Introducción al Álgebra Lineal

Juan Rada
Departamento de Matemáticas
Facultad de Ciencias
Universidad de Los Andes

Primera versión: 19 de febrero del 2005
Segunda versión: 17 de junio del 2005

Índice general

Introducción v

1. Matrices 1
1.1. Definición y terminología 1
1.2. Álgebra de matrices 4
1.3. Sistemas de ecuaciones lineales 11
1.4. Matrices elementales 17
1.5. Equivalencia de matrices 20

2. Espacios vectoriales 25
2.1. El espacio de las n-uplas 25
2.2. Espacios vectoriales 27
2.3. Subespacios 30
2.4. Subespacios generados 33
2.5. Independencia lineal y bases 38
2.6. Dimensión 41
2.7. El rango y la nulidad de una matriz 46

3. Transformaciones lineales 53
3.1. Transformaciones lineales 53
3.2. El núcleo y la imagen 59
3.3. Isomorfismos 63
3.4. Matriz asociada a una transformación lineal 67
3.5. Cambio de base y semejanza de matrices 70

4. Espacios con producto interno 75
4.1. Producto interno 75
4.2. Bases ortonormales 81
4.3. Operadores lineales 85

5. Determinantes 93
5.1. Permutaciones 93
5.2. Determinantes 96
5.3. Funciones n-lineales y alternadas 100

5.4. La expansión de Laplace 102
5.5. La adjunta y la inversa de una matriz 104

6. Diagonalización **109**
6.1. Autovalores y autovectores 109
6.2. Diagonalización 114
6.3. Diagonalización unitaria 118

Introducción

El propósito de escribir este libro es presentar un texto introductorio de álgebra lineal para la asignatura Álgebra 1 de la Licenciatura de Matemáticas en la Universidad de Los Andes. Está dirigido a estudiantes que por primera vez estudian álgebra lineal pero están familiarizados con los conceptos básicos de vectores en el plano y en el espacio, estudiados previamente en cursos de cálculo y geometría analítica.

1. Determinante del producto igual al producto de determinantes puede ir después de las operaciones elementales?

2. La exposición debe ser clean, direct and clear!

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), donde $a, b \in \mathbb{R}$. Decimos que dos números complejos son iguales cuando son iguales coordenada a coordenada:

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

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 a continuación 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. $(a,b)\left[(c,d)+(e,f)\right] = (a,b)(c,d) + (a,b)(e,f)$;

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

Demostración. Probamos 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

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

Si representamos a z en el plano real (ver Figura *) observamos que $z\overline{z}$ es precisamente la longitud al cuadrado del vector $0P$. Denotamos por $|z|$ a esta longitud, es decir,

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

y la llamamos la norma del número complejo z.

Figura *

Si θ es el ángulo que hace el vector $0P$ 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} es el elemento 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).

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 $A + (-B)$.

Teorema 1.2.3 *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. 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 elementos 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 1.2.4 *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 1.2.5 *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 manera 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 1.2.6 *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 1.2.7 *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. Probamos 1. 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 1.2.8 *Si* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ *entonces* $I_m A = A$ *y* $AI_n = A$.

Demostración. Probamos 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 1.2.9 *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 1.2.10 *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 1.2.11 *Sean $A, B \in \mathcal{M}_n(\mathbb{F})$ invertibles. Entonces*

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

Demostración. Probamos 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 entero mayores o iguales a 1, entonces se prueba por inducción que $A^r A^s = A^s A^r = A^{r+s}$.

Definición 1.2.12 *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 1.2.13 *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 1.2.14 *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. 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, la inversa de $A^\top$ es $\left(A^{-1}\right)^\top$. ■

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

Definición 1.2.15 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. La conjungada 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 1.2.16 *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 1.2.17 *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* $(A^*)^{-1} = \left(A^{-1}\right)^*$.

Demostración. 4. $(AB)^* = \overline{(AB)^\top} = \overline{B^\top A^\top} = \overline{B^\top}\,\overline{A^\top} = B^* A^*$.
5. Supongamos que A es invertible. Entonces por la parte 4 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. 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:
$$\text{a) } e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \text{ b) } e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}; \text{ c) } e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ y c) } e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$
5. Encuentre para la matriz B del ejercicio anterior $e_i B$ donde a) $e_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$; b) $e_2 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$ y c) $e_3 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$.

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

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

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

9. 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.

10. 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.

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

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

13. 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?

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

15. 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^*.

16. 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.

17. 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)* $Tr(AB) = Tr(BA)$.

18. 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 $AX = B$. Si $[A]_{ij} = a_{ij}$ y $[B]_{ij} = b_i$ para todo $1 \leq i \leq m$ y $1 \leq j \leq n$, la ecuación $AX = B$ es equivalente a un sistema de m ecuaciones lineales con n incógnitas

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

donde $a_{ij}, b_i \in \mathbb{F}$ para todo $1 \leq i \leq m, 1 \leq j \leq n$ y $x_1, \ldots, x_n$ son las incógnitas o variables. 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 realizar 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 λ 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 aun, 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 $A \underset{f}{\sim} B$.

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

Ejemplo 1.3.2 *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 realizando 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*

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

Definición 1.3.3 *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*} *entonces todos los elementos de* A_{*r} *distinto de* a_{pr} *son ceros.*

El Ejemplo 1.3.2 muestra como 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* E *tal que* $A \underset{f}{\sim} E$.

Demostración. Supongamos que $[A]_{ij} = a_{ij}$ para todo i, j y que r es la primera columna con un elemento distinto de cero. Si $a_{1r} \neq 0$ entonces aplicamos sucesivamente las operaciones elementales por filas

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

colocando ceros en todos los elementos de la columna r debajo del 1 en la posición $1, r$. Si $a_{1r} = 0$ entonces existe algún $a_{pr} \neq 0$. En este caso aplicamos primero f_{1p} y luego procedemos como antes. Así obtenemos una matriz B con elementos $[B]_{ij} = b_{ij}$ que tiene las primeras $r-1$ columnas 0 y los elementos de la columna r son todos cero excepto b_{1r}, que es igual a 1. Ahora sea s la primera columna distinta de cero de la submatriz de B que resulta al eliminar la primera fila. Si $b_{2s} \neq 0$ entonces aplicamos sucesivamente

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

colocando ceros en todos los elementos de la columna s excepto la posición $2, s$ que es 1. Si $b_{2s} = 0$ entonces existe $q > 2$ tal que $b_{qs} \neq 0$. Aplicamos f_{2q} y procedemos como antes.

Después de un número finito de pasos es claro que llegamos a una matriz escalonada reducida. ■

Probamos más adelante que una matriz es equivalente por filas a una *única* matriz escalonada reducida.

Veamos ahora una aplicación de la eliminación de Gauss-Jordan para resolver la ecuación matricial $AX = B$, donde $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y $B \in \mathcal{M}_{m\times 1}(\mathbb{F})$.

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 prueba que $AX = B$. ■

Estos dos últimos teoremas reducen el problema de resolver una ecuación matricial de la forma $AX = B$ a una 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 una matriz (E, C). Por el resultado anterior, 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* $E_{i*} = 0$ *para todo* $r < i \leq n$ *y que* $e_{ik_i} = 1$ *es el primer elemento distinto de cero de la fila* E_{i*} *para todo* $1 \leq i \leq r$, *donde* $k_1 < \cdots < k_r$. *Entonces:*

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

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

2. $r = n$ *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* $(c_1, \ldots, c_n)^\top$.

3. $r < n$ *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, por la fórmula:*

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

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.

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

2. Si $r = n$ entonces como $1 \leq k_1 < k_2 < \cdots < k_n \leq n$ resulta $k_i = i$ 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 sigue inmediatamente que la única solución es $(c_1,\ldots,c_n)^{\top}$.

3. Si $r<n$ entonces $\{1,\ldots,n\}\setminus\{k_1,\ldots,k_r\}\neq\emptyset$. Fijemos un valor de 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 $j\in\{1,\ldots,n\}\setminus\{k_1,\ldots,k_r\}$ y por lo tanto calculamos

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

Seguimos hacia atrás tomando las ecuaciones $s=r-1,r-2,\ldots,1$, 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} :

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

■

Ejemplo 1.3.7 *Encontremos la solución del 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*

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

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 3. del Teorema, deducimos que la ecuación tiene infinitas soluciones porque el número de filas distintas de cero es menor que el número de columnas: $2 < 3$. De hecho, para cualquier valor $z = z_0$ encontramos la solución (x_0, y_0, z_0), donde $y_0 = \frac{1}{2}(b - 2a + 5z_0)$ y $x_0 = a - 2y_0 + 3z_0$. En ningún caso la ecuación tiene solución única.

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 la ecuación matricial $AX = B$ tiene infinitas soluciones. En otras palabras, si un sistema de ecuaciones lineales tiene más incógnitas que ecuaciones entonces existen infinitas soluciones.*

Demostración. Sea E una matriz escalonada reducida equivalente por filas a A. Entonces $(A, B) \underset{f}{\sim} (E, C)$, donde $C \in \mathcal{M}_{m\times 1}(\mathbb{F})$. Luego, $AX = B$ y $EX = C$ tienen la misma solución. Ahora, el número r de filas distintas de cero de E satisface $r \leq m < n$. Se sigue del Teorema 1.3.6 parte 3. que existen infinitas soluciones. ■

EJERCICIOS

1. Encuentre la matriz escalonada reducida equivalente por filas a la matriz:

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

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

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) infinita soluciones; y iii) ninguna solución.

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

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 $AX = 0$.

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. Las primeras dos partes las dejamos como ejercicio. Probaremos 3. Si $i \neq p$ entonces

$$\begin{aligned}
\left[E_{pq}(\lambda) A\right]_{ij} &= \left[f_{pq}(\lambda)(I_m) A\right]_{ij} = \sum_k \left[f_{pq}(\lambda)(I_m)\right]_{ik} [A]_{kj} \\
&= \left[f_{pq}(\lambda)(I_m)\right]_{ii} [A]_{ij} = [A]_{ij}
\end{aligned}$$

Por otra parte,

$$\left[E_{pq}(\lambda) A\right]_{pj} = \sum_k \left[f_{pq}(\lambda)(I_m)\right]_{pk} [A]_{kj}$$

Como $\left[f_{pq}(\lambda)(I_m)\right]_{pp} = 1$, $\left[f_{pq}(\lambda)(I_m)\right]_{pq} = \lambda$ y $\left[f_{pq}(\lambda)(I_m)\right]_{pk} = 0$ para $k \neq p, q$ se concluye que

$$\left[E_{pq}(\lambda) A\right]_{pj} = [A]_{pj} + \lambda [A]_{qj} = \left[f_{pq}(\lambda)(A)\right]_{pj}$$

■

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

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

Demostración. Si $p, q \in \{1, \ldots, m\}$ y $\lambda \in \mathbb{F}$ entonces por el Teorema 1.4.3

$$\begin{aligned}
E_p(\lambda) E_p\left(\frac{1}{\lambda}\right) &= E_p(\lambda)\left[f_p\left(\frac{1}{\lambda}\right)(I_m)\right] \\
&= f_p(\lambda)\left[f_p\left(\frac{1}{\lambda}\right)(I_m)\right] = I_m
\end{aligned}$$

Similarmente se prueba 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 *Un matriz A es invertible si, y sólo si, $A \underset{f}{\sim} I$.*

Demostración. Supongamos que A es una matriz invertible $n \times n$. Por el Teorema 1.3.4 existen matrices elementales $E_1, \ldots, E_p$ tales que $B = E_p \cdots E_1 A$ es escalonada reducida. Luego, por el Teorema 1.4.4, B es invertible y, en consecuencia, $B = I$ (ver ejercicio 8). Recíprocamente, si $A \underset{f}{\sim} I$ entonces existen matrices elementales $F_1, \ldots, F_q$ tales que $A = F_q \cdots F_1 I = F_q \cdots F_1$, esto es , A es invertible. ■

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 como construir matrices elementales $E_1, \ldots, E_k$ tales que $E_k \cdots E_1 A = I$. 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* $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$.

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

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

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

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

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

Finalizamos esta sección con una caracterización de la relación " $\underset{f}{\sim}$ " definida sobre las matrices.

Teorema 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, $Q = E_r \cdots E_1 I = E_r \cdots E_1$, 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.

1.5. Equivalencia de matrices

Muchas de las definiciones y teoremas de las dos secciones anteriores tienen su versión dual en términos de columnas.

Definición 1.5.1 *Una operación elemental por columnas aplicada a una matriz $A \in \mathcal{M}_{m \times n}(\mathbb{F})$ significa realizar una de las siguientes operaciones a la matriz A:*

1. *Intercambiar la columna p por la columna q (denotada por $c_{pq}(A)$);*

2. *Muliplicar por λ a la columna p (denotada por $c_p(\lambda)(A)$);*

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

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

Las demostraciones de los resultados por columnas se deducen de los correspondientes por filas teniendo en cuenta que las columnas de A son las traspuestas de las filas de $A^\top$. Luego, si c es una operación elemental por columna y f es la correspondiente por fila entonces $c(A) = \left[f\left(A^\top\right)\right]^\top$.

Teorema 1.5.2 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$*. Entonces*

1. $c_{pq}(A) = AE_{pq}$;
2. $c_p(\lambda)(A) = AE_p(\lambda)$;
3. $c_{pq}(\lambda)(A) = AE_{qp}(\lambda)$

   donde $E_{pq}, E_p(\lambda)$ *y* $E_{pq}(\lambda)$ *son las matrices elementales de la Definición 1.4.1.*

Demostración. Probamos 3 dejando como ejercicio las primeras dos partes. Por el Teorema 1.4.3 y el Teorema 1.2.14 tenemos que

$$\begin{aligned} c_{pq}(\lambda)(A) &= \left[f_{pq}(\lambda)\left(A^\top\right)\right]^\top = \left[E_{pq}(\lambda)A^\top\right]^\top \\ &= A\left[E_{pq}(\lambda)\right]^\top = AE_{qp}(\lambda) \end{aligned}$$

■

En consecuencia, realizar una operación elemental por columnas equivale a multiplicar a la derecha por una matriz elemental. De esta manera, $A \underset{c}{\sim} B$ si, y sólo si $B = AE_1 \cdots E_k$, donde las E_i son matrices elementales.

El resultado dual al Teorema 1.4.7 es el siguiente:

Teorema 1.5.3 *Sean* $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$*. Entonces* $A \underset{c}{\sim} B$ *si, y sólo si, existe una matriz invertible* $P \in \mathcal{M}_n(\mathbb{F})$ *tal que* $B = AP$*.*

Demostración. Supongamos que $A \underset{c}{\sim} B$. Por el Teorema 1.5.2, existen matrices elementales $E_1, \ldots, E_k$ tales que $B = AE_1 \cdots E_k$. Como las matrices elementales son invertibles entonces $P = E_1 \cdots E_k$ es invertible y $B = AP$. Recíprocamente, sea Q invertible tal que $B = AQ$. Entonces por el Teorema 1.4.5, Q es producto de matrices elementales y así, por el Teorema 1.5.2, $A \underset{c}{\sim} B$.

■

Si permitimos realizar operaciones elementales por filas y columnas es posible convertir una matriz en otra bien sencilla.

Ejemplo 1.5.4 *Consideremos la matriz*

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

Vimos en el Ejemplo 1.3.2 que

$$A \underset{f}{\sim} \begin{pmatrix} 1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Ahora efectuando la operación elemental por columna c_{34} obtenemos la matriz

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

Luego aplicamos $c_{41}(1), c_{51}(-2), c_{42}(-2), c_{52}(1), c_{53}(-3)$ para convertirla en

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

Definición 1.5.5 *Una operación elemental aplicada a una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ significa realizar una operación elemental por fila o por columna a la matriz A.*

Más aun, si B se obtiene a partir de A a través de una sucesión finita de operaciones elementales entonces decimos que A y B son equivalentes, y lo denotamos por $A \sim B$.

El procedimiento usado en el Ejemplo 1.5.4 funciona para cualquier matriz.

Teorema 1.5.6 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. Entonces A es equivalente a una matriz de la forma*

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

donde r es el número de filas distintas de cero en la matriz escalonada reducida equivalente por filas a A.

Demostración. Por el Teorema 1.3.4 existe una matriz escalonada reducida E tal que $A \underset{f}{\sim} E$. Intercambiando columnas transformamos E en una matriz de la forma

$$\begin{pmatrix} I_r & C \\ 0 & 0 \end{pmatrix}$$

donde r es el número de filas distintas de cero de E. Aplicando operaciones elementales por columnas de la forma $c_{pq}(\lambda)$ llevamos a esta matriz a una de la forma

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

■

Finalmente, presentamos una caracterización de la relación $\sim$ definida sobre las matrices.

Teorema 1.5.7 *Sean $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$. Entonces $A \sim B$ si, y sólo si, existen matrices invertibles $P \in \mathcal{M}_m(\mathbb{F})$ y $Q \in \mathcal{M}_n(\mathbb{F})$ tales que $B = PAQ$.*

Demostración. Si $A \sim B$ entonces por los Teoremas 1.4.3 y 1.5.2 existen matrices elementales $E_1, \ldots, E_k$ y $F_1, \ldots, F_l$ tales que $B = E_k \cdots E_1 A F_1 \cdots F_l$. Por el Teorema 1.4.4, las matrices $P = E_k \cdots E_1$ y $Q = F_1 \cdots F_l$ son invertibles y $B = PAQ$. Recíprocamente, supongamos que $B = PAQ$, donde P y Q son invertibles. Entonces por los Teoremas 1.4.7 y 1.5.3 se sigue que

$$A \underset{c}{\sim} AQ \underset{f}{\sim} PAQ = B$$

Esto prueba que $A \sim B$. ■

EJERCICIOS

1. 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 $AE_1, \ldots, E_k = I$.

2. Calcule la inversa de la matriz $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$ usando operaciones elementales por columnas.

3. Demuestre que la relación "equivalencia por columnas" es una relación de equivalencia sobre el espacio de las matrices.

4. Demuestre que $A \underset{c}{\sim} B$ si, y sólo si, $A^\top \underset{f}{\sim} B^\top$.

5. Encuentre la matriz equivalente a la matriz
$$A = \begin{pmatrix} 1 & 1 & 3 & 2 \\ 2 & 3 & 9 & 2 \\ 1 & 2 & 6 & 0 \\ 0 & 1 & 3 & -2 \end{pmatrix}$$
que tiene la forma dada en el Teorema 1.5.6.

6. Demuestre que la relación "equivalencia" es una relación de equivalencia sobre el espacio de las matrices.

Capítulo 2

Espacios vectoriales

2.1. El espacio de las n-uplas

En esta sección generalizamos 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 (ver Figura *).

Figura *

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 2.1.1 *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 2.1.2 *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 2.1.3 *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. Probamos 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 2.1.1 el espacio de n-uplas sobre $\mathbb{F}$. Es claro que si $n = 2$, $n = 3$ y $\mathbb{F} = \mathbb{R}$ entonces recuperamos los espacios conocidos $\mathbb{R}^2$ y $\mathbb{R}^3$.

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 2.1.4 *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

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

EJERCICIOS

1. Considere los vectores $x = (3+i, -1+i, 2), y = (1, -i, 2i)$ de $\mathbb{C}^3$ junto con los escalares $\alpha = 2$ y $\beta = -2i$. Calcule $\alpha x + \beta y$.

2. Complete la demostración del Teorema 2.1.3.

2.2. Espacios vectoriales

El espacio de las matrices y el espacio de las n-uplas son ejemplos de sistemas algebraicos donde 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 2.2.1 *Un espacio vectorial* V *sobre* $\mathbb{F}$ *es un conjunto no vacío* V *junto con dos operaciones. 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 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 2.2.2 *Cada uno de los siguientes conjuntos con las operaciones indicadas es un espacio vectorial.*

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

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

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

2. *El espacio de las matrices $\mathcal{M}_{m \times n}(\mathbb{F})$ 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. *$\mathbb{C}^n$ es un espacio vectorial sobre $\mathbb{R}$. La multiplicación escalar es la misma que en el Ejemplo 1 excepto que los escalares son tomados en $\mathbb{R}$.*

4. *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 $x \in X$.

5. *El conjunto $Suc(\mathbb{F})$ de todas las sucesiones en $\mathbb{F}$ con las operaciones*

$$\begin{aligned} (x_n) + (y_n) &= (x_n + y_n) \\ \alpha (x_n) &= (\alpha x_n) \end{aligned}$$

   donde $(x_n), (y_n) \in Suc(\mathbb{F})$ y $\alpha \in \mathbb{F}$. Este ejemplo es un caso particular del anterior: $Suc(\mathbb{F}) = \mathcal{F}(\mathbb{N}, \mathbb{F})$.

6. *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 2.2.3 *Sea V un espacio vectorial sobre $\mathbb{F}$. Entonces:*

1. *La identidad aditiva es única;*
2. *El inverso aditivo 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 2.2.4 *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 prueba 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. ■

EJERCICIOS

1. Demuestre que cada uno de los conjuntos con las operaciones indicadas en el Ejemplo 2.2.2 son espacios vectoriales.
2. Complete la demostración del Teorema 2.2.4.

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 $\alpha \in \mathbb{F}$.

2.3. Subespacios

En esta sección estudiamos los subconjuntos de un espacio vectorial V que a su vez tienen estructura de espacio vectorial, con las mismas operaciones definidas en V.

Definición 2.3.1 *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 2.3.2 *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, $(0, 1), (1, 2) \in S$ y, sin embargo, $(0, 1) + (1, 2) = (1, 3) \notin S$.

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. Las rectas y planos que pasan por el origen son subespacios de $\mathbb{R}^3$. Para ver esto, supongamos que $0 \neq u \in \mathbb{R}^3$. Entonces la recta que pasa por el origen generada por el vector u es

$$S = \{\lambda u : \lambda \in \mathbb{R}\}$$

Así, dados $x, y \in S$ existen $\lambda, \mu \in \mathbb{R}$ tales que $x = \lambda u$ y $y = \mu u$. Luego,

$$x + y = \lambda u + \mu u = (\lambda + \mu) u \in S$$

Además, si $\alpha \in \mathbb{R}$ entonces

$$\alpha x = \alpha (\lambda u) = (\alpha \lambda) u \in S$$

Esto demuestra que S es un subespacio de $\mathbb{R}^3$.

Por otra parte, el plano que pasa por el origen generado por los vectores $u, v \in \mathbb{R}^3$ (suponemos que v no pertenece a la recta generada por u) viene dado por

$$T = \{\lambda u + \mu v : \lambda, \mu \in \mathbb{R}\}$$

Dejamos al lector probar que T es un subespacio de $\mathbb{R}^3$ (ver Figura *).

Figura *

Más adelante demostramos que estos son los únicos subespacios no triviales de $\mathbb{R}^3$.

6. 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 $\mathbb{F}^n$.

7. 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 $\mathbb{F}^m$.

8. Sea $S = \left\{A \in \mathcal{M}_{m\times n}(\mathbb{F}) : A^\top = A\right\}$. Es decir, S es el conjunto formado por las matrices simétricas. Si $\alpha \in \mathbb{F}$ y $A, B \in S$ entonces por el Teorema 1.2.14
$$(A+B)^\top = A^\top + B^\top = A + B$$
y
$$(\alpha A)^\top = \alpha A^\top = \alpha A$$
Así, S es un subespacio de $\mathcal{M}_{m\times n}(\mathbb{F})$.

9. Consideremos
$$S = \{f \in \mathcal{F}([a,b], \mathbb{F}) : f \text{ es continua}\}.$$
Si $\alpha \in \mathbb{F}$ y $f, g \in S$ entonces sabemos que $f+g$ y αf son continuas. Luego, S es un subespacio de $\mathcal{F}([a,b], \mathbb{F})$.

10. Sea $S = \left\{(x_n) \in Suc(\mathbb{F}) : \lim_{n\to\infty} x_n = 0\right\}$. Si $\alpha \in \mathbb{F}$ y $(x_n), (y_n) \in S$ entonces
$$\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n = 0 + 0 = 0$$
y
$$\lim_{n\to\infty}(\alpha x_n) = \alpha \lim_{n\to\infty} x_n = \alpha 0 = 0$$
Esto prueba que S es un subespacio de $Suc(\mathbb{F})$.

11. 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: dado $x \in S$ tenemos que $0 = 0x \in S$.

Teorema 2.3.3 *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 $0 \in \bigcap_{i\in I} S_i$ y así, $\bigcap_{i\in I} S_i \neq \emptyset$. Sea $\alpha \in \mathbb{F}$ y $x, y \in \bigcap_{i\in I} S_i$. 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 $\bigcap_{i\in I} S_i$, lo que prueba que $\bigcap_{i\in I} S_i$ es un subespacio de V. ■

2.4. Subespacios generados

Si X es un subconjunto no vacío de un espacio vectorial V entonces, por el Teorema 2.3.3, la intersección de todos los subespacios de V que contienen a X es un subespacio de V.

Definición 2.4.1 *Sea V un espacio vectorial y X un subconjunto de V. La intersección de todos lo 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 2.4.2 *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 $\mathbb{F}$.

Teorema 2.4.3 *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, $\mathcal{C} \subseteq \mathcal{S}$. Probamos 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 2.4.4 *Veamos algunos ejemplos de subespacios generados por un subconjunto de un espacio vectorial.*

1. *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.

2. *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.

3. *Consideremos para cada $i = 1, \ldots, n$ los vectores $e_i \in \mathbb{F}^n$ que tienen todas las coordenadas 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 $\mathbb{F}^n$.

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

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

5. *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

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

6. *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:*

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

7. *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 2.4.5 *Un espacio vectorial V es finitamente generado si existe un número finito de vectores $v_1, \ldots, v_n$ de V tales que $V = gen(v_1, \ldots, v_n)$. En este caso, decimos que los vectores $v_1, \ldots, v_n$ generan a V.*

Ejemplo 2.4.6 *El espacio $\mathbb{F}^n$ es un espacio vectorial finitamente generado. De hecho, los vectores canónicos $e_1, \ldots, e_n$ genera a $\mathbb{F}^n$.*

Ejemplo 2.4.7 *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 $g(x) \notin gen(p_1(x), \ldots, p_n(x))$.*

Ejemplo 2.4.8 *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 9). 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 2.4.9 *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 2.4.10 *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 prueba 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 2.4.11 *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* $(0, 0, z) \in S_2$.

En el ejemplo anterior, la descomposición de un vector de $\mathbb{F}^3$ como suma de vectores de S_1 y S_2 no es única. Por ejemplo, también descomponemos a (x, y, z) como

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

donde $(x, 0, 0) \in S_1$ y $(0, y, z) \in S_2$. Es especialmente interesante el caso en que cada vector de un espacio vectorial se descompone de manera única como suma de vectores de cada uno de los subespacios.

Definición 2.4.12 *Decimos que* V *es suma directa de los subespacios* S_1 *y* S_2, *y escribimos* $V = S_1 \oplus S_2$, *si cada vector* $v \in V$ *se expresa de manera única como*

$$v = u_1 + u_2$$

donde cada $u_1 \in S_1$ *y* $u_2 \in S_2$.

Ejemplo 2.4.13 *Consideremos los subespacios de* $\mathbb{F}^3$

$$T_1 = \left\{(x, y, z) \in \mathbb{F}^3 : y = 0\ ,\ z = 0\right\} \text{ y } T_2 = \left\{(x, y, z) \in \mathbb{F}^3 : x = 0\right\}$$

Entonces $\mathbb{F}^3 = T_1 \oplus T_2$.

Un criterio útil para decidir si un espacio es suma directa de dos subespacios viene dado en el siguiente resultado.

Teorema 2.4.14 *Sea* V *un espacio vectorial y* S_1, S_2 *subespacios de* V. *Las siguientes condiciones son equivalentes:*

1. $V = S_1 \oplus S_2$;
2. $V = S_1 + S_2$ *y* $S_1 \cap S_2 = \{0\}$.

Demostración. $1 \Rightarrow 2$. Es claro que $V = S_1 + S_2$. Sea $x \in S_1 \cap S_2$. Entonces $x = 0 + x = x + 0$ y por la unicidad de la descomposición, $x = 0$. Esto prueba que $S_1 \cap S_2 = \{0\}$.

$2 \Rightarrow 1$. Sea $x \in V$ y supongamos que

$$\begin{aligned} x &= u_1 + u_2 \\ &= w_1 + w_2 \end{aligned}$$

donde $u_i, w_i \in S_i$ para todo $i = 1, 2$. Entonces

$$u_1 - w_1 = w_2 - u_2 \in S_1 \cap S_2 = \{0\}$$

Esto es, $u_j = w_j$ para todo $j = 1, 2$ lo que prueba que la descomposición es única. ■

EJERCICIOS

1. ¿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 $S = \{g \in \mathcal{F}(\mathbb{R}, \mathbb{R}) : g(1) = g(0)\}$;

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

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

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

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

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

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

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

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

7. 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 $\mathcal{C}(A)$?

8. 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.

9. 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 $S_2 \subseteq S_1$.

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.

2.5. 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 $\mathcal{F}(A)=gen\left[(1,0,0,1),(2,-2,2,-4)\right]$. Tratemos de precisar mejor esta idea.

Definición 2.5.1 *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 sigue 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 2.5.2 *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 porque*

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

2. *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 $\mathbb{F}^n$.

3. *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 $i=1,\ldots,n$.*

4. *Las matrices*

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

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, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$.

5. *Las funciones $f(t) = e^t$ y $g(t) = e^{2t}$ definidas sobre $[0,1]$ son funciones linealmente independientes de $\mathcal{F}([0,1], \mathbb{R})$. Supongamos que*

$$\alpha f + \beta g = \mathbf{0}$$

Entonces para todo $t \in [0,1]$ se tiene que

$$\alpha e^t + \beta e^{2t} = 0 \tag{2.1}$$

Derivando ambos lados de la igualdad obtenemos que

$$\alpha e^t + 2\beta e^{2t} = 0$$

y restando ambas ecuaciones llegamos a $\beta e^{2t} = 0$ para todo $t \in [0,1]$. Como $e^{2t} > 0$ concluimos que $\beta = 0$ lo que implica $\alpha = 0$.

Ahora estamos en condiciones de introducir uno de los conceptos fundamentales del álgebra lineal.

Definición 2.5.3 *Sea V un espacio vectorial. Si los vectores $v_1, \ldots, v_r$ de V generan a V y son linealmente independientes entonces decimos que los vectores $v_1, \ldots, v_r$ forman una base de V.*

Ejemplo 2.5.4 *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, 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, $\alpha = \beta = 0$. Así, u, v son vectores 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* $(1,0,0,1),(2,-2,2,-4)$ *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 resta probar que*
$$(1,0,0,1),(2,-2,2,-4)$$
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*
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \ y \ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
forman una base para $\mathcal{M}_2(\mathbb{F})$*. Ya vimos que son linealmente independientes. Como cualquier matriz* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{F})$ *se escribe como*
$$A = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
entonces es claro que estas matrices generan a $\mathcal{M}_2(\mathbb{F})$*.*

5. *Los vectores* $1, x, x^2, \ldots, x^k$ *de* $\mathbb{F}[x]$ *forman una base para el espacio*
$$S = \{p(x) \in \mathbb{F}[x] : grad(p(x)) \leq k\}$$

2.6. Dimensión

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 para $\mathbb{R}^2$. Ahora, ¿existirá alguna relación entre dos bases de un espacio vectorial?. Intentamos dar respuesta a estas preguntas en los siguientes resultados.

Comenzamos presentando una caracterización de una base de un espacio vectorial.

Teorema 2.6.1 *Sean* $v_1, \ldots, v_r$ *vectores de un espacio vectorial* V*. Las siguientes condiciones son equivalentes:*

1. *Los vectores* $v_1, \ldots, v_r$ *forman una base de* V*;*

2. *Cada vector* $v \in V$ *se expresa de manera única como*
$$v = \alpha_1 v_1 + \cdots + \alpha_r v_r$$
donde $\alpha_i \in \mathbb{F}$ *para todo* i*;*

3. *$\{v_1, \ldots, v_r\}$ es un conjunto generador minimal de V, es decir, $v_1, \ldots, v_r$ generan a V pero ningún subconjunto propio de $\{v_1, \ldots, v_r\}$ genera a V;*

4. *$\{v_1, \ldots, v_r\}$ es un conjunto linealmente independiente maximal de V, es decir, $\{v_1, \ldots, v_r\}$ es un conjunto linealmente independiente pero ningún subconjunto de V que contenga propiamente a $\{v_1, \ldots, v_r\}$ es linealmente independiente.*

Demostración. $1 \Rightarrow 2$. Sea $v \in V$. Como $v_1, \ldots, v_r$ generan a V existen escalares $\alpha_1, \ldots, \alpha_r$ tales que $v = \sum_{i=1}^{r} \alpha_i v_i$. Si también lo expresamos como $v = \sum_{i=1}^{r} \beta_i v_i$, donde $\beta_i \in \mathbb{F}$, entonces $0 = \sum_{i=1}^{r} \alpha_i v_i - \sum_{i=1}^{r} \beta_i v_i = \sum_{i=1}^{r} (\alpha_i - \beta_i) v_i$. Como $v_1, \ldots, v_r$ son linealmente independientes, $\alpha_i - \beta_i = 0$ para todo i.

$2 \Rightarrow 1$. $v_1, \ldots, v_r$ generan a V. Por otra parte, si $\sum_{i=1}^{r} \alpha_i v_i = 0 = \sum_{i=1}^{r} 0 v_i$, donde $\alpha_i \in \mathbb{F}$ para todo i, entonces por la unicidad $\alpha_i = 0$ para todo i. Es decir, $v_1, \ldots, v_r$ son linealmente independientes.

$1 \Rightarrow 3$. $v_1, \ldots, v_r$ generan a V. Si C es un subconjunto propio de $v_1, \ldots, v_r$ que genera a V entonces existe $x \in \{v_1, \ldots, v_r\} \setminus C$ tal que x es una combinación lineal no trivial de vectores de $C \subset \{v_1, \ldots, v_r\}$. Esto contradice la independencia lineal de $v_1, \ldots, v_r$.

$3 \Rightarrow 1$. Si $v_1, \ldots, v_r$ son linealmente dependientes, existe un vector $x \in \{v_1, \ldots, v_r\}$ que se escribe como combinación lineal de vectores de $\{v_1, \ldots, v_r\} - \{x\}$. Se sigue que $\{v_1, \ldots, v_r\} - \{x\}$ es un subconjunto propio que genera a V.

$1 \Rightarrow 4$. Si $\{v_1, \ldots, v_r\}$ no es maximal, existe un vector $x \in V \setminus \{v_1, \ldots, v_r\}$ tal que $\{v_1, \ldots, v_r\} \cup \{x\}$ es linealmente independiente. Pero entonces $x \notin$ *gen* $(v_1, \ldots, v_r)$ lo que contradice el hecho de que $v_1, \ldots, v_r$ genera a V.

$4 \Rightarrow 1$. Sea $x \in V$. Entonces $\{v_1, \ldots, v_r, x\}$ es linealmente dependiente. Es decir, existe una combinación lineal $\alpha x + \sum_{i=1}^{r} \alpha_i v_i = 0$, donde los escalares no son todos ceros. Como $v_1, \ldots, v_r$ es linealmente independiente, $\alpha \neq 0$. Esto implica que x es una combinación lineal de $v_1, \ldots, v_r$. Así, $v_1, \ldots, v_r$ generan a V. ■

Teorema 2.6.2 *Sea V un espacio vectorial. Supongamos que los vectores $w_1, \ldots, w_p$ de V pertenecen a gen $(v_1, \ldots, v_q)$. Si $p > q$ entonces $w_1, \ldots, w_p$ son linealmente dependientes.*

Demostración. Para cada $j = 1, \ldots, p$ existen escalares a_{ij} tales que $w_j = \sum_{i=1}^{q} a_{ij} v_i$. 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, $\sum_{j=1}^{p} a_{ij}x_j = 0$ 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 prueba que los vectores $w_1, \ldots, w_p$ de V son linealmente dependientes. ■

Teorema 2.6.3 *Sea $V = gen\,(v_1, \ldots, v_n)$ un espacio vectorial finitamente generado. Las siguientes condiciones se cumplen:*

1. *Todo conjunto de vectores linealmente independientes $w_1, \ldots, w_p$ de V se extiende a una base de V.*
2. *El conjunto de generadores $v_1, \ldots, v_n$ de V se reduce a una base de V.*

Demostración. 1. Sean $w_1, \ldots, w_p$ vectores linealmente independientes de V. Por el Teorema 2.6.2, $p \leq n$. Si el conjunto $\{w_1, \ldots, w_p\}$ es maximal entonces $w_1, \ldots, w_p$ es la base que buscamos, por el Teorema 2.6.1. De lo contrario, existe $w_{p+1} \in V$ tal que $w_1, \ldots, w_p, w_{p+1}$ son linealmente independientes. De nuevo, si $w_1, \ldots, w_p, w_{p+1}$ es maximal entonces ya encontramos una base para V. De lo contrario existe $w_{p+2} \in V$ tal que $w_1, \ldots, w_{p+2}$ son vectores linealmente independientes. Como n es una cota superior del número de vectores linealmente independientes de V, este proceso tiene que terminar en un conjunto de vectores linealmente independiente maximal $w_1, \ldots, w_p, \ldots, w_s$. Por lo tanto, $w_1, \ldots, w_p, \ldots, w_s$ es una base de V.

2. Si el conjunto $\{v_1, \ldots, v_n\}$ es minimal entonces, por el Teorema 2.6.1, ya tenemos la base de V. De lo contrario, existe un subconjunto propio G_1 de $\{v_1, \ldots, v_n\}$ que genera a V. De nuevo, si G_1 es minimal entonces tenemos la base de V. De lo contrario existe un subconjunto propio de G_1 que genera a V. Continuando este proceso, después de un número finito de pasos llegamos a un subconjunto de $\{v_1, \ldots, v_n\}$ minimal, y por lo tanto, una base de V. ■

En particular, se sigue del teorema anterior que todo espacio vectorial finitamente generado tiene una base. Ahora, si el espacio V no es finitamente generado ¿existirá una base para V? Primero que todo, observamos que en la definición 2.5.3 se dice cuando un conjunto finito de vectores $v_1, \ldots, v_r$ de un espacio V es una base. Esta definición la generalizamos para subconjuntos arbitrarios de V como sigue:

Definición 2.6.4 *Sea V un espacio vectorial. Un subconjunto X de V es linealmente independiente si todo subconjunto finito de X es linealmente independiente. Luego, decimos que X es una base de V si X es linealmente independiente y $gen\,(X) = V$.*

Por ejemplo, el espacio vectorial $\mathbb{F}[x]$ no tiene bases finitas porque, como vimos en un ejemplo anterior, ningún subconjunto finito de $\mathbb{F}[x]$ genera a $\mathbb{F}[x]$. Sin embargo, $X = \{1, x, x^2, \ldots\}$ es una base (infinita) de $\mathbb{F}[x]$.

Nuestro interés en este curso es estudiar los espacios vectoriales finitamente generados que sabemos, por el Teorema 2.6.3, tienen bases finitas. Además, probamos en el próximo resultado que dos bases de un espacio finitamente generado tienen exactamente el mismo número de vectores. Sin embargo, 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.

Teorema 2.6.5 *Si $v_1, \ldots, v_p$ y $w_1, \ldots, w_q$ son dos bases del espacio vectorial V entonces $p = q$.*

Demostración. Como $v_1, \ldots, v_p$ son linealmente independientes y contenidos en $V = gen(w_1, \ldots, w_q)$ se sigue del Teorema 2.6.2 que $p \leq q$. Por otro lado, $w_1, \ldots, w_q$ son linealmente independientes y contenidos en $V = gen(v_1, \ldots, v_p)$ lo que implica $q \leq p$. En consecuencia, $p = q$. ■

Definición 2.6.6 *La dimensión de un espacio vectorial finitamente generado V, denotado por $\dim(V)$, es el número de vectores que tiene una base de V.*

Si S es el subespacio trivial $\{0\}$ entonces S no tiene base y decimos que tiene dimensión 0.

Definición 2.6.7 *Los espacios vectoriales que tienen una base finita, i.e., los espacios finitamente generados, reciben el nombre de espacios de dimensión finita.*

Ejemplo 2.6.8 *Veamos algunos ejemplos de espacios 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, $\dim(S) = k + 1$.

Teorema 2.6.9 *Sean S y T subespacios de un espacio de dimensión finita V tales que $S \subseteq T$. Entonces $\dim(S) \leq \dim(T)$. Más aun, $\dim(S) = \dim(T)$ si, y sólo si $S = T$.*

Demostración. Sea $v_1, \ldots, v_p$ una base de $S \subseteq T$. Por el Teorema 2.6.3 extendemos esta base a una base de T. Luego $\dim(S) = p \leq \dim(T)$. Para ver la segunda parte, si $\dim(S) = \dim(T)$ entonces por el Teorema 2.6.5, $v_1, \ldots, v_p$ es una base de T. Luego $S = gen(v_1, \ldots, v_p) = T$. ■

En particular, todo subespacio de un espacio de dimensión finita tiene dimensión finita.

EJERCICIOS

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

2. 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 $\mathcal{M}_2(\mathbb{R})$.

3. Sea $V = \{p(x) \in \mathbb{F}[x] : \text{grado de } p(x) \leq 2\}$. Demuestre que
$$\mathcal{B} = \{1, x-1, x^2 - x\} \text{ y } \mathcal{B}' = \{3, x, x^2 - 1\}$$
son bases de V. Escriba cada uno de los vectores de $\mathcal{B}$ como combinación lineal de los vectores de $\mathcal{B}'$.

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

5. Determine si el conjunto
$$\{1, 1+x, x+x^2, x^2+x^3, \ldots, x^{m-1}+x^m\}$$
es una base para el espacio $V = \{p(x) \in \mathbb{F}[x] : \text{grado de } p(x) \leq m\}$

6. 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.

7. 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 $\dim(U)$.

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

2.7. El rango y la nulidad de una matriz

Describimos en esta sección métodos para encontrar una base de los subespacios $\mathcal{C}(A), \mathcal{F}(A), Im(A)$ y $\ker(A)$ 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 2.7.1 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y supongamos que $A \underset{f}{\sim} E$, donde E es escalonada reducida. Si $1 \leq k_1 < k_2 < \cdots < k_r \leq n$ son las columnas que contienen los primeros elementos distintos de cero de cada fila 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 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 Teorema 1.4.7, existe una matriz $P \in \mathcal{M}_m(\mathbb{F})$ invertible tal que $E = PA$. Luego para cada j,

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

Sea A_{*j} la columna j de A. Entonces

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

Esto prueba que 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 $\sum_{i=1}^{r}\beta_i A_{*k_i} = 0$ 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 2.7.2 *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 1.3.2 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 2.7.3 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ y supongamos que $A \underset{f}{\sim} E$, donde E es escalonada reducida. Si r es el número de filas distintas de cero de E entonces* $\dim \mathcal{C}(A) = r$.

Demostración. Es inmediato del Teorema 2.7.1. ■

Para encontrar una base para el subespacio $\mathcal{F}(A)$, calculamos una base para $\mathcal{C}\left(A^{\top}\right)$ usando el procedimiento del Teorema 2.7.1, luego la traspuesta de estos vectores forman una base para $\mathcal{F}(A)$.

Veamos ahora como calcular una base para $\ker(A)$, donde $A \in \mathcal{M}_{m\times n}(\mathbb{F})$.

Teorema 2.7.4 *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 los primeros elementos distintos de cero de cada fila 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 $\mathbb{F}^n$ obtenido al hacer $x_j = 1$ y $x_i = 0$ para todos los otros valores de i en J. Luego se calcula x_q para $q = k_r, \ldots, k_1$, en este orden, por la fórmula:*

$$x_q = -\frac{1}{e_{rq}} \left[\sum_{j>q}^{n} e_{rj} x_j \right]$$

Entonces $\{S_j : j \in J\}$ es una base para $\ker(A)$.

Demostración. La dejamos como ejercicio para el lector. ■

Ejemplo 2.7.5 *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*

$$S_2 = (2,1,0,0,0) \text{ y } S_3 = (1,0,1,0,0)$$

es una base del espacio solución del sistema.

Corolario 2.7.6 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ *y supongamos que* $A \underset{f}{\sim} E$, *donde* E *es escalonada reducida. Si* r *es el número de filas distintas de cero de* E *entonces* $\dim\ker(A) = n - r$.

Demostración. Es inmediato del Teorema 2.7.4. ■

Recordamos 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 para cada matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$

$$Ax = x_1 A_{*1} + x_2 A_{*2} + \cdots + x_n A_{*n}$$

para todo $x = (x_1, \ldots, x_n)^\top \in \mathbb{F}^n$, se deduce que

$$\mathcal{C}(A) = Im(A)$$

Definición 2.7.7 *El rango de una matriz* $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, *denotado por* $rgo(A)$, *es la dimensión de* $Im(A)$. *Esto es,*

$$rgo(A) = \dim(Im(A)) = \dim(\mathcal{C}(A))$$

La nulidad de A, *denotada por* $nul(A)$, *es la dimensión de* $\ker(A)$:

$$nul(A) = \dim(\ker(A))$$

Los Corolarios 2.7.3 y 2.7.6 implican el siguiente importante resultado.

Teorema 2.7.8 *(Teorema de la dimensión para matrices) Sea $A \in \mathcal{M}_{m\times n}(\mathbb{F})$. Entonces*
$$nul(A) + rgo(A) = n.$$

Demostración. Es inmediato. ■

En los próximos resultados probamos que para cualquier matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$
$$rgo(A) := \dim(\mathcal{C}(A)) = \dim(\mathcal{F}(A))$$
Es decir, la dimensión del espacio de columnas de A es igual a la dimensión del espacio de filas de A.

Teorema 2.7.9 *Sean $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$. Las siguientes condiciones se cumplen:*

1. *Si $A \underset{f}{\sim} B$ entonces $rgo(A) = rgo(B)$;*
2. *Si $A \underset{c}{\sim} B$ entonces $\mathcal{C}(A) = \mathcal{C}(B)$. En particular, $rgo(A) = rgo(B)$.*

Demostración. 1. Sea P una matriz invertible tal que $B = PA$. Entonces es claro que $Ax = 0$ si, y sólo si, $Bx = 0$. Esto implica que $\ker A = \ker B$ y así, $\dim\ker A = \dim\ker B$. Se sigue del Teorema 2.7.8 que $\dim\mathcal{C}(A) = \dim\mathcal{C}(B)$.

2. Existe Q invertible tal que $B = AQ$. Luego $Bx = A(Qx) \in \mathcal{C}(A)$. Por lo tanto, $\mathcal{C}(B) \subseteq \mathcal{C}(A)$. Recíprocamente, $Aw = B\left(Q^{-1}w\right) \in \mathcal{C}(B)$. Así, $\mathcal{C}(A) \subseteq \mathcal{C}(B)$. ■

Teorema 2.7.10 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Entonces $rgo(A) = \dim\mathcal{F}(A)$.*

Demostración. Por el Teorema 1.5.6, $A \sim B = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Luego existen matrices invertibles P, Q tales que $A = PBQ$. Como $B^\top = B$ se sigue que $A^\top = Q^\top BP^\top$. Por lo tanto $A^\top \sim B$ y así $A \sim A^\top$. Se sigue del Teorema 2.7.9 que $rgo(A) = rgo\left(A^\top\right) = \dim\mathcal{C}\left(A^\top\right) = \dim\mathcal{F}(A)$. ■

Recordamos que en el Teorema 1.5.7 demostramos que las matrices $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$ son equivalentes si, y sólo si, existen matrices invertibles $P \in \mathcal{M}_m(\mathbb{F})$ y $Q \in \mathcal{M}_n(\mathbb{F})$ tales que $B = PAQ$. Presentamos ahora una nueva caracterización de equivalencia en términos del rango de una matriz.

Teorema 2.7.11 *Sean $A, B \in \mathcal{M}_{m\times n}(\mathbb{F})$. Entonces $A \sim B$ si, y sólo si, $rgo(A) = rgo(B)$.*

Demostración. Si $A \sim B$ entonces por el Teorema 2.7.9 $rgo(A) = rgo(B)$. Recíprocamente, si $rgo(A) = rgo(B) = r$ entonces por el Teorema 1.5.6
$$A \sim \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \sim B$$
■

Como consecuencia de los Teoremas 2.7.8 y 2.7.10 obtenemos una caracterización para las matrices invertibles.

Teorema 2.7.12 *Sea* $A \in \mathcal{M}_n(\mathbb{F})$. *Las siguientes condiciones son equivalentes:*

1. *A es invertible;*
2. $rgo(A) = n$;
3. $nul(A) = 0$;
4. *La ecuación* $Ax = 0$ *tiene única solución* $x = 0$;
5. *Las columnas de* A *forman una base de* $\mathbb{F}^n$;
6. *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 $A \underset{f}{\sim} I$. Se sigue del Teorema 2.7.9 que $rgo(A) = rgo(I) = n$.

$2 \Rightarrow 3$. Si $rgo(A) = n$ entonces por el Teorema 2.7.8, $nul(A) = 0$.

$3 \Rightarrow 4$. Si $nul(A) = 0$ entonces $\{0\} = \ker(A) = \{x \in \mathbb{F}^n : Ax = 0\}$.

$4 \Rightarrow 5$. Supongamos que $\alpha_1 A_{*1} + \cdots + \alpha_n A_{*n} = \mathbf{0}$. Entonces $Ax = 0$, donde $x = (\alpha_1, \ldots, \alpha_n)^\top \in \mathbb{F}^n$. Por hipótesis $x = 0$ y así, $\alpha_1 = \cdots = \alpha_n = 0$. Pero n vectores linealmente independientes en $\mathbb{F}^n$ forman una base.

$5 \Leftrightarrow 6$. Se sigue del Teorema 2.7.10.

$5 \Rightarrow 1$. Supongamos que los vectores $A_{*1}, \ldots, A_{*n}$ forman una base de $\mathbb{F}^n$. Entonces cada vector canónico de $\mathbb{F}^n$ se expresa como combinación lineal de estos vectores: para todo $i = 1, \ldots, n$

$$e_i = \sum_{k=1}^{n} \alpha_{ki} A_{*k}$$

Definamos la matriz X que tiene como columna j a $(\alpha_{1j}, \ldots, \alpha_{nj})^\top$. Es fácil ver que $XA = I$. Además, para cada $i = 1, \ldots, n$

$$\begin{aligned}(AX)_{*i} &= AX_{*i} = A(\alpha_{1i}, \ldots, \alpha_{ni})^\top \\ &= \alpha_{1i} A_{*1} + \alpha_{2i} A_{*2} + \cdots + \alpha_{ni} A_{*n} \\ &= e_i\end{aligned}$$

Esto prueba que $AX = I$ y así, 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 el Teorema 2.7.4.
5. Demuestre que $rgo\,(A) = rgo\left(A^{\top}\right)$.
6. Demuestre que $rgo\,(AB) \leq rgo\,(A)$ y $rgo\,(AB) \leq rgo\,(B)$.

Capítulo 3

Transformaciones lineales

3.1. Transformaciones lineales

Los conjuntos $\mathbb{R}\times\{0\}\subseteq\mathbb{R}^2$ y $\mathbb{R}$ evidentemente son conjuntos diferentes; 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 (ver Figura *)

Figura *

De hecho, es posible establecer una correspondencia biyectiva entre estos dos conjuntos de manera que las operaciones de espacio vectorial son 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 1.2.7, la función $\tau_A:\mathbb{F}^n\rightarrow\mathbb{F}^m$ 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 $\alpha,\beta\in\mathbb{F}$ y $x,y\in\mathbb{F}^n$. Más generalmente tenemos el siguiente concepto fundamental del álgebra lineal:

Definición 3.1.1 *Sean V y W espacios vectoriales sobre $\mathbb{F}$. Una transformación lineal T de V en W es una función $T:V\rightarrow W$ que satisface*

$$\begin{aligned}T(x+y)&=T(x)+T(y)\\ T(\alpha x)&=\alpha T(x)\end{aligned}$$

para todo $\alpha,\beta\in\mathbb{F}$ 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 3.1.2 *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 (ver Figura *).*

   *Figura **

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 (ver Figura *).*

   *Figura **

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 (ver Figura *).*

   *Figura **

Ejemplo 3.1.3 *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

$$\mathcal{I}(\alpha f) = \int_a^b (\alpha f)(x)\, dx = \int_a^b \alpha f(x)\, dx = \alpha \int_a^b f(x)\, dx = \alpha \mathcal{I}(f)$$

3. *La función* $\Psi : Suc(\mathbb{F}) \to Suc(\mathbb{F})$ *definida por* $\Psi(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$. *Sean* $x = (x_1, x_2, \ldots), y = (y_1, y_2, \ldots) \in Suc(\mathbb{F})$ *y* $\alpha \in \mathbb{R}$. *Entonces*

$$\begin{aligned} \Psi(x+y) &= \Psi(x_1+y_1, x_2+y_2, \ldots) = (x_2+y_2, x_3+y_3, \ldots) \\ &= (x_2, x_3, \ldots) + (y_2, y_3, \ldots) = \Psi(x) + \Psi(y) \end{aligned}$$

y

$$\begin{aligned} \Psi(\alpha x) &= \Psi(\alpha(x_1, x_2, \ldots)) = \Psi(\alpha x_1, \alpha x_2, \ldots) \\ &= (\alpha x_2, \alpha x_3, \ldots) = \alpha(x_2, x_3, \ldots) = \alpha \Psi(x) \end{aligned}$$

4. *Sea* V *el espacio de los números complejos sobre* $\mathbb{R}$. *Entonces la función* $\Lambda : V \to V$ *definida por* $\Lambda(z) = \overline{z}$, *donde* $z \in V$, *es una transformación lineal.*

5. *La función* $\Psi : \mathcal{M}_n(\mathbb{F}) \to \mathbb{F}$ *definida por* $\Psi(A) = Tr(A)$, *para todo* $A \in \mathcal{M}_n(\mathbb{F})$, *es una transformación lineal.*

Vamos a establecer algunos hechos simples sobre las transformaciones lineales.

Teorema 3.1.4 *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. Probamos 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 3.1.5 *Sean V, W espacios vectoriales y $v_1, \ldots, v_n$ una base 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 esta 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 3.1.6 *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 3.1.7 *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 3.1.8 *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 3.1.9 *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 3.1.10 *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 si 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 3.1.11 *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 3.1.8.

11. Demuestre el Teorema 3.1.11.

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

$$F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y+z \\ x+y \end{pmatrix},\ G\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+z \\ x+z \end{pmatrix} \text{ y}$$
$$H\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2y \\ x \end{pmatrix}$$

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 $V = \ker(P) \oplus Im(P)$.

3.2. El núcleo y la imagen

Una propiedad importante de las transformaciones lineales es que envía subespacios en subespacios.

Teorema 3.2.1 *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 $x, y \in V$ 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, la imagen de V bajo $T \in \mathcal{L}(V, W)$, que denotamos por $Im(T)$, es un subespacio de W.

Otro subespacio importante asociado a $T \in \mathcal{L}(V, W)$ es el núcleo de T.

Definición 3.2.2 *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 3.2.3 *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 3.2.4 *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. *La transformación lineal $\Psi : Suc(\mathbb{F}) \to Suc(\mathbb{F})$ definida por*

$$\Psi(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$$

   tiene núcleo

$$\begin{aligned} \ker(\Psi) &= \{(x_1, x_2, \ldots) : (x_2, x_3, \ldots) = (0, 0, \ldots)\} \\ &= \{(x_1, 0, , 0, \ldots) : x_1 \in \mathbb{F}\} \end{aligned}$$

   Es claro que $Im(\Psi) = Suc(\mathbb{F})$.

3. *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 3.2.5 *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 sigue 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 prueba 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 3.2.6 *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 esto, probamos 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

$$x = \alpha_1w_1+\cdots+\alpha_kw_k+\beta_1v_1+\cdots+\beta_rv_r$$

y como $w_i\in\ker(T)$ para todo i obtenemos

$$\begin{aligned}y &= T(x) = T(\alpha_1w_1+\cdots+\alpha_kw_k+\beta_1v_1+\cdots+\beta_rv_r)\\ &= \alpha_1T(w_1)+\cdots+\alpha_kT(w_k)+\beta_1T(v_1)+\cdots+\beta_rT(v_r)\\ &= \beta_1T(v_1)+\cdots+\beta_rT(v_r)\end{aligned}$$

Esto prueba 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

$$\gamma_1T(v_1)+\cdots+\gamma_rT(v_r) = 0$$

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 2.7.8).

Una consecuencia inmediata del teorema anterior es la siguiente.

Corolario 3.2.7 *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 3.2.6 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 sigue del Teorema 2.6.9 que $Im(T) = W$. Recíprocamente, si $Im(T) = W$ entonces $\dim[\ker(T)] = 0$ y, por lo tanto, $\ker(T) = \{0\}$. Esto prueba 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 $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

3.3. Isomorfismos

Definición 3.3.1 *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 3.3.2 *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 se sigue del Corolario 3.2.7 y el Teorema 2.7.12.*

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 3.3.3 *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 3.1.9. Deducimos que la relación “ $\cong$ ” es una relación de equivalencia sobre el conjunto de los espacios vectoriales sobre $\mathbb{F}$.

Teorema 3.3.4 *Sean* $S \in \mathcal{L}(V,W)$ *y* $T \in \mathcal{L}(U,V)$ *isomorfismos. Entonces*

1. $\left(T^{-1}\right)^{-1} = T$;

2. ST *es un isomorfismo y* $(ST)^{-1} = T^{-1}S^{-1}$;

Demostración. Probamos 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 $(TS)^{-1} = \left(S^{-1}T^{-1}\right)$. ■

Los conceptos de independencia, generación y bases pasan a través del isomorfismo de un espacio a otro.

Teorema 3.3.5 *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 prueba 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 sigue 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 3.3.6 *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 de V y W, respectivamente. Consideremos la (única) transformación lineal $T \in \mathcal{L}(V, W)$ dada por el Teorema 3.1.5 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 sigue ahora del Teorema 3.2.7. ■

Se desprende del corolario anterior que un espacio vectorial V de dimensión n es isomorfo al espacio $\mathbb{F}^n$. De hecho, si $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base de V entonces para $v \in V$ existen escalares únicos $\alpha_1, \ldots, \alpha_n$ tales que $v = \sum_{i=1}^{n} \alpha_i v_i$. Definamos $[v]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)^\top$.

Definición 3.3.7 *Sea V un espacio vectorial y $\mathcal{B}$ una base de V. Para cada $v \in V$, el vector $[v]_{\mathcal{B}}$ se llama el vector de coordenadas de v con respecto a la base $\mathcal{B}$.*

Teorema 3.3.8 *Sea V un espacio vectorial de dimensión n y $\mathcal{B}$ una base de V. Entonces la función $[\cdot]_{\mathcal{B}} : V \to \mathbb{F}^n$ definida por $v \rightsquigarrow [v]_{\mathcal{B}}$ es un isomorfismo.*

Demostración. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V, $x, y \in V$ y $\alpha, \beta \in \mathbb{F}$. Entonces

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

y

$$y = \beta_1 v_1 + \cdots + \beta_n v_n$$

donde $\alpha_i, \beta_i \in F$ para todo $i = 1, \ldots, n$. Así

$$\alpha x + \beta y = (\alpha\alpha_1 + \beta\beta_1) v_1 + \cdots + (\alpha\alpha_n + \beta\beta_n) v_n$$

Luego,

$$\begin{aligned} [\alpha x + \beta y]_{\mathcal{B}} &= (\alpha\alpha_1 + \beta\beta_1, \ldots, \alpha\alpha_n + \beta\beta_n)^\top \\ &= (\alpha\alpha_1, \ldots, \alpha\alpha_n)^\top + (\beta\beta_1, \ldots, \beta\beta_n)^\top \\ &= \alpha(\alpha_1, \ldots, \alpha_n)^\top + \beta(\beta_1, \ldots, \beta_n)^\top \\ &= \alpha[x]_{\mathcal{B}} + \beta[y]_{\mathcal{B}} \end{aligned}$$

Así tenemos una transformación lineal. La biyectividad la dejamos como ejercicio para el lector. ■

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 $\mathcal{L}(W)$.

3.4. 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 3.4.1 *Sean V y W espacios vectoriales sobre $\mathbb{F}$ de dimensión finita. Supongamos que $\mathcal{B}$ es una base de V y $\mathcal{C}$ es una base W. Para cada $T \in \mathcal{L}(V, W)$ existe una única matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ tal que*

$$[T(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}} \tag{3.1}$$

para todo $v \in V$.

Demostración. Supongamos que $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base de V. Definamos la matriz $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ que tiene 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 que indica que A y B tienen todas las columnas iguales. ∎

Definición 3.4.2 *La matriz A que asociamos a la transformación lineal $T \in \mathcal{L}(V, W)$ en el Teorema 3.4.1 se llama la matriz de T respecto a las bases $\mathcal{B}$ y $\mathcal{C}$. La denotamos por $A = [T]_{\mathcal{BC}}$.*

Ejemplo 3.4.3 *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 3.4.4 *Sea* $V = \{p(x) \in \mathbb{F}[x] : \text{grado de } p(x) \leq 2\}$. *Consideremos las bases* $\mathcal{B} = \{1, x-1, x^2-x\}$ *y* $\mathcal{B}' = \{3, x, x^2-1\}$ *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 3.4.5 *Sea* V *y* W *espacios vectoriales sobre* $\mathbb{F}$ *de dimensión finita. Supongamos que* $\mathcal{B}$ *es una base de* V *y* $\mathcal{C}$ *es una base* W. *La función que asigna a cada transformación lineal* $T \in \mathcal{L}(V,W)$ *su matriz respecto a* $\mathcal{B}$ *y* $\mathcal{C}$ *es un isomorfismo entre los espacios vectoriales* $\mathcal{L}(V,W)$ *y* $\mathcal{M}_{m\times n}(\mathbb{F})$.

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 prueba 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 3.1, $[T(v)]_{\mathcal{C}} = \mathbf{0}$ 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 prueba. ■

La matriz asociada a una composición de transformaciones lineales es el producto de las matrices asociadas a cada una de las transformaciones lineales.

Teorema 3.4.6 *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 (3.1),

$$[Sw]_{\mathcal{D}} = [S]_{\mathcal{CD}} [w]_{\mathcal{C}}$$

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

$$\begin{aligned} [(ST)(v)]_{\mathcal{D}} &= [S(T(v))]_{\mathcal{D}} = [S]_{\mathcal{CD}} [T(v)]_{\mathcal{C}} \\ &= [S]_{\mathcal{BD}} [T]_{\mathcal{BC}} [v]_{\mathcal{B}} \end{aligned}$$

Por la unicidad, $[ST]_{\mathcal{BD}} = [S]_{\mathcal{CD}} [T]_{\mathcal{BC}}$. ■

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 $[T]_{\mathcal{BE}}$, $[T]_{\mathcal{EB}}$, $[T]_{\mathcal{EE}}$ y $[T]_{\mathcal{BB}}$.

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 $[T]_{\mathcal{BE}'}$ y $[T]_{\mathcal{EE}'}$.

3. Sea $T \in \mathcal{L}\left(\mathbb{R}^3\right)$ 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 $T \in \mathcal{L}\left(\mathcal{M}_2(\mathbb{R})\right)$ 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 $\mathcal{L}(\mathbb{F}^n)$.

3.5. Cambio de base y semejanza de matrices

Un caso particular del Teorema 3.4.5 ocurre cuando consideramos operadores lineales sobre V. Si $\mathcal{B}$ es una base de V y $T \in \mathcal{L}(V)$ entonces la matriz $[T]_{\mathcal{B}\mathcal{B}}$ la denotamos por $[T]_{\mathcal{B}}$ y la función $\mathcal{L}(V) \longrightarrow \mathcal{M}_n(\mathbb{F})$ definida por $T \rightsquigarrow [T]_{\mathcal{B}}$ es un isomorfismo. Surge de forma natural la siguiente pregunta: si $\mathcal{C}$ es otra base de V ¿qué relación existe entre $[T]_{\mathcal{B}}$ y $[T]_{\mathcal{C}}$? Para responder a esta pregunta analicemos primero qué ocurre con las coordenadas de un vector cuando cambiamos la base.

Teorema 3.5.1 *Sean $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{C} = \{w_1, \ldots, w_n\}$ bases de V. La matriz $P \in \mathcal{M}_n(\mathbb{F})$ que tiene como columna i al vector columna $[v_i]_{\mathcal{C}}$ es una matriz invertible que satisface*
$$[v]_{\mathcal{C}} = P\,[v]_{\mathcal{B}}$$
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,
$$[v]_{\mathcal{C}} = \left[\sum_{i=1}^{n} \alpha_i v_i\right]_{\mathcal{C}} = \sum_{i=1}^{n} \alpha_i\,[v_i]_{\mathcal{C}} = \sum_{i=1}^{n} \alpha_i P e_i = P \sum_{i=1}^{n} \alpha_i e_i = P\,[v]_{\mathcal{B}}$$
Veamos que P es invertible. Sea Q la matriz que tiene columna i al vector $[w_i]_{\mathcal{B}}$. Entonces por lo anterior $P\,[w_i]_{\mathcal{B}} = [w_i]_{\mathcal{C}} = e_i$ para todo $i = 1, \ldots, n$. Esto es, $PQ = I$. Finalmente, supongamos que $R \in \mathcal{M}_n(\mathbb{F})$ satisface
$$[v]_{\mathcal{C}} = R\,[v]_{\mathcal{B}}$$

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 3.5.2 *Sean $\mathcal{B}$ y $\mathcal{C}$ bases de un espacio vectorial V. La matriz P del teorema anterior recibe el nombre de matriz cambio de base de $\mathcal{B}$ a $\mathcal{C}$.*

Ejemplo 3.5.3 *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 para determinar las columnas de la matriz cambio de base de $\mathcal{E}$ a $\mathcal{B}$ calculamos:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \left(-\frac{1}{2}\right) \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Luego, $\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}$. *Análogamente,*

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Así, $\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}$. *En consecuencia, la matriz cambio de base de $\mathcal{E}$ a $\mathcal{B}$ es*

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

Teorema 3.5.4 *Sea V un espacio vectorial de dimensión finita y $T \in \mathcal{L}(V)$. Supongamos que $\mathcal{B}$ y $\mathcal{C}$ son bases de un espacio vectorial V. Entonces*

$$[T]_{\mathcal{B}} = P^{-1}[T]_{\mathcal{C}} P$$

donde P es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{C}$.

Demostración. Por el Teorema 3.5.1 sabemos que

$$[v]_{\mathcal{C}} = P[v]_{\mathcal{B}}$$

para todo $v \in V$, donde P es la matriz cambio de base de $\mathcal{B}$ a $\mathcal{C}$. En particular,

$$[T(v)]_{\mathcal{C}} = P[T(v)]_{\mathcal{B}}$$

Por otra parte, se sigue de (3.1) que

$$[T(v)]_{\mathcal{C}} = [T]_{\mathcal{C}}[v]_{\mathcal{C}}$$

para todo $v \in V$ y, en consecuencia,

$$\begin{aligned} \left(P^{-1}[T]_{\mathcal{C}} P\right)[v]_{\mathcal{B}} &= P^{-1}[T]_{\mathcal{C}}[v]_{\mathcal{C}} = P^{-1}[T(v)]_{\mathcal{C}} \\ &= P^{-1}P[T(v)]_{\mathcal{B}} = [T(v)]_{\mathcal{B}} \end{aligned}$$

para todo $v \in V$. Por la unicidad del Teorema 3.4.1 concluimos que

$$P^{-1}[T]_{\mathcal{C}} P = [T]_{\mathcal{B}}$$

■

Este resultado nos lleva a la siguiente definición.

Definición 3.5.5 *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 5) y, por lo tanto, divide a $\mathcal{M}_n(\mathbb{F})$ en clases de equivalencia disjuntas. Lo esencial de esta relación es que matrices en una misma clase de equivalencia comparten muchas propiedades importantes, como veremos en secciones posteriores.

Se concluye del teorema 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 3.5.6 *Supongamos que $A, B \in \mathcal{M}_n(\mathbb{F})$ son semejantes. Entonces existen bases $\mathcal{B}$ y $\mathcal{C}$ de un espacio vectorial V de dimensión finita y $T \in \mathcal{L}(V)$ tales que $[T]_{\mathcal{B}} = A$ y $[T]_{\mathcal{C}} = B$.*

Demostración. Sea V un espacio vectorial de dimensión n y $\mathcal{B}$ una base de V. Por el Teorema 3.4.5, 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 $\mathcal{C}$ de V tal que la matriz cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es $P = Q^{-1}$. En tal caso obtenemos que

$$[T]_{\mathcal{C}} = P[T]_{\mathcal{B}} P^{-1} = Q^{-1}AQ = B$$

Recordando que $[\cdot] : V \to \mathbb{F}^n$ es un isomorfismo, consideremos la base $\mathcal{C} = \{v_1, \ldots, v_n\}$ de V determinada por la relación

$$[v_i]_{\mathcal{B}} = Qe_i$$

para todo $i = 1, \ldots, n$. Entonces

$$Q^{-1}[v_i]_{\mathcal{B}} = e_i = [v_i]_{\mathcal{C}}$$

para todo i y, en consecuencia, $Q^{-1}[v]_{\mathcal{B}} = [v]_{\mathcal{C}}$ para todo $v \in V$. Como P es única tal que

$$P[v]_{\mathcal{B}} = [v]_{\mathcal{C}}$$

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. 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 $\mathcal{E}$.

2. 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 $\mathcal{E}$.

3. 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 $T(z) = \overline{z}$.

4. 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 $Q = P^{-1}$.

5. Demuestre que la relación de semejanza de matrices es una relación de equivalencia.

6. Demuestre que matrices semejante tienen la misma traza.

Capítulo 4

Espacios con producto interno

4.1. 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 4.1.1 *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 no-negativo, aun en el caso que $\mathbb{F} = \mathbb{C}$. Por otra parte, las condiciones 2 y 3 establecen linealidad en la primera componente. En 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 4.1.2 *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 4.1.3 *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 4.1.4 *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. Probamos 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 4.1.5 *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 4.1.6 *(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 4.1.7 *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 $x - p$ (ver Figura *).*

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, $\lambda y = p$. ■

Figura *

Definición 4.1.8 *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 4.1.9 *(Desigualdad de Cauchy-Schwarz) Sea V un espacio con producto interno. Si $x, y \in V$ entonces*

$$|\langle x, y\rangle| \leq \|x\| \, \|y\| \, . \tag{4.1}$$

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 (4.1) 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 sigue del Teorema de Pitágoras que

$$\begin{aligned} \|x\|^2 &= \|p\|^2 + \|x-p\|^2 \geq \|p\|^2 \\ &= \left\| \frac{\langle x, y\rangle}{\|y\|^2} y \right\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^4} \|y\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^2} \end{aligned} \tag{4.2}$$

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 (4.1) es una igualdad si, y sólo si (4.2) 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 4.1.2 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 4.1.10 *(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 sigue de la desigualdad de Cauchy-Schwarz que

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \\
&= \|x\|^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + \|y\|^2 \\
&= \|x\|^2 + 2Re\,\langle x,y\rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2\,|\langle x,y\rangle| + \|y\|^2 \\
&\leq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2
\end{aligned}$$

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 se sigue del teorema anterior. ■

EJERCICIOS

1. Compruebe que los espacios vectoriales dados en el Ejemplo 4.1.2 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 $\beta_i \in \mathbb{F}$.

4.2. 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 4.2.1 *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 4.2.2 *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 4.2.3 *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 sigue 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 4.2.4 *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 independiente en vectores ortonormales que generan exactamente el mismo subespacio.

Teorema 4.2.5 *(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 prueba 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, $v_{k+1} \in gen\left(u_1, \ldots, u_k\right) = gen\left(v_1, \ldots, v_k\right)$ 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 $j = 1, \ldots, k$

$$\begin{aligned} \left\langle u_{k+1}, u_j \right\rangle &= \left\langle \frac{v_{k+1} - \left\langle v_{k+1}, u_1 \right\rangle u_1 - \cdots - \left\langle v_{k+1}, u_k \right\rangle u_k}{\left\| v_{k+1} - \left\langle v_{k+1}, u_1 \right\rangle u_1 - \cdots - \left\langle v_{k+1}, u_k \right\rangle u_k \right\|}, u_j \right\rangle \\ &= \frac{\left\langle v_{k+1}, u_j \right\rangle - \left\langle v_{k+1}, u_j \right\rangle}{\left\| v_{k+1} - \left\langle v_{k+1}, u_1 \right\rangle u_1 - \cdots - \left\langle v_{k+1}, u_k \right\rangle u_k \right\|} = 0 \end{aligned}$$

lo que prueba que los vectores $u_1, \ldots, u_{k+1}$ son ortonormales. Además, $v_{k+1} \in gen\left(u_1, \ldots, u_{k+1}\right)$ y, por lo tanto,

$$gen\left(v_1, \ldots, v_{k+1}\right) \subseteq gen\left(u_1, \ldots, u_{k+1}\right)$$

Esto junto con el hecho que

$$\dim\left[gen\left(v_1, \ldots, v_{k+1}\right)\right] = \dim\left[gen\left(u_1, \ldots, u_{k+1}\right)\right]$$

nos lleva a la conclusión que

$$gen\left(v_1, \ldots, v_{k+1}\right) = gen\left(u_1, \ldots, u_{k+1}\right)$$

■

Corolario 4.2.6 *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

$$gen\left(u_1, \ldots, u_k\right) = gen\left(v_1, \ldots, v_k\right) = V$$

Se sigue del Teorema 4.2.3 que $u_1, \ldots, u_k$ es una base ortonormal de V. ■

Ejemplo 4.2.7 *Aplicamos el proceso de Gram-Schmidt a la base $\left\{1, x, x^2\right\}$ del espacio de polinomios de grado ≤ 2 con coeficientes reales y el producto interno*

$$\langle p, q \rangle = \int_0^1 p\left(x\right) q\left(x\right) dx$$

para obtener la base ortonormal

$$\begin{aligned}
u_1 &= 1\\
u_2 &= \frac{x-\int\limits_0^1 x dx}{\left\| x-\int\limits_0^1 x dx \right\|} = \frac{x-\frac{1}{2}}{\left\|x-\frac{1}{2}\right\|} = 2\sqrt{3}\left(x-\frac{1}{2}\right)\\
u_3 &= \frac{x^2-\int\limits_0^1 x^2 dx - 12\left(x-\frac{1}{2}\right)\int\limits_0^1 x^2\left(x-\frac{1}{2}\right)dx}{\left\| x^2-\int\limits_0^1 x^2 dx - 24\left(x-\frac{1}{2}\right)\int\limits_0^1 x^2\left(x-\frac{1}{2}\right)dx \right\|}\\
&= \frac{x^2-x+\frac{1}{6}}{\left\|x^2-x+\frac{1}{6}\right\|} = 6\sqrt{5}\left(x^2-x+\frac{1}{6}\right)
\end{aligned}$$

Definición 4.2.8 *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 4.2.9 *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 4.2.10 *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 prueba que $U \cap U^{\perp} = 0$ y el resultado sigue del Teorema 2.4.14. ■

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 $x,y\in V$
$$\langle x,y\rangle = \sum_{k=1}^{n}\langle x,v_k\rangle\,\overline{\langle y,v_k\rangle}$$

4.3. 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 4.3.1 *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 $1\le i,j\le n$*
$$[A]_{ij}=\langle Tv_j,v_i\rangle$$

Demostración. Por la parte 1 del Teorema 4.2.4, $Tv_j = \langle Tv_j, v_1 \rangle v_1 + \cdots + \langle Tv_j, v_n \rangle v_n$. 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 4.3.2 *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 3.4.5, existe $T^* \in \mathcal{L}(V)$ tal que $[T^*]_{\mathcal{B}} = A^*$. Se sigue 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 4.3.3 *El operador lineal $T^* \in \mathcal{L}(V)$ del teorema anterior recibe el nombre de operador adjunto de T.*

Ejemplo 4.3.4 *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 4.3.5 *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* $\left(R^{-1}\right)^* = (R^*)^{-1}$.

Demostración. Probamos 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 que ocurre si suponemos además que $\mathcal{B}$ y $\mathcal{C}$ son bases ortonormales de un espacio con producto interior V. Si $\mathcal{B} = \{v_1, \ldots, v_n\}$ 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,

$$[v_j]_{\mathcal{C}} = \left(\langle v_j, w_1\rangle, \ldots, \langle v_j, w_n\rangle\right)^{\top}$$

donde $\mathcal{C} = \{w_1, \ldots, w_n\}$. Luego, la fila i de P^* es $\left(\overline{\langle v_i, w_1\rangle}, \ldots, \overline{\langle v_i, w_n\rangle}\right)$ y, por lo tanto,

$$\begin{aligned}
[P^*P]_{ij} &= \overline{\langle v_i, w_1\rangle}\,\langle v_j, w_1\rangle + \cdots + \overline{\langle v_i, w_n\rangle}\,\langle v_j, w_n\rangle \\
&= \langle v_j, \langle v_i, w_1\rangle\, w_1\rangle + \cdots + \langle w_j, \langle w_i, w_n\rangle\, w_n\rangle \\
&= \langle v_j, \langle v_i, w_1\rangle\, w_1 + \cdots + \langle v_i, w_n\rangle\, w_n\rangle \\
&= \langle v_j, v_i\rangle = \left\{ \begin{array}{c} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{array} \right|
\end{aligned}$$

En consecuencia, $P^*P = I_n$.

Definición 4.3.6 *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 4.3.7 *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 4.3.8 *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.

Probamos 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 4.3.9 *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 probada. 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 3.5.6, 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 probar 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 sigue 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 4.3.7 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 4.3.10 *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 prueba al lector. ■

Definición 4.3.11 *Sea $T \in \mathcal{L}(V)$. Decimos que T es unitario si $TT^* = T^*T = I_V$.*

Probamos 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 4.3.12 *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 4.3.2,

$$[T]_{\mathcal{B}} [T]_{\mathcal{B}}^* = [T]_{\mathcal{B}} [T^*]_{\mathcal{B}} = [TT^*]_{\mathcal{B}} = [I_V]_{\mathcal{B}} = I$$

Luego, $[T]_{\mathcal{B}}$ es unitaria. Recíprocamente, supongamos que $[T]_{\mathcal{B}}$ es una matriz unitaria. Entonces de nuevo usando el Teorema 4.3.2 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 4.3.13 *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 sigue del Corolario 3.2.7 que T es un isomorfismo. Luego, si $\mathcal{B} = \{v_1, \dots, v_n\}$ es una base ortonormal de V entonces por el Teorema 3.3.5, $\{Tv_1, \dots, Tv_n\}$ es una base de V. Falta ver que es ortonormal. Pero esto se sigue 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 $\{v_1, \dots, v_n\}$ y $\{Tv_1, \dots, Tv_n\}$ 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

$$\langle Tx, Tx\rangle = \left\langle \sum \alpha_i Tv_i, \sum \alpha_j Tv_j \right\rangle = \sum \alpha_i \overline{\alpha_i} = \langle x, x\rangle$$

$3 \Rightarrow 4$ 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 $V = \{f(x) \in \mathbb{R}[x] :$ grado de $f(x)$ 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.

Capítulo 5

Determinantes

5.1. Permutaciones

Vamos a comenzar con una breve introducción a las permutaciones que necesitamos para desarrollar el tema de los determinantes.

Definición 5.1.1 *Una permutación de* $\{1, \ldots, n\}$ *es una biyección*

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

Usualmente la denotamos por

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \text{ o simplemente por } \sigma(1) \cdots \sigma(n)$$

El conjunto de las permutaciones de $\{1, \ldots, n\}$ *lo denotamos por* S_n*.*

El número de permutaciones en S_n es $n!$ (ver Ejercicio 1).

Ejemplo 5.1.2 *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 5.1.3 *Consideremos las permutaciones* $\sigma = 13524$ *y* $\tau = 12543$ *de* S_5*. Entonces*

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

Teorema 5.1.4 *Las siguientes condiciones se cumplen:*

1. *La función $\Gamma : S_n \to S_n$ definida por $\Gamma(\sigma) = \sigma^{-1}$ es biyectiva;*

2. *Fijada la permutación $\tau \in S_n$, la función $\Lambda : S_n \to S_n$ definida por $\Lambda(\sigma) = \tau\sigma$ es biyectiva.*

Demostración. 1. Si $\sigma, \mu \in S_n$ y $\Gamma(\sigma) = \Gamma(\mu)$ entonces $\sigma^{-1} = \mu^{-1}$. Luego,

$$\sigma = \left(\sigma^{-1}\right)^{-1} = \left(\mu^{-1}\right)^{-1} = \mu$$

Esto prueba que Γ es inyectiva. Por otra parte, dada $\eta \in S_n$ entonces $\eta^{-1} \in S_n$ y $\Gamma\left(\eta^{-1}\right) = \left(\eta^{-1}\right)^{-1} = \eta$, lo que prueba la sobreyectividad.

2. Supongamos que $\Lambda(\sigma) = \Lambda(\mu)$, donde $\sigma, \mu \in S_n$. Entonces $\tau\sigma = \tau\mu$. Componiendo con τ^{-1} ambos lados de esta ecuación obtenemos

$$\tau^{-1}\tau\sigma = \tau^{-1}\tau\mu$$

y, por lo tanto, $\sigma = \mu$. Para ver la sobreyectividad notamos que dada la permutación $\rho \in S_n$ se tiene que $\tau^{-1}\rho \in S_n$ y $\Lambda\left(\tau^{-1}\rho\right) = \tau\left(\tau^{-1}\rho\right) = \rho$. Esto termina la prueba. ■

Definición 5.1.5 *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 5.1.6 *Consideremos la permutación $142365 \in S_6$. Entonces las inversiones de ésta son*

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

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

Ejemplo 5.1.7 *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 5.1.7 recibe el nombre de trasposición. Es claro que si $\sigma = \sigma(1)\cdots\sigma(n) \in S_n$ 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. Luego, existen trasposiciones $\tau_1, \ldots, \tau_m$ tales que

$$\sigma\tau_1\cdots\tau_m = \varepsilon$$

Desde luego, m no es único: hay varias maneras de transformar una permutación en la identidad intercambiando dos posiciones en cada paso. Por ejemplo,

$$\sigma = 1423 \overset{2\leftrightarrow 4}{\rightsquigarrow} 1324 \overset{2\leftrightarrow 3}{\rightsquigarrow} 1234 = \varepsilon$$

o bien

$$\sigma = 1423 \overset{3\leftrightarrow 4}{\rightsquigarrow} 1432 \overset{2\leftrightarrow 3}{\rightsquigarrow} 1342 \overset{2\leftrightarrow 4}{\rightsquigarrow} 1243 \overset{3\leftrightarrow 4}{\rightsquigarrow} 1234 = \varepsilon$$

Sin embargo, probamos en el siguiente resultado que hay unicidad en la paridad de m.

Teorema 5.1.8 *Las siguientes condiciones se cumplen:*

1. $sgn(\tau\sigma) = sgn(\tau)\, sgn(\sigma)$ *para todo* $\tau, \sigma \in S_n$;
2. $sgn(\sigma) = sgn\left(\sigma^{-1}\right)$ *para todo* $\sigma \in S_n$.
3. *Si* $\sigma \in S_n$ *y* $\tau_1, \ldots, \tau_m$ *son trasposiciones tales que* $\sigma\tau_1\cdots\tau_m = \varepsilon$ *entonces*

$$\begin{aligned} sgn(\sigma) &= 1 \Rightarrow m \text{ es par} \\ sgn(\sigma) &= -1 \Rightarrow m \text{ es impar} \end{aligned}$$

Demostración. 1. Consideremos el polinomio $p = p(x_1, \ldots, x_n) = \prod\limits_{i<j}(x_i - x_j)$ y sea $\sigma \in S_n$. Definamos $\sigma(p) = \prod\limits_{i<j}\left(x_{\sigma(i)} - x_{\sigma(j)}\right)$. Como σ es biyectiva entonces vamos a tener $\sigma(p) = p$ 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 probado 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 $sgn(\tau\sigma) = sgn(\tau)\, sgn(\sigma)$.

2. Por una parte sabemos que $sgn(\varepsilon) = (-1)^0 = 1$. Luego, si $\sigma \in S_n$ entonces $\varepsilon = \sigma\sigma^{-1}$ y, por la parte anterior,

$$1 = sgn(\varepsilon) = sgn\left(\sigma\sigma^{-1}\right) = sgn(\sigma)\, sgn\left(\sigma^{-1}\right)$$

Esto prueba que $sgn(\sigma) = sgn\left(\sigma^{-1}\right)$.

3. Por la parte 1. y el Ejemplo 5.1.7,

$$1 = sgn(\varepsilon) = sgn(\sigma\tau_1 \cdots \tau_m) = sgn(\sigma)\, sgn(\tau_1) \cdots sgn(\tau_m) = (-1)^m\, sgn(\sigma)$$

■

EJERCICIOS

1. Demuestre que el número de permutaciones en S_n es $n!$.
2. Determine el signo de las siguientes permutaciones de S_6:
   a) $\sigma = 421635$
   b) $\tau = 124653$
3. Para las permutaciones dadas en el Ejercicio 2, encuentre $sgn(\sigma\tau)$, $sgn(\tau\sigma)$ y $sgn\left(\sigma^{-1}\tau^{-1}\right)$.

5.2. Determinantes

En esta sección introducimos la noción de determinante de una matriz y desarrollamos algunas de sus propiedades básicas.

Definición 5.2.1 *El determinante de una matriz $A \in \mathcal{M}_n(\mathbb{F})$, donde $[A]_{ij} = a_{ij}$, se define como*

$$|A| = \sum_{\sigma \in S_n} sgn(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

Observamos que cada uno de los sumandos $a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ (sin tomar en cuenta el signo) que aparece en el determinante de A es el producto de n elementos de A, tomando uno y sólo uno de cada fila y columna de A.

Ejemplo 5.2.2 *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,*

$$|A| = a_{11}a_{22} - a_{12}a_{21}$$

En particular, $\begin{vmatrix} 2 & 3 \\ -1 & 4 \end{vmatrix} = 2 \cdot 4 - (3)(-1) = 11.$

Ejemplo 5.2.3 *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,

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

Por ejemplo,

$$\begin{aligned} \begin{vmatrix} 3 & 1 & 2 \\ -1 & 0 & 2 \\ 0 & 1 & -1 \end{vmatrix} &= (3 \cdot 0 \cdot 1) + (1 \cdot 2 \cdot 0) + (2 \cdot (-1) \cdot 1) \\ &\quad - (2 \cdot 0 \cdot 0) - (1 \cdot (-1) \cdot (-1)) - (3 \cdot 2 \cdot 1) \\ &= -2 - 1 - 6 = -9 \end{aligned}$$

Comenzamos estudiando algunas propiedades básicas del determinante.

Teorema 5.2.4 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Entonces $|A| = \left|A^\top\right|$.*

Demostración. Para cada $\sigma \in S_n$ tenemos que

$$a_{\sigma(1)1} \cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)}$$

Además, por el Teorema 5.1.8, $sgn\,(\sigma) = sgn\left(\sigma^{-1}\right)$. Luego, si $\left[A^\top\right]_{ij} = b_{ij}$ entonces

$$\begin{aligned} \left|A^\top\right| &= \sum_{\sigma \in S_n} sgn\,(\sigma)\, b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in S_n} sgn\,(\sigma)\, a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} sgn\left(\sigma^{-1}\right) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} = |A| \end{aligned}$$

La última igualdad ocurre porque, en virtud del Teorema 5.1.4, cuando σ recorre a S_n también lo hace σ^{-1}. ■

Ejemplo 5.2.5 *En el ejemplo anterior probamos que para la matriz $A = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 0 & 2 \\ 0 & 1 & -1 \end{pmatrix}$ se tiene que $|A| = -9$. Por otro lado, la traspuesta de A es $A^\top = \begin{pmatrix} 3 & -1 & 0 \\ 1 & 0 & 1 \\ 2 & 2 & -1 \end{pmatrix}$ y*

$$\begin{aligned} \left|A^\top\right| &= (3 \cdot 0 \cdot (-1)) + ((-1) \cdot 1 \cdot 2) + (0 \cdot 1 \cdot 2) \\ &\quad - (0 \cdot 0 \cdot 2) - ((-1) \cdot 1 \cdot (-1)) - (3 \cdot 1 \cdot 2) \\ &= -9 \end{aligned}$$

En el próximo resultado estudiamos el comportamiento del determinante cuando efectuamos operaciones elementales a una matriz.

Teorema 5.2.6 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Entonces se cumplen las siguientes condiciones:*

1. *Si B se obtiene a partir de A intercambiando dos filas (o columnas) entonces $|B| = -|A|$;*

2. *Si A tiene dos filas (o columnas) iguales entonces $|A| = 0$.*

3. *Si B se obtiene a partir de A multiplicando una fila (o columna) de A por $\alpha \in \mathbb{F}$ entonces $|B| = \alpha |A|$;*

4. *Si B se obtiene a partir de A al sumar un múltiplo escalar de la fila i (columna i) a la fila j (a la columna j), donde $i \neq j$, entonces $|B| = |A|$;*

Demostración. 1. Supongamos que B es la matriz que se obtiene al intercambiar la columna i con la columna j de A. Consideremos la trasposición τ tal que $\tau(i) = j$ y $\tau(j) = i$. Entonces, por el Ejemplo 5.1.7, $sgn(\tau) = -1$ y además, por el Teorema 5.1.8, para todo $\sigma \in S_n$

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

Por lo tanto, si $[B]_{ij} = b_{ij}$ entonces

$$\begin{aligned} |B| &= \sum_{\sigma \in S_n} sgn(\sigma)\, b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in S_n} sgn(\sigma)\, a_{1\tau\sigma(1)} a_{2\tau\sigma(2)} \cdots a_{n\tau\sigma(n)} \\ &= -\sum_{\sigma \in S_n} sgn(\tau\sigma)\, a_{1\tau\sigma(1)} a_{2\tau\sigma(2)} \cdots a_{n\tau\sigma(n)} = -|A|. \end{aligned}$$

La última igualdad es por el Teorema 5.1.4 que nos asegura que cuando σ recorre todo S_n también lo hace $\tau\sigma$. Aplicando el Teorema 5.2.4 obtenemos la versión para filas.

2. Supongamos que la columna i y la columna j $(i \neq j)$ de A son iguales. Entonces la matriz que se obtiene al intercambiar la columna i y la columna j es de nuevo la matriz A. Se sigue de la primera parte que $|A| = -|A|$ y, por lo tanto, $|A| = 0$. De nuevo, por el Teorema 5.2.4 obtenemos la versión para filas.

Las partes 3. y 4. son consecuencia del siguiente hecho:

$$\begin{vmatrix} A_1 \\ \vdots \\ A_{j-1} \\ \alpha X + \beta Y \\ A_{j+1} \\ \vdots \\ A_n \end{vmatrix} = \alpha \begin{vmatrix} A_1 \\ \vdots \\ A_{j-1} \\ X \\ A_{j+1} \\ \vdots \\ A_n \end{vmatrix} + \beta \begin{vmatrix} A_1 \\ \vdots \\ A_{j-1} \\ Y \\ A_{j+1} \\ \vdots \\ A_n \end{vmatrix} \tag{5.1}$$

donde $A_i = (a_{i1}, \ldots, a_{in})$, $X = (x_{j1}, \ldots, x_{jn})$ y $Y = (y_{j1}, \ldots, y_{jn})$. En efecto, el determinante en la izquierda es igual a

$$\begin{aligned}
& \sum_{\sigma\in S_n} sgn\,(\sigma)\, a_{1\sigma(1)} \cdots a_{j-1\sigma(j-1)} \left(\alpha x_{j\sigma(j)} + \beta y_{j\sigma(j)}\right) a_{j+1\sigma(j+1)} \cdots a_{n\sigma(n)} \\
= \; & \alpha \sum_{\sigma\in S_n} sgn\,(\sigma)\, a_{1\sigma(1)} \cdots a_{j-1\sigma(j-1)} x_{j\sigma(j)} a_{j+1\sigma(j+1)} \cdots a_{n\sigma(n)} \\
& +\beta \sum_{\sigma\in S_n} sgn\,(\sigma)\, a_{1\sigma(1)} \cdots a_{j-1\sigma(j-1)} y_{j\sigma(j)} a_{j+1\sigma(j+1)} \cdots a_{n\sigma(n)}
\end{aligned}$$

■

Ahora tenemos una técnica para calcular el determinante de una matriz: efectuamos una sucesión finita de operaciones elementales hasta llegar a una matriz triangular. Luego aplicamos el siguiente resultado:

Teorema 5.2.7 *Si A es una matriz triangular entonces $|A| = [A]_{11} \cdots [A]_{nn}$.*

Demostración. Supongamos que A es una matriz triangular inferior, es decir, $a_{kl} = 0$ siempre que $k < l$. Sea $\sigma \in S_n$ y supongamos que $\sigma(1) \neq 1$. Entonces $1 < \sigma(1)$ y, en consecuencia, $a_{1\sigma(1)} = 0$. Luego, los términos de la forma $sgn\,(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ son cero para las permutaciones $\sigma \in S_n$ tales que $\sigma(1) \neq 1$. Consideremos las permutaciones que tienen $\sigma(1) = 1$. Si $\sigma(2) \neq 2$ entonces $2 < \sigma(2)$ y así $a_{2\sigma(2)} = 0$. Luego, los términos

$$sgn\,(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

son ceros cuando $\sigma(1) = 1$ y $\sigma(2) \neq 2$. Tomamos la permutaciones que satisfacen $\sigma(1) = 1$ y $\sigma(2) = 2$ .... Continuando de esta manera es claro que el único posible sumando distinto de cero ocurre cuando σ es la identidad. Por lo tanto, $|A| = [A]_{11} \cdots [A]_{nn}$. ■

Ejemplo 5.2.8 *Consideremos la matriz $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$. Entonces*

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

Luego, usando los Teoremas 5.2.6 y 5.2.7 obtenemos

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

EJERCICIOS

1. Use el Teorema 5.2.6 para calcular 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, $|A + B| \neq |A| + |B|$.

3. Sea $\alpha \in \mathbb{F}$. Demuestre que para cualquier matriz $A \in \mathcal{M}_n(\mathbb{F})$ se tiene que $|\alpha A| = \alpha^n |A|$.

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

5.3. Funciones n-lineales y alternadas

La función determinante es la función

$$|\cdot| : \mathcal{M}_n(\mathbb{F}) \to \mathbb{F}$$

que asigna a cada $A \in \mathcal{M}_n(\mathbb{F})$ su determinante $|A| \in \mathbb{F}$. Notamos que el espacio de las matrices $\mathcal{M}_n(\mathbb{F})$ está en correspondencia biunívoca con el espacio vectorial $(\mathbb{F}^n)^n = \mathbb{F}^n \times \cdots \times \mathbb{F}^n$ formado por n-uplas $(X_1, \ldots, X_n)$, donde cada X_i es un vector fila de $\mathbb{F}^n$, para $i = 1, \ldots, n$. La correpondencia es natural: a cada matriz $A \in \mathcal{M}_n(\mathbb{F})$ le asociamos el vector $(A_1, \ldots, A_n)$, donde A_i es la fila i de A para todo $i = 1, \ldots, n$. Es fácil ver que esta correspondencia es biyectiva y, aun más, es un isomorfismo de espacios vectoriales. De esta manera, podemos considerar a la función determinante como una función

$$|\cdot| : (\mathbb{F}^n)^n \to \mathbb{F}$$

definida en n coordenadas.

Definición 5.3.1 *Sea $f : (\mathbb{F}^n)^n \to \mathbb{F}$ una función. Decimos que*

1. *f es n-lineal si para todo $\alpha, \beta \in \mathbb{F}$ y $X, Y \in \mathbb{F}^n$ se tiene que*

$$f\left(\ldots, \alpha X \overset{j}{+} \beta Y, \ldots\right) = \alpha f\left(\ldots, \overset{j}{X}, \ldots\right) + \beta f\left(\ldots, \overset{j}{Y}, \ldots\right)$$

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

2. *f es alternada si $f(X_1, \ldots, X_n) = 0$ cuando $X_i = X_j$ para algún $i \neq j$.*

En la demostración del Teorema 5.2.6 se obtiene que la función determinante $|\cdot| : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F}$ es n-lineal y alternada. Veremos a continuación que estas dos propiedades junto con el hecho de que el determinante de la identidad es 1 (Teorema 5.2.7) caracterizan a la función determinante.

Lema 5.3.2 *Sea* $f : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F}$ *una función* n*-lineal, alternada y* $X_1, \ldots, X_n \in \mathbb{F}^n$. *Si* $1 \leq i < j \leq n$ *entonces*

$$f\left(\ldots, \overset{i}{X_i}, \ldots, \overset{j}{X_j}, \ldots\right) = -f\left(\ldots, \overset{i}{X_j}, \ldots, \overset{j}{X_i}, \ldots\right)$$

Es decir, f *cambia de signo cuando se intercambian la fila* i *y la fila* j. *Además, si* $\sigma \in S_n$, *entonces*

$$f\left(X_{\sigma(1)}, \ldots, X_{\sigma(n)}\right) = sgn\left(\sigma\right) f\left(X_1, \ldots, X_n\right)$$

Demostración. Al ser f alternada y n-lineal se sigue que

$$\begin{aligned}
0 &= f\left(\ldots, X_i \overset{i}{+} X_j, \ldots, X_i \overset{j}{+} X_j, \ldots\right) \\
&= f\left(\ldots, \overset{i}{X_i}, \ldots, \overset{j}{X_i}, \ldots\right) + f\left(\ldots, \overset{i}{X_i}, \ldots, \overset{j}{X_j}, \ldots\right) \\
&\quad + f\left(\ldots, \overset{i}{X_j}, \ldots, \overset{j}{X_i}, \ldots\right) + f\left(\ldots, \overset{i}{X_j}, \ldots, \overset{j}{X_j}, \ldots\right) \\
&= f\left(\ldots, \overset{i}{X_i}, \ldots, \overset{j}{X_j}, \ldots\right) + f\left(\ldots, \overset{i}{X_j}, \ldots, \overset{j}{X_i}, \ldots\right)
\end{aligned}$$

Supongamos ahora que $\sigma \in S_n$. Entonces existen trasposiciones $\tau_1, \ldots, \tau_m$ tales que $\sigma\tau_1 \cdots \tau_m = \varepsilon$. Luego, por la primera parte y la parte 3 del Teorema 5.1.8,

$$f\left(X_{\sigma(1)}, \ldots, X_{\sigma(n)}\right) = (-1)^m f\left(X_1, \ldots, X_n\right) = sgn\left(\sigma\right) f\left(X_1, \ldots, X_n\right)$$

■

Teorema 5.3.3 *Sea* $f : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F}$ *una función* n*-lineal, alternada y* $f\left(e_1, \ldots, e_n\right) = 1$. *Entonces* f *es la función determinante.*

Demostración. Sean $A_i = \left(a_{i1}, \ldots, a_{in}\right)$ las filas de A, donde $i = 1, \ldots, n$. Entonces

$$A_i = \sum_{j=1}^{n} a_{ij} e_j$$

Luego, por la n-linealidad de f deducimos que

$$\begin{aligned}
f\left(A_1, \ldots, A_n\right) &= f\left(\sum_{j_1=1}^{n} a_{1j_1} e_{j_1}, \ldots, \sum_{j_n=1}^{n} a_{nj_n} e_{j_n}\right) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdots a_{nj_n} f\left(e_{j_1}, \ldots, e_{j_n}\right)
\end{aligned}$$

Ahora, como f es alternada, los términos diferentes de cero ocurren cuando $j_1, \ldots, j_n$ son diferentes. Es decir,

$$f\left(A_1, \ldots, A_n\right) = \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} f\left(e_{\sigma(1)}, \ldots, e_{\sigma(n)}\right)$$

Por otra parte, como f es alternada y n-lineal se sigue del Lema 5.3.2 que

$$f\left(e_{\sigma(1)},\ldots,e_{\sigma(n)}\right)=sgn\left(\sigma\right)f\left(e_1,\ldots,e_n\right)=sgn\left(\sigma\right)$$

Así, concluimos que

$$f\left(A_1,\ldots,A_n\right)=\sum_{\sigma\in S_n}sgn\left(\sigma\right)a_{1\sigma(1)}\cdots a_{n\sigma(n)}=|A|$$

■

EJERCICIOS

1. Demuestre que si $f:\mathbb{F}^n\times\cdots\times\mathbb{F}^n\to\mathbb{F}$ satisface la condición más débil $f\left(X_1,\ldots,X_n\right)=0$ cuando $X_i=X_{i+1}$ para algún i, entonces f es alternada.

2. Sea $f:\mathbb{F}^3\times\mathbb{F}^3\times\mathbb{F}^3\to\mathbb{F}$ definida por

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

Demuestre que f es 3-lineal, alternada y $f\left(I_3\right)=1$.

5.4. La expansión de Laplace

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

Definición 5.4.1 *Sea $A\in\mathcal{M}_n\left(\mathbb{F}\right)$ tal que $[A]_{ij}=a_{ij}$. Consideremos la matriz $A_{ij}\in\mathcal{M}_{n-1}\left(\mathbb{F}\right)$ obtenida a partir de A al eliminar la fila i y la columna j de A. Entonces A_{ij} se llama el menor del elemento a_{ij} de A. El escalar $(-1)^{i+j}\left|A_{ij}\right|$ se llama el cofactor de a_{ij}.*

Ejemplo 5.4.2 *Para la matriz A dada en el Ejemplo 5.2.8, el menor del elemento a_{23} de A es*

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

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

Teorema 5.4.3 *Sea $A\in\mathcal{M}_n\left(\mathbb{F}\right)$ tal que $[A]_{ij}=a_{ij}$. Entonces para cada $i=1,\ldots,n$*

$$|A|=\sum_{j=1}^{n}(-1)^{i+j}a_{ij}\left|A_{ij}\right|$$

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

$$|A|=\sum_{i=1}^{n}(-1)^{i+j}a_{ij}\left|A_{ij}\right|$$

Demostración. Probamos la segunda fórmula. Por el Teorema 5.3.3 basta probar que para j fijo, la función $\Psi : \mathcal{M}_n(\mathbb{F}) \to \mathbb{F}$ dada por $\Psi(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}|$ es n-lineal, alternada y $\Psi(I) = 1$. Supongamos que $A_1, \ldots, A_n$ son las filas de A y $A_k = A_l$, donde $1 \leq k < l \leq n$. Si $i \neq k$ o $i \neq l$ entonces claramente $|A_{ij}| = 0$ 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 sigue de la parte 1 del Teorema 5.2.6 que $|A_{lj}| (-1)^{l-k-1} = |A_{kj}|$. En consecuencia,

$$\begin{aligned}
\Psi(A) &= (-1)^{k+j} a_{kj} |A_{kj}| + (-1)^{l+j} a_{lj} |A_{lj}| \\
&= (-1)^{k+j} a_{kj} |A_{lj}| (-1)^{l-k-1} + (-1)^{l+j} a_{lj} |A_{lj}| \\
&= (-1)^{j+l-1} a_{kj} |A_{lj}| + (-1)^{l+j} a_{kj} |A_{lj}| = 0
\end{aligned}$$

Esto prueba que Ψ es alternada. Probamos ahora que Ψ es n-lineal. Para esto, supongamos que la fila p de A tiene la forma $A_p = \alpha B_p + \beta C_p$ y consideremos las matrices B con elementos $[B]_{ij} = b_{ij}$ y C con elementos $[C]_{ij} = c_{ij}$, que tienen filas

$$A_1, \ldots, A_{p-1}, B_p, A_{p+1}, \ldots, A_n$$

y

$$A_1, \ldots, A_{p-1}, C_p, A_{p+1}, \ldots, A_n$$

respectivamente. En primer lugar, si $i = p$ entonces $|A_{pj}| = |B_{pj}| = |C_{pj}|$ y, por lo tanto,

$$\begin{aligned}
a_{pj} |A_{pj}| &= (\alpha b_{pj} + \beta c_{pj}) |A_{pj}| \\
&= \alpha b_{pj} |B_{pj}| + \beta c_{pj} |C_{pj}|
\end{aligned}$$

Si $i \neq p$, entonces sabemos por la $(n-1)$-linealidad del determinante que

$$|A_{ij}| = \alpha |B_{ij}| + \beta |C_{ij}|$$

Además, en este caso $a_{ij} = b_{ij} = c_{ij}$, por lo tanto,

$$a_{ij} |A_{ij}| = \alpha b_{ij} |B_{ij}| + \beta c_{ij} |C_{ij}|$$

Esto prueba que Ψ es n-lineal. Finalmente, es claro que $\Psi(I) = 1$. ∎

Las fórmulas para el determinante de una matriz encontradas en el Teorema 5.4.3 se llaman expansión de Laplace del determinante de A por la fila i y por la columna j, respectivamente.

Ejemplo 5.4.4 *Sea* $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$. *Consideremos las siguientes operaciones elementales*:

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

Luego, por el Teorema 5.2.6, $|A| = \begin{vmatrix} 1 & 3 & -2 \\ 0 & 2 & 1 \\ 0 & 4 & 3 \end{vmatrix}$. *Usamos ahora la expansión de Laplace por la columna* 1 *y obtenemos*

$$|A| = a_{11} |A_{11}| = 1 \begin{vmatrix} 2 & 1 \\ 4 & 3 \end{vmatrix} = 6 - 4 = 2$$

EJERCICIOS

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

2. 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 5.4.3 demuestre que el determinante de una matriz triangular es el producto de los elementos sobre la diagonal.

4. Demuestre que si $A \in \mathcal{M}_n(\mathbb{F})$ es antisimétrica entonces $\left|A^\top\right| = (-1)^n |A|$.

5. Demuestre que si $A \in \mathcal{M}_n(\mathbb{F})$ es antisimétrica y n es impar entonces $|A| = 0$.

6. Demuestre que $|V| = \prod\limits_{1 \leq i < j \leq n} (x_j - x_i)$, donde V es la matriz de Vandermonde

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

5.5. La adjunta y la inversa de una matriz

Usamos en esta sección los cofactores para construir la inversa de una matriz.

Definición 5.5.1 *La adjunta de* $A \in \mathcal{M}_n(\mathbb{F})$, *que denotamos por* A^a, *la definimos como*

$$[A^a]_{ij} = (-1)^{i+j} |A_{ji}|$$

Ejemplo 5.5.2 *Para la matriz* $A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$ *tenemos que*

$$\begin{array}{lll} [A^a]_{11} = 29 & [A^a]_{12} = -17 & [A^a]_{13} = 7 \\ [A^a]_{21} = -5 & [A^a]_{22} = 3 & [A^a]_{23} = -1 \\ [A^a]_{31} = 6 & [A^a]_{32} = -4 & [A^a]_{33} = 2 \end{array}$$

Luego,

$$A^a = \begin{pmatrix} 29 & -17 & 7 \\ -5 & 3 & -1 \\ 6 & -4 & 2 \end{pmatrix}$$

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

Teorema 5.5.3 *Si* $A \in \mathcal{M}_n(\mathbb{F})$ *entonces* $AA^a = |A| I_n$*. Además, A es invertible si, y sólo si,* $|A| \neq 0$*, en cuyo caso*

$$A^{-1} = \frac{1}{|A|} A^a$$

Demostración. El término i, j de AA^a es

$$[AA^a]_{ij} = \sum_{k=1}^{n} a_{ik} (-1)^{k+j} |A_{jk}|$$

Si $i = j$ entonces por el Teorema 5.4.3

$$[AA^a]_{ii} = \sum_{k=1}^{n} (-1)^{k+i} a_{ik} |A_{ik}| = |A|$$

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 $|A_{jk}| = |C_{jk}|$ para todo k. Además, $|C| = 0$ porque tiene dos filas iguales. Luego,

$$\begin{aligned} [AA^a]_{ij} &= \sum_{k=1}^{n} a_{ik} (-1)^{k+j} |A_{jk}| = \sum_{k=1}^{n} (-1)^{k+j} c_{jk} |C_{jk}| \\ &= |C| = 0 \end{aligned}$$

Esto prueba que $AA^a = |A| I$.

Para ver la segunda parte, si $|A| \neq 0$ entonces por lo anterior,

$$A \frac{A^a}{|A|} = I$$

Además, como $A^\top \left(A^\top\right)^a = \left|A^\top\right| I$ entonces

$$|A|\, I = \left|A^\top\right| I = A^\top \left(A^\top\right)^a = A^\top \left(A^a\right)^\top = \left(A^a A\right)^\top$$

y, en consecuencia, $A^a A = \left(A^a A\right)^{\top\top} = |A|\, I^\top = |A|\, I$. Esto prueba que

$$\frac{A^a}{|A|} A = I$$

Así, A es invertible y $A^{-1} = \frac{A^a}{|A|}$. Recíprocamente, si A es invertible entonces, por el Teorema 1.4.5, A es equivalente por filas a la matriz identidad. Por el Teorema 5.2.6, $|A| = (-1)^t r_1 \cdot \cdots \cdot r_s$ donde $r_1, \ldots, r_s$ son los escalares diferentes de cero que aparecen al multiplicar una fila y $(-1)^t$ es el signo que aparece cuando intercambiamos dos filas. Esto prueba que $|A| \neq 0$. ■

Ejemplo 5.5.4 *En un ejemplo anterior demostramos que* $\begin{vmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{vmatrix} = 2$ *y*

$$\begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}^a = \begin{pmatrix} 29 & -17 & 7 \\ -5 & 3 & -1 \\ 6 & -4 & 2 \end{pmatrix}$$

Se sigue del teorema anterior que

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

Teorema 5.5.5 *Sean* $A, B \in \mathcal{M}_n(\mathbb{F})$. *Entonces:*

1. $|AB| = |A|\,|B|$;
2. *Si* A *es invertible entonces* $\left|A^{-1}\right| = \frac{1}{|A|}$.

Demostración. 1. Si A no es invertible tampoco lo es AB. Se sigue del teorema anterior que $|AB| = 0 = |A|\,|B|$. Luego, suponemos que A es invertible, en cuyo caso

$$A = E_1 \cdots E_k$$

donde E_j es una matriz elemental para cada $j = 1, \ldots, n$. Luego, el problema se reduce a probar que $|EB| = |E|\,|B|$ para cualquier matriz elemental E. Por el Teorema 1.4.3 tenemos

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

Luego, por el Teorema 5.2.6

$$\begin{aligned}|E_{pq}| &= -1\,|I_n| = -1\\ |E_p(\lambda)| &= \lambda\,|I_n| = \lambda\\ |E_{pq}(\lambda)| &= |I_n| = 1\end{aligned}$$

y así,

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

2. Si A es invertible entonces $AA^{-1} = I$. Luego por la parte anterior,

$$1 = |I| = \left|AA^{-1}\right| = |A|\left|A^{-1}\right|$$

En consecuencia, $\left|A^{-1}\right| = \frac{1}{|A|}$. ■

Corolario 5.5.6 *(La regla de Cramer) Sea $A \in \mathcal{M}_n(\mathbb{F})$ y $B = (b_1, \ldots, b_n)^\top \in \mathbb{F}^n$. Si $|A| \neq 0$ entonces la ecuación matricial $AX = B$ tiene una única solución dada por*

$$x_j = |A|^{-1} \sum_{i=1}^{n} (-1)^{i+j}\, b_i\, |A_{ij}|$$

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

Demostración. A es invertible porque $|A| \neq 0$. En consecuencia, $AX = B$ tiene única solución $X = A^{-1}B$. Pero por el Teorema 5.5.3,

$$X = A^{-1}B = \frac{1}{|A|}A^aB$$

■

EJERCICIOS

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

2. Determine si las matrices dadas a continuación son invertibles. Si son invertibles, calcule la inversa.

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

3. Demuestre que si A es ortogonal entonces $|A| = \pm 1$.

4. Demuestre que si U es unitaria entonces $|U| = \pm 1$.

5. 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 $|N| = 0$.

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

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

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

Capítulo 6

Diagonalización

6.1. Autovalores y autovectores

Uno de los objetivos principales de este curso consiste en estudiar los operadores lineales definidos sobre un espacio vectorial V. En vista del isomorfismo entre $\mathcal{L}(V)$ y $\mathcal{M}_n(\mathbb{F})$, para estudiar un operador lineal T es posible elegir una base $\mathcal{B}$ de V y estudiar a $[T]_{\mathcal{B}}$, la matriz asociada a T con respecto a la base $\mathcal{B}$. 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 6.1.1 *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 6.1.2 *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 6.1.3 *Consideremos el operador lineal $T \in \mathcal{L}\left(\mathbb{R}^2\right)$ 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 6.1.4 *Sea $S\in\mathcal{L}\left(\mathbb{R}^2\right)$ 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 $Q^{-1}AQ=D$, donde D es diagonal. Luego,

$$\mathbf{0}=A^2=\left(QDQ^{-1}\right)^2=QD^2Q^{-1}$$

Al ser Q invertible y D diagonal, concluimos que $D=0$ y esto obliga a $A=0$, 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 6.1.5 *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 3.5.4,

$$D=[T]_{\mathcal{C}}=P\,[T]_{\mathcal{B}}\,P^{-1}=PAP^{-1}$$

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 3.5.6, existe una base $\mathcal{C}$ de V tal que $[T]_{\mathcal{C}}=D$. Esto prueba que T es diagonalizable. ■

Definición 6.1.6 *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 6.1.7 *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 6.1.8 *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 6.1.9 *Sea $A \in \mathcal{M}_n(\mathbb{F})$. Las siguientes condiciones son equivalentes:*

1. *λ es un autovalor de A;*
2. *$|\lambda I - A| = 0$.*

Demostración. Esto es consecuencia de los Teoremas 2.7.12 y 5.5.3. ■

Este resultado nos proporciona un método para encontrar los autovalores y autovectores de una matriz. Si tratamos a λ como una variable entonces $|\lambda I - A|$ es un polinomio en λ y los autovalores de A son las raíces de la ecuación polinómica $|\lambda I - A| = 0$. 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 6.1.10 *El polinomio $p_A = p_A(x) = |xI - A| \in \mathbb{F}[x]$ se llama el 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 6.1.11 *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*

$$p_B = \begin{vmatrix} x & -1 \\ 1 & x \end{vmatrix} = x^2 + 1 \in \mathbb{R}[x]$$

Este polinomio no tiene raíces sobre $\mathbb{R}$ y, por lo tanto, B no tiene autovalores.

Ejemplo 6.1.12 *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*

$$p_B = \begin{vmatrix} x & -1 \\ 1 & x \end{vmatrix} = x^2 + 1 = (x + i)(x - i)$$

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 6.1.13 *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

$$p_A = |xI - A| = \left|P^{-1}(xI - A)P\right| = |xI - B| = p_B$$

■

En particular, matrices semejantes tienen los mismos autovalores.

Concluimos esta sección con un importante resultado sobre la independencia lineal de autovectores.

Teorema 6.1.14 *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

$$x_{s+1} = \alpha_1 x_1 + \cdots + \alpha_s x_s \tag{6.1}$$

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 (6.1) 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 probar 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 sigue 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 la 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 prueba. ■

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.

6.2. Diagonalización

En esta sección establecemos la conexión entre la teoría de autovalores y el problema de diagonalización.

Teorema 6.2.1 *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 2.7.12, 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 prueba 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:

$$p_A = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k} \tag{6.2}$$

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 (6.2).

Definición 6.2.2 *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 6.2.3 *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 6.2.4 *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 sigue del Teorema 6.1.13 que

$$p_A = p_{Q^{-1}AQ} = \begin{vmatrix} (x-\lambda_0)\, I_r & -B \\ 0 & xI_{n-r} - C \end{vmatrix}$$

Desarrollando este determinante por las primeras m columnas obtenemos que $p_A = (x-\lambda_0)^r \, |xI_{n-r} - C|$ lo que prueba 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 6.2.5 *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,

$$(\lambda I - A) = P \begin{pmatrix} 0 & 0 \\ 0 & \lambda I_{n-r} - B \end{pmatrix} P^{-1}$$

y así,

$$rgo\,(\lambda I - A) = rgo \begin{pmatrix} 0 & 0 \\ 0 & \lambda I_{n-r} - B \end{pmatrix} = rgo\,(\lambda I_{n-r} - B) = n - r$$

Se sigue 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 6.1.14, 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 sigue del Teorema 6.2.1. ■

Ejemplo 6.2.6 *Sea* $A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 2 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$. *Entonces*

$$p_A = \begin{vmatrix} x-2 & -1 & -1 \\ -2 & x-3 & -2 \\ -1 & -1 & x-2 \end{vmatrix} = (x-5)(x-1)^2$$

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*

$$5I - A = \begin{pmatrix} 3 & -1 & -1 \\ -2 & 2 & -2 \\ -1 & -1 & 3 \end{pmatrix} \underset{f}{\sim} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \tag{6.3}$$

Luego, $rgo(5I - A) = 2$ *y por el Teorema de la dimensión,* $mg_A(5) = 1$. *Similarmente,*

$$1I - A = \begin{pmatrix} -1 & -1 & -1 \\ -2 & -2 & -2 \\ -1 & -1 & -1 \end{pmatrix} \underset{f}{\sim} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{6.4}$$

y así, $rgo(I - A) = 1$ *lo que implica* $mg_A(1) = 2$. *El Teorema 6.2.5 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 (6.3) obtenemos que*

$$\begin{aligned} y - 2z &= 0 \\ x - z &= 0 \end{aligned}$$

Luego, la solución del sistema homogéneo $(5I - A)X = 0$ *viene dado 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 sigue de la relación (6.4) 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 6.2.7 *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?

6.3. Diagonalización unitaria

En las secciones anteriores estudiamos el problema de decidir cuando 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 6.3.1 *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 6.3.2 *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 4.3.9,

$$D = [T]_{\mathcal{C}} = P^* [T]_{\mathcal{B}} P$$

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 4.3.9, existe una base ortonormal $\mathcal{C}$ de V tal que $[T]_{\mathcal{C}} = D$. Esto prueba que T es diagonalizable. ■

Definición 6.3.3 *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 6.3.4 *(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 6.3.5 *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 6.3.6 *Una matriz* $A \in \mathcal{M}_n(\mathbb{F})$ *es hermitiana si* $A^* = A$.

Ejemplo 6.3.7 *La matriz* $A = \begin{pmatrix} 1 & 2i & 1-i \\ -2i & 2 & 1+i \\ 1+i & 1-i & 0 \end{pmatrix}$ *es hermitiana.*

Teorema 6.3.8 *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

$$\begin{aligned} x^* A x &= \lambda x^* x \\ x^* A x &= \overline{\lambda} x^* x \end{aligned}$$

Se sigue 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 $U^\top A U = T$, para una matriz real ortogonal U y una matriz triangular superior T. En particular, $U^\top A U$ es simétrica y triangular, y, por lo tanto, diagonal. ■

Nos planteamos ahora el problema de determinar cuando una matriz $A \in \mathcal{M}_n(\mathbb{F})$ es unitariamente diagonalizable.

Teorema 6.3.9 *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 $U^* A U = D$, donde U es una matriz unitaria. Luego $A = UDU^*$ y $A^* = UD^*U^*$. De aquí se sigue 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 exite una matriz unitaria U y una matriz triangular superior T tal que

$$U^* A U = T = \begin{pmatrix} \lambda_1 & t_{12} & t_{13} & \cdots & t_{1n} \\ 0 & \lambda_2 & t_{23} & \cdots & t_{2n} \\ 0 & 0 & \lambda_3 & \cdots & t_{3n} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{6.5}$$

donde $\lambda_1, \ldots, \lambda_n$ son los autovalores de A. Como U es unitaria y A es normal entonce 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 (6.5). Luego, $A = UTU^*$ y $A^* = UT^*U^*$ y, por lo tanto, $AA^* = UTT^*U^*$. Se sigue 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, $\sum_{i<j}^{n} |t_{ij}|^2 = 0$ y así, $t_{ij} = 0$ para todo $i < j$. Esto prueba que T es diagonal.

1 $\Rightarrow$ 4. Supongamos que $AU = UD$, donde $D = diag\,(\lambda_1, \ldots, \lambda_n)$. Por el Teorema 4.3.7, 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 prueba que las columnas de U son autovectores de A.

4 $\Rightarrow$ 1. Es claro. ■

Definición 6.3.10 *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 anti-hermitianas: $A \in \mathcal{M}_n(\mathbb{F})$ es anti-hermitiana si $A^* = -A$.

Ejemplo 6.3.11 *La matriz* $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ *es normal pero no es unitaria, hermitiana ni anti-hermitiana.*

EJERCICIOS

1. Considere las matrices

$$A = \frac{1}{2}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix}, \ B = \begin{pmatrix} 3 & 1-2i & 4+7i \\ 1+2i & -4 & -2i \\ 4-7i & 2i & 2 \end{pmatrix} \text{ y}$$

$$C = \begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix}$$

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).