

Capítulo 4

Aplicaciones lineales

4.1. Introducción a las aplicaciones lineales

En el capítulo anterior encontramos la aplicación de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ que asignaba, dada una base del espacio vectorial, coordenadas a un vector. Vimos que las coordenadas del vector suma son la suma de las coordenadas, y las de un vector escalado son las coordenadas escaladas:

$$\mathbf{x} + \mathbf{y} \mapsto [\mathbf{x}]_{\mathcal{B}} + [\mathbf{y}]_{\mathcal{B}} \quad \text{y} \quad c\mathbf{x} \mapsto c[\mathbf{x}]_{\mathcal{B}}$$

Estas curiosas propiedades son compartidas por muchas funciones o aplicaciones entre espacios vectoriales. En términos generales, una aplicación T entre dos espacios vectoriales $\mathcal{V}$ y $\mathcal{W}$ es una función

$$\begin{aligned} T : \mathcal{V} &\rightarrow \mathcal{W} \\ \mathbf{x} \in \mathcal{V} &\mapsto \mathbf{y} = T(\mathbf{x}) \in \mathcal{W} \end{aligned}$$

Se describe muchas veces una aplicación con la notación $\mathbf{x} \mapsto T(\mathbf{x})$. El espacio $\mathcal{V}$ se denomina **dominio** de T y el espacio $\mathcal{W}$ **codominio** de T . A cada elemento $\mathbf{v} \in \mathcal{V}$ del dominio le corresponde una **imagen** $\mathbf{w} \in \mathcal{W}$ en el codominio. A cada elemento del codominio puede corresponderle un conjunto entero de **preimágenes**. La **imagen*** (o recorrido) de T es el conjunto de vectores de $\mathcal{W}$ tales que son imagen de algún $\mathbf{x} \in \mathcal{V}$, es decir, son de la forma $T\mathbf{x} \in \mathcal{W}$.

*No confundáis la imagen de un elemento $\mathbf{v}$ con la imagen de toda la aplicación T . En el libro de Lay [1] a la imagen de T se la denomina rango de T .

Ejemplo 4.1.

1. La aplicación $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ que pone a cero la tercera componente:

$$T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

tiene como dominio y codominio a $\mathbf{R}^3$. Su rango es el subespacio de $\mathbf{R}^3$ formado por todos los vectores $(a, b, 0)$ con $a, b \in \mathbf{R}$ arbitrarios (el plano $x_3=0$)

2. La aplicación $T : \mathbf{R}^3 \rightarrow \mathbf{R}$ que elige la segunda componente

$$T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2$$

tiene como dominio $\mathbf{R}^3$, codominio $\mathbf{R}$ y rango todo $\mathbf{R}$.

3. En el caso de la aplicación de coordenadas el dominio es el espacio vectorial $\mathcal{V}$ dado y el codominio es $\mathbf{R}^n$, donde n es la dimensión de $\mathcal{V}$. Es decir, $[\cdot]_{\mathcal{B}} : \mathcal{V} \rightarrow \mathbf{R}^n$ y el rango es todo $\mathbf{R}^n$.

Calcular la imagen de un cierto vector $\mathbf{v}$ del dominio $\mathcal{V}$ es simplemente calcular el valor $\mathbf{w}$ en el codominio $\mathcal{W}$ que le corresponde a $\mathbf{v}$: $\mathbf{w} = T(\mathbf{v})$. Ilustremos el cálculo de la preimagen de un vector del codominio $\mathbf{w} \in \mathcal{W}$: en el ejemplo 4.1, 1., la preimagen de $\mathbf{0}$ es el conjunto de vectores $\{(0, 0, x_3) \in \mathbf{R}^3 : x_3 \in \mathbf{R}\}$.

Ejercicio 4.2. Calculad la preimagen de 1 en la aplicación 2 del ejemplo 4.1 ¿ Es un subespacio vectorial ?

Las propiedades interesantes que tiene la función de coordenadas son las que le confieren el carácter de ser una *aplicación lineal*.

Definición 4.3. Una aplicación $T : \mathcal{V} \rightarrow \mathcal{W}$ es **lineal** si tiene las dos siguientes propiedades:

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para todo $\mathbf{u}, \mathbf{v}$ en el dominio de T .
2. $T(c\mathbf{u}) = c T(\mathbf{u})$ para todo $\mathbf{u}$ y todo escalar c .

La aplicación del ejemplo 4.1, 1. es lineal porque

$$T(\mathbf{x} + \mathbf{y}) = T \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ 0 \end{bmatrix} = T(\mathbf{x}) + T(\mathbf{y})$$

$$T(c\mathbf{x}) = \begin{bmatrix} cx_1 \\ cx_2 \\ 0 \end{bmatrix} = c \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} = cT(\mathbf{x})$$

y la 2. porque

$$T(\mathbf{x} + \mathbf{y}) = x_2 + y_3 = T(\mathbf{x}) + T(\mathbf{y}), \quad T(c\mathbf{x}) = cx_2 = cT(\mathbf{x})$$

Ejemplo 4.4. Las aplicaciones lineales tienen un profundo significado geométrico. Especialmente, aquellas cuyo dominio y codominio son el mismo conjunto $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$, se pueden considerar como *transformaciones** (lineales) del espacio $\mathbf{R}^n$. Veamos algunos ejemplos significativos.

1. Una dilatación en $\mathbf{R}^n$ es una aplicación $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$ que multiplica todos los vectores por un factor real k determinado: $\mathbf{x} \rightarrow T(\mathbf{x}) = k\mathbf{x}$. Si $0 \leq k < 1$ entonces tenemos más bien una contracción, y si $k < 1$ se produce una inversión de todos los vectores respecto al origen, además de la dilatación o contracción.
2. La reflexión respecto a un eje en $\mathbf{R}^2$, por ejemplo el eje x_2 (ver figura 4.1)

Ejercicio 4.5. Demostrad que una dilatación es lineal. Demostrad que la reflexión respecto al eje y es también lineal (sugerencia: observad que en coordenadas la transformación se escribe $(x_1, x_2) \mapsto (-x_1, x_2)$)

Proposición 4.6. Si T es una aplicación lineal

$$T(\mathbf{0}) = \mathbf{0}, \quad T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p).$$

Demostración. Se tiene que $T(\mathbf{0}) = T(\mathbf{x} - \mathbf{x}) = T(\mathbf{x}) - T(\mathbf{x}) = \mathbf{0}$. Con tres vectores $T(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3) = T(c_1\mathbf{x}_1 + (c_2\mathbf{x}_2 + c_3\mathbf{x}_3)) = T(c_1\mathbf{x}_1) + T(c_2\mathbf{x}_2 + c_3\mathbf{x}_3) = c_1T(\mathbf{x}_1) + T(c_2\mathbf{x}_2) + T(c_3\mathbf{x}_3) = c_1T(\mathbf{x}_1) + c_2T(\mathbf{x}_2) + c_3T(\mathbf{x}_3)$. Para más vectores se demuestra por inducción. $\square$

*en el libro de Lay se utiliza el término transformación para denotar una aplicación en general

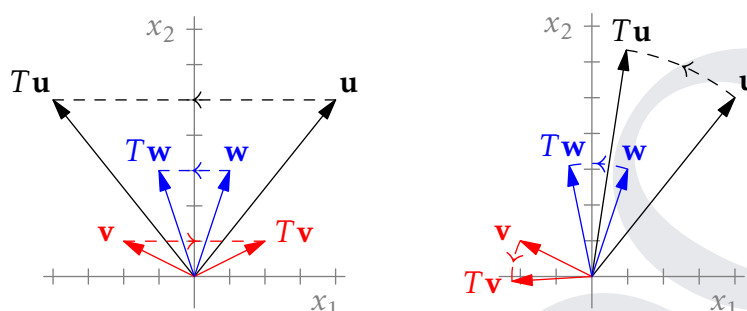


Figura 4.1: Transformaciones geométricas del plano: reflexión respecto al eje x_2 y rotación de 30°

La última fórmula de la proposición 4.6 se parafrasea diciendo que las aplicaciones lineales *respetan* las combinaciones lineales, y en física y tecnología es la formulación matemática del *principio de superposición* que muchos sistemas obedecen: los estados eléctricos de un circuito, o en general las disposiciones de los campos electromagnéticos que obedecen a las ecuaciones de Maxwell, los orbitales electrónicos que obedecen la ecuación de Schrödinger, etc.

Además del rango, un conjunto importante asociado a una aplicación lineal es su *núcleo* o espacio nulo. Repetimos en la siguiente definición la definición de rango, por conveniencia de la exposición.

Definición 4.7. Dada aplicación $T : \mathcal{V} \rightarrow \mathcal{W}$ lineal

- La **imagen** o rango de T es el conjunto de vectores $\mathbf{w}$ de $\mathcal{W}$ tales que son imagen de algún $\mathbf{v} \in \mathcal{V}$, es decir, son de la forma $\mathbf{w} = T(\mathbf{v}) \in \mathcal{W}$. Se denota

$$\text{imagen } T = \{\mathbf{w} \in \mathcal{W} : \exists \mathbf{v} \in \mathcal{V}, \mathbf{w} = T(\mathbf{v})\}$$

- El **núcleo** o espacio nulo de T es el conjunto de vectores de su dominio $\mathbf{v} \in \mathcal{V}$ tales que $T(\mathbf{v}) = \mathbf{0}$. Se denota

$$\text{núcleo } T = \ker T = \{\mathbf{v} \in \mathcal{V} : T(\mathbf{v}) = \mathbf{0}\}$$

En otras palabras, la imagen de una aplicación es el conjunto de las imágenes de todos los elementos de $\mathcal{V}$, y el núcleo es el conjunto preimagen de $\mathbf{0}$. La imagen es un subconjunto de $\mathcal{W}$, mientras que el núcleo está en $\mathcal{V}$.

Teorema 4.8. El núcleo y la imagen de una aplicación lineal son espacios vectoriales.

Demostración. Si $\mathbf{v}_1$ y $\mathbf{v}_2$ están en el núcleo de T , entonces $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) = c_1\mathbf{0} + c_2\mathbf{0} = \mathbf{0}$, por lo que $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ también está en el núcleo

de T . Si $\mathbf{w}_1$ y $\mathbf{w}_2$ están en el rango de T , entonces existen $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ tales que $\mathbf{w}_1 = T\mathbf{v}_1$ y $\mathbf{w}_2 = T\mathbf{v}_2$. Pero entonces $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\mathbf{w}_1 + c_2\mathbf{w}_2$, por lo que $c_1\mathbf{w}_1 + c_2\mathbf{w}_2$ está en el rango de T . $\square$

Ejercicio 4.9. Encontrar la imagen y el núcleo de las aplicaciones lineales del ejemplo 4.1.

Ejercicio 4.10. Sea el conjunto $C^1(\mathbf{R})$ el conjunto de funciones diferenciables $f(x)$ y con derivada continua para todo $x \in \mathbf{R}$. ¿Es un espacio vectorial?. Demostrar que la aplicación

$$T : C^1(\mathbf{R}) \rightarrow C^0(\mathbf{R})$$

$$f(x) \mapsto \frac{d}{dx}f(x)$$

es lineal. ¿Cuál es su núcleo? ¿Y su imagen?

4.2. Aplicaciones matriciales

Se puede asociar a una matriz $m \times n$ una aplicación T sobre los vectores de $\mathbf{R}^n$ utilizando el producto de la matriz por un vector:

$$\begin{aligned} T : \mathbf{R}^n &\rightarrow \mathbf{R}^m \\ \mathbf{x} &\mapsto T(\mathbf{x}) = A\mathbf{x} \end{aligned} \quad (4.1)$$

El dominio de T es $\mathbf{R}^n$ y el codominio es $\mathbf{R}^m$. Llamaremos a este tipo de funciones T **aplicaciones matriciales**, y si es necesario denotaremos por T_A la aplicación matricial asociada a la matriz A . Denotaremos también la aplicación como $\mathbf{x} \mapsto A\mathbf{x}$.

Ejemplo 4.11. La transformación trivial de $\mathbf{R}^n$ es la transformación $Id : \mathbf{R}^n \rightarrow \mathbf{R}^n$ dada por

$$Id(\mathbf{x}) = I_n \mathbf{x} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Ejercicio 4.12. ¿Cuál es el resultado de la transformación trivial?

Ejemplo 4.13. Sea la aplicación $T : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ tal que

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

Dados los vectores

$$\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

i) encontrar la imagen de T de un vector, por ejemplo $\mathbf{u}$, es sencillo. Basta con calcular el producto matriz-vector:

$$T(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ -3 \end{bmatrix};$$

ii) Encontrar las preimágenes de un vector $\mathbf{b}$, es decir, un vector $\mathbf{x}$ tal que $A\mathbf{x} = \mathbf{b}$, requiere calcular la solución de un sistema:

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & -2 & 3 \\ 1 & 1 & 2 \\ 0 & 3 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 \\ 0 & 3 & -1 \\ 0 & 3 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{7}{3} \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

por lo que $x_1 = 7/3$, $x_2 = -1/3$ y $\mathbf{x} = \frac{1}{3} \begin{bmatrix} 7 \\ -1 \end{bmatrix}$ es la preimagen de $\mathbf{b}$. Es importante darse cuenta de que, dado $\mathbf{b}$, en este sistema particular sólo hay una posible preimagen. Y puede incluso suceder que no haya ninguna, es decir, que $\mathbf{b}$ no esté en la imagen de T . Por ejemplo, intentando encontrar la preimagen de $\mathbf{c}$ tenemos

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

y el sistema es incompatible, por lo que $\mathbf{c}$ no tiene preimagen y, por tanto, no está en el rango de T .

La observación fundamental es la siguiente.

Proposición 4.14. *Las aplicaciones matriciales son lineales.*

Demostración. El teorema 1.30 implica que $T_A(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T_A\mathbf{x} + T_A\mathbf{y}$ y $T_A(c\mathbf{x}) = Ac\mathbf{x} = cA\mathbf{x} = cT_A(\mathbf{x})$. $\square$

Ejercicio 4.15. Hallad el núcleo y la imagen de la aplicación matricial del Ejemplo 4.13.

Proposición 4.16. Para una aplicación matricial $T(\mathbf{x}) = A\mathbf{x}$, el núcleo es el espacio nulo de A , y el rango el espacio columna de A

$$\text{núcleo } T = \text{Nul } A, \quad \text{imagen } T = \text{Col } A$$

4.3. La matriz de una aplicación lineal

Ejemplo 4.17. (Rotación de $\pi/6$). Consideremos la transformación de $\mathbf{R}^2$ que a cada vector le asigna su imagen rotada 30° , como se ilustra en la figura 4.1. Si $\mathbf{x} = (x, y) = (r \cos \alpha, r \sin \alpha)$ son las coordenadas de un vector $\mathbf{x}$ en polares, las coordenadas del vector rotado son, ($30^\circ = \pi/6$ radianes)

$$x' = r \cos\left(\alpha + \frac{\pi}{6}\right), \quad y' = r \sin\left(\alpha + \frac{\pi}{6}\right) \quad (4.2)$$

Esta rotación es una aplicación lineal, aunque no sea del todo evidente. La propiedad $T(c\mathbf{x}) = cT(\mathbf{x})$ es sencilla de demostrar: el rotado de $c\mathbf{x}$ es c veces el rotado de $\mathbf{x}$. Sin embargo, ver que el rotado de la suma es la suma de los rotados no es tan sencillo. Geométricamente de la figura 4.2, en donde se ilustra una rotación (de 70° por claridad gráfica) se puede intuir que es cierto. Desarrollando las

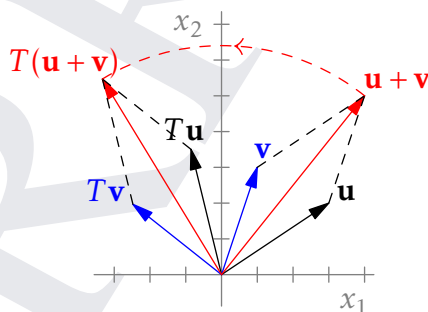


Figura 4.2: Rotación de 70°

coordenadas (4.2)

$$x' = r \cos\left(\alpha + \frac{\pi}{6}\right) = r \cos \alpha \cos \frac{\pi}{6} - r \sin \alpha \sin \frac{\pi}{6} = \frac{\sqrt{3}}{2}x - \frac{1}{2}y$$

$$y' = r \sin\left(\alpha + \frac{\pi}{6}\right) = r \sin \alpha \cos \frac{\pi}{6} + r \cos \alpha \sin \frac{\pi}{6} = \frac{1}{2}x + \frac{\sqrt{3}}{2}y$$

donde se ha usado que $r \cos \alpha = x$ y $r \sin \alpha = y$. Vemos entonces que la rotación es equivalente a una aplicación matricial

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

y por tanto, según la proposición 4.14, es lineal.

Es importante el hecho expuesto en el ejemplo anterior: una transformación lineal, la rotación de 30° , ha sido *implementada* mediante una transformación matricial. Las aplicaciones lineales están definidas por las propiedades de linealidad de la definición 4.3. Todas las aplicaciones matriciales son aplicaciones lineales, por lo que el concepto de aplicación lineal parece más amplio. Pero según el siguiente resultado, el teorema 4.18, las aplicaciones lineales entre $\mathbf{R}^n$ y $\mathbf{R}^m$, son aplicaciones matriciales, con lo que *ambos tipos de aplicación, matricial y lineal, son equivalentes* entre $\mathbf{R}^n$ y $\mathbf{R}^m$. Este hecho se utiliza de forma esencial en el cómputo y la implementación de modelos lineales por ordenador.

Recordemos que los *vectores canónicos* de $\mathbf{R}^n$ son los vectores columna de la matriz identidad I_n , que denotaremos por $\mathbf{e}_i$, $i = 1, \dots, n$. Es decir

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

El teorema de implementación de aplicaciones lineales es el siguiente.

Teorema 4.18. *Si $T : \mathbf{R}^n \rightarrow \mathbf{R}^m$ es una aplicación lineal, entonces existe una única matriz A (de $m \times n$) tal que*

$$T(\mathbf{x}) = A\mathbf{x} \quad \text{para todo } \mathbf{x} \in \mathbf{R}^n$$

La matriz A es aquella que tiene como columna k -ésima el vector $T(\mathbf{e}_j)$, la imagen bajo T del vector canónico $\mathbf{e}_j$:

$$A = \left[T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \dots \mid T(\mathbf{e}_n) \right]$$

Denominaremos a la matriz A de una aplicación lineal T , **matriz canónica de T** . Evidentemente, la matriz canónica de una aplicación matricial T_A es A (la matriz es A y la aplicación T_A)

Ejemplo 4.19. Sea una aplicación lineal $T : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ tal que, si $\mathbf{e}_1$ y $\mathbf{e}_2$ son los vectores canónicos de $\mathbf{R}^2$,

$$T(\mathbf{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad \text{y} \quad T(\mathbf{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

Según el teorema 4.18, podemos escribir directamente la matriz de la aplicación lineal como

$$A = \begin{bmatrix} 5 & -3 \\ -7 & 8 \\ 2 & 0 \end{bmatrix}$$

Razonemos por qué esto es así, dando una idea de la demostración del teorema. Cualquier vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ se puede escribir como combinación lineal de los vectores canónicos:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2.$$

Al ser T una aplicación lineal

$$\begin{aligned} T(\mathbf{x}) &= T(x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) \\ &= x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 & -3 \\ -7 & 8 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

como aseguraba el teorema.

Interesa también observar que una aplicación lineal de $\mathbf{R}^n$ a $\mathbf{R}^m$ se puede considerar como una *función vectorial* de varias variables. El resultado de la aplicación es un vector de m componentes, cada una de las cuales es una función lineal escalar de n variables. Es decir, las componentes son, observemos, combinaciones lineales de las variables. Los coeficientes de la combinación lineal correspondiente a la componente i -ésima son los elementos de la *fila* i -ésima de la matriz A . Por ejemplo,

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

Existen aplicaciones lineales sobre espacios más generales que $\mathbf{R}^n$ que no son representables como aplicaciones matriciales. Veremos ejemplos más adelante.

Ejemplo 4.20. Matriz canónica de la aplicación *dilatación* por un factor 3 $T(\mathbf{x}) = 3\mathbf{x}$. Esta matriz asociada A se encuentra calculando la acción de la aplicación sobre los vectores canónicos

$$T(\mathbf{e}_1) = 3\mathbf{e}_1 = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad T(\mathbf{e}_2) = 3\mathbf{e}_2 = 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

con lo cual

$$T(\mathbf{x}) = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Ejemplo 4.21. La aplicación *rotación* por un ángulo φ se define como aquella que rota todos los vectores un ángulo φ en sentido positivo (contrario a las agujas del reloj) La aplicación sobre los vectores canónicos es

$$T(\mathbf{e}_1) = \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} -\sin \varphi \\ \cos \varphi \end{bmatrix}$$

con lo cual se obtiene una aplicación matricial con una *matriz de rotación*

$$T(\mathbf{x}) = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

4.4. Inyectividad y sobreyectividad

Definición 4.22. Se dice que una aplicación entre dos conjuntos $T : \mathcal{X} \rightarrow \mathcal{Y}$ es **sobreyectiva** (o **sobre** $\mathcal{Y}$)^{*} si todos los elementos del codominio $y \in \mathcal{Y}$ son alcanzados por la aplicación, es decir, si todos los $y \in \mathcal{Y}$ son la imagen de algún $x \in \mathcal{X}$ ($T(x) = y$)

Se dice que $T : \mathcal{X} \rightarrow \mathcal{Y}$ es **inyectiva** si cada $y \in \mathcal{Y}$ es la imagen de un $x \in \mathcal{X}$, como mucho (T podría no ser sobreyectiva) Es decir, cada imagen tiene una sola preimagen.

Si $T : \mathcal{X} \rightarrow \mathcal{Y}$ es inyectiva y sobreyectiva, se dice que es **biyectiva**, y establece una correspondencia **uno a uno** entre $\mathcal{X}$ e $\mathcal{Y}$, es decir, a cada $x \in \mathcal{X}$ le corresponde una sola imagen $y \in \mathcal{Y}$, y a todo $y \in \mathcal{Y}$ una sola preimagen x .

Ejercicio 4.23. Hemos dado previamente un nombre a una aplicación biyectiva entre espacios vectoriales, que sea lineal. ¿Cuál ?

^{*}también se usa el término suprayectiva

Teorema 4.24 (Parte del teorema I). *Una aplicación lineal $T : \mathbf{R}^n \rightarrow \mathbf{R}^m$ es inyectiva si y solo si la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene solo la solución trivial, es decir, núcleo $T = \{\mathbf{0}\}$. Si A es la matriz canónica de T , esto equivale a que las columnas de A formen un conjunto linealmente independiente.*

Demostración. La aplicación T es inyectiva si cada imagen $\mathbf{y} \in \mathbf{R}^m$ tiene solo una (o ninguna) preimagen $\mathbf{x} \in \mathbf{R}^n$, con $T(\mathbf{x}) = \mathbf{y}$. Si T es inyectiva, entonces $\mathbf{0}$ solo puede tener una preimagen, que resulta ser $\mathbf{x} = \mathbf{0} \in \mathbf{R}^n$ (que siempre es preimagen) luego $T(\mathbf{x}) = \mathbf{0}$ solo tiene la solución trivial.

Viceversa, si $T(\mathbf{x}) = \mathbf{0}$ solo tiene la solución trivial, es decir, $\mathbf{0}$ solo tiene una preimagen, todos los $\mathbf{y} \in \mathbf{R}^m$ tienen solo una preimagen. Porque si un $\mathbf{y}$ tuviera dos preimágenes distintas $\mathbf{x}_1, \mathbf{x}_2$, entonces $\mathbf{y} = T(\mathbf{x}_1)$ e $\mathbf{y} = T(\mathbf{x}_2)$. Restando estas ecuaciones $\mathbf{0} = T(\mathbf{x}_1) - T(\mathbf{x}_2) = T(\mathbf{x}_1 - \mathbf{x}_2)$, y como $T(\mathbf{x}) = \mathbf{0}$ solo tiene la solución trivial $\mathbf{x} = \mathbf{0}$, ha de ser $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, es decir, $\mathbf{x}_1 = \mathbf{x}_2$, la misma preimagen: contradicción.

Evidentemente, la ecuación $T(\mathbf{x}) = \mathbf{0}$ equivale a la ecuación $A\mathbf{x} = \mathbf{0}$, que tiene sólo la solución trivial si las columnas de A forman un conjunto linealmente independiente (teorema 1.44) $\square$

Es interesante observar que el teorema 4.24 afirma que si T es lineal basta comprobar el número de preimágenes de $\mathbf{0} \in \mathbf{R}^m$ para saber si es inyectiva, y no es necesario realizar esta comprobación para todos los elementos $\mathbf{y} \in \mathbf{R}^m$.

Teorema 4.25 (Parte del teorema S). *Una aplicación lineal $T : \mathbf{R}^n \rightarrow \mathbf{R}^m$ es sobreyectiva si su imagen es todo $\mathbf{R}^m$, imagen $T = \mathbf{R}^m$. Esto sucede si y solo si las columnas de su matriz estándar A generan todo $\mathbf{R}^m$.*

Demostración. La aplicación T es sobreyectiva si y solo si $T(\mathbf{x}) = \mathbf{y}$ tiene solución para todo $\mathbf{y} \in \mathbf{R}^m$. Esto equivale a que la ecuación $A\mathbf{x} = \mathbf{y}$ tenga solución para todo $\mathbf{y} \in \mathbf{R}^m$. Pero esto equivale (teorema 1.27) a que las columnas de A generen $\mathbf{R}^m$. $\square$

Transformaciones lineales y el teorema de la matriz invertible. En la proposición 3.18 asegurábamos que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución si y sólo si el vector $\mathbf{b}$ está en el espacio columna de A . En términos de aplicaciones lineales, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución si y sólo si el vector $\mathbf{b}$ está en la imagen (rango) de la aplicación matricial $T_A(\mathbf{x}) = A\mathbf{x} = \mathbf{b}$.

Una vez que hemos extendido los teoremas I y S, el teorema B (teorema de la matriz invertible 2.27) se puede extender análogamente. Recopilamos también

aquí algunas conclusiones obtenidas en el capítulo 3.

Teorema 4.26 (Extensión del teorema de la matriz invertible). *Sea A $n \times n$. Entonces es equivalente a que A sea invertible cualquiera de las siguientes afirmaciones.*

- f. La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es biyectiva.*
- i. La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es sobreyectiva.*
- m. Las columnas de A forman una base para $\mathbf{R}^n$.*
- n. $\text{Col } A = \mathbf{R}^n$.*
- o. $\dim \text{Col } A = n$.*
- p. $\text{rango } A = n$.*
- q. $\text{Nul } A = \{\mathbf{0}\}$.*
- r. $\dim \text{Nul } A = 0$.*

Demostración. La única cuestión peculiar es el apartado i. Lo que afirma es que si la transformación es sobreyectiva, entonces es obligatoriamente biyectiva.

El apartado m. se demostró en el ejemplo 3.23. □

4.5. Composición de aplicaciones y multiplicación de matrices.

Recordemos lo que hicimos en el capítulo 2 para definir el producto de matrices. Supongamos que *componemos* dos aplicaciones matriciales consecutivas a un vector, es decir, primero hacemos actuar una aplicación matricial de matriz B sobre un vector $\mathbf{x}$ de $\mathbf{R}^p$ y sobre el vector resultante hacemos actuar una aplicación matricial de matriz A . El vector resultante es $A(B\mathbf{x})$.

Ejercicio 4.27. Demostrad que la aplicación compuesta $\mathbf{x} \mapsto A(B\mathbf{x})$ es lineal.

Como la aplicación compuesta es lineal, existirá una matriz asociada, según el teorema 4.18, que hará que la aplicación compuesta sea matricial. Ya calculamos en la discusión anterior a la definición 2.3 que esta matriz es la que definimos como producto AB . Podemos comprobarlo aquí de otra manera: calculando la matriz canónica con el teorema 4.18. Como Be_i es la columna i -ésima $\mathbf{b}_i$ de B , es decir $Be_i = \mathbf{b}_i$, aplicando la transformación a un vector $\mathbf{e}_i$ de la base canónica

tenemos que

$$A(Be_i) = A(\mathbf{b}_i) = A\mathbf{b}_i$$

La matriz canónica buscada es entonces

$$\left[A(Be_1) \mid A(Be_2) \mid \cdots \mid A(Be_p) \right] = \left[A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_p \right] = A \cdot B$$

Resumiendo, la composición de aplicaciones matriciales corresponde al producto de matrices:

$$A(B\mathbf{x}) = (AB)\mathbf{x}.$$

Transformaciones lineales invertibles. La ecuación $A^{-1}(A\mathbf{x}) = \mathbf{x}$ se puede interpretar en términos de las transformaciones lineales matriciales asociadas a A y A^{-1} : T_A transforma $\mathbf{x}$ en otro vector $A\mathbf{x}$, mientras que $T_{A^{-1}}$ transforma cualquier $A\mathbf{x}$ en el vector $\mathbf{x}$, es decir, transforma un vector $\mathbf{y}$ en aquel vector $\mathbf{x}$ que se transformaría en $\mathbf{y}$ bajo T_A : $A^{-1}\mathbf{y} = \mathbf{x}$ con $A\mathbf{x} = \mathbf{y}$.

Se dice que una transformación S es la inversa de la transformación lineal T si

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{y} \quad T(S(\mathbf{x})) = \mathbf{x} \quad \text{para todo } \mathbf{x} \in \mathbf{R}^n.$$

Teorema 4.28. Sea $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$ una transformación lineal, y A su matriz asociada. Entonces T es invertible si y sólo si A es invertible. En ese caso, la matriz asociada a la aplicación inversa S es A^{-1} , es decir, $S(\mathbf{x}) = A^{-1}\mathbf{x}$.

Demostración. Si T es invertible, $\mathbf{x} = S \circ T(\mathbf{x}) = S(T(\mathbf{x})) = S(A\mathbf{x}) = B A \mathbf{x}$ donde B es la matriz asociada a S , y también $T \circ S(\mathbf{x}) = T(S(\mathbf{x})) = A B \mathbf{x} = \mathbf{x}$. La matriz canónica tanto de $S \circ T$ como de $T \circ S$ es la identidad I_n . Entonces $AB = I$ y $BA = I$ y por tanto la matriz A es invertible y B es su inversa A^{-1} .

Evidentemente, $S(\mathbf{x}) = A^{-1}\mathbf{x}$ es la aplicación inversa de T , si A es invertible. $\square$

4.6. La aplicación de coordenadas

Retomemos el estudio de la aplicación de coordenadas, presentada en la sección 3.4.1, usando los conceptos de las aplicaciones lineales discutidos en este capítulo. La relación $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una aplicación $\mathcal{V} \rightarrow \mathbf{R}^n$ con dominio $\mathcal{V}$ y codominio $\mathbf{R}^n$. que es uno a uno (biyectiva) es decir a todos los vectores $\mathbf{x} \in \mathcal{V}$ les corresponde un vector de $\mathbf{R}^n$, y sólo uno (esto es lo que dice el teorema de representación única 3.34) Como demostramos en la sección 3.4.1, es una aplicación lineal.

Teorema 4.29. *La aplicación $\mathbf{x} \rightarrow [\mathbf{x}]_{\mathcal{B}}$, que da las coordenadas de un vector $\mathbf{x}$ de un espacio vectorial $\mathcal{V}$ en una base $\mathcal{B}$ de $\mathcal{V}$ determinada, es una aplicación uno a uno y lineal.*

A una aplicación $T : \mathcal{V} \rightarrow \mathcal{W}$ entre dos espacios vectoriales $\mathcal{V}$ y $\mathcal{W}$, que es uno a uno y lineal, se la denomina **isomorfismo**. La aplicación mantiene todas las combinaciones lineales y relaciones de dependencia: si $\mathbf{v}_1, \dots, \mathbf{v}_p \in \mathcal{V}$, y $\mathbf{w}_1 = T(\mathbf{v}_1), \dots, \mathbf{w}_p = T(\mathbf{v}_p) \in \mathcal{W}$, entonces

$$T(c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \dots + c_pT(\mathbf{v}_p) = c_1\mathbf{w}_1 + \dots + c_p\mathbf{w}_p.$$

Si dos espacios vectoriales $\mathcal{V}$ y $\mathcal{W}$ se pueden relacionar a través de un isomorfismo, entonces su estructura como espacios vectoriales es idéntica y se dice que son espacios **isomorfos**. Se puede pensar que, aunque pueden estar definidos de modo diferente y contener elementos aparentemente muy distintos, son el mismo espacio vectorial. En el ejemplo 3.43 vimos cómo el espacio de polinomios de hasta tercer grado es isomorfo a $\mathbf{R}^4$.

4.7. Respuestas a los ejercicios

4.10 El núcleo lo constituyen las funciones constantes, que como se puede demostrar usando el teorema del valor medio (enunciado en el curso de Cálculo I) son las únicas de derivada nula. La imagen es el conjunto completo de funciones continuas $C^0(\mathbf{R})$, ya que por el teorema fundamental del cálculo integral toda función continua $g(x)$ tiene una primitiva $f(x) = \int g(x)dx$ cuya derivada $f'(x) = g(x)$ es $g(x)$.

4.8. Resumen

Definición. Una aplicación $T : \mathcal{V} \rightarrow \mathcal{W}$ es **lineal** si

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para todo $\mathbf{u}, \mathbf{v}$ en el dominio de T .
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ para todo $\mathbf{u}$ y todo escalar c .

Proposición. Si T es lineal $T(\mathbf{0}) = \mathbf{0}$ y

$$T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p)$$

Definición. Si $T : \mathcal{V} \rightarrow \mathcal{W}$ es lineal

- La **imagen** de T es

$$\text{imagen } T = \{\mathbf{w} \in \mathcal{W} : \exists \mathbf{v} \in \mathcal{V}, \mathbf{w} = T(\mathbf{v})\}$$

- El **núcleo** o espacio nulo de T es

$$\text{núcleo } T = \ker T = \{\mathbf{v} \in \mathcal{V} : T(\mathbf{v}) = \mathbf{0}\}$$

Teorema. El núcleo y la imagen de una aplicación lineal son espacios vectoriales.

Proposición. Las aplicaciones matriciales $T_A(\mathbf{x}) \equiv A\mathbf{x}$ son lineales.

Proposición. Si $T(\mathbf{x}) = A\mathbf{x}$

$$\text{núcleo } T = \text{Nul } A, \quad \text{imagen } T = \text{Col } A$$

Teorema. Si $T : \mathbf{R}^n \rightarrow \mathbf{R}^m$ es lineal, su **matriz canónica** A (de $m \times n$) es

$$A = \left[T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n) \right]$$

y es tal que $T(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x} \in \mathbf{R}^n$.

Definición. Si $T : \mathcal{X} \rightarrow \mathcal{Y}$

- T es **sobreyectiva** (o **sobre** $\mathcal{Y}$) si $\forall y \in \mathcal{Y} \exists x \in \mathcal{X}$ con $T(x) = y$.
- T es **inyectiva** si $T(x_1) = T(x_2) \Rightarrow x_1 = x_2$.
- T es **biyectiva** si es inyectiva y sobreyectiva.

Teorema (S). Sea A de $m \times n$. Las siguientes afirmaciones equivalen:

1. Para todo $\mathbf{b}$ la ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución.
2. Todos los vectores $\mathbf{b} \in \mathbf{R}^m$ son combinaciones lineales de las columnas de A .
3. Las columnas de A generan $\mathbf{R}^m$.
4. Hay una posición pivote en cada fila de A .
5. $\text{Col } A = \mathbf{R}^m$. $\dim \text{Col } A = m$.
6. $\text{rango } A = m$. $\text{imagen } T_A = \mathbf{R}^m$.
7. T_A es sobreyectiva.

Teorema (I). Sea A de $m \times n$. Las siguientes afirmaciones equivalen:

1. No existe solución no trivial de la ecuación $A\mathbf{x} = \mathbf{0}$.
2. No hay una combinación lineal no trivial tal que $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n = \mathbf{0}$ ($\mathbf{a}_i$ columnas de A)
3. Las columnas de A son linealmente independientes.
4. Hay una posición pivote en cada columna de A .
5. $\text{Nul } A = \{\mathbf{0}\}$. $\dim \text{Nul } A = 0$.
6. $\text{rango } A = n$. $\text{núcleo } T_A = \{\mathbf{0}\}$.
7. T_A es inyectiva.

Teorema (Teorema de la matriz invertible ó B). Sea A una matriz $n \times n$ (cuadrada). Las siguientes afirmaciones son equivalentes.

- a. A es invertible.
- b. A es equivalente por filas a la matriz identidad I_n .
- c. A tiene n posiciones pivote (es decir, n pivotes no nulos).
- d. La ecuación homogénea $A\mathbf{x} = \mathbf{0}$ admite sólo la solución trivial.
- e. Las columnas de A forman un conjunto linealmente independiente.
- f. La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es biyectiva.
- g. La ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos

una solución para todo $\mathbf{b} \in \mathbb{R}^n$.

- h. Las columnas de A generan $\mathbb{R}^n$.
- i. La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es sobreyectiva.
- j. Existe una matriz C $n \times n$ tal que $CA = I$.
- k. Existe una matriz D $n \times n$ tal que $AD = I$.
- l. A^T es invertible.
- m. Las columnas de A forman una base para $\mathbb{R}^n$.
- n. $\text{Col } A = \mathbb{R}^n$.
- o. $\dim \text{Col } A = n$.
- p. $\text{rango } A = n$.
- q. $\text{Nul } A = \{\mathbf{0}\}$.
- r. $\dim \text{Nul } A = 0$.

Teorema. La matriz de la aplicación compuesta $T_A \circ T_B$ es el producto $A \cdot B$, es decir $T_A(T_B(\mathbf{x})) = (A \cdot B) \cdot \mathbf{x}$.

Teorema. Sea $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una transformación lineal, y A su matriz asociada. T es invertible si y sólo si A es invertible. La

matriz de la aplicación inversa es A^{-1} .

Teorema. La aplicación de coordenadas $\mathbf{x} \rightarrow [\mathbf{x}]_{\mathcal{B}}$, que da las coordenadas de un vector $\mathbf{x}$ de un espacio vectorial $\mathcal{V}$ en una base $\mathcal{B}$ es un **isomorfismo** (uno a uno y lineal)