pmatrix} \vdots \\ A_{i*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix} + D\begin{pmatrix} \vdots \\ \alpha A_{j*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix} = D\begin{pmatrix} \vdots \\ A_{i*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix} + \alpha D\begin{pmatrix} \vdots \\ A_{j*} \\ \vdots \\ A_{j*} \\ \vdots \end{pmatrix} = D(A)$$

■

Por el teorema 1.3.4 y el teorema 2.1.2, si $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ es alternada, entonces, para todo $A \in \mathcal{M}_n(\mathbb{F})$, $D(A) = \alpha D(R)$ para $0 \neq \alpha \in \mathbb{F}$, donde R es la matriz escalonada reducida de A. Calcular $D(R)$ es fácil, como veremos en nuestro próximo resultado.

Teorema 2.1.3 *Sea $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ una función determinante y $A \in \mathcal{M}_n(\mathbb{F})$ una matriz triangular. Entonces*

$$D(A) = [A]_{11} \cdot [A]_{22} \cdot \cdots \cdot [A]_{nn}$$

Demostración. Supongamos primero que $[A]_{11}, [A]_{22}, \ldots, [A]_{nn}$ son todos diferentes de cero. Entonces, multiplicando cada fila por $[A]_{ii}^{-1}$ y aplicando la parte 1. del teorema 2.1.2 obtenemos que

$$D(A) = [A]_{11} \cdot [A]_{22} \cdot \cdots \cdot [A]_{nn} D(B)$$

donde B es una matriz triangular con todos los coeficientes en la diagonal igual a 1. Ahora, aplicando la operación elemental en la parte 3. del teorema 2.1.2 repetidas veces nos lleva a que B es equivalente por filas a I_n y $D(B) = D(I_n) = 1$. En consecuencia,

$$D(A) = [A]_{11} \cdot [A]_{22} \cdot \cdots \cdot [A]_{nn}.$$

Si $[A]_{ii} = 0$ para algún i entonces la matriz escalonada reducida R de A tiene una fila 0, lo cual implica que $D(A) = 0 = [A]_{11} \cdot [A]_{22} \cdot \cdots \cdot [A]_{nn}$. ■

Teorema 2.1.4 *Si existe una función determinante $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$, esta es única.*

Demostración. Supongamos que $D, D' : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ son funciones determinantes y sea $A \in \mathcal{M}_n(\mathbb{F})$. Si R es la (única) matriz escalonada reducida de A entonces, por el teorema 2.1.2 tenemos que $D(A) = \alpha D(R)$ y $D'(A) = \alpha D'(R)$ para algún $\alpha \in \mathbb{F}$. Como R es una matriz triangular, se deduce del teorema 2.1.3 que

$$D(R) = D'(R) = [R]_{11} \cdot [R]_{22} \cdot \cdots \cdot [R]_{nn}.$$

y así $D(A) = D'(A)$. ■

A partir de ahora denotaremos por $det : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ a la única función determinante (que existe como veremos en la sección 2.3) y diremos que $det(A)$ es el determinante de la matriz $A \in \mathcal{M}_n(\mathbb{F})$. Los teoremas 2.1.2 y 2.1.3 nos proporcionan un método para calcular el valor $det(A)$ para cualquier matriz $A \in \mathcal{M}_n(\mathbb{F})$.

Ejemplo 2.1.5 *Consideremos la matriz* $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$. *Entonces*

$$\begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix} \overset{f_{21}(-2)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 \\ 0 & 2 & 1 \\ 1 & 7 & 1 \end{pmatrix} \overset{f_{31}(-1)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 \\ 0 & 2 & 1 \\ 0 & 4 & 3 \end{pmatrix} \overset{f_{32}(-2)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Luego, usando los teoremas 2.1.2 y 2.1.3 obtenemos

$$det \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix} = 1 \cdot 2 \cdot 1 = 2$$

EJERCICIOS

1. Calcule el determinante de la matrices A y B dadas a continuación:

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

2. Demuestre que en general, $det(A + B) \neq det(A) + det(B)$.

3. Si una matriz A tiene una fila 0, entonces $det(A) = 0$.

4. Demuestre que $det(E_{pq}) = -1$, $det(E_p(\lambda)) = \lambda$ y $det(E_{pq}(\lambda)) = 1$.

5. Sea $\alpha \in \mathbb{F}$. Demuestre que para cualquier matriz $A \in \mathcal{M}_n(\mathbb{F})$ se tiene que

$$det(\alpha A) = \alpha^n det(A).$$

6. Demuestre que si $A, B \in \mathcal{M}_n(\mathbb{F})$ son matrices triangulares, entonces

$$det(AB) = det(A)\, det(B).$$

2.2. Propiedades del determinante

Vamos ahora a deducir algunas propiedades importantes de la función determinante

$$det : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}.$$

Teorema 2.2.1 *Una matriz $A \in \mathcal{M}_n(\mathbb{F})$ es invertible si, y sólo si, $det(A) \neq 0$.*

Demostración. Si A es invertible, entonces, por el teorema 1.4.5, $A \underset{f}{\sim} I_n$. Luego por el teorema 2.1.2, $det(A) = \alpha det(I_n) = \alpha \neq 0$. Recíprocamente, si A no es invertible, entonces $rgo(A) < n$ por el teorema 1.4.5. Es decir, $A \underset{f}{\sim} R$, donde la matriz escalonada reducida R de A tiene sólo ceros en la última fila. En particular, $det(R) = 0$. Así, $det(A) = \beta det(R) = 0$. ■

Teorema 2.2.2 *Si $A, B \in \mathcal{M}_n(\mathbb{F})$, entonces $det(AB) = det(A)\, det(B)$.*

Demostración. Si A no es invertible, entonces, por el teorema 2.2.1, $det(A) = 0$. Por otro lado, usando el corolario 1.4.7 tenemos que $R = UA$, donde R es la matriz escalonada reducida de A y U una matriz invertible. Además, por el teorema 1.4.5, la última fila R_{n*} de R es 0, esto implica que $(RB)_{n*} = R_{n*}B = 0$. Como $RB = U(AB)$, se deduce del corolario 1.4.7 que $AB \underset{f}{\sim} RB$. Se obtiene entonces por el teorema 2.1.2 que $det(AB) = \alpha det(RB) = \alpha 0 = 0$.

Supongamos ahora que A es invertible. Entonces, por el teorema 1.4.5, $A = E_1 \cdots E_k$, donde E_j es una matriz elemental para cada $j = 1, \ldots, k$. Luego el problema se reduce a demostrar que $det(EB) = det(E)\, det(B)$ para cualquier matriz elemental E. Por el teorema 1.4.3, teorema 2.1.2 y teniendo en cuenta que $det(E_{pq}) = -1$, $det(E_p(\lambda)) = \lambda$ y $det(E_{pq}(\lambda)) = 1$ tenemos que

$$\begin{aligned}
det(E_{pq}B) &= det(f_{pq}(B)) = -1det(B) = det(E_{pq})\, det(B) \\
det(E_p(\lambda)B) &= det(f_p(\lambda)(B)) = \lambda det(B) = det(E_p(\lambda))\, det(B) \\
det(E_{pq}(\lambda)B) &= det(f_{pq}(\lambda)(B)) = det(B) = det(E_{pq}(\lambda))\, det(B)
\end{aligned}$$

■

Corolario 2.2.3 *Si $A \in \mathcal{M}_n(\mathbb{F})$ es invertible, entonces $det(A^{-1}) = \frac{1}{det(A)}$.*

Demostración. Observamos primero que por el teorema 2.2.1, $det(A) \neq 0$. Luego, por el teorema 2.2.2,

$$1 = det(I_n) = det\left(AA^{-1}\right) = det(A)\, det\left(A^{-1}\right)$$

■

Teorema 2.2.4 *Si $A \in \mathcal{M}_n(\mathbb{F})$, entonces $det(A) = det\left(A^{\top}\right)$.*

Demostración. Si A no es invertible, entonces, por el teorema 1.2.15, $A^{\top}$ tampoco es invertible. Luego por el teorema 2.2.1, $0 = det(A) = det\left(A^{\top}\right)$. Supongamos ahora que A es invertible. Entonces, por el teorema 1.4.5 existen matrices elementales $E_1, \ldots, E_k$ tales que $A = E_k \cdots E_1$. Luego por el teorema 1.2.15 parte 4., $A^{\top} = E_1^{\top} \cdots E_k^{\top}$. Así, por el teorema 2.2.2, $det(A) = det(E_k) \cdots det(E_1)$ y $det\left(A^{\top}\right) = det\left(E_k^{\top}\right) \cdots det\left(E_1^{\top}\right)$. Bastaría con demostrar que $det(E) = det\left(E^{\top}\right)$ para toda matriz elemental. En el caso de las matrices elementales del tipo E_{pq} y $E_p(\lambda)$ se cumple porque son matrices simétricas. Finalmente, como $(E_{pq}(\lambda))^{\top} = E_{qp}(\lambda)$, entonces

$$det\left(E_{pq}(\lambda)^{\top}\right) = det(E_{qp}(\lambda)) = 1 = det(E_{pq}(\lambda))$$

■

Como consecuencia de este resultado obtenemos que el teorema 2.1.2 sigue siendo válido cuando sustituimos "fila" por "columna".

Teorema 2.2.5 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Entonces:*

1. *Si B se obtiene a partir de A multiplicando una columna de A por $\alpha \in \mathbb{F}$, entonces $det(B) = \alpha det(A)$;*

2. *Si B se obtiene a partir de A intercambiando dos columnas, entonces $det(B) = -det(A)$;*

3. *Si B se obtiene a partir de A sumando un múltiplo de una columna a otra, entonces $det(B) = det(A)$.*

Demostración. En cada caso aplicamos el teorema 2.1.2 a la matriz traspuesta y luego usamos teorema 2.2.4. ■

EJERCICIOS

1. La matriz $N \in \mathcal{M}_n(\mathbb{F})$ se llama nilpotente si $N^k = 0$ para algún entero $k \geq 1$. Demuestre que si N es nilpotente, entonces $det(N) = 0$.

2. Una matriz $A \in \mathcal{M}_n(\mathbb{F})$ se llama idempotente si $A^2 = A$. ¿Cuáles son los valores posibles para $det(A)$?

3. Demuestre que para cualesquiera dos matrices A y B se tiene que $det(AB) = det(BA)$.

4. Sean A y B matrices $n \times n$ tales que $AB = -BA$. Si n es impar demuestre que A o B no es invertible.

5. Una matriz ortogonal es una matriz $n \times n$ que satisface $AA^{\top} = I_n$. Demuestre que si A es ortogonal entonces $det(A) = \pm 1$.

2.3. Existencia de la función determinante

Vamos a demostrar en esta sección que existe una función determinante $\mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$. Para eso necesitamos primero recordar algunas nociones básicas de la teoría de permutaciones.

Definición 2.3.1 *Una **permutación** de $\{1, \ldots, n\}$ es una biyección*

$$\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

Usualmente la denotamos por $\sigma_1 \cdots \sigma_n$, donde σ_i es la imagen de i bajo σ. S_n es el conjunto de las permutaciones de $\{1, \ldots, n\}$.

Ejemplo 2.3.2 *Las permutaciones de S_2 son 12 y 21. En S_3 hay 6 permutaciones:*

$$123, 132, 231, 213, 312 \text{ y } 321$$

Si $\sigma \in S_n$ entonces la función inversa la denotamos por σ^{-1}, de manera que $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \varepsilon$, donde $\varepsilon = 1 \cdots n$ es la permutación identidad. Claramente $\sigma^{-1} \in S_n$ y $\sigma = \left(\sigma^{-1}\right)^{-1}$. Por otra parte, la composición de permutaciones es también una permutación. Es decir, si $\sigma, \tau \in S_n$, entonces $\sigma\tau \in S_n$.

Ejemplo 2.3.3 *Consideremos las permutaciones $\sigma = 13524$ y $\tau = 12543$ de S_5. Entonces*

$$\sigma^{-1} = 14253 \text{ y } \sigma\tau = 13425$$

Definición 2.3.4 *Una **inversión** de una permutación $\sigma \in S_n$ es un par (j, k) tal que*

$$1 \leq j < k \leq n$$

y $\sigma_j > \sigma_k$. El signo de $\sigma \in S_n$ se define como $sgn(\sigma) = (-1)^m$, donde m es el número de inversiones de σ.

En otras palabras, el signo de una permutación $\sigma \in S_n$ es 1 cuando tiene un número par de inversiones y -1 cuando tiene un número impar de inversiones.

Ejemplo 2.3.5 *Consideremos la permutación $142365 \in S_6$. Entonces las inversiones de esta permutación son*

$$(2, 3), (2, 4), (5, 6)$$

y, en consecuencia, $sgn(142365) = -1$.

Ejemplo 2.3.6 *Fijemos $1 \leq i < j \leq n$ y definamos la permutación $\tau \in S_n$ como $\tau_i = j$, $\tau_j = i$ y $\tau_k = k$ para todo $k \in \{1, \ldots, n\}$ tal que $k \neq i$ y $k \neq j$. Esto es,*

$$\tau = 1 \cdots (i-1)(j)(i+1) \cdots (j-1)(i)(j+1) \cdots n$$

Entonces las inversiones de τ son:

$$(i, i+1), (i, i+2), \ldots, (i, j)$$

y

$$(i+1, j), (i+2, j), \ldots, (j-1, j)$$

En la primera lista aparecen $j-i$ inversiones y en la segunda lista $j-i-1$. Por lo tanto, $sgn(\tau) = (-1)^{2(j-i)-1} = -1$.

La permutación definida en el ejemplo 2.3.6 recibe el nombre de **trasposición**. Observamos que si $\sigma \in S_n$ y τ es la trasposición que mueve a i, j, entonces $\sigma\tau$ es la permutación que se obtiene a partir de σ intercambiando las posiciones i y j.

Teorema 2.3.7 *Si $\tau, \sigma \in S_n$, entonces $sgn(\tau\sigma) = sgn(\tau)\, sgn(\sigma)$.*

Demostración. Consideremos el polinomio $p = p(x_1, \ldots, x_n) = \prod_{i<j} (x_i - x_j)$ y sea $\sigma \in S_n$. Definamos $\sigma(p) = \prod_{i<j} \left(x_{\sigma_i} - x_{\sigma_j}\right)$. Como σ es biyectiva, entonces vamos a tener que $\sigma(p) = p$ o $\sigma(p) = -p$. De hecho, $\sigma(p) = p$ si, y sólo si, hay un número par de términos en $\sigma(p)$ de la forma $(x_r - x_s)$ tales que $r > s$. Es decir, existen un número par de (i, j) tales que $i < j$, $\sigma_i = r > s = \sigma_j$. Este es precisamente el número total de inversiones de σ. Así hemos demostrado que para cada $\sigma \in S_n$

$$\sigma(p) = sgn(\sigma)\, p$$

Ahora sean $\tau, \sigma \in S_n$. Entonces

$$\begin{aligned} sgn(\tau\sigma)\, p &= \tau\sigma(p) = \tau[\sigma(p)] = \tau[sgn(\sigma)\, p] \\ &= sgn(\tau)\, sgn(\sigma)\, p \end{aligned}$$

Esto implica que $sgn(\tau\sigma) = sgn(\tau)\, sgn(\sigma)$. ■

Ahora estamos en condiciones de definir explícitamente la función determinante

$$det : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}.$$

Teorema 2.3.8 *La función $D : \mathcal{M}_n(\mathbb{F}) \longrightarrow \mathbb{F}$ definida por*

$$D(A) = \sum_{\sigma \in S_n} sgn(\sigma)\, [A]_{1\sigma_1} [A]_{2\sigma_2} \cdots [A]_{n\sigma_n}$$

con $A \in \mathcal{M}_n(\mathbb{F})$, es una función determinante.

Demostración. Vamos a demostrar que D satisface las cuatro condiciones de la definición 2.1.1.

1. Supongamos que $B \in \mathcal{M}_n(\mathbb{F})$ se obtiene a partir de $A \in \mathcal{M}_n(\mathbb{F})$ multiplicando la fila i de A por $\alpha \in \mathbb{F}$. Entonces, para todo $j = 1, \ldots, n$ tenemos que $[B]_{kj} = [A]_{kj}$ si $k \neq i$ y $[B]_{ij} = \alpha [A]_{ij}$. Luego,

$$\begin{aligned} D(B) &= \sum_{\sigma \in S_n} sgn(\sigma) [B]_{1\sigma(1)} [B]_{2\sigma(2)} \cdots [B]_{i\sigma(i)} \cdots [B]_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma) [A]_{1\sigma(1)} [A]_{2\sigma(2)} \cdots \alpha [A]_{i\sigma(i)} \cdots [A]_{n\sigma(n)} \\ &= \alpha \sum_{\sigma \in S_n} sgn(\sigma) [A]_{1\sigma(1)} [A]_{2\sigma(2)} \cdots [A]_{i\sigma(i)} \cdots [A]_{n\sigma(n)} \\ &= \alpha D(A) \end{aligned}$$

2. Supongamos que $A, B, C \in \mathcal{M}_n(\mathbb{F})$ son idénticas excepto en la fila i, y la fila i de C es la suma de la fila i de A y la fila i de B. Es decir, para todo $j = 1, \ldots, n$ tenemos que $[C]_{kj} = [A]_{kj} = [B]_{kj}$ si $k \neq i$ y $[C]_{ij} = [A]_{ij} + [B]_{ij}$. Entonces

$$\begin{aligned} D(C) &= \sum_{\sigma \in S_n} sgn(\sigma) [C]_{1\sigma(1)} [C]_{2\sigma(2)} \cdots [C]_{i\sigma(i)} \cdots [C]_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma) [C]_{1\sigma(1)} [C]_{2\sigma(2)} \cdots \left([A]_{i\sigma(i)} + [B]_{i\sigma(i)}\right) \cdots [C]_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma) [A]_{1\sigma(1)} [A]_{2\sigma(2)} \cdots [A]_{i\sigma(i)} \cdots [A]_{n\sigma(n)} \\ &\quad + \sum_{\sigma \in S_n} sgn(\sigma) [B]_{1\sigma(1)} [B]_{2\sigma(2)} \cdots [B]_{i\sigma(i)} \cdots [B]_{n\sigma(n)} \\ &= D(A) + D(B) \end{aligned}$$

3. Supongamos que A tiene dos filas iguales, digamos las filas que están en la posición i y la posición k, donde $1 \leq i < k \leq n$. Es decir, para todo $j = 1, \ldots, n$ tenemos que $[A]_{ij} = [A]_{kj}$. Sea τ la trasposición tal que $\tau_i = k$ y $\tau_k = i$. Entonces, para todo $\sigma \in S_n$ se tiene que

$$\begin{aligned} [A]_{p\sigma_p} &= [A]_{p(\sigma\tau)_p} \text{ si } p \neq i \text{ y } p \neq k \\ [A]_{i\sigma_i} &= [A]_{k(\sigma\tau)_k} \\ [A]_{k\sigma_k} &= [A]_{i(\sigma\tau)_i} \end{aligned}$$

y por el ejemplo 2.3.6 y el teorema 2.3.7,

$$sgn(\tau\sigma) = sgn(\tau)\, sgn(\sigma) = -sgn(\sigma)$$

Consideremos la partición de S_n dada por los dos conjuntos

$$\Gamma = \{\sigma \in S_n : \sigma_i < \sigma_k\}$$

y

$$\Omega = \{\sigma \in S_n : \sigma_i > \sigma_k\}.$$

Observamos que existe una biyección $\Gamma \longrightarrow \Omega$ dada por $\sigma \rightsquigarrow \sigma\tau$, para todo $\sigma \in S_n$. Entonces

$$\begin{aligned}
D(A) &= \sum_{\sigma \in S_n} sgn(\sigma)[A]_{1\sigma_1} \cdots [A]_{i\sigma_i} \cdots [A]_{k\sigma_k} \cdots [A]_{n\sigma_n} \\
&= \sum_{\sigma \in \Gamma} sgn(\sigma)[A]_{1\sigma_1} \cdots [A]_{i\sigma_i} \cdots [A]_{k\sigma_k} \cdots [A]_{n\sigma_n} \\
&\quad + \sum_{\sigma \in \Omega} sgn(\sigma)[A]_{1\sigma_1} \cdots [A]_{i\sigma_i} \cdots [A]_{k\sigma_k} \cdots [A]_{n\sigma_n} \\
&= \sum_{\sigma \in \Gamma} sgn(\sigma)[A]_{1\sigma_1} \cdots [A]_{i\sigma_i} \cdots [A]_{k\sigma_k} \cdots [A]_{n\sigma_n} \\
&\quad + \sum_{\sigma \in \Gamma} sgn(\sigma\tau)[A]_{1(\sigma\tau)_1} \cdots [A]_{i(\sigma\tau)_i} \cdots [A]_{k(\sigma\tau)_k} \cdots [A]_{n(\sigma\tau)_n} \\
&= \sum_{\sigma \in \Gamma} sgn(\sigma)[A]_{1\sigma_1} \cdots [A]_{i\sigma_i} \cdots [A]_{k\sigma_k} \cdots [A]_{n\sigma_n} \\
&\quad - \sum_{\sigma \in \Gamma} sgn(\sigma)[A]_{1\sigma_1} \cdots [A]_{i\sigma_i} \cdots [A]_{k\sigma_k} \cdots [A]_{n\sigma_n} \\
&= 0
\end{aligned}$$

4.
$$D(I_n) = \sum_{\sigma \in S_n} sgn(\sigma)[I_n]_{1\sigma_1} \cdots [I_n]_{n\sigma_n} = sgn(12\cdots n)[I_n]_{11} \cdots [I_n]_{nn} = 1$$

■

Como consecuencia de la propiedad de unicidad que tiene la función determinante (teorema 2.1.4) escribiremos de ahora en adelante

$$det(A) = \sum_{\sigma \in S_n} sgn(\sigma)[A]_{1\sigma_1}[A]_{2\sigma_2} \cdots [A]_{n\sigma_n} \tag{2.1}$$

para todo $A \in \mathcal{M}_n(\mathbb{F})$.

Ejemplo 2.3.9 *Sea $A \in \mathcal{M}_2(\mathbb{F})$ con elementos $[A]_{ij} = a_{ij}$, para todo $1 \leq i, j \leq 2$. Entonces, $S_2 = \{12, 21\}$, donde $sgn(12) = 1$ y $sgn(21) = -1$ y, en consecuencia,*

$$det(A) = a_{11}a_{22} - a_{12}a_{21}$$

Ejemplo 2.3.10 *Sea $A \in \mathcal{M}_3(\mathbb{F})$ con elementos $[A]_{ij} = a_{ij}$, para todo $1 \leq i, j \leq 3$. Entonces*

$$S_3 = \{123, 231, 312, 321, 213, 132\}$$

donde

$$sgn(123) = sgn(231) = sgn(312) = 1$$

y

$$sgn\,(321) = sgn\,(213) = sgn\,(132) = -1$$

Por lo tanto,

$$\begin{aligned} det\,(A) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33 -} a_{11}a_{23}a_{32} \end{aligned}$$

EJERCICIOS

1. Calcule el signo de las siguientes permutaciones: 32154; 21453; 41235.
2. Usando la fórmula (2.1), calcule el determinante para las siguientes matrices:

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

2.4. La expansión de Laplace

Vamos a presentar en esta sección un nuevo método para calcular el determinante de una matriz.

Definición 2.4.1 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Consideremos la matriz $A_{ij} \in \mathcal{M}_{n-1}(\mathbb{F})$ obtenida a partir de A al eliminar la fila i y la columna j de A. Entonces A_{ij} se llama la* ***submatriz principal*** *ij de A. El escalar $(-1)^{i+j}\,det\,(A_{ij})$ se llama el* ***cofactor*** *de $[A]_{ij}$.*

Ejemplo 2.4.2 *Para la matriz $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$, la submatriz principal 23 de A es*

$$A_{23} = \begin{pmatrix} 1 & 3 \\ 1 & 7 \end{pmatrix}$$

y el cofactor de $[A]_{23}$ es $(-1)^{2+3} \cdot 4 = -4$.

Vamos a presentar una forma recursiva para calcular el determinante de una matriz en términos de los cofactores de la matriz.

Teorema 2.4.3 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Entonces para cada $i = 1, \ldots, n$*

$$det\,(A) = \sum_{j=1}^{n} (-1)^{i+j}\,[A]_{ij}\,det\,(A_{ij})$$

y para cada $j = 1, \ldots, n$

$$det\,(A) = \sum_{i=1}^{n} (-1)^{i+j}\,[A]_{ij}\,det\,(A_{ij})$$

Demostración. Demostraremos la segunda fórmula. Por el teorema 2.1.4 basta ver que la función $\Psi : \mathcal{M}_n(\mathbb{F}) \to \mathbb{F}$ dada por $\Psi(A) = \sum_{i=1}^{n} (-1)^{i+j} [A]_{ij} \, det(A_{ij})$ es una función determinante.

1. Supongamos que $B \in \mathcal{M}_n(\mathbb{F})$ se obtiene a partir de $A \in \mathcal{M}_n(\mathbb{F})$ multiplicando la fila k de A por $\alpha \in \mathbb{F}$. Entonces para todo j es claro que $B_{kj} = A_{kj}$ y así,

$$[B]_{kj} \, det(B_{kj}) = \alpha [A]_{kj} \, det(A_{kj}).$$

Si $i \neq k$, entonces, por la $(n-1)$-linealidad del determinante tenemos que para todo j

$$det(B_{ij}) = \alpha det(A_{ij})$$

y como $[A]_{ij} = [B]_{ij}$ concluimos que

$$[B]_{ij} \, det(B_{ij}) = \alpha [A]_{ij} \, det(A_{ij})$$

Esto demuestra que $\Psi(B) = \alpha \Psi(A)$.

2. Supongamos que $A, B, C \in \mathcal{M}_n(\mathbb{F})$ son idénticas excepto en la fila k, y la fila k de C es la suma de la fila k de A y la fila k de B. Para todo j tenemos que $C_{kj} = A_{kj} = B_{kj}$ y así,

$$[C]_{kj} \, det(C_{kj}) = [A]_{kj} \, det(A_{kj}) + [B]_{kj} \, det(B_{kj})$$

Si $i \neq k$, entonces, por la $(n-1)$-linealidad del determinante tenemos que para todo j

$$det(C_{ij}) = det(A_{ij}) + det(B_{ij})$$

y al ser $[C]_{ij} = [A]_{ij} = [B]_{ij}$ entonces

$$[C]_{ij} \, det(C_{ij}) = [A]_{ij} \, det(A_{ij}) + [B]_{ij} \, det(B_{ij})$$

Esto demuestra que $\Psi(C) = \Psi(A) + \Psi(B)$.

3. Supongamos que A tiene dos filas iguales, digamos las filas que están en la posición k y la posición l, donde $1 \leq k < l \leq n$.. Si $i \neq k$ o $i \neq l$, entonces, claramente, $det(A_{ij}) = 0$ para todo j porque es el determinante de una matriz que tiene dos filas iguales. Por otra parte, como A_{lj} se obtiene a partir de A_{kj} intercambiando la fila $l-1$ por la filas $k, k+1, \ldots, l-2$, se deduce de la parte 1 del teorema 2.1.2 que $(-1)^{l-k-1} det(A_{lj}) = det(A_{kj})$. En consecuencia,

$$\begin{aligned}
\Psi(A) &= (-1)^{k+j} [A]_{kj} \, det(A_{kj}) + (-1)^{l+j} [A]_{lj} \, det(A_{lj}) \\
&= (-1)^{k+j} [A]_{kj} (-1)^{l-k-1} det(A_{lj}) + (-1)^{l+j} [A]_{lj} \, det(A_{lj}) \\
&= (-1)^{j+l-1} [A]_{kj} \, det(A_{lj}) + (-1)^{l+j} [A]_{kj} \, det(A_{lj}) = 0
\end{aligned}$$

4.

$$\Psi(I_n) = \sum_{i=1}^{n} (-1)^{i+j} [I_n]_{ij} \, det\left((I_n)_{ij}\right) = (-1)^{2j} [I_n]_{jj} \, det\left((I_n)_{jj}\right) = det(I_{n-1}) = 1$$

■

Las fórmulas para el determinante de una matriz encontradas en el teorema 2.4.3 se llaman **la expansión de Laplace** del determinante de A por la fila i y por la columna j respectivamente.

Ejemplo 2.4.4 *Sea* $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$. *Consideremos las siguientes operaciones elementales*:

$$\begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix} \overset{f_{21}(-2)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 \\ 0 & 2 & 1 \\ 1 & 7 & 1 \end{pmatrix} \overset{f_{31}(-1)}{\rightsquigarrow} \begin{pmatrix} 1 & 3 & -2 \\ 0 & 2 & 1 \\ 0 & 4 & 3 \end{pmatrix}$$

Luego por el teorema 2.1.2, $det\,(A) = det \begin{pmatrix} 1 & 3 & -2 \\ 0 & 2 & 1 \\ 0 & 4 & 3 \end{pmatrix}$. *Usamos ahora la expansión de Laplace por la columna* 1 *y obtenemos*

$$det\,(A) = [A]_{11}\, det\,(A_{11}) = 1 det \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} = 6 - 4 = 2$$

Presentamos a continuación una técnica para construir la inversa de una matriz a través de los cofactores de una matriz.

Definición 2.4.5 *La* ***adjunta*** *de* $A \in \mathcal{M}_n(\mathbb{F})$, *que denotamos por* $adj\,(A)$, *la definimos como*

$$[adj\,(A)]_{ij} = (-1)^{i+j}\, det\,(A_{ji})$$

Ejemplo 2.4.6 *Para la matriz*

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

tenemos, por ejemplo, que $[adj\,(A)]_{32} = (-1)^5\, det \begin{pmatrix} 1 & 3 \\ 1 & 7 \end{pmatrix} = -4$. *De manera similar se calculan todos los coeficientes* $[adj\,(A)]_{ij}$ *de la matriz* $adj\,(A)$ *obteniendo así la matriz*

$$adj\,(A) = \begin{pmatrix} 29 & -17 & 7 \\ -5 & 3 & -1 \\ 6 & -4 & 2 \end{pmatrix}$$

Relacionamos la adjunta de una matriz con la inversa en nuestro próximo resultado.

Teorema 2.4.7 *Si* $A \in \mathcal{M}_n(\mathbb{F})$, *entonces* $A adj\,(A) = det\,(A)\, I_n$. *En particular, si* A *es invertible, entonces*

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

Demostración. El término i, j de $Aadj\,(A)$ es

$$[Aadj\,(A)]_{ij} = \sum_{k=1}^{n} [A]_{ik} (-1)^{k+j} \, det\,(A_{jk})$$

Si $i = j$, entonces, por el teorema 2.4.3

$$[Aadj\,(A)]_{ii} = \sum_{k=1}^{n} (-1)^{k+i} [A]_{ik} \, det\,(A_{ik}) = det\,(A)$$

Supongamos que $i \neq j$, digamos $i < j$. Si $A_{1*}, \ldots, A_{n*}$ son las filas de A construyamos la matriz C con filas $C_{1*}, \ldots, C_{n*}$ dadas por $C_{k*} = A_{k*}$ para todo $k \neq j$ y $C_{j*} = A_{i*}$. Entonces es claro que $[C]_{ik} = [A]_{ik} = [C]_{jk}$ y $det\,(A_{jk}) = det\,(C_{jk})$ para todo k. Además, $det\,(C) = 0$ porque tiene dos filas iguales. Luego,

$$\begin{aligned} [Aadj\,(A)]_{ij} &= \sum_{k=1}^{n} [A]_{ik} (-1)^{k+j} \, det\,(A_{jk}) = \sum_{k=1}^{n} (-1)^{k+j} [C]_{jk} \, det\,(C_{jk}) \\ &= det\,(C) = 0 \end{aligned}$$

Esto demuestra que $Aadj\,(A) = det\,(A)\, I_n$.

Si A es invertible, entonces multiplicamos ambos lados por A^{-1}, obteniendo así

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

■

Ejemplo 2.4.8 *Si* $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$ *entonces ya vimos en el ejemplo 2.4.4 que* $det\,(A) = 2$ *y en el ejemplo 2.4.6 que*

$$adj\,(A) = \begin{pmatrix} 29 & -17 & 7 \\ -5 & 3 & -1 \\ 6 & -4 & 2 \end{pmatrix}$$

Se deduce del teorema anterior que

$$A^{-1} = \begin{pmatrix} 29/2 & -17/2 & 7/2 \\ -5/2 & 3/2 & -1/2 \\ 3 & -2 & 1 \end{pmatrix}$$

Es posible aplicar el teorema 2.4.7 para describir la única solución de un sistema de ecuaciones del tipo $AX = B$ cuando A es una matriz invertible.

Corolario 2.4.9 *(Regla de Cramer) Sea* $A \in \mathcal{M}_n(\mathbb{F})$ *y* $B = (b_1, \ldots, b_n)^\top \in \mathcal{M}_{n\times 1}(\mathbb{F})$. *Si* $det(A) \neq 0$, *entonces la ecuación matricial* $AX = B$ *tiene una única solución dada por*

$$[X]_j = \frac{1}{det(A)} \sum_{i=1}^{n} (-1)^{i+j} b_i det(A_{ij})$$

para todo $j = 1, \ldots, n$.

Demostración. Por el teorema 2.2.1, A es invertible porque $det(A) \neq 0$. En consecuencia, por el teorema 1.4.5, $AX = B$ tiene única solución $X = A^{-1}B$. Pero por el teorema 2.4.7,

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

■

EJERCICIOS

1. Para cada $1 \leq i, j \leq 3$, calcule los cofactores de $[A]_{ij}$ para la matriz $A = \begin{pmatrix} 2 & 8 & -1 \\ -2 & 3 & 1 \\ 3 & 1 & 5 \end{pmatrix}$.

2. Usando el teorema 2.4.3 encuentre el determinante de la matrices
$A = \begin{pmatrix} 1 & -1 & 3 & -2 \\ 2 & 3 & 6 & -1 \\ 3 & 1 & 0 & 1 \\ 2 & 2 & -3 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 2 \\ 1 & -1 & 1 \end{pmatrix}$ y $C = \begin{pmatrix} 0 & 1+i & 1+2i \\ 1-i & 0 & 2-3i \\ 1-2i & 2+3i & 0 \end{pmatrix}$

3. Usando el teorema 2.4.3 demuestre que el determinante de una matriz triangular es el producto de los elementos sobre la diagonal.

4. Demuestre que $det(V) = \prod_{1\leq i<j\leq n} (x_j - x_i)$, donde V es la matriz de Vandermonde

$$V = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$$

5. Considere la matriz $B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 5 & 8 & 9 \end{pmatrix}$. Encuentre *i)* $det(B)$; *ii)* $adj(B)$ y *iii)* B^{-1} usando $adj(B)$.

6. Determine si las matrices dadas a continuación son invertibles. Si son invertibles, calcule la inversa usando el teorema 2.4.7:

$$A = \begin{pmatrix} 3 & 6 \\ -4 & -8 \end{pmatrix}, \; B = \begin{pmatrix} 2 & -1 & 4 \\ -1 & 0 & 5 \\ 19 & -7 & 3 \end{pmatrix}, C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{pmatrix}$$

7. Resuelva cada uno de los sistemas de ecuaciones usando la Regla de Cramer:

$$\text{a)} \quad \left.\begin{array}{c} 2x + y + z = 6 \\ 3x - 2y - 3z = 5 \\ 8x + 2y + 5z = 11 \end{array}\right\} \; ; \; \text{b)} \quad \left.\begin{array}{c} 2x + y - z = 4 \\ x + z = 2 \\ -y + 5z = 1 \end{array}\right\} .$$

8. Demuestre que si A es una matriz $n \times n$, entonces

   a) $det\,(adj\,(A)) = det\,(A)^{n-1}$;
   b) $adj\,(A)^{\top} = adj\left(A^{\top}\right)$;
   c) $adj\,(adj\,(A)) = det\,(A)^{n-2}\,A$.

CAPÍTULO 3

ESPACIOS VECTORIALES

3.1. Espacios vectoriales

Comenzamos este capítulo generalizando los espacios conocidos $\mathbb{R}^2$ y $\mathbb{R}^3$ a espacios de n dimensiones. Recordamos que $\mathbb{R}^2$ está formado por pares ordenados (x_1, x_2), donde x_1 y x_2 son números reales. Los elementos de $\mathbb{R}^2$ los visualizamos como puntos del plano cartesiano y estos a su vez están en correspondencia natural con los vectores del plano que tienen punto inicial en el origen. Similarmente, $\mathbb{R}^3$ consiste en ternas (x_1, x_2, x_3), donde x_1, x_2 y x_3 son números reales y estos los identificamos con puntos del espacio tridimensional o con vectores del espacio con punto inicial en el origen.

Más generalmente, dado un entero $n \geq 1$, denotamos por $\mathbb{F}^n$ ($\mathbb{F} = \mathbb{R}$ o $\mathbb{F} = \mathbb{C}$) al conjunto de todas las n-uplas $(x_1, \ldots, x_n)$, donde cada $x_i \in \mathbb{F}$. Por analogía a los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$, llamamos vectores a los elementos de $\mathbb{F}^n$. Para cada $k = 1, \ldots, n$, decimos que x_k es la coordenada k-ésima de la n-upla. Las n-uplas $x = (x_1, \ldots, x_n)$ y $y = (y_1, \ldots, y_n)$ de $\mathbb{F}^n$ son iguales si, y sólo si, $x_k = y_k$ para todo $k = 1, \ldots, n$.

Extendemos la suma de vectores y multiplicación escalar conocidas para $\mathbb{R}^2$ y $\mathbb{R}^3$ a $\mathbb{F}^n$ de la siguiente manera:

Definición 3.1.1 *Sean $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{F}^n$. La suma de x y y, denotado por $x + y$, es la n-upla*

$$x + y = (x_1 + y_1, \ldots, x_n + y_n).$$

Si $\alpha \in \mathbb{F}$, la multiplicación del escalar α por x se denota por αx y se define como la n-upla

$$\alpha x = (\alpha x_1, \ldots, \alpha x_n)$$

El vector cero de $\mathbb{F}^n$ se denota por $\mathbf{0}$ y se define como la n-upla $\mathbf{0} = (0, \ldots, 0)$. Si

$$x = (x_1, \ldots, x_n) \in \mathbb{F}^n,$$

entonces el inverso aditivo de x se denota por $-x$ y se define como

$$-x = (-x_1, \ldots, -x_n).$$

Ejemplo 3.1.2 *Consideremos los vectores* $x = (3, 2 - i, 1)$, $y = (i, 2 - i, 0)$ *de* $\mathbb{C}^3$ *junto con los escalares* $\alpha = i$ *y* $\beta = -3$. *Entonces*

$$\begin{aligned} \alpha x + \beta y &= i(3, 2 - i, 1) - 3(i, 2 - i, 0) \\ &= (3i, 2i + 1, i) - (3i, 6 - 3i, 0) \\ &= (0, -5 + 5i, i) \end{aligned}$$

El teorema a continuación describe las propiedades básicas que satisfacen estas operaciones.

Teorema 3.1.3 *Sean* $x, y, z \in \mathbb{F}^n$ *y* $\alpha, \beta \in \mathbb{F}$. *Entonces*

1. $x + y = y + x$;
2. $(x + y) + z = x + (y + z)$;
3. $x + \mathbf{0} = x$;
4. $x + (-x) = \mathbf{0}$;
5. $\alpha(x + y) = \alpha x + \alpha y$;
6. $(\alpha + \beta) x = \alpha x + \beta x$;
7. $(\alpha\beta) x = \alpha(\beta x)$;
8. $1x = x$.

Demostración. Demostramos 5. Sean $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{F}^n$ y $\alpha \in \mathbb{F}$. Entonces

$$\begin{aligned} \alpha(x + y) &= \alpha(x_1 + y_1, \ldots, x_n + y_n) \\ &= (\alpha(x_1 + y_1), \ldots, \alpha(x_n + y_n)) \\ &= (\alpha x_1 + \alpha y_1, \ldots, \alpha x_n + \alpha y_n) \\ &= (\alpha x_1, \ldots, \alpha x_n) + (\alpha y_1, \ldots, \alpha y_n) \\ &= \alpha x + \alpha y \end{aligned}$$

■

Llamamos a $\mathbb{F}^n$ junto con las dos operaciones introducidas en la definición 3.1.1 el espacio de las n-uplas sobre $\mathbb{F}$. Es claro que cuando $\mathbb{F} = \mathbb{R}$, si $n = 2$ y $n = 3$, entonces recuperamos los espacios conocidos $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente.

Es importante resaltar que los vectores de $\mathbb{F}^n$ pueden representarse también como vectores columna $(x_1, \ldots, x_n)^\top$, $x_i \in \mathbb{F}$. En este caso, las operaciones toman la forma

$$\begin{aligned} (x_1, \ldots, x_n)^\top + (y_1, \ldots, y_n)^\top &= (x_1 + y_1, \ldots, x_n + y_n)^\top \\ \alpha (x_1, \ldots, x_n)^\top &= (\alpha x_1, \ldots, \alpha x_n)^\top \end{aligned}$$

De esta manera es posible identificar al espacio $\mathbb{F}^n$ con el espacio de las matrices $\mathcal{M}_{n\times 1}(\mathbb{F})$. Bajo esta identificación, si $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $x \in \mathbb{F}^n$, entonces $Ax \in \mathbb{F}^m$.

Ejemplo 3.1.4 *Sea* $A = \begin{pmatrix} 2 & i & -1 \\ i+1 & 0 & 1 \end{pmatrix} \in \mathcal{M}_{n\times 1}(\mathbb{C})$ *y* $x = \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix} \in \mathbb{C}^3$. *Entonces*

$$Ax = \begin{pmatrix} 2 & i & -1 \\ i+1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix} = \begin{pmatrix} 2-i \\ 2i+1 \end{pmatrix} \in \mathbb{C}^2$$

El espacio de las matrices y el espacio de las n-uplas son ejemplos de sistemas algebraicos en los cuales es posible sumar sus elementos y multiplicar por escalares. En distintas áreas de las matemáticas aparecen estructuras similares, por lo tanto, es razonable presentar una noción general que abarque cada uno de estos casos particulares.

Definición 3.1.5 *Un espacio vectorial V sobre $\mathbb{F}$ es un conjunto no vacío V junto con dos operaciones. En primer lugar, la suma que asocia a cada par de vectores u, v de V un vector $u + v$ de V que satisface las condiciones:*

1. *Conmutatividad:* $u + v = v + u$ *para todo* $u, v \in V$;
2. *Asociatividad:* $(u + v) + w = u + (v + w)$ *para todo* $u, v, w \in V$;
3. *Identidad aditiva: existe un vector* 0, *llamada vector cero, tal que* $u + 0 = u$ *para todo* $u \in V$;
4. *Inverso aditivo: para cada vector* $u \in V$ *existe un vector* $z \in V$ *tal que* $u + z = 0$.

La segunda operación, que llamamos multiplicación por escalar, asocia a cada escalar $\alpha \in \mathbb{F}$ y vector $u \in V$ un vector αu de V que satisface las condiciones:

5. $\alpha(\beta u) = (\alpha\beta) u$ *para todo* $\alpha, \beta \in \mathbb{F}$ *y* $u \in V$;
6. $\alpha(u + v) = \alpha u + \alpha v$ *para todo* $\alpha \in \mathbb{F}$ *y* $u, v \in V$;
7. $(\alpha + \beta) u = \alpha u + \beta u$ *para todo* $\alpha, \beta \in \mathbb{F}$ *y* $u \in V$;
8. $1u = u$ *para todo* $u \in V$.

Ejemplo 3.1.6 *Cada uno de los siguientes conjuntos con las operaciones indicadas es un espacio vectorial.*

1. *El espacio de las* n*-uplas* $\mathbb{F}^n$ *es un espacio vectorial sobre* $\mathbb{F}$ *con las operaciones*

$$\begin{aligned}(x_1,\ldots,x_n)^\top + (y_1,\ldots,y_n)^\top &= (x_1+y_1,\ldots,x_n+y_n)^\top \\ \alpha(x_1,\ldots,x_n)^\top &= (\alpha x_1,\ldots,\alpha x_n)^\top\end{aligned}$$

   donde $(x_1,\ldots,x_n)^\top, (y_1,\ldots,y_n)^\top \in \mathbb{F}^n$ *y* $\alpha \in \mathbb{F}$.

2. *El espacio de las matrices* $\mathcal{M}_{m\times n}(\mathbb{F})$ *es un espacio vectorial sobre* $\mathbb{F}$ *con las operaciones*

$$\begin{aligned}[A+B]_{ij} &= [A]_{ij} + [B]_{ij} \\ [\alpha A]_{ij} &= \alpha [A]_{ij}\end{aligned}$$

   donde $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$ *y* $\alpha \in \mathbb{F}$.

3. *Sea* X *un conjunto no vacío y* $\mathcal{F}(X,\mathbb{F})$ *el conjunto de las funciones* $X \to \mathbb{F}$. *Si*

$$f, g \in \mathcal{F}(X,\mathbb{F})$$

   y $\alpha \in \mathbb{F}$ *definimos las operaciones* $f+g$ *y* αf *por*

$$\begin{aligned}(f+g)(x) &= f(x) + g(x) \\ (\alpha f)(x) &= \alpha f(x)\end{aligned}$$

   para todo $x \in X$.

4. *El conjunto* $\mathbb{F}[x]$ *de todos los polinomios con coeficientes en* $\mathbb{F}$ *junto con las operaciones usuales*

$$\begin{aligned}\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i &= \sum_{i=0}^{n} (a_i+b_i)\, x^i \\ \alpha \sum_{i=0}^{n} a_i x^i &= \sum_{i=0}^{n} (\alpha a_i)\, x^i\end{aligned}$$

A partir de este momento, cuando decimos que V es un espacio vectorial entendemos que V es un espacio vectorial sobre $\mathbb{F}$, junto con las dos operaciones que satisfacen las ocho condiciones de la definición anterior.

Comenzamos el estudio sistemático de los espacios vectoriales deduciendo algunas propiedades inmediatas de la definición. Entre los requisitos de un espacio vectorial está el de la existencia de identidad aditiva e inverso aditivo. El próximo resultado establece que son únicos.

Teorema 3.1.7 *Sea* V *un espacio vectorial sobre* $\mathbb{F}$. *Entonces:*

1. La identidad aditiva es única;

2. El inverso aditivo de cada elemento es único.

Demostración. 1. Supongamos que 0 y $0'$ son identidades aditivas. Entonces, por ser 0 identidad aditiva tenemos que $0' = 0 + 0'$, pero al ser $0'$ identidad aditiva $0 + 0' = 0$. En consecuencia, $0 = 0'$.

2. Sea $u \in V$ y supongamos que v, w son inversos aditivos de u. Entonces

$$v = v + 0 = v + (u + w) = (v + u) + w = 0 + w = w$$

■

A partir de ahora denotamos por $-u$ al (único) inverso aditivo de u. Además, escribimos $u - v$ en lugar de $u + (-v)$.

Teorema 3.1.8 *Sea V un espacio vectorial sobre $\mathbb{F}$, $u, v \in V$ y $\alpha, \beta \in \mathbb{F}$. Entonces:*

1. $\alpha 0 = 0$;

2. $0v = 0$;

3. $(-\alpha) v = -(\alpha v) = \alpha(-v)$;

4. $(-1) v = -v$;

5. $\alpha(u - v) = \alpha u - \alpha v$;

6. $(\alpha - \beta) v = \alpha v - \beta v$;

7. $\alpha v = 0$ *si, y sólo si,* $\alpha = 0$ *o* $v = 0$.

Demostración. 2. $0v = (0 + 0) v = 0v + 0v$. Luego sumando el inverso aditivo de $0v$ ambos lados de la ecuación obtenemos $0v = 0$.

3. $(-\alpha) v + \alpha v = (-\alpha + \alpha) v = 0v = 0$. Esto demuestra que $(-\alpha) v = -(\alpha v)$. Para ver la otra igualdad observamos que $\alpha(-v) + \alpha v = \alpha(-v + v) = \alpha 0 = 0$.

7. Supongamos que $\alpha v = 0$ y $\alpha \neq 0$. Entonces

$$0 = \frac{1}{\alpha}(\alpha v) = \left(\frac{1}{\alpha}\alpha\right) v = 1v = v.$$

La recíproca es consecuencia de 1. y 2. ■

Estudiamos a continuación los subconjuntos de un espacio vectorial V que a su vez tienen estructura de espacio vectorial, con las mismas operaciones definidas en V.

Definición 3.1.9 *Sea S un subconjunto no vacío de un espacio vectorial V. Decimos que S es un **subespacio** de V si satisface las siguientes condiciones:*

1. Si $u, v \in S$, entonces $u + v \in S$;

2. Si $\alpha \in \mathbb{F}$ y $u \in S$, entonces $\alpha u \in S$.

Un espacio vectorial V siempre tiene los subespacios triviales $\{0\}$ y V.

Ejemplo 3.1.10 *Los siguientes subconjuntos son subespacios del espacio indicado:*

1. Sea $S = \{(x, y) \in \mathbb{R}^2 : y = 0\}$. Entonces
$$(x, 0) + (x', 0) = (x + x', 0) \in S$$
y si $\alpha \in \mathbb{R}$
$$\alpha (x, 0) = (\alpha x, 0) \in S$$
Luego S es un subespacio de $\mathbb{R}^2$.

2. Cualquier recta de $\mathbb{R}^2$ que pasa por el origen es un subespacio de $\mathbb{R}^2$.

3. Sea $S = \{(x, y) \in \mathbb{R}^2 : y = 2x + 1\}$. S no es un subespacio de $\mathbb{R}^2$. Por ejemplo
$$(0, 1), (1, 3) \in S$$
y, sin embargo, $(0, 1) + (1, 3) = (1, 4) \notin S$.

4. Los únicos subespacios de $\mathbb{R}$ como espacio sobre sí mismo son los triviales. En efecto, supongamos que S es un subespacio de $\mathbb{R}$ distinto de cero y $0 \neq z \in S$. Entonces, para cualquier $x \in \mathbb{R}$ tenemos $x = \frac{x}{z} z \in S$. Luego, $S = \mathbb{R}$.

5. Sea $A \in \mathcal{M}_{m \times n} (\mathbb{F})$. Definimos el **núcleo** de A, denotado por $ker (A)$, al subconjunto de $\mathbb{F}^n$
$$ker (A) = \{x \in \mathbb{F}^n : Ax = 0\}$$
$ker (A)$ es el conjunto solución del sistema de ecuaciones homogéneo $Ax = 0$. Observamos que $0 \in ker (A)$. Supongamos que $x, y \in ker (A)$ y $\alpha \in \mathbb{F}$. Entonces
$$A (x + y) = Ax + Ay = 0 + 0 = 0$$
y
$$A (\alpha x) = \alpha Ax = \alpha 0 = 0$$
Luego $x + y$ y $\alpha x \in ker (A)$ y, por lo tanto, $ker (A)$ es un subespacio de $\mathbb{F}^n$.

6. Sea $A \in \mathcal{M}_{m \times n} (\mathbb{F})$. Definimos la **imagen** de A, denotado por $Im (A)$ como el subconjunto de $\mathbb{F}^m$ dado por
$$Im (A) = \{Ax : x \in \mathbb{F}^n\}$$
Es fácil ver que $Im (A)$ es un subespacio de $\mathbb{F}^m$.

7. Sea k un entero positivo y
$$S = \{p(x) \in \mathbb{F}[x] : grad\,(p(x)) \leq k\}.$$
Si $\alpha \in \mathbb{F}$ y $p(x), q(x) \in \mathbb{F}[x]$, entonces es fácil ver que
$$grad\,[p(x) + q(x)] \leq k$$
y
$$grad\,[\alpha p(x)] \leq k$$
En consecuencia, S es un subespacio de $\mathbb{F}[x]$.

Si V es un espacio vectorial, entonces cualquier subespacio S de V tiene al 0. En efecto, dado $x \in S$ tenemos que $0 = 0x \in S$.

Teorema 3.1.11 *La intersección de cualquier familia de subespacios de un espacio vectorial V es un subespacio de V.*

Demostración. Si $\{S_i : i \in I\}$ es una familia de subespacios de V, entonces $0 \in \bigcap\limits_{i\in I} S_i$ y así, $\bigcap\limits_{i\in I} S_i \neq \emptyset$. Sea $\alpha \in \mathbb{F}$ y $x, y \in \bigcap\limits_{i\in I} S_i$. Luego $x, y \in S_i$ para todo $i \in I$ y, como cada S_i es un subespacio, $x + y$ y $\alpha x \in S_i$ para todo $i \in I$. En consecuencia, $x + y$ y αx pertenecen a $\bigcap\limits_{i\in I} S_i$, lo que demuestra que $\bigcap\limits_{i\in I} S_i$ es un subespacio de V. ■

EJERCICIOS

1. Demuestre que cada uno de los conjuntos con las operaciones indicadas en el ejemplo 3.1.6 son espacios vectoriales.

2. Complete la demostración de los teoremas 3.1.3 y 3.1.8.

3. ¿Es $\mathbb{R}^2$ con las operaciones indicadas a continuación un espacio vectorial sobre $\mathbb{R}$?

   a) $(x, y) + (w, z) = (x + w, y + z)$ y $\alpha(x, y) = (\alpha x, y)$;
   b) $(x, y) + (w, z) = (x + w, 0)$ y $\alpha(x, y) = (\alpha x, 0)$.

4. Sean U y W espacios vectoriales. Demuestre que el producto directo $U \times W$ es un espacio vectorial con las operaciones
$$\begin{aligned} (u, w) + (u', w') &= (u + u', w + w') \\ \alpha(u, w) &= (\alpha u, \alpha w) \end{aligned}$$
donde $u, u' \in U$, $w, w' \in W$ y $\alpha \in \mathbb{F}$.

5. ¿Cuáles de los siguientes subconjuntos S es un subespacio de V?

   a) $V = \mathbb{R}^n$ y $S = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 \leq 0\}$;
   b) $V = \mathbb{R}^n$ y $S = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 \text{ es racional}\}$.
   c) $V = \mathcal{M}_n(\mathbb{C})$ y $S = \{A \in \mathcal{M}_n(\mathbb{C}) : A^* = A\}$;
   d) $V = \mathcal{M}_n(\mathbb{F})$ y $S = \{A \in \mathcal{M}_n(\mathbb{F}) : A \text{ es invertible}\}$;
   e) $V = \mathcal{M}_n(\mathbb{F})$ y $S = \{A \in \mathcal{M}_n(\mathbb{F}) : A \text{ conmuta con } P\}$;
   f) $V = \mathcal{M}_n(\mathbb{F})$ y $S = \{T \in \mathcal{M}_n(\mathbb{F}) : T \text{ es triangular superior}\}$;
   g) $V = \mathcal{M}_n(\mathbb{F})$ y $S = \{D \in \mathcal{M}_n(\mathbb{F}) : D \text{ es diagonal}\}$;
   h) $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ y $S = \{g \in \mathcal{F}(\mathbb{R}, \mathbb{R}) : g(1) = g(0)\}$;

6. Demuestre que S es un subespacio de V si, y sólo si, S tiene estructura de espacio vectorial con las mismas operaciones definidas en V.

3.2. Subespacios generados

Si X es un subconjunto no vacío de un espacio vectorial V, entonces, por el teorema 3.1.11, la intersección de todos los subespacios de V que contienen a X es un subespacio de V.

Definición 3.2.1 *Sea V un espacio vectorial y X un subconjunto de V. La intersección de todos los subespacios de V que contienen a X se llama el* ***subespacio generado*** *por X y lo denotamos por $gen(X)$.*

El subespacio $gen(X)$ es el menor subespacio de V que contiene a X. En el próximo resultado describimos la naturaleza de los elementos de $gen(X)$.

Definición 3.2.2 *Sea V un espacio vectorial sobre $\mathbb{F}$. Una* ***combinación lineal*** *de los vectores $v_1, \ldots, v_k$ de V es un elemento de V de la forma*

$$\alpha_1 v_1 + \cdots + \alpha_k v_k$$

donde $\alpha_1, \ldots, \alpha_k$ son elementos de $\mathbb{F}$.

Ejemplo 3.2.3 *Consideremos los vectores $u = (1, 2)$ y $v = (-1, 1)$ de $\mathbb{R}^2$. Entonces una combinación lineal de estos vectores es, por ejemplo,*

$$-3u + 5v = (-8, -1)$$

Más generalmente, el conjunto de todas las combinaciones lineales de u y v vienen dadas por

$$\lambda u + \mu v = (\lambda - \mu, 2\lambda + \mu)$$

donde $\lambda, \mu \in \mathbb{R}$. Observamos que en este caso, todo vector $(x, y) \in \mathbb{R}^2$ es una combinación lineal de los vectores u, v. En efecto, basta con resolver las ecuaciones $x = \lambda - \mu$ y $y = 2\lambda + \mu$ obteniendo que $\lambda = \frac{x+y}{3}$ y $\mu = \frac{y-2x}{3}$. Así

$$(x, y) = \frac{x+y}{3}u + \frac{y-2x}{3}v$$

Teorema 3.2.4 *Sea X un subconjunto no vacío del espacio vectorial V. Entonces los elementos de $gen\,(X)$ son combinaciones lineales de elementos de X.*

Demostración. Llamemos $\mathcal{C}$ al conjunto de todas las combinaciones lineales de vectores de X. Entonces $\mathcal{C}$ es un subespacio de V. En efecto, si $x, y \in \mathcal{C}$, entonces, tomando algunos escalares igual a cero si es necesario, $x = \sum_{i=1}^{n} \alpha_i x_i$ y $y = \sum_{i=1}^{n} \beta_i x_i$, donde $\alpha_i, \beta_i \in \mathbb{F}$ y $x_i \in X$ para todo $i = 1, \ldots, n$. Así, para $\alpha \in \mathbb{F}$ tenemos que

$$x + y = \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{n} \beta_i x_i = \sum_{i=1}^{n} (\alpha_i + \beta_i)\, x_i \in \mathcal{C}$$

y

$$\alpha x = \alpha \sum_{i=1}^{n} \alpha_i x_i = \sum_{i=1}^{n} (\alpha\alpha_i)\, x_i \in \mathcal{C}$$

Además, si $x \in X$ entonces $x = 1x \in \mathcal{C}$, por lo tanto, $\mathcal{C}$ es un subespacio de V que contiene a X. De manera que $gen\,(X) \subseteq \mathcal{C}$. Recíprocamente, supongamos que S es un subespacio de V que contiene a X. Entonces, por la definición de subespacio, contiene toda combinación lineal de vectores de X. Es decir, $\mathcal{C} \subseteq \mathcal{S}$. Demostramos así que $\mathcal{C}$ está contenido en cualquier subespacio de V que contenga a X y, en consecuencia, $\mathcal{C} \subseteq gen\,(X)$. Por lo tanto, $gen\,(X) = \mathcal{C}$. ■

Cuando X es finito, digamos $X = \{x_1, \ldots, x_n\}$, escribimos directamente

$$gen\,(X) = gen\,(x_1, \ldots, x_n)$$

Ejemplo 3.2.5 *Veamos algunos ejemplos de subespacios generados por un subconjunto de un espacio vectorial.*

1. *Si $u = (1, 2)$ y $v = (-1, 1)$ en $\mathbb{R}^2$ entonces por el ejemplo 3.2.3, $gen\,(u, v) = \mathbb{R}^2$.*

2. *Sea $0 \neq u \in \mathbb{R}^3$. Entonces el subespacio de $\mathbb{R}^3$ generado por u es*

$$gen\,(u) = \{\lambda u : \lambda \in \mathbb{R}\}$$

Esta es la recta en el espacio que pasa por el origen en la dirección de u.

3. *Sean* $u, v \in \mathbb{R}^3$ *diferentes de cero. Entonces*

$$gen(u, v) = \{\lambda u + \mu v : \lambda, \mu \in \mathbb{R}\}$$

Observamos que si $v \in gen(u)$, *entonces*

$$gen(u, v) = gen(u)$$

En efecto, $v = \alpha u$ *y, por lo tanto, para escalares* $\lambda, \mu \in \mathbb{R}$ *tenemos que*

$$\lambda u + \mu v = \lambda u + \mu(\alpha u) = (\lambda + \mu\alpha) u \in gen(u)$$

En caso que $v \notin gen(u)$, *entonces* $gen(u, v)$ *es el plano en* $\mathbb{R}^3$ *que pasa por el origen generado por los vectores* u, v.

4. *Consideremos para cada* $i = 1, \ldots, n$ *los vectores* $e_i \in \mathbb{F}^n$ *que tienen todas las coordenadas iguales a* 0 *excepto la coordenada* i, *que es* 1. *Entonces*

$$\begin{aligned} gen(e_1, \ldots, e_n) &= \{\lambda_1 e_1 + \cdots + \lambda_n e_n : \lambda_1, \ldots, \lambda_n \in \mathbb{F}\} \\ &= \{(\lambda_1, \ldots, \lambda_n) : \lambda_1, \ldots, \lambda_n \in \mathbb{F}\} = \mathbb{F}^n \end{aligned}$$

Los vectores $e_1, \ldots, e_n$ *reciben el nombre de vectores canónicos de* $\mathbb{F}^n$.

5. *El subespacio de* $\mathbb{F}[x]$ *generado por los polinomios* $1, x, x^2, \ldots, x^k$ *es*

$$S = \{f(x) \in \mathbb{F}[x] : f(x) \text{ tiene grado } \leq k\}$$

6. *El subespacio de* $\mathcal{M}_2(\mathbb{F})$ *generado por las matrices*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ y } B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

es

$$gen(A, B) = \left\{ \begin{pmatrix} \lambda & \mu \\ 0 & 0 \end{pmatrix} : \lambda, \mu \in \mathbb{F} \right\}$$

7. *Sea* $A \in \mathcal{M}_{m \times n}(\mathbb{F})$. *El* ***espacio de filas*** *de* A *lo denotamos por* $\mathcal{F}(A)$, *y se define como el subespacio de* $\mathbb{F}^n$ *generado por las filas de* A:

$$\mathcal{F}(A) = gen(A_{1*}, \ldots, A_{m*})$$

8. *Sea* $A \in \mathcal{M}_{m \times n}(\mathbb{F})$. *El* ***espacio de columnas*** *de* A *lo denotamos por* $\mathcal{C}(A)$, *y se define como el subespacio de* $\mathbb{F}^m$ *generado por las columnas de* A:

$$\mathcal{C}(A) = gen(A_{*1}, \ldots, A_{*n})$$

Definición 3.2.6 *Un espacio vectorial V está* ***generado*** *por un subconjunto X si $V = gen\,(X)$. También decimos que X* ***genera*** *al espacio V. Si X es finito, entonces V es* ***finitamente generado****.*

Ejemplo 3.2.7 *El espacio $\mathbb{F}^n$ es un espacio vectorial finitamente generado. De hecho, los vectores canónicos $e_1, \ldots, e_n$ generan a $\mathbb{F}^n$.*

Ejemplo 3.2.8 *El espacio vectorial $\mathbb{F}[x]$ no es finitamente generado. Para ver esto supongamos que $p_1(x), \ldots, p_n(x)$ generan a $\mathbb{F}[x]$. Elijamos k el máximo grado entre los polinomios $p_1(x), \ldots, p_n(x)$. Entonces cualquier combinación lineal de estos polinomios tiene grado menor o igual a k. Por lo tanto, si tomamos un polinomio $g(x) \in \mathbb{F}[x]$ de grado estrictamente mayor que k, entonces es claro que $g(x) \notin gen\,(p_1(x), \ldots, p_n(x))$. Sin embargo, es fácil ver que $\mathbb{F}[x]$ está generado por los polinomios $\left\{1, x, x^2, \ldots, x^k, \ldots\right\}$.*

Ejemplo 3.2.9 *El subespacio de $\mathbb{F}[x]$*

$$S = \{p(x) \in \mathbb{F}[x] : grad\,(p(x)) \leq k\}$$

es finitamente generado porque $S = gen\left(1, x, x^2, \ldots, x^k\right)$.

En general, si V es un espacio vectorial y S_1, S_2 son subespacios de V, entonces $S_1 \cup S_2$ no es necesariamente un subespacio de V (ver Ejercicio 8). Por ejemplo, en $\mathbb{R}^2$,

$$gen\,((1,1)) \cup gen\,((-1,1))$$

contiene a $(1,1)$ y $(-1,1)$ pero no a $(0,2) = (1,1) + (-1,1)$. Describimos a continuación el menor subespacio de V que contiene a $S_1 \cup S_2$, esto es, $gen\,(S_1 \cup S_2)$.

Definición 3.2.10 *Sea V un espacio vectorial y S_1, S_2 subespacios de V. La* ***suma*** *de estos subespacios, denotada por $S_1 + S_2$, se define como*

$$S_1 + S_2 = \{v \in V : v = s_1 + s_2 \text{ donde } s_i \in S_i\}$$

Teorema 3.2.11 *Sea V un espacio vectorial y S_1, S_2 subespacios de V. Entonces*

$$S_1 + S_2 = gen\,(S_1 \cup S_2).$$

En otras palabras, $S_1 + S_2$ es el menor subespacio de V que contiene a cada uno de los S_i.

Demostración. Sean $x, y \in S = S_1 + S_2$ y $\alpha \in \mathbb{F}$. Entonces

$$x = s_1 + s_2 \text{ y } y = s_1' + s_2'$$

donde $s_i, s_i' \in S_i$ para $i = 1, 2$. En consecuencia,

$$\begin{aligned} x + y &= s_1 + s_2 + s_1' + s_2' \\ &= (s_1 + s_1') + (s_2 + s_2') \in S \end{aligned}$$

y

$$\alpha x = \alpha (s_1 + s_2) = \alpha s_1 + \alpha s_2 \in S$$

porque $s_i + s_i', \alpha s_i \in S_i$ para $i = 1, 2$, al ser cada S_i un subespacio. Esto demuestra que S es un subespacio que además contiene claramente a $S_1 \cup S_2$. Por lo tanto, $gen\,(S_1 \cup S_2) \subseteq S$. Por otra parte, como $gen\,(S_1 \cup S_2)$ está formado por todas las combinaciones lineales finitas de elementos de $S_1 \cup S_2$, entonces es claro que $S \subseteq gen\,(S_1 \cup S_2)$. Así, $S = gen\,(S_1 \cup S_2)$. ■

Ejemplo 3.2.12 *Consideremos los subespacios de* $\mathbb{F}^3$

$$S_1 = \left\{(x, y, z) \in \mathbb{F}^3 : z = 0\right\} \text{ y } S_2 = \left\{(x, y, z) \in \mathbb{F}^3 : x = 0\right\}$$

Entonces $\mathbb{F}^3 = S_1 + S_2$. *En efecto, si* $(x, y, z) \in \mathbb{F}^3$, *entonces*

$$(x, y, z) = (x, y, 0) + (0, 0, z)$$

donde $(x, y, 0) \in S_1$ *y* $(0, 0, z) \in S_2$.

EJERCICIOS

1. Demuestre que los vectores $(2, 1, 3), (2, 1, 0)$ y $(1, 5, 5)$ generan a $\mathbb{R}^3$.

2. ¿Pertenece $(-4, -8, 1, -8, 0)$ al subespacio generado por los vectores

$$(1, 2, 0, 3, 0), (0, 0, 1, 4, 0), (0, 0, 0, 0, 1)\,?$$

3. Exprese el polinomio $p(x) = x^2 + 4x - 3$ como combinación lineal de los polinomios

$$u(x) = x^2 - 2x + 5,\ v(x) = 2x^2 - 3x \text{ y } w(x) = x + 3$$

4. Escriba la matriz $P = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$ como combinación lineal de las matrices

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

5. Sea $A = \begin{pmatrix} 2 & -1 & 3 & 2 \\ -1 & 1 & 1 & -3 \\ 1 & 1 & 9 & -5 \end{pmatrix}$. ¿$(3, -1, 0, -1) \in \mathcal{F}(A)$?

6. Sea $A \in \mathcal{M}_{m \times n}(\mathbb{F})$ y $S = \{B \in \mathcal{M}_{m \times 1}(\mathbb{F}) : AX = B \text{ tiene solución}\}$. Demuestre que S es un subespacio de $\mathcal{M}_{m \times 1}(\mathbb{F})$. ¿Qué relación existe entre S y $\mathcal{C}(A)$?

7. Demuestre que la solución del sistema homogéneo

$$\begin{aligned} 6x - 3y + 4w - 3z &= 0 \\ 3x + 2w - 3u &= 0 \\ 3x - y + 2w - w - u &= 0 \end{aligned}$$

es un subespacio finitamente generado.

8. Sean S_1, S_2 son subespacios de un espacio vectorial V. Demuestre que $S_1 \cup S_2$ es un subespacio de V si, y sólo si, $S_1 \subseteq S_2$ o $S_2 \subseteq S_1$.

9. Decimos que V es **suma directa** de los subespacios S_1 y S_2, y escribimos $V = S_1 \oplus S_2$, si $V = S_1 + S_2$ y $S_1 \cap S_2 = \emptyset$. Demuestre que $V = S_1 \oplus S_2$ si, y sólo si para cada $v \in V$ existen elementos únicos $x \in S_1$ y $y \in S_2$ tales que $v = x + y$.

10. En cada uno de los casos determine si $V = S_1 \oplus S_2$:

    a) $V = \mathcal{M}_2(\mathbb{C})$,

    $$\begin{aligned} S_1 &= \left\{ A \in \mathcal{M}_2(\mathbb{C}) : A = \begin{pmatrix} 0 & a \\ b & c \end{pmatrix} \right\} \text{ y} \\ S_2 &= \left\{ A \in \mathcal{M}_2(\mathbb{C}) : A = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \right\} \end{aligned}$$

    donde $a, b, c \in \mathbb{C}$;

    b) $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$,

    $$\begin{aligned} S_1 &= \{ g \in \mathcal{F}(\mathbb{R}, \mathbb{R}) : g(x) = g(-x) \} \text{ y} \\ S_2 &= \{ g \in \mathcal{F}(\mathbb{R}, \mathbb{R}) : g(x) = -g(-x) \} \end{aligned}$$

    c) $V = \mathcal{M}_n(\mathbb{F})$, S_1 el subespacio de las matrices simétricas y S_2 el subespacio de las matrices antisimétricas;

    d) $V = \mathcal{M}_n(\mathbb{F})$, S_1 el subespacio de las matrices triangulares superiores y S_2 el subespacio de las matrices triangulares inferiores.

3.3. Independencia lineal y bases

Consideremos la matriz $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 2 & -2 & 2 & -4 \\ 8 & -6 & 6 & -10 \\ 4 & -3 & 3 & -5 \end{pmatrix}$. Recordamos que el espacio de filas de A es por definición

$$\mathcal{F}(A) = gen\left((1,0,0,1), (2,-2,2,-4), (8,-6,6,-10), (4,-3,3,-5)\right)$$

Algunos vectores generadores de $\mathcal{F}(A)$ son superfluos en el sentido que podemos eliminarlos y los vectores restantes siguen generando a $\mathcal{F}(A)$. Esta situación ocurre porque algunos vectores se escriben como combinación lineal de otros:

$$(8,-6,6,-10)=2(1,0,0,1)+3(2,-2,2,-4)$$

y

$$(4,-3,3,-5)=(1,0,0,1)+\frac{3}{2}(2,-2,2,-4)$$

lo que nos lleva a que $\mathcal{F}(A)=gen[(1,0,0,1),(2,-2,2,-4)]$. Tratemos de precisar mejor esta idea.

Definición 3.3.1 *Sea V un espacio vectorial. Los vectores $v_1,\ldots,v_n$ de V son* ***linealmente dependientes*** *si existen escalares $\alpha_1,\ldots,\alpha_n$ de $\mathbb{F}$, no todos iguales a cero, tales que*

$$\alpha_1 v_1+\cdots+\alpha_n v_n=0$$

Si los vectores $v_1,\ldots,v_n$ de V no son linealmente dependientes decimos que son ***linealmente independientes****.*

Para demostrar que los vectores $v_1,\ldots,v_n$ de V son linealmente independientes debemos ver lo siguiente: si $\alpha_1 v_1+\cdots+\alpha_n v_n=0$, donde $\alpha_1,\ldots,\alpha_n\in\mathbb{F}$, entonces $\alpha_1=\cdots=\alpha_n=0$.

Se deduce fácilmente de la definición que cualquier subconjunto de un conjunto de vectores linealmente independiente es linealmente independiente. Dualmente, cualquier conjunto que contenga a un conjunto de vectores linealmente dependientes es linealmente dependiente.

Ejemplo 3.3.2 *Ilustramos el concepto de independencia en los siguientes ejemplos.*

1. *Los vectores $(1,0,0,1),(2,-2,2,-4)$ y $(8,-6,6,-10)$ de $\mathbb{R}^4$ son linealmente dependientes. En efecto, si $\alpha,\beta,\gamma\in\mathbb{R}$ son tales que*

$$\alpha(1,0,0,1)+\beta(2,-2,2,-4)+\gamma(8,-6,6,-10)=\mathbf{0}$$

entonces obtenemos el sistema de ecuaciones

$$\begin{aligned}\alpha+2\beta+8\gamma&=0\\2\beta+6\gamma&=0\\\alpha-4\beta-10\gamma&=0\end{aligned}$$

que tiene solución $(-2\gamma,-3\gamma,\gamma)$, donde $\gamma\in\mathbb{R}$. Así, por ejemplo, si tomamos $\gamma=1$, entonces

$$-2(1,0,0,1)-3(2,-2,2,-4)+(8,-6,6,-10)=\mathbf{0}$$

2. *Los vectores* $u = (1,1)$ *y* $v = (0,2)$ *de* $\mathbb{R}^2$ *son linealmente independientes. En efecto, supongamos que* $\alpha, \beta \in \mathbb{R}$ *son tales que*

$$\alpha\,(1,1) + \beta\,(0,2) = (0,0)\,.$$

Esto implica $(\alpha, \alpha + 2\beta) = (0,0)$ *y, en consecuencia,* $\alpha = \beta = 0$*.*

3. *Sean* $e_1 = (1,0,0)\,, e_2 = (0,1,0)$ *y* $e_3 = (0,0,1)$ *pertenecientes a* $\mathbb{F}^3$*. Entonces,* e_1, e_2, e_3 *son vectores linealmente independientes de* $\mathbb{F}^3$*. En efecto, si* $\alpha_1, \alpha_1, \alpha_3$ *son escalares tales que*

$$\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0$$

entonces $(\alpha_1, \alpha_2, \alpha_3) = (0,0,0)$ *y, por lo tanto,* $\alpha_1 = \alpha_2 = \alpha_3 = 0$*.*

Más generalmente, los vectores canónicos $e_1, \ldots, e_n$ *son vectores linealmente independientes de* $\mathbb{F}^n$*.*

4. *Los vectores* $1, x, x^2, \ldots, x^k$ *de* $\mathbb{F}\,[x]$ *son linealmente independientes. En efecto, si*

$$\sum_{i=1}^{n} \alpha_i x^i = 0,$$

donde $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$*, entonces es claro que* $\alpha_i = 0$ *para todo* $i = 1, \ldots, n$*.*

5. *Las matrices*

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \; y \; \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

de $\mathcal{M}_2\,(\mathbb{F})$ *son linealmente independientes.*

Supongamos que $\alpha_1, \alpha_1, \alpha_3, \alpha_4 \in \mathbb{F}$ *satisfacen*

$$\alpha_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Entonces

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

y, en consecuencia, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$*.*

Vamos a extender el concepto de independencia lineal dada en la definición 3.3.1 para un subconjunto (no necesariamente finito) de un espacio vectorial.

Definición 3.3.3 *Sea* V *un espacio vectorial. Un subconjunto* S *de* V *se dice que es linealmente independiente si todo subconjunto finito de* S *es linealmente independiente.*

Ejemplo 3.3.4 *Los vectores* $1, x, x^2, \ldots, x^k, \ldots$ *de* $\mathbb{F}\,[x]$ *son linealmente independientes.*

Ahora introducimos uno de los conceptos más importantes del álgebra lineal.

Definición 3.3.5 *Sea V un espacio vectorial. Un subconjunto S de V se dice que es una* ***base*** *de V si S es linealmente independiente y genera a V.*

Ejemplo 3.3.6 *Veamos algunos ejemplos de bases de espacios vectoriales.*

1. *Se deduce de los ejemplos anteriores que los vectores canónicos $e_1, \ldots, e_n$ de $\mathbb{F}^n$ forman una base para $\mathbb{F}^n$. La llamamos a partir de ahora la base canónica de $\mathbb{F}^n$.*

2. *Los vectores $u = (1,1)$ y $v = (0,2)$ forman una base para $\mathbb{R}^2$. En efecto, ya vimos en un ejemplo anterior que son linealmente independientes. Por otra parte, si $z = (z_1, z_2) \in \mathbb{R}^2$, entonces buscamos escalares $\lambda, \mu \in \mathbb{R}$ tales que $z = \lambda u + \mu v$. Esto ocurre para $\lambda = z_1$ y $\mu = \frac{z_2 - z_1}{2}$. Es decir, $\mathbb{R}^2 = gen\,(u, v)$.*

3. *Los vectores $u = (1,0,0,1)$,$v = (2,-2,2,-4)$ forman una base para $\mathcal{F}(A)$, donde*
$$A \in \mathcal{M}_4(\mathbb{R})$$
es la matriz definida al comienzo de la sección. En efecto, ya sabemos que generan a $\mathcal{F}(A)$. Sólo falta demostrar que
$$u = (1,0,0,1)\,, v = (2,-2,2,-4)$$
son linealmente independientes. Si $\alpha, \beta \in \mathbb{R}$ son tales que $\alpha u + \beta v = 0$ entonces
$$(\alpha + 2\beta, -2\beta, 2\beta, \alpha - 4\beta) = (0,0,0,0)$$
lo que implica $\alpha = \beta = 0$.

4. *Las matrices $\{E(i,j) : 1 \leq i \leq m \text{ y } 1 \leq j \leq n\}$, donde $E(i,j) \in \mathcal{M}_{m,n}(\mathbb{F})$ está definida como 1 en la posición ij y 0 en el resto, forman una base para $\mathcal{M}_{m,n}(\mathbb{F})$.*

5. *Los vectores $1, x, x^2, \ldots, x^k$ de $\mathbb{F}[x]$ forman una base para el espacio*
$$S = \{p(x) \in \mathbb{F}[x] : grad\,(p(x)) \leq k\}$$
Veamos ahora un ejemplo de un espacio vectorial que no tiene una base finita.

6. *Se desprende del ejemplo 3.2.8 que $\mathbb{F}[x]$ no tiene una base finita. Sin embargo, el conjunto (infinito) de vectores $1, x, x^2, \ldots, x^k, \ldots$ es una base de $\mathbb{F}[x]$.*

Dos preguntas surgen naturalmente: ¿tendrá todo espacio vectorial una base? Por otra parte, es posible que existan bases diferentes para un espacio vectorial. Por ejemplo, $\{(1,1), (0,2)\}$ y $\{(1,0), (0,1)\}$ son bases diferentes de $\mathbb{R}^2$. Ahora, ¿existirá alguna relación entre dos bases de un espacio vectorial? Intentamos dar respuesta a estas preguntas en los siguientes resultados.

Teorema 3.3.7 *Sea V un espacio vectorial generado por los vectores $v_1, \ldots, v_q$. Entonces todo subconjunto de vectores linealmente independiente de V es finito y tiene a lo sumo q vectores.*

Demostración. Es suficiente demostrar que todo subconjunto de V que tenga más de q vectores es linealmente dependiente. Supongamos que $w_1, \ldots, w_p$ son vectores de V, donde $p > q$. Como $V = gen\,(v_1, \ldots, v_q)$ entonces para cada $j = 1, \ldots, p$ existen escalares a_{ij} tales que $w_j = \sum\limits_{i=1}^{q} a_{ij} v_i$. Sea $A \in \mathcal{M}_{q\times p}(\mathbb{F})$ la matriz tal que $[A]_{ij} = a_{ij}$. Como $p > q$ deducimos del corolario 1.3.8 que la ecuación $Ax = 0$ tiene una solución no trivial $x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$. Es decir, $\sum\limits_{j=1}^{p} a_{ij} x_j = 0$ para todo $i = 1, \ldots, q$. Luego

$$\begin{aligned} \sum_{j=1}^{p} x_j w_j &= \sum_{j=1}^{p} x_j \left(\sum_{i=1}^{q} a_{ij} v_i \right) \\ &= \sum_{i=1}^{q} \left(\sum_{j=1}^{p} a_{ij} x_j \right) v_i = 0 \end{aligned}$$

Esto demuestra que los vectores $w_1, \ldots, w_p$ de V son linealmente dependientes. ∎

Corolario 3.3.8 *Si V es un espacio vectorial con base $\{v_1, \ldots, v_q\}$, entonces toda base de V contiene exactamente q vectores.*

Demostración. Sea S una base de V. Como $V = gen\,(v_1, \ldots, v_q)$, se deduce del teorema 3.3.7 que S tiene a lo sumo q vectores, digamos $w_1, \ldots, w_p$, donde $p \leq q$. Por otro lado, $V = gen\,(w_1, \ldots, w_p)$ y $v_1, \ldots, v_q$, son linealmente independientes, luego otra vez por el teorema 3.3.7, $q \leq p$. En consecuencia, $p = q$. ∎

El número de vectores en una base de un espacio vectorial es importante.

Definición 3.3.9 *Si V es un espacio vectorial que tiene una base finita, entonces la* ***dimensión*** *de V, denotada por $dim\,(V)$, es el número de vectores en cualquier base de V. Cuando V tiene una base finita decimos que V es un espacio vectorial de* ***dimensión finita****.*

Ejemplo 3.3.10 *Veamos algunos ejemplos de espacios vectoriales de dimensión finita.*

1. *El espacio $\mathbb{F}^n$ tiene dimensión n. En efecto, la base canónica $\{e_1, \ldots, e_n\}$ tiene n vectores.*

2. *Tenemos que $dim\mathcal{F}\,(A) = 2$ para la matriz A al comienzo de la sección.*

3. *Los vectores $1, x, x^2, \ldots, x^k$ de $\mathbb{F}\,[x]$ forman una base para el espacio*

$$S = \{p\,(x) \in \mathbb{F}\,[x] : grad\,(p\,(x)) \leq k\}$$

Por lo tanto, $dim\,(S) = k + 1$.

Ya tenemos respuesta sobre la relación existente entre dos bases en un espacio vectorial de dimensión finita. Nuestro interés ahora es abordar el problema sobre la existencia de bases en un espacio vectorial. En esta dirección apunta el próximo resultado.

Teorema 3.3.11 *Sea W un subespacio de un espacio vectorial finitamente generado V. Todo subconjunto linealmente independiente de W es finito y puede extenderse hasta una base de W.*

Demostración. Por el teorema 3.3.7, todo subconjunto linealmente independiente de W es finito. Sean $w_1, \ldots, w_p$ vectores linealmente independientes de W. Si

$$W = gen\,(w_1, \ldots, w_p)\,,$$

entonces $\{w_1, \ldots, w_p\}$ es la base que buscamos. De lo contrario existe $w_{p+1} \in W$ tal que

$$w_{p+1} \notin gen\,(w_1, \ldots, w_p)\,.$$

Observamos que los vectores $w_1, \ldots, w_p, w_{p+1}$ son linealmente independientes. De hecho, si

$$\beta_1 w_1 + \cdots + \beta_p w_p + \beta_{p+1} w_{p+1} = 0$$

entonces, $\beta_{p+1} = 0$, de lo contrario w_{p+1} pertenecería a $gen\,(w_1, \ldots, w_p)$. Luego

$$\beta_1 w_1 + \cdots + \beta_p w_p = 0$$

y como $w_1, \ldots, w_p$ son linealmente independientes, $\beta_i = 0$ para todo $i = 1, \ldots, p$. Ahora, si

$$W = gen\,(w_1, \ldots, w_p, w_{p+1})\,,$$

entonces $\{w_1, \ldots, w_p, w_{p+1}\}$ es la base que buscamos. Si no, repetimos el argumento para construir un $w_{p+2} \notin gen\,(w_1, \ldots, w_{p+1})$ tal que $w_1, \ldots, w_{p+2}$ son linealmente independientes. Continuando este proceso concluimos por el teorema 3.3.7 que

$$\{w_1, \ldots, w_s\}$$

es una base de W, para algún $p \leq s$. ■

En particular, todo espacio vectorial finitamente generado es de dimensión finita. En realidad es posible demostrar que todo espacio vectorial (no necesariamente finitamente generado) tiene una base y cualesquiera dos bases tienen la misma cardinalidad. El lector interesado puede consultar un libro de álgebra lineal más avanzado para explorar estos hechos a mayor profundidad.

Corolario 3.3.12 *Sea W un subespacio propio de un espacio vectorial de dimensión finita V. Entonces W tiene dimensión finita y $dim\,(W) < dim\,(V)$.*

Demostración. Sea $0 \neq w_1 \in W$. Por el teorema 3.3.11 existe una base $w_1, \ldots, w_p$ de W. Por el teorema 3.3.7, $p \leq dim\,(V)$. Como W es un subespacio propio de V existe un vector $v \in V$ tal que $v \notin W$. Luego $w_1, \ldots, w_p, v$ son vectores linealmente independientes de V y, por el teorema 3.3.11, puede extenderse hasta una base de V. Esto implica que $p < dim\,(V)$. ■

EJERCICIOS

1. ¿Cuáles de los siguientes conjuntos de vectores es linealmente independiente?:

   a) Los vectores $(1, -2, 1), (2, 1, -1)$ y $(7, -4, 1)$ en $\mathbb{R}^3$;

   b) Los polinomios $x^3 - 4x^2 + 2x + 3$, $x^3 + 2x^2 + 4x - 1$ y $2x^3 - x^2 - 3x + 5$ en $\mathbb{R}[x]$;

   c) Las funciones $senx, cosx$ y e^x en $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

2. Si $v_1, \ldots, v_n$ son vectores de un espacio vectorial V linealmente independientes, demuestre que $v_1, v_1 + v_2, v_1 + v_2 + v_3, \ldots, v_1 + v_2 + \cdots + v_n$ son vectores linealmente independientes.

3. Sea $A \in \mathcal{M}_n(\mathbb{F})$ invertible. Si $v_1, \ldots, v_n$ son vectores linealmente independientes de $\mathbb{F}^n$, entonces $Av_1, \ldots, Av_n$ son también vectores linealmente independientes.

4. Si $v_1, \ldots, v_n$ son vectores de un espacio vectorial V linealmente independientes y $v \notin gen\,(v_1, \ldots, v_n)$, entonces demuestre que $v_1 + v, v_2 + v, \ldots, v_n + v$ son vectores linealmente independientes.

5. Demuestre que los vectores $(2, 1, 3), (2, 1, 0)$ y $(1, 5, 5)$ forman una base de $\mathbb{R}^3$.

6. Determine si las matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

   constituyen una base para $\mathcal{M}_2(\mathbb{R})$.

7. Sea $V = \{p(x) \in \mathbb{F}[x] : \text{grado de } p(x) \leq 2\}$. Demuestre que

$$\mathcal{B} = \left\{1, x - 1, x^2 - x\right\} \text{ y } \mathcal{B}' = \left\{3, x, x^2 - 1\right\}$$

   son bases de V. Escriba cada uno de los vectores de $\mathcal{B}$ como combinación lineal de los vectores de $\mathcal{B}'$.

8. Demuestre que $\mathbb{C}$ es un espacio vectorial sobre $\mathbb{R}$ de dimensión 2.

9. Determine si el conjunto

$$\left\{1, 1 + x, x + x^2, x^2 + x^3, \ldots, x^{m-1} + x^m\right\}$$

   es una base para el espacio $V = \{p(x) \in \mathbb{F}[x] : \text{grado de } p(x) \leq m\}$

10. Encuentre una base para los espacios indicados a continuación:

    a) $\mathcal{M}_{m\times n}(\mathbb{F})$;
    b) El subespacio de $\mathcal{M}_n(\mathbb{F})$ formado por las matrices diagonales;
    c) El subespacio de $\mathcal{M}_n(\mathbb{F})$ formado por las matrices triangulares superiores;
    d) El subespacio de $\mathcal{M}_n(\mathbb{F})$ formado por las matrices simétricas.

11. Sean U y W espacios vectoriales de dimensión finita. Demuestre que

    a) $\dim(U\times W)=\dim(U)+\dim(W)$;
    b) Si U es un subespacio de V, entonces el subespacio $\{(u,u):u\in U\}$ de $V\times V$ tiene dimensión igual a $\dim(U)$.

12. Demuestre que n vectores linealmente independientes en un espacio de dimensión n forman una base.

13. Demuestre que n vectores generadores de un espacio de dimensión n forma una base.

3.4. El rango y la nulidad de una matriz

Describimos en esta sección métodos para encontrar una base de los subespacios

$$\mathcal{C}(A),\mathcal{F}(A),Im(A),ker(A)$$

asociados a la matriz $A\in\mathcal{M}_{m\times n}(\mathbb{F})$. Para esto nos apoyamos una vez más en la eliminación de Gauss-Jordan.

Teorema 3.4.1 *Sea E la matriz escalonada reducida de $A\in\mathcal{M}_{m\times n}(\mathbb{F})$. Si*

$$1\leq k_1<k_2<\cdots<k_r\leq n$$

son las columnas que contienen los $1's$ principales de E, entonces $A_{*k_1},\ldots,A_{*k_r}$ forman una base para $\mathcal{C}(A)$.

Demostración. Afirmación 1: las columnas $E_{*k_1},\ldots,E_{*k_r}$ forman una base para el subespacio generado por las columnas de E. En efecto, recordamos que para todo $i=1,\ldots,r$ se cumple que $E_{*k_i}=e_i$, el vector i-ésimo vector canónico de $\mathbb{F}^m$. En consecuencia, son linealmente independientes. Por otro lado, dado un vector columna E_{*j} de E, como $[E]_{ij}=0$ para todo $i>r$, entonces es claro que E_{*j} se escribe como combinación lineal de los vectores e_i.

Afirmación 2: los vectores $A_{*k_1},\ldots,A_{*k_r}$ forman una base para el subespacio generado por las columnas de A. Por el Corolario 1.4.7, existe una matriz $P\in\mathcal{M}_m(\mathbb{F})$ invertible tal que $E=PA$. Luego para cada j,

$$E_{*j}=(PA)_{*j}=PA_{*j}$$

Así

$$\begin{aligned} A_{*j} &= P^{-1}E_{*j} = P^{-1}\sum_{i=1}^{r}\alpha_i E_{*k_i} \\ &= \sum_{i=1}^{r}\alpha_i\left(P^{-1}E_{*k_i}\right) = \sum_{i=1}^{r}\alpha_i A_{*k_i} \end{aligned}$$

Esto demuestra que las columnas $A_{*k_1},\ldots,A_{*k_r}$ generan el subespacio generado por las columnas de A. Por otra parte, si $\beta_1,\ldots,\beta_r$ son escalares tales que $\sum\limits_{i=1}^{r}\beta_i A_{*k_i}=0$, entonces

$$0 = P0 = P\sum_{i=1}^{r}\beta_i A_{*k_i} = P\sum_{i=1}^{r}\beta_i\left(P^{-1}E_{*k_i}\right) = \sum_{i=1}^{r}\beta_i E_{*k_i}$$

Como $E_{*k_1},\ldots,E_{*k_r}$ son linealmente independientes concluimos que $\beta_1=\cdots=\beta_r=0$. ■

Ejemplo 3.4.2 *Encontremos una base del subespacio de* $\mathbb{R}^4$

$$S = gen\left(\begin{pmatrix}2\\6\\0\\2\end{pmatrix},\begin{pmatrix}4\\4\\4\\2\end{pmatrix},\begin{pmatrix}6\\2\\8\\2\end{pmatrix},\begin{pmatrix}2\\2\\2\\2\end{pmatrix},\begin{pmatrix}6\\14\\2\\8\end{pmatrix}\right)$$

Primero formamos la matriz A con las columnas dadas por lo vectores generadores de S :

$$A = \begin{pmatrix}2&4&6&2&6\\6&4&2&2&14\\0&4&8&2&2\\2&2&2&2&8\end{pmatrix}$$

Vimos en el ejemplo 1.3.3 que

$$A \underset{f}{\sim} \begin{pmatrix}1&0&-1&0&2\\0&1&2&0&-1\\0&0&0&1&3\\0&0&0&0&0\end{pmatrix}$$

Por el teorema anterior, E_{*1},E_{*2} y E_{*4} forman una base para el subespacio generado por las columnas de E. Por lo tanto, los vectores

$$\begin{pmatrix}2\\6\\0\\2\end{pmatrix},\begin{pmatrix}4\\4\\4\\2\end{pmatrix},\begin{pmatrix}2\\2\\2\\2\end{pmatrix}$$

forman una base para S.

Corolario 3.4.3 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. *Entonces* $dim\mathcal{C}(A) = rgo(A)$.

Demostración. Es inmediato del teorema 3.4.1 y la definición del rango de una matriz. ■
Ahora pasamos a calcular una base para el subespacio $\mathcal{F}(A)$.

Teorema 3.4.4 *Sean* $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$. *Si* $A \underset{f}{\sim} B$, *entonces* $\mathcal{F}(A) = \mathcal{F}(B)$.

Demostración. Es suficiente demostrar que si B es una matriz que se obtiene a partir de A a través de una operación elemental por fila, entonces $\mathcal{F}(B) = \mathcal{F}(A)$. En efecto, como cada fila B_{i*} de la matriz B es una combinación lineal de las filas $A_{1*}, \ldots, A_{m*}$ de la matriz A, es claro que

$$gen(B_{1*}, \ldots, B_{m*}) \subseteq gen(A_{1*}, \ldots, A_{m*})$$

Invirtiendo la operación elemental por fila, es claro que A también se obtiene a partir de B a través de una operación elemental por fila. Luego, un argumento similar demuestra que

$$gen(A_{1*}, \ldots, A_{m*}) \subseteq gen(B_{1*}, \ldots, B_{m*})$$

Así, $\mathcal{F}(B) = \mathcal{F}(A)$. ■

Corolario 3.4.5 *Sea E la matriz escalonada reducida de $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. Entonces las filas distintas de cero de E forman una base de $\mathcal{F}(A)$. En particular,* $dim\mathcal{F}(A) = rgo(A) = dim\mathcal{C}(A)$.

Demostración. Por el teorema 3.4.4, $\mathcal{F}(A) = \mathcal{F}(E)$. Es claro que las filas distintas de cero $E_{1*}, \ldots, E_{r*}$ de E generan a $\mathcal{F}(E)$. Para ver que son linealmente independientes, supongamos que

$$\alpha_1 E_{1*} + \cdots + \alpha_r E_{r*} = 0,$$

donde $\alpha_1, \ldots, \alpha_r$ son elementos de $\mathbb{F}$. Sean $1 \leq k_1 < k_2 < \cdots < k_r \leq n$ las columnas de E que contienen a los $1's$ principales. Como $E_{*k_i} = e_i$, el vector canónico de $\mathbb{F}^m$, para todo $i = 1, \ldots, r$, se deduce que la coordenada en la posición k_i de $\alpha_1 E_{1*} + \cdots + \alpha_r E_{r*}$ es precisamente α_i. Por lo tanto, $\alpha_i = 0$ para todo i y así, $E_{1*}, \ldots, E_{r*}$ son linealmente independientes. Esto demuestra que $E_{1*}, \ldots, E_{r*}$ forma una base para $\mathcal{F}(E) = \mathcal{F}(A)$.

La última parte es consecuencia del corolario 3.4.3 y del hecho que $rgo(A)$ es precisamente el número de filas distintas de cero de E. ■

Ejemplo 3.4.6 *En el ejemplo 1.3.3 vimos que*

$$A = \begin{pmatrix} 2 & 4 & 6 & 2 & 6 \\ 6 & 4 & 2 & 2 & 14 \\ 0 & 4 & 8 & 2 & 2 \\ 2 & 2 & 2 & 2 & 8 \end{pmatrix} \underset{f}{\sim} \begin{pmatrix} 1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Luego los vectores $(1, 0, -1, 0, 2), (0, 1, 2, 0, -1), (0, 0, 0, 1, 3)$ *es una base para* $\mathcal{F}(A)$.

Como consecuencia de los resultados anteriores llegamos al sorprendente resultado:

Corolario 3.4.7 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$*. Entonces* $rgo(A) = rgo\left(A^{\top}\right)$*.*

Demostración. Claramente $\mathcal{C}(A) = \mathcal{F}\left(A^{\top}\right)$. Luego por los corolarios 3.4.3 y 3.4.5,

$$rgo(A) = dim\mathcal{C}(A) = dim\mathcal{F}\left(A^{\top}\right) = rgo\left(A^{\top}\right)$$

■

Veamos ahora cómo calcular una base para $ker(A)$, donde $A \in \mathcal{M}_{m\times n}(\mathbb{F})$.

Teorema 3.4.8 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ *y supongamos que* $A \underset{f}{\sim} E$*, donde* E *es escalonada reducida. Supongamos que* $1 \leq k_1 < k_2 < \cdots < k_r \leq n$ *son las columnas que contienen a los* $1's$ *principales de* E*. Entonces:*

1. *Si* $r = n$ *entonces* $ker(A) = \{0\}$*;*
2. *Si* $r < n$ *entonces para cada* $j \in J = \{1,\ldots,n\} \setminus \{k_1,\ldots,k_r\}$ *sea* S_j *el vector columna de* $\mathbb{F}^n$

   $$S_j = e_j - \sum_{s=1}^{r} [E]_{sj}\, e_{k_s}$$

   donde e_i *es el vector canónico de* $\mathbb{F}^n$*. Entonces* $\{S_j : j \in J\}$ *es una base de* $ker(A)$*.*

Demostración. 1. Si $r = n$, entonces, por el teorema 1.4.5, $ker(A) = \{0\}$.
2. Sea $X = (x_1,\ldots,x_n)^{\top}$ una solución de la ecuación $EX = 0$. Entonces

$$\sum_{j=1}^{n} [E]_{ij}\, x_j = 0$$

para todo $i = 1,\ldots,r$.

Como $[E]_{ik_s} = \delta_{is}$ (delta de Kronecker), se deduce

$$x_{k_s} = -\sum_{j\in J} [E]_{sj}\, x_j$$

para todo $s = 1,\ldots,r$. Luego

$$\begin{aligned} X &= \sum_{j=1}^{n} x_j e_j = \sum_{j\in J} x_j e_j + \sum_{s=1}^{r} x_{k_s} e_{k_s} \\ &= \sum_{j\in J} x_j e_j - \sum_{s=1}^{r}\left[\sum_{j\in J} [E]_{sj}\, x_j\right] e_{k_s} \\ &= \sum_{j\in J} x_j \left[e_j - \sum_{s=1}^{r} [E]_{sj}\, e_{k_s}\right] = \sum_{j\in J} x_j S_j \end{aligned}$$

En particular, para cada $j \in J$, S_j es la solución obtenida al hacer $x_j = 1$, $x_i = 0$ para los otros i de J y luego calculando $x_{k_r}, \ldots, x_{k_1}$ como en el teorema 1.3.6 (parte 3). Así, $\{S_j : j \in J\}$ es un conjunto generador para $ker\,(A)$. Vamos a ver que también es linealmente independiente. Observamos que $[S_j]_{i1} = \delta_{ij}$ para todo $i, j \in J$. Luego si $\lambda_1, \ldots, \lambda_s$ son escalares tales que $\sum_{j \in J} \lambda_j S_j = 0$, entonces $\lambda_i = \sum_{j \in J} \lambda_j [S_j]_{i1} = 0$ para cada $i \in J$. Esto termina la demostración. ■

Ejemplo 3.4.9 *Encontremos una base del espacio solución del sistema homogéneo*

$$\begin{aligned} x + 2y - w + 3z - 4u &= 0 \\ 2x + 4y - 2w - z + 5u &= 0 \\ 2x + 4y - 2w + 4z - 2u &= 0 \end{aligned}$$

Tenemos que

$$\begin{pmatrix} 1 & 2 & -1 & 3 & -4 \\ 2 & 4 & -2 & -1 & 5 \\ 2 & 4 & -2 & 4 & -2 \end{pmatrix} \underset{f}{\sim} \begin{pmatrix} 1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

En este caso $k_1 = 1$, $k_2 = 4$, $k_3 = 5$, $r = 3$ *y* $J = \{2, 3\}$. *Entonces*

$$\begin{aligned} S_2 &= \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \left[2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right] = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad y \\ S_3 &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \left[-1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right] = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

y así, $\{S_2, S_3\}$ *es una base del espacio solución del sistema.*

Definición 3.4.10 *Sea* $A \in \mathcal{M}_{m \times n}(\mathbb{F})$. *La* ***nulidad*** *de* A*, denotada por* $nul\,(A)$*, se define como*

$$nul\,(A) = dimker\,(A)$$

Ejemplo 3.4.11 *Si* $A = \begin{pmatrix} 1 & 2 & -1 & 3 & -4 \\ 2 & 4 & -2 & -1 & 5 \\ 2 & 4 & -2 & 4 & -2 \end{pmatrix}$, *entonces se deduce del ejemplo 3.4.9 que* $nul\,(A) = 2$.

Ahora podemos demostrar uno de los teoremas importantes del álgebra lineal.

Corolario 3.4.12 *(Teorema de la dimensión) Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. *Entonces*

$$rgo\,(A) + nul\,(A) = n.$$

Demostración. Consecuencia inmediata del teorema 3.4.8. ■

Recordamos ahora que dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, la imagen de A es el subespacio de $\mathbb{F}^m$

$$Im\,(A) = \{Ax : x \in \mathbb{F}^n\}$$

Por el hecho de que

$$Ax = x_1 A_{*1} + x_2 A_{*2} + \cdots + x_n A_{*n}$$

para todo $x = (x_1, \ldots, x_n)^\top \in \mathbb{F}^n$, se deduce que

$$\mathcal{C}\,(A) = Im(A)$$

Finalizamos esta sección extendiendo la lista de condiciones equivalentes dadas en el teorema 1.4.5 para que una matriz sea invertible.

Teorema 3.4.13 *Sea* $A \in \mathcal{M}_n(\mathbb{F})$. *Las siguientes condiciones son equivalentes:*

1. *A es invertible;*
2. *$nul\,(A) = 0$;*
3. *La ecuación $Ax = 0$ tiene única solución $x = 0$;*
4. *Las columnas de A forman una base de $\mathbb{F}^n$;*
5. *Las filas de A forman una base de $\mathbb{F}^n$.*

Demostración. 1. $\Rightarrow$ 2. Si A es invertible, entonces, por el teorema 1.4.5, $rgo\,(A) = n$. Luego, por el Teorema de la dimensión, $nul\,(A) = 0$.

2. $\Rightarrow$ 3. Si $nul\,(A) = 0$, entonces $\{0\} = ker\,(A) = \{x \in \mathbb{F}^n : Ax = 0\}$.

3. $\Rightarrow$ 4. Supongamos que $\alpha_1 A_{*1} + \cdots + \alpha_n A_{*n} = \mathbf{0}$ para escalares $\alpha_1, \ldots, \alpha_n$. Entonces $Ax = 0$, donde $x = (\alpha_1, \ldots, \alpha_n)^\top \in \mathbb{F}^n$. Por hipótesis $x = 0$ y así, $\alpha_1 = \cdots = \alpha_n = 0$. Es decir, $A_{*1}, \ldots, A_{*n}$ son linealmente independientes. Pero n vectores linealmente independientes en $\mathbb{F}^n$ forman una base.

4. $\Rightarrow$ 5. Si las columnas de A forman una base de $\mathbb{F}^n$, entonces $\mathcal{C}\,(A) = \mathbb{F}^n$. Luego por el corolario 3.4.5,

$$dim\mathcal{F}\,(A) = dim\mathcal{C}\,(A) = n$$

y así $\mathcal{F}\,(A) = \mathbb{F}^n$. En consecuencia, los n vectores generadores $A_{1*}, \ldots, A_{n*}$ de $\mathbb{F}^n$ forman una base de $\mathbb{F}^n$.

5. $\Rightarrow$ 1. Si las filas de A forman una base de $\mathbb{F}^n$, entonces $\mathcal{F}\,(A) = \mathbb{F}^n$. Se deduce del corolario 3.4.5 que

$$n = dim\mathcal{F}\,(A) = rgo\,(A)$$

. Luego por el teorema 1.4.5, A es invertible. ■

EJERCICIOS

1. Demuestre que los vectores $(1,0,-i),(1+i,1-i,1)$ y (i,i,i) forman una base para $\mathbb{C}^3$.

2. Encuentre una base para $\mathcal{F}(A)$, una base para $\mathcal{C}(A)$ y una base para $ker(A)$ si
$$A=\begin{pmatrix}1 & 7 & -6 & 2\\ 2 & 8 & -2 & 4\\ 2 & 5 & 3 & 4\end{pmatrix}$$

3. Encuentre una base para el espacio solución del sistema homogéneo
$$\begin{aligned}6x-3y+4w-3z &= 0\\ 3x+2w-3u &= 0\\ 3x-y+2w-w-u &= 0\end{aligned}$$

4. Demuestre que $rgo(AB)\leq rgo(A)$ y $rgo(AB)\leq rgo(B)$.

3.5. Coordenadas en un espacio vectorial

Las coordenadas de un vector $x=(x_1,\ldots,x_n)^{\top}\in\mathbb{F}^n$ satisfacen la siguiente relación con respecto a la base canónica $e_1,\ldots,e_n$ de $\mathbb{F}^n$:

$$x=x_1e_1+\cdots+x_ne_n$$

Vamos a ver en esta sección que es posible definir las coordenas de un vector de un espacio vectorial de dimensión finita con respecto a una base dada.

Teorema 3.5.1 *Sea V un espacio vectorial con base $v_1,\ldots,v_n$. Entonces para cada vector $v\in V$ existen escalares únicos $\alpha_1,\ldots,\alpha_n$ tales que $v=\alpha_1v_1+\cdots+\alpha_nv_n$.*

Demostración. Como los vectores $v_1,\ldots,v_n$ forman una base de V, todo elemento $v\in V$ se puede escribir como combinación lineal de $v_1,\ldots,v_n$. Supongamos que

$$v=\alpha_1v_1+\cdots+\alpha_nv_n=\beta_1v_1+\cdots+\beta_nv_n$$

Entonces $(\alpha_1-\beta_1)v_1+\cdots+(\alpha_n-\beta_n)v_n=0$ y como los vectores $v_1,\ldots,v_n$ son linealmente independiente, concluimos que $\alpha_i-\beta_i=0$ para todo $i=1,\ldots,n$. ■

La unicidad de los escalares $\alpha_1,\ldots,\alpha_n$ en el teorema anterior nos permite definir las coordenadas del vector $v\in V$. Estas coordenadas dependen de la base que elijamos y también del orden en que aparezcan los vectores de la base.

Definición 3.5.2 *Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ordenada de un espacio vectorial V. Entonces para cada $v \in V$ existen escalares únicos $\alpha_1, \ldots, \alpha_n$ tales que $v = \sum_{i=1}^{n} \alpha_i v_i$. El **vector coordenadas** de v con respecto a $\mathcal{B}$, denotado por $[v]_{\mathcal{B}}$, se define como*

$$[v]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)^{\top} \in \mathbb{F}^n$$

α_i se llama la coordenada i-ésima de v, para cada $i = 1, \ldots, n$.

Ejemplo 3.5.3 *Consideremos la base ordenada*

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$$

de $\mathbb{R}^2$. Entonces para encontrar $\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]_{\mathcal{B}}$ expresamos al vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ como combinación lineal de los vectores $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}$:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \left(-\frac{1}{2} \right) \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Por lo tanto, $\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}$. Similarmente, $\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}$ porque

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \left(\frac{1}{2} \right) \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Ejemplo 3.5.4 *Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ordenada del espacio vectorial V. Entonces para cada $i = 1, \ldots, n$*

$$[v_i]_{\mathcal{B}} = e_i$$

donde $e_i = \Big(0, \ldots, 0, \underset{i}{1}, 0, \ldots, 0\Big)^{\top}$ es el vector canónico de $\mathbb{F}^n$. En efecto,

$$v_i = 0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_n$$

Teorema 3.5.5 *Sea V un espacio vectorial de dimensión finita con base ordenada $\mathcal{B} = \{v_1, \ldots, v_n\}$. Entonces la función $V \longrightarrow \mathbb{F}^n$ dada por $v \rightsquigarrow [v]_{\mathcal{B}}$, para cada $v \in V$, es biyectiva. Además,*

$$[v + w]_{\mathcal{B}} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}$$

y

$$[\alpha v]_{\mathcal{B}} = \alpha \, [v]_{\mathcal{B}}$$

para todo $v, w \in V$ y $\alpha \in \mathbb{F}$.

Demostración. La función es claramente biyectiva. Supongamos que

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

y

$$w = \sum_{i=1}^{n} \beta_i v_i.$$

Entonces

$$\begin{aligned} [v+w]_{\mathcal{B}} &= \left[\sum_{i=1}^{n} (\alpha_i + \beta_i) v_i\right]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}} \end{aligned}$$

y

$$[\alpha v]_{\mathcal{B}} = \sum_{i=1}^{n} (\alpha \alpha_i) v_i = \begin{pmatrix} \alpha\alpha_1 \\ \vdots \\ \alpha\alpha_n \end{pmatrix} = \alpha \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha [v]_{\mathcal{B}}$$

■

Nos planteamos ahora la siguiente pregunta: sea V un espacio vectorial de dimensión finita con bases ordenadas $\mathcal{B}$ y $\mathcal{B}'$, y $v \in V$. ¿Existe alguna relación entre $[v]_{\mathcal{B}}$ y $[v]_{\mathcal{B}'}$?

Teorema 3.5.6 *Sean $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{B}' = \{w_1, \ldots, w_n\}$ bases ordenadas del espacio vectorial V. La matriz $P \in \mathcal{M}_n(\mathbb{F})$ que tiene como columna i al vector columna $[v_i]_{\mathcal{B}'}$ es una matriz invertible que satisface*

$$[v]_{\mathcal{B}'} = P [v]_{\mathcal{B}}$$

para todo $v \in V$. Además, P es única con esta propiedad.

Demostración. Sea $v \in V$. Entonces existen escalares $\alpha_1, \ldots, \alpha_n$ tales que $v = \sum\limits_{i=1}^{n} \alpha_i v_i$. Luego por el teorema 3.5.5,

$$[v]_{\mathcal{B}'} = \left[\sum_{i=1}^{n} \alpha_i v_i\right]_{\mathcal{B}'} = \sum_{i=1}^{n} \alpha_i [v_i]_{\mathcal{B}'} = \sum_{i=1}^{n} \alpha_i P e_i = P \sum_{i=1}^{n} \alpha_i e_i = P [v]_{\mathcal{B}}$$

Por otra parte observamos que por el teorema 3.5.5, como $\{v_1, \ldots, v_n\}$ es una base de V, entonces las columnas de P forman una base de $\mathbb{F}^n$ (ver el ejercicio 1 de esta sección). Luego aplicamos el teorema 3.4.13 para concluir que P es invertible. Finalmente, supongamos que $R \in \mathcal{M}_n(\mathbb{F})$ satisface

$$[v]_{\mathcal{B}'} = R\,[v]_{\mathcal{B}}$$

para todo $v \in V$. Entonces, en particular,

$$Pe_i = P\,[v_i]_{\mathcal{B}} = R\,[v_i]_{\mathcal{B}} = Re_i$$

para todo i y así, $P = R$. ■

Definición 3.5.7 *Sean $\mathcal{B}$ y $\mathcal{B}'$ bases ordenadas de un espacio vectorial V. La matriz P del teorema anterior recibe el nombre de* ***matriz cambio de base*** *de $\mathcal{B}$ a $\mathcal{B}'$.*

Ejemplo 3.5.8 *Consideremos las bases*

$$\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \text{ y } \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$$

de $\mathbb{R}^2$. Entonces la matriz cambio de base de $\mathcal{E}$ a $\mathcal{B}$ tiene como primera columna a $\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_{\mathcal{B}}$ y como segunda columna a $\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]_{\mathcal{B}}$. Se deduce del ejemplo 3.5.3 que la matriz cambio de base de $\mathcal{E}$ a $\mathcal{B}$ es

$$P = \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

EJERCICIOS

1. Sea V un espacio vectorial sobre $\mathbb{F}$ con bases ordenadas $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{B}'$. Demuestre que $\{[v_1]_{\mathcal{B}'}, \ldots, [v_n]_{\mathcal{B}'}\}$ es una base para $\mathbb{F}^n$.

2. Sea V el subespacio de $\mathbb{R}^3$ que tiene como base ordenada a

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Encuentre $\left[\begin{pmatrix} 7 \\ -2 \\ 3 \end{pmatrix}\right]_{\mathcal{B}}$.

3. Sea S el subespacio de $\mathbb{F}[x]$

$$S = \{p(x) \in \mathbb{F}[x] : grad\,(p(x)) \leq 4\}$$

Encuentre las coordenadas de cualquier polinomio de S con respecto a la base ordenada

$$\mathcal{B} = \left\{1, x+1, (x+1)^2, (x+1)^3, (x+1)^4\right\}$$

4. Dadas las bases de $\mathbb{R}^2$

$$\mathcal{E}=\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\} \text{ y } \mathcal{B}=\left\{\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}2\\5\end{pmatrix}\right\}$$

   Encuentre la matriz cambio de base P de $\mathcal{E}$ a $\mathcal{B}$, la matriz cambio de base Q de $\mathcal{B}$ a $\mathcal{E}$ y compruebe que $Q=P^{-1}$.

5. Considere las bases de $\mathbb{R}^3$

$$\mathcal{E}=\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right\} \text{ y } \mathcal{B}=\left\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$$

   Encuentre la matriz cambio de base P de $\mathcal{E}$ a $\mathcal{B}$, la matriz cambio de base Q de $\mathcal{B}$ a $\mathcal{E}$ y compruebe que $Q=P^{-1}$.

6. Considere las bases $\mathcal{B}=\{1,i\}$ y $\mathcal{C}=\{1+i,1+2i\}$ del espacio vectorial $\mathbb{C}$ sobre $\mathbb{R}$.

   a) Encuentre las matrices cambio de base P y Q de $\mathcal{B}$ a $\mathcal{C}$ y de $\mathcal{C}$ a $\mathcal{B}$ respectivamente;

   b) Compruebe que $Q=P^{-1}$.

7. Sea V un espacio vectorial con bases $\mathcal{B}$ a $\mathcal{E}$. Si P es la matriz cambio de base de $\mathcal{E}$ a $\mathcal{B}$ y Q es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{E}$ demuestre que $Q=P^{-1}$.

8. Suponga que $\mathcal{B}$ y $\mathcal{B}'$ son bases ordenadas de un espacio vectorial V de dimensión finita. Demuestre que si $[v]_{\mathcal{B}}=[v]_{\mathcal{B}'}$ para todo $v\in V$, entonces $\mathcal{B}=\mathcal{B}'$.

CAPÍTULO 4

TRANSFORMACIONES LINEALES

4.1. Transformaciones lineales

Los conjuntos $\mathbb{R}\times\{0\}\subseteq\mathbb{R}^2$ y $\mathbb{R}$ son evidentemente distintos; el primero está formado por pares ordenados de $\mathbb{R}^2$ que tienen la segunda coordenada 0, mientras que el segundo consiste de números reales. Sin embargo, como espacios vectoriales sobre $\mathbb{R}$, son esencialmente los mismos. De hecho, es posible establecer una correspondencia biyectiva entre estos dos conjuntos de manera que las operaciones de espacio vectorial sean preservadas. Cuando esto ocurre decimos que los espacios vectoriales son isomorfos. Precisamos mejor esta idea a lo largo de este capítulo.

Sea $A\in\mathcal{M}_{m\times n}(\mathbb{F})$. Por el teorema 1.2.8, la función $\tau_A:\mathbb{F}^n\to\mathbb{F}^m$ definida por $\tau_A(x)=Ax$ $(x\in\mathbb{F}^n)$, satisface las propiedades:

$$\begin{aligned}\tau_A(x+y) &= \tau_A(x)+\tau_A(y)\\ \tau_A(\alpha x) &= \alpha\tau_A(x)\end{aligned}$$

para todo $\alpha\in\mathbb{F}$ y $x,y\in\mathbb{F}^n$. Más generalmente tenemos el siguiente concepto fundamental del álgebra lineal:

Definición 4.1.1 *Sean V y W espacios vectoriales sobre $\mathbb{F}$. Una **transformación lineal** T de V en W es una función $T:V\to W$ que satisface*

$$\begin{aligned}T(x+y) &= T(x)+T(y)\\ T(\alpha x) &= \alpha T(x)\end{aligned}$$

*para todo $\alpha\in\mathbb{F}$ y $x,y\in V$. Denotamos por $\mathcal{L}(V,W)$ al conjunto de todas las transformaciones lineales de V en W. Cuando $V=W$ escribimos $\mathcal{L}(V,V)=\mathcal{L}(V)$ y una transformación lineal de $\mathcal{L}(V)$ la llamamos un **operador lineal** sobre V.*

Un ejemplo trivial de transformación lineal entre espacios vectoriales V y W es la transformación nula, denotada por $\mathbf{0}:V \to W$ y definida por $\mathbf{0}(v) = 0$ para todo $v \in V$.

Ejemplo 4.1.2 *Como vimos antes, si $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, entonces $\tau_A : \mathbb{F}^n \to \mathbb{F}^m$ es una transformación lineal. Consideremos algunas matrices en $\mathcal{M}_2(\mathbb{R})$:*

1. *Reflexión respecto al eje x: Sea $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Entonces $\tau_A : \mathbb{R}^2 \to \mathbb{R}^2$ está definida por $\tau_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$. Geométricamente, τ_A toma un vector de $\mathbb{R}^2$ y lo refleja respecto al eje x.*

2. *Proyección sobre el eje x: Si $P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, entonces $\tau_P : \mathbb{R}^2 \to \mathbb{R}^2$ está definida por $\tau_P \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$. Esto es, τ_P toma un vector de $\mathbb{R}^2$ y lo proyecta sobre el eje x.*

3. *Rotación: sea $0 \leq \theta \leq 2\pi$ y consideremos la matriz*

$$U = \begin{pmatrix} cos\theta & -sen\theta \\ sen\theta & cos\theta \end{pmatrix}.$$

Si $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ entonces $x = rcos\alpha$ y $y = rsen\alpha$. Así,

$$U\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cos\theta & -sen\theta \\ sen\theta & cos\theta \end{pmatrix}\begin{pmatrix} rcos\alpha \\ rsen\alpha \end{pmatrix} = \begin{pmatrix} rcos(\theta+\alpha) \\ rsen(\theta+\alpha) \end{pmatrix}$$

En otras palabras, la transformación lineal $\tau_U : \mathbb{R}^2 \to \mathbb{R}^2$ toma un vector de $\mathbb{R}^2$ y lo rota un ángulo θ en sentido contrario a las agujas del reloj.

Ejemplo 4.1.3 *Las siguientes funciones son ejemplos de transformaciones lineales entre espacios vectoriales:*

1. *Sea $\mathcal{D}$ el subespacio de $\mathcal{F}(\mathbb{R},\mathbb{R})$ formado por las funciones $f : \mathbb{R} \to \mathbb{R}$ que tienen derivadas de todos los órdenes y $D : \mathcal{D} \to \mathcal{D}$ definida por $D(f) = f'$. Para cualesquiera $f, g \in \mathcal{F}(\mathbb{R},\mathbb{R})$ y $\alpha \in \mathbb{R}$ tenemos que*

$$\begin{aligned} D(f+g) &= (f+g)' = f' + g' = D(f) + D(g) \\ D(\alpha f) &= (\alpha f)' = \alpha f' = \alpha D(f) \end{aligned}$$

2. *Sea $S = \{f \in \mathcal{F}([a,b],\mathbb{R}) : f \text{ es continua}\}$ e $\mathcal{I} : S \to \mathbb{R}$ la función definida por*

$$\mathcal{I}(f) = \int_a^b f(x)\,dx.$$

Si $f,g\in S$ y $\alpha\in\mathbb{R}$, entonces

$$\begin{aligned}\mathcal{I}(f+g) &= \int_a^b (f+g)(x)\,dx = \int_a^b [f(x)+g(x)]\,dx \\ &= \int_a^b f(x)\,dx + \int_a^b g(x)\,dx = \mathcal{I}(f)+\mathcal{I}(g)\end{aligned}$$

y

$$\mathcal{I}(\alpha f) = \int_a^b (\alpha f)(x)\,dx = \int_a^b \alpha f(x)\,dx = \alpha\int_a^b f(x)\,dx = \alpha\mathcal{I}(f)$$

Vamos a establecer algunos hechos simples sobre las transformaciones lineales.

Teorema 4.1.4 *Sea $T\in\mathcal{L}(V,W)$. Entonces se cumplen las siguientes condiciones:*

1. $T(0)=0$;
2. $T(-v)=-T(v)=(-1)T(v)$.

Demostración. Demostramos la parte 1, la segunda la dejamos como ejercicio. Tenemos que

$$T(0)=T(0+0)=T(0)+T(0)$$

Luego sumando el inverso aditivo de $T(0)$ a ambos lados de la relación anterior obtenemos $T(0)=0$. ■

El siguiente resultado nos proporciona muchos ejemplos de transformaciones lineales. Además, establece que una transformación lineal $T\in\mathcal{L}(V,W)$ queda completamente determinada por los valores que toma en una base de V.

Teorema 4.1.5 *Sean V,W espacios vectoriales y $v_1,\ldots,v_n$ una base ordenada de V. Consideremos elementos $w_1,\ldots,w_n\in W$. Entonces existe una única $T\in\mathcal{L}(V,W)$ tal que $T(v_i)=w_i$ para todo i.*

Demostración. Definamos la función $T:V\to W$ por $T(v_i)=w_i$ para cada i y extendamos linealmente a V por la fórmula

$$T\left(\sum_{i=1}^n \alpha_i v_i\right)=\sum_{i=1}^n \alpha_i T(v_i).$$

T está bien definida por el hecho de que cada vector de V tiene una representación única como combinación lineal de vectores de la base. La linealidad de T es clara por la forma que la definimos. Para ver la unicidad, supongamos que $S \in \mathcal{L}(V,W)$ satisface $S(v_i) = w_i$ para todo i. Entonces

$$S\left(\sum_{i=1}^{n}\alpha_i v_i\right) = \sum_{i=1}^{n}\alpha_i S(v_i) = \sum_{i=1}^{n}\alpha_i T(v_i) = T\left(\sum_{i=1}^{n}\alpha_i v_i\right)$$

∎

Ejemplo 4.1.6 *Consideremos el espacio vectorial*

$$V = \{p(x) \in \mathbb{F}[x] : grad(p(x)) \leq k\}$$

y sean $\alpha_0, \ldots, \alpha_k$ elementos de $\mathbb{F}$. Entonces existe una única transformación lineal $T : V \to \mathbb{F}$ tal que por $T(x^i) = \alpha_i$, para todo $i = 0, \ldots, k$. De hecho, está definida por

$$T\left(\sum_{i=0}^{k}\beta_i x^i\right) = \sum_{i=0}^{k}\beta_i\alpha_i.$$

Dotamos a continuación al conjunto de las transformaciones lineales $\mathcal{L}(V,W)$ de una estructura de espacio vectorial.

Definición 4.1.7 *Dadas $S,T \in \mathcal{L}(V,W)$ definimos la suma de S y T, denotada por $S+T$, a la función $S+T : V \to W$ definida por*

$$(S+T)(x) = S(x) + T(x)$$

para todo $x \in V$. Si $\alpha \in \mathbb{F}$ entonces la multiplicación del escalar α por $T \in \mathcal{L}(V,W)$, denotado por αT, es la función $\alpha T : V \to W$ definida por

$$(\alpha T)(x) = \alpha T(x)$$

Teorema 4.1.8 *Sean V y W espacios vectoriales sobre $\mathbb{F}$. Si $S,T \in \mathcal{L}(V,W)$ y $\alpha \in \mathbb{F}$ entonces $S+T$ y αS pertenecen a $\mathcal{L}(V,W)$. Más aun, $\mathcal{L}(V,W)$ es un espacio vectorial sobre $\mathbb{F}$ con estas operaciones.*

Demostración. Dejamos como ejercicio al lector. ∎

A $\mathcal{L}(V,W)$ lo llamamos el espacio de las transformaciones lineales de V en W.

Consideramos ahora la composición entre transformaciones lineales.

Teorema 4.1.9 *Sean*

$$S \in \mathcal{L}(V,W)$$

y

$$T \in \mathcal{L}(U,V)$$

. Entonces

$$ST \in \mathcal{L}(U,W).$$

Demostración. Sean $\alpha, \beta \in \mathbb{F}$ y $x, y \in U$. Entonces

$$\begin{aligned}(ST)(\alpha x + \beta y) &= S[T(\alpha x + \beta y)] \\ &= S[\alpha T(x) + \beta T(y)] \\ &= \alpha S(T(x)) + \beta S(T(y)) \\ &= \alpha (ST)(x) + \beta (ST)(y)\end{aligned}$$

■

La transformación lineal identidad es la función $I_V : V \to V$ definida por $I(x) = x$ para todo $x \in V$. Es fácil ver que $I_V T = T I_V = T$ para cualquier $T \in \mathcal{L}(V)$.

Al igual que las matrices, la composición de transformaciones lineales no es conmutativa.

Ejemplo 4.1.10 *Sean $S : \mathbb{R}^3 \to \mathbb{R}^3$ y $T : \mathbb{R}^3 \to \mathbb{R}^3$ los operadores lineales definidos por*

$$S\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \text{ y } T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ z \\ 0 \end{pmatrix}.$$

Entonces

$$ST\begin{pmatrix} x \\ y \\ z \end{pmatrix} = S\begin{pmatrix} x \\ z \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ z \\ 0 \end{pmatrix} \text{ y } TS\begin{pmatrix} x \\ y \\ z \end{pmatrix} = T\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}$$

Usamos la notación de potencias para representar la composición de un operador lineal consigo mismo: si $T \in \mathcal{L}(V)$, entonces $T^0 = I$, y para un entero $n \geq 1$, $T^n = TT^{n-1}$. Las potencias de un operador lineal sí conmutan, i.e.,

$$T^r T^s = T^s T^r = T^{r+s}.$$

En el próximo resultado presentamos una lista de propiedades básicas que satisface la composición de transformaciones lineales. Suponemos que los espacios de partida y de llegada son los adecuados.

Teorema 4.1.11 *Sean R, S y T transformaciones lineales. Entonces*

1. $(RS)T = R(ST)$;

2. $R(S+T) = RS + RT$;

3. $(R+S)T = RT + ST$;

4. $\lambda(RS) = (\lambda R)S = R(\lambda S)$ *para todo* $\lambda \in \mathbb{F}$.

Demostración. La dejamos como ejercicio para el lector. ■

EJERCICIOS

1. Sea $U \in \mathcal{M}_n(\mathbb{F})$ invertible y considere la función $T : \mathcal{M}_n(\mathbb{F}) \to \mathcal{M}_n(\mathbb{F})$ definida por $T(A) = U^{-1}AU$, para todo $A \in \mathcal{M}_n(\mathbb{F})$. ¿Es T un operador lineal? ¿Es inyectivo? ¿Es sobreyectivo?

2. Sea $M \in \mathcal{M}_n(\mathbb{F})$ y defina la función $T : \mathcal{M}_n(\mathbb{F}) \to \mathcal{M}_n(\mathbb{F})$ por $T(A) = AM - MA$. Demuestre que T es un operador lineal.

3. Sea $F : \mathbb{R}^2 \to \mathbb{R}$ definida por $F\begin{pmatrix} x \\ y \end{pmatrix} = xy$. ¿Es F una transformación lineal?

4. Encuentre un operador lineal $R : \mathbb{R}^2 \to \mathbb{R}^2$ que tome un vector de $\mathbb{R}^2$ y:

   a) lo refleje respecto a la recta $y = x$;

   b) lo refleje respecto al eje y;

   c) lo proyecte sobre el eje y.

5. Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ el único operador lineal que satisface
$$F\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \text{ y } F\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}.$$
   Encuentre una expresión explícita para $F\begin{pmatrix} x \\ y \end{pmatrix}$, donde $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$.

6. ¿Existe una transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que
$$T\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ y } T\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}?$$

7. ¿Existe un operador lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que
$$T\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, T\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ y } T\begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}?$$

8. Sea V el espacio de los números complejos considerado como espacio vectorial sobre $\mathbb{R}$. Encuentre una transformación lineal de V en V pero que no sea lineal cuando considere a V como espacio vectorial sobre $\mathbb{C}$.

9. Sea $F : V \to W$ una transformación lineal y suponga que $v_1, \ldots, v_n \in V$ son tales que $F(v_1), \ldots, F(v_n)$ son linealmente independientes. Demuestre que los vectores $v_1, \ldots, v_n$ también son linealmente independientes.

10. Demuestre el teorema 4.1.8.

11. Demuestre el teorema 4.1.11.

12. Considere las transformaciones lineales $F : \mathbb{R}^3 \to \mathbb{R}^2, G : \mathbb{R}^3 \to \mathbb{R}^2$ y $H : \mathbb{R}^3 \to \mathbb{R}^2$ definidas por

$$\begin{aligned} F\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} x+y+z \\ x+y \end{pmatrix}, \ G\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+z \\ x+z \end{pmatrix} \text{ y} \\ H\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 2y \\ x \end{pmatrix} \end{aligned}$$

Demuestre que F, G y H son linealmente independientes como vectores de $\mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$.

13. Encuentre dos operadores lineales S, T en $\mathbb{R}^2$ tales que $ST = 0$ pero $TS \neq 0$.

14. Sea $P \in \mathcal{L}(V)$ tal que $P^2 = P$. Demuestre que $V = \ker(P) \oplus Im(P)$.

4.2. El núcleo y la imagen

Una propiedad importante de las transformaciones lineales es que envía subespacios en subespacios.

Teorema 4.2.1 *Sea $T \in \mathcal{L}(V, W)$ y U un subespacio de V. Entonces*

$$T(U) = \{T(u) : u \in U\}$$

es un subespacio de W.

Demostración. Sean $x, y \in U$ y $\alpha \in \mathbb{F}$. Entonces

$$T(x) + T(y) = T(x+y) \in T(U)$$

y

$$\alpha T(x) = T(\alpha x) \in T(U)$$

■

En particular, si $T \in \mathcal{L}(V, W)$, entonces $T(V)$, llamado la **imagen** de T y denotado por $Im(T)$, es un subespacio de W. Otro subespacio importante asociado a T es el núcleo de T.

Definición 4.2.2 *Sea $T \in \mathcal{L}(V,W)$. El **núcleo** de T, denotado por $ker(T)$, es el subconjunto de V*

$$ker(T) = \{v \in V : T(v) = 0\}.$$

Observamos que $ker(T)$ es no vacío porque $T(0) = 0$.

Teorema 4.2.3 *Sea $T \in \mathcal{L}(V,W)$. Entonces $ker(T)$ es un subespacio de V.*

Demostración. Sean $x, y \in ker(T)$ y $\alpha \in \mathbb{F}$. Entonces, al ser T una transformación lineal,

$$T(x+y) = T(x) + T(y) = 0 + 0 = 0$$

y

$$T(\alpha x) = \alpha T(x) = \alpha 0 = 0$$

porque $x, y \in ker(T)$. En consecuencia, $x + y$ y $\alpha x \in ker(T)$. ∎

Ejemplo 4.2.4 *Veamos algunos ejemplos.*

1. *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $\tau_A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ la transformación lineal inducida por A. Entonces*

$$\begin{aligned} ker(\tau_A) &= \{x \in \mathbb{F}^n : \tau_A(x) = 0\} \\ &= \{x \in \mathbb{F}^n : Ax = 0\} \end{aligned}$$

   y

$$\begin{aligned} Im(\tau_A) &= \{\tau_A(x) : x \in \mathbb{F}^n\} \\ &= \{Ax : x \in \mathbb{F}^n\} \end{aligned}$$

   Esto es, $ker(\tau_A)$ y $Im(\tau_A)$ coincide con lo que llamamos en secciones anteriores el núcleo y la imagen de A. Para calcular estos subespacios usamos las técnicas estudiadas en el capítulo 1.

2. *Sea $\mathcal{D}$ el subespacio de $\mathcal{F}(\mathbb{R},\mathbb{R})$ formado por las funciones $f : \mathbb{R} \to \mathbb{R}$ que tienen derivadas de todos los órdenes y $D : \mathcal{D} \to \mathcal{D}$ definida por $D(f) = f'$. Entonces*

$$\begin{aligned} ker(D) &= \{f \in \mathcal{D} : D(f) = 0\} \\ &= \{f \in \mathcal{D} : f' = 0\} \\ &= \{f \in \mathcal{D} : f \text{ es constante}\} \end{aligned}$$

Para decidir si una transformación lineal es inyectiva basta con revisar el núcleo.

Teorema 4.2.5 *Sea $T \in \mathcal{L}(V,W)$. Entonces T es inyectiva si, y sólo si, $ker(T) = \{0\}$.*

Demostración. Supongamos que T es inyectiva y sea $x \in ker\,(T)$. Entonces

$$T\,(0) = 0 = T\,(x)\,.$$

De aquí se concluye que $x = 0$ y, por lo tanto, $ker\,(T) = \{0\}$. Recíprocamente, sean $x, y \in V$ y supongamos que $T\,(x) = T\,(y)$. Como T es lineal entonces

$$T\,(x - y) = 0$$

y así, $x - y \in ker\,(T) = \{0\}$. Esto demuestra que $x = y$ y T es inyectiva. ■

El resultado central de esta sección establece una relación entre las dimensiones del núcleo y la imagen de una transformación lineal definida sobre un espacio vectorial de dimensión finita.

Teorema 4.2.6 *Sea $T \in \mathcal{L}\,(V, W)$ y V un espacio vectorial de dimensión finita. Entonces $Im\,(T)$ es de dimensión finita y*

$$dim\,(V) = dim\,[ker\,(T)] + dim\,[Im\,(T)]$$

Demostración. Consideremos una base $w_1, \ldots, w_k$ de $ker\,(T)$. Extendamos este conjunto de vectores a una base $w_1, \ldots, w_k, v_1, \ldots, v_r$ de V. Entonces es claro que $dim\,(V) = k + r$. Faltaría ver que $dim\,[Im\,(T)] = r$. Para eso demostramos que $T\,(v_1)\,, \ldots, T\,(v_r)$ es una base de $Im\,(T)$. Primero, si $y \in Im\,(T)$, entonces $y = T\,(x)$ para algún $x \in V$. Luego existen escalares $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_r$ tales que

$$x = \alpha_1 w_1 + \cdots + \alpha_k w_k + \beta_1 v_1 + \cdots + \beta_r v_r$$

y como $w_i \in ker\,(T)$ para todo i obtenemos

$$\begin{aligned} y &= T\,(x) = T\,(\alpha_1 w_1 + \cdots + \alpha_k w_k + \beta_1 v_1 + \cdots + \beta_r v_r) \\ &= \alpha_1 T\,(w_1) + \cdots + \alpha_k T\,(w_k) + \beta_1 T\,(v_1) + \cdots + \beta_r T\,(v_r) \\ &= \beta_1 T\,(v_1) + \cdots + \beta_r T\,(v_r) \end{aligned}$$

Esto demuestra que $T\,(v_1)\,, \ldots, T\,(v_r)$ generan a $Im\,(T)$. Por otro lado, si $\gamma_1, \ldots, \gamma_r$ son escalares tales que

$$\gamma_1 T\,(v_1) + \cdots + \gamma_r T\,(v_r) = 0$$

entonces

$$0 = \gamma_1 T\,(v_1) + \cdots + \gamma_r T\,(v_r) = T\,(\gamma_1 v_1 + \cdots + \gamma_r v_r)$$

En particular, $\gamma_1 v_1 + \cdots + \gamma_r v_r \in ker\,(T)$ y así, existen escalares $\zeta_1, \ldots, \zeta_k$ tales que

$$\gamma_1 v_1 + \cdots + \gamma_r v_r = \zeta_1 w_1 + \cdots + \zeta_k w_k$$

o equivalentemente,

$$\zeta_1 w_1 + \cdots + \zeta_k w_k - \gamma_1 v_1 - \cdots - \gamma_r v_r = 0$$

Pero $w_1, \ldots, w_k, v_1, \ldots, v_r$ es una base de V, por lo tanto, $\zeta_1 = \cdots = \zeta_k = \gamma_1 = \cdots = \gamma_r = 0$. En consecuencia, $T(v_1), \ldots, T(v_r)$ son vectores linealmente independientes. ■

En particular, si $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $\tau_A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ es la transformación lineal inducida por A, entonces

$$dim(\mathbb{F}^n) = dim(ker(\tau_A)) + dim(Im(\tau_A))$$

o, equivalentemente,

$$n = nul(A) + rgo(A)$$

Este es el teorema de la dimensión para matrices (teorema 3.4.12).

Una consecuencia inmediata del teorema anterior es la siguiente.

Corolario 4.2.7 *Sean V y W espacios de dimensión finita tales que $dim(V) = dim(W)$ y $T \in \mathcal{L}(V, W)$. Entonces T es inyectiva si, y sólo si, T es sobreyectiva.*

Demostración. Por el teorema 4.2.6 sabemos que

$$dim(V) = dim[ker(T)] + dim[Im(T)]$$

Si T es inyectiva entonces $ker(T) = \{0\}$ y así, $dim(W) = dim(V) = dim[Im(T)]$. Se concluye del teorema 3.3.12 que $Im(T) = W$. Recíprocamente, si $Im(T) = W$, entonces $dim[ker(T)] = 0$ y, por lo tanto, $ker(T) = \{0\}$. Esto demuestra que T es inyectiva. ■

EJERCICIOS

1. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y + 2z \\ 2x + y \\ -x - 2y + 2z \end{pmatrix}$. Encuentre una base para $\ker(T)$ y $Im(T)$.

2. Describa explícitamente un operador lineal de $\mathbb{R}^3$ en $\mathbb{R}^3$ que tenga como imagen al subespacio generado por $(1, 0, -1)^\top$ y $(1, 2, 2)^\top$.

3. Sea $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ y considere el operador lineal $T : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ definida por $T(A) = AM - MA$, para todo $A \in \mathcal{M}_2(\mathbb{R})$. Encuentre una base del núcleo de T.

4. Sean $A, B \in \mathcal{M}_n(\mathbb{F})$. Demuestre que si $AB = I_n$ entonces $BA = I_n$.

5. Sea $T : V \to W$ una transformación lineal y S un subespacio de V. Demuestre (a) Si $\dim(S) = 1$ entonces $T(S)$ es un punto o una recta; (b) Si $\dim(S) = 2$ entonces $T(S)$ es un plano, una recta o un punto.

6. Sea $T : V \to W$ una transformación lineal y S un subespacio de V. Demuestre (a) Si S es un recta arbitraria de V, entonces $T(S)$ es una recta o un punto; (b) Si S es un plano arbitrario de V, entonces $T(S)$ es un plano, una recta o un punto.

7. Sea $T : V \to W$ una transformación lineal. Demuestre que si $\dim(V) > \dim(W)$, entonces T no es inyectiva.

8. Sea V el subespacio de $\mathcal{F}(\mathbb{R}, \mathbb{R})$ formado por las funciones que tienen derivadas de todos los órdenes y sea $D^n : V \to V$ el operador derivada n-ésima. ¿Cuál es el núcleo de D^n?

9. ¿Cuál es la dimensión del espacio $S = \{A \in \mathcal{M}_n(\mathbb{F}) : Tr(A) = 0\}$?

10. Demuestre que la función $P : \mathcal{M}_n(\mathbb{F}) \to \mathcal{M}_n(\mathbb{F})$ definida por $P(A) = \frac{1}{2}\left(A + A^\top\right)$ es un operador lineal. Además, $\ker(P)$ consiste en el espacio de las matrices antisimétricas y $Im(P)$ el espacio de las matrices simétricas. ¿Cuál es la dimensión del espacio de las matrices antisimétricas? Encuentre una base para el espacio de las matrices antisimétricas.

11. Suponga que U, W son subespacios de V.

    a) Demuestre que la función $T : U \times W \to V$, definida por $T(u, w) = u - w$, es una transformación lineal.

    b) Demuestre que $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

4.3. Isomorfismos

Definición 4.3.1 *Un **isomorfismo** de V a W es una transformación lineal $T : V \to W$ biyectiva. Si existe un isomorfismo de V a W decimos que V es **isomorfo** a W y escribimos $V \cong W$.*

Dos espacios isomorfos son esencialmente los mismos, lo único que cambia es la naturaleza de sus elementos.

Ejemplo 4.3.2 *Presentamos algunos ejemplos de isomorfismos.*

1. *$A \in \mathcal{M}_n(\mathbb{F})$ es invertible si, y sólo si, $\tau_A : \mathbb{F}^n \to \mathbb{F}^n$ es un isomorfismo. Esto es consecuencia del corolario 4.2.7 y el teorema 3.4.13.*

2. *Sea $\mathbb{F}\times\{0\}$ el subespacio de $\mathbb{F}^2$ definido como*

$$\mathbb{F}\times\{0\} = \{(x, 0) : x \in \mathbb{F}\}$$

   Entonces $\varphi : \mathbb{F} \to \mathbb{F}\times\{0\}$ definido por $\varphi(x) = (x, 0)$ es un isomorfismo. Supongamos que $x, y \in \mathbb{F}$ y $\alpha, \beta \in \mathbb{F}$. Entonces

$$\begin{aligned} \varphi(\alpha x + \beta y) &= (\alpha x + \beta y, 0) = \alpha(x, 0) + \beta(y, 0) \\ &= \alpha\varphi(x) + \beta\varphi(y) \end{aligned}$$

   Así, φ es una transformación lineal. Además se verifica fácilmente que φ es biyectiva.

Si $T \in \mathcal{L}(V, W)$ es un isomorfismo, entonces, en particular, existe la función inversa $T^{-1} : W \to V$ tal que

$$TT^{-1} = I_W \text{ y } T^{-1}T = I_V$$

Teorema 4.3.3 *Si $T \in \mathcal{L}(V, W)$ es un isomorfismo, entonces $T^{-1} \in \mathcal{L}(W, V)$ es un isomorfismo.*

Demostración. Sabemos que T^{-1} es biyectiva. Sean $w, z \in W$ y $\alpha, \beta \in \mathbb{F}$. Como T es sobreyectiva, $w = T(x)$ y $z = T(y)$, donde $x, y \in V$. Luego

$$\begin{aligned} T^{-1}(\alpha w + \beta z) &= T^{-1}(\alpha T(x) + \beta T(y)) = T^{-1}(T(\alpha x + \beta y)) \\ &= \alpha x + \beta y = \alpha T^{-1}(w) + \beta T^{-1}(z) \end{aligned}$$

■

Como consecuencia de este resultado tenemos que la relación " $\cong$ " es simétrica: si V es isomorfo a W, entonces W es isomorfo a V. Por lo tanto, a partir de ahora decimos simplemente que los espacios V y W son isomorfos. También es reflexiva porque la I_V es un isomorfismo para cualquier espacio V, y transitiva por el teorema 4.1.9. Deducimos que la relación " $\cong$ " es una relación de equivalencia sobre el conjunto de los espacios vectoriales sobre $\mathbb{F}$.

Teorema 4.3.4 *Sean $S \in \mathcal{L}(V, W)$ y $T \in \mathcal{L}(U, V)$ isomorfismos. Entonces*

1. $(T^{-1})^{-1} = T$;
2. *ST es un isomorfismo y* $(ST)^{-1} = T^{-1}S^{-1}$;

Demostración. Demostramos 2.

$$\begin{aligned} (ST)\left(T^{-1}S^{-1}\right) &= S\left(TT^{-1}\right)S^{-1} = SI_VS^{-1} = I_W \\ \left(T^{-1}S^{-1}\right)(ST) &= T^{-1}\left(S^{-1}S\right)T = T^{-1}I_VT = I_U \end{aligned}$$

Luego $(TS)^{-1} = (S^{-1}T^{-1})$. ■

Los conceptos de independencia, generación y bases pasan a través del isomorfismo de un espacio a otro.

Teorema 4.3.5 *Supongamos que $T \in \mathcal{L}(V, W)$ es un isomorfismo y X un subconjunto de V. Entonces:*

1. *X es linealmente independiente si, y sólo si, $T(X)$ es linealmente independiente*;
2. *X genera a V si, y sólo si, $T(X)$ genera a W*;
3. *X es una base de V si, y sólo si, $T(X)$ es una base de W.*

Demostración. 1. Supongamos que X es un conjunto linealmente independiente y

$$\sum_{i=1}^{n} \alpha_i T(x_i) = 0,$$

donde los $\alpha_i \in \mathbb{F}$ y los $x_i \in X$. Como T es lineal, entonces $T\left(\sum_{i=1}^{n} \alpha_i x_i\right) = 0$ y en consecuencia, $\sum_{i=1}^{n} \alpha_i x_i \in ker(T) = \{0\}$ porque T es inyectiva. Esto implica que $\alpha_i = 0$ para todo i. Recíprocamente, supongamos que $T(X)$ es linealmente independiente. Sean $\beta_1, \ldots, \beta_k$ escalares tales que $\sum_{i=1}^{n} \beta_i x_i = 0$ para ciertos $x_i \in X$. Entonces $0 = T\left(\sum_{i=1}^{n} \beta_i x_i\right) = \sum_{i=1}^{n} \beta_i T(x_i)$ y así, $\beta_i = 0$ para todo i.

2. Supongamos que X genera a V y sea $w \in W$. Entonces existe $v \in V$ tal que $T(v) = w$ porque T es sobreyectiva. Sean $\gamma_1, \ldots, \gamma_r$ escalares y $x_1, \ldots, x_r \in X$ tales que $v = \sum_{i=1}^{r} \gamma_i x_i$. Luego

$$w = T(v) = T\left(\sum_{i=1}^{r} \gamma_i x_i\right) = \sum_{i=1}^{r} \gamma_i T(x_i)$$

Esto demuestra que $T(X)$ genera a W. Recíprocamente, supongamos que $T(X)$ genera a W y sea $x \in V$. Entonces

$$T(x) = \sum_{i=1}^{r} \zeta_i T(x_i) = T\left(\sum_{i=1}^{r} \zeta_i x_i\right)$$

para ciertos $x_i \in X$ y escalares $\zeta_i \in \mathbb{F}$. Se deduce por la inyectividad de T que $x = \sum_{i=1}^{r} \zeta_i x_i$. Así, X genera a V.

3. Es consecuencia inmediata de las partes 1 y 2. ■

Corolario 4.3.6 *Sean V y W espacios vectoriales de dimensión finita sobre $\mathbb{F}$. Entonces $V \cong W$ si, y sólo si, $dim(V) = dim(W)$.*

Demostración. Si $V \cong W$ existe un isomorfismo $T \in \mathcal{L}(V, W)$. Si X es una base de V, entonces sabemos por el teorema anterior que $T(X)$ es una base de W. Como T es biyectiva, $T(X)$ tiene exactamente el mismo número de vectores que X. Por lo tanto, $dim(V) = dim(W)$.

Recíprocamente, sean $v_1, \ldots, v_n$ y $w_1, \ldots, w_n$ bases ordenadas de V y W respectivamente. Consideremos la (única) transformación lineal $T \in \mathcal{L}(V, W)$ dada por el teorema 4.1.5 tal que $T(v_i) = w_i$. Entonces T es sobreyectiva porque $W = gen(w_1, \ldots, w_n) \subseteq Im(T)$. El resultado se deduce ahora del teorema 4.2.7. ■

Se desprende del corolario anterior que un espacio vectorial V de dimensión n es isomorfo al espacio $\mathbb{F}^n$.

Ejemplo 4.3.7 *De hecho, si $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base ordenada de V, entonces, por el teorema 3.5.5, la función $V \longrightarrow \mathbb{F}^n$ dada por $v \rightsquigarrow [v]_{\mathcal{B}}$, para cada $v \in V$, es un isomorfismo.*

EJERCICIOS

1. Sea $S \in \mathcal{L}(\mathbb{R}^2)$ definida por $S\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ 3x - 5y \end{pmatrix}$. Demuestre que S es invertible.
2. Sea $T \in \mathcal{L}(\mathbb{R}^3)$ definida por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x \\ x - y \\ 2x + y + z \end{pmatrix}$. ¿Es T invertible? De serlo, encuentre T^{-1}.
3. Sea $S \in \mathcal{L}(V)$ tal que $S^r = 0$. Demuestre que $I - S$ es invertible.
4. Suponga que $m > n$ y sean $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ y $U \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Demuestre que UT no es invertible.
5. Sea $T \in \mathcal{L}(V)$, donde V es un espacio vectorial de dimensión finita. Suponga que existe $U \in \mathcal{L}(V)$ tal que $TU = I$. Demuestre que T es invertible y $T^{-1} = U$. Por medio de un ejemplo demuestre que el resultado no es cierto si V no tiene dimensión finita.
6. Suponga que el espacio vectorial V es suma directa de los subespacios U y W. Demuestre que V es isomorfo al espacio producto directo $U \times W$.
7. Sean V y W espacios vectoriales y sea $U \in \mathcal{L}(V, W)$ un isomorfismo. Demuestre que $T \rightsquigarrow UTU^{-1}$ es un isomorfismo de $\mathcal{L}(V)$ en $\mathcal{L}(W)$.

4.4. Matriz asociada a una transformación lineal

En esta sección demostramos que si V y W son espacios vectoriales de dimensión n y m respectivamente, entonces $\mathcal{L}(V, W) \cong \mathcal{M}_{m\times n}(\mathbb{F})$.

Teorema 4.4.1 *Sean V y W espacios vectoriales sobre $\mathbb{F}$ de dimensión n y m respectivamente. Supongamos que $\mathcal{B}$ es una base ordenada de V y $\mathcal{C}$ es una base ordenada W. Para cada $T \in \mathcal{L}(V, W)$ existe una única matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ tal que*

$$[T(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}} \tag{4.1}$$

para todo $v \in V$.

Demostración. Supongamos que $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base ordenada de V. Definamos la matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ que tiene como columna i a $[T(v_i)]_{\mathcal{C}}$, para todo $i = 1, \ldots, n$. Sea $v \in V$

y supongamos que $[v]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)^{\top}$. Entonces

$$\begin{aligned} [T(v)]_{\mathcal{C}} &= \left[T\left(\sum_{i=1}^{n} \alpha_i v_i\right)\right]_{\mathcal{C}} = \left[\sum_{i=1}^{n} \alpha_i T(v_i)\right]_{\mathcal{C}} \\ &= \sum_{i=1}^{n} \alpha_i [T(v_i)]_{\mathcal{C}} = \sum_{i=1}^{n} \alpha_i A e_i = A \sum_{i=1}^{n} \alpha_i e_i \\ &= A [v]_{\mathcal{B}} \end{aligned}$$

Supongamos que $B \in \mathcal{M}_{m \times n}(\mathbb{F})$ satisface

$$[T(v)]_{\mathcal{C}} = B [v]_{\mathcal{B}}$$

para todo $v \in V$. En particular, $Ae_i = A[v_i]_{\mathcal{B}} = B[v_i]_{\mathcal{B}} = Be_i$ para todo $i = 1, \ldots, n$, lo cual indica que A y B tienen todas las columnas iguales. ■

Definición 4.4.2 *La matriz A que asociamos a la transformación lineal $T \in \mathcal{L}(V, W)$ en el teorema 4.4.1 se llama la **matriz de T respecto a las bases $\mathcal{B}$ y $\mathcal{C}$**. La denotamos por $A = [T]_{\mathcal{BC}}$.*

Ejemplo 4.4.3 *Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ el operador lineal definido por*

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ 2x - 5y \end{pmatrix}$$

Vamos a calcular la matriz $[T]_{\mathcal{BB}}$ donde $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$. Para esto calculamos

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \left(-\frac{5}{2}\right)\begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Por otra parte,

$$T\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -10 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-6)\begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

En consecuencia,

$$[T]_{\mathcal{BB}} = \begin{pmatrix} 2 & 2 \\ -\frac{5}{2} & -6 \end{pmatrix}$$

Ejemplo 4.4.4 *Sea $V = \{p(x) \in \mathbb{F}[x] : \text{grado de } p(x) \leq 2\}$. Consideremos las bases*

$$\mathcal{B} = \left\{1, x - 1, x^2 - x\right\}$$

y

$$\mathcal{B}' = \left\{3, x, x^2 - 1\right\}$$

de V y la transformación lineal $D : V \to V$ definida por $D(p(x)) = p'(x)$, la derivada de $p(x)$. Vamos a construir la matriz $[D]_{\mathcal{BB}'}$.

$$\begin{aligned} D(1) &= 0 = 0(3) + 0(x) + 0\left(x^2-1\right) \\ D(x-1) &= 1 = \frac{1}{3}(3) + 0(x) + 0\left(x^2-1\right) \\ D\left(x^2-x\right) &= 2x-1 = -\frac{1}{3}(3) + 2(x) + 0\left(x^2-1\right) \end{aligned}$$

En consecuencia,

$$[D]_{\mathcal{BB}'} = \begin{pmatrix} 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Teorema 4.4.5 *Sean V y W espacios vectoriales sobre $\mathbb{F}$ de dimensión n y m respectivamente. Supongamos que $\mathcal{B}$ es una base ordenada de V y $\mathcal{C}$ es una base ordenada W. La función*

$$[\cdot]_{\mathcal{BC}} : \mathcal{L}(V,W) \longrightarrow \mathcal{M}_{m\times n}(\mathbb{F}),$$

definida por $T \rightsquigarrow [T]_{\mathcal{BC}}$, es un isomorfismo.

Demostración. Sean $S, T \in \mathcal{L}(V,W)$ y $\alpha \in \mathbb{F}$. Entonces la columna i de la matriz asociada a $S+T$ es

$$[(S+T)(v_i)]_{\mathcal{C}} = [S(v_i) + T(v_i)]_{\mathcal{C}} = [S(v_i)]_{\mathcal{C}} + [T(v_i)]_{\mathcal{C}}$$

y la columna i de la matriz asociada a αS es

$$[(\alpha S)(v_i)]_{\mathcal{C}} = [\alpha S(v_i)]_{\mathcal{C}} = \alpha [S(v_i)]_{\mathcal{C}}$$

Luego $[S+T]_{\mathcal{BC}} = [S]_{\mathcal{BC}} + [T]_{\mathcal{BC}}$ y $[\alpha S]_{\mathcal{BC}} = \alpha [S]_{\mathcal{BC}}$. Esto demuestra que la función definida por $T \rightsquigarrow [T]_{\mathcal{BC}}$ es lineal. Además, es biyectiva. En efecto, supongamos que $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{C} = \{w_1, \ldots, w_m\}$. Si $T \in \mathcal{L}(V,W)$ y $[T]_{\mathcal{BC}} = \mathbf{0}$, entonces, por (4.1), $[T(v)]_{\mathcal{C}} = \mathbf{0}$ para todo $v \in V$. Esto implica que $T = \mathbf{0}$ y, por lo tanto, la función es inyectiva. Para ver la sobreyectividad sea $M \in \mathcal{M}_{m\times n}(\mathbb{F})$ y supongamos que $Me_i = (x_{1i}, \ldots, x_{mi})^\top$ para cada $i = 1, \ldots, n$. Definamos $Q : V \to W$ por $Q(v_i) = x_{i1}w_1 + \cdots + x_{im}w_m$ y extendamos linealmente a todo V. Entonces $Q \in \mathcal{L}(V,W)$ y $[Qv_i]_{\mathcal{C}} = Me_i$ para todo $i = 1, \ldots, n$. Por lo tanto, $[Q]_{\mathcal{BC}} = M$. Esto termina la demostración. ■

La matriz asociada a una composición de transformaciones lineales es el producto de las matrices asociadas a cada una de las transformaciones lineales.

Teorema 4.4.6 *Sea $T \in \mathcal{L}(V,W)$ y $S \in \mathcal{L}(W,Z)$. Supongamos que $\mathcal{B}, \mathcal{C}$ y $\mathcal{D}$ son bases de V, W y Z respectivamente. Entonces*

$$[ST]_{\mathcal{BD}} = [S]_{\mathcal{CD}} [T]_{\mathcal{BC}}$$

Demostración. Por la relación (4.1),

$$[S(w)]_{\mathcal{D}} = [S]_{\mathcal{CD}} [w]_{\mathcal{C}}$$

para todo $w \in W$ y

$$[T(v)]_{\mathcal{C}} = [T]_{\mathcal{BC}} [v]_{\mathcal{B}}$$

para todo $v \in V$. Luego para $v \in V$ tenemos que

$$\begin{aligned} [(ST)(v)]_{\mathcal{D}} &= [S(T(v))]_{\mathcal{D}} = [S]_{\mathcal{CD}} [T(v)]_{\mathcal{C}} \\ &= [S]_{\mathcal{CD}} [T]_{\mathcal{BC}} [v]_{\mathcal{B}} \end{aligned}$$

Por la unicidad, $[ST]_{\mathcal{BD}} = [S]_{\mathcal{CD}} [T]_{\mathcal{BC}}$. ■

El isomorfismo dado en el teorema 4.4.5 entre los espacios vectoriales $\mathcal{L}(V, W)$ y $\mathcal{M}_{m\times n}(\mathbb{F})$ depende de las bases elegidas $\mathcal{B}$ y $\mathcal{C}$. Surge de forma natural la siguiente pregunta: si $\mathcal{B}'$ y $\mathcal{C}'$ son otras bases de V y W respectivamente, ¿qué relación existe entre $[T]_{\mathcal{BC}}$ y $[T]_{\mathcal{B}'\mathcal{C}'}$?

Recordamos que si $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{B}'$ son bases ordenadas del espacio vectorial V, la matriz cambio de base de $\mathcal{B}$ a $\mathcal{B}'$ es la matriz invertible $P \in \mathcal{M}_n(\mathbb{F})$ que tiene como columna i al vector columna $[v_i]_{\mathcal{B}'}$.

Teorema 4.4.7 *Sean V, W espacios vectoriales de dimensión finita y $T \in \mathcal{L}(V, W)$. Supongamos que $\mathcal{B}$,$\mathcal{B}'$ son bases ordenadas de V y $\mathcal{C}$,$\mathcal{C}'$ son bases ordenadas de W. Sea P la matriz cambio de base de $\mathcal{B}$ a $\mathcal{B}'$ y Q la matriz cambio de base de $\mathcal{C}$ a $\mathcal{C}'$. Entonces*

$$[T]_{\mathcal{B}'\mathcal{C}'} = Q [T]_{\mathcal{BC}} P^{-1}$$

Demostración. Por los teoremas 3.5.6 y 4.4.1 tenemos que

$$\begin{aligned} \left(Q [T]_{\mathcal{BC}} P^{-1}\right) [v]_{\mathcal{B}'} &= Q [T]_{\mathcal{BC}} [v]_{\mathcal{B}} = Q [T(v)]_{\mathcal{C}} \\ &= [T(v)]_{\mathcal{C}'} \end{aligned}$$

para todo $v \in V$. Por la unicidad del teorema 4.4.1 deducimos que

$$Q [T]_{\mathcal{BC}} P^{-1} = [T]_{\mathcal{B}'\mathcal{C}'}$$

■

Ejemplo 4.4.8 *Sea $U : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la transformación lineal definida por*

$$U\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x-y \\ y \end{pmatrix}$$

Consideremos la base $\mathcal{B} = \left\{(1,1,1)^\top, (1,1,0)^\top, (1,0,0)^\top\right\}$ de $\mathbb{R}^3$ y $\mathcal{C} = \left\{(1,1)^\top, (1,-1)^\top\right\}$ de $\mathbb{R}^2$. Un cálculo demuestra que $[U]_{\mathcal{CB}} = \begin{pmatrix} 1 & -1 \\ -1 & 3 \\ 1 & -1 \end{pmatrix}$ y $[U]_{\mathcal{E}'\mathcal{E}} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$, donde $\mathcal{E}'$

y $\mathcal{E}$ son las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente. Por el teorema anterior sabemos que $[U]_{\mathcal{E}'\mathcal{E}} = Q\,[U]_{\mathcal{CB}}\,P^{-1}$, donde $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ es la matriz cambio de base de $\mathcal{C}$ a $\mathcal{E}'$ y $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{E}$.

Para simplificar la notación, si $T \in \mathcal{L}(V)$ y $\mathcal{B}$ es una base ordenada de V, escribiremos $[T]_{\mathcal{B}}$ en lugar de $[T]_{\mathcal{BB}}$ y diremos que $[T]_{\mathcal{B}}$ es la matriz de T con respecto a $\mathcal{B}$.

Corolario 4.4.9 *Sea V un espacio vectorial de dimensión finita y $T \in \mathcal{L}(V)$. Supongamos que $\mathcal{B}$,$\mathcal{B}'$ son bases ordenadas de V y P es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{B}'$. Entonces*

$$[T]_{\mathcal{B}'} = P\,[T]_{\mathcal{B}}\,P^{-1}$$

Demostración. Esto es consecuencia inmediata del teorema 4.4.7, tomando $V = W$, $\mathcal{B} = \mathcal{C}$ y $\mathcal{B}' = \mathcal{C}'$. ■

Este resultado motiva la siguiente definición.

Definición 4.4.10 *Sean $A, B \in \mathcal{M}_n(\mathbb{F})$. Decimos que A es semejante a B si existe una matriz invertible $Q \in \mathcal{M}_n(\mathbb{F})$ tal que $B = Q^{-1}AQ$.*

La relación de semejanza es una relación de equivalencia sobre el espacio $\mathcal{M}_n(\mathbb{F})$ (ver ejercicio 12) y, por lo tanto, divide a $\mathcal{M}_n(\mathbb{F})$ en clases de equivalencia disjuntas. Se concluye del corolario anterior que las matrices asociadas a un operador lineal con respecto a bases diferentes son matrices semejantes. El recíproco también es cierto, es decir, si $A, B \in \mathcal{M}_n(\mathbb{F})$ son semejantes, entonces A, B son matrices asociadas a un operador lineal $T \in \mathcal{L}(V)$ con respecto a dos bases de V. Ese es nuestro próximo resultado.

Teorema 4.4.11 *Supongamos que $A, B \in \mathcal{M}_n(\mathbb{F})$ son semejantes. Entonces existen bases $\mathcal{B}$ y $\mathcal{B}'$ de un espacio vectorial V de dimensión finita y $T \in \mathcal{L}(V)$ tales que $[T]_{\mathcal{B}} = A$ y $[T]_{\mathcal{B}'} = B$.*

Demostración. Sea V un espacio vectorial de dimensión n y $\mathcal{B}$ una base de V. Por el teorema 4.4.5, existe $T \in \mathcal{L}(V)$ tal que $[T]_{\mathcal{B}} = A$. Sea $Q \in \mathcal{M}_n(\mathbb{F})$ una matriz invertible tal que $B = Q^{-1}AQ$. Veamos que es posible construir una base $\mathcal{B}'$ de V tal que la matriz cambio de base de $\mathcal{B}$ a $\mathcal{B}'$ es $P = Q^{-1}$. En tal caso obtenemos que

$$[T]_{\mathcal{B}'} = P\,[T]_{\mathcal{B}}\,P^{-1} = Q^{-1}AQ = B$$

Recordando que $[\cdot] : V \to \mathbb{F}^n$ es un isomorfismo, consideremos la base $\mathcal{B}' = \{v_1, \ldots, v_n\}$ de V determinada por la relación

$$[v_i]_{\mathcal{B}} = Qe_i$$

para todo $i = 1, \ldots, n$. Entonces

$$Q^{-1}\,[v_i]_{\mathcal{B}} = e_i = [v_i]_{\mathcal{B}'}$$

para todo i y, en consecuencia, $Q^{-1}[v]_{\mathcal{B}} = [v]_{\mathcal{B}'}$ para todo $v \in V$. Como P es única tal que

$$P[v]_{\mathcal{B}} = [v]_{\mathcal{B}'}$$

para todo $v \in V$, concluimos que $P = Q^{-1}$. ■

En resumen, concluimos que las diferentes representaciones de un operador lineal $T \in \mathcal{L}(V)$ forman una de las clases de equivalencias determinadas por la relación de semejanza sobre $\mathcal{M}_n(\mathbb{F})$.

EJERCICIOS

1. Sea $T : \mathbb{C}^2 \to \mathbb{C}^2$ el operador lineal definido por $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$. Considere la base $\mathcal{B} = \left\{ \begin{pmatrix} i \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -i \end{pmatrix} \right\}$ de $\mathbb{C}^2$ y $\mathcal{E}$ la base canónica de $\mathbb{C}^2$. Calcule $[T]_{\mathcal{B}\mathcal{E}}$, $[T]_{\mathcal{E}\mathcal{B}}$, $[T]_{\mathcal{E}\mathcal{E}}$ y $[T]_{\mathcal{B}\mathcal{B}}$.

2. Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ el operador lineal definido por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+z \\ 2y-x \end{pmatrix}$. Considere la base $\mathcal{B} = \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$ de $\mathbb{R}^3$ y las bases canónicas $\mathcal{E}$ y $\mathcal{E}'$ de $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente. Calcule $[T]_{\mathcal{B}\mathcal{E}'}$ y $[T]_{\mathcal{E}\mathcal{E}'}$.

3. Sea $T \in \mathcal{L}(\mathbb{R}^3)$ cuya matriz en la base canónica es

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

   Encuentre una base de $Im(T)$ y una base de $\ker(T)$.

4. Sea $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y $T \in \mathcal{L}(\mathcal{M}_2(\mathbb{R}))$ definida por $T(A) = MA$, para todo $A \in \mathcal{M}_2(\mathbb{R})$. Si $\mathcal{B} = \{I_{11}, I_{12}, I_{21}, I_{22}\}$ es la base canónica de $\mathcal{M}_2(\mathbb{R})$ encuentre $[T]_{\mathcal{B}\mathcal{B}}$.

5. Demuestre que $\dim Im(T) = rgo(A)$ y $\dim\ker(T) = nul(A)$, donde $T \in \mathcal{L}(V)$ y $A = [T]_{\mathcal{B}\mathcal{B}}$ para una base arbitraria $\mathcal{B}$ de V.

6. Demuestre que la función definida por $A \rightsquigarrow \tau_A$ es un isomorfismo entre los espacios vectoriales $\mathcal{M}_n(\mathbb{F})$ y $\mathcal{L}(\mathbb{F}^n)$.

7. Sea $\mathcal{B}=\left\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$ una base ordenada de $\mathbb{R}^3$ y $\mathcal{C}=\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\right\}$ una base ordenada de $\mathbb{R}^2$. Considere la aplicación lineal $T:\mathbb{R}^3\longrightarrow\mathbb{R}^2$ definida por

$$T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x-3z\\2x+y\end{pmatrix}$$

Encuentre:

a) $[T]_{\mathcal{E}\mathcal{E}'}$, donde $\mathcal{E}$ es la base standard de $\mathbb{R}^3$ y $\mathcal{E}'$ es la base standard de $\mathbb{R}^2$;
b) $[T]_{\mathcal{B}\mathcal{C}}$;
c) Matrices invertibles P,Q tales que $[T]_{\mathcal{B}\mathcal{C}}=Q\,[T]_{\mathcal{E}\mathcal{E}'}\,P^{-1}$.

8. Dada $T\in\mathcal{L}(\mathbb{R}^2)$ definida por $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}4x-y\\2x+y\end{pmatrix}$ y las bases de $\mathbb{R}^2$

$$\mathcal{E}=\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}\text{ y }\mathcal{B}=\left\{\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}2\\5\end{pmatrix}\right\}$$

a) Encuentre la matriz cambio de base P de $\mathcal{E}$ a $\mathcal{B}$, la matriz cambio de base Q de $\mathcal{B}$ a $\mathcal{E}$ y compruebe que $Q=P^{-1}$.
b) Encuentre $[T]_{\mathcal{E}}$, $[T]_{\mathcal{B}}$ y compruebe que

$$[T]_{\mathcal{B}}=Q^{-1}\,[T]_{\mathcal{E}}\,Q$$

donde Q es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{E}$.

9. Sea $S\in\mathcal{L}(\mathbb{R}^3)$ definida por $S\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}2y+z\\x-4y\\3x\end{pmatrix}$ y considere las bases de $\mathbb{R}^3$

$$\mathcal{E}=\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}\text{ y }\mathcal{B}=\left\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$$

a) Encuentre la matriz cambio de base P de $\mathcal{E}$ a $\mathcal{B}$, la matriz cambio de base Q de $\mathcal{B}$ a $\mathcal{E}$ y compruebe que $Q=P^{-1}$.
b) Encuentre $[T]_{\mathcal{E}}$, $[T]_{\mathcal{B}}$ y compruebe que

$$[T]_{\mathcal{B}}=Q^{-1}\,[T]_{\mathcal{E}}\,Q$$

donde Q es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{E}$.

10. Considere las bases $\mathcal{B} = \{1, i\}$ y $\mathcal{C} = \{1+i, 1+2i\}$ del espacio vectorial $\mathbb{C}$ sobre $\mathbb{R}$.

    a) Encuentre las matrices cambio de base P y Q de $\mathcal{B}$ a $\mathcal{C}$ y de $\mathcal{C}$ a $\mathcal{B}$ respectivamente;

    b) Compruebe que $Q = P^{-1}$;

    c) Demuestre que $[T]_{\mathcal{B}} = P^{-1} [T]_{\mathcal{C}} P$, donde T es el operador lineal $T : \mathbb{C} \to \mathbb{C}$ definido por $T(z) = \overline{z}$.

11. Sea V un espacio vectorial con bases $\mathcal{B}$ a $\mathcal{E}$. Si P es la matriz cambio de base de $\mathcal{E}$ a $\mathcal{B}$ y Q es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{E}$ demuestre que $Q = P^{-1}$.

12. Demuestre que la relación de semejanza de matrices es una relación de equivalencia.

13. Demuestre que matrices semejantes tienen la misma traza.

14. Demuestre que matrices semejantes tienen el mismo rango.

4.5. Autovalores y autovectores

Uno de los objetivos principales de este curso consiste en estudiar los operadores lineales definidos sobre un espacio vectorial V. Teniendo en cuenta que un operador lineal tiene distintas representaciones matriciales, es natural elegir una base de V de forma que la representación resulte lo más simple posible. Las matrices diagonales son bastante sencillas.

Definición 4.5.1 *Una matriz $D \in \mathcal{M}_n(\mathbb{F})$ es **diagonal** si cumple $[D]_{ij} = 0$ para todo $i \neq j$.*

Es decir, todos los elementos fuera de la diagonal son 0. Si los elementos sobre la diagonal son $[D]_{ii} = \lambda_i$, para $i = 1, \ldots, n$, entonces escribimos

$$D = diag(\lambda_1, \ldots, \lambda_n)$$

Presentamos en este capítulo criterios para decidir cuando existe una base $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ es una matriz diagonal.

Definición 4.5.2 *Sea V un espacio vectorial. Un operador lineal $T \in \mathcal{L}(V)$ es **diagonalizable** si existe una base $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ es una matriz diagonal.*

Ejemplo 4.5.3 *Consideremos el operador lineal $T \in \mathcal{L}(\mathbb{R}^2)$ definido por*

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6x - y \\ 3x + 2y \end{pmatrix}$$

Entonces $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ es una base de $\mathbb{R}^2$. Además,

$$T\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

y

$$T\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}5\\5\end{pmatrix}=0\begin{pmatrix}1\\3\end{pmatrix}+5\begin{pmatrix}1\\1\end{pmatrix}$$

Por lo tanto,

$$[T]_{\mathcal{B}}=\begin{pmatrix}3&0\\0&5\end{pmatrix}$$

y así, T es diagonalizable.

Ejemplo 4.5.4 *Sea $S\in\mathcal{L}(\mathbb{R}^2)$ definida por*

$$S\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\0\end{pmatrix}$$

La matriz asociada a S con respecto a la base canónica es $A=\begin{pmatrix}0&1\\0&0\end{pmatrix}$. Si S fuera diagonalizable existiría una matriz invertible Q tal que $A=QDQ^{-1}$, donde D es diagonal. Luego,

$$\mathbf{0}=A^2=\left(QDQ^{-1}\right)^2=QD^2Q^{-1}$$

Pero entonces $D^2=0$ porque Q es invertible, y esto implica que $D=0$ porque D es diagonal. Esto implica que $A=0$, una contradicción.

El problema de decidir si un operador lineal es diagonalizable es equivalente a determinar si una matriz es semejante a una matriz diagonal.

Teorema 4.5.5 *Sea V un espacio vectorial y $A\in\mathcal{M}_n(\mathbb{F})$ una matriz asociada a $T\in\mathcal{L}(V)$. T es diagonalizable si, y sólo si, A es semejante a una matriz diagonal.*

Demostración. Sea $\mathcal{B}$ una base de V tal que $[T]_{\mathcal{B}}=A$. Si T es diagonalizable, entonces existe una base $\mathcal{C}$ de V tal que $[T]_{\mathcal{C}}=D$, donde D es una matriz diagonal. Por el teorema 4.4.7,

$$D=[T]_{\mathcal{C}}=P\,[T]_{\mathcal{B}}\,P^{-1}=PAP^{-1}$$

Así, A es semejante a una matriz diagonal. Recíprocamente, si $A=[T]_{\mathcal{B}}$ es semejante a una matriz diagonal D, entonces por el teorema 4.4.11, existe una base $\mathcal{C}$ de V tal que $[T]_{\mathcal{C}}=D$. Esto demuestra que T es diagonalizable. ■

Definición 4.5.6 *Decimos que la matriz $A\in\mathcal{M}_n(\mathbb{F})$ es* ***diagonalizable*** *si existe una matriz diagonal D tal que A es semejante a D.*

Introducimos la teoría de autovalores con el fin de estudiar el problema de diagonalización de una matriz.

Definición 4.5.7 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Un **autovalor** de A es un escalar $\lambda \in \mathbb{F}$ tal que*

$$Ax = \lambda x$$

*para algún $0 \neq x \in \mathbb{F}^n$. Decimos que x es un **autovector** de A asociado a λ.*

Ejemplo 4.5.8 *Consideremos la matriz $A = \begin{pmatrix} 7 & -2 \\ 4 & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$. Entonces 3 es un autovalor de A porque*

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

El vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ es un autovector de A asociado a 3.

Teorema 4.5.9 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Las siguientes condiciones son equivalentes:*

1. *λ es un autovalor de A;*
2. *$det(\lambda I - A) = 0$.*

Demostración. Esto es consecuencia de los teoremas 3.4.13 y 2.4.7 porque λ es un autovalor de A si, y sólo si, $(\lambda I - A)x = 0$ para $0 \neq x \in \mathbb{F}^n$ si, y sólo si, $\lambda I - A$ no es invertible si, y sólo si, $det(\lambda I - A) = 0$. ■

Este resultado nos proporciona un método para encontrar los autovalores y autovectores de una matriz. Si tratamos a λ como una variable, entonces $det(\lambda I - A)$ es un polinomio en λ y los autovalores de A son las raíces de la ecuación polinómica $det(\lambda I - A) = 0$. Después de encontrar los autovalores de A encontramos los autovectores como la solución no trivial del sistema de ecuaciones homogéneas $(\lambda I - A)x = 0$. Es decir, los vectores diferentes de cero en $ker(\lambda I - A)$. Llamamos a $ker(\lambda I - A)$ el **autoespacio** de A asociado al autovalor λ.

Definición 4.5.10 *El polinomio $p_A = p_A(x) = det(xI - A) \in \mathbb{F}[x]$ se llama **polinomio característico** de A.*

El hecho de que los autovalores de una matriz A se obtienen como raíces de $p_A(x)$ nos indica que hay diferencias cuando tomamos $\mathbb{F} = \mathbb{R}$ o $\mathbb{F} = \mathbb{C}$

Ejemplo 4.5.11 *Consideremos la matriz $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$. Entonces el polinomio característico de B es*

$$p_B = det \begin{pmatrix} x & -1 \\ 1 & x \end{pmatrix} = x^2 + 1 \in \mathbb{R}[x]$$

Este polinomio no tiene raíces sobre $\mathbb{R}$ y, por lo tanto, B no tiene autovalores.

Ejemplo 4.5.12 *Consideremos la matriz* $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$. *Entonces el polinomio característico de* B *es*

$$p_B = det \begin{pmatrix} x & -1 \\ 1 & x \end{pmatrix} = x^2 + 1 = (x+i)(x-i)$$

En este ejemplo notamos que los autovalores de B *son* $\{-i, i\}$. *El autoespacio asociado al autovalor* $\lambda = -i$ *es el subespacio generado por el vector* $\begin{pmatrix} 1 \\ -i \end{pmatrix}$. *El autoespacio asociado a* $\lambda = i$ *es el subespacio generado por el vector* $\begin{pmatrix} 1 \\ i \end{pmatrix}$.

El próximo resultado muestra que matrices semejantes tienen el mismo polinomio característico.

Teorema 4.5.13 *Sean* $A, B \in \mathcal{M}_n(\mathbb{F})$ *matrices semejantes. Entonces* $p_A = p_B$.

Demostración. Sea $P \in \mathcal{M}_n(\mathbb{F})$ una matriz invertible tal que $B = P^{-1}AP$. Entonces

$$p_A = det(xI - A) = det\left(P^{-1}(xI - A)P\right) = det(xI - B) = p_B$$

■

En particular, matrices semejantes tienen los mismos autovalores.

Concluimos esta sección con un importante resultado sobre la independencia lineal de autovectores.

Teorema 4.5.14 *Sea* $A \in \mathcal{M}_n(\mathbb{F})$ *y* $\lambda_1, \ldots, \lambda_k$ *los diferentes autovalores de* A.

1. *Si* x_i *es un autovector asociado a* λ_i *para cada* $i = 1, \ldots, k$, *entonces el conjunto* $x_1, \ldots, x_k$ *es linealmente independiente;*

2. *Si para cada* $i = 1, \ldots, k$, $\mathcal{B}_i$ *es una base de* $ker(\lambda_i I - A)$ *entonces* $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ *es un conjunto linealmente independiente.*

Demostración. 1. Supongamos que el conjunto $x_1, \ldots, x_k$ es linealmente dependiente. Entonces para algún entero $1 \leq s < k$ tenemos que $x_1, \ldots, x_s$ es linealmente independiente pero $x_1, \ldots, x_{s+1}$ es linelamente dependiente. Luego

$$x_{s+1} = \alpha_1 x_1 + \cdots + \alpha_s x_s \tag{4.2}$$

donde no todos los $\alpha_i \in \mathbb{F}$ son ceros. Multiplicando ambos lados por A obtenemos

$$\lambda_{s+1} x_{s+1} = \alpha_1 \lambda_1 x_1 + \cdots + \alpha_s \lambda_s x_s$$

y multiplicando ambos lados de (4.2) por λ_{s+1} obtenemos

$$\lambda_{s+1}x_{s+1} = \lambda_{s+1}\alpha_1 x_1 + \cdots + \lambda_{s+1}\alpha_s x_s$$

Por la independencia lineal de $x_1, \ldots, x_s$ concluimos que $\alpha_j\lambda_j = \lambda_{s+1}\alpha_j$ para todo $j = 1, \ldots, s$. Sea r tal que $\alpha_r \neq 0$. Entonces $\alpha_r(\lambda_r - \lambda_{s+1}) = 0$ y así $\lambda_r = \lambda_{s+1}$, una contradicción.

2. Vamos a demostrar que

$$ker(\lambda_j I - A) \cap [ker(\lambda_1 I - A) + \cdots + ker(\lambda_{j-1} I - A)] = \{0\}$$

para todo $j = 2, \ldots, k$. En efecto, supongamos que $x \in ker(\lambda_j I - A)$ y $x = x_1 + \cdots + x_{j-1}$, donde $x_i \in ker(\lambda_i I - A)$ para todo $i = 1, \ldots, j-1$. Entonces

$$\sum_{i=1}^{j-1}(\lambda_i - \lambda_j)x_i = \sum_{i=1}^{j-1}\lambda_i x_i - \lambda_j\sum_{i=1}^{j-1}x_i = Ax - \lambda_j x = 0$$

Se deduce de la primera parte que todos los x_i son cero y, por lo tanto, $x = 0$.

Ahora supongamos que $\mathcal{B}_i = \{x_{i,1}, \ldots, x_{i,t_i}\}$ y consideremos una combinación lineal

$$\alpha_{1,1}x_{1,1} + \cdots + \alpha_{1,t_1}x_{1,t_1} + \cdots + \alpha_{k,1}x_{k,1} + \cdots + \alpha_{k,t_k}x_{k,t_k} = 0$$

Si llamamos $y_i = \alpha_{i,1}x_{i,1} + \cdots + \alpha_{i,t_i}x_{i,t_i} \in ker(\lambda_i I - A)$, entonces $y_1 + \cdots + y_k = 0$. Sea j el máximo entero positivo tal que $y_j \neq 0$. Luego

$$y_j = y_1 + \cdots + y_{j-1} \in ker(\lambda_j I - A) \cap [ker(\lambda_1 I - A) + \cdots + ker(\lambda_{j-1} I - A)] = \{0\}$$

Esto es una contradicción. Por lo tanto, $y_i = 0$ para todo i y así, $\alpha_{i,t} = 0$ para todo $1 \leq i \leq k$ y $1 \leq t \leq t_i$. Esto termina la demostración. ■

EJERCICIOS

1. Encuentre el polinomio característico de la matriz $A = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$, los autovalores y los autoespacios correspondientes. ¿Qué ocurre si considera a $A = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$?

2. Sea $A \in \mathcal{M}_n(\mathbb{F})$ una matriz triangular. Demuestre que los autovalores de A son los elementos de la diagonal de A.

3. Encuentre el polinomio característico de la matriz $A = \begin{pmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$, los autovalores y los autoespacios correspondientes.

4. Sea $T \in \mathcal{L}(V)$ donde V es un espacio vectorial de dimensión n. Si T tiene n autovalores diferentes, demuestre que T es diagonalizable.

5. Sean $A, B \in \mathcal{M}_n(\mathbb{F})$. Demuestre que si $I - AB$ es invertible, entonces $I - BA$ es invertible y que
$$(I - BA)^{-1} = I + B(I - AB)^{-1} A$$
Concluya que AB y BA tienen los mismos autovalores.

6. Sea V el espacio de las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$ y $T : V \to V$ el operador lineal sobre V definido por
$$(Tf)(x) = \int_0^x f(t)\, dt$$
Demuestre que T no tiene autovalores.

7. Sea $D \in \mathcal{M}_n(\mathbb{F})$ una matriz diagonal con polinomio característico
$$(x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$$
donde $c_1, \ldots, c_k$ son distintos. Sea $V = \{A \in \mathcal{M}_n(\mathbb{F}) : AD = DA\}$. Demuestre que la dimensión de V es $d_1^2 + \cdots + d_k^2$.

8. Demuestre que 0 es un autovalor de A si, y sólo si, A no es invertible.

9. Suponga que λ es un autovalor de la matriz invertible A. Demuestre que λ^{-1} es un autovalor de A^{-1}.

10. Suponga que v es un autovector de las matrices $A, B \in \mathcal{M}_n(\mathbb{F})$. Demuestre que v es un autovector de $\alpha A + \beta B$, donde $\alpha, \beta \in \mathbb{F}$.

11. Suponga que λ es un autovalor de la matriz A con autovector asociado v. Demuestre que para todo entero $n > 0$, λ^n es un autovalor de A^n con autovector asociado v.

12. Sean $A, B \in \mathcal{M}_n(\mathbb{F})$ tales que $AB = BA$. Sea λ un autovalor de B y W el autoespacio de B asociado a λ. Demuestre que $AW \subseteq W$.

13. Sea $T \in \mathcal{L}(V)$. Un autovalor de T es un escalar $\lambda \in \mathbb{F}$ tal que
$$Tx = \lambda x$$
para algún $0 \neq x \in V$. Se dice que x es un autovector de T asociado a λ. Sea $A = [T]_{\mathcal{B}}$ la matriz asociada a T con respecto a la base $\mathcal{B}$. Demuestre que $\lambda \in \mathbb{F}$ es un autovalor de T si, y sólo si, λ es un autovalor de A.

4.6. Diagonalización

En esta sección establecemos la conexión entre la teoría de autovalores y el problema de diagonalización.

Teorema 4.6.1 *La matriz $A \in \mathcal{M}_n(\mathbb{F})$ es diagonalizable si, y sólo si, existe una base de $\mathbb{F}^n$ formada por autovectores de A.*

Demostración. Sea P una matriz invertible tal que $AP = PD$, donde

$$D = diag(\lambda_1, \ldots, \lambda_n).$$

Por el teorema 3.4.13, las columnas de $P_{*1}, \ldots, P_{*n}$ de P forman una base de $\mathbb{F}^n$. Por otra parte, un cálculo simple demuestra que

$$AP_{*j} = (AP)_{*j} = (PD)_{*j} = \lambda_j P_{*j}$$

para cada $j = 1, \ldots, n$. Así, cada P_{*j} es un autovector asociado al autovalor λ_j. Recíprocamente, supongamos que $y_1, \ldots, y_n$ es una base de $\mathbb{F}^n$ formada por autovectores de A, digamos $Ay_j = \lambda_j y_j$. Sea P la matriz cuya columna j es el vector y_j. Entonces P es invertible y para todo $j = 1, \ldots, n$

$$(AP)_{*j} = Ay_j = \lambda_j y_j = (PD)_{*j}$$

donde $D = diag(\lambda_1, \ldots, \lambda_n)$. Esto demuestra que $P^{-1}AP = diag(\lambda_1, \ldots, \lambda_n)$. ■

Damos a continuación un criterio útil para decidir cuando una matriz $A \in \mathcal{M}_n(\mathbb{F})$ es diagonalizable, en el caso que el polinomio característico p_A de A se expresa como producto de factores lineales:

$$p_A = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k} \tag{4.3}$$

donde $\lambda_1, \ldots, \lambda_k$ son los autovalores diferentes de A y $m_1 + \cdots + m_k = n$. Para cada $j = 1, \ldots, k$, recordamos que la multiplicidad algebraica de λ_j, que denotamos por $ma_A(\lambda_j)$, es el número m_j. Esto es, $ma_A(\lambda_j) = m_j$.

Es importante notar que por el Teorema Fundamental del Álgebra, si tomamos $\mathbb{F} = \mathbb{C}$, entonces cualquier matriz $A \in \mathcal{M}_n(\mathbb{C})$ tiene polinomio p_A en la forma dada en (4.3).

Definición 4.6.2 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. La dimensión del autoespacio $ker(\lambda I - A)$ asociado a un autovalor λ de A se llama la **multiplicidad geométrica** de λ y la denotamos por $mg_A(\lambda)$.*

En general, la multiplicidad algebraica y geométrica de un autovalor son diferentes.

Ejemplo 4.6.3 *Consideremos la matriz $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Entonces $\Phi_A = x^2$ y, por lo tanto, el único autovalor de A es el cero. En este caso, $dim[ker(0I - A)] = dim[ker(A)] = 1$. Esto es, $mg_A(0) = 1 \neq 2 = ma_A(0)$.*

Teorema 4.6.4 *Sea $A \in \mathcal{M}_n(\mathbb{F})$ y λ_0 un autovalor de A. Entonces $mg_A(\lambda_0) \leq ma_A(\lambda_0)$.*

Demostración. Supongamos que $mg_A(\lambda_0) = r$ y sea $Q_1, \ldots, Q_r$ una base de $ker(\lambda_0 I - A)$. Extendamos hasta una base

$$Q_1, \ldots, Q_r, Q_{r+1}, \ldots, Q_n$$

de $\mathbb{F}^n$ y consideremos la matriz invertible Q que tiene a Q_j como columna j, para $j = 1, \ldots, n$. Entonces

$$\left(Q^{-1}AQ\right)_{*j} = Q^{-1}AQ_{*j} = Q^{-1}AQ_j = Q^{-1}\lambda_0 Q_j = \lambda_0 e_j$$

para todo $j = 1, \ldots, r$. Por lo tanto

$$Q^{-1}AQ = \begin{pmatrix} \lambda_0 I_r & B \\ 0 & C \end{pmatrix}$$

para ciertas $B \in \mathcal{M}_{r,n-r}(\mathbb{F})$ y $C \in \mathcal{M}_{n-r}(\mathbb{F})$. Se concluye del teorema 4.5.13 que

$$p_A = p_{Q^{-1}AQ} = det\begin{pmatrix} (x - \lambda_0) I_r & -B \\ 0 & xI_{n-r} - C \end{pmatrix}$$

Desarrollando este determinante por las primeras m columnas obtenemos que $p_A = (x - \lambda_0)^r |xI_{n-r} - C|$ lo cual demuestra que

$$ma_A(\lambda_0) \geq r = mg_A(\lambda_0)$$

■

Ahora estamos en condiciones de presentar un criterio para decidir cuando una matriz es diagonalizable.

Teorema 4.6.5 *Sea $A \in \mathcal{M}_n(\mathbb{F})$ tal que p_A se expresa como producto de factores lineales. Entonces A es diagonalizable si, y sólo si, $mg_A(\lambda) = ma_A(\lambda)$ para todo $\lambda \in \sigma(A)$.*

Demostración. Supongamos que A es diagonalizable y sea λ un autovalor de A tal que $ma_A(\lambda) = r$. Entonces

$$P^{-1}AP = \begin{pmatrix} \lambda I_r & 0 \\ 0 & B \end{pmatrix}$$

donde $B \in \mathcal{M}_{n-r}(\mathbb{F})$ es diagonal y λ no es autovalor de B. En consecuencia,

$$(\lambda I - A) = P\begin{pmatrix} 0 & 0 \\ 0 & \lambda I_{n-r} - B \end{pmatrix}P^{-1}$$

Como matrices semejantes tienen igual rango se obtiene que

$$rgo(\lambda I - A) = rgo\begin{pmatrix} 0 & 0 \\ 0 & \lambda I_{n-r} - B \end{pmatrix} = rgo(\lambda I_{n-r} - B) = n - r$$

Se deduce del Teorema de la Dimensión que

$$mg_A(\lambda) = dim\left[ker(\lambda I - A)\right] = n - rgo(\lambda I - A) = n - (n-r) = r$$

Recíprocamente, supongamos que $\lambda_1, \ldots, \lambda_k$ son los diferentes autovalores de A y supongamos que $mg_A(\lambda_i) = ma_A(\lambda_i)$ para todo i. Elijamos para cada $i = 1, \ldots, k$ una base $\mathcal{B}_i$ de $ker(\lambda_i I - A)$. Entonces, por el teorema 4.5.14, el conjunto $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ es linealmente independiente. Además, el número de vectores en $\mathcal{B}$ es $\sum_{i=1}^{k} mg_A(\lambda_i) = \sum_{i=1}^{k} ma_A(\lambda_i) = n$. En consecuencia, $\mathcal{B}$ es una base de $\mathbb{F}^n$ formada por autovectores de A y el resultado se deduce del teorema 4.6.1. ■

Ejemplo 4.6.6 *Sea* $A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 2 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$. *Entonces*

$$p_A = det\begin{pmatrix} x-2 & -1 & -1 \\ -2 & x-3 & -2 \\ -1 & -1 & x-2 \end{pmatrix} = (x-5)(x-1)^2$$

Por lo tanto, el conjunto de autovalores de A *es el conjunto* $\{1,5\}$, $ma_A(1) = 2$ *y* $ma_A(5) = 1$. *Calculemos ahora las multiplicidades geométricas de los autovalores de* A. *Usando las técnicas del primer capítulo tenemos*

$$5I - A = \begin{pmatrix} 3 & -1 & -1 \\ -2 & 2 & -2 \\ -1 & -1 & 3 \end{pmatrix} \underset{f}{\sim} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.4}$$

Luego $rgo(5I - A) = 2$, *y por el Teorema de la Dimensión,* $mg_A(5) = 1$. *Similarmente,*

$$1I - A = \begin{pmatrix} -1 & -1 & -1 \\ -2 & -2 & -2 \\ -1 & -1 & -1 \end{pmatrix} \underset{f}{\sim} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.5}$$

y así, $rgo(I - A) = 1$ *lo cual implica* $mg_A(1) = 2$. *El teorema 4.6.5 nos garantiza que* A *es diagonalizable. Para encontrar la matriz* P *invertible tal que* $P^{-1}AP$ *es diagonal debemos construir bases de los autoespacios* $ker(5I - A)$ *y* $ker(I - A)$. *De la relación (4.4) obtenemos que*

$$\begin{aligned} y - 2z &= 0 \\ x - z &= 0 \end{aligned}$$

Luego la solución del sistema homogéneo $(5I - A)X = 0$ *viene dada por los vectores de la forma* $\begin{pmatrix} x \\ 2x \\ x \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. *Es decir,* $\left\{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\right\}$ *es una base de* $ker(5I - A)$. *Por otra parte, se deduce de la relación (4.5) que*

$$x + y + z = 0$$

Luego el conjunto $\left\{\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right\}$ *forma una base de* $\ker(I-A)$. *La matriz invertible* P *que buscamos es*

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

En efecto,

$$\begin{aligned} AP &= (AP_{*1}, AP_{*2}, AP_{*3}) = (5P_{*1}, P_{*2}, P_{*3}) \\ &= (P_{*1}, P_{*2}, P_{*3})\, diag\,(5,1,1) = P diag\,(5,1,1) \end{aligned}$$

Ejemplo 4.6.7 *Sea* $B = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$. *Entonces*

$$p_B = \begin{vmatrix} x-3 & -1 & 0 \\ 0 & x-3 & -1 \\ 0 & 0 & x-3 \end{vmatrix} = (x-3)^3$$

Así, 3 *es el único autovalor de* B *y* $ma_B(3) = 3$. *Por otro lado,*

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

En consecuencia, $rgo\,(3I-B) = 2$ *y* $mg_B(3) = 1$. *Concluimos que la matriz* B *no es diagonalizable.*

EJERCICIOS

1. Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ el operador lineal definido por $F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6x-y \\ 3x+2y \end{pmatrix}$. ¿Es F diagonalizable? En caso de serlo, encuentre una base $\mathcal{B}$ de $\mathbb{R}^2$ tal que $[F]_{\mathcal{B}}$ es diagonal.

2. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ el operador lineal definido por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+y \\ y-z \\ 2y+4z \end{pmatrix}$. ¿Es T diagonalizable? En caso de serlo, encuentre una base $\mathcal{B}$ de $\mathbb{R}^3$ tal que $[T]_{\mathcal{B}}$ es diagonal.

3. Sea $A = \begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{pmatrix}$. Es A semejante, sobre $\mathbb{C}$, a una matriz diagonal?

4. Sea $A \in \mathcal{M}_2(\mathbb{R})$ simétrica. Demuestre que A es diagonalizable.

5. ¿Bajo que condiciones es la matriz $A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \end{pmatrix}$ diagonalizable?

CAPÍTULO 5

ESPACIOS VECTORIALES CON PRODUCTO INTERNO

5.1. Producto interno

Es nuestro interés ahora extender la geometría básica de los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$ a espacio vectoriales de dimensión finita. En lo que sigue, V es un espacio vectorial de dimensión finita sobre $\mathbb{F}$, donde $\mathbb{F} = \mathbb{R}$ o $\mathbb{F} = \mathbb{C}$.

Definición 5.1.1 *Un* ***producto interno*** *sobre V es una función que asigna a cada par ordenado $x, y \in V$ un elemento $\langle x, y\rangle \in \mathbb{F}$ que satisface las siguientes condiciones:*

1. $\langle x, y\rangle = \overline{\langle y, x\rangle}$ *para todo* $x, y \in V$;
2. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ *para todo* $x, y, z \in V$;
3. $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$ *para todo* $\lambda \in \mathbb{F}$ *y* $x, y \in V$;
4. $\langle x, x\rangle \geq 0$. *Además,* $\langle x, x\rangle = 0$ *si, y sólo si,* $x = 0$.

Un espacio vectorial V junto con un producto interno definido sobre V se llama un ***espacio con producto interno****.*

La condición 4 de la definición de producto interno establece que $\langle x, x\rangle$ es un número real nonegativo, aun en el caso que $\mathbb{F} = \mathbb{C}$. Por otra parte, las condiciones 2 y 3 establecen linealidad en la primera componente. Con relación a la segunda componente tenemos que para todo $x, y, z \in V$ y $\lambda, \mu \in \mathbb{F}$

$$\langle x, \lambda y + \mu z\rangle = \overline{\langle \lambda y + \mu z, x\rangle} = \overline{\lambda \langle y, x\rangle + \mu \langle z, x\rangle} = \overline{\lambda} \langle x, y\rangle + \overline{\mu} \langle x, z\rangle$$

Veamos algunos ejemplos importantes de espacios con producto interno.

Ejemplo 5.1.2 *Los siguientes espacios vectoriales son espacios con producto interno:*

a *El espacio* $\mathbb{F}^n$ *con el producto interno*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$$

donde $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{F}^n$. *Este producto interno recibe el nombre de producto interno canónico de* $\mathbb{F}^n$.

b *El espacio* $S = \{f \in \mathcal{F}([a,b], \mathbb{R}) : f \text{ es continua}\}$ *con el producto interno*

$$\langle f, g \rangle = \int_a^b f(x)\, g(x)\, dx$$

donde $f, g \in S$.

c *El espacio* $\mathcal{M}_{m \times n}(\mathbb{F})$ *con el producto interno*

$$\langle A, B \rangle = Tr(AB^*)$$

donde $A, B \in \mathcal{M}_{m \times n}(\mathbb{F})$.

d *Sea*

$$S = \{(x_n) \in Suc(\mathbb{F}) : x_n = 0 \text{ para todo } n, \text{ excepto un número finito}\}.$$

Es fácil ver que S *es un espacio con producto interno*

$$\langle (x_n), (y_n) \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$$

En un espacio con producto interno es posible definir los conceptos de norma, ortogonalidad y proyecciones de la misma manera como se realiza en los espacios conocidos $\mathbb{R}^2$ y $\mathbb{R}^3$.

Definición 5.1.3 *Sea* V *un espacio con producto interno. La* ***norma*** *de un vector* $v \in V$, *denotado por* $\|v\|$, *está definida como*

$$\|v\| = \sqrt{\langle v, v \rangle}$$

Teorema 5.1.4 *Sea* V *un espacio con producto interno. Si* $x, y \in V$ *y* $\lambda \in \mathbb{F}$, *entonces*

1. $\|x\| \geq 0$. *Además,* $\|x\| = 0$ *si, y sólo si,* $x = 0$;
2. $\|\lambda x\| = |\lambda| \, \|x\|$.

Demostración. Demostramos 2. Si $x \in V$ y $\lambda \in \mathbb{F}$ entonces

$$\|\lambda x\|^2 = \langle \lambda x, \lambda x\rangle = \lambda\overline{\lambda}\langle x, x\rangle = |\lambda|^2 \langle x, x\rangle$$

Ahora tomamos raíz cuadrada en ambos miembros de la igualdad. ■

Definición 5.1.5 *Sea V un espacio con producto interno. Los vectores $x, y \in V$ son* ***ortogonales*** *si $\langle x, y\rangle = 0$.*

En la definición anterior no importa el orden de los vectores porque $\langle x, y\rangle = 0$ si, y sólo si, $\langle y, x\rangle = 0$, tomando en cuenta que $\langle x, y\rangle = \overline{\langle y, x\rangle}$. El vector 0 es ortogonal a todo vector $v \in V$:

$$\langle 0, v\rangle = \langle 0 + 0, v\rangle = \langle 0, v\rangle + \langle 0, v\rangle$$

Por lo tanto, $\langle 0, v\rangle = 0$ para todo $v \in V$. De hecho, es el único con esta propiedad (ver Ejercicio 9).

Teorema 5.1.6 *(Teorema de Pitágoras) Sea V un espacio con producto interno. Si $x, y \in V$ son ortogonales, entonces*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Demostración. Si $x, y \in V$ son ortogonales, entonces $\langle x, y\rangle = 0$. Luego

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 \\ &= \|x\|^2 + \|y\|^2 \end{aligned}$$

■

Consideramos ahora el concepto de proyección de un vector sobre otro en un espacio con producto interno V.

Teorema 5.1.7 *Sean $x, y \in V$ y $y \neq 0$. Existe un único vector $p \in V$ que es un múltiplo escalar de y y que es ortogonal a $x - p$.*

Demostración. Sea $p = \frac{\langle x,y\rangle}{\|y\|^2} y$. Entonces

$$\langle x - p, p\rangle = \langle x, p\rangle - \|p\|^2 = \frac{\langle x, y\rangle^2}{\|y\|^2} - \frac{\langle x, y\rangle^2}{\|y\|^2} = 0.$$

Para ver la unicidad, supongamos que λy es ortogonal a $x - \lambda y$. Entonces $0 = \langle x - \lambda y, y\rangle = \langle x, y\rangle - \lambda \|y\|^2$, de donde deducimos que $\lambda = \frac{\langle x,y\rangle}{\|y\|^2}$ y, en consecuencia, $\lambda y = p$. ■

Definición 5.1.8 *Sea V un espacio con producto interno, $x, y \in V$ y $y \neq 0$. El vector $p = \frac{\langle x,y\rangle}{\|y\|^2} y$ recibe el nombre de* ***proyección de** x **sobre** y. El escalar $\frac{\langle x,y\rangle}{\|y\|^2}$ se llama el* ***coeficiente de Fourier*** *de x con respecto a y.*

Teorema 5.1.9 *(Desigualdad de Cauchy-Schwarz) Sea V un espacio con producto interno. Si $x, y \in V$, entonces*

$$|\langle x, y\rangle| \leq \|x\| \, \|y\| \, . \tag{5.1}$$

La igualdad se satisface si, y sólo si, uno de los vectores x, y es un múltiplo escalar del otro.

Demostración. Si $y = 0$, entonces ambos lados de la desigualdad (5.1) son iguales a 0. Así suponemos que $y \neq 0$. Sea p la proyección de x sobre y. Entonces $x = p + (x - p)$ donde p y $x - p$ son ortogonales. Se deduce del Teorema de Pitágoras que

$$\begin{aligned} \|x\|^2 &= \|p\|^2 + \|x - p\|^2 \geq \|p\|^2 \\ &= \left\| \frac{\langle x, y\rangle}{\|y\|^2} y \right\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^4} \|y\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^2} \end{aligned} \tag{5.2}$$

Multiplicando ambos lados de la desigualdad por $\|y\|^2$ y luego tomando raíz cuadrada se obtiene la desigualdad de Cauchy-Schwarz.

Para ver la segunda parte notamos que (5.1) es una igualdad si, y sólo si, (5.2) es una igualdad. Esto ocurre si, y sólo si, $\|x - p\| = 0$, o equivalentemente, x es un múltiplo escalar de y. ■

Aplicando la desigualdad de Cauchy-Schwarz a los espacios con producto interno del Ejemplo 5.1.2 obtenemos

1. $\left| \sum\limits_{i=1}^{n} x_i \overline{y_i} \right| \leq \left(\sum\limits_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}} \left(\sum\limits_{i=1}^{n} |y_i|^2 \right)^{\frac{1}{2}}$, para todo $x_i, y_i \in \mathbb{F}$, $i = 1, \ldots, n$;

2. $\left| \int\limits_a^b f(x) g(x) \, dx \right| \leq \left(\int\limits_a^b [f(x)]^2 \, dx \right)^{\frac{1}{2}} \left(\int\limits_a^b [g(x)]^2 \, dx \right)^{\frac{1}{2}}$, para toda las funciones continuas $f, g \in \mathcal{F}([a, b], \mathbb{R})$;

3. $|Tr(AB^*)| \leq [Tr(AA^*)]^{\frac{1}{2}} [Tr(BB^*)]^{\frac{1}{2}}$, para todas las matrices $A, B \in \mathcal{M}_n(\mathbb{F})$;

4. $\left| \sum\limits_{i=1}^{\infty} x_i \overline{y_i} \right| \leq \left(\sum\limits_{i=1}^{\infty} |x_i|^2 \right)^{\frac{1}{2}} \left(\sum\limits_{i=1}^{\infty} |y_i|^2 \right)^{\frac{1}{2}}$, para todas las sucesiones $(x_n), (y_n) \in Suc(\mathbb{F})$ tales que $x_n = 0$ y $y_n = 0$ para todo n excepto un número finito.

Teorema 5.1.10 *(Desigualdad triangular) Sea V un espacio con producto interno. Si $x, y \in V$, entonces $\|x + y\| \leq \|x\| + \|y\|$. La igualdad se cumple si, y sólo si, uno de los vectores x, y es un múltiplo no negativo del otro.*

Demostración. Si $x, y \in V$, entonces se obtiene de la desigualdad de Cauchy-Schwarz que

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \\
&= \|x\|^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + \|y\|^2 \\
&= \|x\|^2 + 2Re\langle x,y\rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2\left|\langle x,y\rangle\right| + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2
\end{aligned}$$

Tomando raíz cuadrada ambos lados obtenemos la desigualdad triangular.

La igualdad ocurre si, y sólo si,

$$\|x\|^2 + 2\langle x,y\rangle + \|y\|^2 = \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$$

o equivalentemente, $\langle x,y\rangle = \|x\|\|y\|$. El resultado es consecuencia del teorema anterior. ■

EJERCICIOS

1. Compruebe que los espacios vectoriales dados en el ejemplo 5.1.2 son espacios con producto interno.

2. Considere el espacio $\mathbb{C}^4$ con el producto interno canónico y los vectores $x = (1+i, 3i, 3-2i, 5)$ y $y = (3, 4-i, 0, 2i)$. Calcule $\langle x,y\rangle, \|x\|$ y $\|y\|$.

3. Considere el espacio $\mathcal{M}_{2\times 3}(\mathbb{R})$ con el producto interno

$$\langle A,B\rangle = Tr\left(B^\top A\right)$$

   donde $A, B \in \mathcal{M}_{2\times 3}(\mathbb{R})$. Sean

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

   Calcule $\langle 2A+3B, 4C\rangle, \|A\|$ y $\|B\|$.

4. Demuestre que si $x = (x_1, x_2)$ y $y = (y_1, y_2)$ son elementos de $\mathbb{R}^2$, entonces

$$\langle x,y\rangle = x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2$$

   define un producto interno sobre $\mathbb{R}^2$.

5. Demuestre que si V es un espacio sobre $\mathbb{R}$ con producto interno, entonces, para todo $x, y \in V$

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$$

6. Demuestre que si V es un espacio sobre $\mathbb{C}$ con producto interno, entonces, para todo $x, y \in V$
$$\langle x, y \rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) + \frac{i}{4}\left(\|x+iy\|^2 - \|x-iy\|^2\right)$$

7. Sea V un espacio sobre $\mathbb{R}$ con producto interno. Demuestre que dados $x, y \in V$

   a) $\|x\| = \|y\|$ si, y sólo si, $\langle x+y, x-y \rangle = 0$;
   b) $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ si, y sólo si, $\langle x, y \rangle = 0$.
   Demuestre por medio de contraejemplos que las afirmaciones anteriores no son ciertas para $\mathbb{C}^2$.

8. En el espacio $\mathbb{C}^2$ con producto interno canónico calcule el coeficiente de Fourier y la proyección de $(3+4i, 2-3i)$ sobre $(5+i, 2i)$.

9. Sea V un espacio con producto interno. Demuestre que si $\langle u, v \rangle = 0$ para todo $v \in V$, entonces $u = 0$.

10. Sea V un espacio con producto interno y $x, y \in V$. Demuestre que $x = y$ si, y sólo si, $\langle x, z \rangle = \langle y, z \rangle$ para todo $z \in V$.

11. Sea V un espacio vectorial con producto interno y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V. Demuestre que si $\alpha_1, \ldots, \alpha_n$ son escalares de $\mathbb{F}$, entonces existe un único $v \in V$ tal que $\langle v, v_j \rangle = \alpha_j$ para todo $j = 1, \ldots, n$.

12. Sea V un espacio vectorial sobre $\mathbb{F}$ con producto interno. Demuestre la ley del paralelogramo:
$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

13. Sea V un espacio con producto interno. Definimos la **distancia** entre los vectores x y y en V por
$$d(x, y) = \|x-y\|$$
Demuestre que

   a) $d(x, y) \geq 0$;
   b) $d(x, y) = 0$ si, y sólo si, $x = y$;
   c) $d(x, y) = d(y, x)$;
   d) $d(x, y) \leq d(x, z) + d(z, y)$.

14. Sean $v_1, \ldots, v_n$ vectores mutuamente ortogonales y tales que $\|v_i\| \neq 0$. Sea $v \in V$ y α_i el coeficiente de Fourier de v con respecto a v_i. Demuestre que

   a) $v - \sum \alpha_i v_i$ es ortogonal a cada uno de los v_i;
   b) $\|v - \sum \alpha_i v_i\| \leq \|v - \sum \beta_i v_i\|$ para todo $\beta_i \in \mathbb{F}$.

5.2. Bases ortonormales

Si V es un espacio vectorial y $v_1, \ldots, v_n$ es una base de V, entonces, para cada vector $v \in V$ existen escalares únicos $\alpha_1, \ldots, \alpha_n$ tales que

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

El procedimiento para calcular los escalares resulta en general complicado. Por ejemplo, si los vectores $v_1, \ldots, v_n$ forman una base de $\mathbb{R}^n$, entonces debemos resolver la ecuación $Ax = v$, donde A es la matriz (invertible) cuya columna j es el vector v_j. La solución de esta ecuación viene dada por $x = A^{-1}v$, por lo tanto, debemos calcular la inversa de la matriz A de tamaño $n \times n$. En esta sección construimos bases para espacios con producto interno que, entre otras cosas, facilitan este cálculo.

Definición 5.2.1 *Sea V un espacio con producto interno. Un conjunto de vectores $v_1, \ldots, v_k$ de V es **ortogonal** si $\langle v_i, v_j \rangle = 0$ para todo $i \neq j$. Si además cada uno de los vectores tiene norma 1, entonces los vectores $v_1, \ldots, v_k$ son **ortonormales**.*

Ejemplo 5.2.2 *Los siguientes conjuntos de vectores son ortonormales en el espacio indicado:*

1. *Los vectores $e_1, \ldots, e_n$ de $\mathbb{F}^n$ con el producto interno usual;*

2. *Las funciones $\frac{1}{\sqrt{2\pi}}, \frac{sen x}{\sqrt{\pi}}, \frac{sen 2x}{\sqrt{\pi}}, \frac{cos x}{\sqrt{\pi}}, \frac{cos 2x}{\sqrt{\pi}}$ son ortonormales en el espacio*

$$S = \{f \in \mathcal{F}([0, 2\pi], \mathbb{R}) : f \text{ es continua}\}$$

con el producto interno

$$\langle f, g \rangle = \int_0^{2\pi} f(x)\, g(x)\, dx$$

3. *Las matrices simétricas $I_{11}, \ldots, I_{nn}$ con el producto interno dado por la traza.*

Teorema 5.2.3 *Sea V un espacio vectorial con producto interno y $u_1, \ldots, u_k$ un conjunto de vectores ortogonales diferentes de cero de V. Entonces los vectores $u_1, \ldots, u_k$ son linealmente independientes.*

Demostración. Supongamos que

$$\alpha_1 u_1 + \cdots + \alpha_k u_k = 0$$

donde $\alpha_1, \ldots, \alpha_k$ pertenecen a $\mathbb{F}$. Entonces para cada $j = 1, \ldots, k$ tenemos que

$$\begin{aligned} 0 &= \langle 0, u_j \rangle = \langle \alpha_1 u_1 + \cdots + \alpha_k u_k, u_j \rangle \\ &= \alpha_1 \langle u_1, u_j \rangle + \cdots + \alpha_k \langle u_k, u_j \rangle = \alpha_j \langle u_j, u_j \rangle \end{aligned}$$

Como $u_j \neq 0$ se obtiene que $\langle u_j, u_j \rangle > 0$ y, en consecuencia, $\alpha_j = 0$ para todo $j = 1, \ldots, k$. ■

En el siguiente resultado presentamos un método fácil para encontrar las coordenadas de un vector que se escribe como combinación lineal de vectores en un conjunto ortonormal.

Teorema 5.2.4 *Sea V un espacio con producto interno. Si $v \in V$ es una combinación lineal de los vectores ortonormales $u_1, \ldots, u_k$, entonces*

1. $v = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_k \rangle u_k$;
2. $\|v\|^2 = |\langle v, u_1 \rangle|^2 + \cdots + |\langle v, u_k \rangle|^2$.

Demostración. Sea $v \in V$ y supongamos que existen escalares $\alpha_1, .., \alpha_k$ pertenecientes a $\mathbb{F}$ tales que
$$v = \alpha_1 u_1 + \cdots + \alpha_k u_k$$
Tomando el producto interno ambos lados con u_j obtenemos
$$\langle v, u_j \rangle = \langle \alpha_1 u_1 + \cdots + \alpha_k u_k, u_j \rangle = \alpha_j \langle u_j, u_j \rangle = \alpha_j$$
Además, por el Teorema de Pitágoras,
$$\begin{aligned} \|v\|^2 &= \|\alpha_1 u_1 + \cdots + \alpha_k u_k\|^2 = \|\alpha_1 u_1\|^2 + \cdots + \|\alpha_k u_k\|^2 \\ &= |\alpha_1|^2 \|u_1\|^2 + \cdots + |\alpha_k|^2 \|u_k\|^2 = |\alpha_1|^2 + \cdots + |\alpha_k|^2 \\ &= |\langle v, u_1 \rangle|^2 + \cdots + |\langle v, u_k \rangle|^2 \end{aligned}$$

■

La pregunta que nos planteamos ahora es ¿Cuándo existen bases ortonormales en espacios con producto interno? El siguiente teorema es crucial en este sentido porque produce un algoritmo que permite convertir vectores linealmente independientes en vectores ortonormales que generan exactamente el mismo subespacio.

Teorema 5.2.5 *(Gram-Schmidt) Sean $v_1, \ldots, v_k$ vectores linealmente independientes en un espacio con producto interno V. Entonces existen vectores $u_1, \ldots, u_k$ ortonormales de V tales que $gen(v_1, \ldots, v_k) = gen(u_1, \ldots, u_k)$.*

Demostración. La demostración es por inducción sobre k. Si $k = 1$, entonces $v_1 \neq 0$ y tomamos $u_1 = \frac{v_1}{\|v_1\|}$. Supongamos que el resultado es cierto para $k \geq 1$. Sean $v_1, \ldots, v_k, v_{k+1}$ vectores linealmente independientes. Como $v_1, \ldots, v_k$ son vectores linealmente independientes, existen vectores $u_1, \ldots, u_k$ ortonormales de V tales que $gen(v_1, \ldots, v_k) = gen(u_1, \ldots, u_k)$. Consideremos el vector de norma 1
$$u_{k+1} = \frac{v_{k+1} - \langle v_{k+1}, u_1 \rangle u_1 - \cdots - \langle v_{k+1}, u_k \rangle u_k}{\|v_{k+1} - \langle v_{k+1}, u_1 \rangle u_1 - \cdots - \langle v_{k+1}, u_k \rangle u_k\|}$$
En primer lugar notamos que el denominador de esta expresión es diferente de cero, de lo contrario, $v_{k+1} \in gen(u_1, \ldots, u_k) = gen(v_1, \ldots, v_k)$ pero esto no es posible por la independencia lineal de los vectores $v_1, \ldots, v_k, v_{k+1}$. Por otra parte, para cada $j = 1, \ldots, k$
$$\begin{aligned} \langle u_{k+1}, u_j \rangle &= \left\langle \frac{v_{k+1} - \langle v_{k+1}, u_1 \rangle u_1 - \cdots - \langle v_{k+1}, u_k \rangle u_k}{\|v_{k+1} - \langle v_{k+1}, u_1 \rangle u_1 - \cdots - \langle v_{k+1}, u_k \rangle u_k\|}, u_j \right\rangle \\ &= \frac{\langle v_{k+1}, u_j \rangle - \langle v_{k+1}, u_j \rangle}{\|v_{k+1} - \langle v_{k+1}, u_1 \rangle u_1 - \cdots - \langle v_{k+1}, u_k \rangle u_k\|} = 0 \end{aligned}$$

lo que demuestra que los vectores $u_1, \ldots, u_{k+1}$ son ortonormales. Además,

$$v_{k+1} \in gen\,(u_1, \ldots, u_{k+1})$$

y, por lo tanto,

$$gen\,(v_1, \ldots, v_{k+1}) \subseteq gen\,(u_1, \ldots, u_{k+1})$$

Esto, junto con el hecho de que

$$dim\,[gen\,(v_1, \ldots, v_{k+1})] = dim\,[gen\,(u_1, \ldots, u_{k+1})]$$

nos lleva a la conclusión de que

$$gen\,(v_1, \ldots, v_{k+1}) = gen\,(u_1, \ldots, u_{k+1})$$

■

Corolario 5.2.6 *Si V es un espacio vectorial de dimensión finita con producto interno entonces V tiene una base ortonormal.*

Demostración. Seleccionamos una base $v_1, \ldots, v_k$ de V. Por el teorema anterior existen vectores ortonormales $u_1, \ldots, u_k$ tales que

$$gen\,(u_1, \ldots, u_k) = gen\,(v_1, \ldots, v_k) = V$$

Se deduce del teorema 5.2.3 que $u_1, \ldots, u_k$ es una base ortonormal de V. ■

Ejemplo 5.2.7 *Aplicamos el proceso de Gram-Schmidt a la base $\{1, x, x^2\}$ del espacio de polinomios de grado ≤ 2 con coeficientes reales y el producto interno*

$$\langle p, q\rangle = \int_0^1 p\,(x)\,q\,(x)\,dx$$

para obtener la base ortonormal

$$\begin{aligned}
u_1 &= 1 \\
u_2 &= \frac{x - \int_0^1 x dx}{\left\| x - \int_0^1 x dx \right\|} = \frac{x - \frac{1}{2}}{\left\| x - \frac{1}{2} \right\|} = 2\sqrt{3}\left(x - \frac{1}{2}\right) \\
u_3 &= \frac{x^2 - \int_0^1 x^2 dx - 12\left(x - \frac{1}{2}\right)\int_0^1 x^2\left(x - \frac{1}{2}\right)dx}{\left\| x^2 - \int_0^1 x^2 dx - 24\left(x - \frac{1}{2}\right)\int_0^1 x^2\left(x - \frac{1}{2}\right)dx \right\|} \\
&= \frac{x^2 - x + \frac{1}{6}}{\left\| x^2 - x + \frac{1}{6} \right\|} = 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right)
\end{aligned}$$

Definición 5.2.8 *Sea V un espacio con producto interno. Si S es un subconjunto de V, entonces el* ***complemento ortogonal*** *de S, denotado por $S^{\perp}$, lo definimos como*

$$S^{\perp} = \{v \in V : \langle v, s\rangle = 0 \text{ para todo } s \in S\}$$

Es un ejercicio fácil demostrar que para cualquier subconjunto S de V, el complemento ortogonal $S^{\perp}$ es un subespacio de V.

Ejemplo 5.2.9 *En $\mathbb{R}^2$ con el producto interno usual, el complemento ortogonal de la recta gen (u), donde $0 \neq u \in \mathbb{R}^2$, es la recta gen (v), donde v es un vector de $\mathbb{R}^2$ diferente de cero tal que $\langle u, v\rangle = 0$.*

Teorema 5.2.10 *Sea V un espacio de dimensión finita con producto interno. Si U es un subespacio de V, entonces $V = U \oplus U^{\perp}$.*

Demostración. Sea $u_1, \ldots, u_k$ una base ortonormal de U. Para cada $v \in V$ definamos el elemento

$$\overline{v} = \langle v, u_1\rangle u_1 + \cdots + \langle v, u_k\rangle u_k \in U$$

Entonces $v = \overline{v} + (v - \overline{v})$. Además, para cada $i = 1, \ldots, k$

$$\begin{aligned} \langle v - \overline{v}, u_i\rangle &= \langle v, u_i\rangle - \langle \overline{v}, u_i\rangle \\ &= \langle v, u_i\rangle - \langle \langle v, u_1\rangle u_1 + \cdots + \langle v, u_k\rangle u_k, u_i\rangle \\ &= \langle v, u_i\rangle - \langle v, u_i\rangle = 0 \end{aligned}$$

En consecuencia, $v - \overline{v} \in U^{\perp}$. Por otra parte, si $x \in U \cap U^{\perp}$ entonces $0 = \langle x, x\rangle$ y, por lo tanto, $x = 0$. Esto demuestra que $U \cap U^{\perp} = 0$. ■

EJERCICIOS

1. Encuentre una base ortonormal para el subespacio W de $\mathbb{R}^4$ generado por los vectores

$$(1, 1, 1, 1), (1, 1, 2, 4), (1, 2, -4, -3)$$

2. Considerando a $\mathbb{C}^3$ con el producto interno canónico, encuentre una base ortonormal para el subespacio generado por $(1, 0, i)$ y $(2, 1, 1 + i)$.

3. Encuentre una base ortonormal para el espacio de soluciones del sistema homogéneo

$$\begin{aligned} 3x - 2y + w + z &= 0 \\ x + y + 2z &= 0 \end{aligned}$$

4. Sea $V = \{f(x) \in \mathbb{R}[x] \,/\, \text{grado de } f(x) \text{ es menor o igual a } 3\}$ con el producto interno

$$\langle f, g\rangle = \int_0^1 f(x)\, g(x)\, dx$$

a) Halle el complemento ortogonal del subespacio W de V formado por polinomios de grado 0;

b) Aplique el proceso de Gram-Schmidt a la base $\{1, x, x^2, x^3\}$.

5. Considere el espacio $\mathcal{M}_n(\mathbb{C})$ con el producto interno

$$\langle A, B\rangle = Tr(AB^*)$$

Halle el complemento ortogonal del subespacio W de $\mathcal{M}_n(\mathbb{C})$ formado por las matrices diagonales.

6. Sea V un espacio con producto interno y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de V. Demuestre que para vectores $x, y \in V$

$$\langle x, y\rangle = \sum_{k=1}^{n} \langle x, v_k\rangle \overline{\langle y, v_k\rangle}$$

5.3. Operadores lineales

Comenzamos esta sección introduciendo el concepto de **adjunto** de un operador lineal sobre un espacio con producto interno. Dado $T \in \mathcal{L}(V)$, para abreviar escribimos Tx en lugar de $T(x)$, donde $x \in V$.

Lema 5.3.1 *Sea V un espacio con producto interno y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de V. Si $T \in \mathcal{L}(V)$ y $A = [T]_{\mathcal{B}}$, entonces, para todo $1 \leq i, j \leq n$*

$$[A]_{ij} = \langle Tv_j, v_i\rangle$$

Demostración. Por la parte 1 del teorema 5.2.4, $Tv_j = \langle Tv_j, v_1\rangle v_1 + \cdots + \langle Tv_j, v_n\rangle v_n$. Luego

$$A_{*j} = [Tv_j]_{\mathcal{B}} = (\langle Tv_j, v_1\rangle, \ldots, \langle Tv_j, v_n\rangle)^{\top}$$

Así, $[A]_{ij} = \langle Tv_j, v_i\rangle$. ■

Teorema 5.3.2 *Sea V un espacio con producto interno de dimensión finita y $T \in \mathcal{L}(V)$. Entonces existe un único operador lineal $T^* \in \mathcal{L}(V)$ que satisface*

$$\langle Tx, y\rangle = \langle x, T^* y\rangle$$

para todo $x, y \in V$. Además, para cualquier base ortonormal $\mathcal{B}$ de V se cumple

$$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^*$$

Demostración. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de V y $A = [T]_{\mathcal{B}}$. Por el teorema 4.4.5 existe $T^* \in \mathcal{L}(V)$ tal que $[T^*]_{\mathcal{B}} = A^*$. Se deduce del teorema anterior que

$$[A]_{ij} = \langle Tv_j, v_i \rangle \text{ y } [A^*]_{ij} = \langle T^* v_j, v_i \rangle$$

Luego para todo $1 \leq i, j \leq n$

$$\langle v_i, T^* v_j \rangle = \overline{\langle T^* v_j, v_i \rangle} = \overline{[A^*]_{ij}} = \overline{\overline{[A]}}_{ji} = [A]_{ji} = \langle Tv_i, v_j \rangle$$

Se deduce de aquí que $\langle Tx, y \rangle = \langle x, T^* y \rangle$ para todo $x, y \in V$.

Supongamos que $R \in \mathcal{L}(V)$ satisface $\langle Tx, y \rangle = \langle x, Ry \rangle$ para todo $x, y \in V$. En particular,

$$\langle Rv_j, v_i \rangle = \overline{\langle v_i, Rv_j \rangle} = \overline{\langle Tv_i, v_j \rangle} = \langle v_j, Tv_i \rangle = \langle T^* v_j, v_i \rangle$$

Luego $[R]_{\mathcal{B}} = [T^*]_{\mathcal{B}}$ y, en consecuencia, $R = T^*$. ∎

Definición 5.3.3 *El operador lineal $T^* \in \mathcal{L}(V)$ del teorema anterior recibe el nombre de* ***operador adjunto*** *de T.*

Ejemplo 5.3.4 *Encontremos el adjunto del operador lineal $T : \mathbb{C}^3 \to \mathbb{C}^3$ definido por*

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + (1-i)\,y \\ (3+2i)\,x - 4iz \\ 2ix + (4-3i)\,y - 3z \end{pmatrix}$$

Si $\mathcal{E}$ es la base canónica de $\mathbb{C}^3$, entonces

$$[T]_{\mathcal{E}} = \begin{pmatrix} 2 & 1-i & 0 \\ 3+2i & 0 & -4i \\ 2i & 4-3i & -3 \end{pmatrix} \text{ y } [T]^*_{\mathcal{E}} = \begin{pmatrix} 2 & 3-2i & -2i \\ 1+i & 0 & 4+3i \\ 0 & 4i & -3 \end{pmatrix}$$

Por lo tanto, $T^* : \mathbb{C}^3 \to \mathbb{C}^3$ está definido por

$$T^* \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + (3-2i)\,y - 2iz \\ (1+i)\,x + (4+3i)\,z \\ 4iy - 3z \end{pmatrix}$$

El siguiente teorema contiene las propiedades básicas del adjunto de un operador lineal.

Teorema 5.3.5 *Sean R, S operadores lineales sobre el espacio con producto interno V de dimensión finita. Entonces*

1. $(R+S)^* = R^* + S^*$;
2. $(\alpha R)^* = \overline{\alpha} R^*$;

3. $(RS)^* = S^*R^*$;

4. $(R^*)^* = R$;

5. *Si R es un isomorfismo, entonces* $(R^{-1})^* = (R^*)^{-1}$.

Demostración. Demostramos 1. Sean $x, y \in V$. Entonces

$$\begin{aligned}\langle (R+S)\, x, y\rangle &= \langle Rx + Sx, y\rangle = \langle Rx, y\rangle + \langle Sx, y\rangle = \langle x, R^*y\rangle + \langle x, S^*y\rangle \\ &= \langle x, R^*y + S^*y\rangle = \langle x, (R^* + S^*)\, y\rangle\end{aligned}$$

Por lo tanto, $(R+S)^* = R^* + S^*$. ■

Sabemos que si $\mathcal{B}$ y $\mathcal{C}$ son bases del espacio vectorial V y $T \in \mathcal{L}(V)$, entonces existe una matriz invertible P, la matriz cambio de base de $\mathcal{B}$ a $\mathcal{C}$, tal que

$$[T]_{\mathcal{B}} = P^{-1} [T]_{\mathcal{C}} P$$

Esto nos llevó a considerar la relación de semejanza sobre $\mathcal{M}_n(\mathbb{F})$:

$$A \text{ es semejante a } B \Leftrightarrow B = P^{-1}AP$$

donde $P \in \mathcal{M}_n(\mathbb{F})$ es una matriz invertible. Veamos ahora qué ocurre si suponemos además que $\mathcal{B}$ y $\mathcal{C}$ son bases ortonormales de un espacio con producto interior V. Si $\mathcal{B} = \{v_1, \dots, v_n\}$, entonces recordamos que la columna j de P viene dada por $[v_j]_{\mathcal{C}}$, la coordenada de v_j con respecto a la base $\mathcal{C}$. Esto es,

$$[v_j]_{\mathcal{C}} = (\langle v_j, w_1\rangle, \dots, \langle v_j, w_n\rangle)^\top$$

donde $\mathcal{C} = \{w_1, \dots, w_n\}$. Luego la fila i de P^* es $\left(\overline{\langle v_i, w_1\rangle}, \dots, \overline{\langle v_i, w_n\rangle}\right)$ y, por lo tanto,

$$\begin{aligned}[P^*P]_{ij} &= \overline{\langle v_i, w_1\rangle}\, \langle v_j, w_1\rangle + \cdots + \overline{\langle v_i, w_n\rangle}\, \langle v_j, w_n\rangle \\ &= \langle v_j, \langle v_i, w_1\rangle\, w_1\rangle + \cdots + \langle v_j, \langle v_i, w_n\rangle\, w_n\rangle \\ &= \langle v_j, \langle v_i, w_1\rangle\, w_1 + \cdots + \langle v_i, w_n\rangle\, w_n\rangle \\ &= \langle v_j, v_i\rangle = \left\{\begin{array}{l} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{array}\right|\end{aligned}$$

En consecuencia, $P^*P = I_n$.

Definición 5.3.6 *Una matriz $U \in \mathcal{M}_n(\mathbb{F})$ es* ***unitaria*** *si $U^*U = I_n$. Si $U^\top U = I_n$ decimos que U es* ***ortogonal****.*

Es claro que si $U \in \mathcal{M}_n(\mathbb{R})$, entonces $U^* = U^\top$ y así, matrices unitarias y ortogonales coinciden.

Teorema 5.3.7 *Sea $U \in \mathcal{M}_n(\mathbb{F})$. Las siguientes condiciones son equivalentes:*

1. *U es unitaria ;*
2. *U es invertible y $U^{-1} = U^*$;*
3. *U^* es unitaria;*
4. *Las columnas de U forman una base ortonormal de $\mathbb{F}^n$;*
5. *Las filas de U forman una base ortonormal de $\mathbb{F}^n$;*

Demostración. $1 \Rightarrow 2$. Si U es unitaria, entonces $U^*U = I_n$. Luego por el ejercicio 4, $UU^* = I_n$. En consecuencia, U y U^* son invertibles y $U^* = U^{-1}$.

$2 \Leftrightarrow 1$. Es claro.

$1 \Leftrightarrow 3$. Esto es consecuencia de $U^{**} = U$ y la equivalencia $U^*U = I \Leftrightarrow UU^* = I$.

$1 \Leftrightarrow 4$. El elemento ij de la matriz U^*U es

$$[U^*U]_{ij} = (U^*)_{i*} (U)_{*j} = \left(\overline{U}\right)_{*i} (U)_{*j}$$

Por lo tanto, $U^*U = I$ si, y sólo si, $\left(\overline{U}\right)_{*i} (U)_{*j} = \delta_{ij}$, donde δ_{ij} es el delta de Kronecker.

$3 \Leftrightarrow 5$ Es similar. ∎

Tenemos también una versión de este teorema para matrices reales ortogonales sustituyendo unitaria por ortogonal, U^* por $U^\top$ y $\mathbb{F}^n$ por $\mathbb{R}^n$.

Como vimos antes, si $\mathcal{B}$ y $\mathcal{C}$ son bases ortonormales de V y $T \in \mathcal{L}(V)$, entonces

$$[T]_{\mathcal{B}} = P^* [T]_{\mathcal{C}} P$$

donde P es una matriz unitaria. Esto sugiere introducir una nueva relación de equivalencia sobre $\mathcal{M}_n(\mathbb{F})$.

Definición 5.3.8 *Sean $A, B \in \mathcal{M}_n(\mathbb{F})$. Decimos que la matriz A es **unitariamente equivalente** a la matriz B si existe una matriz unitaria $U \in \mathcal{M}_n(\mathbb{F})$ tal que $B = U^*AU$. Decimos que A es **ortogonalmente equivalente** a B si existe una matriz O ortogonal tal que $B = O^\top AO$.*

Por el teorema anterior y el hecho que producto de matrices unitarias (resp. ortogonales) es unitaria (resp. ortogonal) (ver ejercicio 6), se deduce que las equivalencias unitaria y ortogonal son relaciones de equivalencia.

Demostramos que las matrices asociadas a un operador lineal con respecto a dos bases ortonormales son unitariamente equivalentes. El recíproco también es cierto, como vemos en nuestro próximo resultado.

Teorema 5.3.9 *Sean $A, B \in \mathcal{M}_n(\mathbb{F})$. Las matrices A y B son unitariamente equivalentes si, y sólo si, A y B son matrices asociadas a un operador lineal de $\mathcal{L}(V)$ con respecto a dos bases ortonormales de V.*

Demostración. Una de las implicaciones ya fue demostrada. Supongamos que U es una matriz unitaria y $B = U^*AU$. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de un espacio vectorial V y $T \in \mathcal{L}(V)$ tal que $[T]_{\mathcal{B}} = A$. Por el teorema 4.4.11, la base $\mathcal{C} = \{w_1, \ldots, w_n\}$ que satisface

$$[w_i]_{\mathcal{B}} = Ue_i$$

para cada $i = 1, \ldots n$, cumple con $[T]_{\mathcal{C}} = B$. Sólo debemos demostrar que $\mathcal{C}$ es una base ortonormal de V. Para cada i tenemos que

$$w_i = u_{1i}v_1 + u_{2i}v_2 + \cdots + u_{ni}v_n$$

donde $Ue_i = (u_{1i}, u_{2i}, \ldots, u_{ni})^\top$ es la columna i de la matriz unitaria U. Luego al ser $\mathcal{B}$ una base ortonormal de V se deduce que

$$\langle w_i, w_j \rangle = \left\langle \sum_{k=1}^{n} u_{ki}v_k, \sum_{l=1}^{n} u_{lj}v_l \right\rangle = \sum_{r=1}^{n} u_{ri}\overline{u_{rj}} = \langle U_{*i}, U_{*j} \rangle$$

Como U es unitaria, se deduce del teorema 5.3.7 que $\{U_{*1}, \ldots, U_{*n}\}$ es una base ortonormal de $\mathbb{F}^n$. En consecuencia, $\mathcal{C}$ es una base ortonormal de V. ■

El resultado para equivalencia ortogonal es similar.

Teorema 5.3.10 *Sean $A, B \in \mathcal{M}_n(\mathbb{R})$. Las matrices A y B son ortogonalmente equivalentes si, y sólo si, A y B son matrices asociadas a un operador lineal real de $\mathcal{L}(V)$ con respecto a dos bases ortonormales de $\mathbb{R}^n$.*

Demostración. Dejamos la demsotración al lector. ■

Definición 5.3.11 *Sea $T \in \mathcal{L}(V)$. Decimos que T es* ***unitario*** *si $TT^* = T^*T = I_V$.*

Demostramos en nuestro próximo resultado que en el isomorfismo $\mathcal{L}(V) \to \mathcal{M}_n(\mathbb{F})$ definido por $T \rightsquigarrow [T]_{\mathcal{B}}$, los operadores unitarios se corresponden con las matrices unitarias.

Teorema 5.3.12 *Sea $T \in \mathcal{L}(V)$ y $\mathcal{B}$ una base ortonormal de V. T es unitario si, y sólo si, $[T]_{\mathcal{B}}$ es unitaria.*

Demostración. Si T es unitario entonces $TT^* = T^*T = I_V$. Luego por el teorema 5.3.2,

$$[T]_{\mathcal{B}} [T]_{\mathcal{B}}^* = [T]_{\mathcal{B}} [T^*]_{\mathcal{B}} = [TT^*]_{\mathcal{B}} = [I_V]_{\mathcal{B}} = I$$

Luego, $[T]_{\mathcal{B}}$ es unitaria. Recíprocamente, supongamos que $[T]_{\mathcal{B}}$ es una matriz unitaria. Entonces, de nuevo usando el teorema 5.3.2 obtenemos

$$[TT^*]_{\mathcal{B}} = [T]_{\mathcal{B}} [T^*]_{\mathcal{B}} = [T]_{\mathcal{B}} [T]_{\mathcal{B}}^* = I = [I_V]_{\mathcal{B}}$$

En consecuencia, $TT^* = I_V$ y así, T es un operador unitario. ■

Finalizamos esta sección con una caracterización de los operadores unitarios.

Teorema 5.3.13 *Sea V un espacio con producto interno de dimensión finita y $T \in \mathcal{L}(V)$. Las siguientes condiciones son equivalentes:*

1. *T es un operador unitario;*
2. *T envía bases ortonormales en bases ortonormales;*
3. *$\|Tx\| = \|x\|$ para todo $x \in V$;*
4. *$\langle Tx, Ty\rangle = \langle x, y\rangle$ para todo $x, y \in V$.*

Demostración. $1 \Rightarrow 2$ Primero que todo observamos que si T es unitario, entonces T es inyectiva. En efecto, si $Tx = 0$, entonces $\langle x, x\rangle = \langle Tx, Tx\rangle = 0$ lo que implica $x = 0$. Se deduce del corolario 4.2.7 que T es un isomorfismo. Luego si $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base ortonormal de V, entonces por el teorema 4.3.5, $\{Tv_1, \ldots, Tv_n\}$ es una base de V. Falta ver que es ortonormal. Pero esto se deduce del hecho de que T es unitario: para todo i, j tenemos

$$\langle Tv_i, Tv_j\rangle = \langle v_i, T^*Tv_j\rangle = \langle v_i, v_j\rangle = \delta_{ij}$$

$2 \Rightarrow 3$ Supongamos que $\{v_1, \ldots, v_n\}$ y $\{Tv_1, \ldots, Tv_n\}$ son bases ortonormales de V. Entonces para todo i, j

$$\langle v_i, v_j\rangle = \delta_{ij} = \langle Tv_i, Tv_j\rangle$$

Luego, si $x \in V$ entonces $x = \sum \alpha_i v_i$ y así

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \left\langle \sum \alpha_i Tv_i, \sum \alpha_j Tv_j \right\rangle = \sum \alpha_i \overline{\alpha_i} = \langle x, x\rangle = \|x\|^2$$

$3 \Rightarrow 4$ Sean $x, y \in V$. Entonces

$$\langle T(x+y), T(x+y)\rangle = \langle x+y, x+y\rangle$$

lo que implica

$$\langle Tx, Ty\rangle + \overline{\langle Tx, Ty\rangle} = \langle x, y\rangle + \overline{\langle x, y\rangle}$$

De aquí se deduce que $Re(\langle Tx, Ty\rangle) = Re(\langle x, y\rangle)$. Por otra parte,

$$\langle T(x+iy), T(x+iy)\rangle = \langle x+iy, x+iy\rangle$$

Luego

$$-i\langle Tx, Ty\rangle + i\overline{\langle Tx, Ty\rangle} = -i\langle x, y\rangle + i\overline{\langle x, y\rangle}$$

y, en consecuencia,

$$Im(\langle Tx, Ty\rangle) = Im(\langle x, y\rangle)$$

$4 \Rightarrow 1$ Sea $x \in V$. Entonces para todo $y \in V$ tenemos

$$\begin{aligned} \langle T^*Tx - x, y\rangle &= \langle T^*Tx, y\rangle - \langle x, y\rangle = \langle Tx, Ty\rangle - \langle x, y\rangle \\ &= \langle x, y\rangle - \langle x, y\rangle = 0 \end{aligned}$$

Por lo tanto, $T^*Tx = x$. ■

Se desprende de este teorema que un operador unitario T preserva las distancias:

$$\|Tx - Ty\| = \|x - y\|$$

para todo $x, y \in V$. Por esta razón, T se también recibe el nombre de isometría.

EJERCICIOS

1. Encuentre la matriz cambio de base de

$$\{(1,0,0),(0,1,0),(0,0,1)\} \text{ a } \left\{\left(\frac{i}{\sqrt{2}},0,\frac{-1}{\sqrt{2}}\right),(0,1,0),(0,0,i)\right\}$$

en $\mathbb{C}^3$. Verifique que la matriz obtenida es unitaria.

2. Encuentre el adjunto de cada uno de los operadores lineales siguientes:

$a)$ $S:\mathbb{R}^3 \to \mathbb{R}^3$ definido por $S\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x+4y-5z \\ 2x-6y+7z \\ 5x-9y+z \end{pmatrix}$;

$b)$ $T:\mathbb{C}^3 \to \mathbb{C}^3$ definido por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+(1-i)\,y \\ (3+2i)\,x-4iz \\ 2ix+(4-3i)\,y-3z \end{pmatrix}$.

3. Sea $V=\{f(x)\in\mathbb{R}[x]:\ \text{grado de } f(x) \text{ es menor o igual a } 3\}$ con el producto interno

$$\langle f,g\rangle = \int_0^1 f(x)\,g(x)\,dx$$

Encuentre el adjunto del operador derivación D.

4. Sea $V=\mathcal{M}_n(\mathbb{C})$ con el producto interno $\langle A,B\rangle = Tr(AB^*)$. Sea P una matriz invertible en V y sea $\Gamma_P : V\to V$ definido por $\Gamma_P(A)=P^{-1}AP$. Encuentre el adjunto de Γ_P.

5. Sea $V=\mathcal{M}_n(\mathbb{C})$ con el producto interno $\langle A,B\rangle = Tr(AB^*)$. Para cada M en V, sea $\Psi_M : V\to V$ definida por $\Psi_M(A)=MA$. Demuestre que Ψ_M es unitario si, y sólo si, M es una matriz unitaria.

6. Suponga que $A,B\in\mathcal{M}_n(\mathbb{C})$ son unitarias. Demuestre que A^{-1} y AB son unitarias. Similarmente para matrices ortogonales.

7. Demuestre que $T\in\mathcal{L}(\mathbb{R}^2)$ es unitario si, y sólo si, T es una rotación o una reflexión sobre el eje x seguida de una rotación. (Sugerencia: si e_1,e_2 es la base canónica de $\mathbb{R}^2$ entonces $u_1=Te_1, u_2=Te_2$ es una base ortonormal de $\mathbb{R}^2$. Existe $0\le\theta\le 2\pi$ tal que $u_1=\begin{pmatrix} cos\theta \\ sen\theta \end{pmatrix}$. Como u_2 es ortogonal a u_1, entonces $u_2=\begin{pmatrix} cos\left(\theta\pm\frac{\pi}{2}\right) \\ sen\left(\theta\pm\frac{\pi}{2}\right) \end{pmatrix}$).

8. Demuestre que si $T^*T=0$, entonces $T=0$.

9. Sea $T \in \mathcal{L}(V)$ y suponga que W es un subespacio de V tal que $T(W) \subseteq W$. Demuestre que $T^*\left(W^\top\right) \subseteq W^\top$.

10. Sea T un operador unitario sobre V y W un subespacio de V tal que $T(W) \subseteq W$. Demuestre que $T\left(W^\top\right) \subseteq W^\top$.

11. Sea V un espacio con producto interno de dimensión finita. Demuestre que

$$Im\,(T^*) = (kerT)^\top$$

En particular, $rgo\,(T) = rgo\,(T^*)$.

12. Demuestre que las relaciones equivalencia unitaria y ortogonal son relaciones de equivalencia.

5.4. Diagonalización unitaria

En secciones anteriores estudiamos el problema de decidir cuándo un operador lineal $T \in \mathcal{L}(V)$ (o, equivalentemente una matriz) es diagonalizable, donde V es un espacio vectorial. En esta sección nos planteamos el mismo problema pero con la condición adicional de que V es un espacio con producto interno. En este caso es posible elegir una matriz asociada al operador lineal $T \in \mathcal{L}(V)$ con respecto a una base ortonormal. Luego la pregunta que nos planteamos es la siguiente: ¿cuándo existe una base ortonormal de V de tal manera que la matriz asociada a T es diagonal?

Definición 5.4.1 *Sea V un espacio con producto interno. Decimos que $T \in \mathcal{L}(V)$ es **unitariamente diagonalizable** si existe una base ortonormal $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ es diagonal.*

Esta sección tiene como objetivo determinar cuándo un operador $T \in \mathcal{L}(V)$ es unitariamente diagonalizable. Como antes, trasladamos el problema a las matrices.

Teorema 5.4.2 *Sea V un espacio con producto interno, $T \in \mathcal{L}(V)$ y A una matriz asociada a T con respecto a una base ortonormal de V. Entonces T es unitariamente diagonalizable si, y sólo si, A es unitariamente equivalente a una matriz diagonal.*

Demostración. Sea $\mathcal{B}$ una base ortonormal de V tal que $[T]_{\mathcal{B}} = A$. Si T es unitariamente diagonalizable, entonces existe una base ortonormal $\mathcal{C}$ de V tal que $[T]_{\mathcal{C}} = D$, donde D es una matriz diagonal. Por el teorema 5.3.9,

$$D = [T]_{\mathcal{C}} = P^*\,[T]_{\mathcal{B}}\,P$$

donde P es una matriz unitaria. Así, A es unitariamente equivalente a una matriz diagonal. Recíprocamente, si $A = [T]_{\mathcal{B}}$ es unitariamente equivalente a una matriz diagonal D, entonces, por el teorema 5.3.9 existe una base ortonormal $\mathcal{C}$ de V tal que $[T]_{\mathcal{C}} = D$. Esto demuestra que T es diagonalizable. ■

Definición 5.4.3 *Decimos que una matriz $A \in \mathcal{M}_n(\mathbb{F})$ es **unitariamente (ortogonalmente) diagonalizable** si A es unitariamente (ortogonalmente) equivalente a una matriz diagonal.*

Uno de los resultados más importantes en la teoría elemental de matrices establece que cualquier matriz $A \in \mathcal{M}_n(\mathbb{F})$ es unitariamente equivalente a una matriz triangular superior que tiene los autovalores de A en la diagonal.

Teorema 5.4.4 *(Schur) Sea $A \in \mathcal{M}_n(\mathbb{F})$ con autovalores $\lambda_1, \ldots, \lambda_n$. Existe una matriz unitaria $U \in \mathcal{M}_n(\mathbb{F})$ tal que $U^{-1}AU = T$, donde T es una matriz triangular superior con elementos sobre la diagonal $[T]_{ii} = \lambda_i$ para cada $i = 1, \ldots, n$. Además, si $A \in \mathcal{M}_n(\mathbb{R})$ y si todos los autovalores de A son reales, entonces U la tomamos real ortogonal.*

Demostración. Sea $x_1 \in \mathbb{F}^n$ un autovector de A asociado a λ_1 y supongamos que $\|x_1\| = 1$. Extendemos este vector a una base de $\mathbb{F}^n$ y luego aplicamos el proceso de Gram-Schmidt para obtener una base ortonormal $x_1, y_2, \ldots, y_n$ de $\mathbb{F}^n$. Consideremos la matriz $U_1 \in \mathcal{M}_n(\mathbb{F})$ que tiene primera columna a x_1 y para $2 \leq j \leq n$, la columna j de U_1 es y_j. Entonces, un cálculo simple muestra que

$$U_1^{-1}AU_1 = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}$$

donde $A_1 \in \mathcal{M}_{n-1}(\mathbb{F})$ tiene autovalores $\lambda_2, \ldots, \lambda_n$. Sea $x_2 \in \mathbb{F}^{n-1}$ un autovector de A_1 asociado al autovalor λ_2 y supongamos que $\|x_2\| = 1$. Por el mismo procedimiento anterior construimos una matriz unitaria $U_2 \in \mathcal{M}_{n-1}(\mathbb{F})$ tal que

$$U_2^{-1}A_1U_2 = \begin{pmatrix} \lambda_2 & * \\ 0 & A_2 \end{pmatrix}$$

donde $A_2 \in \mathcal{M}_{n-2}(\mathbb{F})$ tiene autovalores $\lambda_3, \ldots, \lambda_n$. Sea $W_2 = \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}$. Entonces U_1W_2 es unitaria y

$$\begin{aligned}
(U_1W_2)^{-1} AU_1W_2 &= W_2^{-1}U_1^{-1}AU_1W_2 = W_2^{-1} \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix} W_2 \\
&= \begin{pmatrix} \begin{matrix} \lambda_1 & * \\ 0 & \lambda_2 \end{matrix} & * \\ 0 & A_2 \end{pmatrix}
\end{aligned}$$

Continuamos este procedimiento para obtener matrices unitarias $U_k \in \mathcal{M}_{n-(k-1)}(\mathbb{F})$ y matrices $W_k \in \mathcal{M}_n(\mathbb{F})$ para $k = 1, \ldots, n-1$. La matriz $U = U_1W_2 \cdots W_{n-1}$ es unitaria y $U^{-1}AU$ tiene la forma deseada.

Si $A \in \mathcal{M}_n(\mathbb{R})$ y si todos los autovalores de A son reales, entonces los autovectores correspondientes los tomamos reales y el procedimiento anterior se repite. ■

Notamos que en el Teorema de Schur, ni la matriz unitaria U ni la matriz triangular T son únicas, como vemos en el siguiente ejemplo.

Ejemplo 5.4.5 *Sea* $A = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$. *Entonces*

$$U^{-1}AU = \begin{pmatrix} 2 & -1 & 3\sqrt{2} \\ 0 & 1 & \sqrt{2} \\ 0 & 0 & 3 \end{pmatrix} \text{ e } I^{-1}AI = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

Una primera aplicación del Teorema de Schur demuestra que las matrices reales simétricas son ortogonalmente diagonalizables.

Definición 5.4.6 *Una matriz* $A \in \mathcal{M}_n(\mathbb{F})$ *es* ***hermitiana*** *si* $A^* = A$.

Ejemplo 5.4.7 *La matriz* $A = \begin{pmatrix} 1 & 2i & 1-i \\ -2i & 2 & 1+i \\ 1+i & 1-i & 0 \end{pmatrix}$ *es hermitiana.*

Teorema 5.4.8 *Los autovalores de una matriz hermitiana son reales. En particular, una matriz simétrica real es ortogonalmente equivalente a una matriz diagonal.*

Demostración. Sea A una matriz hermitiana y λ un autovalor de A. Sea $x \neq 0$ un autovector asociado. Como $Ax = \lambda x$, entonces $x^*A = \overline{\lambda}x^*$. Multiplicamos por x^* a la izquierda de la primera relación y por x a la derecha en la segunda para obtener

$$\begin{aligned} x^*Ax &= \lambda x^*x \\ x^*Ax &= \overline{\lambda}x^*x \end{aligned}$$

Se deduce que $\lambda x^*x = \overline{\lambda}x^*x$ y como $x \neq 0$, $\lambda = \overline{\lambda}$. Es decir, λ es un número real.

Para ver la segunda parte, si A es simétrica real (con autovalores reales), entonces $U^\top AU = T$ para una matriz real ortogonal U y una matriz triangular superior T. En particular, $U^\top AU$ es simétrica y triangular y por lo tanto diagonal. ■

Nos planteamos ahora el problema de determinar cuándo una matriz $A \in \mathcal{M}_n(\mathbb{F})$ es unitariamente diagonalizable.

Teorema 5.4.9 *Sea* $A \in \mathcal{M}_n(\mathbb{F})$ *con autovalores* $\lambda_1, \ldots, \lambda_n$. *Las siguientes condiciones son equivalentes:*

1. *A es unitariamente diagonalizable;*
2. $AA^* = A^*A$;
3. $\sum_{i,j} \left|[A]_{ij}\right|^2 = \sum_{i=1}^{n} |\lambda_i|^2$;
4. *Existe una base ortonormal de autovectores de* $\mathbb{F}^n$.

Demostración. $1 \Rightarrow 2$. Si $A \in \mathcal{M}_n(\mathbb{F})$ es unitariamente equivalente a la matriz diagonal D, entonces $U^*AU = D$, donde U es una matriz unitaria. Luego $A = UDU^*$ y $A^* = UD^*U^*$. De aquí se concluye que

$$\begin{aligned} AA^* &= UDU^*UD^*U^* = UDD^*U^* = UD^*DU^* \\ &= UD^*U^*UDU^* = A^*A \end{aligned}$$

y así, A es normal.

$2 \Rightarrow 1$. Por el Teorema de Schur existe una matriz unitaria U y una matriz triangular superior T tal que

$$U^*AU = T = \begin{pmatrix} \lambda_1 & t_{12} & t_{13} & \cdots & t_{1n} \\ 0 & \lambda_2 & t_{23} & \cdots & t_{2n} \\ 0 & 0 & \lambda_3 & \cdots & t_{3n} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{5.3}$$

donde $\lambda_1, \ldots, \lambda_n$ son los autovalores de A. Como U es unitaria y A es normal, entonces es fácil ver que T es normal. Por lo tanto,

$$(T^*T)_{11} = \overline{\lambda_1}\lambda_1 = (TT^*)_{11} = \overline{\lambda_1}\lambda_1 + \overline{t_{12}}t_{12} + \cdots + \overline{t_{1n}}t_{1n}$$

En consecuencia,

$$t_{12} = \cdots = t_{1n} = 0$$

Análogamente,

$$(T^*T)_{22} = \overline{\lambda_2}\lambda_2 = (TT^*)_{22} = \overline{\lambda_2}\lambda_2 + \overline{t_{23}}t_{23} + \cdots + \overline{t_{2n}}t_{2n}$$

y por lo tanto,

$$t_{23} = \cdots = t_{2n} = 0$$

Continuando así concluimos que $T = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

$1 \Rightarrow 3$. Supongamos que $D = U^*AU$, donde U es unitaria y D diagonal. Entonces, usando el hecho de que la traza es invariante por semejanza obtenemos

$$\begin{aligned} \sum_{i=1}^{n} |\lambda_i|^2 &= Tr\,(D^*D) = Tr\,(U^*A^*UU^*AU) \\ &= Tr\,(U^*A^*AU) = Tr\,(A^*A) = \sum_{i,j} \left|[A]_{ij}\right|^2 \end{aligned}$$

$3 \Rightarrow 1$. Sabemos por el Teorema de Schur que existe una matriz unitaria U tal que $U^*AU = T$, donde T tiene la forma dada en (5.3). Luego $A = UTU^*$ y $A^* = UT^*U^*$ y, por lo tanto, $AA^* = UTT^*U^*$. Se deduce que

$$\begin{aligned} \sum_{i=1}^{n} |\lambda_i|^2 &= \sum_{i,j} \left|[A]_{ij}\right|^2 = Tr\,(AA^*) = Tr\,(TT^*) \\ &= \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i<j}^{n} |t_{ij}|^2 \end{aligned}$$

En consecuencia, $\sum\limits_{i<j}^{n} |t_{ij}|^2 = 0$ y así, $t_{ij} = 0$ para todo $i < j$. Esto demuestra que T es diagonal.

$1 \Rightarrow 4$. Supongamos que $AU = UD$, donde $D = diag(\lambda_1, \ldots, \lambda_n)$. Por el teorema 5.3.7, las columnas de U forman una base ortonormal de $\mathbb{C}^n$. Además,

$$A[U]_{*j} = [AU]_{*j} = [UD]_{*j} = U[D]_{*j} = U\lambda_j e_j = \lambda_j [U]_{*j}$$

lo que demuestra que las columnas de U son autovectores de A.

$4 \Rightarrow 1$. Es claro. ■

Definición 5.4.10 *Una matriz $A \in \mathcal{M}_n(\mathbb{F})$ es **normal** si $AA^* = A^*A$.*

Es claro que las matrices unitarias y las matrices hermitianas son matrices normales. También lo son las matrices antihermitianas: $A \in \mathcal{M}_n(\mathbb{F})$ es **antihermitiana** si $A^* = -A$.

Ejemplo 5.4.11 *La matriz $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ es normal pero no es unitaria, hermitiana ni antihermitiana.*

EJERCICIOS

1. Considere las matrices

$$\begin{aligned} A &= \frac{1}{2}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix}, \; B = \begin{pmatrix} 3 & 1-2i & 4+7i \\ 1+2i & -4 & -2i \\ 4-7i & 2i & 2 \end{pmatrix} \text{ y} \\ C &= \begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix} \end{aligned}$$

   Demuestre que A es unitaria, B es hermitiana y C es normal.

2. Sea $A \in \mathcal{M}_2(\mathbb{R})$ una matriz normal. Demuestre que A es simétrica o A es suma de una matriz múltiplo de la identidad y una antisimétrica.

3. Suponga que $A \in \mathcal{M}_n(\mathbb{C})$. Demuestre que AA^* y A^*A son ambas hermitianas.

4. Suponga que $A \in \mathcal{M}_n(\mathbb{C})$. Demuestre que $A+A^*$ es hermitiana y $A-A^*$ es antihermitiana. En particular, cualquier matriz de $\mathcal{M}_n(\mathbb{C})$ se escribe como suma de una matriz hermitiana y una antihermitiana.

5. Suponga que $A, U \in \mathcal{M}_n(\mathbb{C})$, A es normal y U es unitaria. Demuestre que U^*AU es normal.

6. Calcule una forma de Schur para cada una de las siguientes matrices:

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

7. Dada la matriz

$$A = \begin{pmatrix} 1 & -14 & -10 \\ -14 & -2 & -4 \\ -10 & -4 & 10 \end{pmatrix}$$

encuentre una matriz ortogonal Q tal que $Q^\top AQ$ es diagonal.

8. Compruebe que la matriz $A = \begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix}$ es normal. Encuentre una matriz unitaria U tal que U^*AU es diagonal.

9. Sea V un espacio con producto interno y $T \in \mathcal{L}(V)$. Defina

   a) T es un **operador hermitiano o autoadjunto** si $T^* = T$;

   b) T es un **operador unitario** si $T^* = T^{-1}$;

   c) T es un **operador normal** si $TT^* = T^*T$.

   Sea $\mathcal{B}$ una base ortonormal del espacio V sobre $\mathbb{F}$. Demuestre que $T \in \mathcal{L}(V)$ es hermitiano (resp. unitario o normal) si, y sólo si, $[T]_{\mathcal{B}}$ es una matriz hermitiana (resp. unitaria o normal).